University of Naples Federico II

Department of Agricultural Sciences

PhD in Sustainable agricultural and forestry systems and food security

XXXVII Cycle

Behind the colourful surface: a study of anthocyanin pathway regulation in *Solanaceae*

Annalisa Staiti

Supervisor:
Prof. Domenico Carputo

Co-supervisor:
Prof. Vincenzo D'Amelia

INDEX

Abstract

Anthocyanins are a group of polyphenolic water-soluble pigments responsible for the colouration of many flowers, fruits and vegetables in several plant families. Anthocyanins draw an outstanding interest not only for their role in pigmentation but also for their protective functions in plants against environmental and biotic stresses. Moreover, due to their role in human health, research is becoming more addressed in identifying health-promoting anthocyanins and in validating their functional activity against several non-communicable human diseases. For this reason, obtaining a rife resource of anthocyanins *in vitro* and *in planta* is an enticing achievement for food, pharmaceutical and cosmetic industries. The levels of anthocyanins in plants are controlled by both synthesis and degradation processes, which are genetically regulated in a time-specific and tissue-specific manner across different stages of development. In this thesis, we focused on anthocyanin metabolism, with a particular interest in the transcriptional factors implicated in the regulation of this pathway in *Solanaceae*. The first aim of this research was to knock-out the gene Inducer Silencing of Anthocyanins in Cell culture (StISAC), a novel R3-MYB repressor, using CRISPR-Cas9 editing, obtaining a stable platform for anthocyanin production. The second aim was to use the edited potato cell cultures as a unique platform for investigating the possible impact of editing on all the anthocyanin metabolism. In particular, we paid attention to anthocyanin catabolism, mainly focusing on enzymes involved in an active degradation of these pigments, and to post-biosynthetic modifications mediated by decorative enzymes. The third and last aim of this thesis was to use a traditional breeding method to identify transcriptional factors involved in anthocyanin retaining in pepper. This was included as model crop due to its simpler (diploid) genetic inheritance of traits. As for the first aim, the first part of this research focused on the knock-out of the gene *Inducer Silencing of Anthocyanins in Cell culture* (*StISAC*), using CRISPR-Cas9 editing, to increase the production of anthocyanins (**Chapter 2**). The results obtained provided evidence that mutant cell lines doubled the accumulation level and the production of these pigments was stabilized over time. As for the second aim of this thesis, we used the wildtype and the obtained mutant cell lines to deepen the regulation of anthocyanin biosynthesis and degradation in cell culture platform. Since anthocyanins in potato displays a valuable chemical diversity, we focused on post-biosynthetic enzymes, particularly Anthocyanin O-Methyltransferases (AOMTs), carrying out a comprehensive genome-wide analysis and then elucidating expression differences among four selected AOMT-encoding genes using potato cell cultures (**Chapter 3**). To fill the knowledge gap in the study of anthocyanin degradation, we investigated this process at a transcriptome and metabolome level (**Chapter 4**). We identified and proposed candidate peroxidases and β-glucosidases involved in degradation based on the phylogenetic relationships with functionally identified genes from other species (*Brunfelsia calycina*

and *Arabidopsis thaliana*) and RNA-seq transcript profiles of potato cell cultures. The third part of this work was mainly concentrated on fully understanding the anthocyanin genetic regulation *in planta*. A gene-mining and genomics-assisted breeding was chosen to move in the intricated gene regulation that influences anthocyanin retention in pepper fruits (**Chapter 5**). The most significant outcome of this research was the identification of a unique QTL containing multiple transcriptional factors involved in the anthocyanin pathway. The function of these genes was studied using the reverse-genetics approach VIGS (Virus-Induced Gene Silencing). This last chapter is the result of 1-year research conducted in collaboration with Wageningen University and Research, Plant Breeding department (The Netherlands), under the supervision of Dr. Arnaud Bovy and Dr. Yury Tikunov. Our findings contribute to a better understanding of the counterbalance between anthocyanin biosynthesis and degradation in *Solanaceae*. Altogether, results obtained in this thesis provide different strategies for obtaining *in vitro* and *in planta* stable resources of these pigments.

Chapter 1. General introduction

Partially published in: Staiti, A., D'Amelia, V. and Carputo, D. (2022) 'Exploring and exploiting anthocyanins for human needs: potato as a research case study', *Italus Hortus*, 29(2), pp. 18-34. doi: 10.26353/j.itahort/2022.2.18

Chapter 1. General introduction

1.1. The eye-catching *Solanaceae* family

The *Solanaceae* family, commonly known as the nightshade family, has a nearly global distribution except Antarctica, with the greatest concentration of diversity in Central and South America. This diverse family encompasses approximately 3000 species spread across 100 genera (Olmstead et al., 2008), ranking among the largest and most economically significant angiosperm families. Certainly, the Solanaceae are considered one of the most important to humanity, containing a substantial number of essential crops. Many of them belong to the genus *Solanum*, including *Solanum lycopersicum* L. (tomato), *Solanum tuberosum* L. (potato) and *Solanum melongena* L. (eggplant). Further species hold significant importance as the one belonging to the genus *Capsicum* spp. (sweet and chilli peppers), *Physalis* spp. (tomatillo, cape gooseberry) and *Nicotiana* spp. (*N. tabacum*, *N. benthamiana*) (Motti, 2021). Many Solanaceae are also grown for their ornamental features such as *Petunia*, *Brunfelsia*, *Datura* and *Atropa*. Moreover, some *Solanaceae* as tomato, tobacco and *Petunia* serve as crucial biological model systems with diverse applications, including studies on plant development, biotic and abiotic stress response, hormone biology and signalling. Due to its significant diversity, the *Solanaceae* family is also an ideal model for comparative genomics, allowing the exploration of a broad spectrum of genetic traits and evolutionary adaptations (Gebhardt, 2016). Nowadays, as living standards rise, research aimed at enhancing the nutritional quality of horticultural crops has gained significant focus. While taste plays a key role in regulating food intake, various sensory cues - including appearance, odour, texture, temperature, and flavour- also influence consumption, with visual perception often being the initial sensory interaction with food (Wadhera & Capaldi-Phillips, 2014). The Solanaceae family shows a variegated colour palette in their leaves, flowers, fruits and tubers (Figure 1). Among pigments, anthocyanins are considered one of the most important due to their diverse functions and wide-ranging benefits. They are produced by most of the *Solanaceous* plants in different tissues and they became hotspots in the fields of phytochemistry, medicine and breeding programs, helping to develop new varieties with desirable traits. Given the importance of anthocyanins in plant pigmentation, stress response and human health benefits, it is crucial to deepen our understanding of the regulatory mechanisms that control their accumulation. Using *Solanaceae* species as model crops for this research presents a unique opportunity: species such as potato and pepper not only vary widely in anthocyanin content but they also offer well-established genetic tools and resources that facilitate in-depth studies.

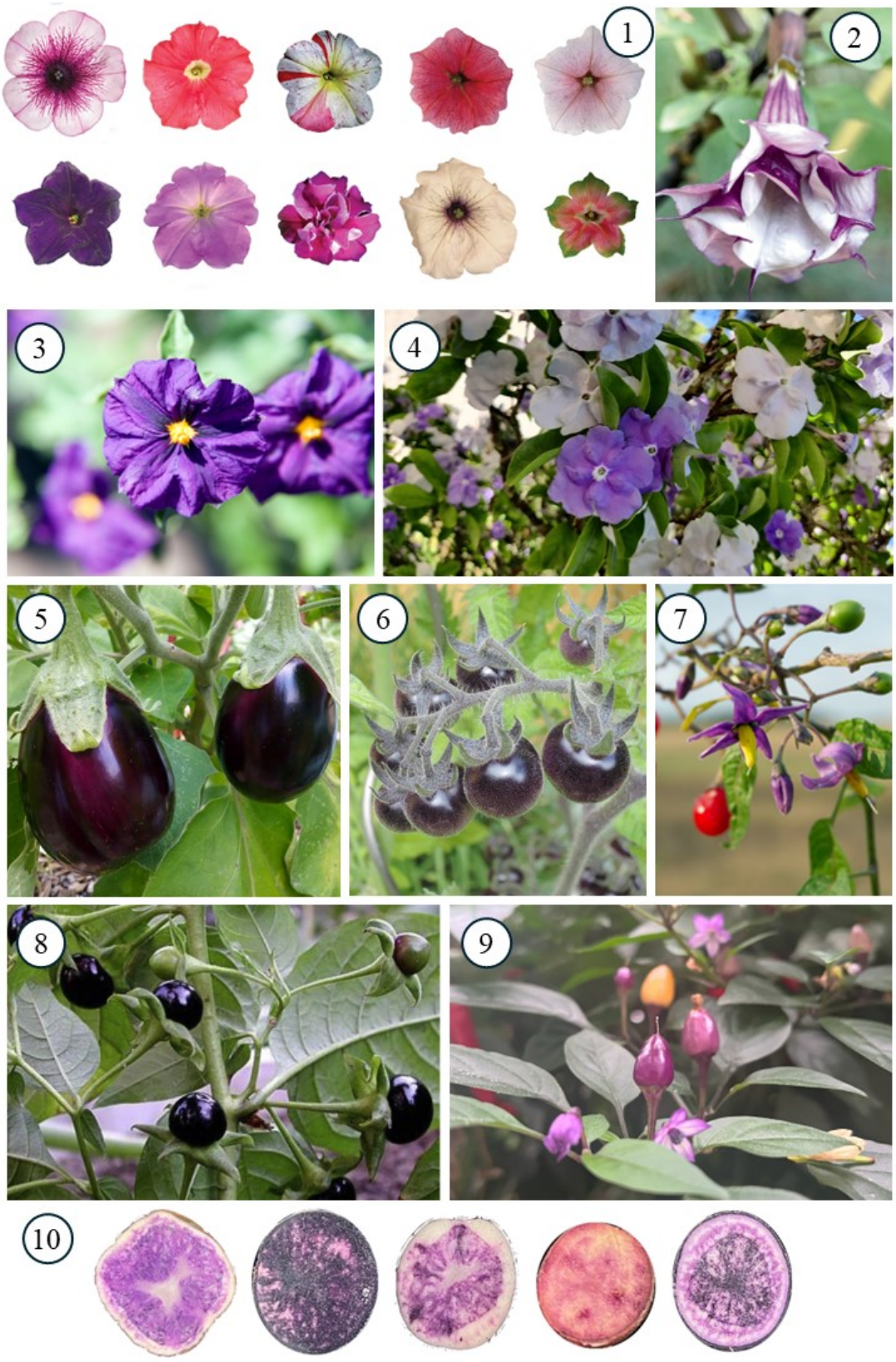

Figure 1. Anthocyanin diversity within the *Solanaceae* flowers, fruits and tubers: 1) *Petunia hybrida* accessions and mutants; 2) *Datura metel* L.; 3) *Lycianthes rantonnetii* L.; 4) *Brunfelsia* spp.; 5) *Solanum melongena* L.; 6) *Solanum lycopersicum* L.; 7) *Solanum dulcamara* L.; 8) *Atropa belladonna* L.; 9) *Capsicum annuum* L.; 10) *Solanum tuberosum* L. (Amiryousefi et al., 2018; Bombarely et al., 2016; Kwakye et al., 2018; Motti, 2021; Staiti et al., 2022).

1.2. Anthocyanin biochemistry and genetic regulation

Plant secondary metabolites are a large group of chemical compounds which are distributed in almost all plant organs. Despite their definition as “secondary”, these molecules play a fundamental role in plant development, propagation and survival. For example, the colouration of flowers and fruits helps to attract pollinators and seed dispersers and is crucial for plant reproduction under natural conditions (Shen et al., 2024). For their important role in plant physiology today the term “specialized” is preferred over secondary metabolites (Dixon & Dickinson, 2024). Among specialized metabolites, anthocyanins (from Greek νθός (anthos) = flower and κυανός (kyanos) = blue), belonging to flavonoids family, are intensely studied for their antimicrobial, antioxidant, and free radical scavenging properties that have applications in food technology, human health and plant protection (Saini et al., 2024). They accumulate in a wide array of plant families including *Rosaceae* (e.g. *Fragaria vesca* and *Rubus fruticosus*), *Moraceae* (e.g. *Ficus carica* and *Morus nigra*), *Brassicaceae* (e.g. *Brassica oleracea* and *Raphanus sativus*), *Poaceae* (e.g. *Oryza sativa* and *Zea mays*), *Solanaceae* (e.g. *Solanum melongena*, *Capsicum annuum,* and *Solanum tuberosum*).

1.2.1. Structural variation and classification of anthocyanins

More than 600 anthocyanins have been identified in nature (J. Liu et al., 2021). They derive from six main anthocyanidins, namely pelargonidin, cyanidin, delphinidin, peonidin, petunidin, and malvidin (Sendri & Bhandari, 2024). Among them, the primary anthocyanidins are: pelargonidin, which produces the colours orange, pink and red; delphinidin, which produces the colours purple, blue, or dark blue; and cyanidin, which produces red or mauve colours and is found mainly in fruits. Anthocyanins consist of a molecule of benzene fused with a molecule of pyran, connected in turn to a phenolic group, which can have different substituent groups: hydroxylic (-OH) or methylic (-CH3). This complex molecule takes the name of "flavylium cation" and it is the basic structure of all anthocyanins. In addition to the different structures of the anthocyanidins, the molecule can be conjugated to different sugar moieties which transform anthocyanidins into anthocyanins. Finally, the hydroxylic group (-OH) of the glycosyl substituents can be acylated with organic acids (as p-coumaric, caffeic, and ferulic acids) via ester bonds, thus obtaining acylated anthocyanins. This aspect is important as some decorations (as acylation, glycosylation and methylation) to the anthocyanin aglycone provide specific functions that can overcome important functions and drawbacks that anthocyanins can have. Indeed, several studies have confirmed that glycosyl acylations increase the ability of anthocyanins to tolerate various physicochemical (light, heat, pH changes) and biochemical (digestive enzymes) factors, contributing to greater stability in the colouring potential of these pigments (Carrillo et al., 2020; Fei et al., 2021; Vidana Gamage et al., 2022; Zhao et al., 2017). In

potato, all six most common anthocyanidins have been found in the tubers, giving them a large complexity of colours and nuances (Figure 2), while in *Capsicum* spp. Delphinidin-3-O-glucoside is the predominant anthocyanin produced (Aza-González et al., 2012; Tang et al., 2020).

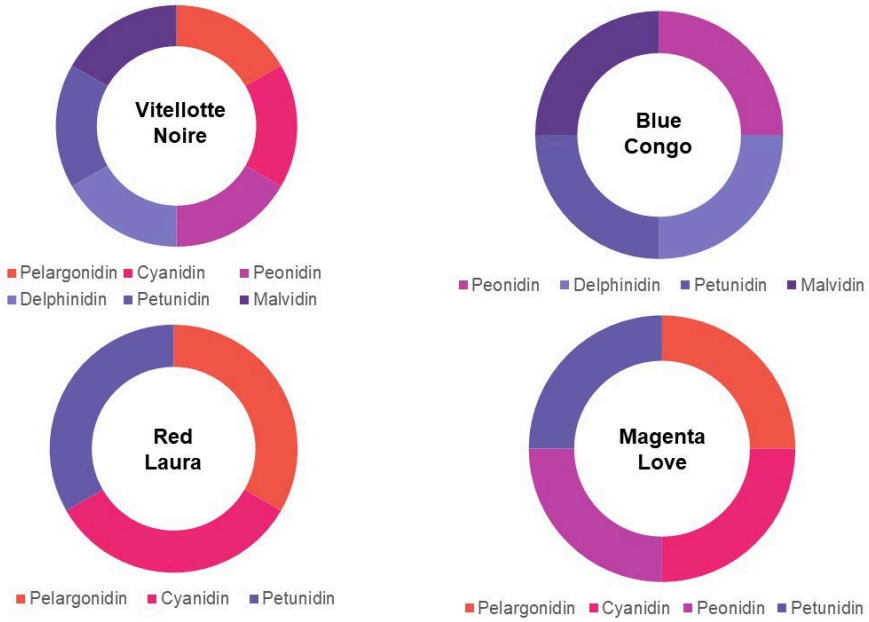

Figure 2. Representative image of the principal anthocyanidin patterns in red and blue pigmented potatoes. The presence of all six most common anthocyanidins gives potato tubers a unique and wide range of colours characteristic of skin and flesh (Burmeister et al., 2011; De Masi et al., 2020; Lachman et al., 2009).

1.2.2. Genetic regulation of anthocyanin biosynthesis

The biosynthetic pathway of anthocyanins is well characterized and conserved in many plant species, including *Solanaceous* plants. The enzymes involved in the biosynthesis of anthocyanidins are mainly localized in the endoplasmic reticulum, organized into a multi-enzyme complex named flavonoid metabolon (Fujino et al., 2018). The biosynthetic pathway can be divided into two sections, the basic flavonoid upstream pathway, which includes early biosynthetic genes (EBGs), and the specific anthocyanin downstream branch, which includes late biosynthetic genes (LBGs) (Lachman et al., 2009). It begins with the synthesis of naringenin chalcone starting from 4-coumaroyl-CoA and malonyl-CoA mediated by the enzyme chalcone synthase (CHS) (Figure 3). Naringenin chalcone is then isomerized by chalcone isomerase (CHI) into naringenin. Flavanone 3-hydroxylase (FH3) enzyme converts naringenin into dihydrokaempferol which can be hydroxylated by flavonoid 3'-hydroxylase (F3'H) or flavonoid 3'-5'-hydroxylase (F3'5'H) in others two dihydroflavonols, respectively dihydroquercetin or dihydromyricetin. Subsequently, the three dihydroflavonols are converted to colourless leukoanthocyanidins by dihydroflavonol 4-reductase (DFR) and later to anthocyanidins coloured by anthocyanidin synthase (ANS). Finally, the sugar moieties are attached

to anthocyanidins by various members of the glycosyltransferase family, such as the flavonoid 3-O glucosyltransferase (UFGT) and can be further acetylated with acyl-aromatic groups by the acyltransferases. CHS is therefore the initial key enzyme for the biosynthesis of flavonoids. F3'H and F3'5'H are the main enzymes responsible for the diversification of anthocyanins, determined by the hydroxylation of the catechol, with consequent variation in colour (Y. Liu et al., 2018). A positive correlation between LBG expression levels and anthocyanin content has been observed (Andre et al., 2009; F. Liu et al., 2018; Tengkun et al., 2019). In fact, in potato peel with high anthocyanin pigmentation, the genes *StF3'H*, *StF3'5'H*, *StDRF*, *StANS,* and *StUFGT* were highly expressed (Andre et al., 2009; Jung et al., 2009; Tengkun et al., 2019). Furthermore, *StDFR* was found to be highly expressed in the pulp of red and purple cultivars Colorado Rose, Mountain Rose, Purple Majesty, and Russet Nugget, indicating a strong correlation between this gene and pigmentation (Stushnoff et al., 2010). LBG as *CaF3'5'H*, *CaDFR*, *CaANS*, and *CaUFGT* were also upregulated in anthocyanin-pigmented peppers (*C. annuum*), reaching a maximum at the late unripe stage before ripening, and then downregulated thereafter, which corresponded to the transient anthocyanin accumulation pattern of these fruits (Aza-González et al., 2012; Borovsky et al., 2004; Zhou et al., 2022).

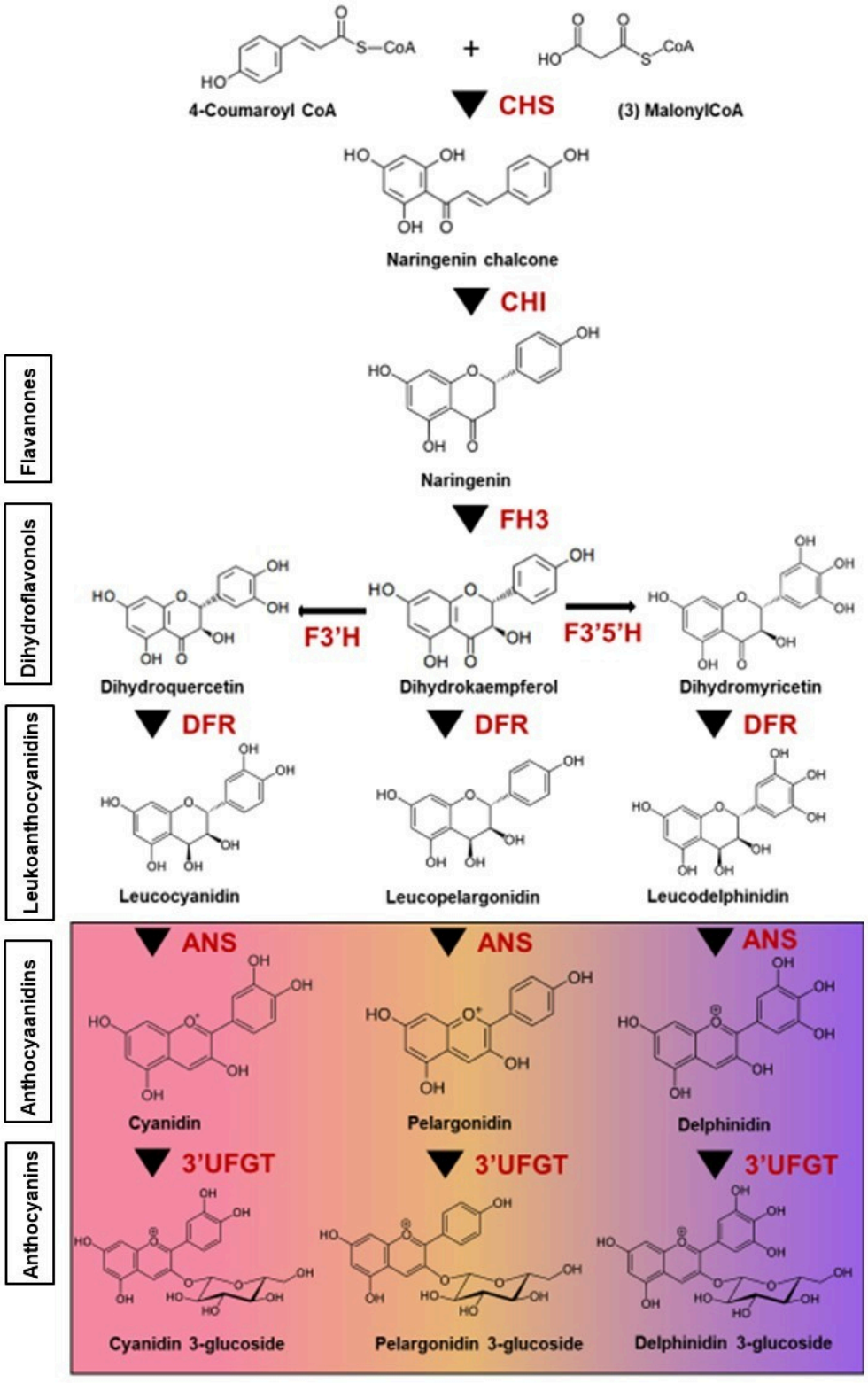

Figure 3. Schematic representation of the anthocyanin biosynthetic pathway. In red are the structural genes regulating each step of the pathway: *CHS*, chalcone synthase; *CHI*, chalcone isomerase; *F3H*, flavanone 3-hydroxylase; *F3'H*, flavonoid 3′-hydroxylase; *F3'5'H*, flavonoid 3′,5′-hydroxylase; *DFR*, dihydroflavonol 4-reductase; *ANS*, anthocyanidin synthase; *3'UFGT*, flavonoid 3-O-glucosyltransferase. Modified from Mekapogu et al. (2020).

The structural genes of the anthocyanin biosynthetic pathway are under the control of a regulatory complex, called MYB-bHLH-WD40 (MBW). The expression of the MYB genes, mainly belonging to the R2R3-MYB subfamily, and of the basic helix loop helix bHLH genes, belonging to subdivision IIIf (Heim et al., 2003), is largely specific for pigmented tissues. By contrast, the expression of the genes encoding the WD40 factors, involved in the stabilization of MBW complexes, is similar between pigmented and non-pigmented tissues (Koes et al., 2005). The R2R3-MYB transcription factors consist of a conserved domain (R2R3) binding DNA and are the key element for the tissue-specific accumulation of anthocyanins. They are considered the most crucial component of this complex and they can induce anthocyanin accumulation by themselves (Kiferle et al., 2015; Spelt et al., 2000). In potato, the gene that codes for an activator of anthocyanins of the R2R3-MYB class is called StAN1 (Jung et al., 2009). Among the differently pigmented varieties, there is a variation in the number of repeats in a 10 amino acid motif at the C-terminus of StAN1 (Strygina et al., 2019). In pepper, CaAN2 and CaAN3 have emerged as prominent players within this regulatory network. Byun et al. (2022) identified CaAN3 as a key regulator of fruit-specific anthocyanin accumulation in *Capsicum annuum*, highlighting its distinct expression profile compared to CaAN2, which is more broadly expressed across different tissues due to the activation mediated by a non-LTR retrotransposon (Borovsky et al., 2004). BHLH proteins are the second largest class of transcription factors (Duan et al., 2021). Generally, the first 200 amino acids of this protein interact with the MYB factors, while the next 200 amino acids interact with the WD40 protein (D'Amelia et al., 2014). The bHLH domains are involved in the formation of homo- or heterodimers with other bHLH proteins; this appears to be a prerequisite for the recognition of binding sites on DNA, contributing to binding specificity (Montefiori et al., 2015). In *Solanaceae*, there are two main clades involved in the regulation of anthocyanin biosynthesis, which are orthologues of the *PhAN1* and *PhJAF13* genes identified in Petunia. In potato, the *StbHLH1* gene, ortholog of *PhAN1*, is widely expressed in red and purple tubers (Payyavula et al., 2013). *StbHLH1* expression alone is not associated with anthocyanin biosynthesis in the tuber and co-expression of the *StAN1* gene is required (Liu et al., 2015). In pepper, high expression levels of *CabHLH*, ortholog of *PhAN1*, were found in anthocyanin-pigmented fruits and it was positively correlated with an upregulation of regulatory genes and increased anthocyanin content (Stommel et al., 2009). WD40 proteins provide a stable platform for bHLH and MYB proteins during the establishment of the MBW complex. In *Solanaceae*, in addition to these transcriptional factors activating anthocyanins, MYB repressors are present, which reduce the biosynthesis of anthocyanins. These repressors are divided into two categories: R2R3-MYB repressors, containing a typical motif at the C-terminus, change the function of the MBW complex from activator to repressor of downstream gene transcription; R3-MYB repressors, on the other hand,

do not have a repressor motif and are unable to act directly on the target genes (Y. Liu et al., 2018). They show "passive suppression", which means that they compete with MYB activators for interaction with bHLH, thus reducing the pool of MBW complexes that can bind the promoters of genes for anthocyanin biosynthesis. An R3-MYB repressor, named StISAC, has been identified in potato (Figure 4) (D'Amelia et al., 2022; D'Amelia et al., 2020). In pepper, functional analyses of the R2-R3 MYB CaMYB101 showed its role as a repressor of anthocyanin accumulation in pepper leaves and ovaries (Y. Liu et al., 2021).

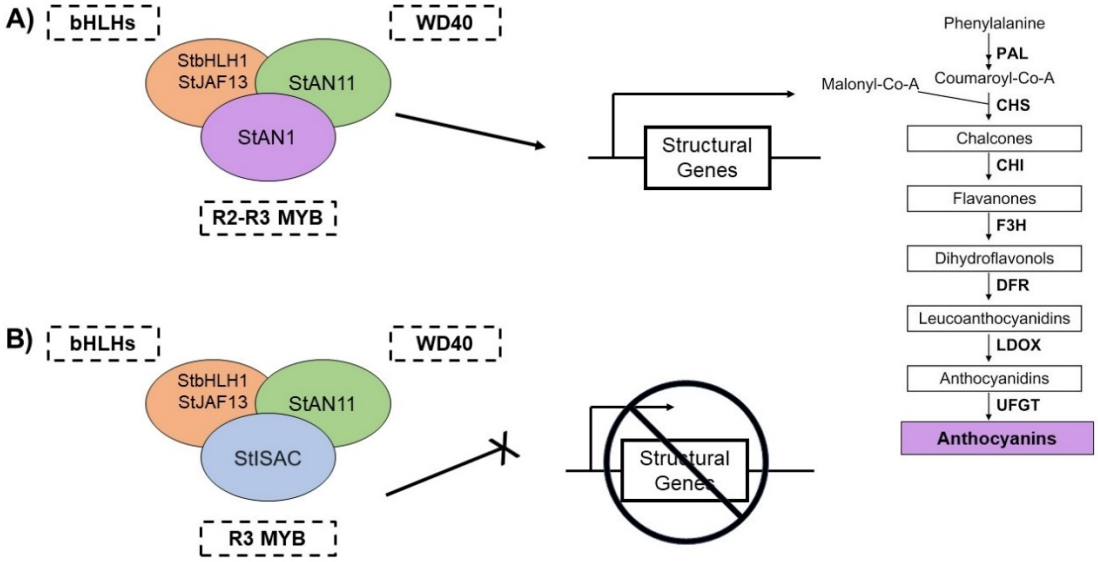

Figure 4. A simplified model outlining the regulatory mechanism of potato transcription factors involved in the MBW complex, MYB, bHLH, and WD40. They modulate the expression of structural genes of the anthocyanin biosynthetic pathway. **(A)** Active regulation of anthocyanin biosynthesis. **(B)** Repressive regulation of anthocyanin biosynthesis. MYB repressors compete with MYB activators for *St*bHLH1/*St*JAF13. "→ " means activation and "—X" means repression.

1.2.3. Mechanisms and factors influencing anthocyanins degradation

Anthocyanin biosynthesis has been investigated in great detail, with several studies focused on the manipulation of regulatory genes in transgenic plants with either increased concentration or altered composition of anthocyanins in flowers, foliage or fruit (Ben Zvi et al., 2008; Butelli et al., 2008; Li et al., 2022; Zhang et al., 2016). Nonetheless, increasing anthocyanin production does not certainly correspond to a higher yield of these pigments, due to the role played by degradation (Passeri et al., 2016). Indeed, anthocyanins frequently accumulate transiently, accumulating and disappearing throughout plant development or as environmental conditions change (Oren-Shamir, 2009). Little is known about anthocyanin catabolism, which seems to be attributed to enzymatic activity, spontaneous reactions or a combination of both (Bar-Akiva et al., 2010; Fang et al., 2015; Oren-Shamir, 2009; Zipor et al., 2015). Three different enzyme families have been suggested to be

implicated in anthocyanin degradation: peroxidase (Prx), polyphenol oxidase (PPO) and β-glycosidases (Fang et al., 2015; Oren-Shamir, 2009). Flowers turned out to be an excellent model to study anthocyanin fading. Indeed, the active enzymatic *in planta* degradation of anthocyanins was first suggested in *Brunfelsia calycina*, an ornamental *Solanacea* from South America known as *yesterday-today-tomorrow*, whose flowers change colour rapidly from dark purple to complete white after opening (Vaknin et al., 2005). Further evidence for active enzymatic anthocyanin degradation was obtained from a petunia mutant whose petal colour completely faded after bud opening (Quattrocchio et al., 2006). Two mechanisms have been suggested to explain anthocyanin degradation *in planta*. The first pathway is comprised by a two-step degradation, where anthocyanins are hydrolysed by β-glycosidase resulting in breaking the glycosidic bond and yielding highly unstable anthocyanidins that subsequently are easily degraded by polyphenol oxidase and/or peroxidase. This pathway has been suggested in blood orange and litchi fruits during their final ripe stage (Habibi et al., 2023; Zhang et al., 2001). The second pathway consists of a direct oxidation of anthocyanins mediated by Prx. For instance, a vacuolar class III peroxidase, BcPrx01, was suggested to be responsible for this anthocyanin degradation in petunia flowers in the presence of H2O2, when the glucosidase activity was completely inhibited (Zipor et al., 2015). Besides enzymatic factors, non-enzymatic factors could affect anthocyanin colour and stability, increasing the vulnerability to degradative enzymes. For example, the aforementioned decorations (as glycosylation, acylation and methylation) can affect anthocyanin characteristics. In particular, glycosylation at the C3 enhances the stability of the molecule and shifts the colour towards red. On the other hand, glycosylation at C5 reduces pigment intensity (Woodward et al., 2009). Furthermore, co-pigmentation of anthocyanins with organic compounds to form complexes results in darker and more stable pigments (Cao et al., 2023). These organic compounds are known as co-pigments, and they include a variety of alkaloids, amino acids, flavonoids, phenolic acids, tannins and even anthocyanins themselves (Tang et al., 2023). Also metal ions, particularly trivalent metal ions ($Fe^{3+} \approx Ga^{3+} > Al^{3+} > Cr^{3+}$) form complexes with the anthocyanins causing bathochromic shifts and improving their stability (Sigurdson et al., 2016). Lastly, research on anthocyanin degradation has only begun to uncover the complexity of its catabolism, making it crucial to consider all the several factors that influence this process.

1.3. Plant breeding and biotechnological methods for the design of high-value anthocyanins

Over the past decades, in response to consumer trends prioritizing health-positive foods, breeders and biotechnologists have focused on enhancing the antioxidant content of plant-derived foods and nutraceuticals (Sorrenti et al., 2023). Certainly, among the various antioxidant molecules, anthocyanins have been the most extensively studied due to their high demand and significant value

in the global market. The global anthocyanin market size was worth USD 318 million in 2020 and is expected to grow at a compound annual rate of 4.6%, to reach USD 388 million in 2026 (Market Data Forecast, 2021). Several economically interesting plant species, both food crops and ornamentals, have been obtained by traditional breeding or transgenesis with an improved anthocyanin content, mainly manipulating the biosynthetic pathway (Figure 5). Most of the knowledge on anthocyanin chemistry and genetics comes from studies on flowers belonging to model species as *Petunia* spp. (Passeri et al., 2016). It was possible to determine that, while uneven patterns as the spot-pattern is the result of an unstable transposon insertion in structural and/or regulatory genes (Spelt et al., 2000; van Houwelingen et al., 1999), the colouration of petals, anthers and tube are due to differential expression of MBW transcription complex genes (Zhang et al., 2021). Insights into anthocyanin metabolism from flower studies have laid the foundation for research in crops, where genomic tools are crucial for enhancing quality traits through increased pigmentation linked to health benefits. The use of wild relatives and the development of specific markers is undoubtedly one of the leading approaches for breeding programs. For example, wild potato relatives possess a set of traits with great potential for use in potato breeding, including a broad range of highly-pigmented tubers (Behn et al., 2023). Using the fingerprint of 742 wild and cultivated landraces, it was possible to develop an exhaustive Potato Genetic Identity (PGI) kit, used for a variety of purposes like identity verification, gene flow and linkage QTL mapping (Ghislain et al., 2009). The most important genes related to anthocyanin are located in the major QTLs named the Developer (D), Purple (P) and Red (R) loci. The molecular markers included in these loci were used to illustrate the genetic distinctiveness of the landrace “Purple Pelisse”, a high-pigmented ‘fingerling’ potato with the wild-relative *Solanum vernei* in the pedigree (Vales et al., 2012). Recent findings on anthocyanins and the potential and limits of new applications in crop breeding and management are reviewed by (Marone et al., 2022).

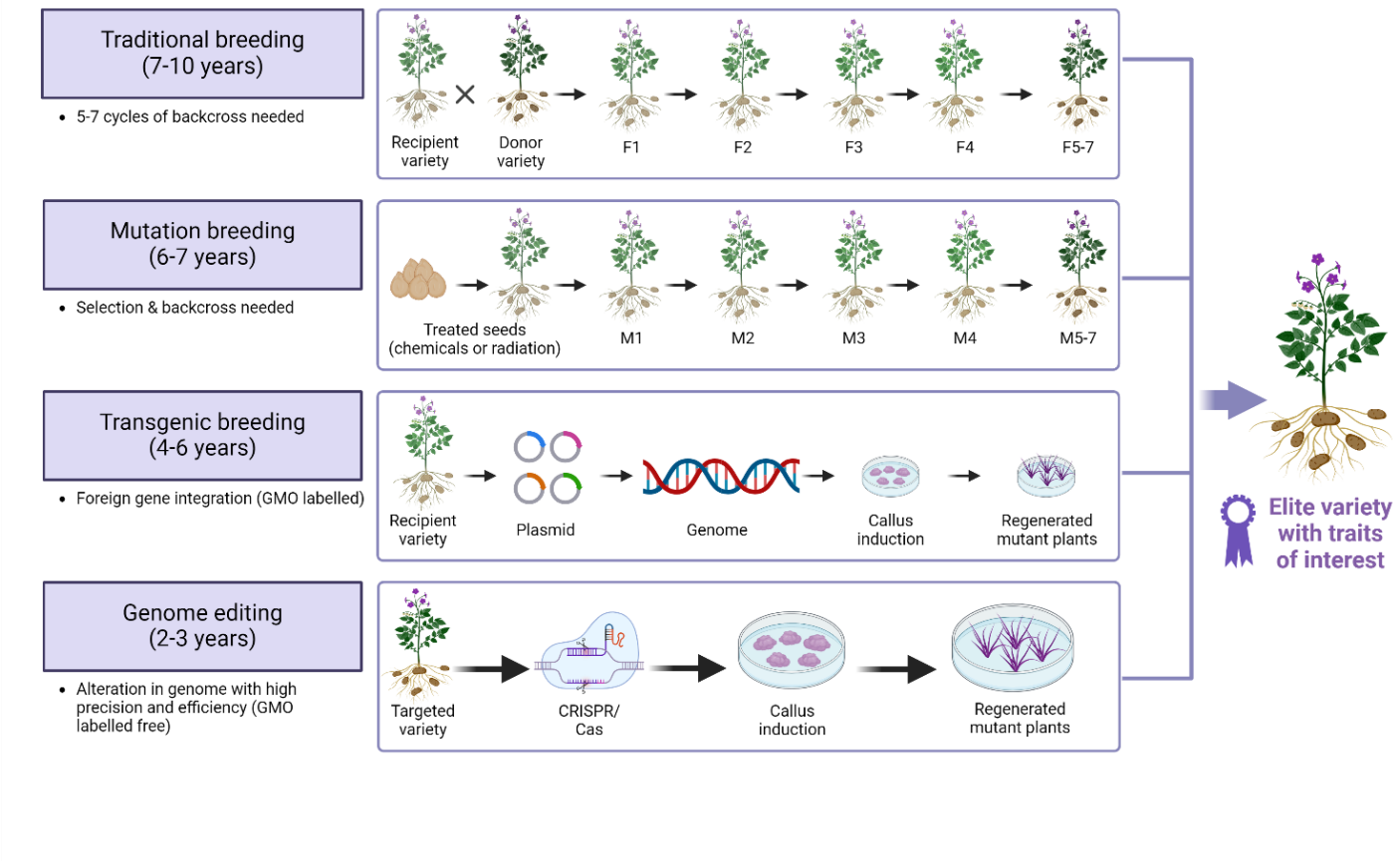

Figure 5. Comparison of the traditional breeding and modern breeding methods in potato: cross-breeding, mutation breeding, transgenic breeding and new breeding techniques. Created with BioRender.com

Besides the new avenues for crop improvement and the accelerated time-to-market of new anthocyanin-rich varieties, one of the main goals in plant biotechnology is how to obtain a stable and efficient system for anthocyanin production used for food and drug applications. As anthocyanins are found in a large variety of plant tissues, it stands to reason that one of the main approaches used for their production is through extraction and isolation from fresh matrices or from food industry by-products (Silva et al., 2017). However, many factors as seasonality, unstableness due to environmental conditions, and large space requirements for cropping often negatively impact the production of these metabolites. In addition, the composition of anthocyanins from these natural resources varies qualitatively and quantitatively, according to the genetic background, the region of origin, and the growing season of the plant (Timmers et al., 2017). In this scenario, biotechnologies may help to overcome limits and constraints, providing reliable ways to efficiently and effectively produce and extract anthocyanins. For example, microbial production by engineered microorganisms is an alternative strategy that has been described in detail by Marienhagen and Bott (2013), (Cress et al., 2017) and (Levisson et al., 2018). It combines fast growth and easy cultivation with a trouble-free

manipulation technique (Zha et al., 2020). However, microorganisms used for this purpose, mainly yeasts (as *Saccharomyces cerevisiae*) and *Escherichia coli*, require the expression of at least 11 transgenes to produce the simplest anthocyanin molecule, the pelargonidin 3-O-glucoside, starting from phenylalanine (Dudnik et al., 2017). The number of transgenes required increases with the level of decorations for the target anthocyanins; for example, more than 20 transgenes are required for acetylated anthocyanins with a blue tint (Appelhagen et al., 2018). An oldie but goodie strategy that overcomes these limits is the use of plant cell platforms. Indeed, unlike microbial cells, plant cells contain all the genetic information required to produce and store anthocyanins at the vacuole level. In the last few years, several European biotech companies as Arterra Biosciences (www.arterrabio.it) and Mibelle Biochemistry (www.mibellebiochemistry.com) developed robust plant cell cultures, obtaining a consistent amount of anthocyanins used as additives in cosmetic products. Further plant cell culture technology-derived active cosmetic ingredients, currently available on the market are reviewed by (Georgiev et al., 2018). Plant cells culture technologies have been explored also to produce anthocyanin with pharmaceutical and clinical interest. Appelhagen et al. (2018) suggested that cell cultures in stir tank bioreactors can be a suitable system for the production of ^{13}C-labelled anthocyanins, which are required for human stable-isotope tracer and bioavailability studies. To advance breeding and biotechnological efforts aimed at developing plant varieties and cell culture platforms with enhanced economic and nutraceutical value, a deeper understanding of the mechanisms governing anthocyanin production and stabilization is essential. Based on the groundwork based by previous research, this thesis specifically focuses on the regulatory genes involved in anthocyanin metabolism. By targeting these genes, our goal is to develop more stable and consistent anthocyanin products, which can overcome the limitations of current methods and lead to more reliable sources of these valuable compounds.

1.4. Aim of the thesis

As sessile organisms, plants had to develop a depot of molecular and biochemical strategies to adapt and survive to environmental changes. Thus, they evolved the ability to synthesize specific compounds (i.e. specialized metabolites) to facilitate successful interactions with the surrounding environment, as well as other organisms (Huang & Dudareva, 2023). The biosynthesis and accumulation of these metabolites are temporally and developmentally modulated by transcriptional factors, which form intricate regulatory networks that are the basis for their production. Knowing in detail all the key actors involved in these networks is essential for manipulating specialized metabolites and enhancing quality traits essential for the fitness of plants but also for human consumption. Indeed, many of these compounds possess biological activities that have been used for centuries as active natural ingredients for medicinal purposes. Among specialized metabolites,

anthocyanins are among the most significant due to their diverse functions and wide-ranging benefits. Although great progress in research on anthocyanin structures and biosynthesis has been made, the molecular regulatory mechanisms of anthocyanin metabolism remain less clear. Therefore, in recent years they have become hotspots in several genetics and biotechnology studies.

Given the importance of anthocyanins for plant and human health, and the need for research in the molecular dynamics underlying their formation, in this thesis we carried out a multifaceted study to better understand the positive and negative players and dynamics involved in the transcriptional regulation of the anthocyanin pathway. As model crops for this study we used two Solanaceae, potato and pepper, an autotetraploid (2n=4x=48) and a diploid (2n=2x=24) crop, respectively. While the potato has been utilized for genetics and transcriptional studies, pepper (due to its simpler inheritance patterns) has been included for breeding purposes. Indeed, its fruit quality is determined by multiple traits of breeding value, of which fruit colour is the most significant since plant pigments are linked to flavour, nutrition, and health. In detail, the main objectives pursued can be summarized as follows:

- To knock-out the gene Inducer Silencing of Anthocyanins in Cell culture (StISAC), a R3-MYB repressor, using CRISPR-Cas9 editing (Chapter 2). The starting material consisted of the anthocyanin-rich potato variety "Blue Star", from which both wild-type and stisac cell cultures were established. The results obtained provided evidence that mutant cell lines doubled the accumulation level and the production of these pigments was stabilized over time.

- To understand the effect of StISAC knockout on anthocyanin metabolism. Once obtained stisac culture platform, we investigated the post-biosynthetic pathway of anthocyanins, particularly focusing on Anthocyanin-O-methyltransferases (AOMTs). We carried out a comprehensive genome-wide analysis and elucidated the expression differences among four selected AOMT-encoding genes (Chapter 3). We also exploited our cell culture lines as a unique platform for elucidating the anthocyanin degradation, which is still largely unknown. We investigated this process at a transcriptome and metabolome level and found out that the increment of anthocyanins might be also due to a misregulation of the anthocyanin catabolic system (Chapter 4).

- To identify, based on previous sequencing data, a specific pepper QTL that correlates with anthocyanin retention in ripe fruits, which harbours three different TFs involved in anthocyanin metabolism. For this reason, a transient silencing (using VIGS- virus induced gene silencing method) was carried out to verify gene functions (Chapter 5). Filling the knowledge of anthocyanin metabolism regulators has an important impact on obtaining new varieties that can acquire the label of functional food.

REFERENCES

Amiryousefi, A., Hyvönen, J., & Poczai, P. (2018). The chloroplast genome sequence of bittersweet: Plastid genome structure evolution in Solanaceae. *PLoS One*, *13*(4). https://doi.org/ARTN e019606910.1371/journal.pone.0196069

Andre, C. M., Schafleitner, R., Legay, S., Lefevre, I., Aliaga, C. A. A., Nomberto, G., Hoffmann, L., Hausman, J. F., Larondelle, Y., & Evers, D. (2009). Gene expression changes related to the production of phenolic compounds in potato tubers grown under drought stress. *Phytochemistry*, *70*(9), 1107-1116. https://doi.org/10.1016/j.phytochem.2009.07.008

Appelhagen, I., Wulff-Vester, A. K., Wendell, M., Hvoslef-Eide, A. K., Russell, J., Oertel, A., Martens, S., Mock, H. P., Martin, C., & Matros, A. (2018). Colour bio-factories: Towards scale-up production of anthocyanins in plant cell cultures. *Metabolic Engineering*, *48*, 218-232. https://doi.org/10.1016/j.ymben.2018.06.004

Aza-González, C., Núñez-Palenius, H. G., & Ochoa-Alejo, N. (2012). Molecular Biology of Chili Pepper Anthocyanin Biosynthesis. *Journal of the Mexican Chemical Society*, *56*(1), 93-98. <Go to ISI>://WOS:000303853300015

Bar-Akiva, A., Ovadia, R., Rogachev, I., Bar-Or, C., Bar, E., Freiman, Z., Nissim-Levi, A., Gollop, N., Lewinsohn, E., Aharoni, A., Weiss, D., Koltai, H., & Oren-Shamir, M. (2010). Metabolic networking in Brunfelsia calycina petals after flower opening. *J Exp Bot*, *61*(5), 1393-1403. https://doi.org/10.1093/jxb/erq008

Behn, A., Lizana, C., Zapata, F., Gonzalez, A., Reyes-Diaz, M., & Fuentes, D. (2023). Phenolic and anthocyanin content characterization related to genetic diversity analysis of subsp. Chilotanum Group in southern Chile. *Frontiers in Plant Science*, *13*. https://doi.org/ARTN 104589410.3389/fpls.2022.1045894

Ben Zvi, M. M., Negre-Zakharov, F., Masci, T., Ovadis, M., Shklarman, E., Ben-Meir, H., Tzfira, T., Dudareva, N., & Vainstein, A. (2008). Interlinking showy traits: co-engineering of scent and colour biosynthesis in flowers. *Plant Biotechnol J*, *6*(4), 403-415. https://doi.org/10.1111/j.1467-7652.2008.00329.x

Bombarely, A., Moser, M., Amrad, A., Bao, M., Bapaume, L., Barry, C. S., Bliek, M., Boersma, M. R., Borghi, L., Bruggmann, R., Bucher, M., D'Agostino, N., Davies, K., Druege, U., Dudareva, N., Egea-Cortines, M., Delledonne, M., Fernandez-Pozo, N., Franken, P., . . . Kuhlemeier, C. (2016). Insight into the evolution of the Solanaceae from the parental genomes of. *Nature Plants*, *2*(6). https://doi.org/Artn 1607410.1038/Nplants.2016.74

Borovsky, Y., Oren-Shamir, M., Ovadia, R., De Jong, W., & Paran, I. (2004). The A locus that controls anthocyanin accumulation in pepper encodes a MYB transcription factor homologous to Anthocyanin2 of Petunia. *Theor Appl Genet*, *109*(1), 23-29. https://doi.org/10.1007/s00122-004-1625-9

Burmeister, A., Bondiek, S., Apel, L., Kühne, C., Hillebrand, S., & Fleischmann, P. (2011). Comparison of carotenoid and anthocyanin profiles of raw and boiled Solanum tuberosum and Solanum phureja tubers. *Journal of Food Composition and Analysis*, *24*(6), 865-872. https://doi.org/10.1016/j.jfca.2011.03.006

Butelli, E., Titta, L., Giorgio, M., Mock, H. P., Matros, A., Peterek, S., Schijlen, E. G., Hall, R. D., Bovy, A. G., Luo, J., & Martin, C. (2008). Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nat Biotechnol*, *26*(11), 1301-1308. https://doi.org/10.1038/nbt.1506

Byun, J., Kim, T. G., Lee, J. H., Li, N., Jung, S., & Kang, B. C. (2022). Identification of as a fruit-specific regulator of anthocyanin biosynthesis in pepper. *Theoretical and Applied Genetics*, *135*(7), 2197-2211. https://doi.org/10.1007/s00122-022-04106-y

Cao, Y., Xia, Q., Aniya, Chen, J., & Jin, Z. (2023). Copigmentation effect of flavonols on anthocyanins in black mulberry juice and their interaction mechanism investigation. *Food Chem*, *399*, 133927. https://doi.org/10.1016/j.foodchem.2022.133927

Carrillo, C., Kamiloglu, S., Grootaert, C., Van Camp, J., & Hendrickx, M. (2020). Co-Ingestion of Black Carrot and Strawberry. Effects on Anthocyanin Stability, Bioaccessibility and Uptake. *Foods*, *9*(11). https://doi.org/10.3390/foods9111595

Cress, B. F., Leitz, Q. D., Kim, D. C., Amore, T. D., Suzuki, J. Y., Linhardt, R. J., & Koffas, M. A. (2017). CRISPRi-mediated metabolic engineering of E. coli for O-methylated anthocyanin production. *Microb Cell Fact*, *16*(1), 10. https://doi.org/10.1186/s12934-016-0623-3

D'Amelia, V., Aversano, R., Batelli, G., Caruso, I., Moreno, M. C., Castro-Sanz, A. B., Chiaiese, P., Fasano, C., Palomba, F., & Carputo, D. (2014). High variability and interaction with basic helix-loop-helix co-factors related to anthocyanin biosynthesis in potato leaves. *Plant Journal*, *80*(3), 527-540. https://doi.org/10.1111/tpj.12653

D'Amelia, V., Staiti, A., D'Orso, F., Maisto, M., Piccolo, V., Aversano, R., & Carputo, D. (2022). Targeted mutagenesis of StISAC stabilizes the production of anthocyanins in potato cell culture. *Plant Direct*, *6*(8), e433. https://doi.org/10.1002/pld3.433

D'Amelia, V., Villano, C., Batelli, G., Cobanoglu, O., Carucci, F., Melito, S., Chessa, M., Chiaiese, P., Aversano, R., & Carputo, D. (2020). Genetic and epigenetic dynamics affecting anthocyanin biosynthesis in potato cell culture. *Plant Sci*, *298*, 110597. https://doi.org/10.1016/j.plantsci.2020.110597

De Masi, L., Bontempo, P., Rigano, D., Stiuso, P., Carafa, V., Nebbioso, A., Piacente, S., Montoro, P., Aversano, R., D'Amelia, V., Carputo, D., & Altucci, L. (2020). Comparative Phytochemical Characterization, Genetic Profile, and Antiproliferative Activity of Polyphenol-Rich Extracts from Pigmented Tubers of Different Solanum tuberosum Varieties. *Molecules*, *25*(1). https://doi.org/10.3390/molecules25010233

Dixon, R. A., & Dickinson, A. J. (2024). A century of studying plant secondary metabolism—from "what?" to "where, how, and why?". *Plant Physiology*, *195*(1), 48-66.

Duan, Z. F., Tian, S. Y., Yang, G. B., Wei, M., Li, J., & Yang, F. J. (2021). The Basic Helix-Loop-Helix Transcription Factor SmbHLH1 Represses Anthocyanin Biosynthesis in Eggplant. *Frontiers in Plant Science*, *12*. https://doi.org/ARTN 75793610.3389/fpls.2021.757936

Dudnik, A., Almeida, A. F., Andrade, R., Avila, B., Bañados, P., Barbay, D., Bassard, J.-E., Benkoulouche, M., Bott, M., Braga, A., Breitel, D., Brennan, R., Bulteau, L., Chanforan, C., Costa, I., Costa, R. S., Doostmohammadi, M., Faria, N., Feng, C., Forster, J. (2017). BacHBerry: BACterial Hosts for production of Bioactive phenolics from bERRY fruits. *Phytochemistry Reviews*, *17*(2), 291-326. https://doi.org/10.1007/s11101-017-9532-2

Fang, F., Zhang, X. L., Luo, H. H., Zhou, J. J., Gong, Y. H., Li, W. J., Shi, Z. W., He, Q., Wu, Q., Li, L., Jiang, L. L., Cai, Z. G., Oren-Shamir, M., Zhang, Z. Q., & Pang, X. Q. (2015). An Intracellular Laccase Is Responsible for Epicatechin-Mediated Anthocyanin Degradation in Litchi Fruit Pericarp. *Plant Physiol*, *169*(4), 2391-2408. https://doi.org/10.1104/pp.15.00359

Fei, P., Zeng, F. S., Zheng, S. Y., Chen, Q. L., Hu, Y. H., & Cai, J. (2021). Acylation of blueberry anthocyanins with maleic acid: Improvement of the stability and its application potential in intelligent color indicator packing materials. *Dyes and Pigments*, *184*. https://doi.org/ARTN 10885210.1016/j.dyepig.2020.108852

Fujino, N., Tenma, N., Waki, T., Ito, K., Komatsuzaki, Y., Sugiyama, K., Yamazaki, T., Yoshida, S., Hatayama, M., Yamashita, S., Tanaka, Y., Motohashi, R., Denessiouk, K., Takahashi, S., & Nakayama, T. (2018).

Physical interactions among flavonoid enzymes in snapdragon and torenia reveal the diversity in the flavonoid metabolon organization of different plant species. *Plant J*, *94*(2), 372-392. https://doi.org/10.1111/tpj.13864

Gebhardt, C. (2016). The historical role of species from the Solanaceae plant family in genetic research. *Theor Appl Genet*, *129*(12), 2281-2294. https://doi.org/10.1007/s00122-016-2804-1

Georgiev, V., Slavov, A., Vasileva, I., & Pavlov, A. (2018). Plant cell culture as emerging technology for production of active cosmetic ingredients. *Eng Life Sci*, *18*(11), 779-798. https://doi.org/10.1002/elsc.201800066

Ghislain, M., Nunez, J., Herrera, M. R., Pignataro, J., Guzman, F., Bonierbale, M., & Spooner, D. M. (2009). Robust and highly informative microsatellite-based genetic identity kit for potato. *Molecular Breeding*, *23*(3), 377-388. https://doi.org/10.1007/s11032-008-9240-0

Habibi, F., García-Pastor, M. E., Puente-Moreno, J., Garrido-Auñón, F., Serrano, M., & Valero, D. (2023). Anthocyanin in blood oranges: a review on postharvest approaches for its enhancement and preservation. *Critical Reviews in Food Science and Nutrition*, *63*(33), 12089-12101. https://doi.org/10.1080/10408398.2022.2098250

Heim, M. A., Jakoby, M., Werber, M., Martin, C., Weisshaar, B., & Bailey, P. C. (2003). The basic helix-loop-helix transcription factor family in plants: A genome-wide study of protein structure and functional diversity. *Molecular Biology and Evolution*, *20*(5), 735-747. https://doi.org/10.1093/molbev/msg088

Jung, C. S., Griffiths, H. M., De Jong, D. M., Cheng, S., Bodis, M., Kim, T. S., & De Jong, W. S. (2009). The potato developer (D) locus encodes an R2R3 MYB transcription factor that regulates expression of multiple anthocyanin structural genes in tuber skin. *Theor Appl Genet*, *120*(1), 45-57. https://doi.org/10.1007/s00122-009-1158-3

Kiferle, C., Fantini, E., Bassolino, L., Povero, G., Spelt, C., Buti, S., Giuliano, G., Quattrocchio, F., Koes, R., Perata, P., & Gonzali, S. (2015). Tomato R2R3-MYB Proteins SlANT1 and SlAN2: Same Protein Activity, Different Roles. *PLoS One*, *10*(8). https://doi.org/ARTN e013636510.1371/journal.pone.0136365

Koes, R., Verweij, W., & Quattrocchio, F. (2005). Flavonoids: a colorful model for the regulation and evolution of biochemical pathways. *Trends in Plant Science*, *10*(5), 236-242. https://doi.org/10.1016/j.tplants.2005.03.002

Kwakye, G. F., Jiménez, J., Jiménez, J. A., & Aschner, M. (2018). neurotoxicity: Implications to neurological disorders. *Food and Chemical Toxicology*, *116*, 346-353. https://doi.org/10.1016/j.fct.2018.04.022

Lachman, J., Hamouz, K., Šulc, M., Orsák, M., Pivec, V., Hejtmánková, A., Dvořák, P., & Čepl, J. (2009). Cultivar differences of total anthocyanins and anthocyanidins in red and purple-fleshed potatoes and their relation to antioxidant activity. *Food Chemistry*, *114*(3), 836-843. https://doi.org/10.1016/j.foodchem.2008.10.029

Levisson, M., Patinios, C., Hein, S., de Groot, P. A., Daran, J. M., Hall, R. D., Martens, S., & Beekwilder, J. (2018). Engineering de novo anthocyanin production in. *Microbial Cell Factories*, *17*. https://doi.org/ARTN 10310.1186/s12934-018-0951-6

Li, C., Yang, K., Yang, J., Wu, H., Chen, H., Wu, Q., & Zhao, H. (2022). Tartary buckwheat FtF3'H1 as a metabolic branch switch to increase anthocyanin content in transgenic plant. *Front Plant Sci*, *13*, 959698. https://doi.org/10.3389/fpls.2022.959698

Liu, F., Yang, Y., Gao, J., Ma, C., & Bi, Y. (2018). A comparative transcriptome analysis of a wild purple potato and its red mutant provides insight into the mechanism of anthocyanin transformation. *PLoS One*, *13*(1), e0191406. https://doi.org/10.1371/journal.pone.0191406

Liu, J., Zhou, H., Song, L., Yang, Z., Qiu, M., Wang, J., & Shi, S. (2021). Anthocyanins: Promising Natural Products with Diverse Pharmacological Activities. *Molecules*, *26*(13). https://doi.org/10.3390/molecules26133807

Liu, Y., Lin-Wang, K., Deng, C., Warran, B., Wang, L., Yu, B., Yang, H., Wang, J., Espley, R. V., Zhang, J., Wang, D., & Allan, A. C. (2015). Comparative Transcriptome Analysis of White and Purple Potato to Identify Genes Involved in Anthocyanin Biosynthesis. *PLoS One*, *10*(6), e0129148. https://doi.org/10.1371/journal.pone.0129148

Liu, Y., Tikunov, Y., Schouten, R. E., Marcelis, L. F. M., Visser, R. G. F., & Bovy, A. (2018). Anthocyanin Biosynthesis and Degradation Mechanisms in Solanaceous Vegetables: A Review. *Frontiers in Chemistry*, *6*. https://doi.org/10.3389/fchem.2018.00052

Liu, Y., Wu, Y., Wang, Z., Zhang, S., Liu, X., Tikunov, Y., Visser, R. G. F., Schouten, R. E., Marcelis, L. F. M., Zhang, Z., & Bovy, A. (2021). *Genome-wide characterization of the R2R3-MYB transcription factors in pepper (Capsicum spp.) unveils the role of CaMYB101 as repressor in anthocyanin biosynthesis*. Cold Spring Harbor Laboratory. https://dx.doi.org/10.1101/2021.08.24.457473

Marienhagen, J., & Bott, M. (2013). Metabolic engineering of microorganisms for the synthesis of plant natural products. *Journal of Biotechnology*, *163*(2), 166-178. https://doi.org/10.1016/j.jbiotec.2012.06.001

Marone, D., Mastrangelo, A. M., Borrelli, G. M., Mores, A., Laidò, G., Russo, M. A., & Ficco, D. B. M. (2022). Specialized metabolites: Physiological and biochemical role in stress resistance, strategies to improve their accumulation, and new applications in crop breeding and management. *Plant Physiology and Biochemistry*, *172*, 48-55. https://doi.org/10.1016/j.plaphy.2021.12.037

Mekapogu, M., Vasamsetti, B. M. K., Kwon, O. K., Ahn, M. S., Lim, S. H., & Jung, J. A. (2020). Anthocyanins in Floral Colors: Biosynthesis and Regulation in Chrysanthemum Flowers. *International Journal of Molecular Sciences*, *21*(18). https://doi.org/ARTN 653710.3390/ijms21186537

Montefiori, M., Brendolise, C., Dare, A. P., Lin-Wang, K., Davies, K. M., Hellens, R. P., & Allan, A. C. (2015). In the Solanaceae, a hierarchy of bHLHs confer distinct target specificity to the anthocyanin regulatory complex. *J Exp Bot*, *66*(5), 1427-1436. https://doi.org/10.1093/jxb/eru494

Motti, R. (2021). The Solanaceae Family: Botanical Features and Diversity. In (pp. 1-9). Springer International Publishing. https://doi.org/10.1007/978-3-030-30343-3_1

Olmstead, R. G., Bohs, L., Migid, H. A., Santiago-Valentin, E., Garcia, V. F., & Collier, S. M. (2008). A molecular phylogeny of the Solanaceae. *TAXON*, *57*(4), 1159-1181. https://doi.org/10.1002/tax.574010

Oren-Shamir, M. (2009). Does anthocyanin degradation play a significant role in determining pigment concentration in plants? *Plant Science*, *177*(4), 310-316. https://doi.org/10.1016/j.plantsci.2009.06.015

Passeri, V., Koes, R., & Quattrocchio, F. M. (2016). New Challenges for the Design of High Value Plant Products: Stabilization of Anthocyanins in Plant Vacuoles. *Front Plant Sci*, *7*, 153. https://doi.org/10.3389/fpls.2016.00153

Payyavula, R. S., Singh, R. K., & Navarre, D. A. (2013). Transcription factors, sucrose, and sucrose metabolic genes interact to regulate potato phenylpropanoid metabolism. *J Exp Bot*, *64*(16), 5115-5131. https://doi.org/10.1093/jxb/ert303

Quattrocchio, F., Verweij, W., Kroon, A., Spelt, C., Mol, J., & Koes, R. (2006). PH4 of Petunia is an R2R3 MYB protein that activates vacuolar acidification through interactions with basic-helix-loop-helix transcription factors of the anthocyanin pathway. *Plant Cell*, *18*(5), 1274-1291. https://doi.org/10.1105/tpc.105.034041

Saini, R. K., Khan, M. I., Shang, X. M., Kumar, V., Kumari, V., Kesarwani, A., & Ko, E. Y. (2024). Dietary Sources, Stabilization, Health Benefits, and Industrial Application of Anthocyanins-A Review. *Foods*, *13*(8). https://doi.org/ARTN 122710.3390/foods13081227

Sendri, N., & Bhandari, P. (2024). Anthocyanins: a comprehensive review on biosynthesis, structural diversity, and industrial applications. *Phytochemistry Reviews*, 1-62.

Shen, Y., Rao, Y., Ma, M., Li, Y., He, Y., Wang, Z., Liang, M., & Ning, G. (2024). Coordination among flower pigments, scents and pollinators in ornamental plants. *Horticulture Advances*, *2*(1), 6.

Sigurdson, G. T., Robbins, R. J., Collins, T. M., & Giusti, M. M. (2016). Evaluating the role of metal ions in the bathochromic and hyperchromic responses of cyanidin derivatives in acidic and alkaline pH. *Food Chemistry*, *208*, 26-34. https://doi.org/10.1016/j.foodchem.2016.03.109

Silva, S., Costa, E. M., Calhau, C., Morais, R. M., & Pintado, M. E. (2017). Anthocyanin extraction from plant tissues: A review. *Critical Reviews in Food Science and Nutrition*, *57*(14), 3072-3083. https://doi.org/10.1080/10408398.2015.1087963

Sorrenti, V., Buro, I., Consoli, V., & Vanella, L. (2023). Recent Advances in Health Benefits of Bioactive Compounds from Food Wastes and By-Products: Biochemical Aspects. *Int J Mol Sci*, *24*(3). https://doi.org/10.3390/ijms24032019

Spelt, C., Quattrocchio, F., Mol, J. N. M., & Koes, R. (2000). of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell*, *12*(9), 1619-1631. https://doi.org/DOI 10.1105/tpc.12.9.1619

Staiti, A., D'Amelia, V., & Carputo, D. (2022). Exploring and exploiting anthocyanins for human needs: potato as a research case study. *Italus Hortus*, *29*(2), 18-34. https://doi.org/10.26353/j.itahort/2022.2.1834

Stommel, J. R., Lightbourn, G. J., Winkel, B. S., & Griesbach, R. J. (2009). Transcription Factor Families Regulate the Anthocyanin Biosynthetic Pathway in. *Journal of the American Society for Horticultural Science*, *134*(2), 244-251. https://doi.org/Doi 10.21273/Jashs.134.2.244

Strygina, K. V., Kochetov, A. V., & Khlestkina, E. K. (2019). Genetic control of anthocyanin pigmentation of potato tissues. *BMC Genet*, *20*(Suppl 1), 27. https://doi.org/10.1186/s12863-019-0728-x

Stushnoff, C., Ducreux, L. J., Hancock, R. D., Hedley, P. E., Holm, D. G., McDougall, G. J., McNicol, J. W., Morris, J., Morris, W. L., Sungurtas, J. A., Verrall, S. R., Zuber, T., & Taylor, M. A. (2010). Flavonoid profiling and transcriptome analysis reveals new gene-metabolite correlations in tubers of Solanum tuberosum L. *J Exp Bot*, *61*(4), 1225-1238. https://doi.org/10.1093/jxb/erp394

Tang, B. Y., Li, L., Hu, Z. L., Chen, Y. N., Tan, T. T., Jia, Y. H., Xie, Q. L., & Chen, G. P. (2020). Anthocyanin Accumulation and Transcriptional Regulation of Anthocyanin Biosynthesis in Purple Pepper. *Journal of Agricultural and Food Chemistry*, *68*(43), 12152-12163. https://doi.org/10.1021/acs.jafc.0c02460

Tang, R., He, Y., & Fan, K. (2023). Recent advances in stability improvement of anthocyanins by efficient methods and its application in food intelligent packaging: A review. *Food Bioscience*, *56*, 103164. https://doi.org/10.1016/j.fbio.2023.103164

Tengkun, N., Dongdong, W., Xiaohui, M., Yue, C., & Qin, C. (2019). Analysis of Key Genes Involved in Potato Anthocyanin Biosynthesis Based on Genomics and Transcriptomics Data. *Front Plant Sci*, *10*, 603. https://doi.org/10.3389/fpls.2019.00603

Timmers, M. A., Grace, M. H., Yousef, G. G., & Lila, M. A. (2017). Inter- and intra-seasonal changes in anthocyanin accumulation and global metabolite profiling of six blueberry genotypes. *Journal of Food Composition and Analysis*, *59*, 105-110. https://doi.org/10.1016/j.jfca.2017.02.019

Vaknin, H., Bar-Akiva, A., Ovadia, R., Nissim-Levi, A., Forer, I., Weiss, D., & Oren-Shamir, M. (2005). Active anthocyanin degradation in Brunfelsia calycina (yesterday--today--tomorrow) flowers. *Planta*, *222*(1), 19-26. https://doi.org/10.1007/s00425-005-1509-5

Vales, M. I., Brown, C. R., Yilma, S., Hane, D. C., James, S. R., Shock, C. C., Charlton, B. A., Karaagac, E., Mosley, A. R., Culp, D., Feibert, E., Stark, J. C., Pavek, M. J., Knowles, N. R., Novy, R. G., & Whitworth, J. L. (2012). Purple Pelisse: A Specialty 'Fingerling' Potato with Purple Skin and Flesh and Medium Specific Gravity. *American Journal of Potato Research*, *89*(4), 306-314. https://doi.org/10.1007/s12230-012-9254-3

van Houwelingen, A., Souer, E., Mol, J., & Koes, R. (1999). Epigenetic interactions among three transposons in two homologous chromosomes activate a new excision-repair mechanism in petunia. *Plant Cell*, *11*(7), 1319-1336. <Go to ISI>://WOS:000081674300012

Vidana Gamage, G. C., Lim, Y. Y., & Choo, W. S. (2022). Sources and relative stabilities of acylated and nonacylated anthocyanins in beverage systems. *J Food Sci Technol*, *59*(3), 831-845. https://doi.org/10.1007/s13197-021-05054-z

Wadhera, D., & Capaldi-Phillips, E. D. (2014). A review of visual cues associated with food on food acceptance and consumption. *Eat Behav*, *15*(1), 132-143. https://doi.org/10.1016/j.eatbeh.2013.11.003

Woodward, G., Kroon, P., Cassidy, A., & Kay, C. (2009). Anthocyanin stability and recovery: implications for the analysis of clinical and experimental samples. *J Agric Food Chem*, *57*(12), 5271-5278. https://doi.org/10.1021/jf900602b

Zha, J., Wu, X., & Koffas, M. A. G. (2020). Making brilliant colors by microorganisms. *Current Opinion in Biotechnology*, *61*, 135-141. https://doi.org/10.1016/j.copbio.2019.12.020

Zhang, B., Xu, X. J., Huang, R. W., Yang, S., Li, M. Y., & Guo, Y. L. (2021). CRISPR/Cas9-mediated targeted mutation reveals a role for rather than in regulating venation formation in the corolla tube of. *Horticulture Research*, *8*(1). https://doi.org/ARTN 11610.1038/s41438-021-00555-6

Zhang, Y., Chu, G., Hu, Z., Gao, Q., Cui, B., Tian, S., Wang, B., & Chen, G. (2016). Genetically engineered anthocyanin pathway for high health-promoting pigment production in eggplant. *Molecular Breeding*, *36*(5). https://doi.org/10.1007/s11032-016-0454-2

Zhang, Z., Pang, X., Ji, Z., & Jiang, Y. (2001). Role of anthocyanin degradation in litchi pericarp browning. *Food Chemistry*, *75*(2), 217-221. https://doi.org/10.1016/s0308-8146(01)00202-3

Zhao, C. L., Yu, Y. Q., Chen, Z. J., Wen, G. S., Wei, F. G., Zheng, Q., Wang, C. D., & Xiao, X. L. (2017). Stability-increasing effects of anthocyanin glycosyl acylation. *Food Chem*, *214*, 119-128. https://doi.org/10.1016/j.foodchem.2016.07.073

Zhou, Y., Mumtaz, M. A., Zhang, Y., Shu, H., Hao, Y., Lu, X., Cheng, S., Zhu, G., & Wang, Z. (2022). Response of Anthocyanin Accumulation in Pepper (Capsicum annuum) Fruit to Light Days. *Int J Mol Sci*, *23*(15). https://doi.org/10.3390/ijms23158357

Zipor, G., Duarte, P., Carqueijeiro, I., Shahar, L., Ovadia, R., Teper-Bamnolker, P., Eshel, D., Levin, Y., Doron-Faigenboim, A., Sottomayor, M., & Oren-Shamir, M. (2015). In planta anthocyanin degradation by a vacuolar class III peroxidase in Brunfelsia calycina flowers. *New Phytol*, *205*(2), 653-665. https://doi.org/10.1111/nph.13038

Published paper front-page

ITALUS HORTUS
Review

Vol. 29 (2022), Pages 18-34
doi: 10.26353/j.itahort/2022.2.1834

Exploring and exploiting anthocyanins for human needs: potato as a research case study

Annalisa Staiti[1], Vincenzo D'Amelia[2*], Domenico Carputo[1]

[1] Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, Italy
[2] National Research Council (CNR), Institute of Biosciences and Bioresources (IBBR), Via Università 133, 80055 Portici, Italy
[*] Corresponding author: vincenzo.damelia@ibbr.cnr.it; Tel.; +39-081-2539587

Received: 20 June 2022; Accepted: 7 September; Published: 21 September 2022

Abstract: Anthocyanins are a group of polyphenolic water-soluble pigments, largely distributed in the plant kingdom. Being natural pigments with a health-promoting bioactivity, there is an increasing interest for their application in food and drugs industry. The increasing in anthocyanins production is an enticing achievement of industries. In the first part of this review, we highlight key concepts related to the biochemistry, biological function and genetics of these important pigments in the potato. We chose this crop because it displays a valuable anthocyanin chemical diversity and it is also highly amenable to various biotechnological applications. In this latter regard, we present and briefly discuss the potential of cell cultures as a suitable method for the production of highly decorated anthocyanins. The final message of this review is to underline the impact of bio-factories as a customizable and sustainable strategy for anthocyanin production and their high potential in industrial and medical applications.

Keywords: Metabolic engineering, plant cell cultures, MYB transcription factors, flavonoids biosynthesis, *Solanum tuberosum*.

1. Anthocyanins: an overture on the biotechnological methods for their production

Plant secondary metabolites are a large group of chemical compounds which are distributed in almost all plant organs. Despite their definition as "secondary", these molecules play a fundamental role in plant development, propagation and survival. For example, the coloration of flowers and fruits helps to attract pollinators and seed dispersers and is crucial for plant reproduction under natural conditions (Davies et al., 2012; Landi et al., 2015). For their important role in plant physiology today the term "specialized" is preferred over secondary metabolites (Pichersky et al., 2011). Among specialized metabolites, anthocyanins (from Greek ανθός (anthos) '= flower and κυανός (kyanos) '= blue), belonging to flavonoids family, are intensely studied for their antimicrobial, antioxidant, and free radical scavenging properties that have applications in food technology, human health and plant protection (D'Amelia et al., 2018). They accumulate in a wide array of plant families including *Rosaceae* (e.g. *Fragaria vesca* and *Rubus fruticosus*), *Moraceae* (e.g. *Ficus carica* and *Morus nigra*), *Brassicaceae* (e.g. *Brassica oleracea* and *Raphanus sativus*), *Poaceae* (e.g. *Oryza sativa* and *Zea mays*), *Solanaceae* (e.g. *Solanum melongena*, *Capsicum annuum,* and *Solanum tuberosum*). As anthocyanins are found in a large variety of plant tissues, it stands to reason that one of the main approaches used for their production is through extraction and isolation from fresh matrices or from food industry by-products (Silva et al., 2017). However, many factors as seasonality, unstableness due to environmental conditions, and large space requirements for cropping often negatively impact the production of these metabolites. In this scenario, biotechnologies may help to overcome limits and constraints, providing reliable ways to efficiently and effectively produce and extract anthocyanins. For example, microbial production by

Chapter 2. Targeted mutagenesis of StISAC stabilizes the production of anthocyanins in potato cell culture

D'Amelia, V., Staiti, A., D'Orso, F., Maisto, M., Piccolo, V., Aversano, R., Carputo, D.

Published: 15 July 2022, *Plant Direct*

Chapter 2. Targeted mutagenesis of StISAC stabilizes the production of anthocyanins in potato cell culture

ABSTRACT

To increase the production of complex anthocyanins in potato cell cultures, we knocked out a novel potato gene, named *Inducer Silencing of Anthocyanins in Cell culture* (*StISAC*) using CRISPR/Cas9 editing. Our results provided evidence that mutant cell lines doubled the accumulation level of anthocyanins biosynthesized. Moreover, the production of these important pigments was stabilized over time. Our study accomplished overcame challenges in the efficient biotechnological production of these valuable pigments and reported the function of a novel potato anthocyanin repressor gene.

2.1 Introduction

Anthocyanins are flavonoid pigments mostly responsible for the red, blue and purple colours of plant tissues. They are used as coloured additives and bioactive ingredients in processed foods (Dong and Lin, 2020; Neri-Numa et al., 2020). Among plant species, the cultivated potato Solanum tuberosum offers a wide and unique range of anthocyanins that can be extracted from pigmented tubers (Oertel et al., 2017). According to a report from the market data forecast (www.marketdataforecast.com), the market size of anthocyanins was estimated at 318 million USD in 2020 and is expected to increase at a compound annual rate of 4.6%. Since direct extraction from plant matrices does not guarantee a sustainable and efficient production on demand, innovative methodologies and technologies can be used. Among them, plant cell cultures represent a very versatile and powerful system for metabolite production. The main drawback is that cell cultures are quite unstable in producing anthocyanins over time. This has a strong negative impact on their production in bioreactors (Appelhagen et al., 2018; D'Amelia et al., 2020). CRISPR/Cas9-based gene editing is the emerging technique that can promote the accumulation of anthocyanin in plants (Yan et al., 2020). In the study presented here, we exploited CRISPR/Cas9 technology to generate an edited potato cell culture line that stably accumulates a high level of complex anthocyanins. With this in mind, we undertook a two-stage study. In the first stage, we pointed out StISAC as a novel promising target gene encoding for a negative R3-MYB regulator of anthocyanin biosynthesis. Then, we knocked out StISAC by means CRISPR/Cas9 induced mutagenesis. The obtained edited cells showed not only an increased amount of anthocyanins, but a stable production of these pigments during successive subcultures.

2.2 Materials and Methods

2.2.1 Generation of potato calli and suspension cell culture

Potato calli were obtained from leaf explants of the potato variety "Blue Star" (2n = 4x = 48) as described in D'Amelia et al. (2020). Briefly, callus formation was induced in young leaves and internodes for approximately 15–20 days on MS agar plates supplemented with 3 mg/L 2,4-dichlorophenoxyacetic acid (2,4-D, Sigma, St. Louis, USA) and 30 g/L sucrose; pH 5.8. Calli were cultured at 24°C in the dark and were subcultured every 2 weeks to a fresh medium. Cell suspension cultures were initiated by culturing approximately 1 g callus in 250-Erlenmeyer's flasks. Cell suspension cultures were maintained under continuous agitation at 110 rpm on a shaker (Infors AG, CH-4103, Switzerland) and incubated under the same culture conditions as described for the induction of callus culture.

2.2.2 Vector construction

The final CRISPR construct was developed by using the Golden Gate cloning method (Engler et al., 2008). Two sgRNAs (Sg1 and Sg2) were CCT-PAM1 and CCC-PAM2. These chosen sequences (5'-TCATAGACTCCCCTTGTAAG-PAM1-3' and 5'-CTTGTAAGTTCTTAATTTTC–PAM2-3') were assembled by PCR to a sgRNA backbone contained in pICSL70001 plasmid, which has been used as an amplification template. The two different amplicons were assembled, respectively, in vectors pICH47732 and pICH47742 along with pU6 promoter contained in pICSL90001. The two obtained vectors containing the expression cassettes pU6:sgRNA were used in a Golden Gate digestion-ligation reaction to make the final vector in pICSL002203 backbone containing the Cas9 driven by the 2 _ 35S CaMV and Kanamycin selection cassette according to the Golden Gate methodology. All the assembly steps to obtain a complete transcription unit was obtained by means Golden Gate digestion-ligation reactions (http://synbio.tsl.ac.uk/). The final vector was expressed in potato cells via Agrobacterium tumefaciens strain LBA4404-pAL4404 by electroporation.

2.2.3 Potato transformation

A. tumefaciens grew in YEP broth supplemented with 100 mg L^{-1} of streptomycin, rifampicin, and kanamycin at 28°C. An overnight culture was centrifuged at 7000 rpm for 10 min. The pellet was resuspended in callus culture liquid medium with the addition of 100 μM acetosyringone (Sigma, St. Louis, USA) and diluted to OD600 = 0.7. Potato explants were infected by immerging in *A. tumefaciens* inoculum for 20 min. Then, explants were blotted on a sterile filter paper and placed on a solidified callus induction medium, followed by 2-day cocultivation at 26°C in the dark. Subsequently, for Agrobacterium elimination, explants were rinsed with sterile water containing 400 mg L^{-1} cefotaxime and were transferred to a fresh callus culture medium supplemented 100 mg L^{-1} kanamycin for transformed callus selection. Single, small calli developed after 3 weeks of selection were picked up and transferred periodically to the same selection medium until whole callus tissues were obtained.

2.2.4 High-resolution melting analysis, amplicon sequencing and CRISPR mediated indels detection

Genomic DNA was extracted from five calli using the DNeasy Plant Mini Kit (Qiagen, www.qiagen.com). A nucleotide region flanking the target region of sgRNAs was amplified by PCR and sequenced (Eurofins, www.eurofins.com) using Mix2Seq kit. To identify indels directly from Sanger sequencing traces without sub-cloning, we used the web tool CRISPR-ID. Reads were manually aligned to the reference sequence using the Geneious software to detect induced

mutations. HRM analysis was conducted with Rotor-Gene™ 6000® (Corbett, Life Science). Reaction was performed in of 20 µl containing 10 ng of gDNA, 0.4 µM of each primer, 0.8 µM dNTPs, Phusion™ High-Fidelity DNA Polymerase (Thermo Scientific), 1x PCR buffer, and 1x Eva- Green Dye (Biotium). PCR products (149 bp) were melted from 65 to 90°C with a melting rate of 0.1°C $^{s-1}$. All the HRM data collected were analyzed by the dedicated HRM software (Rotor-Gene 6000

Series Software 1.7).

2.2.5 Spectrophotometric quantification

Five independent pieces of edited and wild-type calli were used. Five hundred milligrams of calli were dispersed in ethanol/water blend (70:30 v/v, 25 ml). Then, an ultrasonic extraction was carried out for 30 min at the power of 50 W. After centrifugation (15 min, 5000 rpm), the extractive solution (supernatant) was evaporated under reduced pressure at 40C to remove ethanol and then freeze-dried, producing a powder stored under vacuum until use. The pH differential method was employed to quantify the total content of monomeric anthocyanins from cell culture.3 Anthocyanin profiles of purple potato extracts were obtained using high-performance liquid chromatography coupled to mass spectrometry with a heated electrospray ionization interface (HPLC-HESI-MS). LTQ XL mass spectrometer (Thermo Fisher Scientific) equipment, coupled with HPLC DIONEX UltiMate 3000 (Thermo Fisher Scientific) instrument, were used for the analysis. The samples' analytes were separated using a C18 reversed-phase column (250 x 4.6 mm, 5 µm, Phenomenex). The separation conditions were as follows: Column temperature was set at 30°C, inject volume was 5 µl, and the solvent flow rate was set at 1 ml/min. The mobile phase consisted of A (formic acid/water at the ratio of 0.3/99.7 v/v) and B (formic acid/acetonitrile/water at the ratio of 0.3/30/69.7 v/v/v). Elution gradient was as follows: 0 min: 20% (B), 3 min: 20% (B), 10 min: 30% (B), 40 min: 90% (B), 45 min: 90% (B), 45.1 min: 30% (B), 53 min: 30% (B). The source was a heated electrospray interface (HESI), operated in positive ionization mode using the following parameters: sheath gas flow rate: 35 ml/min; auxiliary gas flow rate: 10 ml/min; capillary temperature: 320°C; source heater temperature: 150°C; source voltage 3.5 kV; source current 100 µA; capillary voltage: 32 V; tube lens: 80 V. Data acquisition was performed using X-Calibur software and HPLC system was controlled by Chromeleon software.

2.3 Results and Discussion

2.3.1 Re-annotation of *StMYBATV* in *StISAC*

In D'Amelia et al. (2020), we isolated a R3-MYB transcription factor we named StMYBATV due to its phylogenetic relatedness to the tomato anthocyanin repressor SlMYBATV (MF197509). Based

on the potato genome annotation (v. 6.1., DM 1-3 516 R44), we found that StMYBATV localized on chromosome 12 (*StMYBATV-chr12*; Soltu.DM.12G023200), whereas *SlMYBATV* is placed on tomato chromosome 7. These data indicated that StMYBATV is not the direct ortholog of SlMYBATV, given the conserved genomic synteny between the two Solanaceae (Peters et al., 2012). Instead, StMYBATV-chr12 is the direct ortholog of SlMYBATV-like (MF197514), another tomato R3-MYB encoding gene. It is localized on chromosome 12 (Cao et al. 2017), but it has not been functionally characterized yet. The phylogenetic analysis of the encoded proteins (Figure 1a) indicated that the tomato the tomato SlMYBATV (with its different alleles, SlMYBATV-X1, SlMYBATV-X2, SlMYBATV-X3; Cao et al., 2017) grouped (cluster IIa) with a potato protein encoded by a gene placed on chr07 (StMYBATV-chr07, Soltu.DM.07G017910), while StMYBATV-chr12 (StMYBATV in D'Amelia et al., 2020) and SlMYBATV-like grouped apart (cluster IIb). A signature motif Asp-Ser-Thr-Arg-Val-Val within the R3 domain contradistinguished the two classes of proteins (Ia and IIb). Indeed, it was present only in proteins of cluster IIa (Figure 1b). Differences in exon-intron structure suggest possible distinct evolutionary histories among these MYBs (Figure 1c), where the protein encoding genes of cluster IIa have four exons rather than the three found in the cluster IIb encoding-proteins. Based on the functional analyses described below, we re-named StMYBATV-chr12 as *ISAC* (*Inducer Silencing of Anthocyanins in Cell culture*).

2.3.2 StISAC editing increases and stabilizes anthocyanin accumulation in potato cells

StISAC has been proposed to act against anthocyanin accumulation similarly to SlMYBATV by re-recruiting anthocyanin bHLH co-partners. This might not allow the formation of the MYB-bHLH-WD40 (MBW) anthocyanin promoter complex (D'Amelia et al., 2020). Therefore, to knock out StISAC we edited its sequence in cells of the potato variety 'Blue Star' through CRISPR/Cas9. This system has been efficiently used to mutate alleles of the tetraploid potato for several desirable traits such as biotic stress tolerance and starch quality (reviewed in Tiwari et al., 2022). However, it has never been used to improve anthocyanin content in potato cells. In total, more than 20 kanamycin-resistant callus lines were obtained and subcultured for more than 12 months along with the wild type (Wt). The target region for the putative mutations was amplified and sequenced on the gDNA extracted from five randomly selected calli. Allele discrimination was carried out by CRISPR-ID® software analysis (Dehairs et al., 2016). As represented in Figure 1d, the end of the first exon was the target region for two SgRNA guides, and five type mutations were characterized in this region. Overall, deletions of nucleotides between PAMs of Sg1 and Sg2 were produced. *Mut_1*, *mut_2*, and *mut_3* led to a frameshift mutation and a premature stop codon, which was predicted to truncate the encoded protein after about the first 20 amino acids. *Mut_4* and *mut_5* were predicted to produce

sense mutations (i.e., D11N, F14S, I15S, or I15T) and to delete five amino acids. High resolution melting-based genotyping (HRM) revealed the heterogenic composition of the five selected calli as regards the StISAC allelic composition. Most alleles showed heteroallelic mutations (Figure S1). They were gathered together in the main peaks in Figure S1B. All cultures derived from edited calli were fully dark purple, in contrast to a "mosaic" (white and coloured cells and pale purple) culture generated from Wt potato explants (Figure 1e). The edited cells (Stisac) produced about twice as many anthocyanins as Wt cells produced and 14-fold more anthocyanins than tubers of "Blue Star" (Figure 1f,g). Compared with other sources of anthocyanins, the total amount of pigment accumulated by the Stisac-line was similar to that produced by blackberries (Figure 1g). However, the rich potato metabolite biodiversity may cover all of the six main aglycon anthocyanidins, adding further value to the potato-edited cell platforms. Moreover, potato accumulates several acylated anthocyanin forms (Oertel et al., 2017). This is felt to be important since acylation provides increased stability during technological processes (Zhao et al., 2017). Anthocyanin profiling of Stisac-edited lines highlighted a different increase of anthocyanin fractions favoring p-coumarylated forms. Malvidin-3-(p-coumaroyl)-rutinoside-5-glucoside increased more than other fractions and about 150% more than Wt cells (Figure 1h). We hypothesize that the lack of functionality of the R3-MYB repressor had positively impacted the main flux leading to the biosynthesis of flavonoid precursors (including p-coumaroyl-CoA) and affected the side routes less (e.g., feruoyl-CoA). Wt lines contained subpopulations of cells that accumulated no (or very small amounts of) anthocyanins. In contrast, edited cell culture, grown at the same conditions, showed more stable and homogeneously coloured cell populations (Figure 1e,i). Associated with the increased total amount of anthocyanins, anthocyanoplasts were observed within vacuoles of Stisac cell lines (Figure 1i). In addition, cells from edited cultures showed cylindrical forms with a mean length of 120 µm, whereas wt cells were round-shaped (diameter about 50 µm; Figure 1j). This difference might be due to the augmented volume of the vacuolar lumen of Stisac-cells determined by an improved anthocyanin vacuolar uptake. We cannot exclude that StISAC can directly affect cell shape. Indeed, members of clade II are closely related to CPC-type R3-MYBs (Figure 1a, LaFountain & Yuan, 2021), which are known to be involved in cell determination (Ma & Constabel, 2019). Future experimental works are needed to shed light on the potential involvement of Clade IIb members in cell shape determination.

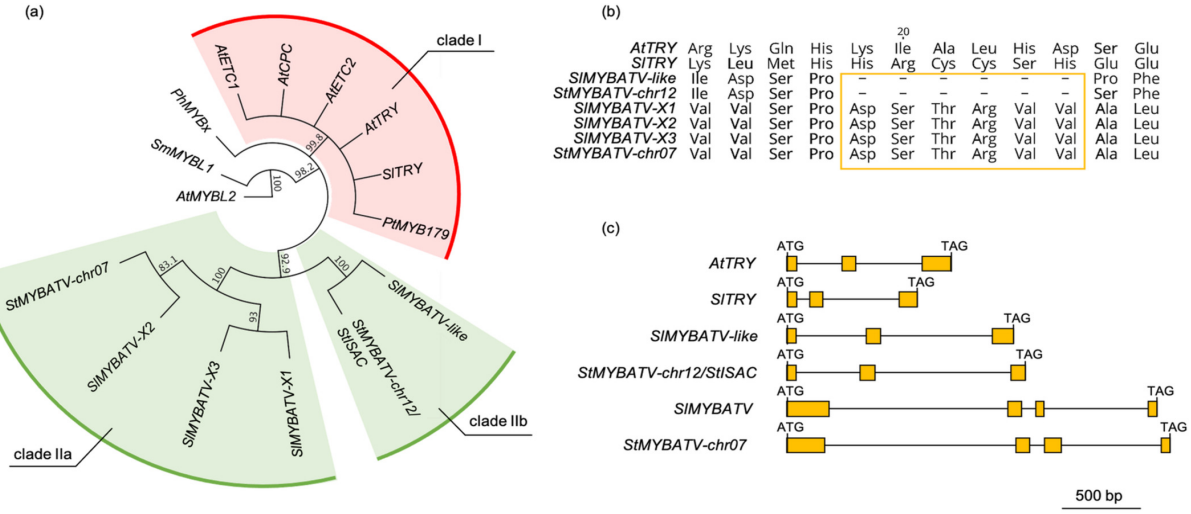
(a)
clade I
clade IIa
clade IIb
AtETC1
AtCPC
AtETC2
AtTRY
SlTRY
PtMYB179
PhMYBx
SmMYBL1
AtMYBL2
StMYBATV-chr07
SlMYBATV-X2
SlMYBATV-X3
SlMYBATV-X1
SlMYBATV-like
StISAC
SlMYBATV-chr12/
(b)
20
AtTRY Arg Lys Gln His Lys Ile Ala Leu His Asp Ser Glu
SlTRY Lys Leu Met His His Arg Cys Cys Ser His Glu Glu
SlMYBATV-like Ile Asp Ser Pro – – – – – – Pro Phe
StMYBATV-chr12 Ile Asp Ser Pro – – – – – – Ser Phe
SlMYBATV-X1 Val Val Ser Pro Asp Ser Thr Arg Val Val Ala Leu
SlMYBATV-X2 Val Val Ser Pro Asp Ser Thr Arg Val Val Ala Leu
SlMYBATV-X3 Val Val Ser Pro Asp Ser Thr Arg Val Val Ala Leu
StMYBATV-chr07 Val Val Ser Pro Asp Ser Thr Arg Val Val Ala Leu
(c)
AtTRY
SlTRY
SlMYBATV-like
StMYBATV-chr12/StISAC
SlMYBATV
StMYBATV-chr07
ATG
TAG
500 bp
(d)
Target sequence
StISAC
100 bp
PAM1
PAM2
Intron
Wt CATCAGATAATGCCTTCATAGACTCCCCTTGTAAGTTGTAAGTTCTTAATTTTCAAA → MADLDSSSTSDNAFIDSPSFEAKKEESS…
mut_1 CATCAGATAATGCCTTCAT ---------- CTTGTAAGTTGTAAGTTCTTAATTTTCAAA → MADLDSSSTSNNAFIFI*
mut_2 CATCAGATAATGCCTTC ---------- TTGTAAGTTGTAAGTTCTTAATTTTCAAA → MADLDSSSTSDNAFILGGH*
mut_3 CATCAGATAATGCCTTCAT ---------- TTGTAAGTTGTAAGTTCTTAATTTTCAAA → MADLDSSSTSNNAFII*
mut_4 CATCAGATAATGCCT ---------- CTTGTAAGTTGTAAGTTCTTAATTTTCAAA → MADLDSSSTSNNASS-------FEAKKEESS…
mut_5 CATCAGATAATGCCTTCAT ---------- TAAGTTGTAAGTTCTTAATTTTCAAA → MADLDSSSTSNNAFT------- FEAKKEESS…
(e)
Wt
stisac
(f)
Wt
stisac
(g)
mg PetAcd3G equ g-1 DW
Tubers 'Blue Star'
Wild Type Cells
Stisac cell lines
Grape Pomace
Raspberry
Blackberry
Blueberry
(h)
Petunidin-3-(p-coumaroyl)-rutinoside-5-glucoside*
Malvidin-3-(p-coumaroyl)-rutinoside-5-glucoside
Petunidin-3-(feruloyl)-rutinoside-5-glucoside
Malvidin-3-(feruloyl)-rutinoside-5-glucoside
Percentage of increment (%)
(i)
50 µm
(j)
µm
Wild Type Cells
Stisac cell lines

Figure 1. CRISPR/Cas9-mediated editing to stabilize anthocyanin production in potato cells. (a) Phylogenetic tree of CPC-like R3-MYBs. Values lower than 7% of bootstrap are not shown. (b) Particular of the alignment of R3-MYBs of clade I and II of the phylogenetic tree in (a). Yellow box underlines an amino acid motif which is held by components of clade IIa, but which lacks in clade IIb. (c) Genomic structures components of phylogenetic tree in (a). (d) A schematic map of the SgRNA targeted sites on StISAC and sequencing results of the Stisac cell lines (mut) obtained after transformation. The predicted mutation at amino acid level was also reported. In (c) and (d), introns (not drawn to scale) are shown as lines and exons are shown as yellow boxes. (d, f) Phenotype of a representative Stisac cell line (Stisac) at about 12 months after the transformation, respectively, as calli and liquid suspension culture. (g) Total anthocyanin content of Wt and Stisac edited cell lines in comparison with other sources of anthocyanins. (h) Relative percentage of anthocyanins in three representative Stisac cell lines compared with Wt, 1petunidin-3-(p-coumaroyl)-rutinoside-5-glucoside is present as two different stereoisomers. (i, j) Microscopic images of the suspension cultures and average cell length. * $P < 0.05$ according to Student's t-test.

2.4 Conclusions

To date, gene editing has never been used in potato or in cell cultures to change anthocyanin content. In this work, we functionalized a new potato anthocyanin repressor named StISAC and provided a method (based on CRISPR-Cas9) for stabilizing anthocyanin production in cell culture. Potato cell culture is advantageous for anthocyanin production because coloured potatoes display all the six most common anthocyanidins with different acylation levels. The method presented here is transferable to other potato varieties and related Solanaceae species.

REFERENCES

Appelhagen, I., Wulff-Vester, A.K., Wendell, M., Hvoslef-Eide, A.K., Russell, J., Oertel, A., Martens, S., Mock, H.P., Martin, C. and Matros, A. (2018). Colour bio-factories: Towards scale-up production of anthocyanins in plant cell cultures. *Metabolic Engeneering 48*, 218-232.

Cao, X., Qiu, Z., Wang, X., Van Giang, T., Liu, X., Wang, J., Wang, X., Gao, J., Guo, Y., Du, Y. and Wang, G. (2017). A putative R3 MYB repressor is the candidate gene underlying atroviolacium, a locus for anthocyanin pigmentation in tomato fruit. *Jornal of Experimental Bot*any *68*, 5745-5758.

D'Amelia, V., Villano, C., Batelli, G., Çobanoğlu, Ö., Carucci, F., Melito, S., Chessa, M., Chiaiese, P., Aversano, R. and Carputo, D. (2020). Genetic and epigenetic dynamics affecting anthocyanin biosynthesis in potato cell culture. *Plant Science 298*, 110597.

Dehairs, J., Talebi, A., Cherifi, Y. and Swinnen, J.V. (2016). CRISP-ID: decoding CRISPR mediated indels by Sanger sequencing. *Scientific Reports 6*, 28973.

Dong, N.Q. and Lin, H.X. (2021). Contribution of phenylpropanoid metabolism to plant development and plant–environment interactions. *Journal of Integrative Plant Biology 63*, 180-209.

LaFountain, A.M. and Yuan, Y.W. (2021). Repressors of anthocyanin biosynthesis. *New Phytologist 231*, 933–949.

Ma, D. and Constabel, C.P. (2019). MYB repressors as regulators of phenylpropanoid metabolism in plants. *Trends in Plant Science 24*, 275-289.

Neri-Numa, I.A., Arruda, H.S., Geraldi, M.V., Júnior, M.R.M. and Pastore, G.M. (2020). Natural prebiotic carbohydrates, carotenoids and flavonoids as ingredients in food systems. *Current Opinion in Food Science 33*, 98-107.

Oertel, A., Matros, A., Hartmann, A., Arapitsas, P., Dehmer, K.J., Martens, S. and Mock, H.P. (2017). Metabolite profiling of red and blue potatoes revealed cultivar and tissue specific patterns for anthocyanins and other polyphenols. *Planta 246*, 281-297.

Yan, S., Chen, N., Huang, Z., Li, D., Zhi, J., Yu, B., Liu, X., Cao, B. and Qiu, Z. (2020). Anthocyanin Fruit encodes an R2R3-MYB transcription factor, SlAN2-like, activating the transcription of SlMYBATV to fine-tune anthocyanin content in tomato fruit. *New Phytologist 225*, 2048-2063.

Zhao, C.L., Yu, Y.Q., Chen, Z.J., Wen, G.S., Wei, F.G., Zheng, Q., Wang, C.D. and Xiao, X.L. (2017). Stability-increasing effects of anthocyanin glycosyl acylation. *Food Chemistry 214*, 119-128.

Supplemental Material

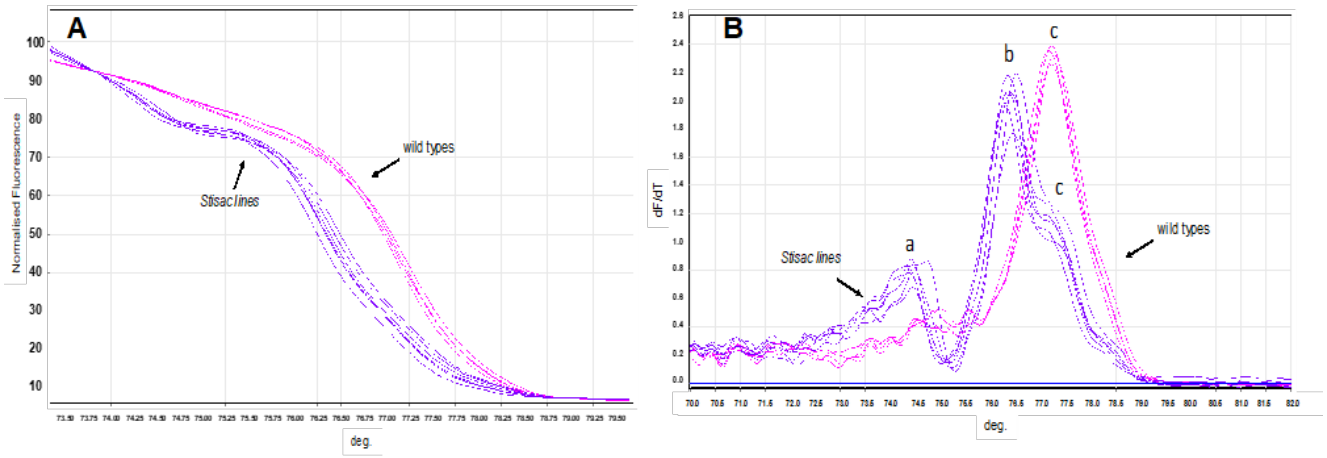

Figure S1. HRM curve analyses wild type and stisac cell populations. A) Changes in curve shape in Stsiac lines identify the heteroduplex (annealing of different alleles) melt curve. Heteroduplexes are also identified in B) where peak 'a' identified the presence of heteroduplex. Peak 'b' identified the presence of mutants compared to the wild type peak 'c'. The presence in stisac melt curve of a shoulder that resembles a peak c (same melting temperature) cannot exclude the presence of a low percentage of wild type alleles.

Published paper front-page

Received: 14 June 2022 | Accepted: 15 July 2022

DOI: 10.1002/pld3.433

ORIGINAL RESEARCH

Targeted mutagenesis of *StISAC* stabilizes the production of anthocyanins in potato cell culture

Vincenzo D'Amelia[1] | Annalisa Staiti[2] | Fabio D'Orso[3] | Maria Maisto[4] | Vincenzo Piccolo[4] | Riccardo Aversano[2] | Domenico Carputo[2]

[1]Institute of Biosciences and Bioresources (IBBR), National Research Council of Italy, Portici, Italy

[2]Department of Agricultural Sciences, University of Naples Federico II, Portici, Italy

[3]Research Centre for Genomics and Bioinformatics (CREA-GB), Council for Agricultural Research and Economics, Rome, Italy

[4]Department of Pharmacy, University of Naples Federico II, Naples, Italy

Correspondence
Domenico Carputo, Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, Italy.
Email: carputo@unina.it

Funding information
This research was supported by NUTR-AGE (FOE-2019-DSB.AD004.271-CNR) and ExPoPe-(STAR-Line 1, 2017-Unina- Intesa San Paolo). FD was supported by BIOTECH-Cisget project from the Italian Ministry of Agriculture (MiPAAF).

Abstract

To increase the production of decorated anthocyanins in potato cell cultures, we knocked out a novel potato gene, named *Inducer Silencing of Anthocyanins in Cell culture* (*StISAC*), using CRISPR-Cas9 editing. Our results provided evidence that mutant cell lines doubled the accumulation level of anthocyanins biosynthesized. Moreover, the production of these important pigments was stabilized over time. Our study overcame important challenges in the efficient biotechnological production of these valuable pigments and reported the function of a novel anthocyanin biosynthesis repressor gene.

KEYWORDS
calli, CRISPR-Cas9, R3-MYB, *Solanum tuberosum*

1 | INTRODUCTION

Anthocyanins are flavonoid pigments mostly responsible for the red, blue, and purple colors of plant tissues. They are used as colored additives and bioactive ingredients in processed foods (Neri-Numa et al., 2020). Among plant species, the cultivated potato *Solanum tuberosum* offers a wide and unique range of anthocyanins that can be extracted from pigmented tubers (Oertel et al., 2017). According to a *market data forecast* (www.marketdataforecast.com) report, the market size of anthocyanins was estimated at 318 million USD in 2020 and is expected to increase at a compound annual rate of 4.6%. Since direct extraction from plant matrices does not guarantee a sustainable and efficient production on demand, innovative methodologies and technologies can be used. Among them, plant cell cultures represent a very versatile and powerful system for metabolite production. The main drawback is that cell cultures are quite unstable in producing anthocyanins over time. This has a strong negative impact on their production in bioreactors (Appelhagen et al., 2018; D'Amelia et al., 2020). CRISPR-Cas9-based gene editing is the emerging technique that can promote the accumulation of anthocyanins in plants

Chapter 3. Genome-wide characterization of the anthocyanin O-methyltransferases in the cultivated potato

Annalisa Staiti †, Carmine Fruggiero †, Domenico Carputo, Vincenzo D'Amelia, Nunzio D'Agostino

Submitted to *Plant Gene*

Chapter 3. Genome-wide characterization of the anthocyanin O-methyltransferases in the cultivated potato

ABSTRACT

O-methyltransferases (OMTs) are a group of enzymes involved in the methylation of various secondary metabolites, including anthocyanins. This secondary modification, together with hydroxylation and glycosylation, affects the chromatic properties, stability and reactivity of these pigments. Meanwhile, no detailed identification or genome-wide analysis of the OMT gene family members in potato (*Solanum tuberosum*) has been reported. We conducted a genome-wide identification and characterization of potato OMT family members, identifying 65 OMT members. These were characterized based on gene structure, evolutionary relationships, and promoter motifs. Public database analysis revealed expression patterns, suggesting tissue and temporal specificity. Finally, we further elucidated expression differences among four selected OMT-encoding genes using potato cell cultures, which represent a smart and gainful alternative for anthocyanin production and deep knowledge about these post-biosynthetic modifications.

3.1 Introduction

Flowers, leaves, and tubers of the cultivated potato (*Solanum tuberosum* L.) exhibit a captivating array of pigmentations (Mattoo et al., 2022). In addition to orange and yellow colours of carotenoids, anthocyanin molecules offer vibrant blue, red, and purple hues. Anthocyanins, classified as flavonoids, boast diverse functionalities in plant defence. A recent review by (Lakshmikanthan et al., 2024) discussed the antimicrobial, antioxidant, and free radical scavenging properties exhibited by anthocyanins. This highlights their versatile nature and underscores their importance not only in plant biology but also in enhancing human well-being and extensive applications in human health and food technology. The chemical structure of anthocyanins consists of an anthocyanidin core, known as the flavium cation, which can undergo conjugation with different sugar moieties. An extensive literature reports over 600 anthocyanidins, of which six commonly occurring in nature: cyanidin, delphinidin, pelargonidin, malvidin, peonidin, and petunidin (Ayvaz et al., 2022). Anthocyanidins are distinguished by changes in the number and positioning of the hydroxyl (-OH) and methoxyl (-OCH3) groups on the B-ring. Post-synthesis modifications include methylation of the 3′ hydroxyl group of the cyanidin, producing peonidin, and the methylation of the 3′ hydroxyl group of delphinidin, resulting in petunidin, or both its 3′ and 5′ hydroxyl groups, forming malvidin (Fournier-Level et al., 2011). The anthocyanidin core, particularly the B- and C-rings, can undergo additional chemical modifications, including acylation, glycosylation, prenylation and methylation. Referred to as side chain decorations, these modifications significantly increase the variability of anthocyanins, even within different tissues (Alappat & Alappat, 2020). Chemical decorations exert a profound influence on the properties of these pigments, encompassing factors such as colour intensity, stability, interactions with other compounds, bioavailability, and potentially their health-promoting effects (Zhang et al., 2014). The methylation of the 3′-hydroxyl group of the B-ring is facilitated by Anthocyanin O-methyltransferases (AOMTs), which exhibit a pronounced affinity for anthocyanins (Du et al., 2015). These enzymes represent a specialized group within the broader category of S-adenosyl-L-methionine (SAM)-dependent O-methyltransferases (OMTs). Within various biosynthetic pathways, including the anthocyanin pathway OMTs serve as a crucial branch point, enhancing the water solubility and structural stability of these pigments. This process mitigates the chemical reactivity of phenolic hydroxyl groups, ultimately inducing a reddening effect (Y. T. Liu et al., 2022; Xie et al., 2022). Plant OMTs fall into two primary classes. Type I OMTs consist of homodimeric enzymes with molecular weights ranging from 38 to 43 kDa and operate independently of cations. Type II OMTs are cation-dependent and exhibit lower molecular mass of 23 and 27 kD (Khalil et al., 2015). Over the past decade, extensive *in silico* investigations of OMTs has been performed across various plant species, including citrus (Liu et al., 2016; Liu et al., 2020), tomato

(Gomez Roldan et al., 2014), pomegranate (Zhang et al., 2021) and grape (Lu et al., 2022). However, in the case of potato, there remain gaps in identifying putative genes encoding for methyltransferases, necessitating a more systematic characterization. Notably, OMTs exhibit versatility in substrate-specificity, with different enzymes potentially contributing to similar methylation processes on distinct substrates. For instance, (Do et al., 2007) reported that different class I Flavonoid O-methyltransferases (FOMTs), involved in alternative biosynthetic pathways like lignin production, can also catalyze reactions relevant to anthocyanin biosynthesis. Furthermore, it is hypothesized that different OMTs may function differently across different tissues or in response to distinct environmental conditions (Y. Liu et al., 2022). Considering these factors, we undertook a comprehensive genome-wide identification and characterization of the OMT family members. Initially, we utilized 57 tomato OMTs as baits, identifying 65 OMT members. Subsequently, we thoroughly characterized these members, examining their gene structure, evolutionary relationships and motifs present in their putative promoters. Then we interrogated public databases to ascertain the expression patterns of OMT candidates, aiming to elucidate their potential tissue and/or temporal specificity. In particular, all analyses were guided by our classification into four arbitrary groups based on discriminating conserved motifs identified within the OMTs. Finally, we used cell cultures to uncover expression differences among four selected OMT-encoding genes. Altogether, this research offers an overview of OMTs in potato and provides new insights into OMT genes as valuable tools for tailored anthocyanin production.

3.2 Materials and Methods

3.2.1 *In silico* identification of OMT gene family members

The 57 *Solanum lycopersicum* OMT sequences described by Gomez Roldan et al. (2014) were extracted and used to construct a multiple sequence alignment using MUSCLE (Edgar, 2004) with default settings. From this alignment, a hidden Markov model (HMM) profile was constructed using the *hmmbuild* utility from the HMMER 3.4 software package (Eddy, 2011). The *hmmsearch* utility form the same software package was then used to search for OMT candidates within the *Solanum tuberosum* Group phureja DM1-3 (v6.1) proteome (high confidence gene models) (Pham et al., 2020) using an e-value threshold of $< 10^{-3}$. Protein domains were identified by querying the PFAM (Mistry et al., 2021) and PROSITE (Hulo et al., 2006) databases via InterProScan (Quevillon et al., 2005). This step was crucial to confirm the domain integrity and exclude any incorrectly predicted OMT family members.

3.2.2 Analysis of gene structure and gene duplication

Chromosome positions and exon-intron structures of all identified OMT genes were extracted from the GFF3 file available for download at http://spuddb.uga.edu/data/DM_1-3_516_R44_potato.v6.1.hc_gene_models.gff3.gz. The gene structures and chromosome distribution were graphically represented using the Gene Structure Display Server 2.0 (Hu et al., 2015) and Inkscape software version 1.3 (https://inkscape.org/). Duplicated OMT genes were identified by comparing the protein sequences against each other using *blastp* with an e-value threshold of $< e^{-10}$. Duplication was confirmed based on two criteria: (a) alignment length > 75% of the longest sequence, and (b) a percentage of similarity > 80%. Genes located within a 200 kb region on the same chromosome were classified as tandem duplicates (Holub, 2001), while those separated by more than 200 kb were classified as segmental duplicates. Protein subcellular localization was predicted using the WoLF PSORT algorithm (https://wolfpsort.hgc.jp/).

3.2.3 Phylogenetic analysis

The identified potato sequences, along with the tomato OMT protein sequences, were aligned using MAFFT (Katoh & Standley, 2013). A maximum-likelihood phylogenetic tree was constructed using IQ-TREE (Nguyen et al., 2015) with 1,000 bootstrap replicates. The best-fitting substitution model was JTT+F+R4. Finally, the tree was visualized using the R packages Treeio (Wang et al., 2020) and ggtree (Yu, 2020). The sequence identifiers used to build the tree are listed in Supplementary Table 1.

3.2.4 Analysis of OMT gene expression patterns

Gene expression values, expressed as TPM (Transcript Per Million), were retrieved from RNA-Seq data (DM_1-3_516_R44_potato.v6.1) available on SpudDB (http://spuddb.uga.edu/index.shtml). The libraries were selected to fall into three different categories: “tissue-specific”, “abiotic stress” and “biotic stress”. TPM values > 0.5 were retained and transformed into $\log_2(TPM+1)$. Heatmaps were generated using ComplexHeatmap version 2.17.0 (Gu et al., 2016).

3.2.5 Analysis of transcription factor binding motifs in OMT gene promoter regions

The putative proximal regulatory regions (1,672 bp upstream of the annotated transcription start site as suggested by Kristiansson et al. (2009)) were extracted from the DM 1-3 516 R44 reference chromosome sequences. Transcription factor binding motifs (TFBMs) within these regions were identified using PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/) (Lescot et al., 2002) with default settings. The TFBMs were categorized into “stress responsive”, “light

responsive" and "MYB responsive" groups. To visualize the occurrences of these TFBMs, a heatmap was generated using the R package ComplexHeatmap version 2.17.0 (Gu et al., 2016).

3.2.6 Cell culture production

In this study, cell lines derived from calli of the potato variety "Blue Star" were used, which over-accumulate methylated anthocyanins, specifically malvidin and petunidin, due to the mutation of *StISAC* (D'Amelia et al., 2022). *StISAC* is a regulatory repressor that interferes with the anthocyanin activator complex during biosynthesis. Wild-type cells were also used as controls. All cell cultures were maintained in 100 mm tissue culture plates containing 25 ml of Murashige and Skoog (MS) medium (including vitamins), supplemented with 3 mg/L of 2,4-dichlorophenoxyacetic acid (2,4-D) (Sigma-Aldrich), 3% sucrose, and 8 mg/L of micro agar (Duchefa Biochemie) adjusted to pH 5.8. Cell cultures were grown in the dark at 24 °C ± 2 and subcultured every two weeks.

3.2.7 Primer design

Real-time RT-PCR primers were designed using Primer3Plus (https://www.primer3plus.com/) (Untergasser et al., 2012) following these criteria: primer length between 18 and 24 bp, melting temperature between 58 °C and 62 °C, and GC content between 40% and 60%. The absence of hairpin and dimer formation was verified, and primer specificity was tested using Primer Blast (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). Primer sequences are listed in Supplementary Table 2.

3.2.8 RNA isolation and quantitative RT-PCR

DNA-free total RNA was extracted from 100 mg of frozen calli using the Spectrum Plant Total RNA Kit and the On-Column DNase I Digestion Set (Sigma-Aldrich), following the manufacturer's instructions. One µg of total RNA was reverse transcribed into complementary DNA (cDNA) using an oligo-dT(20) primer and SuperScript III reverse transcriptase (Invitrogen) in a final reaction volume of 20 µl, according to the manufacturer's protocol. Quantitative RT-PCR experiments were performed in biological triplicates, with each replicate consisting of pooled calli collected from a single Petri dish. The reactions used the 2x QuantiFast SYBR Green PCR Master Mix (Qiagen) and were conducted on the ABI PRISM 7900 H T instrument (Thermo Fisher Scientific). Each 20 µl reaction mixture included 10 mM of each primer and cDNA diluted 1:20. The cycling conditions followed those specified in the QuantiFast SYBR Green PCR Kit handbook (Qiagen). The elongation factor 1 alpha (EF-1α, Gene ID: 102600444) was used as an endogenous control to normalize the gene expression data. Melting curves of the PCR products were generated to confirm their specificity and identity. Each Assay was performed in triplicate to ensure technical reproducibility. The obtained

data were analyzed using the ABI PRISM 7900 H T Sequence Detection System Version 2.1 (Thermo Fisher Scientific). Relative expression levels were estimated using the ΔCt method.

3.2.9 Statistical analysis

Statistical analysis was performed using SPSS software (v29.0.1, SPSS Inc., Chicago). The statistical significance of the experimental means was assessed using independent samples t-test, with a threshold of $*p < 0.05$ considered significant. Mean and standard deviation (SD) values were calculated from three independent biological and technical replicates.

3.3 Results

3.3.1 Identification and characterization of OMT encoding genes

A total of 65 genes encoding OMT family members were identified (Table 1). The length of these genes varied significantly, ranging from 328 bp to 14,818 bp. By querying protein signature databases, we verified that all predicted proteins contained an O-methyltransferase domain. Specifically, four InterPro signatures were found: IPR002935 (Class I-like SAM-dependent O-methyltransferase), IPR016461 (O-methyltransferase COMT-type), IPR001077 (O-methyltransferase), and IPR012967 (Plant methyltransferase dimerization). Based on the distribution of these domains, the sequences were categorized into four distinct groups. Group 1 consisted of 13 genes harboring the IPR002935 domain; group 2 comprised 10 genes with both IPR016461 and IPR001077 domains; group 3 included 2 genes housing IPR016461 and IPR012967 domains; group 4 contained 40 genes with IPR016461, IPR012967 and IPR001077 domains (Table 1). The majority of the proteins (63%) were predicted to be localized in the cytoplasm. Approximately 21% were predicted to localize to the cytoskeleton, 7 proteins to the chloroplast, and only 3 proteins to the nucleus (Table 1).

Table 1. AOMT gene family members identified with their chromosomal and protein product subcellular localizations, number of exons, InterPro indentified domains and the relative grouping.

Chr	Start	Stop	ID	Length	Predicted subcellular localization	Number of exons	InterPro Domain	Label
chr01	65303485	65304636	Soltu.DM.01G025740.1	1150	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr01	65322987	65324394	Soltu.DM.01G025760.1	1406	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr01	65334346	65335482	Soltu.DM.01G025770.1	1135	nucleus	2	IPR016461, IPR001077, IPR012967	Group 4
chr01	84638275	84639979	Soltu.DM.01G047320.1	1703	cytoplasm	5	IPR002935	Group 1
chr01	88004850	88007025	Soltu.DM.01G051460.1	2174	cytoskeleton	3	IPR016461, IPR001077, IPR012967	Group 4
chr01	88022223	88023473	Soltu.DM.01G051480.1	1249	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr02	31387859	31389043	Soltu.DM.02G016860.1	1183	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr02	31394610	31396114	Soltu.DM.02G016870.2	1503	chloroplast	2	IPR016461, IPR001077, IPR012967	Group 4
chr02	31397851	31400258	Soltu.DM.02G016880.1	2406	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr02	31401264	31403167	Soltu.DM.02G016890.1	1902	cytoplasm	3	IPR016461, IPR001077	Group 2
chr02	41170188	41171593	Soltu.DM.02G028520.1	1404	cytoplasm	5	IPR002935	Group 1
chr02	41172855	41177532	Soltu.DM.02G028530.1	4676	cytoplasm	3	IPR002935	Group 1
chr02	41184782	41186195	Soltu.DM.02G028540.1	1412	cytoskeleton	5	IPR002935	Group 1
chr02	41188339	41189774	Soltu.DM.02G028550.1	1434	cytoskeleton	5	IPR002935	Group 1
chr03	3323083	3337902	Soltu.DM.03G003400.1	14818	cytoplasm	10	IPR002935	Group 1
chr03	44489463	44490613	Soltu.DM.03G019770.1	1149	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr03	44500072	44500753	Soltu.DM.03G019790.1	680	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4
chr03	44505766	44507013	Soltu.DM.03G019800.1	1246	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4
chr03	44539967	44540418	Soltu.DM.03G019810.1	450	chloroplast	2	IPR016461, IPR001077	Group 2
chr03	44546497	44548106	Soltu.DM.03G019820.1	1608	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4
chr03	44563991	44565830	Soltu.DM.03G019830.1	1838	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4
chr03	44586417	44586839	Soltu.DM.03G019850.1	421	cytoplasm	1	IPR016461, IPR001077	Group 2
chr03	44654942	44655271	Soltu.DM.03G019870.1	328	cytoskeleton	1	IPR016461, IPR012967	Group 3
chr03	44655313	44656123	Soltu.DM.03G019880.1	809	cytoplasm	2	IPR016461, IPR001077	Group 2
chr03	44668067	44669969	Soltu.DM.03G019890.1	1901	cytoplasm	2	IPR016461, IPR001077	Group 2
chr03	44672155	44673718	Soltu.DM.03G019900.1	1562	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4
chr03	44674512	44677841	Soltu.DM.03G019910.1	3328	cytoskeleton	4	IPR016461, IPR001077, IPR012967	Group 4
chr03	44698208	44699293	Soltu.DM.03G019920.1	1084	nucleus	2	IPR016461, IPR001077	Group 2
chr03	46325427	46328884	Soltu.DM.03G021440.1	3456	chloroplast	4	IPR016461, IPR001077, IPR012967	Group 4
chr04	11232516	11234902	Soltu.DM.04G010660.1	2385	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr04	54993819	54995289	Soltu.DM.04G025040.1	1469	cytoplasm	2	IPR002935	Group 1
chr04	54996590	54998928	Soltu.DM.04G025050.3	2337	chloroplast	3	IPR002935	Group 1
chr05	21946475	21947000	Soltu.DM.05G014010.1	524	cytoplasm	3	IPR002935	Group 1
chr06	2630617	2633858	Soltu.DM.06G002210.1	3240	chloroplast	4	IPR016461, IPR001077, IPR012967	Group 4
chr06	43392751	43395881	Soltu.DM.06G017160.1	3129	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr06	43505619	43509408	Soltu.DM.06G017290.2	3788	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr06	47947191	47948601	Soltu.DM.06G021280.1	1409	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4

chr06	47955403	47956543	Soltu.DM.06G021290.1	1139	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr06	47961794	47963637	Soltu.DM.06G021300.1	1842	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr06	57996700	57999161	Soltu.DM.06G033770.1	2460	chloroplast	4	IPR016461, IPR001077, IPR012967	Group 4
chr06	58001852	58003853	Soltu.DM.06G033780.1	2000	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr06	58007118	58008874	Soltu.DM.06G033790.1	1755	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr06	58012126	58013782	Soltu.DM.06G033800.1	1655	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr08	2391445	2401186	Soltu.DM.08G001680.1	9740	chloroplast	10	IPR002935	Group 1
chr08	2428548	2430904	Soltu.DM.08G001740.1	2355	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr09	44025484	44026666	Soltu.DM.09G016040.1	1181	cytoplasm	3	IPR016461, IPR001077	Group 2
chr09	44953283	44953732	Soltu.DM.09G016210.1	448	cytoplasm	1	IPR016461, IPR012967	Group 3
chr09	46667427	46667825	Soltu.DM.09G016500.1	397	cytoplasm	1	IPR016461, IPR001077	Group 2
chr09	60970955	60974845	Soltu.DM.09G025040.1	3889	cytoplasm	6	IPR002935	Group 1
chr10	98895	101356	Soltu.DM.10G000070.1	2460	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr10	3202771	3204188	Soltu.DM.10G003870.2	1416	cytoplasm	2	IPR016461, IPR001077, IPR012967	Group 4
chr10	3254414	3256757	Soltu.DM.10G003940.1	2342	cytoskeleton	2	IPR016461, IPR001077, IPR012967	Group 4
chr10	42911223	42914010	Soltu.DM.10G014940.1	2786	cytoplasm	5	IPR002935	Group 1
chr10	53440964	53443968	Soltu.DM.10G021470.1	3003	cytoskeleton	4	IPR016461, IPR001077, IPR012967	Group 4
chr10	53446126	53449270	Soltu.DM.10G021480.1	3143	cytoskeleton	4	IPR016461, IPR001077	Group 2
chr10	59032584	59034046	Soltu.DM.10G027870.1	1461	nucleus	2	IPR016461, IPR001077, IPR012967	Group 4
chr11	965704	968368	Soltu.DM.11G000820.1	2663	cytoplasm	4	IPR016461, IPR001077, IPR012967	Group 4
chr12	16982808	16983310	Soltu.DM.12G013090.1	501	cytoplasm	2	IPR002935	Group 1
chr12	56058349	56061661	Soltu.DM.12G026090.2	3311	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr12	56072222	56079011	Soltu.DM.12G026100.1	6788	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr12	56095470	56099511	Soltu.DM.12G026120.1	4040	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr12	56110433	56116582	Soltu.DM.12G026130.1	6148	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr12	56127619	56130826	Soltu.DM.12G026140.1	3206	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr12	56161844	56164880	Soltu.DM.12G026150.1	3035	cytoplasm	3	IPR016461, IPR001077, IPR012967	Group 4
chr12	58031121	58033159	Soltu.DM.12G028310.1	2037	cytoplasm	2	IPR016461, IPR001077	Group 2

3.3.2 Gene structure analysis, duplication events, and phylogenetic relationships of OMT gene family members

The 65 OMT genes exhibited various intron/exon organizations (Table 1; Figure 1). Within the four previously identified groups, distinct patterns of intron/exon organization were observed (Table 1; Figure 1). Specifically, group 1 genes had between 2 and 10 exons; group 2 genes had between 1 and 4 exons; group 3 genes contained only one exon; group 4 genes had between 2 and 4 exons. Figure 1 illustrates these four groups to facilitate a better understanding of potential functional relationships among the potato genes. Interestingly, group 1 genes, despite being phylogenetically related, exhibited the greatest variability in exon number. This group also included the genes with the highest number of exons, namely Soltu.DM.03G003400.1 and Soltu.DM.08G001680.1. Despite some protein sequences being phylogenetically close, their gene structure showed divergences placing them in different groups. For example, Soltu.DM.09G016210.1(group 3) and Soltu.DM.09G016040.1 (group 2), as well as Soltu.DM.03G019910.1 (group 4) and Soltu.DM.03G019880.1 (group 2), illustrate this phenomenon.

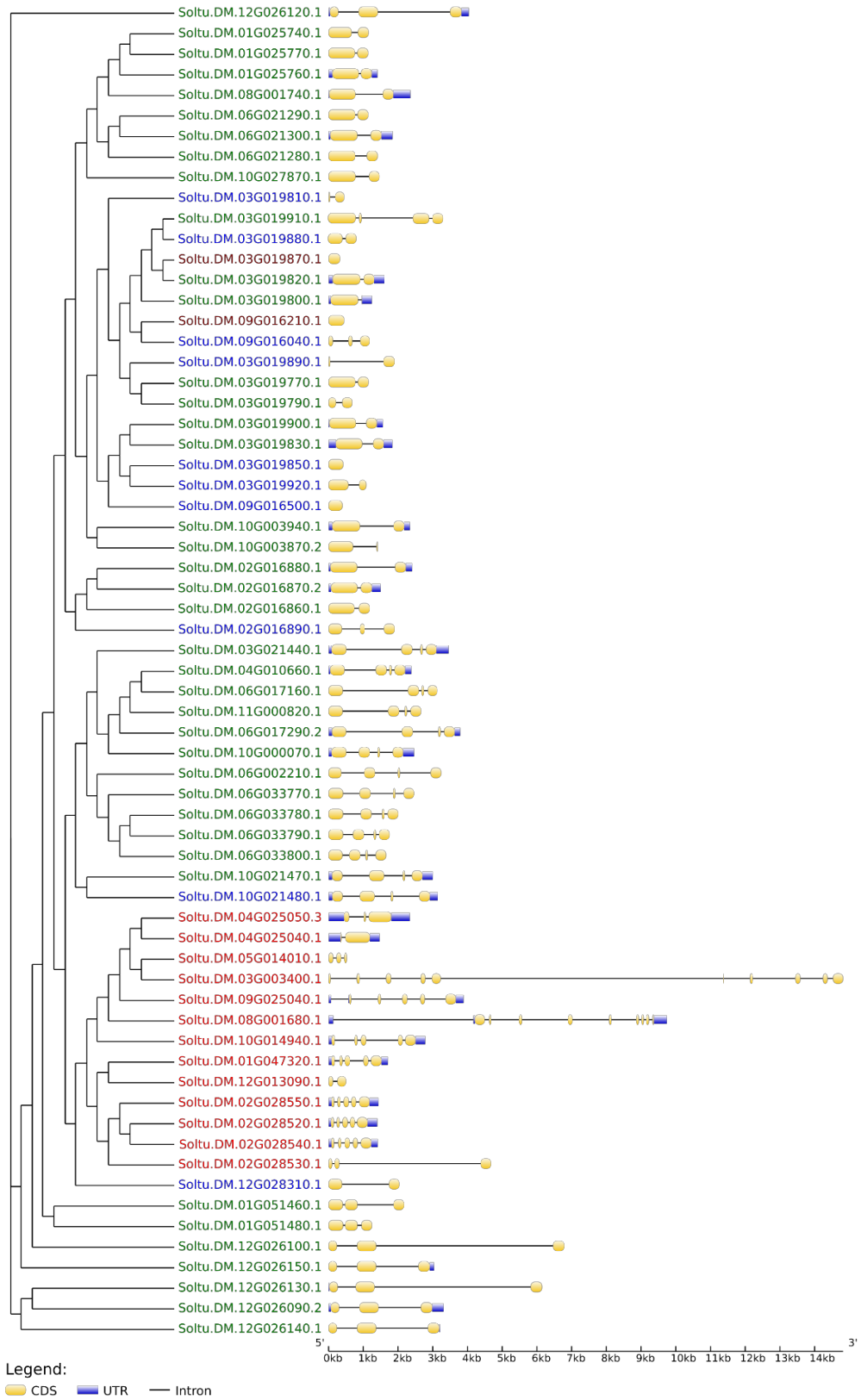

Figure 1. Exon-intron organization of AOMT gene family members (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4"). Round-corner boxes representes exons (blue box = UTR region, yellow box = CDS region) and black line represents introns. The relative size of the full transcript, intron, exon, and upstream region could be inferred from the supplied scale in kilobase pair (kb).

The genes were distributed across all 12 chromosomes of the potato genome, except for the chromosome 7 (Table 2). Chromosome 3 contained the largest number of OMT-encoding genes (N. = 15), followed by chromosome 6 (N.= 10) and chromosome 2 (N. = 8). Fifty-three OMT genes (shown in Table 2) exhibited sequence similarity greater than 80% and are believed to have originated from gene duplication events. These duplicated genes were distributed across all potato chromosomes, except for chromosomes 5, 7 and 8. Table 2 also details the types of duplication events. Some genes evolved through both segmental and tandem duplication events, indicating that they may be on different chromosomes or on the same chromosome, respectively.

Specifically, 16 genes were presumably involved in segmental duplication events, while 48 genes were implicated in tandem duplication events. Typically, tandem duplicated genes belonged to the same group with few exceptions. For instance, on chr 3, Soltu.DM.03G019850.1, belonging to group 2, was duplicated in tandem with Soltu.DM.03G019830.1 and Soltu.DM.03G019900.1, both belonging to group 4. The evolutionary history of the OMT potato genes was further investigated by analyzing their phylogenetic relationships with orthologs in tomato (Figure 2). The sequences used to construct the phylogenetic tree are listed in Supplementary Table 1, and the resulting tree is depicted in Figure 2. Notably, among these sequences, only “group 1”, distinguished by the presence of the IPR002935 domain, formed a monophyletic clade, even when supplemented with additional 14 OMT genes from tomato (Figure 2). Adjacent to this clade, a paraphyletic group containing only tomato sequences was observed, flanked by the sequences Solyc12g014500.1.1 and Solyc00g029190.2.1. No specific relationships were evident for the other three groups, as their members were randomly distributed across the remaining clades.

Table 2. Gene duplication analysis.

Gene A	Chromosome location gene A	Gene B	Chromosome location gene B	Distance (bp)	Duplication type	Percentage Identity
Soltu.DM.01G025740.1	chr01	Soltu.DM.01G025770.1	chr01	30861	Tandem duplication	80.94
Soltu.DM.01G025760.1	chr01	Soltu.DM.01G025770.1	chr01	11359	Tandem duplication	87.1
Soltu.DM.01G047320.1	chr01	Soltu.DM.10G014940.1	chr10	NA	Segmental duplication	88.66
Soltu.DM.01G047320.1	chr01	Soltu.DM.02G028550.1	chr02	NA	Segmental duplication	89.03
Soltu.DM.01G047320.1	chr01	Soltu.DM.02G028540.1	chr02	NA	Segmental duplication	89.03
Soltu.DM.01G047320.1	chr01	Soltu.DM.02G028520.1	chr02	NA	Segmental duplication	88.19
Soltu.DM.01G051460.1	chr01	Soltu.DM.01G051480.1	chr01	17373	Tandem duplication	91.62
Soltu.DM.02G016860.1	chr02	Soltu.DM.02G016880.1	chr02	9992	Tandem duplication	90.08
Soltu.DM.02G016870.2	chr02	Soltu.DM.02G016880.1	chr02	3241	Tandem duplication	85.67
Soltu.DM.02G016870.2	chr02	Soltu.DM.02G016860.1	chr02	6751	Tandem duplication	83.1
Soltu.DM.02G028520.1	chr02	Soltu.DM.02G028540.1	chr02	14594	Tandem duplication	97.94
Soltu.DM.02G028520.1	chr02	Soltu.DM.02G028550.1	chr02	18151	Tandem duplication	97.12

Soltu.DM.02G028530.1	chr02	Soltu.DM.02G028540.1	chr02	11927	Tandem duplication	97.16
Soltu.DM.02G028530.1	chr02	Soltu.DM.02G028550.1	chr02	15484	Tandem duplication	97.16
Soltu.DM.02G028530.1	chr02	Soltu.DM.02G028520.1	chr02	2667	Tandem duplication	95.46
Soltu.DM.02G028540.1	chr02	Soltu.DM.02G028550.1	chr02	3557	Tandem duplication	99.18
Soltu.DM.03G019800.1	chr03	Soltu.DM.03G019820.1	chr03	40731	Tandem duplication	93.23
Soltu.DM.03G019800.1	chr03	Soltu.DM.03G019910.1	chr03	168746	Tandem duplication	92.45
Soltu.DM.03G019820.1	chr03	Soltu.DM.03G019910.1	chr03	128015	Tandem duplication	94.81
Soltu.DM.03G019830.1	chr03	Soltu.DM.03G019900.1	chr03	108164	Tandem duplication	96.89
Soltu.DM.03G019850.1	chr03	Soltu.DM.03G019900.1	chr03	85738	Tandem duplication	88.89
Soltu.DM.03G019850.1	chr03	Soltu.DM.09G016500.1	chr09	NA	Segmental duplication	87.88
Soltu.DM.03G019850.1	chr03	Soltu.DM.03G019830.1	chr03	22426	Tandem duplication	88.15
Soltu.DM.03G019850.1	chr03	Soltu.DM.03G019920.1	chr03	111791	Tandem duplication	83.7
Soltu.DM.03G019870.1	chr03	Soltu.DM.03G019820.1	chr03	108445	Tandem duplication	93.58
Soltu.DM.03G019870.1	chr03	Soltu.DM.03G019910.1	chr03	19570	Tandem duplication	92.66
Soltu.DM.03G019880.1	chr03	Soltu.DM.03G019910.1	chr03	19199	Tandem duplication	97.89
Soltu.DM.03G019890.1	chr03	Soltu.DM.03G019770.1	chr03	178604	Tandem duplication	83.17
Soltu.DM.03G019920.1	chr03	Soltu.DM.03G019900.1	chr03	26053	Tandem duplication	85.66
Soltu.DM.03G019920.1	chr03	Soltu.DM.03G019830.1	chr03	134217	Tandem duplication	84.93
Soltu.DM.04G010660.1	chr04	Soltu.DM.11G000820.1	chr11	NA	Segmental duplication	93.92
Soltu.DM.04G010660.1	chr04	Soltu.DM.06G017160.1	chr06	NA	Segmental duplication	91.46
Soltu.DM.04G025040.1	chr04	Soltu.DM.04G025050.3	chr04	2771	Tandem duplication	84.9
Soltu.DM.06G017160.1	chr06	Soltu.DM.11G000820.1	chr11	NA	Segmental duplication	91.44
Soltu.DM.06G017290.2	chr06	Soltu.DM.11G000820.1	chr11	NA	Segmental duplication	89.44
Soltu.DM.06G017290.2	chr06	Soltu.DM.04G010660.1	chr04	NA	Segmental duplication	87.29
Soltu.DM.06G017290.2	chr06	Soltu.DM.06G017160.1	chr06	112868	Tandem duplication	85.08
Soltu.DM.06G021280.1	chr06	Soltu.DM.06G021300.1	chr06	14603	Tandem duplication	84.72
Soltu.DM.06G021280.1	chr06	Soltu.DM.06G021290.1	chr06	8212	Tandem duplication	82.96
Soltu.DM.06G021290.1	chr06	Soltu.DM.06G021300.1	chr06	6391	Tandem duplication	90.83
Soltu.DM.06G033770.1	chr06	Soltu.DM.06G033780.1	chr06	5152	Tandem duplication	86.78
Soltu.DM.06G033770.1	chr06	Soltu.DM.06G033790.1	chr06	10418	Tandem duplication	85.12
Soltu.DM.06G033770.1	chr06	Soltu.DM.06G033800.1	chr06	15426	Tandem duplication	84.87
Soltu.DM.06G033780.1	chr06	Soltu.DM.06G033790.1	chr06	5266	Tandem duplication	94.23
Soltu.DM.06G033780.1	chr06	Soltu.DM.06G033800.1	chr06	10274	Tandem duplication	90.5
Soltu.DM.09G016500.1	chr09	Soltu.DM.03G019900.1	chr03	NA	Segmental duplication	87.79
Soltu.DM.09G016500.1	chr09	Soltu.DM.03G019830.1	chr03	NA	Segmental duplication	87.02
Soltu.DM.09G016500.1	chr09	Soltu.DM.03G019920.1	chr03	NA	Segmental duplication	80.92
Soltu.DM.10G003870.2	chr10	Soltu.DM.10G003940.1	chr10	51643	Tandem duplication	95.85
Soltu.DM.10G014940.1	chr10	Soltu.DM.02G028550.1	chr02	NA	Segmental duplication	87.55

Soltu.DM.10G014940.1	chr10	Soltu.DM.02G028520.1	chr02	NA	Segmental duplication	84.36
Soltu.DM.10G014940.1	chr10	Soltu.DM.02G028540.1	chr02	NA	Segmental duplication	86.69
Soltu.DM.10G021470.1	chr10	Soltu.DM.10G021480.1	chr10	5162	Tandem duplication	85.55
Soltu.DM.12G013090.1	chr12	Soltu.DM.01G047320.1	chr01	NA	Segmental duplication	85
Soltu.DM.12G026090.2	chr12	Soltu.DM.12G026140.1	chr12	69270	Tandem duplication	96.22
Soltu.DM.12G026090.2	chr12	Soltu.DM.12G026120.1	chr12	37121	Tandem duplication	95.95
Soltu.DM.12G026090.2	chr12	Soltu.DM.12G026130.1	chr12	52084	Tandem duplication	96.22
Soltu.DM.12G026090.2	chr12	Soltu.DM.12G026100.1	chr12	13873	Tandem duplication	95.14
Soltu.DM.12G026090.2	chr12	Soltu.DM.12G026150.1	chr12	103495	Tandem duplication	94.05
Soltu.DM.12G026100.1	chr12	Soltu.DM.12G026150.1	chr12	89622	Tandem duplication	94.6
Soltu.DM.12G026100.1	chr12	Soltu.DM.12G026120.1	chr12	23248	Tandem duplication	94.32
Soltu.DM.12G026100.1	chr12	Soltu.DM.12G026140.1	chr12	55397	Tandem duplication	94.6
Soltu.DM.12G026100.1	chr12	Soltu.DM.12G026130.1	chr12	38211	Tandem duplication	93.78
Soltu.DM.12G026120.1	chr12	Soltu.DM.12G026140.1	chr12	32149	Tandem duplication	96.76
Soltu.DM.12G026120.1	chr12	Soltu.DM.12G026130.1	chr12	14963	Tandem duplication	95.95
Soltu.DM.12G026120.1	chr12	Soltu.DM.12G026150.1	chr12	66374	Tandem duplication	94.32
Soltu.DM.12G026130.1	chr12	Soltu.DM.12G026140.1	chr12	17186	Tandem duplication	96.76
Soltu.DM.12G026130.1	chr12	Soltu.DM.12G026150.1	chr12	51411	Tandem duplication	94.05
Soltu.DM.12G026140.1	chr12	Soltu.DM.12G026150.1	chr12	34225	Tandem duplication	93.78

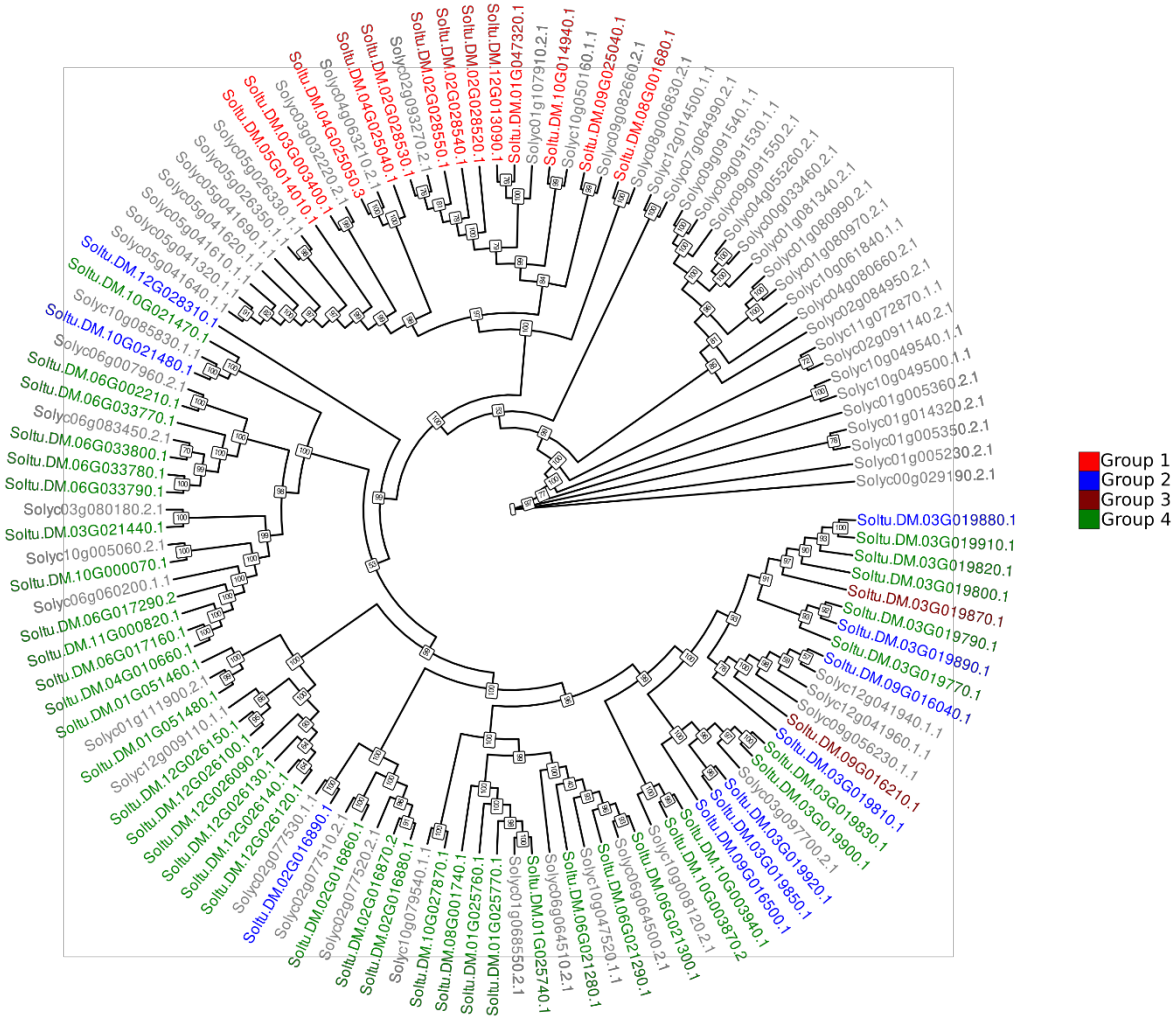

Figure 2. Phylogenetic tree of the 65 identified 65 potato AOMT and the 57 identified tomato AOMT by Gomez Roldan et al., 2014. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4") while all tomato AOMT ID were coloured in black.

3.3.3 Expression patterns of OMT genes

The RNA-Seq libraries utilized are detailed in Supplementary Table 3 and were derived from various tissues sampled across 7 potato cultivars (RH89-039-16, Longshu3, DM 1-3 R44, Russet Burbank, Group Phujera, Atlantic, Superior). Among the analyzed genes, 12 from either group 1 or group 4 exhibited distinct expression patterns across different plant tissues, displaying higher expression compared to other OMT members (Supplementary Figure 1). In particular: Soltu.DM.10G014940.1 (group 1) and Soltu.DM.03G021440.1 (group 4) showed elevated expression levels in the stem relative to other OMT members; Soltu.DM.08G001740.1 (group 4) exhibited heightened expression in petals; Soltu.DM.02G028550.1 (group 1) displayed increased expression in in-vitro callus;

Soltu.DM.10G000070.1 (group 4) demonstrated enhanced expression in the flower; Soltu.DM.02G028520.1 (group 1) exhibited elevated expression in the stamens; Soltu.DM.09G025040.1 (group 1) showed heightened expression in the leaves; and Soltu.DM.10G003940 (group 4) displayed increased expression in the meristem. The remaining OMTs did not exhibit specific expression patterns, being lowly represented in the considered tissues. Additionally, genes from either group 1 or group 4 displayed differential expression under both abiotic and biotic stress conditions (Supplementary Figure 3; Supplementary Table 4). Under various abiotic stress conditions (Supplementary Figure 2), we observed distinct expression patterns. For instance, Soltu.DM.10G014940.1, Soltu.DM.09G025040.1 (group 1) and Soltu.DM.03G021440.1 (group 4) showed significantly higher expression levels under salt stress compared to other OMT members. Conversely, Soltu.DM.02G028550.1 (group 1) was highly expressed under mannitol stress condition (Supplementary Table 5). Similarly, four genes in group 4, namely Soltu.DM.03G019800.1, Soltu.DM.03G019820.1, Soltu.DM.03G019900.1, and Soltu.DM.03G019830.1, were highly expressed in wounded leaves compared to other OMT members (Supplementary Table 6).

3.3.4 TFBMs in the promoter region of OMT genes

TFBMs considered for this analysis focused on three categories (Table 3): “stress responsive”, “light-responsive”, “MYB responsive”, presumed to be involved in activating the anthocyanin metabolic pathway. The heatmap depicted in Figure 3 illustrates the presence of these motifs in the regulatory regions of OMT genes. Notably, most motifs were found in low copy numbers across the regulatory regions. Hierarchical clustering highlighted that only two motifs, STRE (activated under heat stress) and BOX4 (light responsive), had a significant number of occurrences, reaching up to 9 copies in several sequences primarily belonging to groups 1 and 4. In particular, the STRE motif was detected in multiple copies in Soltu.DM.12G028310.1 (group 2), Soltu.DM.09G025040.1 (group 1), Soltu.DM.10G003940.1 (group 4), Soltu.DM.04G025040.1 (group 1). On the other hand, the BOX4 motif was identified in Soltu.DM.02G028530.1 (group 1), and in Soltu.DM.10G000070.1, Soltu.DM.06G033780.1, Soltu.DM.06G033790.1, and Soltu.DM.06G021290.1. all belonging to group 4. A few motifs, each present in up to 8 copies, were identified in the upstream sequences of selected genes, with three belonging to group 1, one to group 2, and one to group 4 (Figure 3; Supplementary Table 7). Specifically, Soltu.DM.02G016870.2 (group 4) harbored 6 copies of the G-BOX element (a cis-acting regulatory element associated with light responsiveness), Soltu.DM.02G028530.1 (group 1) contained 7 copies of the BOX4 element (a component of a conserved DNA module linked to light responsiveness), Soltu.DM.03G019810.1 (group 2) possessed 6 copies of the WUN-MOTIF (a wound-responsive element), Soltu.DM.05G014010.1 (group 1)

contained 6 copies of the LTR element (a cis-acting element implicated in low-temperature responsiveness), and Soltu.DM.09G025040.1 (group 1) contained 7 copies of the STRE element (a stress response element). Over half of the upstream regions of the 65 candidate genes contained multiple regulatory motifs, with certain genes, such as Soltu.DM.01G025740.1 and Soltu.DM.02G016880.1, exhibiting notably high motif counts (Table 3).

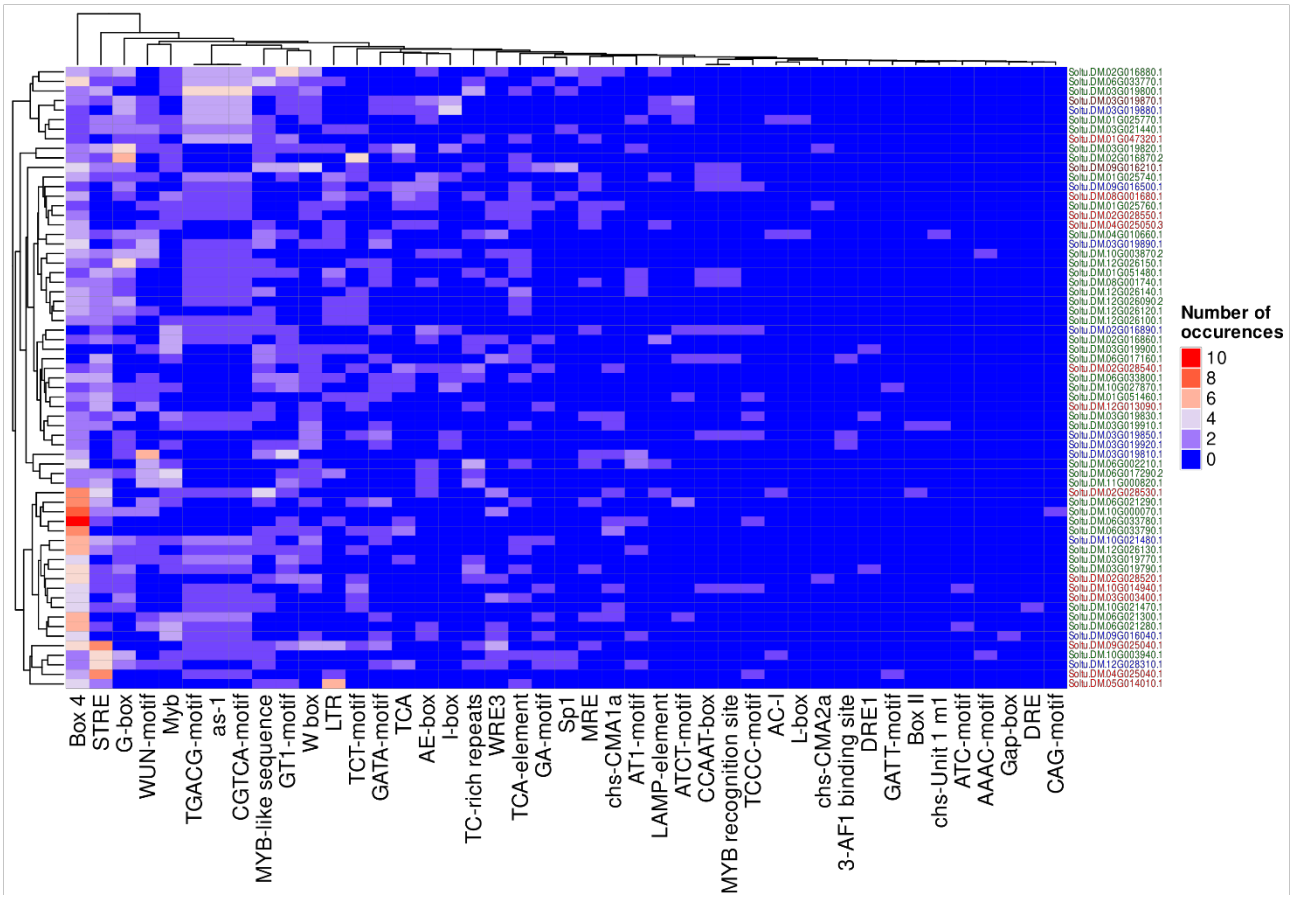

Figure 3. Predicted TFBMs (transcription factor binding motifs) in AOMT promoters. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

Table 3. Promoter TFBMs (transcription factor binding motifs) selected.

Site name	Sequence	Organism	Function
3-AF1 binding site	TAAGAGAGGAA	*Solanum tuberosum*	light responsive element
AAAC-motif	CAATCAAAACCT	*Spinacia oleracea*	light responsive element
AC-I	(T/C)C(T/C)(C/T)ACC(T/C)ACC	*Phaseolus vulgaris*	NA
AE-box	AGAAACTT	*Arabidopsis thaliana*	part of a module for light response
AE-box	AGAAACAA	*Arabidopsis thaliana*	part of a module for light response
as-1	TGACG	*Arabidopsis thaliana*	NA
AT1-motif	AATTATTTTTTATT	*Solanum tuberosum*	part of a light responsive module
ATC-motif	AGTAATCT	*Spinacia oleracea*	part of a conserved DNA module involved in light responsiveness
ATCT-motif	AATCTAATCC	*Pisum sativum*	part of a conserved DNA module involved in light responsiveness
Box 4	ATTAAT	*Petroselinum crispum*	part of a conserved DNA module involved in light responsiveness
Box II	TGGTAATAA	*Solanum tuberosum*	part of a light responsive element
CAG-motif	GAAAGGCAGAC	*Arabidopsis thaliana*	part of a light response element
CCAAT-box	CAACGG	*Hordeum vulgare*	MYBHv1 binding site
CGTCA-motif	CGTCA	*Hordeum vulgare*	cis-acting regulatory element involved in the MeJA-responsiveness
chs-CMA1a	TTACTTAA	*Daucus carota*	part of a light responsive element
chs-CMA2a	TCACTTGA	*Petroselinum crispum*	part of a light responsive element
chs-Unit 1 m1	ACCTAACCCGC	*Hordeum vulgare*	part of a light responsive element
chs-Unit 1 m1	ACCTAACCCGG	*Zea mays*	part of a light responsive element
DRE	TACCGACAT	*Arabidopsis thaliana*	\"cis-acting element involved in dehydration
DRE1	ACCGAGA	*Zea mays*	NA
G-box	TACGTG	*Arabidopsis thaliana*	cis-acting regulatory element involved in light responsiveness
G-box	CACGTG	*Arabidopsis thaliana*	cis-acting regulatory element involved in light responsiveness
G-box	CACGTC	*Zea mays*	cis-acting regulatory element involved in light responsiveness
G-box	TAACACGTAG	*Brassica oleracea*	cis-acting regulatory element involved in light responsiveness
G-box	CACGAC	*Zea mays*	cis-acting regulatory element involved in light responsiveness
G-box	CCACGTAA	*Brassica napus*	cis-acting regulatory element involved in light responsiveness
G-box	ACACGTGT	*Brassica napus*	cis-acting regulatory element involved in light responsiveness
G-box	TAAACGTG	*Brassica oleracea*	cis-acting regulatory element involved in light responsiveness
G-box	GCCACGTGGA	*Arabidopsis thaliana*	cis-acting regulatory element involved in light responsiveness
G-box	tgACACGTGGCA	*Lycopersicon esculentum*	cis-acting regulatory element involved in light responsiveness
G-box	ACACGTG(G/t)CACC	*Lycopersicon esculentum*	cis-acting regulatory element involved in light responsiveness
GA-motif	ATAGATAA	*Arabidopsis thaliana*	part of a light responsive element
Gap-box	CAAATGAA(A/G)A	*Arabidopsis thaliana*	part of a light responsive element
GATA-motif	GATAGGA	*Arabidopsis thaliana*	part of a light responsive element
GATA-motif	AAGATAAGATT	*Arabidopsis thaliana*	part of a light responsive element
GATA-motif	AAGGATAAGG	*Solanum tuberosum*	part of a light responsive element
GATA-motif	GATAGGG	*Pisum sativum*	part of a light responsive element
GATT-motif	CTCCTGATTAGC	*Zea mays*	part of a light responsive element
GATT-motif	CTCCTGATTGGA	*Oryza sativa*	part of a light responsive element
GT1-motif	GGTTAA	*Arabidopsis thaliana*	light responsive element
GT1-motif	GGTTAAT	*Avena sativa*	light responsive element
GT1-motif	GTGTGTGAA	*Solanum tuberosum*	light responsive element
I-box	TGATAATGT	*Solanum tuberosum*	part of a light responsive element
I-box	atGATAAGGTC	*Helianthus annuus*	part of a light responsive element
I-box	AGATAAGG	*Triticum aestivum*	part of a light responsive element
I-box	cCATATCCAAT	*Flaveria trinervia*	part of a light responsive element
I-box	AAGATAAGGCT	*Gossypium hirsutum*	part of a light responsive element
I-box	GATAAGGGT	*Arabidopsis thaliana*	part of a light responsive element
L-box	attctcaCCTACCA	*Petroselinum hortense*	part of a light responsive element
L-box	ATCCCACCTAC	*Petroselinum crispum*	part of a light responsive element
LAMP-element	CTTTATCA	*Pisum sativum*	part of a light responsive element
LTR	CCGAAA	*Hordeum vulgare*	cis-acting element involved in low-temperature responsiveness
MRE	AACCTAA	*Petroselinum crispum*	MYB binding site involved in light responsiveness
Myb	CAACTG	*Arabidopsis thaliana*	NA
Myb	TAACTG	*Arabidopsis thaliana*	NA
MYB recognition site	CCGTTG	*Arabidopsis thaliana*	NA
MYB-like sequence	TAACCA	*Arabidopsis thaliana*	NA
Sp1	GGGCGG	*Oryza sativa*	light responsive element
STRE	AGGGG	*Arabidopsis thaliana*	stress response element
TC-rich repeats	GTTTTCTTAC	*Nicotiana tabacum*	cis-acting element involved in defense and stress responsiveness
TC-rich repeats	ATTCTCTAAC	*Nicotiana tabacum*	cis-acting element involved in defense and stress responsiveness
TCA	TCATCTTCAT	*Pisum sativum*	NA
TCA-element	CCATCTTTTT	*Nicotiana tabacum*	cis-acting element involved in salicylic acid responsiveness
TCA-element	TCAGAAGAGG	*Brassica oleracea*	cis-acting element involved in salicylic acid responsiveness
TCCC-motif	TCTCCCT	*Spinacia oleracea*	part of a light responsive element
TCT-motif	TCTTAC	*Arabidopsis thaliana*	part of a light responsive element
TGACG-motif	TGACG	*Hordeum vulgare*	cis-acting regulatory element involved in the MeJA-responsiveness
W box	TTGACC	*Arabidopsis thaliana*	NA
WRE3	CCACCT	*Pisum sativum*	NA
WUN-motif	AAATTACT	*Nicotiana glutinosa*	wound-responsive element
WUN-motif	AAATTTCCT	*Brassica oleracea*	wound-responsive element

WUN-motif	AAATTTCTT	*Nicotiana glutinosa*	wound-responsive element
WUN-motif	AAATTACTA	*Nicotiana glutinosa*	wound-responsive element
WUN-motif	CAATTACAT	*Nicotiana glutinosa*	wound-responsive element
WUN-motif	TTATTACAT	*Nicotiana glutinosa*	wound-responsive element
WUN-motif	CCATTTCAA	*Nicotiana glutinosa*	wound-responsive element

3.3.5 High expression of OMT2 could potentially impact anthocyanin profile in edited cell cultures

The expression patterns of OMT genes were examined using RT-qPCR in both wild-type and *stisac* cell cultures, which contrasted in quality and content of anthocyanins. From the pool of 65 candidate genes, 4 genes were randomly chosen for verification. They all belong to group 1 OMTs which clustered with the tomato AOMT Solyc09g082660.2.1, an OMT gene functionally verified to methylate anthocyanins (Gomez Roldan et al., 2014). The selected genes exhibited distinct gene structures and TFBMs, along with a notable expression pattern in various RNA-Seq libraries. While they all displayed unique expression patterns in the selected RNA-Seq libraries, only one demonstrated significant expression differences in the RT-qPCR results. Specifically, the expression of OMT2 (Soltu.DM.09G025040.1) was 4.3-fold higher in edited and high-pigmented calli compared to wild-type calli (Figure 4), indicating a strong correlation with anthocyanin accumulation. In contrast, although the remaining candidates showed higher expression levels in edited cell cultures compared to wild-type (1.5-fold higher for OMT1- Soltu.DM.10G014940.1; 0.5-fold higher for OMT3 - Soltu.DM.01G047320.1; 0.4-fold higher for OMT4 - Soltu.DM.05G014010.1), these differences were not statistically significant.

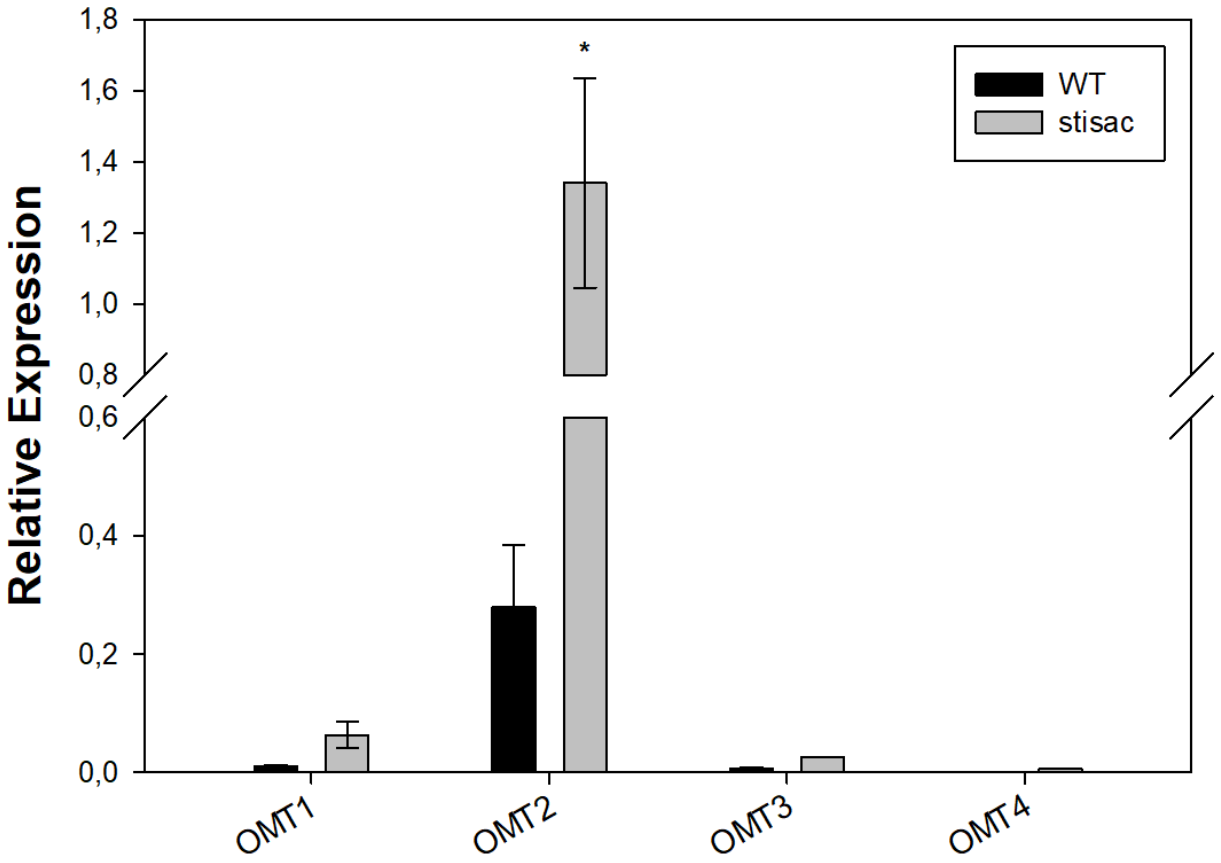

Figure 4. Quantitative analysis of transcript levels of differentially expressed genes (DEGs) in wild-type and edited calli. Each column represents the mean value ± standard error (n = 3). *p<0.05

3.4 Discussion

The cultivated potato is not only the third most important crop worldwide (Schilling et al., 2024) but may also serve as a rich source of diverse of polyphenolic compounds, particularly flavonoids and hydroxycinnamic acids, with a large diversity also of anthocyanins compared to other anthocyanin-rich species (Staiti et al., 2022). Enzymatic processes, including methylation mediated by OMTs, play a crucial role in expanding the diversity of polyphenolic molecules and their associated bioactivities toward plant physiology and human health. Enzymatic modifications, particularly methylation mediated by OMTs, play a pivotal role in the synthesis of lignin, aroma compounds, and pigments in plants. OMTs, characterized by their diverse functionalities, facilitate the transfer of methyl groups from SAM to various substrates. While a minority of OMTs exhibit strict enzymatic specificity, most plant OMTs have broad substrate acceptance. For instance, in *Camptotheca acuminata*, the 10-hydroxycamptothecin O-methyltransferase (Ca10OMT) displays catalytic promiscuity, accommodating both alkaloid and phenylpropanoid substrates with similar kinetics (Salim et al., 2018). Identifying and characterizing anthocyanin-specific OMTs presents challenges, potentially

leading to overlooked candidates. To address this, we conducted a genome-wide analysis to identify a comprehensive set of OMTs and prioritized candidates specifically for anthocyanin modifications. This study revealed 65 OMT encoding genes distributed across all potato chromosomes except for chromosome 7. Recent research indicates a widespread distribution of OMTs throughout the *Solanaceae* genomes, with considerable variation in the gene count among genera (Pezzi et al., 2023). Notably, a significant portion of these OMTs share a high degree of similarity, exceeding 80%. Our analysis also suggests that more than half of these sequences likely originated from tandem duplication events, underscoring the role of unequal crossing-over in the expansion of the OMT gene family in potato, rather than whole-genome duplication or gene transposition events. This observation aligns with findings in Arabidopsis (Kim et al., 2005) and grape (Lu et al., 2022), where a majority of OMTs were either tandemly duplicated or located within segmental duplication blocks. Natural selection capitalizes on gene duplication as a mechanism to generate novel or specialized gene functions, thereby providing organisms with the flexibility to adapt to changing environmental conditions (Magadum et al., 2013). Various mechanisms contribute to the generation of duplicated genes, with polyploidy emerging as a significant process shaping the evolution of higher organisms (Levy & Feldman, 2002). *S. tuberosum*, is an autotetraploid ($2n = 4x = 48$). It likely acquired its polyploid nature through duplication events (Kyriakidou et al., 2020). Given the genomic characteristics of the potato and the substrate specificity and region selectivity of OMTs, it becomes apparent why the OMT family is significantly expanded. The same OMT can potentially be involved in methylating different substrates, making it challenging to predict the specific functions of these proteins. Previous studies often grouped OMTs based on phylogenetic relationships. In this work, we aimed to group OMT-encoding genes based on functional traits identified through InterPro domains, using the *in silico* functional classification to describe our results. The four identified domains shared the same molecular function (O-methyltransferase activity; GO:000817). In contrast, IPR012967 is associated with a different molecular function, classified as protein dimerization activity (GO:0046983). This domain is found at the N terminus of various plant O-methyltransferases and mediates the dimerization of these proteins (Zubieta et al., 2001). Previous studies have used accurate identification of OMT domains to cross-validate and enhance gene annotation accuracy (Ali et al., 2024). The arbitrary classification of OMT groups based on InterPro signatures revealed variations in the number of characteristic domains for each group. Specifically, group 1 comprised proteins with a single domain, while the remaining groups contained proteins with two or three domains. While proteins with a single domain in plants may exhibit specificity, the presence of multiple domains allows for increased functional diversity and adaptability. Plant genomes are believed to contain a core set of conserved single-domain proteins (Esch et al., 2022), whereas multidomain proteins have

undergone complex evolutionary processes involving domain evolution, recruitment, recombination, and losses, leading to functional convergence (Zhang et al., 2012). Phylogenetic distribution analysis showed that among the four distinct OMTs subgroups identified, group 1 formed a singular clade alongside the tomato baits utilized. Different tissues in plants typically harbor distinct specialized metabolites, and diverse OMTs perform multiple functions for specific metabolites (Y. Liu et al., 2022). Consequently, the differential expression patterns of OMTs across various tissues provide valuable insights into their contributions to plant growth and development (see Supplementary Figure 1). Following phylogenic analysis together with domain classification, OMTs belonging to the same subclades exhibited similar tissue-specific and stress-responsive expression patterns, particularly within group 1. To comprehensively understand the regulation of potato OMT gene expression, cis-regulatory elements were analyzed. These elements encode the genomic blueprints that ensure precise patterns of gene expression, aligning with specific roles, conditions, and developmental stages (Marand et al., 2023). Results revealed the presence of conserved motifs (BOX4, STRE, TGACG-motif) implicated in stress and defense responsiveness. The genome-wide identification of OMTs prompted us to focus more closely on those potentially involved in anthocyanin modification. The characterizing domain "Class I-like SAM-dependent O-methyltransferase" found in group 1 suggests features specific to anthocyanin-OMTs, differing from flavonoid-OMTs in their requirement for a cation for catalytic activity (Lashley et al., 2022). Also, group 1 OMTs exhibit stress-responsive elements and patterns associated with anthocyanin production. For instance, AOMT2 harbored 7 copies of STRE (STress Response Element; AGGGG), which are rapidly induced following various stress treatments such as heat, nitrogen starvation, low external pH, osmotic stress, and oxidative stress (Walley et al., 2007). Reactive oxygen species (ROS), acting as signaling molecules, participate in the stress signaling pathway by regulating transcription factors like MYB, bHLH and WD40 through the mitogen-activated protein kinase (MAPK) pathway, ultimately leading to anthocyanin synthesis (Naing & Kim, 2021). Previous studies have indicated the involvement of multiple OMTs in anthocyanin methylation in potato. Zhang et al. (2022) observed significantly lower expression of the methyltransferase gene OMT30376 (Soltu.DM.09G009390.1) in red tubers compared to purple ones, alongside the absence of methylated forms of anthocyanins such as malvidin-3-O-glucoside and peonidin-3-O-glucoside. Previous studies have primarily focused on the role of OMTs in anthocyanin methylation, contributing to colour changes in fruits or flowers (Akita et al., 2011; Nakamura et al., 2015; Xie et al., 2022). However, variations in expression of different OMTs in plant cell cultures have not been extensively explored. Our study sheds light on this aspect by revealing the differential expression of AOMT2 (Soltu.DM.09G025040.1) in edited and highly-pigmented cell cultures, suggesting its involvement in anthocyanin transformation (see D'Amelia et al. (2022)).

This underscores the significance of investigating decorative genes within biotechnological platforms such as cell cultures, offering new perspectives on OMT genes as valuable tools for tailored anthocyanin production. Previous research has demonstrated the potential of OMTs in engineering transgenic purple roses (Nakamura et al., 2015), highlighting their capacity for innovation in artificial germplasm.

3.5 Conclusion

In summary, this study identified and characterized a total of 65 OMTs in potato, providing a comprehensive understanding of their phylogenetic relationships, conserved domains, and gene structures. Through expression profiling analysis, we uncovered tissue-preferential and stress-related OMT genes, suggesting functional divergence within this gene family. Notably, specific expression patterns observed in highly-pigmented cell cultures, along with a strong correlation with methylated anthocyanins and responsiveness to both biotic and abiotic stresses, highlight the involvement of AOMT2 (Soltu.DM.09G025040.1) in anthocyanin methylation. Further insights into the regulation of anthocyanin biosynthesis were gained through the analysis of cis-acting elements of Soltu.DM.09G025040.1. These findings hold significant implications, not only for enhancing quality through genetic engineering in potatoes but also for the production of variegated and stable natural pigments in cell cultures. This opens avenues for the production of tailor-made anthocyanins, which hold promise for various applications in food, cosmetics, and pharmaceutical industries.

Data availability

The 57 tomato OMT sequences were available in the Sol Genomics Network (SGN) database, https://solgenomics.net/ftp/tomato_genome/Heinz1706/annotation/ITAG2.3_release/. The genome assembly and annotation (v6.1) for the doubled monoploid potato DM 1-3 516 R44 were available in the Spud DB database, http://spuddb.uga.edu/dm_v6_1_download.shtml. All data generated or analysed during this study are included in this published article and its supplementary information files.

REFERENCES

Akita, Y., Kitamura, S., Hase, Y., Narumi, I., Ishizaka, H., Kondo, E., Kameari, N., Nakayama, M., Tanikawa, N., Morita, Y., & Tanaka, A. (2011). Isolation and characterization of the fragrant cyclamen O-methyltransferase involved in flower coloration. *Planta*, *234*(6), 1127-1136. https://doi.org/10.1007/s00425-011-1466-0

Alappat, B., & Alappat, J. (2020). Anthocyanin Pigments: Beyond Aesthetics. *Molecules*, *25*(23). https://doi.org/10.3390/molecules25235500

Ali, A., Raza, A., Janiad, S., Rehman, S., Riaz, A., Khan, A., Deeba, F., Alalawy, A. I., Sakran, M., Alasmari, A., Moosa, A., & Zulfiqar, F. (2024). Characterization of rice O-methyltransferase genes and their presumed homologs in and. *Gene Reports*, *35*. https://doi.org/ARTN 10189410.1016/j.genrep.2024.101894

Ayvaz, H., Cabaroglu, T., Akyildiz, A., Pala, C. U., Temizkan, R., Agcam, E., Ayvaz, Z., Durazzo, A., Lucarini, M., Direito, R., & Diaconeasa, Z. (2022). Anthocyanins: Metabolic Digestion, Bioavailability, Therapeutic Effects, Current Pharmaceutical/Industrial Use, and Innovation Potential. *Antioxidants (Basel)*, *12*(1). https://doi.org/10.3390/antiox12010048

D'Amelia, V., Staiti, A., D'Orso, F., Maisto, M., Piccolo, V., Aversano, R., & Carputo, D. (2022). Targeted mutagenesis of StISAC stabilizes the production of anthocyanins in potato cell culture. *Plant Direct*, *6*(8), e433. https://doi.org/10.1002/pld3.433

Do, C. T., Pollet, B., Thevenin, J., Sibout, R., Denoue, D., Barriere, Y., Lapierre, C., & Jouanin, L. (2007). Both caffeoyl Coenzyme A 3-O-methyltransferase 1 and caffeic acid O-methyltransferase 1 are involved in redundant functions for lignin, flavonoids and sinapoyl malate biosynthesis in Arabidopsis. *Planta*, *226*(5), 1117-1129. https://doi.org/10.1007/s00425-007-0558-3

Du, H., Wu, J., Ji, K. X., Zeng, Q. Y., Bhuiya, M. W., Su, S., Shu, Q. Y., Ren, H. X., Liu, Z. A., & Wang, L. S. (2015). Methylation mediated by an anthocyanin, O-methyltransferase, is involved in purple flower coloration in Paeonia. *J Exp Bot*, *66*(21), 6563-6577. https://doi.org/10.1093/jxb/erv365

Eddy, S. R. (2011). Accelerated Profile HMM Searches. *PLoS Comput Biol*, *7*(10), e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics*, *5*, 113. https://doi.org/10.1186/1471-2105-5-113

Esch, L., Kirsch, C., Vogel, L., Kelm, J., Huwa, N., Schmitz, M., Classen, T., & Schaffrath, U. (2022). Pathogen Resistance Depending on Jacalin-Dirigent Chimeric Proteins Is Common among Poaceae but Absent in the Dicot Arabidopsis as Evidenced by Analysis of Homologous Single-Domain Proteins. *Plants (Basel)*, *12*(1). https://doi.org/10.3390/plants12010067

Fournier-Level, A., Hugueney, P., Verries, C., This, P., & Ageorges, A. (2011). Genetic mechanisms underlying the methylation level of anthocyanins in grape (Vitis vinifera L.). *BMC Plant Biol*, *11*, 179. https://doi.org/10.1186/1471-2229-11-179

Gomez Roldan, M. V., Outchkourov, N., van Houwelingen, A., Lammers, M., Romero de la Fuente, I., Ziklo, N., Aharoni, A., Hall, R. D., & Beekwilder, J. (2014). An O-methyltransferase modifies accumulation of methylated anthocyanins in seedlings of tomato. *Plant J*, *80*(4), 695-708. https://doi.org/10.1111/tpj.12664

Gu, Z., Eils, R., & Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, *32*(18), 2847-2849. https://doi.org/10.1093/bioinformatics/btw313

Holub, E. B. (2001). The arms race is ancient history in Arabidopsis, the wildflower. *Nature Reviews Genetics*, *2*(7), 516-527. https://doi.org/10.1038/35080508

Hulo, N., Bairoch, A., Bulliard, V., Cerutti, L., De Castro, E., Langendijk-Genevaux, P. S., Pagni, M., & Sigrist, C. J. (2006). The PROSITE database. *Nucleic Acids Res*, *34*(Database issue), D227-230. https://doi.org/10.1093/nar/gkj063

Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol*, *30*(4), 772-780. https://doi.org/10.1093/molbev/mst010

Khalil, M. N., Brandt, W., Beuerle, T., Reckwell, D., Groeneveld, J., Hansch, R., Gaid, M. M., Liu, B., & Beerhues, L. (2015). O-Methyltransferases involved in biphenyl and dibenzofuran biosynthesis. *Plant J*, *83*(2), 263-276. https://doi.org/10.1111/tpj.12885

Kim, B.-G., Kim, D.-H., Hur, H.-G., Lim, J., Lim, Y.-H., & Ahn, J.-H. (2005). O-Methyltransferases from Arabidopsis thaliana. *Applied Biological Chemistry*, *48*(3), 113-119.

Kristiansson, E., Thorsen, M., Tamás, M. J., & Nerman, O. (2009). Evolutionary Forces Act on Promoter Length: Identification of Enriched Regulatory Elements. *Molecular Biology and Evolution*, *26*(6), 1299-1307. https://doi.org/10.1093/molbev/msp040

Kyriakidou, M., Anglin, N. L., Ellis, D., Tai, H. H., & Strömvik, M. V. (2020). Genome assembly of six polyploid potato genomes. *Scientific Data*, *7*(1). https://doi.org/ARTN 88 10.1038/s41597-020-0428-4

Lakshmikanthan, M., Muthu, S., Krishnan, K., Altemimi, A. B., Haider, N. N., Govindan, L., Selvakumari, J., Alkanan, Z. T., Cacciola, F., & Francis, Y. M. (2024). A Comprehensive Review on Anthocyanin-rich foods: Insights into Extraction, Medicinal Potential, and Sustainable Applications. *Journal of Agriculture and Food Research*, 101245. https://doi.org/10.1016/j.jafr.2024.101245

Lashley, A., Miller, R., Provenzano, S., Jarecki, S.-A., Erba, P., & Salim, V. (2022). Functional Diversification and Structural Origins of Plant Natural Product Methyltransferases. *Molecules*, *28*(1), 43. https://doi.org/10.3390/molecules28010043

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., Rouzé, P., & Rombauts, S. (2002). PlantCARE, a database of plant acting regulatory elements and a portal to tools for analysis of promoter sequences. *Nucleic Acids Research*, *30*(1), 325-327. https://doi.org/DOI 10.1093/nar/30.1.325

Levy, A. A., & Feldman, M. (2002). The impact of polyploidy on grass genome evolution. *Plant Physiology*, *130*(4), 1587-1593. https://doi.org/10.1104/pp.015727

Liu, X. G., Luo, Y., Wu, H. K., Xi, W. P., Yu, J., Zhang, Q. Y., & Zhou, Z. Q. (2016). Systematic analysis of methyltransferase gene family and identification of potential members involved in the formation of methylated flavonoids in. *Gene*, *575*(2), 458-472. https://doi.org/10.1016/j.gene.2015.09.048

Liu, X. J., Zhao, C. N., Gong, Q., Wang, Y., Cao, J. P., Li, X., Grierson, D., & Sun, C. D. (2020). Characterization of a caffeoyl-CoA-methyltransferase-like enzyme involved in biosynthesis of polymethoxylated flavones in. *Journal of Experimental Botany*, *71*(10), 3066-3079. https://doi.org/10.1093/jxb/eraa083

Liu, Y., Fernie, A. R., & Tohge, T. (2022). Diversification of Chemical Structures of Methoxylated Flavonoids and Genes Encoding Flavonoid-O-Methyltransferases. *Plants (Basel)*, *11*(4). https://doi.org/10.3390/plants11040564

Liu, Y. T., Fernie, A. R., & Tohge, T. (2022). Diversification of Chemical Structures of Methoxylated Flavonoids and Genes Encoding Flavonoid-Methyltransferases. *Plants-Basel*, *11*(4). https://doi.org/ARTN 56410.3390/plants11040564

Lu, S. W., Zhuge, Y. X., Hao, T. Y., Liu, Z. J., Zhang, M. W., & Fang, J. G. (2022). Systematic analysis reveals O-methyltransferase gene family members involved in flavonoid biosynthesis in grape. *Plant Physiology and Biochemistry*, *173*, 33-45. https://doi.org/10.1016/j.plaphy.2022.01.007

Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., & Ravikesavan, R. (2013). Gene duplication as a major force in evolution. *Journal of Genetics*, *92*(1), 155-161. https://doi.org/10.1007/s12041-013-0212-8

Marand, A. P., Eveland, A. L., Kaufmann, K., & Springer, N. M. (2023). -Regulatory Elements in Plant Development, Adaptation, and Evolution. *Annual Review of Plant Biology*, *74*, 111-137. https://doi.org/10.1146/annurev-arplant-070122-030236

Mattoo, A. K., Dwivedi, S. L., Dutt, S., Singh, B., Garg, M., & Ortiz, R. (2022). Anthocyanin-Rich Vegetables for Human Consumption-Focus on Potato, Sweetpotato and Tomato. *Int J Mol Sci*, *23*(5). https://doi.org/10.3390/ijms23052634

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., Finn, R. D., & Bateman, A. (2021). Pfam: The protein families database in 2021. *Nucleic Acids Research*, *49*(D1), D412-D419. https://doi.org/10.1093/nar/gkaa913

Naing, A. H., & Kim, C. K. (2021). Abiotic stress-induced anthocyanins in plants: Their role in tolerance to abiotic stresses. *Physiol Plant*, *172*(3), 1711-1723. https://doi.org/10.1111/ppl.13373

Nakamura, N., Katsumoto, Y., Brugliera, F., Demelis, L., Nakajima, D., Suzuki, H., & Tanaka, Y. (2015). Flower color modification in by expressing the adenosylmethionine: anthocyanin 3′,5′-methyltransferase gene from. *Plant Biotechnology*, *32*(2), 109-117. https://doi.org/10.5511/plantbiotechnology.15.0205a

Nguyen, L. T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268-274. https://doi.org/10.1093/molbev/msu300

Pezzi, P. H., Goncalves, L. T., Depra, M., & Freitas, L. B. (2023). Evolution and diversification of the O-methyltransferase (OMT) gene family in Solanaceae. *Genet Mol Biol*, *46*(3 Suppl 1), e20230121. https://doi.org/10.1590/1678-4685-GMB-2023-0121

Pham, G. M., Hamilton, J. P., Wood, J. C., Burke, J. T., Zhao, H. N., Vaillancourt, B., Ou, S. J., Jiang, J. M., & Buell, C. R. (2020). Construction of a chromosome-scale long-read reference genome assembly for potato. *Gigascience*, *9*(9). https://doi.org/ARTN giaa10010.1093/gigascience/giaa100

Quevillon, E., Silventoinen, V., Pillai, S., Harte, N., Mulder, N., Apweiler, R., & Lopez, R. (2005). InterProScan: protein domains identifier. *Nucleic Acids Research*, *33*, W116-W120. https://doi.org/10.1093/nar/gki442

Salim, V., Jones, A. D., & DellaPenna, D. (2018). Camptotheca acuminata 10-hydroxycamptothecin O-methyltransferase: an alkaloid biosynthetic enzyme co-opted from flavonoid metabolism. *Plant J*, *95*(1), 112-125. https://doi.org/10.1111/tpj.13936

Schilling, F., Schumacher, C., Köhl, K., Sprenger, H., Kopka, J., Peters, R., Haas, M., Zuther, E., & Horn, R. (2024). Whole-genome sequencing of tetraploid potato varieties reveals different strategies for drought tolerance. *Scientific Reports*, *14*(1). https://doi.org/ARTN 547610.1038/s41598-024-55669-3

Staiti, A., D'Amelia, V., & Carputo, D. (2022). Exploring and exploiting anthocyanins for human needs: potato as a research case study. *Italus Hortus*, *29*(2), 18-34. https://doi.org/10.26353/j.itahort/2022.2.1834

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3--new capabilities and interfaces. *Nucleic Acids Res*, *40*(15), e115. https://doi.org/10.1093/nar/gks596

Walley, J. W., Coughlan, S., Hudson, M. E., Covington, M. F., Kaspi, R., Banu, G., Harmer, S. L., & Dehesh, K. (2007). Mechanical stress induces biotic and abiotic stress responses via a novel cis-element. *PLoS Genet*, *3*(10), 1800-1812. https://doi.org/10.1371/journal.pgen.0030172

Wang, L. G., Lam, T. T., Xu, S., Dai, Z., Zhou, L., Feng, T., Guo, P., Dunn, C. W., Jones, B. R., Bradley, T., Zhu, H., Guan, Y., Jiang, Y., & Yu, G. (2020). Treeio: An R Package for Phylogenetic Tree Input and Output with Richly Annotated and Associated Data. *Mol Biol Evol*, *37*(2), 599-603. https://doi.org/10.1093/molbev/msz240

Xie, L., Lu, Y., Zhou, Y., Hao, X., & Chen, W. (2022). Functional Analysis of a Methyltransferase Involved in Anthocyanin Biosynthesis from Blueberries (Vaccinium corymbosum). *J Agric Food Chem*, *70*(51), 16253-16262. https://doi.org/10.1021/acs.jafc.2c06743

Yu, G. (2020). Using ggtree to Visualize Data on Tree-Like Structures. *Curr Protoc Bioinformatics*, *69*(1), e96. https://doi.org/10.1002/cpbi.96

Zhang, H. L., Zhao, Y. N., Zhao, X. J., Zhang, Z. H., Liu, J., Shi, M. H., & Song, B. T. (2022). Methylation level of potato gene regulates tuber anthocyanin transformations. *Frontiers in Plant Science*, *13*. https://doi.org/ARTN 102161710.3389/fpls.2022.1021617

Zhang, X., Yuan, W., Zhao, Y., Ren, Y., Zhao, X., & Yuan, Z. (2021). Genome-Wide Identification and Evolutionary Analysis of AOMT Gene Family in Pomegranate (Punica granatum). *Agronomy*, *11*(2), 318. https://doi.org/10.3390/agronomy11020318

Zhang, X. C., Wang, Z., Zhang, X., Le, M. H., Sun, J., Xu, D., Cheng, J., & Stacey, G. (2012). Evolutionary dynamics of protein domain architecture in plants. *BMC Evol Biol*, *12*, 6. https://doi.org/10.1186/1471-2148-12-6

Zhang, Y., Butelli, E., & Martin, C. (2014). Engineering anthocyanin biosynthesis in plants. *Curr Opin Plant Biol*, *19*, 81-90. https://doi.org/10.1016/j.pbi.2014.05.011

Zubieta, C., He, X. Z., Dixon, R. A., & Noel, J. P. (2001). Structures of two natural product methyltransferases reveal the basis for substrate specificity in plant O-methyltransferases. *Nat Struct Biol*, *8*(3), 271-279. https://doi.org/10.1038/85029

Supplemental Material

Figure S1. Expression profiles of AOMT gene members in different tissues and organs. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

Figure S2. Expression profiles of AOMT gene members in different abiotic stress. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

Figure S3. Expression profiles of AOMT gene members in different biotic stress. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

Table S1. Sequence ID of AOMT protein sequences used to build the phylogenetic tree.

Table S2. Primer sequences used for PCR and RT-qPCR analyses.

Table S3. RNA-Seq public data used to extract expression patterns of AOMT gene family members.

SUPPLEMENTARY TABLE IN CSV

Supplementary table 4. TPM (transcripts per million) values related to tissue-specific category.

Supplementary table 5. TPM (transcripts per million) values related to abiotic stress category.

Supplementary table 6. TPM (transcripts per million) values related to biotic stress category.

Supplementary table 7. Cis-acting elements scanned into the OMT promoter sequences.

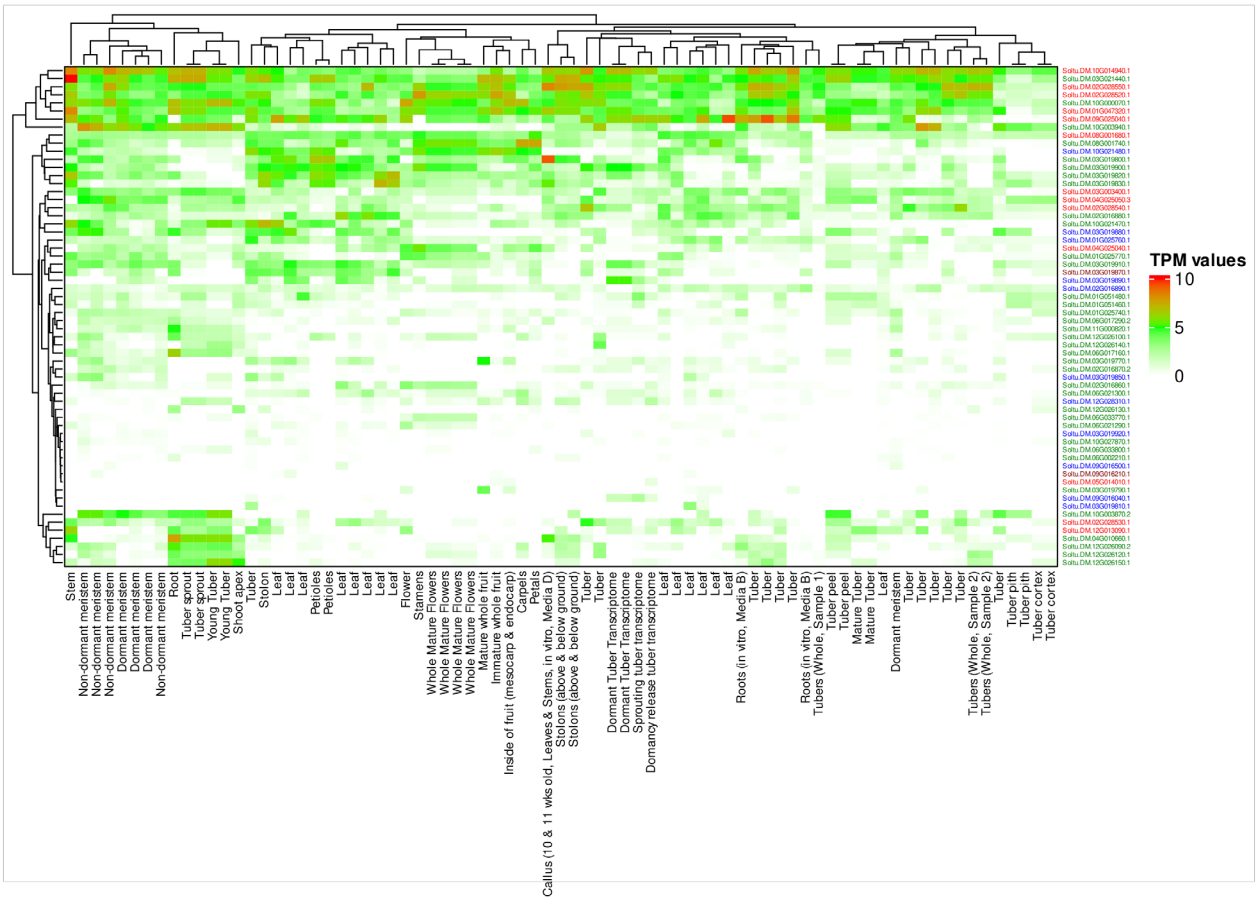

Figure S1. Expression profiles of AOMT gene members in different tissues and organs. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

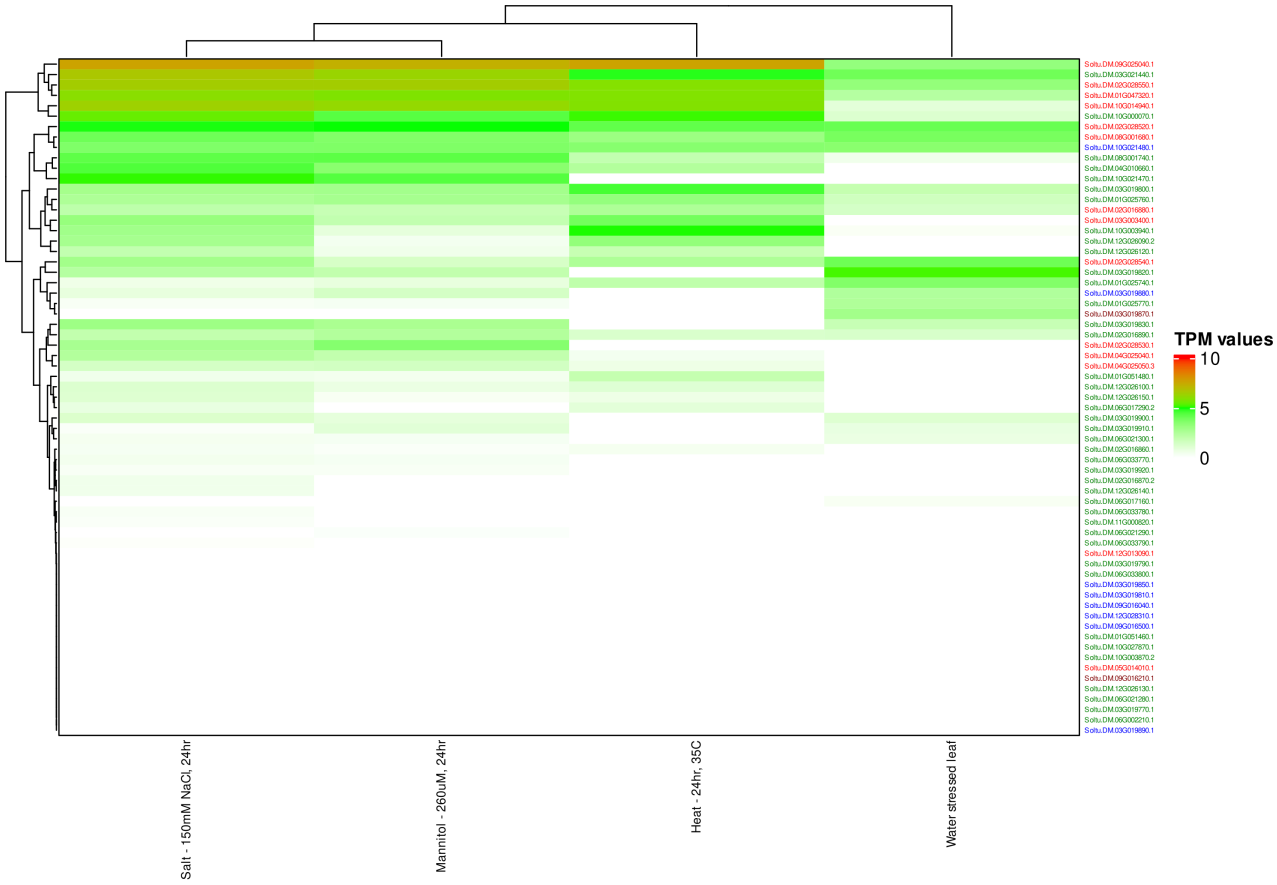

Figure S2. Expression profiles of AOMT gene members in different abiotic stress. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

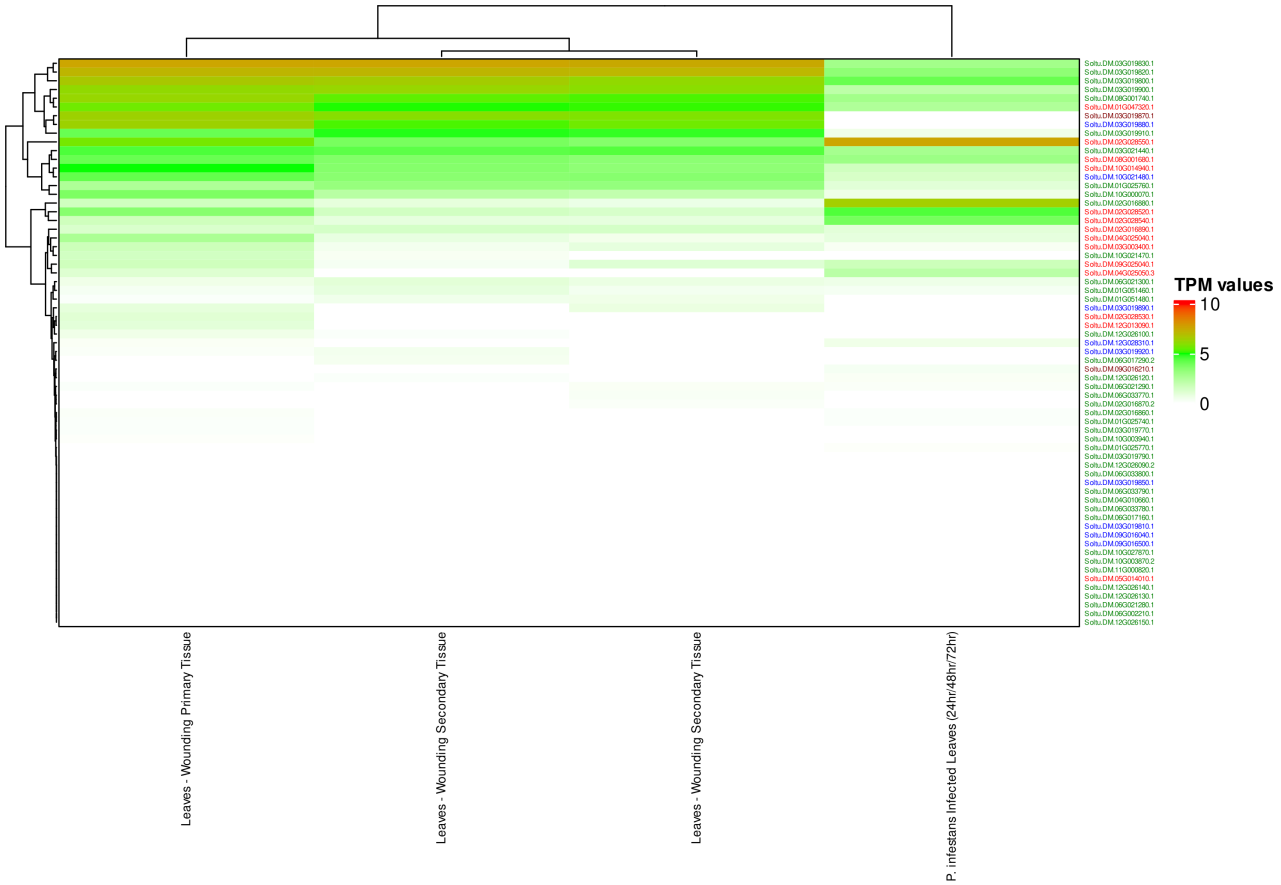

Figure S3. Expression profiles of AOMT gene members in different biotic stress. Potato AOMT were divided into four different categories and their ID were coloured with different colours (red ID = "Group 1", blue ID = "Group 2", brown ID = "Group 3", green ID = "Group 4").

Table S1. Sequence ID of AOMT protein sequences used to build the phylogenetic tree.

Sequence ID		
Soltu.DM.01G025740.1	Soltu.DM.06G033780.1	Solyc02g084950.2.1
Soltu.DM.01G025760.1	Soltu.DM.06G033790.1	Solyc02g091140.2.1
Soltu.DM.01G025770.1	Soltu.DM.06G033800.1	Solyc03g097700.2.1
Soltu.DM.01G047320.1	Soltu.DM.08G001680.1	Solyc04g055260.2.1
Soltu.DM.01G051460.1	Soltu.DM.08G001740.1	Solyc04g063210.2.1
Soltu.DM.01G051480.1	Soltu.DM.09G016040.1	Solyc04g080660.2.1
Soltu.DM.02G016860.1	Soltu.DM.09G016210.1	Solyc05g026330.1.1
Soltu.DM.02G016870.2	Soltu.DM.09G016500.1	Solyc05g026350.1.1
Soltu.DM.02G016880.1	Soltu.DM.09G025040.1	Solyc05g041320.1.1
Soltu.DM.02G016890.1	Soltu.DM.10G000070.1	Solyc05g041610.1.1
Soltu.DM.02G028520.1	Soltu.DM.10G003870.2	Solyc05g041620.1.1
Soltu.DM.02G028530.1	Soltu.DM.10G003940.1	Solyc05g041640.1.1

Soltu.DM.02G028540.1	Soltu.DM.10G014940.1	Solyc05g041690.1.1
Soltu.DM.02G028550.1	Soltu.DM.10G021470.1	Solyc06g007960.2.1
Soltu.DM.03G003400.1	Soltu.DM.10G021480.1	Solyc06g060200.1.1
Soltu.DM.03G019770.1	Soltu.DM.10G027870.1	Solyc06g064500.2.1
Soltu.DM.03G019790.1	Soltu.DM.11G000820.1	Solyc06g064510.2.1
Soltu.DM.03G019800.1	Soltu.DM.12G013090.1	Solyc06g083450.2.1
Soltu.DM.03G019810.1	Soltu.DM.12G026090.2	Solyc07g064990.2.1
Soltu.DM.03G019820.1	Soltu.DM.12G026100.1	Solyc08g006830.2.1
Soltu.DM.03G019830.1	Soltu.DM.12G026120.1	Solyc09g056230.1.1
Soltu.DM.03G019850.1	Soltu.DM.12G026130.1	Solyc09g082660.2.1
Soltu.DM.03G019870.1	Soltu.DM.12G026140.1	Solyc09g091530.1.1
Soltu.DM.03G019880.1	Soltu.DM.12G026150.1	Solyc09g091540.1.1
Soltu.DM.03G019890.1	Soltu.DM.12G028310.1	Solyc09g091550.2.1
Soltu.DM.03G019900.1	Solyc00g029190.2.1	Solyc10g005060.2.1
Soltu.DM.03G019910.1	Solyc00g033460.2.1	Solyc10g008120.2.1
Soltu.DM.03G019920.1	Solyc01g005230.2.1	Solyc10g047520.1.1
Soltu.DM.03G021440.1	Solyc01g005350.2.1	Solyc10g049500.1.1
Soltu.DM.04G010660.1	Solyc01g005360.2.1	Solyc10g049540.1.1
Soltu.DM.04G025040.1	Solyc01g014320.2.1	Solyc10g050160.1.1
Soltu.DM.04G025050.3	Solyc01g068550.2.1	Solyc10g061840.1.1
Soltu.DM.05G014010.1	Solyc01g080970.2.1	Solyc10g079540.1.1
Soltu.DM.06G002210.1	Solyc01g080990.2.1	Solyc10g085830.1.1
Soltu.DM.06G017160.1	Solyc01g081340.2.1	Solyc11g072870.1.1
Soltu.DM.06G017290.2	Solyc01g107910.2.1	Solyc12g009110.1.1
Soltu.DM.06G021280.1	Solyc01g111900.2.1	Solyc12g014500.1.1
Soltu.DM.06G021290.1	Solyc02g077510.2.1	Solyc12g041940.1.1
Soltu.DM.06G021300.1	Solyc02g077520.2.1	Solyc12g041960.1.1
Soltu.DM.06G033770.1	Solyc02g077530.1.1	

Table S2. Primer sequences used for PCR and RT-qPCR analyses.

Primer name	Primer sequence	Genebank ID
AOMT1-F	TTATCCAAGAGAGCATGAGCAA	Soltu.DM.10G014940.1
AOMT1-R	AAGAAATTGGCCTTCATCTTGA	
AOMT2-F	TGATGCACCGATGAGAAAATAC	Soltu.DM.09G025040.1
AOMT2-R	ATCTCGACCCTCTGATCTTCTG	
AOMT3-F	AACATCTGCTGATGAAGGACAA	Soltu.DM.01G047320.1
AOMT3-R	CCAGTGTAGACACCAATCTCCA	
AOMT4-F	TCGAGATGCATACGAGATGG	Soltu.DM.05G014010
AOMT4-R	TCATCAAGGGCTGATAATGC	
Ef1α-F	ATTGGAAACGGATATGCTCCA	Soltu.DM.06G005620.1
Ef1α-R	TCCTTACCTGAACGCCTGTCA	

Table S3. RNA-Seq public data used to extract expression patterns of AOMT gene family members.

Library accession	Cultivar	Sample
ERR029909	RH89-039-16	Flower
ERR029910	RH89-039-16	Leaf
ERR029912	RH89-039-16	Shoot apex
ERR029913	RH89-039-16	Stem
ERR029914	RH89-039-16	Stolon
ERR029915	RH89-039-16	Young Tuber
ERR029916	RH89-039-16	Mature Tuber
ERR029917	RH89-039-16	Root
ERR029919	RH89-039-16	Water stressed leaf
ERR029920	RH89-039-16	Tuber pith
ERR029921	RH89-039-16	Tuber peel
ERR029923	RH89-039-16	Tuber sprout
ERR029924	RH89-039-16	Tuber cortex
SRR1039535	Longshu3	Dormant Tuber Transcriptome
SRR1103933	Longshu3	Domancy release tuber transcriptome
SRR1103934	Longshu3	Sprouting tuber transcriptome

SRR122110	DM 1-3 R44	Inside of fruit (mesocarp & endocarp)
SRR122111	DM 1-3 R44	Stamens
SRR122115	DM 1-3 R44	Heat - 24hr, 35C
SRR122116	DM 1-3 R44	P. infestans Infected Leaves (24hr/48hr/72hr)
SRR122120	DM 1-3 R44	Salt - 150mM NaCl, 24hr
SRR122124	DM 1-3 R44	Whole Mature Flowers
SRR122127	DM 1-3 R44	Immature whole fruit
SRR122128	DM 1-3 R44	Mannitol - 260uM, 24hr
SRR122129	DM 1-3 R44	Leaves - Wounding Secondary Tissue
SRR122130	DM 1-3 R44	Petioles
SRR122132	DM 1-3 R44	Mature whole fruit
SRR122134	DM 1-3 R44	Leaves - Wounding Secondary Tissue
SRR122137	DM 1-3 R44	Carpels
SRR122139	DM 1-3 R44	Petioles
SRR124131	DM 1-3 R44	Whole Mature Flowers
SRR124132	DM 1-3 R44	Petals
SRR124138	DM 1-3 R44	Leaves - Wounding Primary Tissue
SRR1584263	Russet Burbank	Dormant meristem
SRR1584264	Russet Burbank	Dormant meristem
SRR1584265	Russet Burbank	Dormant meristem
SRR1584266	Russet Burbank	Dormant meristem
SRR1584267	Russet Burbank	Non-dormant meristem
SRR1584268	Russet Burbank	Non-dormant meristem
SRR1584269	Russet Burbank	Non-dormant meristem
SRR1584270	Russet Burbank	Non-dormant meristem
SRR2067490	Group Phujera	Leaf
SRR2067491	Group Phujera	Leaf
SRR2067492	Group Phujera	Leaf
SRR2067493	Group Phujera	Leaf
SRR2068271	Group Phujera	Leaf
SRR2068274	Group Phujera	Leaf
SRR2068278	Group Phujera	Leaf
SRR2068282	Group Phujera	Leaf
SRR2068285	Group Phujera	Leaf

SRR2068288	Group Phujera	Leaf
SRR2068291	Group Phujera	Leaf
SRR2068294	Group Phujera	Leaf
SRR2068315	Group Phujera	Tuber
SRR2068320	Group Phujera	Tuber
SRR2068323	Group Phujera	Tuber
SRR2068327	Group Phujera	Tuber
SRR2068328	Group Phujera	Tuber
SRR2068333	Group Phujera	Tuber
SRR2068335	Group Phujera	Tuber
SRR2069933	Group Phujera	Tuber
SRR2069934	Group Phujera	Tuber
SRR2069936	Group Phujera	Tuber
SRR2069948	Atlantic	Leaf
SRR2070071	Atlantic	Tuber
SRR2070075	Superior	Leaf
SRR2070076	Superior	Tuber
SRR2071272	Group Phujera	Tuber