Planta (2023) 258:50
https://doi.org/10.1007/s00425-023-04200-5

ORIGINAL ARTICLE

Functional characterization of DcMYB11, an R2R3 MYB associated with the purple pigmentation of carrot petiole

Vincenzo D'Amelia[1,5] · Julien Curaba[2] · Muhammad Ali Abid[2] · Salvatore Esposito[3] · Pablo Cavagnaro[4] · Domenico Carputo[5] · Massimo Iorizzo[2]

Received: 29 May 2023 / Accepted: 4 July 2023 / Published online: 24 July 2023

Main conclusion ***DcMYB11*****, an R2R3 MYB gene associated with petiole anthocyanin pigmentation in carrot, was functionally characterized. A putative enhancer sequence is able to increase DcMYB11 activity.**

Abstract The accumulation of anthocyanin pigments can exhibit different patterns across plant tissues and crop varieties. This variability allowed the investigation of the molecular mechanisms behind the biosynthesis of these pigments in several plant species. Among crops, carrots have a well-defined anthocyanin pigmentation pattern depending on the genic background. In this work, we report on the discovery of DNA structural differences affecting the activity of an R2R3 MYB (encoded by *DcMYB11*) involved in anthocyanin regulation in carrot petiole. To this end, we first verified the function of *DcMYB11* using heterologous systems and identified three different alleles which may explain differences in petiole pigmentation. Characterization of the *DcMYB11* alleles at the 5' upstream sequence unveiled a sequence that functions as a putative enhancer. In conclusion, this study provides novel insight into the molecular mechanisms controlling anthocyanin accumulation in carrot. By these outcomes, we expanded our knowledge on the cis-regulatory sequences in plants.

Keywords Anthocyanins · Carrot petiole · *Daucus carota* · Regulatory sequences · Transcription factor

Abbreviations

Gp Green petiole
Pp Purple petiole

Communicated by Dorothea Bartels.

Vincenzo D'Amelia and Julien Curaba have contributed equally to this work.

✉ Massimo Iorizzo
miorizz@ncsu.edu

1 Institute of Bioscience and BioResources, National Research Council, Via Università 100, 80055 Portici, Italy

2 Horticulture Science Department, Plants for Human Health Institute, North Carolina State University, North Carolina Research Campus, Kannapolis, NC, USA

3 CREA Research Centre for Cereal and Industrial Crops (CREA-CI), 71122 Foggia, Italy

4 National Scientific and Technical Research Council (CONICET), National Institute of Agricultural Technology (INTA) E.E.A. La Consulta, La Consulta CC8, San Carlos, 5567 Mendoza, Argentina

5 Department of Agricultural Sciences, University of Naples Federico II, Via Università 100, 80055 Portici, Italy

Introduction

Anthocyanins are pink, red, purple, and blue natural pigments widely spread within the land plants. These flavonoid compounds have various roles in the eco-physiology of plants due to their antioxidant and light-absorbing properties. Also, beneficial effects on human health have been found and reported for different types of anthocyanins extracted from different plant matrices. The role of anthocyanins in plants and their bioactivity in human cells have been reviewed in recent papers of Davies et al. (2018), Agati et al. (2021), and He et al. (2022). Anthocyanins' accumulation in different plant tissues creates a wide range of pigmentation patterns, which has been used as an easy-to-phenotype trait in plant genetic studies. As a result, extensive progresses were made to understand the genes involved in their biosynthesis and molecular mechanisms regulating this pathway (Passeri et al. 2016). The expression of anthocyanin biosynthetic genes is regulated by a combination of *cis*- and *trans*-regulatory elements. For the majority of seed plants, a ternary complex (MBW) of transcription factors (TFs) formed by R2R3 myeloblastosis (MYB), basic helix–loop–helix (bHLH), and WD40

repeat proteins control the expression of the genes encoding for enzymes involved in anthocyanin biosynthesis (Baudry et al. 2004; Koes et al. 2005; Ramsey and Glover 2005; Gonzalez et al. 2008). Especially, for dicot species, the R2R3 MYBs (belonging to subgroup 6; Stracke et al. 2001) are fundamental for the MBW complex activation, according to the hierarchical regulatory mechanism described by Alberts et al. (2014). In this mechanism, the R2R3 MYBs activate the expression of the bHLH interacting co-partners to trigger tissue pigmentation. Generally, the genes encoding for R2R3 MYBs, controlling anthocyanin biosynthesis, are present in multiple paralogs within plant genomes and each paralog often regulate the pigmentation of a specific tissue. They can also differently condition the amount and the timing of anthocyanin accumulation, depending upon the surrounding environment and/or the developmental stage (Zhang et al. 2000; Chagné et al. 2013; D'Amelia et al. 2018; Zhang et al. 2019). For example, four main types of R2R3 MYB-encoding genes have been described in *Petunia*. *PhAN2* (Anthocyanin 2) is responsible for the red coloration of corolla, *PhAN4* controls pigmentation of the anther and corolla tubes, *PhDPL* (Deep Purple) determines the color pattern venation, and *PhPHZ* (Purple Haze) is responsible for pigmentation of stems, leaves, and sepals in response to light (Spelt et al. 2000; Albert et al. 2011, 2014). Clustered MYB genes, placed in the loci *Developer* (*D*) or *Anthocyanin fruit (Aft)*, contribute to the pigmentation of specific tissues in potato (locus *D*) and tomato (locus *Aft*) plants. In potato, *StAN1* contributes to tuber skin and leaves pigmentation, while *StFlAN2* regulates anthocyanin production in corollas (Jung et al. 2009; D'Amelia et al. 2014; Laimbeer et al. 2020). In tomato, *SlAN2* is associated with the accumulation of anthocyanins in vegetative tissues, while *SlAN2-like* regulates anthocyanin accumulation in the peel of fruits (Colanero et al. 2020a, b).

The cultivated carrot (*Daucus carota* subsp. *sativus*) exhibits a wide and well-defined combination of anthocyanin pigmentation patterns within varieties and across tissues. Carrots with purple taproot represent the first form of domesticated carrot. It has been documented that carrot grown in Central Asia in the tenth AD century, where the "purple trait" was likely among traits subjected to the first human selection (Ellison et al. 2018). Today, the cultivated germplasm includes several purple/black pigmented varieties. In the carrot taproot, anthocyanins can accumulate in specific tissues including the epidermis, outer phloem, inner phloem, and xylem (Iorizzo et al. 2020). Besides the taproot, anthocyanin pigmentation can be present in other organs and tissues, including the bracts of inflorescences, the flower petals, the seeds, the leaf petiole and lamina, and the nodes and internodes of the flower stalk (Cavagnaro et al. 2019). Such visual diversity suggests a relevant variability in the regulatory mechanisms controlling the activation of the anthocyanin biosynthetic pathway in different tissues.

Genetic studies indicated that the dominant P3 locus, initially identified by Cavagnaro et al. (2014), is the main locus controlling anthocyanin accumulation in the root and petiole (Iorizzo et al. 2019). Within the region harboring the P3 locus, a cluster of six genes encoding for anthocyanin R2R3 MYBs, named as *DcMYB 6*, *7*, *8*, 9, *10,* and *11*, was identified (Iorizzo et al. 2019, 2020). Clustered genes are prone to an evolutionary divergent functionalization (Zhang et al. 2019). It is likely that also the structural diversification of the MYBs, within the P3 locus, affects the timing, the localization, and the degree of anthocyanin accumulation across the tissue of carrots. Studies using the conventional genetic, transgenic and gene editing approaches provided evidence that *DcMYB7* is the master regulator of anthocyanin accumulation in the taproot (Iorizzo et al. 2019; Bannoud et al. 2019; Xu et al. 2019). The ability of *DcMYB6* to activate the anthocyanin pathway has also been demonstrated using heterologous system, but its activity has not been associated to specific carrot tissues (Xu et al. 2017). The genes *DcMYB8* and *9* appeared to be not expressed (Iorizzo et al. 2020), while *DcMYB11* was identified as the best candidate gene controlling anthocyanin accumulation in the carrot petiole (Iorizzo et al. 2019). The role of *DcMYB11* in controlling petiole pigmentation was studied in a mapping population that segregated 3:1 for purple petiole (Pp) *vs.* green petiole (Gp), while the roots were fixed for orange color (Iorizzo et al. 2019). However, the functionality of this gene has not been proven, and polymorphisms between dominant and recessive alleles have not been identified nor functionally characterized yet. To close this gap, the objectives of this study were to evaluate the ability of *DcMYB11* to activate the anthocyanin pathway and to identify and test the effects of putative causal mutations. The study was conducted using tobacco and Arabidopsis as heterologous systems.

Materials and methods

Plant materials

To characterize *DcMYB11 (*hereafter representing DCAR_010751), sequencing and genotyping were conducted on phenotypically contrasting plants (i.e., plants with either purple or green petioles) from a mapping population named 5723. It is an F_5 family segregating for purple petiole (Pp) *vs.* green petiole (Gp) and derived from an initial cross between a purple rooted carrot with purple petioles and an orange-rooted carrot with green petioles derived from diverse European and South American sources. The population was previously used by Iorizzo et al. (2019) to map the

purple petiole locus/phenotype and to identify *DcMYB11* as candidate gene controlling this trait. P5723 carrots were grown in 2018 at the Sandhills Research Station, Jackson Springs, NC, USA. Plants were grown using the conventional agricultural practices for carrot. gDNA and RNA were isolated from the leaf lamina and petioles, respectively.

DcMYB11 sequencing

Primers used to characterize the region spanning the *DcMYB11* gene (genic and genomic regions) were designed using Primer3 (Koressar et al. 2018) and the carrot reference genome DH1v2 (Iorizzo et al. 2016) (Table S1). Genotypic data presented by Iorizzo et al. (2019) were used to select 5723 plants that were homozygous dominant and homozygous recessive at the *DcMYB11* locus. Amplification of promoter regions was performed by PrimeStar Max DNA Polymerase (Takara, Cat. R045A) with general PCR conditions (98 °C for 10 s, 56 °C for 15 s, and 68 °C for 30 s). The *DcMYB11* coding region was amplified using two-step PCR conditions (98 °C for 10 s, 68 °C for 1 min). PCR products were cleaned using PCR cleanup and Gel extraction kit (Takara, Cat. 740609) and Sanger sequencing was performed at Genomic Sciences Laboratory, North Carolina State University.

DcMYB11 isolation and plasmid construction

The *DcMYB11* CDS obtained from 5723 plants with purple petiole phenotype (homozygous dominant) was cloned within pF35SE plasmid vector (TLB::masT::BAR::Pmas-P35S::MCS::ocsT::TRB) and placed under the control 35S CaMV (p35S::) into HindIII and XbaI sites. The expression clones were introduced into *Agrobacterium tumefaciens* strain GV3101 pMP90 for transformation. Two primers were designed to amplify about 750 bp of the 5' upstream sequence of *DcMYB11* using two different genotypes from P5723, with Gp and Pp, respectively. *DcMYB11* was placed under two different regulatory sequences containing, or not, the I481 insertion (described in this work). The obtained recombinant sequences were inserted in the XhoI and XbaI restriction sites of the plasmid pFAGM (TLB::masT::BAR::Pmas-MCS::attR::ocsT::TRB). All the cloning procedures were carried out using the NEBuilder® HiFi DNA Assembly methodology with primers listed in Table S1. Plant genotyping for the I481 insertion was carried out by the conventional PCR using as template gDNA of the 5723 segregating population (Table S1).

Real-time quantitative reverse transcriptase PCR

Total RNA isolation was performed using the Spectrum Plant Total RNA Kit (Sigma-Aldrich, www.sigmaaldrich.com) following the manufacturer's guidelines. One microgram of total RNA was reverse transcribed to cDNA using an oligo-dT(20) primer and SuperScript III reverse transcriptase (Invitrogen, www.invitronge.com). Real-time qPCR on endogenous bHLH tobacco genes (*NtAN1* and *NtJAF13*) and DcMYB11/DcMYB7 was performed as reported by Curaba et al. (2020). Primers used are reported in Table S1. Five biological replicates and three technical replicates were used for the analyses. α-Tubulin was used as endogenous control for tobacco (D'Amelia et al. 2018).

Plant transformation

Agrobacterium tumefaciens strain GV3101 pMP90 carrying the expression vector was used to stably and transiently transform Arabidopsis and tobacco (*Nicotiana tabacum*, cv Samsun) plants. Arabidopsis plants were transformed via the floral-dip method. T1 transgenic Arabidopsis plants containing the p35S::*DcMYB11* or p35S::empty vectors were identified by selection on soil using a solution of glufosinate ammonium (Basta) at 100 μg mL^{-1}. Ten independent T_1 transgenic lines showing a dark purple pigmentation phenotype were selected to generate T_2 seeds and three of these lines were used for further phenotypic analysis. Agroinfiltration of tobacco leaves was carried out in fully expanded leaves of tobacco as previously reported (D'Amelia et al. 2020). Five leaves at a similar stage from different plants were used. The same concentration (OD600) of Agrobacterium was used for the infiltration.

Yeast one-hybrid

The coding sequences (CDSs) of *DcMYB11* and *DcMYB7* were cloned into pGADT7 using the XhoI and XbaI sites and fused in frame with GAL4-activation domain (AD) to generate pGADT7-*DcMYB11Pp* and pGADT7-*DcMYB7*. The promoters of carrot plants exhibiting the purple (pPp; from – 614 to – 24 of the ATG) and the green petiole (Gp; from -147 to -24 from the ATG) phenotypes were cloned into a pAbAi Y1H bait vector and introduced into the yeast strain Y1HGold (Clontech), generating two pAbAI-specific bait-reporter yeast strains (pAbAI-pPp and pAbAI-pGp), which were obtained selecting on SD-Ura medium. Selected yeast clones (pAbAI-pPp and pAbAI-pGp) were transformed with pGADT7-*DcMYB11* and pGADT7-*DcMYB7*. Transformants were selected on SD-Leu medium and selected clones were plated on SD-Leu medium supplemented with Aureobasidin A (150 ng mL^{-1}) to check promoter-binding ability. Yeast transformation was performed according to the lithium acetate/polyethylene glycol method (Bai and Elledge 1997).

Total anthocyanin analysis

The total content of monomeric anthocyanins was determined following the pH differential method previously described by Giusti and Wrolstad (2001) and slightly modified by D'Amelia et al. (2018). Total anthocyanin concentration was expressed as mg of cyanidin-3-rutinoside equivalents per gram of dry weight.

In silico and bioinformatics analyses

All the PCR fragments were sequenced by Sanger sequencing. Alignments and phylogenetic trees were obtained using Geneious software vR11 (Biomatters, http://www.geneious.com/). Blastn and blastx searches were carried out with default parameters to perform homology research against both the carrot genome assembly DH1 v2 (Iorizzo et al. 2016) and the nucleotide collection database implemented in NCBI (http://www.ncbi.nlm.nih.gov/blast). The resulting hits were plotted on carrot chromosomes using the R library karyoploteR (Gel and Serra 2017) with default parameters. The nearest gene to each hit was retrieved using the toolset bedtools closest (Quinlan et al. 2010). Cis-acting regulatory elements were identified using the online tool PlantCare (Lescott et al. 2002). Statistical analyses were performed using XLSTAT-PRO 7.5.3 software, considering P values < 0.001 statistically significant, according to Student's t test.

Results and discussion

DcMYB11 is a functional anthocyanin-activating R2R3 MYB

To perform the functional characterization of *DcMYB11*, its full-length coding sequence was cloned from RNA extracted from 5723 plants that were homozygous for the dominant allele and had purple petiole. This sequence, named here as *DcMYB11Pp*, encoded a 286-amino acid protein containing a complete R2R3 MYB domain (Fig. 1). Its predicted amino acid sequence clustered with members of the R2R3 MYB subgroup 6 (SG6) known to regulate the anthocyanin biosynthesis in dicotyledonous plants, such as NtAN2, PhAN2, CsRuby, and VvMYBA1 (Quattrocchio et al. 1999; Stracke et al. 2001; Kobayashi et al. 2002; Pattanaik et al. 2010; Butelli et al. 2012; Fig. 1a). The alignment of DcMYB11Pp with members of the same phylogenetic group indicated that the encoded protein has all the residues necessary for its interaction with bHLH proteins from clade 3f (DL(2)xR(3)xL(6)xL(3)xR (Heim et al. 2003; Zimmermann et al. 2004; Lin-Wang et al. 2010) and the conservative protein motif ANDV described by Lin-Wang et al. (2010), (Fig. 1b). It also contains a conserved sequence motif KPxPxxL in its C-terminal domain which is typical for R2R3 MYB regulators of anthocyanin biosynthesis (Stracke et al. 2001; Lin-Wang et al. 2010). The ability of *DcMYB11* to activate anthocyanin biosynthesis was verified in heterologous systems (Fig. 2; Fig. 3). Transgenic Arabidopsis plants overexpressing (p35S) *DcMYB11Pp* had darker pigmentation on both vegetative and reproductive tissues and higher accumulation of anthocyanins than wild-type (Wt) plants (Fig. 2). In tobacco, the overexpression of *DcMYB11Pp* (p35S::*DcMYB11Pp*) induced a homogenous and consistent anthocyanin biosynthesis (Fig. 3a). It is assumed that anthocyanin R2R3 MYBs (belonging to subgroup 6; Stracke et al. 2001) need bHLH co-factors (subgroup IIIf; Heim et al. 2003) to sustain pigment accumulation by triggering the expression of anthocyanin structural genes (Gonzalez et al. 2008; Alberts et al. 2014). In the case of ectopic anthocyanin production in tobacco leaves, the exogenous MYB may recruit the endogenous bHLHs by activating their expression according to the hierarchical system (Alberts et al. 2014; Montefiori et al. 2015). In such mechanism, the R2R3 MYB (*DcMYB11Pp*) activates the expression of bHLH-type-2 genes before interacting with their proteins. Indeed, we found that the overexpression of *DcMYB11Pp* promoted an increase of the *NtAN1* transcripts, which encodes for the endogenous tobacco bHLH-type-2 (Fig. 3b). In contrast, with respect to its expression in Wt leaves, the transcript level of *NtJAF13* (bHLH type-1) did not increase in tissues expressing p35S*::DcMYB11Pp* (Fig. 3b). Although not fully demonstrated yet, it is likely that in the homologous carrot system, the anthocyanin R2R2-MYBs also operate with bHLHs in a hierarchical manner. Indeed, Xu et al. (2019) verified that *DcbHLH3* transcript levels were greatly increased when *DcMYB7* was overexpressed in carrots. Overall, these results showed that *DcMYB11Pp* is able to activate the anthocyanin pathway, supporting the hypothesis proposed by Iorizzo et al. (2019) that *DcMYB11* is functional and might be responsible for the purple petiole pigmentation of carrots. Within the P3 locus, along with *DcMYB6* and *DcMYB7, DcMYB11* is the third anthocyanin R2R3-MYB activator gene proven to be functional. These results further elucidate the diversity of the anthocyanin regulatory system that carrot established during its evolution.

Three different DcMYB11 alleles are associated with purple/green petiole pigmentation in carrot germplasm

To assess differences in the alleles of *DcMYB11Pp* that could explain the green/purple petiole phenotypes in carrot, we retrieved its sequence from the DH1 carrot reference genome (DCARv2; Iorizzo et al. 2016). DH1 is a double

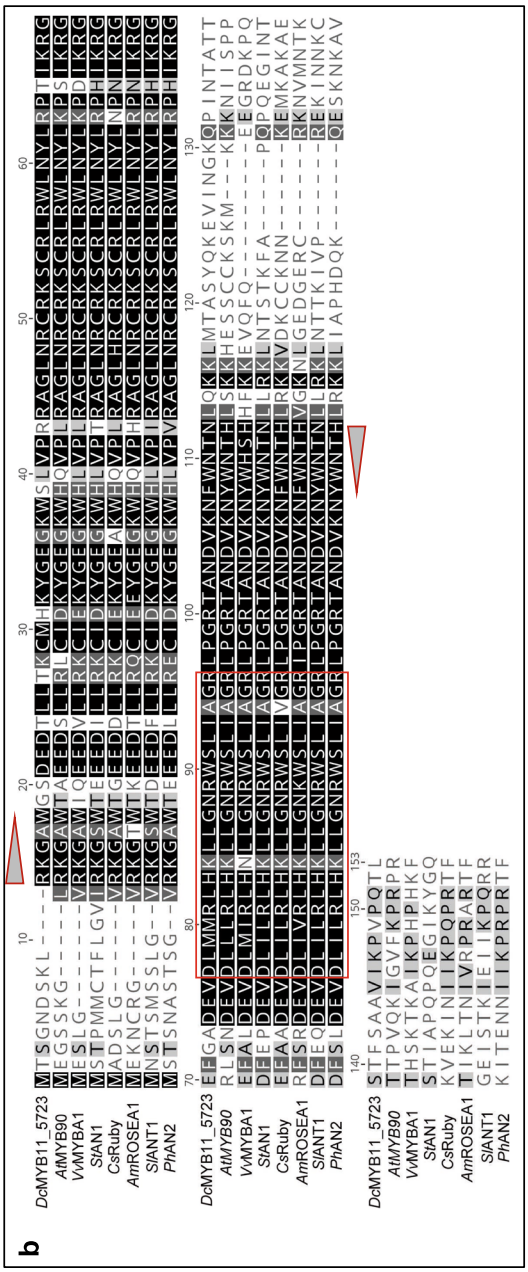

Fig. 1 *DcMYB11Pp* (isolated from purple petiole genotype of P5723) encodes for an anthocyanin R2R3 MYB protein. **a** Phylogenetic analysis indicating that predicted protein of DcMYB11Pp clusters with R2R3 MYB anthocyanin-specific members of subgroup 6. Bootstrap values at or above the 90% confidence significance are illustrated. **b** Protein alignment of DcMYB11Pp (isolated from P5723) with characterized members of the anthocyanin-specific family of MYB factors reviewed by Liu et al. (2015). Colors indicate: 100% similarity (black shading), 80–100% similarity (dark-grey shading), 60–80% similarity (light-grey shading), and less than 60% similarity (white shading). The motif needed for the interaction with bHLH proteins within the R3 MYB DNA-binding domain is highlighted with the ref box, while R2R3 motifs are highlighted by two triangles

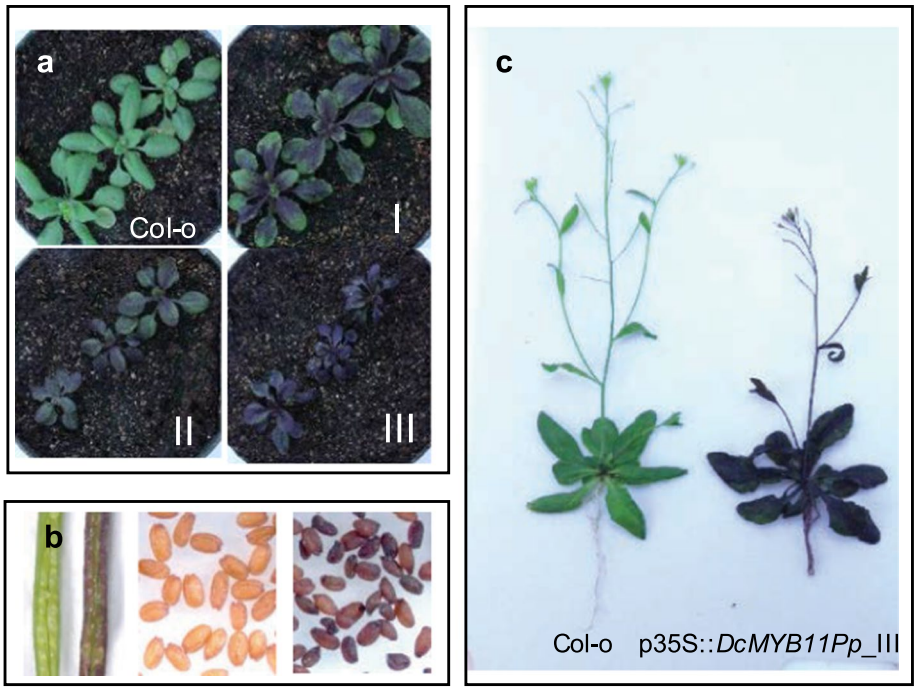

Fig. 2 Functional analysis of *DcMYB11Pp* (isolated from purple petiole genotype of P5723) in Arabidopsis. **a** Three-week-old seedlings of Col-0 and 3 independent Arabidopsis transgenic lines p35S::*DcMYB11Pp* (I, II, III). **b** Immature siliques and mature seeds of Col-0 (right) and p35S::*DcMYB11Pp* (left). **c** Five-week-old plants of Col-0 and p35S::DcMYB11Pp_III

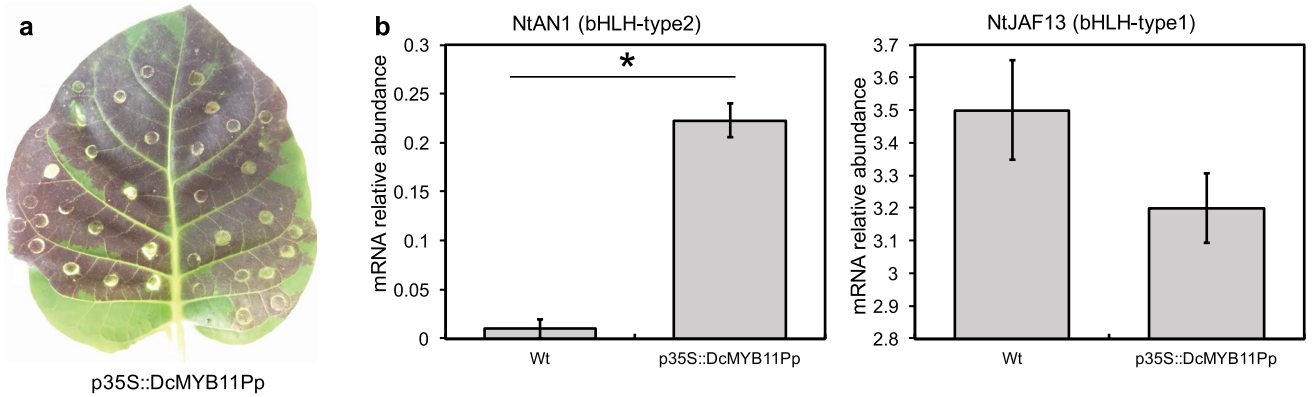

Fig. 3 Functional analysis of *DcMYB11Pp* (isolated from purple petiole genotype of P5723) in tobacco leaves. **a** Tobacco leaf infiltrated with *Agrobacterium tumefaciens* carrying p35S::*DcMYB11Pp*. **b** Transcript levels of tobacco endogenous anthocyanin bHLH co-factors *NtJAF13* (bHLH-type1) and *NtAN1* (bHLH-type2) in leaves of wild-type tobacco plants transiently overexpressing (p35S) *DcMYB11Pp*. The error bars represent SD ($n=3$). *, P values <0.05 according to Student's t test

haploid orange Nantes type carrot known to have green petioles. The domain structure of the translated sequence of *DcMYB11DH1* (DCAR_010751) appeared to represent an incorrect gene prediction (Fig. S1). Indeed, in DH1, the R2R3-binding domain was located at the C-terminal region rather than in the N-terminal part, which is the typical position of the R2R3 domain of plant MYBs controlling anthocyanin biosynthesis (Dubos et al. 2010; Fig. S2). Also, the R2R3 domain was incomplete, as evidenced by the alignment of the newly isolated DcMYB11Pp with that of the predicted protein of *DcMYB7* (Fig. S1; Fig. S2; Fig. S3). The presence of a premature stop codon after 129 nts (the stop codon replaced the TCA codon found in *DcMYB11* isolated from P5723) suggests that the predicted DcMYB11 protein in DH1 encodes for a truncated form of R2R3 MYB protein (Fig. 4; Fig. S1; Fig. S2; Fig. S3). Furthermore, the remaining 475 nts at the 3' end of *DcMYB11Pp* did not map to DCAR_010751 but aligned about 38 kb downstream, in a gene annotated as DCAR_010750 (Fig. 4a, b; Fig. S4). PCR and sequence analysis of the genomic region flanking each side (spanning a ~ 1 kb region in each side) of this large insertion confirmed its presence in DH1 (data not shown). Based on these results, we hypothesize that, in DH1, an insertion within the gene body has interrupted the

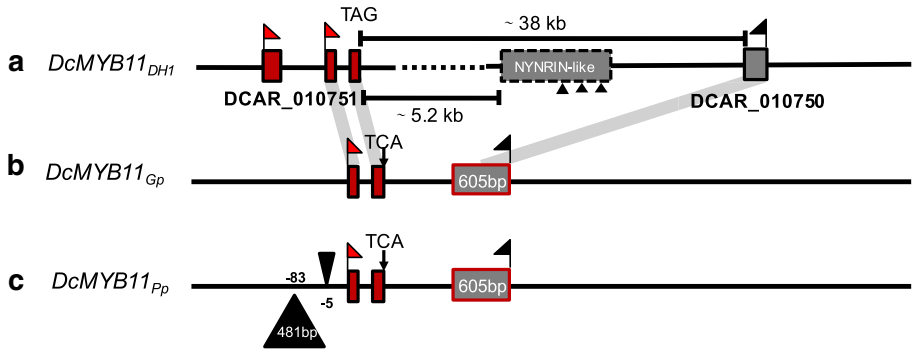

Fig. 4 Structure of *DcMYB11* locus in DH1 compared with that of 5723 individuals that were segregating for green petiole (Gp) and purple petiole (Pp). **a** Structure of *DcMYB11* (DCAR_010751) as predicted on the DH1 genome (DCAR v.2) and a putative NYNRIN-like retrotransposon which trunked the DcMYB11 sequence. **b** Structure of the locus containing the allele of *DcMYB11Gp* isolated from carrots of P5723 with green petiole. **c** Putative structure of *DcMYB11* allele sequenced from carrots of P5723 with purple petiole. Boxes represent exon sequences, separated by thin lines representing introns. Red boxes in **a**, **b,** and **c** indicate exons of DCAR_010752 as predicted on the DH1 genome. Red line around grey box (605 bp) in **b** and **c** indicates reconstructed third exon for *DcMYB11*. Red and black flags indicate start and stop codons, respectively. Black triangles (NYNRIN-like) indicate the presence of three sequences encoding for retrotransposon motifs. Similarity (80–100%) between sequence of **a** and **b** is indicated by light-grey shading

DcMYB11 encoding sequence. To support this hypothesis, alignment of the nucleotide sequence (about 38 kb) between DCAR_010751 and DCAR_010750, found its best match with a 2.4 kb sequence (95.16% homology and e-value 0.0) that is part of *D. carota* NYNRIN-like retrotransposon (NCBI#XM_017372419.1; Fig. 4a). We conclude that DH1 is homozygous for an allele of *DcMYB11* (*DcMYB11*DH1) that encodes a truncated (and likely a non-functional protein) R2R3 MYB, due to the insertion of a retrotransposon (NYNRIN-like) between its second and third exons (annotated in the genome sequence as DCAR_010750; Fig. 4a, b).

Next, we investigated whether Gp plants from the 5723 population hold a DH1-similar allele at *DcMYB11*, thus explaining the lower expression observed in our previous work (Iorizzo et al. 2019) and consequently producing the non-purple (green) petiole phenotype. We compared the sequences spanning the genomic *DcMYB11* region (promoter, coding sequence, and introns) isolated from the 5723 Gp and Pp plants. Analyses allowed the identification of 41 SNPs and 19 indels, with size ranging from 1 to 8 bp and a 481 bp insertion. Of these polymorphisms, 22 SNPs were located within the coding sequence (Fig. S5). Annotation of these SNPs indicated that nine were non-synonymous SNPs and none of them caused disruption (e.g., early stop codon) of the coding sequence (Fig. S5; Fig. S6). Also, the analysis revealed the absence of the retrotransposon identified in DH1. This indicated that 5723 Gp plants harbor a different *DcMYB11* allele as comparted to the truncated version of DH1 (*DcMYB11DH1*) and that the new allele has a full-length coding sequence similar to that characterized in Pp (Fig. 4; Fig. S7). The most striking difference between the sequence of Gp and Pp was a 481 bp insertion (named I481) located in the promoter of Pp plants; it is placed at position -83 bp from the *DcMYB11* start codon (Fig. 4b, c; Fig. S8). As illustrated in Fig. S9, two amplicons differing for the presence or absence of the 481 bp insertion were obtained by amplifying 5723 plants with Gp and Pp. As expected, based on the dominant effect of this gene, heterozygous plants amplified both amplicons and had a purple petiole phenotype. Overall, these results suggest I485 as a possible causal mutation affecting the activity of *DcMYB11* in regulating anthocyanin accumulation in carrot plants with purple petiole. In the following experiments, the impact of the presence/absence of I481 in *DcMYB11* function was tested.

I481 is a novel potential regulatory sequence in the carrot genome

We decided to verify if the presence of the 481 bp insertion, present in the 5' upstream sequence of *DcMYB11Pp* isolated from carrots with purple petiole, could drive a stronger pigmentation (Fig. 5). In these experiments, *DcMYB11Pp* CDS was placed under the action of two native upstream sequences, one of 664 bp that included the I481 insertion (*promoterPp*::*DcMYB11Pp*) and the other of 184 bp, that lacked the I481 insertion (p*Gp::DcMYB11Pp*). Leaves of tobacco infiltrated with Agrobacterium carrying the *pPp*::*DcMYB11Pp* construct showed visual and consistent anthocyanin accumulation about 5 days post-infiltration (Fig. 5a; Fig. S10). In contrast, those infiltrated with *pGp::DcMYB11Pp* showed the first anthocyanin appearance at least 14 days post-infiltration, and in some cases, no anthocyanins were produced (Fig. 5a; Fig. S10). The quantification of total monomeric anthocyanins from leaves at 20 days post-infiltration showed that *pPp*::*DcMYB11Pp* was able to induce a threefold increase in the level of

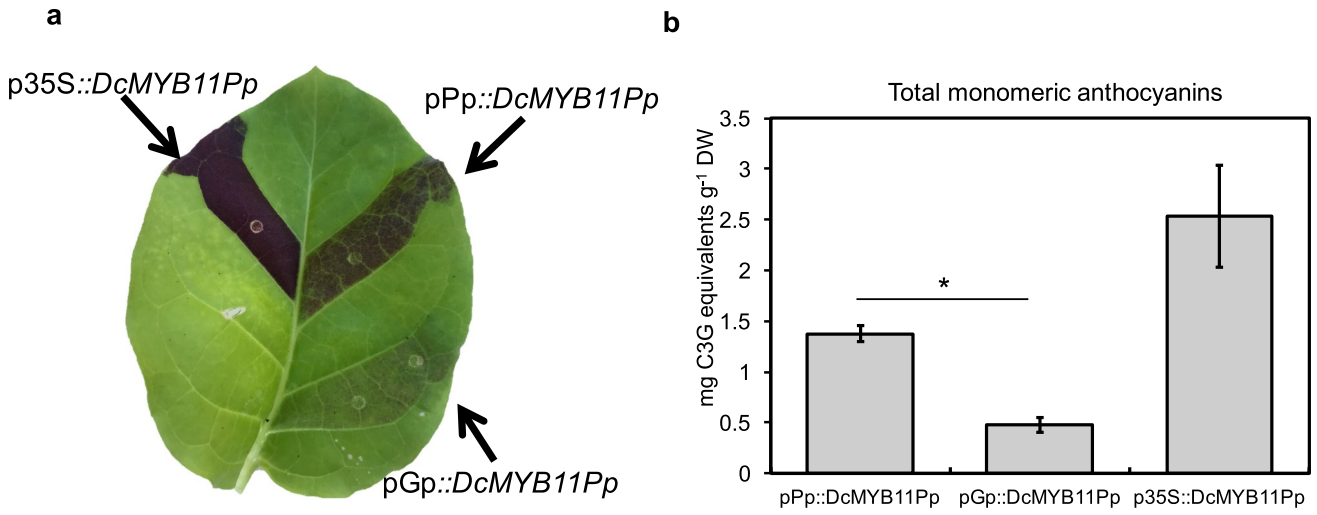

Fig. 5 Functional analysis of I485 sequence in tobacco leaves. **a** Tobacco (*N. tabacum*) leaves after transient expression of p35S::*DcMYB11Pp*, pPp::*DcMYB11Pp* and pGp::*DcMYB11Pp*. Pictures were taken 14 days post-infiltration. **b** Total anthocyanin content in tobacco leaves transformed with pPp::*DcMYB11Pp* and pGp::*DcMYB11Pp*. In **b,** each value represents the mean of three determinations (± SD). *,*P* values < 0.05 according to Student's t test

anthocyanins as compared to *pGp::DcMYB11Pp*, and about 1.6-fold lower pigment concentration than when driven by the constitutive promoter 35S (Fig. 5b). Altogether, these results demonstrated that I481 is a functional cis-regulatory element able to strongly promote *DcMYB11* activity. However, these results pointed out that the Gp promoter might have a basal, though inconstant, activity.

Different hypotheses could explain the origin and function of I481. Some preliminary work was performed here to test them. Our first hypothesis was that I481 might have introduced transcription factor-binding sites that increased *DcMYB11* activity. Indeed, several common TATA and CAAT-box motif were found in I481 along with plant-specific ARE (5´-AAACCA-3´), TCA-element (5´-CCATCT TTTT-3´), and re2f-1 (5´-GCGGGAAA- 3´) motifs associated with anaerobic induction and salicylic acid responsiveness and activation of cell division, respectively. For example, in apple, enhanced auto-regulation and, consequently, increased anthocyanin accumulation, was explained by the presence of an insertion in the promoter of an anthocyanin R2R3 MYB (*MdMYB10*) (Espley et al. 2009; Brendolise et al. 2017). To test if a similar mechanism occurs in *DcMYB11*, a yeast one-hybrid experiment was carried out. Results indicated that the *DcMYB11* promoter was not activated by *DcMYB11,* or its paralog *DcMYB7*, in neither the presence nor absence of I481 (Fig. 6). These results excluded our first hypothesis. Another possible hypothesis is that I481 acts as an enhancer. This would imply that the sequence would have some features that are typical of enhancers, such as containing sequence motifs that are present in the other parts of the genome (Zhu et al. 2015; Schimitz et al. 2022). To gain some preliminary insights into this hypothesis, we conducted a homology research to look at the presence of I481 in other regions of the carrot genome. Using stringent blast parameters (*e*-value = 1 × 10–10), we obtained 71 different matches distributed on all carrot chromosomes (Fig. S11). About 91% of the total matches had high homology with the first 200 bp at the 5' end of I481 (identity values ranging between 77 and 93%), suggesting that this region is the most conserved within the I481 insertion, which contains basic TATA and Cat-box (Table S2). I481 sequence is widespread in the genome, being located in different positions with respect to gene sequences (upstream or downstream and in a distance ranging between 140 bp and 11.7 kb); in two cases, it was within the gene body (Fig. S11; Table S2). Annotation of all the nearest genes indicated that they were involved in different functions (e.g., resistance genes, primary metabolism, and basic cellular and molecular function), thus suggesting a general role of I481 as a cis-activator of transcription. All the above features could be attributed to an enhancer (Weber et al. 2016; Schimitz et al. 2022). Future works in homologous systems are needed to validate this hypothesis and also to understand how and if this sequence could impact its tissue-specific activity.

Conclusions

Overall, the work presented here demonstrated for the first time that *DcMYB11*, a gene previously associated with purple petiole pigmentation in carrot, is able to activate the anthocyanin pathway in heterologous systems. These findings expand our understanding about the function of a cluster of six R2R3 MYB activators that co-localize with

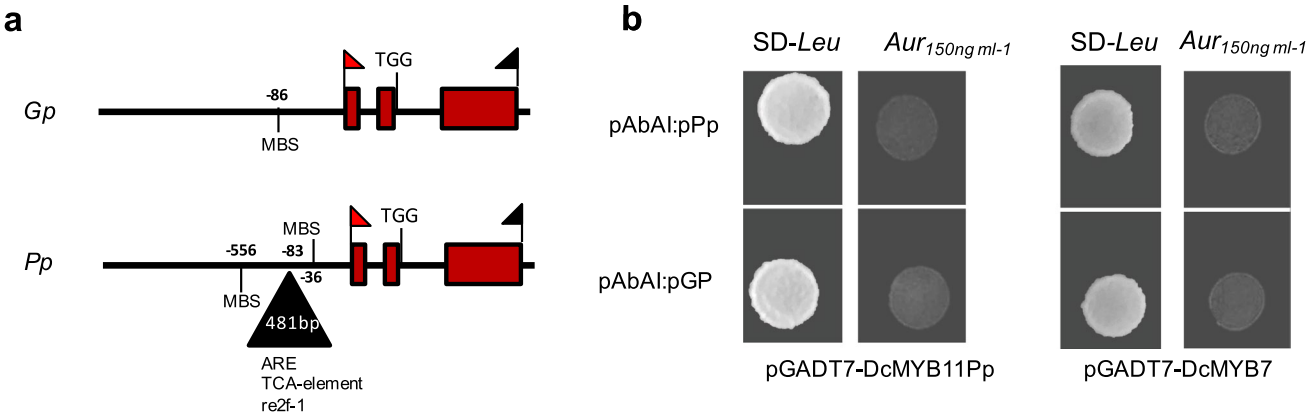

Fig. 6 Evaluation of binding ability of DcMYB11Pp and DcMYB7 with 5' upstream sequence (baits) isolate from pruple (pPp) and green (pGp) petiole. **a** The insertion of I481 moved a putative MYB-binding site (MBS) from -86 to – 556 nts from the start codon of DcMYB11. I481 introduced novel cis-regulatory elements in the upstream region of *DcMYB11Pp*, besides general TATA and CAAT boxes. Boxes represent exons, and the thin line represents introns and intergenic regions. Red and black flags represent the start and stop codons, respectively. The triangles indicate insertions. **b** The plasmids pGADT7-*DcMYB11Pp* and pGADT7-*DcMYB7* were transferred into the bait-reporter yeast strains Y1Hgold [pBait-AbAi:pPp ~664 bp] and Y1Hgold [pBait-AbAi:pGP ~187 bp], respectively, and then selected on SD/ − Leu/ + AbA150 agar plates. The transformants from both the combinations Y1Hgold [pBait-AbAi/pGADT7-*DcMYB11Pp* or /*DcMYB7*] could not grow on the SD/ − Leu/ + AbA150 indicating that there was not activation of the resistance gene (*AbAr*)

DcMYB11 in the carrot genome. Comparative analysis between plants with purple and non-purple pigmentation identified three possible alleles. A 481 bp insertion in the promoter of *DcMYB11* in purple petiole plants enhanced its activity and may be responsible for the purple phenotype. Preliminary analysis to characterize this sequence suggests that I481 is a putative enhancer. Enhancers are still quite unknown in plants and remain largely uncharacterized (Jores et al. 2020; Jain and Garg 2021). Future work in homologous system, characterizing the function of this sequence, can contribute to expanding knowledge about cis-regulatory sequences in plants and understanding if I481 can also influence tissue-specific expression patterns of anthocyanins.

Author contribution statement VDA, JC, and MI planned and designed the research. VDA, MAA and JC performed experiments. VDA, JC, PC and DC analyzed data. MI contributed reagents/materials/analysis tools. VDA and MI wrote the manuscript with the input from all the authors. All authors read and approved the manuscript.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00425-023-04200-5.

Acknowledgements The authors would like to thank Changqing Zhang from N.C. State University Plants for Human Health Institute for providing us the plasmid vectors (pF35SE and pFAGM) used in this study.

Funding This work was supported by the United States Department of Agriculture National Institute of Food and Agriculture, Hatch Project 1008691 and Award No. 2022–67013-36389. VDA was supported by ExPoPe-STAR-Line 1, 2017-Unina-Intesa San Paolo and Short Mobility of program of University of Naples Federico II.

Data availability Detailed data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Competing interests The authors have no competing interest to declare.

References

Agati G, Guidi L, Landi M, Tattini M (2021) Anthocyanins in photoprotection: knowing the actors in play to solve this complex ecophysiological issue. New Phytol 232:2228–2235. https://doi.org/10.1111/nph.17648

Albert NW, Lewis DH, Zhang H, Schwinn KE, Jameson PE, Davies KM (2011) Members of an R2R3-MYB transcription factor family in *Petunia* are developmentally and environmentally regulated to control complex floral and vegetative pigmentation patterning. Plant J 65:771–784. https://doi.org/10.1111/j.1365-313X.2010.04465.x

Albert NW, Davies KM, Lewis DH, Zhang H, Montefiori M, Brendolise C, Boase MR, Ngo H, Jameson PE, Schwinn KE (2014) A conserved network of transcriptional activators and repressors regulates anthocyanin pigmentation in eudicots. Plant Cell 26:962–980. https://doi.org/10.1105/tpc.113.122069

Bai C, Elledge SJ (1997) Gene identification using the yeast two-hybrid system. Meth Enzymol 283:141–156. https://doi.org/10.1016/s0076-6879(97)83013-3

Bannoud F, Ellison S, Paolinelli M, Horejsi T, Senalik D, Fanzone M, Iorizzo M, Simon PW, Cavagnaro PF (2019) Dissecting the genetic control of root and leaf tissue-specific anthocyanin

pigmentation in carrot (*Daucus carota* L.). Theor Appl Genet 132:2485–2507. https://doi.org/10.1007/s00122-019-03366-5

Baudry A, Heim MA, Dubreucq B, Caboche M, Weisshaar B, Lepiniec L (2004) TT2, TT8, and TTG1 synergistically specify the expression of *BANYULS* and proanthocyanidin biosynthesis in *Arabidopsis thaliana*. Plant J 39:366–380. https://doi.org/10.1111/j.1365-313X.2004.02138.x

Brendolise C, Espley RV, Lin-Wang K, Laing W, Peng Y, McGhie T, Dejnoprat S, Tomes S, Hellens RP, Allan AC (2017) Multiple copies of a simple MYB-binding site confers trans-regulation by specific flavonoid-related R2R3 MYBs in diverse species. Front Plant Sci 8:1864. https://doi.org/10.1111/10.3389/fpls.2017.01864

Butelli E, Licciardello C, Zhang Y, Liu J, Mackay S, Bailey P, Reforgiato-Recupero G, Martin C (2012) Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell 24:1242–1255. https://doi.org/10.1105/tpc.111.095232

Butelli E, Garcia-Lor A, Licciardello C, Las Casas G, Hill L, Recupero GR, Keremane ML, Ramadugu C, Krueger R, Xu Q, Deng X (2017) Changes in anthocyanin production during domestication of *Citrus*. Plant Physiol 173:2225–2242. https://doi.org/10.1104/pp.16.01701

Cavagnaro PF, Iorizzo M, Yildiz M, Senalik D, Parsons J, Ellison S, Simon P (2014) A gene-derived SNP-based high resolution linkage map of carrot including the location of QTL conditioning root and leaf anthocyanin pigmentation. BMC Genom 15:1118. https://doi.org/10.17660/10.1186/1471-2164-15-1118

Cavagnaro PF, Bannoud F, Senalik D, Ellison SL, Simon PW (2019) Carrot anthocyanins: nutrition, diversity and genetics. Acta Hortic 2019:101–106. https://doi.org/10.17660/ActaHortic.2019.1264.11

Chagné D, Lin-Wang K, Espley RV et al (2013) An ancient duplication of apple MYB transcription factors is responsible for novel red fruit-flesh phenotypes. Plant Physiol 161:225–239. https://doi.org/10.1104/pp.112.206771

Colanero S, Perata P, Gonzali S (2020a) What's behind purple tomatoes? Insight into the mechanisms of anthocyanin synthesis in tomato fruits. Plant Physiol 182:1841–1853. https://doi.org/10.1104/pp.19.01530

Colanero S, Tagliani A, Perata P, Gonzali S (2020b) Alternative splicing in the anthocyanin fruit gene encoding an R2R3 MYB transcription factor affects anthocyanin biosynthesis in tomato fruits. Plant Commun 1:100006. https://doi.org/10.1016/j.xplc.2019.100006

D'Amelia V, Villano C, Batelli G, Çobanoğlu Ö, 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

D'Amelia V, Aversano R, Batelli G, Caruso I, Castellano Moreno M, Castro-Sanz AB, Chiaiese P, Fasano C, Palomba F, Carputo D (2014) High AN1 variability and interaction with basic helix-loop-helix co-factors related to anthocyanin biosynthesis in potato leaves. Plant J 80:527–540. https://doi.org/10.1111/tpj.12653

D'Amelia V, Aversano R, Ruggiero A, Batelli G, Appelhagen I, Dinacci C, Hill L, Martin C, Carputo D (2018) Subfunctionalization of duplicate MYB genes in *Solanum commersonii* generated the cold-induced *ScAN2* and the anthocyanin regulator *ScAN1*. Plant Cell Environ 41:1038–1051. https://doi.org/10.1111/pce.12966

Davies KM, Albert NW, Zhou Y, Schwinn KE (2018) Functions of flavonoid and betalain pigments in abiotic stress tolerance in plants. In: Roberts JA (ed) Annual plant reviews online, https://doi.org/10.1002/9781119312994.apr0604

Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L (2010) MYB transcription factors in Arabidopsis. Trends Plant Sci 15:573–581

Ellison SL, Luby CH, Corak KE, Coe KM, Senalik D, Iorizzo M, Goldman IL, Simon PW, Dawson JC (2018) Carotenoid presence is associated with the *Or* gene in domesticated carrot. Genetics 210:1497–1508. https://doi.org/10.1534/genetics.118.301299

Gel B, Serra E (2017) karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33:3088–3090. https://doi.org/10.1093/bioinformatics/btx346

Giusti MM, Wrolstad RE (2001) Characterization and measurement of anthocyanins by UV-visible spectroscopy. Curr Protoc Food Anal Chem 1:F1-2. https://doi.org/10.1002/0471142913.faf0102s00

Gonzalez A, Zhao M, Leavitt JM, Lloyd AM (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. Plant J 53:814–827. https://doi.org/10.1111/j.1365-313X.2007.03373.x

He J, Ye S, Correia P, Fernandes I, Zhang R, Wu M, Freitas V, Mateus N, Oliveira H (2022) Dietary polyglycosylated anthocyanins, the smart option? A comprehensive review on their health benefits and technological applications. Compr Rev Food Sci Food Saf 21:3096–3128. https://doi.org/10.1111/1541-4337.12970

Heim MA, Jakoby M, Werber M, Martin C, Weisshaar B, Bailey PC (2003) The basic helix–loop–helix transcription factor family in plants: a genome-wide study of protein structure and functional diversity. Mol Biol Evol 20:735–747. https://doi.org/10.1093/molbev/msg088

Huang D, Wang X, Tang Z, Yuan Y, Xu Y, He J, Jiang X, Peng SA, Li L, Butelli E, Deng X (2018) Subfunctionalization of the *Ruby2*–Ruby1 gene cluster during the domestication of citrus. Nature Plants 4:930–941. https://doi.org/10.1038/s41477-018-0287-6

Iorizzo M, Ellison S, Senalik D et al (2016) A high-quality carrot genome assembly provides new insights into carotenoid accumulation and asterid genome evolution. Nat Genet 48:657–666. https://doi.org/10.1038/ng.3565

Iorizzo M, Cavagnaro PF, Bostan H, Zhao Y, Zhang J, Simon PWA (2019) Cluster of MYB transcription factors regulates anthocyanin biosynthesis in carrot (Daucus carota L.) root and petiole. Front Plant Sci 9:1927. https://doi.org/10.3389/fpls.2018.01927

Iorizzo M, Curaba J, Pottorff M, Ferruzzi MG, Simon P, Cavagnaro PF (2020) Carrot anthocyanins genetics and genomics: Status and perspectives to improve its application for the food colorant industry. Genes 11:906. https://doi.org/10.3390/genes11080906

Jung CS, Griffiths HM, De Jong DM, Cheng S, Bodis M, Kim TS, De Jong WS (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 120:45–57

Kobayashi S, Ishimaru M, Hiraoka K, Honda C (2002) Myb-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis. Planta 215:924–933. https://doi.org/10.1007/s00425-002-0830-5

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

Koressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M (2018) Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics 34:1937–1938. https://doi.org/10.1093/bioinformatics/bty036

Laimbeer FPE, Bargmann BO, Holt SH, Pratt T, Peterson B, Doulis AG, Buell CR, Veilleux RE (2020) Characterization of the f locus responsible for floral anthocyanin production in potato. G3. https://doi.org/10.1534/g3.120.401684

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

Lin-Wang K, Bolitho K, Grafton K, Kortstee A, Karunairetnam S, McGhie TK, Espley RV, Hellens RP, Allan AC (2010) An R2R3 MYB transcription factor associated with regulation of the anthocyanin biosynthetic pathway in Rosaceae. BMC Plant Biol 10:50. https://doi.org/10.1186/1471-2229-10-50

Montefiori M, Brendolise C, Dare AP, Lin-Wang K, Davies KM, Hellens RP, Allan A (2015) In the Solanaceae, a hierarchy of bHLHs confer distinct target specificity to the anthocyanin regulatory complex. J Exp Bot 66:427–1436. https://doi.org/10.1093/jxb/eru494

Passeri V, Koes R, Quattrocchio FM (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

Quattrocchio F, Wing J, van der Woude K, Souer E, de Vetten N, Mol J, Koes R (1999) Molecular analysis of the *anthocyanin2* gene of petunia and its role in the evolution of flower color. Plant Cell 11:1433–1444. https://doi.org/10.1105/tpc.11.8.1433

Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 6:841–842. https://doi.org/10.1093/bioinformatics/btq033

Ramsay NA, Glover BJ (2005) MYB–bHLH–WD40 protein complex and the evolution of cellular diversity. Trends Plant Sci 10:63–70. https://doi.org/10.1016/j.tplants.2004.12.011

Schwinn K, Venail J, Shang Y, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C (2006) A small family of *MYB*-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. Plant Cell 18:831–851. https://doi.org/10.1105/tpc.105.039255

Spelt C, Quattrocchio F, Mol JN, Koes R (2000) *anthocyanin1* of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell 12:1619–1631. https://doi.org/10.1105/tpc.12.9.1619

Stracke R, Werber M, Weisshaar B (2001) The *R2R3-MYB* gene family in *Arabidopsis thaliana*. Curr Opin Plant Biol 4:447–456. https://doi.org/10.1016/s1369-5266(00)00199-0

Weber B, Zicola J, Oka R, Stam M (2016) Plant enhancers: a call for discovery. Trends Plant Sci 21:974–987. https://doi.org/10.1016/j.tplants.2016.07.013

Xu ZS, Feng K, Que F, Wang F, Xiong AS (2017) A MYB transcription factor, DcMYB6, is involved in regulating anthocyanin biosynthesis in purple carrot taproots. Sci Rep 7:1–9. https://doi.org/10.1038/srep45324

Xu ZS, Yang QQ, Feng K, Xiong AS (2019) Changing carrot color: insertions in *DcMYB7* alter the regulation of anthocyanin biosynthesis and modification. Plant Physiol 181:195–207. https://doi.org/10.1104/pp.19.00523

Zhang P, Chopra S, Peterson T (2000) A segmental gene duplication generated differentially expressed *myb*-homologous genes in maize. Plant Cell 12:2311–3232. https://doi.org/10.1105/tpc.12.12.2311

Zhang H, Koes R, Shang H, Fu Z, Wang L, Dong X, Zhang J, Passeri V, Li Y, Jiang H, Gao J (2019) Identification and functional analysis of three new anthocyanin R2R3-MYB genes in *Petunia*. Plant Direct 3:1–13. https://doi.org/10.1002/pld3.114

Zhu B, Zhang W, Zhang T, Liu B, Jiang J (2015) Genome-wide prediction and validation of intergenic enhancers in Arabidopsis using open chromatin signatures. Plant Cell 27:2415–2426

Zimmermann IM, Heim MA, Weisshaar B, Uhrig JF (2004) Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. Plant J 40:22–34