



OPEN ACCESS

EDITED BY
Elsayed Nishawy,
Desert Research Center, Egypt

REVIEWED BY
Kaifeng Zheng,
Beijing Normal University, China
Baoying Shen,
Fujian Agriculture and Forestry University,
China

*CORRESPONDENCE
Cinzia Comino
✉ cinzia.comino@unito.it

RECEIVED 16 March 2026
REVISED 07 May 2026
ACCEPTED 09 May 2026
PUBLISHED 28 May 2026

CITATION

De Marchi F, Acquadro A, Moglia A, Gaccione L, Mannino G, Gramazio P, Barchi L, Berteà CM, Portis E, Prohens J and Comino C (2026) Transcriptomic landscapes of tissue-specific color transition in eggplant reveal regulatory roles of lncRNAs and alternative splicing in anthocyanin biosynthesis. *Front. Plant Sci.* 17:1832029. doi: 10.3389/fpls.2026.1832029

COPYRIGHT

© 2026 De Marchi, Acquadro, Moglia, Gaccione, Mannino, Gramazio, Barchi, Berteà, Portis, Prohens and Comino. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Transcriptomic landscapes of tissue-specific color transition in eggplant reveal regulatory roles of lncRNAs and alternative splicing in anthocyanin biosynthesis

Federica De Marchi¹, Alberto Acquadro¹, Andrea Moglia¹, Luciana Gaccione¹, Giuseppe Mannino², Pietro Gramazio³, Lorenzo Barchi¹, Cinzia Margherita Berteà², Ezio Portis¹, Jaime Prohens³ and Cinzia Comino^{1*}

¹Department of Agricultural, Forest and Food Science, University of Turin, Turin, Italy, ²Department of Life Sciences and Systems Biology, Innovation Centre, University of Turin, Turin, Italy, ³Instituto de Conservación y Mejora de la Agrodiversidad Valenciana, Universitat Politècnica de València, Valencia, Spain

Background: Color transitions from green to purple in eggplant (*Solanum melongena* L.) flowers and berries result from chlorophyll and anthocyanin accumulation. While structural genes (*CHS*, *CHI*, *F3'H*, *DFR*, *ANS*, *UGT*) and regulatory transcription factors (*MYB*, *bHLH*, *WD40*) in the anthocyanin pathway are well characterized, fine-scale transcriptional regulation is less understood. Long non-coding RNAs (lncRNAs) contribute to pre- and post-transcriptional regulation, whereas alternative splicing (AS) events control exclusively post-transcriptional modifications.

Methods: Here we performed an integrated transcriptomic analysis of mRNAs, lncRNAs and AS events to investigate fine regulation of anthocyanin biosynthesis.

Results: A PacBio assembly of the eggplant cultivar “Black Beauty” genome allowed the computational prediction and expression analysis of all transcripts in fruit peel and flower corolla. We identified 4, 126 lncRNAs and 34, 745 novel transcript isoforms, and tissue-specific regulatory networks revealed multiple lncRNAs and protein-coding genes AS events associated with anthocyanin biosynthesis. Quantification of anthocyanins in both peel and corolla enabled the construction of putative tissue-specific models of anthocyanin transcriptional regulation.

Discussion: This study provides a starting point to improve the knowledge of fine-tuning anthocyanin regulation at the transcriptional level, highlighting tissue-specific lncRNAs and AS events linked to regulatory genes, and offering, to our knowledge, the first evidence of lncRNA involvement in anthocyanin biosynthesis in eggplant.

KEYWORDS

delphinidin-3-rutinoside, MYB, PacBio HiFi, RNA-seq, *Solanum melongena*, Spearman correlation

1 Introduction

Anthocyanins are water-soluble secondary metabolites responsible for the pigmentation of different plants' tissues, including fruits and flowers (Tanaka et al., 2008). Among their functional roles, anthocyanins are essential in protecting plants against abiotic and biotic stressors (An et al., 2020), in the attraction of pollinators (Cappellini et al., 2021) and in modulating the cellular redox balance.

The biosynthetic pathway of anthocyanins (Supplementary Figure 1) is a branch of the phenylpropanoid biosynthetic grid and entails a cascade of enzymatic reactions starting from the conversion of phenylalanine into coumaroyl-CoA, catalyzed by the enzymes phenylalanine ammonia-lyase (PAL), cinnamate 4-hydroxylase (C4H) and 4-coumarate-CoA ligase (4CL) (Sunil and Shetty, 2022). The cascade continues toward the flavonoid structural genes, classified into early and late biosynthetic genes (EBGs and LBGs, respectively), leading to stable anthocyanins (Shoeva et al., 2017).

The EBGs lead to the synthesis of dihydroflavonol precursors and other flavonolic compounds, and include *chalcone synthase* (CHS), *chalcone isomerase* (CHI) and *flavanone 3-hydroxylase* (F3H). Specifically, CHS catalyzes the synthesis of naringenin chalcone from 4-coumaroyl-CoA and malonyl-CoA, which is then converted by CHI into naringenin. Subsequently, F3H hydroxylates naringenin to form dihydrokaempferol, the central substrate for downstream diversification. The genes responsible for the ultimate modifications in anthocyanin production are included in the LBGs, comprising *flavonoid 3'-hydroxylase* (F3'H), *flavonoid 3', 5'-hydroxylase* (F3'5'H), which transform dihydrokaempferol into two other dihydroflavonols, dihydroquercetin or dihydromyricetin, respectively, thus enabling structural diversification of anthocyanins (Manickam et al., 2025). *Dihydroflavonol 4-reductase* (DFR) then catalyzes the conversion of these dihydroflavonols into colorless leucoanthocyanidins, which are subsequently transformed by *anthocyanidin synthase* (ANS) into the pigmented anthocyanidins. Finally, *flavonoid 3-O-glucosyltransferase* (UGT, also known as GT) stabilizes these molecules by the conjugation of sugar residues, yielding stable anthocyanins, which may undergo additional modifications such as acylation (Lo Scalzo et al., 2021; Manickam et al., 2025).

In dicotyledons, the transcriptional regulation of the anthocyanin pathway is mostly mediated by transcription factors (TFs) forming the MYB-bHLH-WD (MBW) ternary complex, consisting of R2-R3-MYB, basic helix-loop-helix (bHLH) and WD40-repeat

proteins (Jaakola, 2013). Besides functioning within the MBW complex, tissue-specific MYB TFs can also positively regulate anthocyanin biosynthesis and accumulation alone. In addition, these TFs could also influence anthocyanin production indirectly through hormone signaling. Particularly, the increase in ethylene hormone and its signaling pathway mediated by *Ethylene Response Factors* (ERFs) has been linked to anthocyanin accumulation in several plant species (Yu et al., 2022). In contrast, auxin biosynthesis was found to be antithetical to that of anthocyanins (Li et al., 2024a).

Beyond transcriptional and hormone-mediated regulation, recent advances in high-throughput sequencing technologies and computational analytical approaches have also highlighted the importance of alternative splicing (AS) and long non-coding RNAs (lncRNAs) as key regulators of gene expression. The AS events are generated from the selection of a different splice site by the spliceosome to produce multiple mature mRNAs. AS plays a key role in post-transcriptional regulation, influencing several biological processes in plants, such as both biotic and abiotic stress responses, tissues' development and morphogenesis (Yan et al., 2021; Mandadi et al., 2023). It has been estimated that up to 60% of intron-containing genes in plants undergo alternative splicing (AS) events, a proportion that has risen to nearly 70% in more recent transcriptomic analyses (Szakonyi and Duque, 2018). These events expand transcript diversity and can therefore alter protein function or create premature termination codons that trigger nonsense-mediated decay, leading to mRNA degradation (Syed et al., 2012; Sibley et al., 2016). In recent studies, some protein-coding genes AS events have been reported as anthocyanin-related among the *Solanaceae* family. In tomato, the alternative 5' splicing sites of *SLAN2-like* R2-R3 MYB transcription factor lead to the partial loss of the R3 domain that causes a no-anthocyanin phenotype (Colanero et al., 2020). Non-coding RNAs with a length greater than 200 nucleotides are identified as lncRNAs (Heo et al., 2013). These molecules can recruit factors for epigenetic and chromatin modification and can act as pre- and post-transcriptional regulators of target genes, being characterized by a lower expression and conservation between different organisms compared to mRNAs (Wang et al., 2023). The lncRNAs mode of action includes cis- and trans-acting mechanisms, depending on the location of their target genes. lncRNAs regulate several biological processes in plants, including flowering time, crop yield, responses to abiotic and biotic stresses, reproduction, and organogenesis, by interacting with DNA, RNA, and proteins (Li et al., 2023a; Mattick et al., 2023; Li et al., 2024b). In recent years, some works have focused on anthocyanin accumulation and biosynthesis in several plant species. In flowers of *Melastoma candidum*, different lncRNAs have been identified as regulators of *4CL*, *CHS*, *CHI*, *F3H*, and *MYB4* genes (Li et al., 2023b), and in apple fruits, the lncRNA *MdLNC610* has been positively correlated with *MdACO1* transcript, enhancing light-induced anthocyanin accumulation by activating ethylene synthesis (Yu et al., 2022). Similarly, in carrot roots, the lncRNAs *asDcMYB6* and *asDcMYB7* were found to have an expression pattern consistent with the two MYB TFs from which they originate, possibly leading to a positive role in the anthocyanin pathway regulation (Chialva et al., 2021).

Abbreviations: A3, Alternative 3' Splice Sites; A5, Alternative 5' Splice Sites; AF, Alternative First Exon; AL, Alternative Last Exon; AS, Alternative Splicing; BB, Black Beauty; DEA, Differentially Expression Analysis; DEGs, Differentially Expressed Genes; DELncRNA, Differentially Expressed lncRNA; DSE, Differentially Spliced event; HPLC-ESI-DAD-MS/MS, High-Performance Liquid Chromatography-Diode Array Detector-Electrospray Ionization-Tandem Mass Spectrometry; lncRNA, long non-coding RNA; log2FC, Log2 fold change; MX, Mutually Exclusive Exons; PCA, Principal Component Analysis; RI, Retained Intron; SE, Skipping Exon; SPL-DA, Sparse Partial Least Squares-Discriminant Analysis; TFs, Transcription Factors.

Eggplant (*Solanum melongena* L.) is the third most cultivated crop in the *Solanaceae* family, after potato and tomato (Martina et al., 2024; FAO, 2025). Unlike tomato and potato, eggplant exhibits several anthocyanin accumulation patterns in its fruit peel and flower petals, ranging from white to purple. Specifically, in eggplant varieties with violet pigmentation, anthocyanin accumulation is restricted to derivatives of delphinidin, which are also detectable in vegetative tissues (Filyushin et al., 2024). In fruit, the abundance of delphinidin is maximal at early developmental stages and progressively declines during fruit maturation, ultimately becoming undetectable at full physiologically ripeness (Panda et al., 2025). The predominant anthocyanins identified are delphinidin-3-(*p*-coumaroyl-rutinoside)-5-glucoside (commonly referred to as nasunin) and delphinidin-3-rutinoside (commonly referred to as tulipanin) (Gaccione et al., 2023). Purple eggplant extracts have been reported to show a high superoxide anion radical scavenging activity, underlying the importance of anthocyanins' role not only in the regulation of plant physiological processes, but also in human health and nutrition (Noda et al., 2000; Putta et al., 2017; Mattioli et al., 2020). The structural genes and regulatory transcription factors (TFs) of the anthocyanin pathway have been deeply characterized in eggplant, where two major QTLs have been identified on chromosomes 5 and 10 (Barchi et al., 2012; Toppino et al., 2020). In the fruit peel of *S. melongena* accessions, the *MYB1* gene (also known as *MYB113*) has been characterized as the activator of anthocyanin biosynthesis, acting solo or interacting with the bHLH transcription factor *Transparent Testa 8 (TT8)* in the MBW complex; whereas, in flowers, the MYBs involved in anthocyanin production are reported to be *MYB2* (or *AN2*) and *MYB75* (Shi et al., 2021; Filyushin et al., 2023; Ai et al., 2025). In fruits, *MYBL1*, a R3 type MYB, has also been linked to anthocyanin biosynthesis as a negative regulator in *S. melongena* (Moglia et al., 2020). MYB TFs expression could be regulated by miRNAs, such as miRNA156/157, which are involved in the regulation of anthocyanin biosynthesis in eggplant by directly interacting with MYBs or indirectly by promoting *SQUAMOSA PROMOTER BINDING-LIKE (SPL)* genes (Panda et al., 2025).

Despite the extensive characterization of the structural genes and TFs involved in the production of anthocyanins, the fine-tuning at the transcriptomic level, mediated by alternative splicing (AS) events, and especially long non-coding RNAs (lncRNAs), remains largely unexplored in eggplant. An alternative 5' splicing site of the *DFR* gene in two anthocyanin-free *S. melongena* mutants has been associated with the variation in anthocyanin accumulation observed during the domestication of the spiny *Solanum* group (Wang et al., 2022a).

To our knowledge, no studies have reported the involvement of lncRNAs in the biosynthesis of anthocyanins in eggplant, and protein-coding genes AS has been only partially explored.

Here, we present a comprehensive transcriptomic analysis, focusing on mRNA, lncRNA and AS events to elucidate the fine regulation of the anthocyanin biosynthetic pathway in eggplant fruits and flowers. To this purpose, we performed a PacBio assembly of the "Black Beauty" (BB) cultivar genome, which has purple berries and flowers, and created regulatory networks to explore the

role of lncRNA and AS in the biosynthesis and accumulation of anthocyanins, associating all collected data with metabolic analysis.

2 Materials and methods

2.1 Plant material

Plants of the eggplant cultivar "Black Beauty" (BB) were grown in a greenhouse at a mean temperature of 25 °C for four months. The peel of three berries from three plants was collected and pooled at three developmental stages, yielding a total of nine fruit samples: fruit set surrounded by sepals to ensure no light exposure, displaying green coloration (ST1); enlarged fruits exhibiting color shifting from green to purple (ST2); large commercially mature fruits with fully purple peel (ST3). Similarly, five flower corollas were collected from each sampled plant and were pooled at two stages of development: Stage green (STG), corresponding to close buds with green petals, and Stage purple (STP), corresponding to open buds with purple petals, resulting in a total of six flower samples.

2.2 Extraction and analysis of anthocyanin compounds

Anthocyanins were extracted from plant material using an acidified hydroalcoholic solvent. Briefly, samples were homogenized and extracted with 50% (v/v) ethanol acidified with 1% (v/v) formic acid, using a solid-to-solvent ratio of 1:10 (w/v). The extraction was carried out under constant agitation at room temperature, with samples protected from light. Subsequently, the extracts were clarified by centrifugation, and the supernatants were collected for further analysis (Mannino and Maffei, 2022; Mannino et al., 2022). The analysis of anthocyanins was performed using an Agilent Technologies 1200 HPLC system coupled with a DAD and an LC-MS 6330 Series Ion Trap (Agilent Technologies, USA) equipped with an electrospray ionization (ESI) source. Chromatographic separation was performed at a constant flow rate of 0.2 ml min⁻¹ on a reverse-phase C18 Luna column (3.00 µm, 150 × 3.0 mm i.d., Phenomenex, USA), maintained at 25 °C using the Agilent 1100 HPLC G1316A column compartment. The solvents employed for the analysis were: (A) water acidified with 0.1% (v/v) formic acid and (B) 50% (v/v) methanol acidified with 0.1% (v/v) formic acid. The linear gradient adopted was: 0–5 min 5% (v/v) B, 5–25 min 5–25% (v/v) B, 25–35 min 25–40% (v/v) B, 35–40 min 40–5% (v/v) B, followed by column re-equilibration. UV–VIS spectra were recorded between 220 and 650 nm, while chromatographic profiles were acquired at 220, 280, 360, and 520 nm, the latter specifically for the detection of anthocyanins. Tandem mass spectrometry analyses were performed in positive mode for anthocyanins, using nitrogen as the nebulization gas (flow 5.0 ml min⁻¹, temperature 325 °C) and capillary voltage of 1.5 kV; helium was used as the collision gas. The compounds were identified by comparing the retention times and UV–VIS/MS spectra with those of authentic reference compounds or based on data in the literature. The data obtained from HPLC-ESI-DAD-MS/MS analyses

were processed and normalized using the MetaboAnalyst platform vs. 24 (www.metaboanalyst.ca). Since no missing values were detected, the Low-quality filter was not applied. However, the low-variance filter based on the interquartile range (IQR) with a threshold of 5% was used to exclude features that were almost constant between experimental conditions. In addition, a low-abundance filter based on the average intensity value of the features was applied. To compensate for systematic differences between samples, the data were normalized by median and transformed using $\log_2$, a mathematical transformation that also allows negative values to be handled using simple log and square root approaches. Subsequently, each variable was centered on the mean and divided by the square root of the standard deviation (Pareto scaling) to balance the influence of features with different dispersion.

2.3 Nucleic acids extraction, quality control, sequencing and qPCR

High-molecular-weight nuclear DNA was extracted from the previously ground and frozen peel of ST2 fruits using a modified CTAB method (Li et al., 2024c). Briefly, 800 μ L of extraction buffer (2% (w/v) CTAB, 0.1 M Tris-HCl and 20 mM EDTA disodium salt, 1.4 M NaCl) supplemented with PVP40 1% (w/v) and 0.2% (v/v) β -Mercaptoethanol right before use, were added to 100 mg of frozen tissue powder in a 2 mL Eppendorf tube. The mixture was gently mixed by inverting the tube and 3 μ L of RNase A solution (4 mg/mL, Promega) were added. The tube was incubated at 37 °C for 12 minutes. Then, 10 μ L of Proteinase K (20 mg/mL, Promega) were added, the tube was inverted a couple of times and incubated at 56 °C for 15 minutes. The sample was kept in ice for 5 minutes, and then 600 μ L of cold Phenol-chloroform-isoamyl alcohol (25:24:1) (v/v/v) was added. The tube was kept in agitation at room temperature for 10 minutes and centrifuged at 11,500 rpm for 10 minutes. The supernatant was collected using cut-end tips and placed in a new 2 mL Eppendorf tube, where 42 μ L of $\text{NH}_4\text{OAC}-\text{CH}_3\text{COONH}_4$ 5M and, subsequently, 800 μ L of pure isopropanol were added. The tube was gently inverted and centrifuged at 13,000 rpm for 10 minutes at 4 °C to guarantee the DNA pellet formation. The supernatant was discarded and the pellet washed twice in ethanol 70% (v/v) followed by a 13,000 rpm centrifugation for one minute at 4 °C. After the ethanol residue evaporated, 100 μ L of ultrapure water was added and the pellet was resuspended at room temperature overnight and then conserved at -20 °C. The total RNA of the six BB corolla and the nine peel samples was extracted using a commercial kit (SpectrumTM Plant Total RNA Kit, Sigma-Aldrich, Milan, Italy), according to the manufacturer's instructions. DNA and RNA concentration, integrity and quality were assessed by Qubit 2.0 (Life Technologies, Carlsbad, CA, USA), using Qubit dsDNA BR Assay and Qubit dsRNA BR Assay (Life Science), and NanoDrop[®] ND-1000 UV-Vis Spectrophotometer (Thermo Fisher Scientific) (Martina et al., 2023). To obtain CCS reads (HiFi) for contig assembly, the BB genome was sequenced using a PacBio

Revo sequencer (Pacific Biosciences) at Novogene (Beijing, China). The extracted RNA samples were sequenced using Illumina NovaSeq platforms (PE150) at Novogene (Beijing, China). Raw samples' reads were then cleaned using fastp tool (Chen, 2023).

For the qPCR validation, for the six BB corolla and the nine peel samples, cDNA was synthesized through reverse transcription using 1 μ g of extracted RNA and the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA), following protocol's instructions. The produced cDNA was diluted 1:6 and used for the following qPCR analyses. qPCR reactions were performed with a final volume of 10 μ L using a StepOnePlus Real-Time PCR system (Applied Biosystem), with technical duplicates. Each reaction contained 1, 5 μ L of starting cDNA, 2X Power SYBR Green PCR Master Mix (Applied Biosystem, USA) and specific primers (0.2 μ M. MYB75 primer For: 5'-CCGCATATCAAGAGAGGTGAC-3'; MYB75 primer Rev: 5'-TGCCTGTTTCTAGAAAGCTTATG-3'. EIL primer For:

5'-TGGGTGTTAGCTGCTAATTCTG-3'; EIL primer Rev: 5'-TCCCCACAGAACCCCATTC-3'. MSTRG.970.1 primer For: 5'-AGAAGGTGAACGTGGCAGAG-3'; MSTRG.970.1 primer Rev: 5'-ACCTCGATTCTCAGCTTGG-3'. MSTRG.3938.1 primer For: 5'-TCAATCACCGCCCTTAACC-3'; MSTRG.3938.1 primer Rev: 5'-TGGGGAAGAAGGGTTTGGTG-3'). The elongation factor (*SmEF*, primer For: 5'-ACCAGCATCACCATTCTTCA-3'; primer Rev: 5'-ACTGCCATACTTCCACATT-3') was used as eggplant's housekeeping gene. The following PCR protocol was used: 1 cycle of 95 °C for 5 min; 40 cycles of 95 °C for 15 s, 60 °C for 60 s. The $2^{-\Delta\Delta\text{Ct}}$ method was used for expression quantification based on Ct values of the target genes and the Ct value of the housekeeping gene (*SmEF*). Statistical significance between groups was assessed on ΔCt values using a two-tailed Welch's t-test, with $p < 0.05$ considered significant.

2.4 Genome assembly and annotation

Hifiasm v0.24 (Cheng et al., 2024) was used to construct a *de novo* assembly of the PacBio HiFi reads and Ragtag scaffold (Alonge et al., 2022) was employed for the assembly scaffolding, using the eggplant genome accession GPE001970 V5 (Gaccione et al., 2025). After using QUAST (Gurevich et al., 2013) to assess the final assembly quality of BB genome, chloroplast and mitochondrial genomes were removed based on their alignment to available eggplant plastidial and mitochondrial genome from NCBI (OP688482.1, OR187866.1), using minimap2 (v2.26). EDTA (Ou et al., 2019) was used to construct a custom repeat library and RepeatMasker was employed to low-mask the BB assembled genome. Subsequently, Helixer v0.3.4 (Holst et al., 2023) and Braker3 v3.0.3n (Gabriel et al., 2024) were used to perform genome annotation. The functional annotation was then conducted on the predicted proteins using the *Viridiplantae* NCBI Reference Sequence Database with diamond blastp algorithm (Buchfink et al., 2021).

BUSCO v6.0 (Manni et al., 2021), with the ‘*solanales_odb10*’ dataset, and Merqury v1.3 (Rhie et al., 2020) were used to assess the completeness of the genome assembly.

Minimap2 v2.24 (Li, 2021) was used to align the BB genome against the GPE001970 reference genome and the resulting PAF file was used by pafr R package to produce the dotplot, while GENESPACE v1.4 (Lovell et al., 2022) was used to perform the synteny analysis between the two genomes. CIRCOS (Krzywinski et al., 2009) was used to visualize the topography of the BB genome.

2.5 Transcripts alignment and isoform assembly

Filtered reads were aligned with STAR tool (Dobin et al., 2013), using options alignIntronMin 20 and alignIntronMax 20000, and BB assembled genome as a reference. Subsequently, the aligned reads were used as input for StringTie (Shumate et al., 2022) program to obtain the assembled transcripts of peel and corolla replicates. Then, Stringtie merge with default parameters was used to generate the integrated transcript information of all peel and corolla samples. The resulting GTF file was compared with the annotation of BB genome and integrated with the annotation of the assembled transcriptome using GffCompare (Pertea and Pertea, 2020).

2.6 Identification of putative lncRNAs and new protein-coding transcripts

To find putative long non-coding RNAs (lncRNAs), the final set of transcripts from GffCompare and BB annotated transcripts were filtered for length > 200 nt and class code “i”, “o”, “s”, “u” and “x”, excluding Chr0. The protein-coding-score was calculated for the resulting set of transcripts using Coding Potential Calculator 2 (CPC2) (Kang et al., 2017), FIEExible Extraction of lncRNAs (FEELnc) v0.2.1 (Wucher et al., 2017) and Coding Potential Assessment Tool (CPAT) v3.0.0 (Wang et al., 2013). Moreover, all transcripts with protein-coding domains found in Pfam databases (Paysan-Lafosse et al., 2025) were removed. The transcripts identified as noncoding by all the mentioned tools were considered as the total set of lncRNAs present in BB berries’ peel and flowers’ corolla, while all the transcripts which have passed the filter to be identified as lncRNAs that were conversely classified as coding were listed as novel coding mRNAs.

2.7 Detection of protein-coding gene AS events

For the Alternative Splicing (AS) events detection, transcripts with class code “j”, “m” and “n” were selected and used to run SUPPA2 v2.4 (Trincado et al., 2018), excluding Chr0. The local event types generated were the seven main ones: (I) Skipping Exon (SE); (II) Alternative 5’ Splice Sites (A5); (III) Alternative 3’ Splice

Sites (A3); (IV) Mutually Exclusive Exons (MX); (V) Retained Intron (RI); (VI) Alternative First Exons (AF); (VII) Alternative Last Exons (AL). The proportion spliced-in (PSI) values were extracted from the ‘psiPerEvent’ of local AS events for each sample and ‘diffSplice’ was performed to calculate the differentially spliced events between two conditions at a time (p-value cutoff: < 0.05).

2.8 Phylogenetic analysis of regulatory MYBs and protein domains assessment

BB genome V-myb myeloblastosis viral oncogene homologs (MYBs) predicted protein sequences were obtained using the BLASTP tool to identify known MYBs involved in the regulatory process of anthocyanin biosynthesis as activators or repressors in eggplant (Cammalleri et al., 2024). Multiple sequence alignment was performed, including protein sequences of the main anthocyanins activators: MYB1 (BB_10g017140.1), principal activator in fruits, AN2 (BB_10g017120.1) and MYB75 (BB_10g017150.1), principal activators in flowers, MYB94 (BB_03g025270.1), MYB19 (BB_09g019020.1), MYB5 (BB_08g001920.1) and MYB35 (BB_01g009850.1), other minor activators, detected as positive regulators (Supplementary File S1). As negative regulators, MYBL1 (BB_10g000390.1), the first documented MYB repressor, MYB44 (BB_02g022950.1), and MYB86 (BB_06g021820.1), minor putative repressors, were included (Supplementary File S1). The alignment was performed using MUSCLE algorithm (Edgar, 2004). The visualization of the phylogenetic tree was done by the neighbor-joining algorithm in MEGA11 (Tamura et al., 2021), applying a bootstrap analysis with 1,000 replicates to assess the statistical significance of nodes.

All MYBs TFs protein from BB were also evaluated for the presence of the R2R3 DNA-binding domains using InterPro online tool (<https://www.ebi.ac.uk/interpro/>).

2.9 Differential expression analysis, correlation network construction and evaluation of RNA-RNA interactions

Reads per transcript counts were obtained by re-assembling the lncRNAs, novel coding mRNAs and known coding mRNAs transcripts using StringTie with -e option. The Python script prepDE.py3 (provided by StringTie) was used to generate the raw count matrix. Subsequently, The Differential Expression Analysis (DEA) was conducted using Deseq2 package in R (Love et al., 2014), with a padj threshold set < 0.05 and a value of log2FC set to 1 for coding and novel coding transcripts and to 0.5 for lncRNAs.

The final filtered and normalized count matrix was used to perform the Principal Component Analysis (PCA) for the six corolla and nine peel samples separately.

Heatmaps were generated in R using the pheatmap package (<https://CRAN.R-project.org/package=pheatmap>) based on

variance-stabilized transformed (VST) expression values of significantly differentially expressed transcripts and lncRNAs, and were visualized by hierarchical clustering.

Next, the Transcript Per Million (TPM) normalized matrix was obtained by the function `convertCounts` from Differential Gene Expression (DGE) Analysis Utility Toolkit (Law et al., 2018) in R. This matrix was employed for the regulatory networks constructions using Spearman correlation coefficient, with a p-value cutoff set < 0.05 , to clarify the role of lncRNAs and AS in the biosynthesis and accumulation of anthocyanins. IntaRNA (Mann et al., 2017) was used to evaluate the interaction between strongly correlated alternative spliced transcripts and lncRNAs (correlation cutoff $> |0.98|$ and $> |0.9|$ for corolla and peel, respectively). The temperature parameter was set at 25 °C, reflecting the mean temperature in the greenhouse where the sampled plants were grown, and the results were filtered for a free energy < -30 kcal/mol. RNAfold and RNAcifold Web Servers (Lorenz et al., 2011) were used to visualize the RNAs secondary structures and RNA-RNA interactions. Transcription factors (TFs) known to be involved in anthocyanin regulation that presented an AS event were selected with their

respective correlated and interacting lncRNAs (the ten lncRNAs which presented the highest interaction), structural genes of the anthocyanins' pathway and other known regulatory TFs to obtain the final peel and corolla networks. The correlation values were filtered to keep only interactions with an absolute value greater than 0.8 and 0.7 for corolla and peel, respectively. Cytoscape v3.10.3 (Shannon et al., 2003) was used for networks visualization.

3 Results

3.1 Black Beauty genome sequencing, assembly and annotation

To achieve the most accurate transcripts identification, especially for long non-coding RNAs (lncRNAs), the complete genome assembly of the cultivar “Black Beauty” (BB) was generated. A total of 84 Gb of PacBio HiFi data were obtained (70X). The assembly generated 1,399 contigs, with an N50 of 57.2 Mb and an L50 of

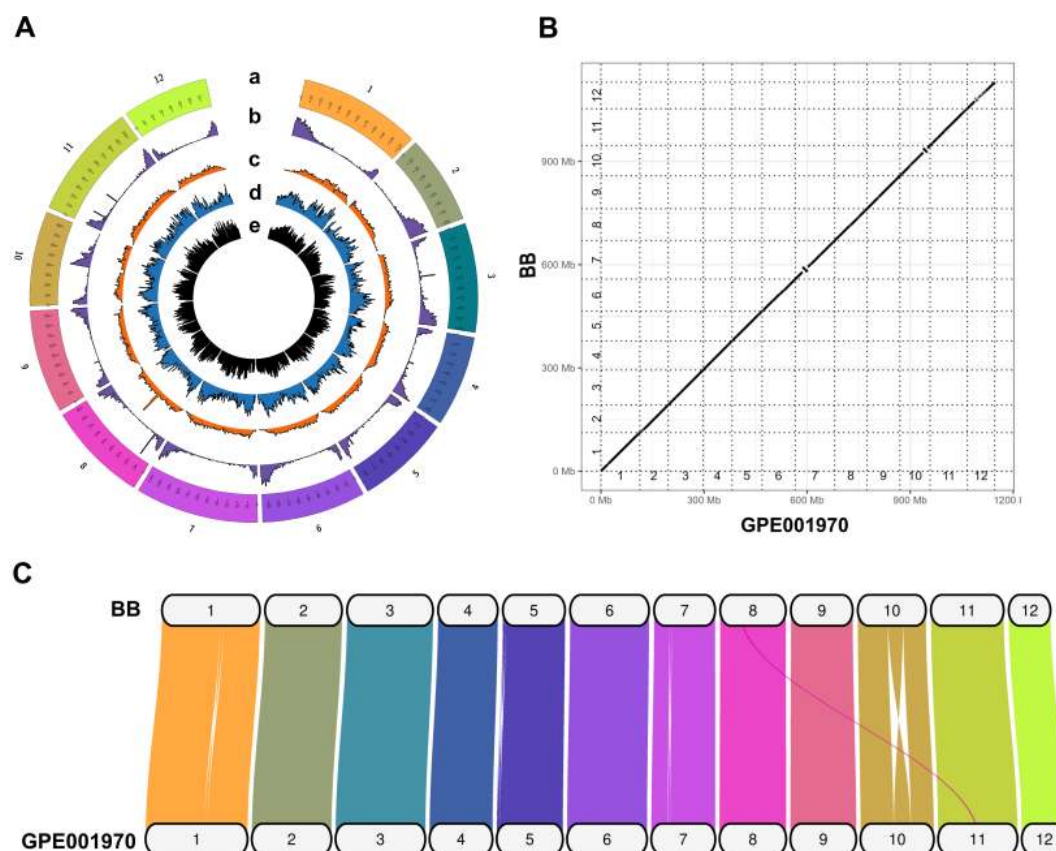


FIGURE 1

Graphical representation of BB genome and its collinearity with GPE001970. (A) Genomic landscape of the eggplant genome. Track (a) Eggplant chromosomes; track (b) gene density; track (c) LTR-Gypsy transposon density; track (d) LTR-Copia transposon density; track (e) DNA transposon density. For tracks (b–e), densities are shown in 1-Mb intervals. (B) Dotplot of the comparisons between BB and GPE001970 genomes (gene rank). (C) Synteny plot of the comparisons between BB and GPE001970 genomes (gene rank).

eight scaffolds, covering a total of 97.13% of the assembled genome. Following RagTag scaffolding using the latest version of the purple accession GPE001970 ("67/3") as reference, the 12 chromosomes were obtained, for a total genome length of 1.15 Gb, with 20.37 Mb of unplaced sequences (Supplementary Table 1).

BUSCO analysis returned 99.3% complete BUSCOs gene copies (97.0% single-copy, 2.3% duplicated), 0.0% fragmented, and 0.7% missing out of 7,934 genes searched. Merqury k-mer completeness analysis showed that 99.68% of distinct k-mers from the sequencing reads are present in the final assembly, with a consensus quality (QV) of 65,215 and an error rate of 3.00956×10^{-7} (Supplementary Table 1). Following the annotation, 74.76% of the genome was masked, and a total of 31,316 genes were predicted. To have an overview of BB genomic asset, gene, LTR Gypsy transposon, LTR Copia transposon and DNA transposon density are reported in Figure 1A.

We then compared our genome to the reference GPE001970 by synteny analysis, revealing strong collinearity. By investigating whether anthocyanin-related genes conserved their position and order between the two accessions, we found that the only structural variations that could comprehend anthocyanin-related genes was an inversion on chromosome 10 (Figure 1B). In particular, the inversion (~7.24 Mb, Figure 1C) contained three

chalcone isomerase genes (BB_10g014680.1, BB_10g014690.1, and BB_10g014700.1) with no sequence differences relative to the GPE001970 accession.

3.2 RNA-seq QC metrics and transcriptome assembly

To deeply characterize the anthocyanin biosynthesis and regulation in BB, both berries' peel and flowers' corolla tissues were sampled, following the progressive anthocyanin accumulation. RNA-sequencing (RNA-seq) was performed on (i) nine samples of peel (three biological replicates for each of the three stages, ST1 green, ST2 color shifting and ST3 completely purple peels, Figure 2A), and (ii) six flower's corolla samples (three biological replicates of STG, green corolla, and STP, purple corolla, Figure 2A), yielding a total of 138 Gb of raw reads. After trimming the adaptors and removing low-quality reads, the number of cleaned reads retained per sample ranged from 46 to 72 million (Supplementary Table 2). All 15 samples showed high sequencing quality, with Q30 scores between 91% and 94%, ranging from 6.9 and 10.7 Gb of high-quality reads per sample, which yield a percentage between 90% and 99% of mapped reads on BB genome. After transcriptome assembly, the integrated set of

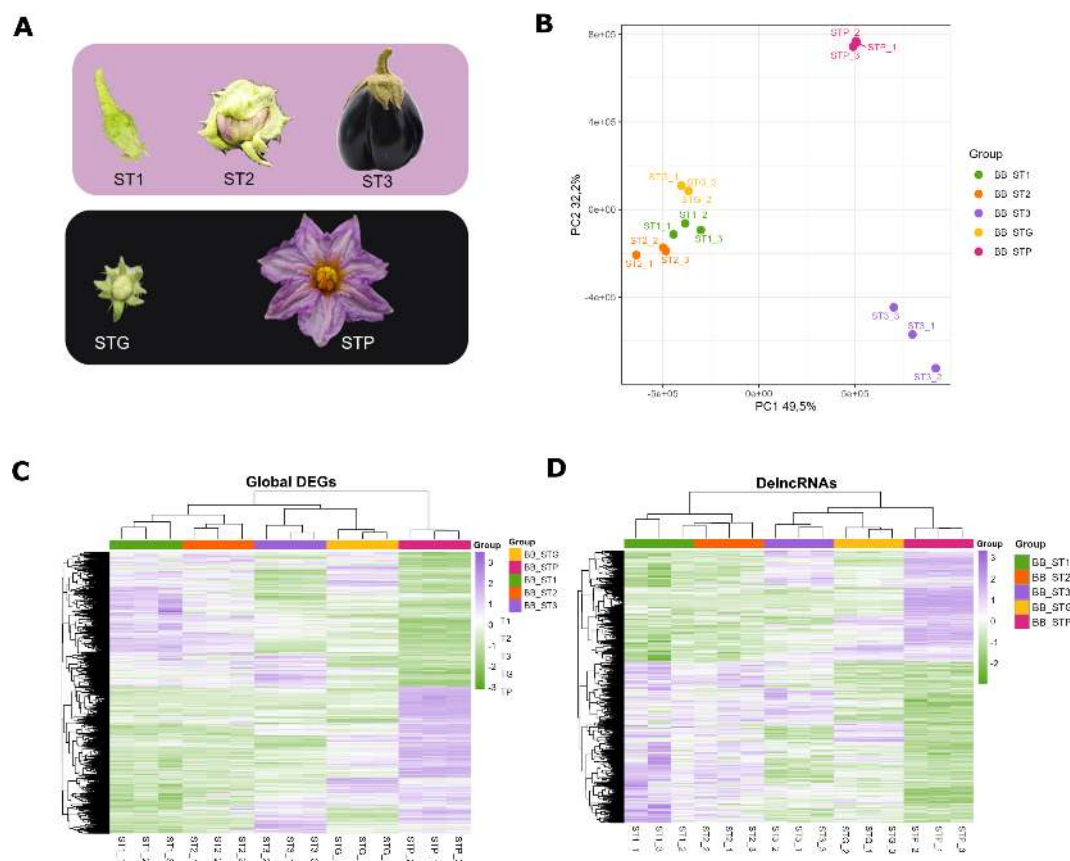


FIGURE 2

Sample clustering and expression patterns in fruit peel and corolla across developmental stages. (A) Developmental stages analyzed in fruits and flowers. Fruit stages are shown in the upper panel: Stage 1, Stage 2, and Stage 3 (ST1, ST2, and ST3). Flower stages are shown in the lower panel: green and purple stages (ST G and ST P). (B) Principal component analysis (PCA) of the samples. (C) Heatmap of the scaled TPM matrix of all genes across peel and corolla developmental stages. (D) Heatmap of the scaled TPM matrix of differentially expressed lncRNAs (DElncRNAs) across peel and corolla developmental stages.

transcripts obtained from all 15 samples included 71, 243 transcripts (Supplementary Table 2).

3.3 Identification of lncRNAs, new protein-coding genes and protein-coding genes AS events

The transcriptome assembly of the 15 BB RNA samples was obtained to identify lncRNAs, new protein-coding genes and AS events, based on a comparison of the primary BB annotation. The integrated transcriptome assembly set included 71, 243 transcripts, which were then filtered on the basis of their position in the genome.

For lncRNAs, the transcripts were filtered for length > 200 nt, and class codes “i”, “o”, “s”, “u” and “x”, obtaining 5, 604 putative lncRNAs. Of these transcripts, only the ones that presented a coding potential < 0 according to Coding Potential Calculator 2 (CPC2), FLEExible Extraction of lncRNAs (FEELnc) and Coding Potential Assessment Tool (CPAT) were selected. After PfamScan filtering on these transcripts, 4, 126 lncRNAs were predicted (Supplementary Figure 2A). Of these lncRNAs, 208 belonged to the class code “i”, 946 to the class code “o”, 4, 449 to the class code “u” and only 1 to the class code “x”, while no transcripts were found in class code “s” (Supplementary Table 3). Transcripts with coding potential > 0 were also evaluated for protein domains, and a final number of 179 transcripts were considered as novel coding genes (Supplementary Figure 2B). For protein-coding genes AS events, the integrated transcriptome was filtered for class codes “n”, “m”, and “j”, amounting to 34, 745 new isoforms detected. After SUPPA analysis, AS events were divided for their AS site type and are listed in Supplementary Table 3. The A3 was the most frequent AS type in BB’s peels and corolla, followed by SE, A5 and RI.

The presence of some lncRNAs and protein-coding genes AS events relevant for further analysis has been verified by qPCR (Supplementary Table 4).

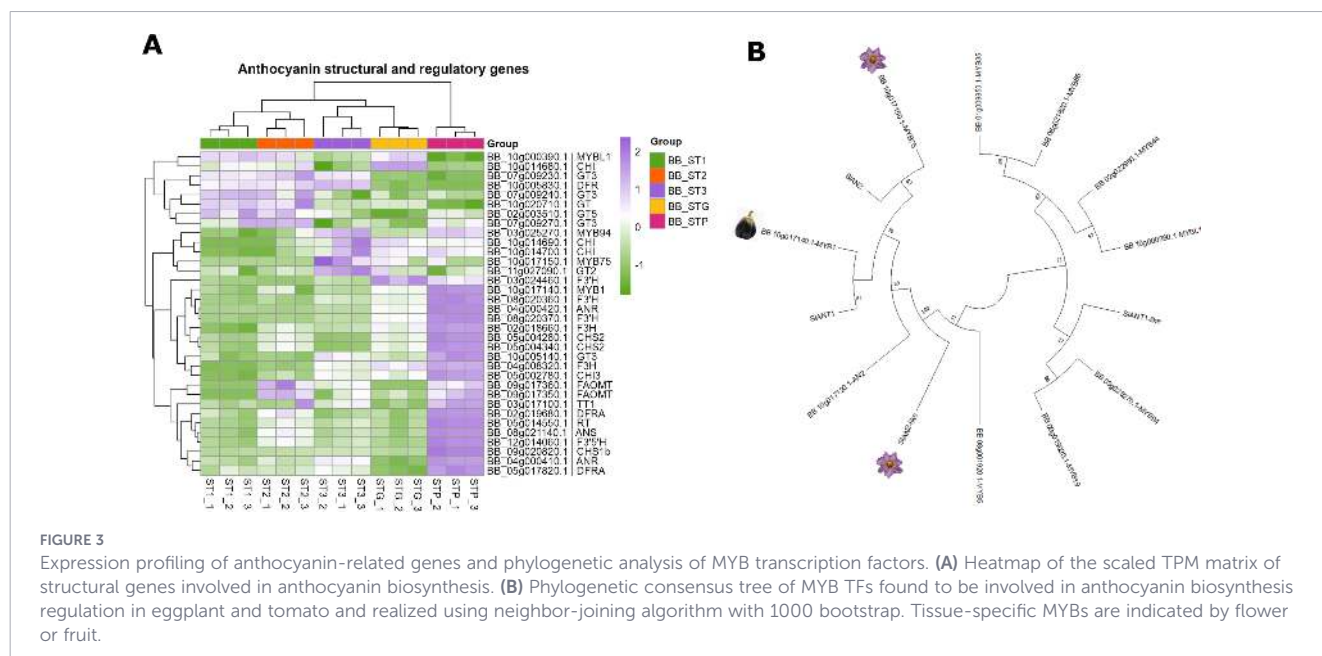
3.4 Identification of differentially expressed genes and lncRNAs

A principal component analysis (PCA), performed on the 15 RNA samples normalized and filtered counts, clearly separated them into the five different groups (ST1, ST2, ST3, STG and STP, Figure 2B). The first two principal components (PC1 and PC2) accounted for 81.7% of the variance (Figure 2B). For the Differential expression analysis (DEA), coding and novel coding transcripts were filtered for a Log2 fold change (log2FC) of 1, while lncRNAs for a log2FC of 0.5.

Specifically, 855 transcripts were up-regulated and 889 were down-regulated in ST1 compared to ST2 in the peel, whereas 3, 035 transcripts were up-regulated and 2, 854 down-regulated in ST2 compared to ST3 in the same tissue. In the corolla, comparison between STG and STP revealed 5, 685 up-regulated and 5, 227 down-regulated transcripts (Figure 2C, Supplementary Table 5). Among the DEGs, 28 transcripts in the peel and 51 in the corolla belonged to the novel coding transcripts’ category (Supplementary Table 5).

For differentially expressed lncRNAs (DElncRNAs), a total of 683 DElncRNAs were detected in the peel and 1, 120 DElncRNAs in the corolla (Figure 2D). Of these lncRNAs, 326 were common DElncRNAs in both peel and corolla (Supplementary Table 6). The expression patterns of DElncRNAs were consistent within each sample group, highlighting pronounced differences between peel ST3, corolla’s STG and STP. In contrast, ST1 and ST2 exhibited a similar expression pattern, with only 25 lncRNAs down-regulated and 22 up-regulated between the two ripening stages Supplementary Table 6.

The heatmap in Figure 3A illustrates the expression profiles of structural genes involved in anthocyanin biosynthesis. At least one copy of each structural gene was expressed in the purple stages (*i.e.* ST2, ST3, and STP). In STP all structural genes were up-regulated, except *dihydroflavonol 4-reductase* (DFR; BB_10g005830.1), which



maintained high expression across all peel stages, and some of the *glucosyltransferases* (GT; GT2: BB_11g027090.1; GT3: BB_07g009230.1 and BB_07g009240.1).

In the peel, clear stage-dependent differences emerged from the hydroxylation step. In fact, *flavonoid 3'-hydroxylase* (F3'H, BB_08g020370.1) was up-regulated in ST3 compared to ST2, whereas, *flavonoid 3'5'-hydroxylase* (F3'5'H, BB_12g014060.1) was up-regulated in ST2 compared to ST1. For *chalcone synthase* (CHS), BB_05g004280.1 and BB_05g004340.1 followed the same expression trend of MYB1 in ST2, while CHS1b (BB_09g020820.1) was more expressed in ST3. Similarly, among chalcone isomerase (CHI) copies, BB_10g014700.1, BB_10g014690.1 and BB_10g014680.1 were more abundant in ST2, while CHI3 (BB_05g002780.1) increased its expression through the three stages, being up-regulated in ST3. A comparable pattern was observed for the *flavanone 3-hydroxylases* (F3H), with BB_02g01860.1 predominantly expressed in ST2 and BB_04g008320.1 in ST3. Likewise, different *dihydroflavonol 4-reductase* (DFRA) copies were preferentially expressed in distinct stages, with BB_02g019680.1 dominating in ST2 and BB_05g017820.1 in ST3.

Other structural genes showed limited variation: anthocyanidin synthase (ANS) expression remains high and constant from ST1 to ST3, likewise *O-methyltransferases* (FAOMT) copies, which catalyze the methylation of delphinidin leading to petunidin and malvidin formation, and cyanidin, leading to peonidin formation (Cammareri et al., 2024), maintained the same trend, suggesting no variation in their expression in the three stages (Figure 3A).

Regarding transcription factors, MYB1 did not show significant regulation in any of the peel purple stages, although a slight, non-significant increase in expression was detectable between ST1 and ST2 (p-value 0.064). By contrast, MYB1 was significantly up-regulated in the purple stage of the corolla compared to the green one. Similarly, AN2 and MYB75 were up-regulated in STP, but showed no expression in the peel, confirming their tissue-specific function. Among minor regulators (Cammareri et al., 2024), MYB94 was the only activator up-regulated in both ST3 compared to ST2 and STP compared to STG, suggesting its potential involvement in the pathway regulation in both tissues. Conversely, MYBL1 was down-regulated in ST3 compared to ST2 and in STP compared to STG, while no other minor repressors displayed differential expression (Figure 3A).

Because transcription factors involved in anthocyanin biosynthesis may influence pigment accumulation indirectly via hormone-related pathways, genes associated with ethylene and auxin signaling were also examined. In the purple stages of both peel and corolla, particularly in ST3 and STP, several *ERFs* and *Ethylene Insensitive3-Like* (EIL, BB_06g023930.1), known enhancers of *ERF* transcription (Feng et al., 2025), were up-regulated compared to green stages, along with *1-Aminocyclopropane-1-Carboxylate Synthase* (ACS) and *1-Aminocyclopropane-1-Carboxylic Acid Oxidase* (ACO) transcripts, which are responsible for ethylene biosynthesis. In contrast, transcripts responsible for auxin biosynthesis, *Tryptophan-pyruvate Aminotransferase 1* (TAR1) and *Indole-3-pyruvate Monooxygenase* (YUCCA), along with *Auxin*

Response Factors (ARFs) exhibited higher expression levels in ST1, ST2 and STG, while the *Auxin/Indole-3-Acetic Acid* (IAA) transcripts, known as ARFs repressors, were predominantly expressed in ST2 and STP (Supplementary Table 5).

3.5 Phylogenetic analysis of Myb TFs involved in anthocyanin biosynthesis

Since MYB94 was specifically activated in ST3 and STP, we investigated its phylogenetic placement. Therefore, the main MYB TFs previously reported as activators or repressors in eggplant (Yang et al., 2022; Cammareri et al., 2024) were identified in the BB genome by their protein sequences and annotation. A neighbor-joining phylogenetic tree based on their protein sequences was constructed, including *S. lycopersicum* MYB proteins previously characterized as activators (SIAN2, SIAN2-like, SIANT1 and SIANT1-like, Supplementary File S1). The resulting tree revealed two major clades corresponding to the known functional MYB types. The first clade included the principal activator MYBs in both eggplant and tomato, with MYB1 clustering closely to SIANT1 and MYB75 with SIAN2, as expected. The second clade comprised two subclades: the first one included SIANT1-like, for which no homolog has yet been found in eggplant, clustering with MYB94 and MYB19; the second subclade grouped the repressor MYBL1 with the two minor repressors and MYB35 (Figure 3B).

3.6 Differentially spliced events in protein-coding genes

Concerning AS events, the differentially spliced events (DSE) and their respective alternatively spliced genes were detected: a total number of 1, 915 and 1, 709 DSE were found in the peel and corolla stages, respectively, with 1, 318 and 1, 244 alternative spliced genes. The number of DSE and of differentially spliced genes for each type of event is reported in Figure 4A, B for the peel and corolla, respectively.

Among these differentially spliced genes, those linked to anthocyanin biosynthesis or its regulation in peel and in corolla are represented in Figures 4C, D, respectively, and listed in Supplementary Table 7. Additionally, it was also investigated whether, among these events, differentially spliced genes could overlap with DEGs, particularly focusing on TFs known to be involved in anthocyanin pathway regulation. In the corolla, MYB75 generated a retained intron (RI) event, retaining both its introns and leading to the loss of the R3 DNA-binding domain and an imperfect version of the R2 domain (Supplementary Figure 3). This transcript was up-regulated in STP, although its alternative spliced isoform was more frequent in STG. In the peel, no AS MYBs involved in the anthocyanin regulation were detected, however, an *Ethylene Insensitive3-Like* (EIL; BB_06g023930.1) gene generating an alternative 5' splice site (A5) event was identified. The canonical isoform was up-regulated in ST3, whereas the A5-spliced variant was more abundant in ST2. Protein domain analysis revealed no structural differences between the two isoforms.

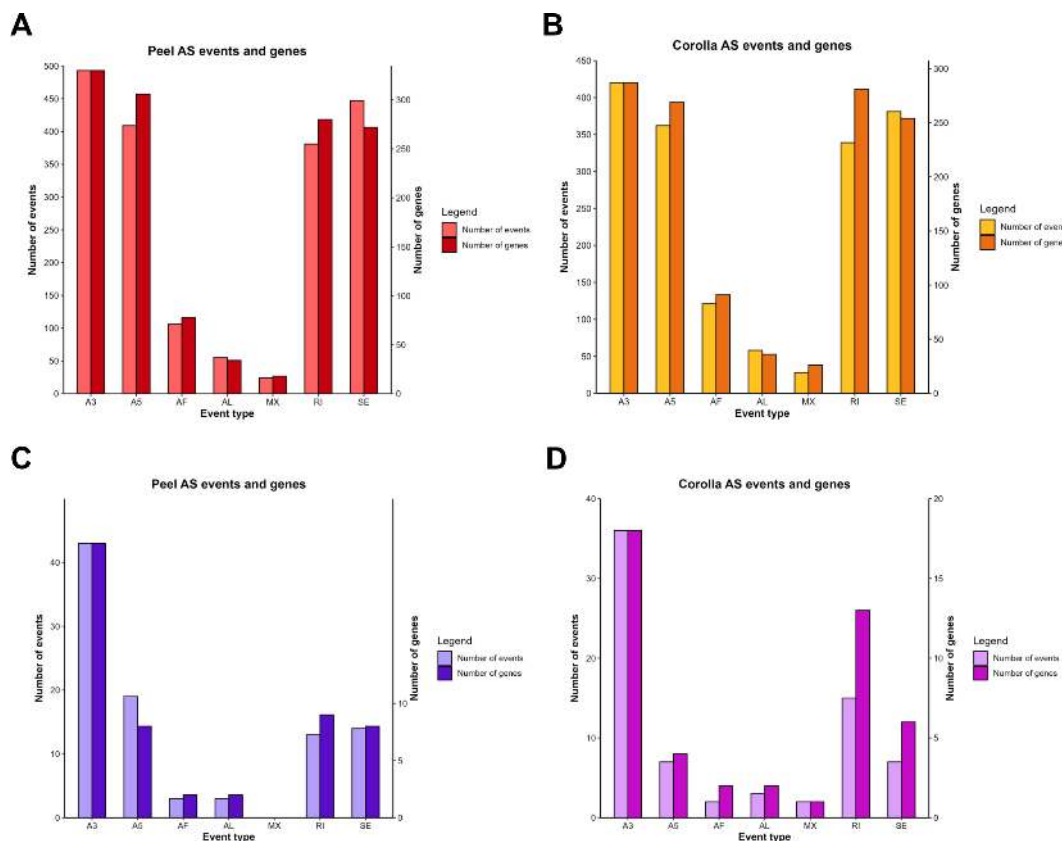


FIGURE 4

Visual representation DSE of protein-coding genes in BB fruit peel and flower corolla. (A) DSE and originating genes detected in the peel. (B) DSE and originating genes detected in the corolla. (C) DSE and originating genes linked to anthocyanin biosynthesis and regulation detected in the peel. (D) DSE and originating genes linked to biosynthesis and regulation detected in the corolla. Alternative 3' Splice Sites (A3); Alternative 5' Splice Sites (A5); Alternative First Exons (AF); Alternative Last Exons (AL); Mutually Exclusive Exons (MX); Retained Intron (RI); Skipping Exon (SE).

3.7 Network construction and lncRNA interaction with target AS events

To investigate the relationships between DEGs, DSE and DElncRNAs involved in anthocyanin biosynthesis and regulation, a regulatory network was constructed using the Spearman correlation coefficient in order to evaluate potential trans-interaction between transcripts.

First, we selected all the DElncRNAs that displayed a strong correlation with *MYB1*, *MYB94*, *MYB75*, *AN2* and *MYBL1* to understand if they could be involved in the pathway regulation by interacting with these main TFs. In the peel (Figure 5A), *MYB94* showed the majority of correlated DElncRNAs, while *MYB1* was in an independent cluster with only six DElncRNAs. Interestingly, the structural genes also clustered differently between the two *MYBs*.

To construct the final network, which involved DElncRNAs and AS events, the ten DElncRNAs with the highest interaction free energy with *EIL* were selected from the 94 initial set (Supplementary Table 8). Among them, *MSTRG.3938.1* showed the highest interaction free energy (-88.55 kcal/mol) with both the canonical spliced version and the A5 spliced isoform of *EIL*, determining that the bond took place in a region present in both transcripts (Figures 5B, C). *MSTRG.3938.1* was negatively correlated with *EIL* and *MYB94*, while exhibiting a positive correlation with *MYBL1*. *MYB94* showed a strong positive correlation with *EIL*, its homologues, several *ERFs*

and structural genes of the anthocyanin biosynthesis, including *CHS1b*, *CHI3*, *F3'H* and *DFRA*, all of which were more expressed in ST3 compared to ST2. In contrast, *MYB94* was negatively correlated with *MSTRG.3938.1* and *MYBL1*, which displayed negative correlations with the structural genes as well as with some of the *IAA* and *ARF* TFs down-regulated in ST3 (Supplementary Table 9). *MYB1* was only positively correlated to other copies of the structural genes, confirming the presence of a distinct cluster from the *MYB94* one (Figure 5D).

In the corolla, the DElncRNAs correlated with the *MYB* TFs create a more homogeneous network, with several DElncRNAs correlated with more than one TF (Figure 6A).

To construct the final corolla network we also selected other TFs, identified as DSE, interacting with *MYB75*. The latter showed a negative correlation with *CONSTANS16-like* (BB_03g028630.1, A3), whose family plays a key role in regulating flowering time and can promote anthocyanin biosynthesis (Liu et al., 2025), and a positive correlation with *IAA13* (BB_09g011750.1, A3). The three TFs were correlated to 39 common DElncRNAs, from which the ten with the highest interaction free energy were selected with the DEGs directly or indirectly involved in anthocyanin biosynthesis regulation (Supplementary Table 8). Among them, the lncRNA *MSTRG.970.1* was negatively correlated with *MYB75* and displayed a free energy interaction of -57.22 kcal/mol only with its RI isoform (Figures 6B, C). *MSTRG.970.1* also showed the potential to interact

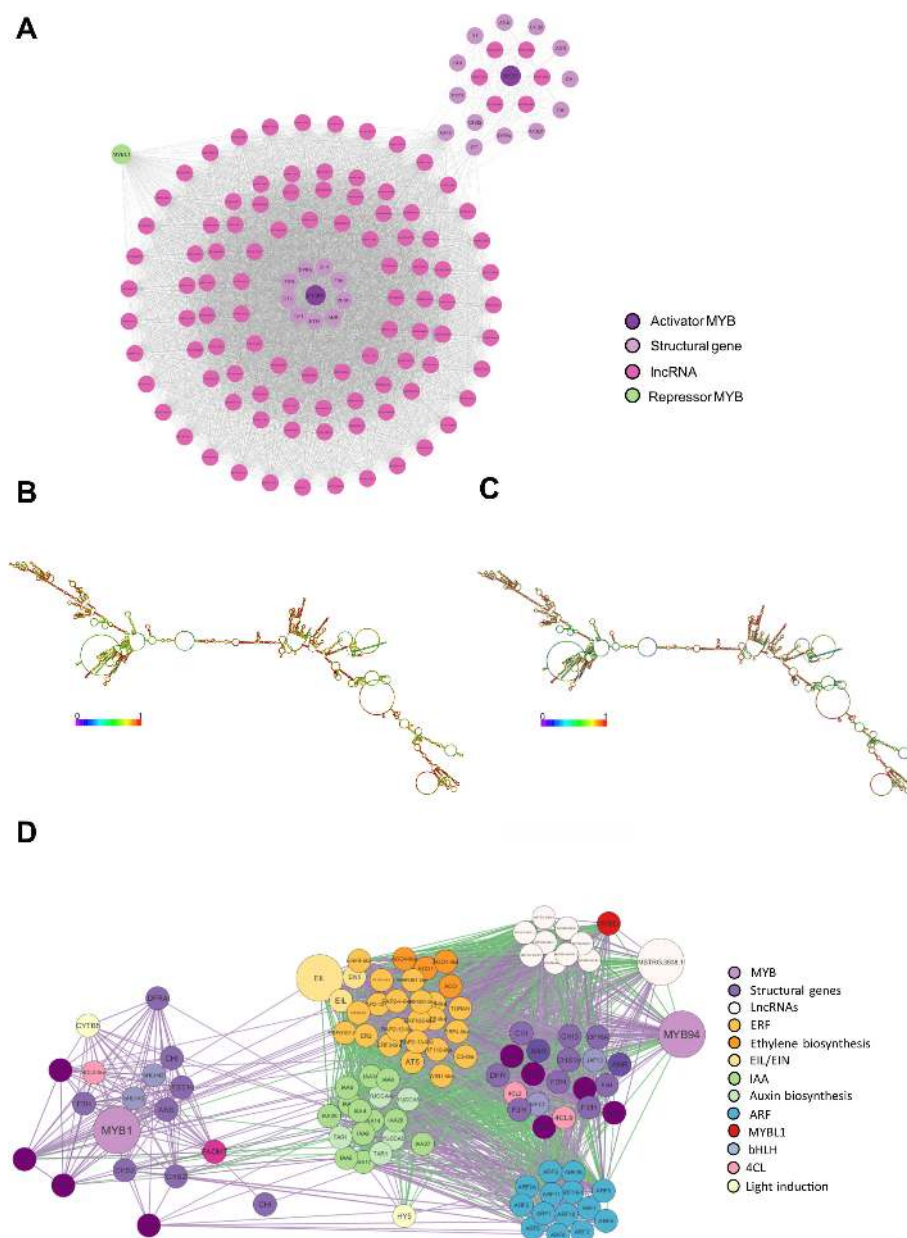


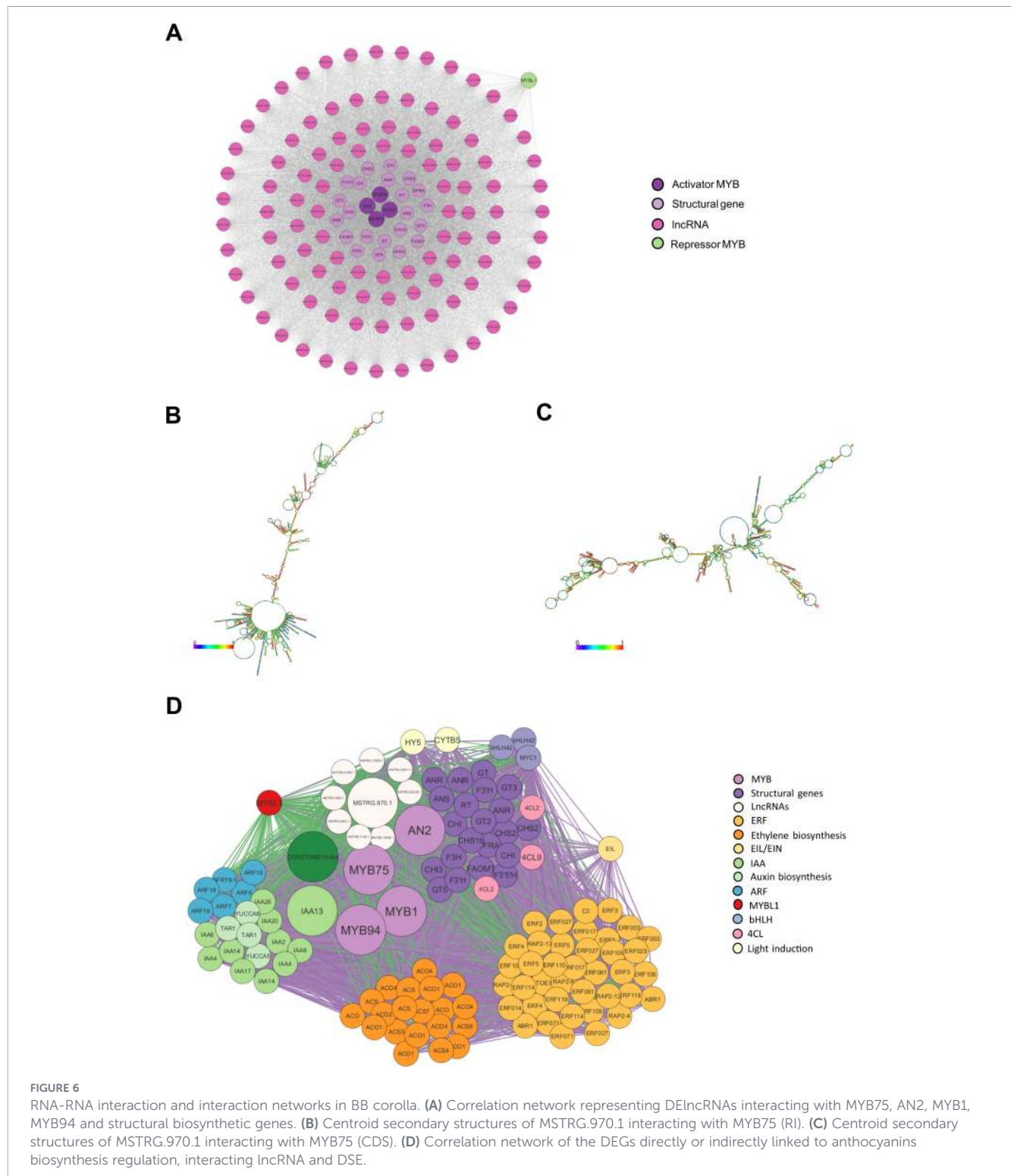
FIGURE 5
RNA-RNA interaction and interaction networks in BB peel. **(A)** Correlation network representing DElncRNAs interacting with MYB1 and MYB94 and structural biosynthetic genes. **(B)** Centroid secondary structures of MSTRG.3938.1 interacting with EIL (A5). **(C)** Centroid secondary structures of MSTRG.3938.1 interacting with EIL (CDS). **(D)** Correlation network of the DEGs directly or indirectly linked to anthocyanin biosynthesis regulation, interacting DElncRNA and DSE.

with the canonical spliced isoform of MYB75, but with a significantly lower free energy interaction (Supplementary Table 9). Furthermore, this lncRNA was negatively correlated with *IAA13*, and positively correlated with *CONSTANS16-like*.

MYB75 also displayed positive correlation with *AN2*, *MYB94* and *MYB1*, all of which were positively related to different copies of the structural genes. In contrast, *MYB75* was negatively correlated with *MYBL1*, which displayed a negative correlation with *IAA13* and a positive one with *CONSTANS16-like* and *MSTRG.970.1*. Moreover, in the corolla, several *ERFs* were positively correlated with *MYB* activators, along with *IAA* TFs (Figure 6D).

3.8 Isolation, identification and quantification of anthocyanin compounds

Overall, anthocyanin-related metabolites displayed clear stage-dependent accumulation patterns, with distinct profiles characterizing corolla and peel samples across development. Specifically, C3OG and Pt3OG are primarily responsible for discriminating between the early stages (STG and STP), while diglucosides and oxidized derivatives, such as D3OR, characterize the more advanced stages (ST1 and ST2). Indeed, in STG samples, Pt3OG and C3OG represent the most significant proportion of total anthocyanins



(21.16% and 16.94%, respectively), followed by Pn3OG, M3OG, Pl3OG, and diglucosides, which together contribute less than 15%. In the STP samples, there is a further increase in Pt3OG (27.63%) and C3OG (22.76%), while D3OR shows a slight increase (6.61%), indicating the beginning of the metabolic transition. Concerning the stages ST2 and ST3, the anthocyanin profile undergoes a drastic change. D3OR becomes the predominant compound, increasing from 5–6% in the early stages to over 38–40% in the advanced

stages, while C3OG and Pt3OG decrease significantly (C3OG 10.3–8.09%; Pt3OG 18.63–16.11%). Diglucosides show different trends: M3dG increases to 7.22% in ST2 before decreasing slightly in ST3, while D3dG and other diglucosides maintain intermediate values. The forms derived from peonidin and petunidin (Pn3dG and Pt3dG) remain marginal (<3% in ST2 and ST3). As shown in the heatmap coupled with cluster analysis (Figure 7A), thanks to different chemical compositions, samples grouped into two different

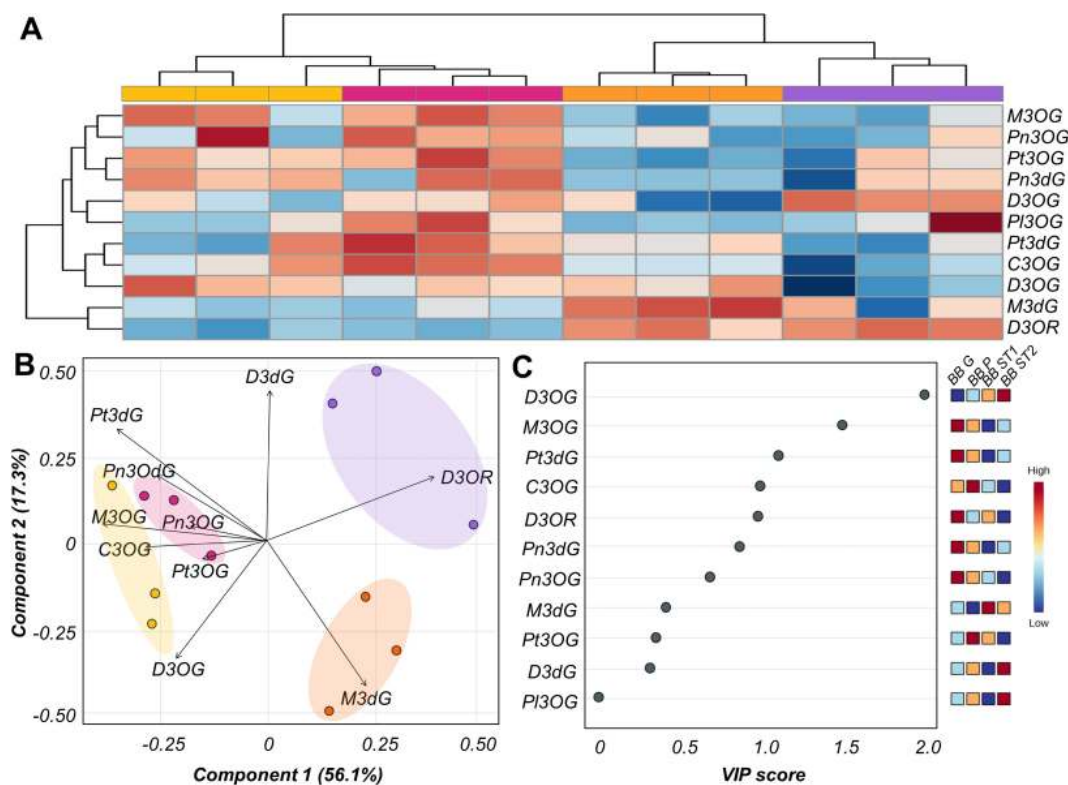


FIGURE 7

Stage-dependent variation of anthocyanin-related metabolites identified by HPLC-DAD-ESI-MS/MS analysis. (A) Heatmap coupled with hierarchical clustering showing the relative accumulation patterns of anthocyanins in BB flower corolla (BB_STG and BB_STP) and fruit peel (BB_ST2 and BB_ST3), highlighting stage- and tissue-dependent differences in metabolite abundance. (B) SPL-DA biplot illustrating the separation of samples based on their anthocyanin profiles across developmental stages and tissues. (C) VIP score plot identifying the metabolites that contribute most to the observed discrimination among samples.

roots. This separation was further confirmed by the biplot originated from SPL-DA analysis (Figure 7B), which highlighted the different organization of the samples and further reinforced these observations, allowing for maximum separation between groups based on anthocyanin content. The supervised model showed a clear distinction between the four stages considered. In particular, the division of the samples is clearly delineated along the first two principal components: PC1, which explained 56.1% of the variability, clearly discriminates STG and STP (characterized by negative scores) from ST2 and ST3 (positive scores), while PC2, which explained 17.3% of the variability, further distinguishes ST2 (negative scores) from ST3 (positive scores). Furthermore, VIP (Variable Importance in Projection) values (Figure 7C) allowed the metabolites most responsible for the separation to be identified, highlighting how simple glycosides are mainly associated with the early stages, while diglycoside and substituted derivatives characterize the mature stages.

4 Discussion

The eggplant anthocyanins biosynthetic pathway has been extensively characterized in terms of both structural genes and their transcription factors (TFs). These TFs belong to the MBW (MYB-bHLH-WD40) complex that has also been deeply studied,

leading to the identification of major activators and some repressors within the MYB TF family (Camareri et al., 2024). However, to our knowledge, only a few studies (Wang et al., 2022a) have explored the alternative splicing (AS) events associated with anthocyanins biosynthesis in eggplant, and no studies have addressed the role of long non-coding RNAs (lncRNAs) in this process.

To expand the understanding of anthocyanin regulation in this species, a *de novo* genome assembly of the “Black Beauty” (BB) cultivar was generated. This high-quality, contiguous assembly provided a robust reference framework for accurate transcript identification and expression profiling within a consistent genetic background. The integrated analysis of BB peel and corolla samples enabled a comprehensive reconstruction of the transcriptome, leading to the identification of 71,243 transcripts, including 31,316 genes predicted after the genome annotation, 4,126 lncRNAs, 179 novel protein-coding genes, and 34,745 alternative splicing events.

Moreover, the genomic positions of the main structural genes and transcription factors involved in the anthocyanin biosynthetic pathway were found to be conserved between the BB genome and the eggplant reference line “67/3”, except for an inversion of chromosome 10, encompassing three *chalcone isomerase* (*CHI*) gene copies, which did not present any variation in their sequences compared to the reference. Therefore, these structural variations are not associated with anthocyanin production, as both cultivars display purple pigmentation in peel and corolla.

4.1 Regulation in fruit peel

The differential expression analysis (DEA) suggested that ST2, corresponding to the early color transition from green to purple in the peel, differed more from ST3 than from ST1. Indeed, only a subset of the anthocyanin biosynthesis structural genes was up-regulated in ST2, correlating with *MYB1* transcription pattern. In fact, *MYB1* expression decreased between ST2 and ST3, as observed for all copies of its correlated structural genes, including *F3'5'H*. These data are consistent with the content of delphinidin-3-O-rutinoside we measured, which was the most present anthocyanin compound in the peel, as expected.

In contrast, *MYB94* expression increases significantly between ST2 and ST3, showing a strong correlation with alternative copies of both EBGs (*CHS1b*, *CHI*, *F3H*) and LBGs (*F3'H* and *DFRA*). Consistently, phylogenetic analysis placed *MYB94* in a subclade together with *SLANT1-like*, a previously characterized activator in tomato (Colanero et al., 2018), suggesting that *MYB94* may represent its eggplant ortholog.

The differential expression of *F3'H* (BB_03g024460.1, BB_08g020360.1, and BB_08g020370.1) and *F3'5'H* (BB_08g020370.1) as well as the different paralogues of *DFRs* (BB_02g019680.1 in ST2 and BB_05g017820.1 in ST3) between ST2 and ST3 could determine a shift in the type of anthocyanin compound produced, as reflected by the slight increase of peonidin-3-O-glucoside and petunidin-3-O-glucoside observed in ST3 samples. These findings are consistent with the recent observations in *Passiflora* spp. where *F3'H* seems to regulate cyanidin accumulation, *F3'5'H* determines blue pigmentation, and *DFR* enhances the biosynthesis of pelargonidins (Nutricati et al., 2025). Moreover, in *Clarkia gracilis*, the spatial and temporal differential expression of the *DFR* paralogues has been reported to influence the anthocyanin-type composition (Martins et al., 2013).

Therefore, we hypothesize that in eggplant fruit peel, two complementary pathways are activated by two distinct *MYB* transcription factors, *MYB1* and *MYB94*, leading to the synthesis of different anthocyanin compounds: delphinidin and cyanidin at two different developmental stages, ST2 and ST3, respectively. This is in accordance with the temporal *MYB*-mediated regulation of anthocyanin synthesis reported in most plants (Schwinn et al., 2006; Albert et al., 2011; Azuma et al., 2015). The possibility that two *MYB* activators regulate specific paralogous copies of structural genes has already been reported in other species, such as *C. gracilis*, where two *MYB* activators were found to be expressed at distinct developmental stages, following the same trend as the structural genes activated in each stage (Lin and Rausher, 2021).

Similarly, in *Gerbera hybrida*, the *CHS* gene most closely associated with cyanidin production was *CHS4* (Deng et al., 2014), whose protein sequence shares 88.1% similarity with *CHS1b* in eggplant. This finding supports the idea that *CHS1b*, which is associated with the pathway regulated by *MYB94*, may play a key role in eggplant cyanidin biosynthesis.

Ethylene has been identified as a positive regulator of the anthocyanin biosynthetic pathway in several species, such as apple and strawberry, both producing cyanidin (Wang et al., 2022b), suggesting a specific involvement of this hormone in its

synthesis. Although eggplant is a non-climacteric fruit, ethylene is marginally produced during the berries' maturation, and can contribute to the ripening process (Fan et al., 2016). In our study, transcriptomic evidence indicated a potential increase in ethylene biosynthesis at ST3, as suggested by the up-regulation of several ethylene biosynthesis structural genes (e.g. 1-Aminocyclopropane-1-Carboxylate Synthase, *ACS*, and 1-Aminocyclopropane-1-Carboxylic Acid Oxidase, *ACO*) and *Ethylene Response Factor* (ERF) transcription factors (TFs). These transcripts were positively correlated with *Ethylene Insensitive3-Like* (*EIL*) and *MYB94*, suggesting a role for ethylene in activating the alternative pathway proposed (Figure 8A). Our analysis revealed that, in the peel, *EIL* underwent an Alternative 5' Splice Site (A5) event. The A5 isoform conserved all the protein domains of its canonical version and was predominant in ST2, where *EIL* expression is down-regulated compared to ST3. *EIL* genes are known as activators of *ERF* TFs, which can also promote the ripening stage in several plant species, including tomato (Gu et al., 2017). Recently, it has been proposed that *ERF* proteins can interact with *MYB* TFs in eggplant, forming complexes aiming to activate anthocyanin structural genes in response to light signaling (Li et al., 2025). In ST3 we have not observed an increase in *HY5* transcript, which was instead strongly expressed in ST2, therefore, we cannot exclude a light-dependent mechanism regulating the increase of ethylene biosynthesis detected in ST3.

In apple, it has been reported that lncRNA could influence the anthocyanin production, in particular the expression of the lncRNA *MdLNC610* regulates the transcription of the ethylene biosynthetic gene *MdACO1*, resulting in the accumulation of cyanidin in fruit peel. In our study we identified a lncRNA (*MSTRG.3938.1*) upregulated in ST2 that is predicted to bind both the A5 and the canonical *EIL* isoforms, potentially influencing its transcription and consequently the ethylene-signaling cascade. Moreover, *MSTRG.3938.1* was found negatively correlated also with *MYB94*, but positively with the R3-MYB repressor *MYBL1* which is down-regulated in ST3. These findings suggest that the lncRNA-mediated modulation of *EIL* and the *MYBL1* activity may indirectly influence *MYB94* expression and, consequently, the transcription of its correlated structural gene.

4.2 Regulation in flower

In corolla, the DEA highlighted the expression of *AN2* and *MYB75* in STP, which were absent in the peel, confirming their tissue-specific activity (Cammareri et al., 2024). Also, *MYB1* and *MYB94* were expressed in the STP, but in lower level compared to the other two *MYBs*. Among the structural genes and *MYBs*, their differences in the expression are not correlated with the composition of the final anthocyanin compounds produced, since all the paralogues of the structural genes were positively correlated with at least one of the *MYBs*. The metabolic analysis confirmed that, also in the corolla, the total anthocyanin content increased in STP compared to STG, identifying cyanidin-3-O-glucoside and petunidin-3-O-glucoside as the most relevant compounds.

In plants, pigmentation has been demonstrated to be regulated by AS of the TF responsible for the activation of colored molecule

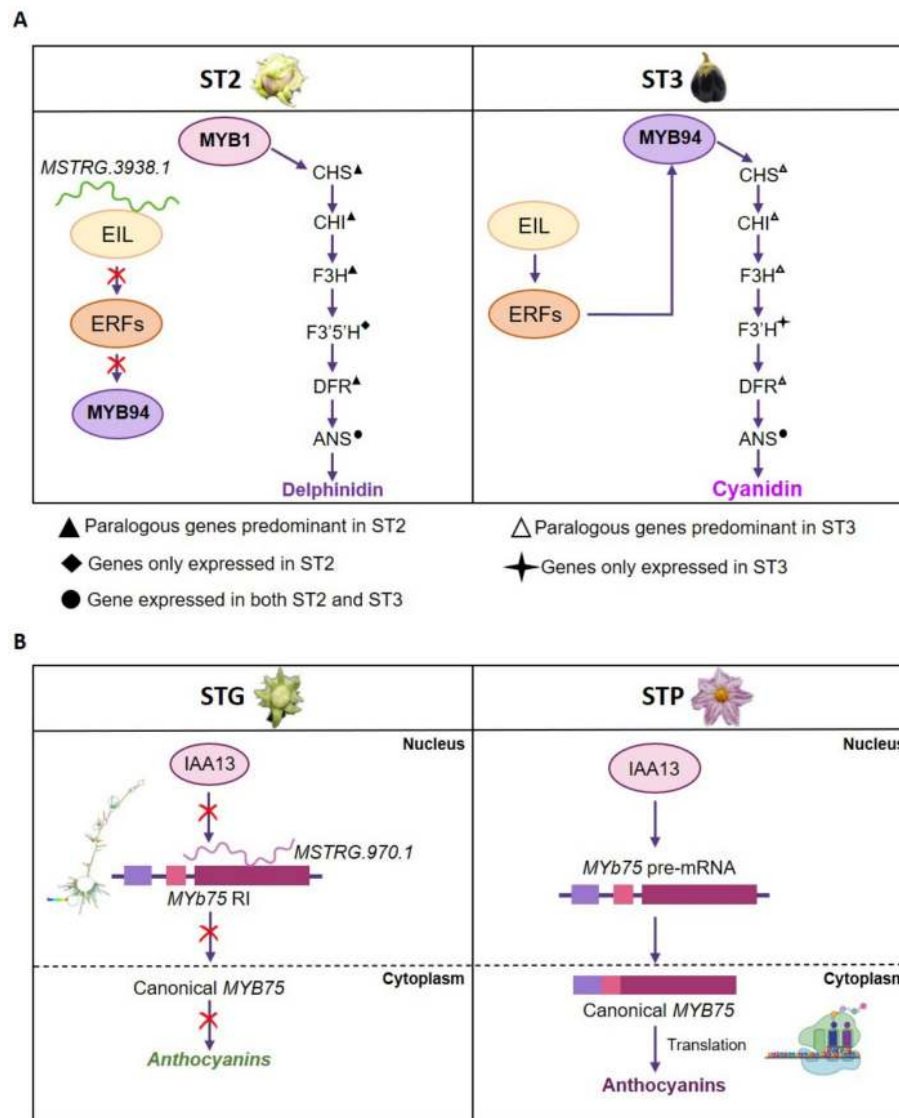


FIGURE 8

Models of anthocyanins transcriptional regulation in BB fruit peel and flower corolla. **(A)** Peel: in ST2, the lncRNA *MSTRG.3938.1* binds EIL inhibiting ethylene signaling cascade and, therefore, MYB94 expression. In this stage, MYB1 and its correlated structural paralogous genes are activated, determining delphinidin production. In ST3, the *MSTRG.3938.1* expression decreases significantly, therefore MYB94 is expressed and triggers the activation of other paralogous copies of the structural genes, which implies cyanidin production. **(B)** Corolla: in STG, when the auxin level is higher, the lncRNA *MSTRG.970.1* binds to the RI form of MYB75 by sequestering it in the nucleus. The splicing of MYB75 transcript cannot occur, therefore it is degraded or stored until a change in external conditions, preventing anthocyanins biosynthesis. Conversely, in STP, where IAA enzymes are active and the auxin level decreases, the level of *MSTRG.970.1* transcript drastically drops, freeing MYB75 transcript and allowing the splicing to occur. The canonical MYB75 determines the production of anthocyanins.

biosynthetic genes. For instance, in tomato, *SLAN2like* gene displayed a splicing mutation responsible for the loss of function of its protein in wild-type plants, leading to the lack of anthocyanin pigmentation in cultivated genotypes (Colanero et al., 2020).

Similarly, in grapevine, the exon skipping splicing event of *VvMYBA1-L* resulted in decreased accumulation of anthocyanin in grape berry flesh (Gao et al., 2023).

Interestingly, in our corolla samples we observed a differential AS event in *MYB75*, specifically a retained-intron (RI) isoform that was enriched in STP.

In plants, RI represents a major regulatory mechanism that influences TF expression by controlling the fate of their pre-mRNA. In fact, this regulation can occur through nuclear retention of RI transcripts that prevent TF translation (Boothby et al., 2013). This mechanism has been demonstrated to be controlled by the action of lncRNAs that can hybridize with the pre-mRNAs to modulate their splicing dynamics (Pisignano and Ladomery, 2021).

Notably, *MYB75* showed a negative correlation with the lncRNA *MSTRG.970.1*, whose RI isoform displayed a predicted strong interaction with, suggesting a regulatory relationship in

which *MSTRG.970.1* binds the RI isoform, hindering its maturation in STG. This interaction could lead to the degradation of IR *MYB75* transcripts via nonsense-mediated decay, or could be responsible for the storage in the nucleus of the IR *MYB75* isoform. This transcript could be converted into mature spliced *MYB75* at a specific time during development, enabling translation to proceed. In fact, in STP, the lncRNA *MSTRG.970.1* is downregulated and this could allow the proper *MYB75* splicing, and, consequently, the activation of anthocyanin biosynthesis in flowers.

Previous studies in various systems support this model (Jia et al., 2020). Zhou and collaborators demonstrated that in *Arabidopsis*, *COP1* is involved in light-regulated IR to modulate seedling photomorphogenesis. IR leads to the nuclear detainment of IR transcript as storage and, upon changes in environmental stimuli or developmental phases, they can be fully spliced and translocated into the cytoplasm for translation (Zhou et al., 2024). Moreover, in *Marsilea vestita*, a large number of intron-containing, polyadenylated pre-mRNAs have been demonstrated to be stored in the microspore and their maturation during development enabled their translation (Boothby et al., 2013).

Moreover, in apple, the regulatory effects of auxin signaling on anthocyanin metabolism have been investigated showing that exogenous auxins repress anthocyanin biosynthesis, most likely through the MdIAA121–MdARF13 signal transduction pathway (Wang et al., 2018). In our analyses, we found that *IAA13* is positively correlated with *MYB75*, and both share correlations with the same lncRNAs, suggesting a common regulatory pathway.

Taken together, these data are consistent with a model in which *MSTRG.970.1* modulates *MYB75* post-transcriptionally and possibly influences a broader hormone-linked regulatory module by the interactions with auxin repressors. This mechanism likely contributes to the switch from a *MYB75*-associated state in green stages to an active anthocyanin-production state in STP (Figure 8B).

5 Conclusion

This work confirms the tissue-specific roles of eggplant key TFs and sheds light on the involvement of the minor activator *MYB94* in regulating anthocyanin biosynthesis, supporting transcriptional data with complementary metabolite analyses. LncRNAs and AS events directly correlated with anthocyanin accumulation were identified, enabling the delineation of tissue-specific interaction models for fruit peel and flower corolla with the involvement of hormone signaling components. Altogether, these results provide a coherent systems-level view of how transcriptional and metabolic layers converge to control anthocyanin biosynthesis and regulation in BB, starting to uncover the fine-tuned regulatory mechanisms underlying this pathway. Future work will further dissect the functional interactions between the identified lncRNAs and their target TFs and experimentally confirm the regulatory roles suggested by our transcriptomic analyses.

Data availability statement

Raw sequencing data have been deposited in the NCBI Sequence Read Archive (SRA) under BioProject accession number PRJNA1374051. Genomic data, including genome assembly and annotation, have been deposited in the Zenodo repository (<https://zenodo.org/records/17790745>).

Author contributions

FD: Data curation, Formal Analysis, Validation, Methodology, Writing – original draft, Investigation, Conceptualization, Software. AA: Conceptualization, Data curation, Writing – review & editing, Software. AM: Writing – review & editing, Conceptualization. LG: Data curation, Writing – review & editing, Software. GM: Writing – original draft, Formal Analysis, Data curation, Investigation, Methodology. PG: Data curation, Supervision, Writing – review & editing. LB: Writing – review & editing, Data curation, Software. CB: Writing – review & editing, Supervision, Formal Analysis, Data curation. EP: Conceptualization, Writing – review & editing, Supervision. JP: Writing – review & editing, Data curation, Supervision. CC: Writing – review & editing, Supervision, Resources, Conceptualization.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Acknowledgments

We are grateful to Dr. Laura Bassolino (CREA-CI, Bologna) for the critical reading of the manuscript.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial

intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any

product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2026.1832029/full#supplementary-material>.

References

- Ai, J., Wang, W., Hu, T., Hu, H., Yan, Y., Wang, J., et al. (2025). Genome-wide identification and expression analysis of eggplant reveals the key MYB transcription factor involved in anthocyanin synthesis. *Horticulturae* 11, 12. doi: 10.3390/horticulturae11010012
- Albert, N. W., Lewis, D. H., Zhang, H., Schwinn, K. E., Jameson, P. E., and Davies, K. M. (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. doi: 10.1111/j.1365-313X.2010.04465.x
- Alonge, M., Lebeigle, L., Kirsche, M., Jenike, K., Ou, S., Aganezov, S., et al. (2022). Automated assembly scaffolding using RagTag elevates a new tomato system for high-throughput genome editing. *Genome Biol.* 23, 258. doi: 10.1186/s13059-022-02823-7
- An, J.-P., Zhang, X.-W., Bi, S.-Q., You, C.-X., Wang, X.-F., and Hao, Y.-J. (2020). The ERF transcription factor MdERF38 promotes drought stress-induced anthocyanin biosynthesis in apple. *Plant J.* 101, 573–589. doi: 10.1111/tpj.14555
- Azuma, M., Morimoto, R., Hirose, M., Morita, Y., Hoshino, A., Iida, S., et al. (2015). A petal-specific InMYB1 promoter from Japanese morning glory: a useful tool for molecular breeding of floricultural crops. *Plant Biotechnol. J.* 14, 354–363. doi: 10.1111/pbi.12389
- Barchi, L., Lanteri, S., Portis, E., Valè, G., Volante, A., Pulcini, L., et al. (2012). A RAD tag derived marker based eggplant linkage map and the location of QTLs determining anthocyanin pigmentation. *PLoS One* 7, e43740. doi: 10.1371/journal.pone.0043740
- Boothby, T. C., Zipper, R. S., van der Wee, C. M., and Wolniak, S. M. (2013). Removal of retained introns regulates translation in the rapidly developing gametophyte of *Marsilea vestita*. *Dev. Cell* 24, 517–525. doi: 10.1016/j.devcel.2013.01.015
- Buchfink, B., Reuter, K., and Drost, H.-G. (2021). Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18, 366–368. doi: 10.1038/s41592-021-01101-x
- Cammareri, M., Frary, A., Frary, A., and Grandillo, S. (2024). Genetic and biotechnological approaches to improve fruit bioactive content: a focus on eggplant and tomato anthocyanins. *Int. J. Mol. Sci.* 25, 6811. doi: 10.3390/ijms25126811
- Cappellini, F., Marinelli, A., Toccaceli, M., Tonelli, C., and Petroni, K. (2021). Anthocyanins: From mechanisms of regulation in plants to health benefits in foods. *Front. Plant Sci.* 12, 748049. doi: 10.3389/fpls.2021.748049
- Chen, S. (2023). Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta* 2, e107. doi: 10.1002/imt2.107
- Cheng, H., Asri, M., Lucas, J., Koren, S., and Li, H. (2024). Scalable telomere-to-telomere assembly for diploid and polyploid genomes with double graph. *Nat. Methods* 21, 967–970. doi: 10.1038/s41592-024-02269-8
- Chialva, C., Blein, T., Crespi, M., and Lijavetzky, D. (2021). Insights into long non-coding RNA regulation of anthocyanin carrot root pigmentation. *Sci. Rep.* 11, 4093. doi: 10.1038/s41598-021-83514-4
- Colanero, S., Perata, P., and Gonzali, S. (2018). The atrovioacea gene encodes an R3-MYB protein repressing anthocyanin synthesis in tomato plants. *Front. Plant Sci.* 9, 9. doi: 10.3389/fpls.2018.00830
- Colanero, S., Tagliani, A., Perata, P., and Gonzali, S. (2020). Alternative splicing in the Anthocyanin Fruit gene encoding an R2R3 MYB transcription factor affects anthocyanin biosynthesis in tomato fruits. *Plant Commun.* 1, 100006. doi: 10.1016/j.xplc.2019.100006
- Deng, X., Bashandy, H., Ainasoja, M., Kontturi, J., Pietiäinen, M., Laitinen, R. A. E., et al. (2014). Functional diversification of duplicated chalcone synthase genes in anthocyanin biosynthesis of *Gerbera hybrida*. *New Phytol.* 201, 1469–1483. doi: 10.1111/nph.12610
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635
- Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340
- Fan, L., Shi, J., Zuo, J., Gao, L., Lv, J., and Wang, Q. (2016). Methyl jasmonate delays postharvest ripening and senescence in the non-climacteric eggplant (*Solanum melongena* L.) fruit. *Postharvest Biol. Technol.* 120, 76–83. doi: 10.1016/j.postharvbio.2016.05.010
- FAO (2025). FAOSTAT: food and agriculture organization of the united nations. Available online at: <https://www.fao.org/home/en/> (Accessed September 20, 2025).
- Feng, J., Wang, M. Y., Chen, X., Tomes, S., Tao, J., Atkinson, R. G., et al. (2025). EIL (ethylene-insensitive 3-like) transcription factors in apple affect both ethylene- and cold response-dependent fruit ripening. *Plant J.* 121, e70059. doi: 10.1111/tpj.70059
- Filyushin, M. A., Dzhos, E. A., Shchennikova, A. V., and Kochieva, E. Z. (2024). Metabolite concentrations and the expression profiles of the corresponding metabolic pathway genes in eggplant (*Solanum melongena* L.) fruits of contrasting colors. *Vavilov J. Genet. Breed.* 28, 619–627. doi: 10.18699/vjgb-24-69
- Filyushin, M. A., Shchennikova, A. V., and Kochieva, E. Z. (2023). Coexpression of structural and regulatory genes of the flavonoid pathway reveals the characteristics of anthocyanin biosynthesis in eggplant organs (*Solanum melongena* L.). *Russ. J. Plant Physiol.* 70, 27. doi: 10.1134/S1021443722603147
- Gabriel, L., Bruna, T., Hoff, K. J., Ebel, M., Lomsadze, A., Borodovsky, M., et al. (2024). BRAKER3: Fully automated genome annotation using RNA-seq and protein evidence with GeneMark-ETP, AUGUSTUS and TSEBRA. *bioRxiv*, 2023.06.10.544449. doi: 10.1101/2023.06.10.544449
- Gaccione, L., Martina, M., Barchi, L., and Portis, E. (2023). A compendium for novel marker-based breeding strategies in eggplant. *Plants* 12, 1016. doi: 10.3390/plants12051016
- Gaccione, L., Toppino, L., Bolger, M., Schmidt, M., Tassone, M. R., Sulli, M., et al. (2025). Graph-based pangenomes and pan-phenome provide a cornerstone for eggplant biology and breeding. *Nat. Commun.* 16, 9919. doi: 10.1038/s41467-025-64866-1
- Gao, L., Wang, W., Li, H., Li, H., Yang, Y., Zheng, H., et al. (2023). Anthocyanin accumulation in grape berry flesh is associated with an alternative splicing variant of VvMYBA1. *Plant Physiol. Biochem.* 195, 1–13. doi: 10.1016/j.plaphy.2022.12.025
- Gu, C., Guo, Z.-H., Hao, P.-P., Wang, G.-M., Jin, Z.-M., and Zhang, S.-L. (2017). Multiple regulatory roles of AP2/ERF transcription factor in angiosperm. *Bot. Stud.* 58, 6. doi: 10.1186/s40529-016-0159-1
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Heo, J. B., Lee, Y.-S., and Sung, S. (2013). Epigenetic regulation by long noncoding RNAs in plants. *Chromosome Research: An. Int. J. Molecular Supramolecular Evolutionary Aspects. Chromosome Biol.* 21, 685–693. doi: 10.1007/s10577-013-9392-6
- Holst, F., Bolger, A., Günther, C., Maß, J., Triesch, S., Kindel, F., et al. (2023). Helixer—de novo prediction of primary eukaryotic gene models combining deep learning and a hidden Markov model. *bioRxiv*, 2023.02.06.527280. doi: 10.1101/2023.02.06.527280
- Jaakola, L. (2013). New insights into the regulation of anthocyanin biosynthesis in fruits. *Trends Plant Sci.* 18, 477–483. doi: 10.1016/j.tplants.2013.06.003
- Jia, J., Long, Y., Zhang, H., Li, Z., Liu, Z., Zhao, Y., et al. (2020). Post-transcriptional splicing of nascent RNA contributes to widespread intron retention in plants. *Nat. Plants* 6, 780–788. doi: 10.1038/s41477-020-0688-1
- Kang, Y.-J., Yang, D.-C., Kong, L., Hou, M., Meng, Y.-Q., Wei, L., et al. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Res.* 45, W12–W16. doi: 10.1093/nar/gkx428

- Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: An information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109
- Law, C. W., Alhamdoosh, M., Su, S., Dong, X., Tian, L., Smyth, G. K., et al. (2018). RNA-seq analysis is easy as 1–2–3 with limma, Glimma and edgeR. *F1000Research* 5, ISCB Comm J–1408. doi: 10.12688/f1000research.9005.3
- Li, H. (2021). New strategies to improve minimap2 alignment accuracy. *Bioinformatics* 37, 4572–4574. doi: 10.1093/bioinformatics/btab705
- Li, R., Cui, L., Martina, M., Bracuto, V., Meijer-Dekens, F., Wolters, A.-M., et al. (2024c). Less is more: CRISPR/Cas9-based mutations in DND1 gene enhance tomato resistance to powdery mildew with low fitness costs. *BMC Plant Biol.* 24, 763. doi: 10.1186/s12870-024-05428-3
- Li, Q., Duncan, S., Li, Y., Huang, S., and Luo, M. (2024b). Decoding plant specialized metabolism: new mechanistic insights. *Trends Plant Sci.* 29, 535–545. doi: 10.1016/j.tplants.2023.11.015
- Li, C., Lai, X., Yu, X., Xiong, Z., Chen, J., Lang, X., et al. (2023a). Plant long noncoding RNAs: Recent progress in understanding their roles in growth, development, and stress responses. *Biochem. Biophys. Res. Commun.* 671, 270–277. doi: 10.1016/j.bbrc.2023.05.103
- Li, B.-J., Shi, Y.-N., Xiao, Y.-N., Jia, H.-R., Yang, X.-F., Dai, Z.-R., et al. (2024a). AUXIN RESPONSE FACTOR 2 mediates repression of strawberry receptacle ripening via auxin-ABA interplay. *Plant Physiol.* 196, 2638–2653. doi: 10.1093/plphys/kiad510
- Li, H., Wang, W., Liu, R., Tong, B., Dai, X., Lu, Y., et al. (2023b). Long non-coding RNA-mediated competing endogenous RNA regulatory network during flower development and color formation in *Melastoma candidum*. *Front. Plant Sci.* 14, 1215044. doi: 10.3389/fpls.2023.1215044
- Li, D., Zhu, M., Li, S., Wang, P., Li, J., Ge, H., et al. (2025). Eggplant transcription factor SmERF118 mediates light signaling regulation of anthocyanin synthesis. *Plant Physiol. Biochem.* 221, 109584. doi: 10.1016/j.plaphy.2025.109584
- Lin, R.-C., and Rausher, M. D. (2021). R2R3-MYB genes control petal pigmentation patterning in *Clarkia gracilis* ssp. *sonomensis* (Onagraceae). *New Phytol.* 229, 1147–1162. doi: 10.1111/nph.16908
- Liu, S., Lyu, S., Zhang, Y., Liu, S., and Deng, S. (2025). Tomato CONSTANS-Like1 promotes anthocyanin biosynthesis under short day and suboptimal low temperature. *Plant Physiol.* 198, kiaf190. doi: 10.1093/plphys/kiad190
- Lorenz, R., Bernhart, S. H., Höner zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P. F., et al. (2011). ViennaRNA package 2.0. *Algorithms Mol. Biol.* 6, 26. doi: 10.1186/1748-7188-6-26
- Lo Scalzo, R., Florio, F. E., Fibiani, M., Speranza, G., Rabuffetti, M., Gattolin, S., et al. (2021). Scrapped but not neglected: Insights into the composition, molecular modulation and antioxidant capacity of phenols in peel of eggplant (*Solanum melongena* L.) fruits at different developmental stages. *Plant Physiol. Biochem.* 167, 678–690. doi: 10.1016/j.plaphy.2021.08.037
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8
- Lovell, J. T., Sreedasyam, A., Schranz, M. E., Wilson, M., Carlson, J. W., Harkess, A., et al. (2022). GENESPACE tracks regions of interest and gene copy number variation across multiple genomes. *eLife* 11, e78526. doi: 10.7554/eLife.78526
- Mandadi, K. K., Petrillo, E., Dubrovina, A. S., and Kiselev, K. V. (2023). Editorial: Regulation of alternative splicing in plant stress responses. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1120961
- Manickam, G. P., L. A., V. A., S. D., V., and K. K. K. (2025). Anthocyanin pathway in eggplant: genetic regulation and future directions for metabolic engineering. *Mol. Biol. Rep.* 52, 522. doi: 10.1007/s11033-025-10597-x
- Mann, M., Wright, P. R., and Backofen, R. (2017). IntaRNA 2.0: enhanced and customizable prediction of RNA–RNA interactions. *Nucleic Acids Res.* 45, W435–W439. doi: 10.1093/nar/gkx279
- Manni, M., Berkeley, M. R., Seppey, M., and Zdobnov, E. M. (2021). BUSCO: assessing genomic data quality and beyond. *Curr. Protocol.* 1, e323. doi: 10.1002/cpz1.323
- Mannino, G., and Maffei, M. E. (2022). Metabolomics-based profiling, antioxidant power, and uropathogenic bacterial anti-adhesion activity of SP4TM, a formulation with a high content of type-A proanthocyanidins. *Antioxidants* 11, 1234. doi: 10.3390/antiox11071234
- Mannino, G., Ricciardi, M., Gatti, N., Serio, G., Vigliante, I., Contartese, V., et al. (2022). Changes in the phytochemical profile and antioxidant properties of *Prunus persica* fruits after the application of a commercial biostimulant based on seaweed and yeast extract. *Int. J. Mol. Sci.* 23, 15911. doi: 10.3390/ijms232415911
- Martina, M., Acquadro, A., Portis, E., Barchi, L., and Lanteri, S. (2023). Diversity analyses in two ornamental and large-genome Ranunculaceae species based on a low-cost Klenow NGS-based protocol. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1187205
- Martina, M., De Rosa, V., Magon, G., Acquadro, A., Barchi, L., Barcaccia, G., et al. (2024). Revitalizing agriculture: next-generation genotyping and -omics technologies enabling molecular prediction of resilient traits in the Solanaceae family. *Front. Plant Sci.* 15. doi: 10.3389/fpls.2024.1278760
- Martins, T. R., Berg, J. J., Blinka, S., Rausher, M. D., and Baum, D. A. (2013). Precise spatio-temporal regulation of the anthocyanin biosynthetic pathway leads to petal spot formation in *Clarkia gracilis* (Onagraceae). *New Phytol.* 197, 958–969. doi: 10.1111/nph.12062
- Mattick, J. S., Amaral, P. P., Carninci, P., Carpenter, S., Chang, H. Y., Chen, L.-L., et al. (2023). Long non-coding RNAs: definitions, functions, challenges and recommendations. *Nat. Rev. Mol. Cell Biol.* 24, 430–447. doi: 10.1038/s41580-022-00566-8
- Mattioli, R., Francioso, A., Mosca, L., and Silva, P. (2020). Anthocyanins: a comprehensive review of their chemical properties and health effects on cardiovascular and neurodegenerative diseases. *Molecules* 25, 3809. doi: 10.3390/molecules25173809
- Moglia, A., Florio, F. E., Iacopino, S., Guerrieri, A., Milani, A. M., Comino, C., et al. (2020). Identification of a new R3 MYB type repressor and functional characterization of the members of the MBW transcriptional complex involved in anthocyanin biosynthesis in eggplant (*S. melongena* L.). *PLoS One* 15 (5), e0232986. doi: 10.1371/journal.pone.0232986
- Noda, Y., Kneyuki, T., Igarashi, K., Mori, A., and Packer, L. (2000). Antioxidant activity of nasunin, an anthocyanin in eggplant peels. *Toxicology* 148, 119–123. doi: 10.1016/S0300-483X(00)00202-X
- Nutricati, E., Sabella, E., Negro, C., Min Allah, S., Luvisi, A., De Bellis, L., et al. (2025). Anthocyanins and anthocyanin biosynthesis gene expression in *Passiflora* flower corona filaments. *Plants* 14, 1050. doi: 10.3390/plants14071050
- Ou, S., Su, W., Liao, Y., Chougule, K., Agda, J. R. A., Hellings, A. J., et al. (2019). Benchmarking transposable element annotation methods for creation of a streamlined, comprehensive pipeline. *Genome Biol.* 20, 275. doi: 10.1186/s13059-019-1905-y
- Panda, S., Chappell-Maor, L., Alejandro de Haro, L., Jozwiak, A., Gharat, S. A., Kazachkova, Y., et al. (2025). Molecular mechanisms driving the unusual pigmentation shift during eggplant fruit development. *Plant Commun.* 6, 101321. doi: 10.1016/j.xplc.2025.101321
- Paysan-Lafosse, T., Andreeva, A., Blum, M., Chuguransky, S. R., Grego, T., Pinto, B. L., et al. (2025). The Pfam protein families database: embracing AI/ML. *Nucleic Acids Res.* 53, D523–D534. doi: 10.1093/nar/gkac997
- Pertea, G., and Pertea, M. (2020). GFF utilities: gffRead and gffcompare. *F1000Research*. doi: 10.12688/f1000research.23297.1
- Posignano, G., and Ladomery, M. (2021). Epigenetic regulation of alternative splicing: how lncRNAs tailor the message. *Non-Coding RNA* 7, 21. doi: 10.3390/ncrna7010021
- Putta, S., Yarla, N. S., Peluso, I., Tiwari, D. K., Reddy, G. V., Giri, P. V., et al. (2017). Anthocyanins: multi-target agents for prevention and therapy of chronic diseases. *Curr. Pharm. Des.* 23, 6321–6346. doi: 10.2174/1381612823666170519151801
- Rhie, A., Walenz, B. P., Koren, S., and Phillippy, A. M. (2020). Merqure: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* 21, 245. doi: 10.1186/s13059-020-02134-9
- Schwinn, K., Venail, J., Shang, Y., Mackay, S., Alm, V., Butelli, E., et al. (2006). A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell* 18, 831–851. doi: 10.1105/tpc.105.039255
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303
- Shi, S., Liu, Y., He, Y., Li, L., Li, D., and Chen, H. (2021). R2R3-MYB transcription factor SmMYB75 promotes anthocyanin biosynthesis in eggplant (*Solanum melongena* L.). *Sci. Hort.* 282, 110020. doi: 10.1016/j.scienta.2021.110020
- Shoeva, O., Glagoleva, A., and Khlestkina, E. K. (2017). The factors affecting the evolution of the anthocyanin biosynthesis pathway genes in monocot and dicot plant species. *BMC Plant Biol.* 17, 256. doi: 10.1186/s12870-017-1190-4
- Shumate, A., Wong, B., Pertea, G., and Pertea, M. (2022). Improved transcriptome assembly using a hybrid of long and short reads with StringTie. *PLoS Comput. Biol.* 18, e1009730. doi: 10.1371/journal.pcbi.1009730
- Sibley, C. R., Blazquez, L., and Ule, J. (2016). Lessons from non-canonical splicing. *Nat. Rev. Genet.* 17, 407–421. doi: 10.1038/nrg.2016.46
- Sunil, L., and Shetty, N. P. (2022). Biosynthesis and regulation of anthocyanin pathway genes. *Appl. Microbiol. Biotechnol.* 106, 1783–1798. doi: 10.1007/s00253-022-11835-z
- Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J. W. S. (2012). Alternative splicing in plants – coming of age. *Trends Plant Sci.* 17, 616–623. doi: 10.1016/j.tplants.2012.06.001
- Szakonyi, D., and Duque, P. (2018). Alternative splicing as a regulator of early plant development. *Front. Plant Sci.* 9, 1174. doi: 10.3389/fpls.2018.01174
- Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120
- Tanaka, Y., Sasaki, N., and Ohmiya, A. (2008). Biosynthesis of plant pigments: anthocyanins, betalains and carotenoids. *Plant J.* 54, 733–749. doi: 10.1111/j.1365-3113.2008.03447.x
- Toppino, L., Barchi, L., Mercati, F., Acciarri, N., Perrone, D., Martina, M., et al. (2020). A new intra-specific and high-resolution genetic map of eggplant based on a RIL

- population, and location of QTLs related to plant anthocyanin pigmentation and seed vigour. *Genes* 11, 745. doi: 10.3390/genes11070745
- Trincado, J. L., Entizne, J. C., Hysenaj, G., Singh, B., Skalic, M., Elliott, D. J., et al. (2018). SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome Biol.* 19, 40. doi: 10.1186/s13059-018-1417-1
- Wang, X., Chen, X., Luo, S., Ma, W., Li, N., Zhang, W., et al. (2022a). Discovery of a DFR gene that controls anthocyanin accumulation in the spiny *Solanum* group: roles of a natural promoter variant and alternative splicing. *Plant J.* 111, 1096–1109. doi: 10.1111/tpj.15877
- Wang, X., Fan, H., Wang, B., and Yuan, F. (2023). Research progress on the roles of lncRNAs in plant development and stress responses. *Front. Plant Sci.* 14, 1138901. doi: 10.3389/fpls.2023.1138901
- Wang, S., Li, L.-X., Zhang, Z., Fang, Y., Li, D., Chen, X.-S., et al. (2022b). Ethylene precisely regulates anthocyanin synthesis in apple via a module comprising MdEIL1, MdMYB1, and MdMYB17. *Hortic. Res.* 9, uhac034. doi: 10.1093/hr/uhac034
- Wang, L., Park, H. J., Dasari, S., Wang, S., Kocher, J.-P., and Li, W. (2013). CPAT: coding-potential assessment tool using an alignment-free logistic regression model. *Nucleic Acids Res.* 41, e74. doi: 10.1093/nar/gkt006
- Wang, Y., Wang, N., Xu, H., Jiang, S., Fang, H., Su, M., et al. (2018). Auxin regulates anthocyanin biosynthesis through the Aux/IAA–ARF signaling pathway in apple. *Hortic. Res.* 5, 59. doi: 10.1038/s41438-018-0068-4
- Wucher, V., Legeai, F., Hédan, B., Rizk, G., Lagoutte, L., Leeb, T., et al. (2017). FEELnc: a tool for long non-coding RNA annotation and its application to the dog transcriptome. *Nucleic Acids Res.* 45, e57. doi: 10.1093/nar/gkw1306
- Yan, X., Bai, D., Song, H., Lin, K., and Pang, E. (2021). Alternative splicing during fruit development among fleshy fruits. *BMC Genomics* 22, 762. doi: 10.1186/s12864-021-08111-1
- Yang, G., Li, L., Wei, M., Li, J., and Yang, F. (2022). SmMYB113 is a key transcription factor responsible for compositional variation of anthocyanin and color diversity among eggplant peels. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.843996
- Yu, J., Qiu, K., Sun, W., Yang, T., Wu, T., Song, T., et al. (2022). A long noncoding RNA functions in high-light-induced anthocyanin accumulation in apple by activating ethylene synthesis. *Plant Physiol.* 189, 66–83. doi: 10.1093/plphys/kiac049
- Zhou, H., Zeng, H., Yan, T., Chen, S., Fu, Y., Qin, G., et al. (2024). Light regulates nuclear detainment of intron-retained transcripts through COP1-spliceosome to modulate photomorphogenesis. *Nat. Commun.* 15, 5130. doi: 10.1038/s41467-024-49571-9