Transcriptome Analysis of MYB Genes and Patterns of Anthocyanin Accumulation During Seed Development in Wheat

Paulina Calderon Flores1, Jin Seok Yoon2, Dae Yeon Kim3 and Yong Weon Seo1*

1Department of Plant Biotechnology, Korea University, Seoul, Korea. 2Ojeong Plant Breeding Research Center, Korea University, Seoul, Korea. 3Department of Biotechnology, Korea University, Seoul, Korea.

ABSTRACT: Plants accumulate key metabolites as a response of biotic/abiotic stress conditions. In seed coats, anthocyanins, carotenoids, and chlorophylls can be found. They have been associated as important antioxidants that affect germination. In wheat, anthocyanins can impart the seed coat color which have been recognized as health-promoting nutrients. Transcription factors act as master regulators of cellular processes. Transcription complexes such as MYB-bHLH-WD40 (MBW) regulate the expression of multiple target genes in various plant species. In this study, the spatiotemporal accumulation of seed coat pigments in different developmental stages (10, 20, 30, and 40 days after pollination) was analyzed using cryo-cuts. Moreover, the accumulation of phenolic, anthocyanin, and chlorophyll contents was quantified, and the expression of flavonoid biosynthetic genes was evaluated. Finally, transcriptome analysis was performed to analyze putative MYB genes related to seed color followed by further characterization of putative genes. TaTCL2, an MYB gene, was cloned and sequenced. It was determined that TaTCL2 contains a SANT domain, which is often present in proteins participating in the response to anthocyanin accumulation. Moreover, TaTCL2 transcript levels were shown to be influenced by anthocyanin accumulation during grain development. Interaction network analysis showed interactions with GL2 (HD-ZIP IV), EGL3 (bHLH), and TTG1 (WD40). The findings of this study elucidate the mechanisms underlying color formation in *Triticum aestivum* L. seed coats.

KEYWORDS: Colored wheat, MBW complex, MYB, purple wheat seed, transcription factors

Introduction

Domestication has made, through the selection, to narrow the genetic base and the variance for traits that are important for continued crop improvement.1 Highlighting the breeding lines, “lost” traits and trait combinations that once were discarded by traditional breeding programs, can find unidentified traits that can, nowadays, be favorable for crop performance.1-3

Domestication of cereals made changes in cereal metabolites, such as anthocyanins. In wheat, barley, maize, and rice, most of the genotypes that can be found, have lost the capability of synthesizing anthocyanins in grain. However, due to the aim to improve health, the study of the anthocyanins in grains have been associated as important antioxidants that affect germination. In wheat, anthocyanins can impart the colorless PAs that accumulate exclusively in the pericarp to the seed coat and in the embryo. Colorless PAs that accumulate exclusively in the inner integument, are oxidized, and polymerized into brown pigments during seed maturation, causing the formation of brown seed coat color.19-22

The flavonoid pathway has been studied elsewhere.22-25 The genes are divided in early anthocyanin biosynthesis genes (EBGs), and late anthocyanin biosynthesis genes (LBGs). EBGs include *chalcone synthase* (*CHS*), *chalcone isomerase* (*CHI*),
flavonone 3-hydroxylase (F3H), flavanone 3′-hydroxylase (F3’H), and flavonol synthase (FLS) which lead to the production of flavonols, whereas LBGs include dihydroflavonol reductase (DFR), anthocyanidin synthase/leucoanthocyanidin dioxygenase (ANS/LOX), UDP-flavonoid glucosyl transferase (UGFT), and anthocyanidin reductase (ANR) which lead to the production of proanthocyanidins and anthocyanins.22,26

Generally, the LBGs for anthocyanin biosynthesis is regulated by a MYB-bHLH-WD40 (MBW) complex.27 The MBW complex has been reported to regulate the expression of flavonoid biosynthesis genes.28 MYBs PAP1 and PAP2 in Arabidopsis,29 ZmCI in maize,26 and MYB3 in petunias30 interact with bHLHs such as GL3, EGL3, and TT829 to form a ternary complex with WD40, such as TTG1 in Arabidopsis, thus affecting proanthocyanin accumulation in seeds and anthocyanin accumulation in plant species.28,29,31

In this study, we analyzed recombinant inbred lines (RILs) of wheat to get a deeper analysis of the pigmentation accumulation in seed coats of deep purple (DP) and yellow (Ye) seeds. Phenotyping and RNA sequencing was used to determine the transcriptome difference between DP and Ye wheat seeds and isolate a dominant gene for the deep purple grain trait. The findings of this study elucidate the mechanisms underlying color formation in Triticum aestivum L. seed coats.

Materials and Methods

Plant materials and growth conditions

RILs with different seed coat phenotypes, Ye (accession no. 10DS1673, Korea University wheat sub–gene bank) and DP (accession no. 10DS1674), were used.32 Seeds were germinated on moistened filter paper at room temperature for 24 hours and vernalized at 4°C in a dark chamber for 4 weeks. Each seedling was then transplanted to a pot (5 × 5 × 16 cm) filled with soil (Sunshine mix #1, Sun Gro Horticulture, Vancouver, BC, Canada) in a well-controlled glasshouse at Korea University with a photoperiod of 16:8 hours (day:night) and temperatures between 20°C and 25°C. Spikes were harvested at the flowering stage at 10, 20, 30, and 40 days after pollination (DAP) in green house. Seeds of the same RILs were planted in the field at Advanced Radiation Technology Institute, Korea Atomic Energy Research Institute (Jeongeup, Jeollabuk-do, Republic of Korea), and the spikes were harvested at the flowering stage at 10, 20, 30, and 40 DAP. Samples were harvested in triplicate biological replicates.

Whole caryopses and cryo-cuts

To observe anthocyanin deposition during seed development in Ye and DP seeds, cryo-cuts (Cryotome CM3050S, Leica, Wetzlar, Germany) were used for sectioning the seeds of the harvested Ye and DP at 10, 20, 30, and 40 DAP. The seeds were embedded in optimum cutting temperature compound (Sakura® FineTek USA, Inc., Torrance, CA, USA). Cross-sections (0.04 mm) were prepared from the central part of the developing caryopses and observed with a Leica EZ4D microscope (Leica). Photographs were taken using Leica Acquire software (version 3.4.1).

Quantification of phenolics and anthocyanins

Free phenolics (FP) and total phenolics (TP) contents were measured using Folin–Ciocalteu reagent using 20 mg of grain sample ground in liquid nitrogen.33 Anthocyanin content was determined according to Hong et al34 using 100 mg from each sample. The absorbance was read at 765 nm for FP and TP and at 530 and 657 nm for anthocyanins using a microplate reader (HIDEX-Sense 425-301, Hidex, Turku, Finland). The results are expressed as the average of 3 biological replicates of phenolics and anthocyanins.

Chlorophyll content

Chlorophyll content was measured using 20 mg of ground grains in liquid nitrogen following the protocol of Warren35 with 3 biological replicates. The absorbance was measured at 652 and 665 nm for chlorophyll content using a microplate reader (HIDEX-Sense 425-301).

Gene expression analysis

Total RNA was extracted in triplicate from samples taken during the seed development stages described in the sampling section. Seed samples were immediately frozen in liquid nitrogen and stored at −80°C until further use. Samples were ground in liquid nitrogen using a pestle and mortar, and RNA was extracted using the TRizol method (Invitrogen, Waltham, MA, USA). cDNA was synthesized using the cDNA Takara PrimeScript™ 1st strand cDNA Synthesis Kit (Takara, Tokyo, Japan). Quantitative real-time PCR (qRT-PCR) was performed to analyze the flavonoid pathway genes during seed development of Ye and DP seeds. Gene-specific primers were designed using Primer-BLAST (NCBI, https://www.ncbi.nlm.nih.gov/tools/primer-blast/) or obtained from previous publications as described in Supplemental Table S1. qRT-PCR was performed in triplicate using EvaGreen 2X qPCR MasterMix (ABM, Milton, ON, Canada) on a CFX-96 RT-PCR machine (Bio-Rad, Hercules, CA, USA). β-actin (Accession no. AB181991) was used as an internal control. The 2−∆∆CT method was used to calculate expression levels in fold changes as previously described.36

cDNA library construction and RNA sequencing

Total RNA was extracted from whole grains using TRizol reagent (Invitrogen, Waltham, MA, USA) and then used for
library construction with the Ri-bospinTMII Kit (Geneall Biotechnology, Seoul, Korea) according to the manufacturer’s instructions. A total of 9 RNA-seq paired-end libraries at 3 developmental stages (10, 20, and 30 DAP, Supplemental Figure S1) in DP and Ye seeds were constructed from each DAP sample consisting of 3 biological replicates each, using the SMARTer Stranded RNA-Seq Kit (Clontech Laboratories Inc., Mountain View, CA, USA) following the manufacturer’s instructions. Each library was then loaded onto the Illumina Hiseq2000 platform, high-throughput sequencing was performed, and paired-end reads were generated. The high-quality reads were obtained after several steps of quality checks, including trimming, removal of adaptor/primer, and low-quality reads using Trimmomatic v 0.35, and extraction of sequence data with quality scores of $Q \geq 20$ using SolexaQA.

**Read mapping and annotation**

The hexaploid common wheat (*Triticum aestivum* L.) draft reference genome sequence (IWGSC1 + popseq.31) was directly downloaded from EnsemblPlants (http://plants.ensembl.org/Triticum_aestivum/Info/Index). To map our reads to the draft genome sequence, an index of the reference genome was constructed using hisat2 (https://ccb.jhu.edu/software/hisat2/index.shtml), and paired-end clean reads

---

**Figure 1.** Images during grain development of the yellow (Ye; sample 1673) and deep purple (DP; sample 1674) parental seeds: (a) frontal and posterior sides at 10, 20, 30, and 40 days after pollination (DAP) in Ye and DP seeds and (b) cross-sections during grain development of Ye and DP parental seeds at 10 and 20DAP.
were aligned to the reference genome, again using hsiast2, with all parameters set to their default values. HTSeq v0.6.1 was used to count the reads mapped to each gene and to the exons of each gene. Differential expression was analyzed for each treatment group using DESeq2, and genes with an adjusted absolute log2 fold change of <1 were delineated as differentially expressed genes (DEGs). DEGs were annotated by blasting (blastx) against the NCBI non-redundant protein database for the Poaceae family and against the UniProt protein sequence databases for rice, Arabidopsis, and Brachypodium (http://www.uniprot.org), with an E-value cutoff of 10^-4.

Selection of putative genes

For the flavonoid biosynthetic pathway, using the Arabidopsis_unipro annotations, we selected for DEGs related to the flavonoid pathway such as CHS (chalcone synthase), CHI (chalcone isomerase), F3H (flavonol 3-hydroxylase), F3′H (flavonol 3′-hydroxylase), FLS (flavonol synthase), OMT1 (O-methyltransferase 1), UGT (UDP-glycosyltransferase), DFR (dihydroflavonol reductase), LDOX (leucoanthocyanidin dioxygenase), ANR (anthocyanidin reductase), OMT (O-methyltransferase), and GT (glycosyltransferase). Similarly, we selected the data to identify MYB TFs expressed during the grain developmental stages using Poaceae_NR, Rice_unipro, Brachypodium_unipro, and Arabidopsis_unipro databases. Heat maps were created using Gene-E (https://software.broad-institute.org/GENE-E/). Highly expressed transcripts at 20 DAP DP seeds were selected for qRT-PCR analysis, as shown in Supplemental Table S2.

For the isolation of MYB putative genes, highly expressed putative genes specifically at 20 DAP in DP seeds were selected from transcriptome analysis (Supplemental Table S2). Primers specific to these genes were designed after referencing the sequences in the Ensembl Plants database (https://plants.ensembl.org/Triticum_aestivum/Info/Index) or Plant Transcription Factor Database v5.0 (http://planttfdb.gao-lab.org). The specific primers are shown in Supplemental Table S1.

Cloning and sequencing of putative MYB genes

The PCR fragments of genes were purified and cloned into pLUG-T Prime® TA-Cloning Vector (Intron Biotechnology, Seongnam, Korea). The vector and PCR-amplified product were mixed and ligated overnight at 4°C and transformed into Escherichia coli DH5α competent cells (YB Biotech, Taipei, Taiwan) using the manufacturer’s instructions. The pLUG plasmids were sequenced from both ends at Bionics (Seoul, South Korea). The sequences were compared with genes in the GenBank database using the BLAST program from the National Center for Biotechnology Information (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

Evolutionary and bioinformatics analysis

Sequences from the cloned genes were used in Expasy software (https://web.expasy.org/translate/) to translate the coding sequence into amino acids. For the cluster analysis MYB genes, the amino acid sequences of Traes_3DL_30CF35BB3 were analyzed with BLASTP from NCBI. Among the putative MYB genes, MYB genes of monocot plants were selected for phylogenetic analysis. The phylogenetic tree was constructed with fast minimum evolution method.

The conserved domains were analyzed with the Conserved Domains tool from NCBI (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi). A specific interaction network was constructed using STRING 11.0b (https://string-db.org), as seen in Zhao et al. with an option value >0.700, and both alignment and phylogenetic tree construction were performed in Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/); accession numbers used are shown in Supplemental Table S3.

Statistic analysis and software

All data are presented as means from at least 3 biological replicates. Significant differences were subjected to analysis of variance and t-test using IBM® SPSS® Statistics for MAC version 25 (IBM Corp., Armonk, NY, USA). All tests were performed at 95%, 99%, and 99.9% confidence.

Results

Whole caryopses and cryo-cuts color deposition

At 10 DAP, Ye and DP seeds were similar in color (Figure 1), indicating that chlorophyll had fully covered both seeds at this stage. Similarly, in the cryo-cuts at 10 DAP, Ye and DP seeds were still in an early stage of development, and no color deposition was observed in either type. At 20 DAP, Ye seeds exhibited little variation, whereas in DP seeds, a purple color had spread gradually almost throughout the seed coat from the pericarp to the aleurone layer. At 20 DAP in Ye seeds, the cryo-cuts showed 2 layers with visible chlorophyll in the seed coat, whereas in DP, the seed coat began to develop a purple color on the top seed coat layer (pericarp) and inside (aleurone layer), and chlorophyll remained under the inner layer (aleurone) or, in some areas, completely disappeared. The seed coat of Ye and DP seeds clearly appeared yellow and purple at 30 and 40 DAP, respectively.

Phenolic, anthocyanin, and chlorophyll quantification

FP and TP contents were quantified during seed development. In both Ye and DP seeds, FP peaked at 20 DAP. As seeds
continued to develop, FP decreased in Ye seeds, whereas in DP seeds, there was a significant increase compared to Ye seeds at 40 DAP. TP showed a significant increase in DP seeds compared to that in Ye seeds. In contrast, TP was significantly reduced in DP seeds at 20 and 30 DAP. At 40 DAP, TP was reduced in Ye seeds and significantly increased in DP seeds (Figure 2a and b).

Similarly, the anthocyanin content was lower in Ye seeds than in DP seeds at the same stage. DP seeds had the highest anthocyanin content at 30 DAP and showed a reduction at 40 DAP; however, the content in all stages was significant compared to that in Ye seeds in the corresponding stages (Figure 2c).

We measured the chlorophyll (Chl) contents, Chla, Chlb, total Chl, and Chla/Chlb ratio during seed development in both Ye and DP seeds. During the early stages (10 and 20 DAP), the Chl contents were high in Ye and DP seeds, whereas at 10 DAP, Chla, Chlb, total Chl, and the Chla/Chlb ratio were higher in DP seeds than in Ye seeds. At 20 DAP, the Ye seeds had higher contents than DP seeds. As expected, 30 and 40 DAP had the lowest chlorophyll content in both Ye and DP seeds, as displayed in Figure 1, where chlorophyll cannot be seen. Interestingly at 40 DAP in both Ye and DP seeds, there was an increase in Chla, Chlb, and total Chl contents, whereas the Chla/Chlb ratio was reduced (Table 1).

**Flavonoid biosynthetic gene expression**

We analyzed the flavonoid biosynthetic gene expression during color deposition in each seed developmental stage. As shown in Figure 3, the early biosynthetic genes (EBGs; CHS, CHI, and F3H) and late biosynthetic genes (LBS; F’3’5H, DFR, ANS, and UFGT) were mostly highly expressed in DP seeds at 20 (CHI, F3H, DFR, ANS, and UFGT) or 30 (CHS and ANR) DAP, with the exception of F’3’5H, which was highly expressed at 40 DAP compared to that in Ye seeds. These results are consistent with our images of color deposition and total anthocyanin quantification (Figure 1). Moreover, ANR had similar expression levels during 10 and 20DAP and during 30 and 40 DAP, whereas F’3’5H had similar expression from 20 to 40 DAP. Furthermore, based on transcriptome analysis, we created a schematic representation and analyzed expression of flavonoid biosynthetic genes using the *Arabidopsis* unipro annotations in Ye and DP seeds during seed color deposition. The DEG names are listed next to the heat map in Figure 4.
Evolutionary Bioinformatics

The genes shown are CHS, CHI, F3H, F3’5’H, FLS, OMT1UGT, DFR, LDOX, ANR, OMT, and GT, which have been shown to be expressed in the flavonoid pathway.21

Transcriptome analysis and qRT-PCR expression levels

Based on transcriptome analysis (Figure 5a, Supplemental Tables S2, and S4), in general, more MYB TFs showed expression (up or down regulated) during the developmental process from 10 to 30 DAP in DP seeds, resulting in 60 MYB TFs expressed in total. More MYBs (>2 fold change) were detected at 10 and 20 DAP in Ye (DP10 DAP_Ye10DAP and DP10DAP_Ye20DAP) than in DP. Moreover, we selected genes for qRT-PCR analysis that showed high expression in DP20DAP_Ye20DAP: Traes_3DL_30CF35BB3 (2.017 fold), Traes_2BL_EED456A17 (2.796 fold), Traes_6BS_44456AE22 (3.828 fold), and Traes_2BL_3041037F1 (4.077 fold). The qRT-PCR results (Figure 5b) revealed that among the genes studied, Traes_3DL_30CF35BB3 showed the highest expression in DP seeds at early stages (DP10DAP and DP20DAP) compared to the other genes. Thus, based on these results, we selected Traes_3DL_30CF35BB3 for further experiments.

Isolation of TaTCL2

Our sequence clone of Traes_3DL_30CF35BB3 was cloned using primers shown Supplemental Table S1. After sequencing, the sequence was subjected to BLAST in the NCBI database, which indicated that the PCR product had 99% identity with a predicted sequence of *Aegilops tauschii* subsp. tauschii MYB-like TF TCL2 (LOC109750608), transcript variant X2, mRNA (sequence ID: XM_020309568.1), which confirmed that the obtained gene corresponded to the MYB TF TCL2.

Phylogenetic analysis and sequence alignments

To determine the evolutionary relationship of TaTCL2, phylogenetic analyses were performed by fast minimum evolution algorithm. The phylogenetic tree revealed evolutionary distance between TaTCL2 and 28 MYB-like protein from monocot species (Supplemental Figure S2). TaTCL2 and 28 MYB-like protein showed various sequence similarity (>70% amino acid identity). Interestingly, TaTCL2 and HvTCL2 displayed a very high sequence similarity (>92% amino acid identity) (Supplemental Figure S2).

To examine the relationship between TaTCL2 and MYB TFs in other plants that regulated anthocyanin biosynthesis, amino acid sequences or translated products were downloaded from the NCBI database to construct a phylogenetic tree. TaTCL2 was closest to AtTCL2, PhODO1, SCAN2like, AtMYB, FaMYB5, PhPH4, VvMYB5b, and VvMYBCS1 (Figure 6a).

Domain search using Pfam database showed that it was a MYB-like DNA binding domain. This family contains the DNA binding domains from MYB proteins, as well as the SANT domain family.44 Moreover, our sequence showed the motifs PF00249 and PF13921, both from the SANT domain (a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-CoR and TFIIB). Likewise, both motifs were identified as MYB-like DNA-binding domains (Supplemental Table S5). Figure 6b shows that the SANT domain is also found in the genes that TaTCL2 was clustered with in our phylogenetic tree.

Network interaction analysis of TaTCL2 response to anthocyanin biosynthesis

The above results suggest a high possibility that the TaTCL2 gene is involved in the anthocyanin biosynthetic pathway for...
wheat seed coat color deposition. Network interaction analysis has been recently demonstrated as a powerful method to study gene function.42 STRING 11.0b software was used to reconstruct the interaction network of TaTCL2 based on the orthologous genes of Arabidopsis. Figure 7a shows that TaTCL2 (homologous to AtTCL2 in Arabidopsis) can be associated with GL2, EGL3, and TTG1, which could combine with several bHLH proteins during anthocyanin biosynthesis.

We performed qRT-PCR of genes showing direct interaction in the interaction network analysis with TCL2 (Figure 7a). Results showed that GL2 (a bHLH) expression was higher in DP seeds at 10 and 20 DAP compared to Ye seeds. At 20 DAP, DP seeds showed the highest expression. At 30 DAP, GL2 expression in DP seeds was reduced compared to that in other stages. EGL3 (a bHLH) and TTG1 (a WD40) had similar expression patterns, with the highest expression in DP 10 DAP seeds. At 20 DAP, EGL3 and TTG1 exhibited decreased expression in both Ye and DP seeds, with higher expression in Ye than in DP seeds. Furthermore, at 30 DAP, the DP seeds had a gene expression reduction compared to Ye seeds in both, EGL3 and TTG1 (Figure 7b).

Discussion
Whole caryopses and cryo-cuts deposition
Our results revealed visible purple color at 20 DAP. Followed by an uniform pigmentation that was seen between 20 and 30 DAP (Figure 1a), in contrast to early stages where pigment distribution was not homogeneous. Similar results were reported...
Evolutionary Bioinformatics

by Trojan et al.\textsuperscript{15} in purple and blue wheat seeds where in the early stages the chlorophyll pigmentation was covering the grains, whereas at 20 or 30 DAP the pigmentation started to be distributed uniformly.

Moreover, in Figure 1b, the purple color was first deposited in the peripheral parts of the pericarp, and the remnants of chlorophyll were detected in the aleurone layer. Chlorophyll disappeared faster in DP seeds than in Ye seeds.

Figure 5. Heat map and transcript levels of MYB putative genes: (a) Heat map of the differentially expressed genes of MYB. A positive value indicates upregulation in the late stage; for example, in DP10DAP_DP20DAP, the plus value of log2FoldChange means upregulation in DP20DAP stage. We used 4 annotation Databases: Poaceae\_NR, Rice\_unipro, Brachypodium\_unipro, and Arabidopsis\_unipro. Numbers represent the fold changes. Colors are based using row minimum and maximum values. The mean values were obtained from 3 biological replicates. (b) qRT-PCR transcription analysis of 4 selected MYB putative genes during seed coat color deposition in 3 different stages of deep purple (DP) and yellow (Ye) samples. Ten days after pollination (DAP), 20 DAP, and 30 DAP (early, middle, and late stages, respectively). Ye is used as an internal control within a stage. Data are means ± SD of 3 biological replicates; (b) transcript levels of putative MYB genes. Data are means ± SEM of 3 biological replicates. Significant differences, evaluated by t-test, are indicated by *$P \leq 0.05$, and **$P \leq 0.01$ when comparing Ye versus DP within the DAP stage.
studies have reported that in purple seeds, chlorophyll disappeared from the pericarp to the aleurone layer. Furthermore, as the seeds continued to develop, there was a decrease in chlorophyll content, allowing purple and/or yellow pigments to be deposited. Moreover, the pericarp became thinner, color intensity increased, and the whole seed became dry (Figure 1).

Phenolic, anthocyanin, and chlorophyll quantification

Results showed that DP seeds had higher FP, TP, and anthocyanin contents than Ye seeds (Figure 2a–c). However, at 40 DAP, the anthocyanin content decreased (Figure 2c). We suggest that the decrease in anthocyanin content at 40 DAP may have been caused by the dilution of the total anthocyanin content by starch in the endosperm, resulting in a decrease in the anthocyanin content over time, as seen in a previous study in 6 wheat cultivars. Similarly, our results showed higher phenolic and anthocyanin contents in darker-colored seeds as reported previously.

As expected, chlorophyll content reduced with the development of seeds, as seen in Figure 1b and Table 1. This reduction allowed color deposition in both Ye and DP seeds. However, our results also show that there was an increase in chlorophyll content even after the seeds maturation as seen in Figure 1a, where it is shown that there is no more chlorophyll. However, these results can be explained by the method used in chlorophyll quantification and anthocyanin quantification. The wavelength used for measuring chlorophyll a, b, and anthocyanin were 652, 665, and 657 nm, respectively. Thus, anthocyanins might have also been quantified in the chlorophyll content in the matured seeds since the wavelength is juxtaposed, explaining the higher chlorophyll content in Ye and DP seeds 30 and 40 DAP.

qRT-PCR analysis of the flavonoid pathway

CHI, F3H, DFR, ANS, and UFGT gene expression peaked at 20 DAP in DP seeds. In Brassica napus, it was reported that CHI exhibited an expression pattern similar to that of CHS, since the activity of this enzyme is closely related to CHS. However, our results showed that CHS peaked at 30 DAP, which might be due to the genotype used. Moreover, during grain development, CHS has been shown to decrease over time. Our results showed that CHS reduced at 40 DAP. F3H and DFR had similar expression patterns to those of CHS and CHI, but they peaked at different developmental stages. Similar results have been shown for DFR, which peaked at 18 days after flowering. Between our 2 cultivars, DFR expression was similar to that in previous reports. CHS and ANR peaked at 30 DAP in DP seeds.
Additionally, F′3′5H peaked at 40 DAP in DP seeds, in contrast to previously reported results, where it peaked in mid-grain development.45 The expression levels of genes involved in pigment biosynthesis were found to vary greatly between the Ye and DP seed lines of *Triticum aestivum* L. However, all these genes were also expressed in the yellow-seeded lines. These data indicate that the absence of pigment synthesis in the yellow-seeded line of *Triticum aestivum* L. involves the downregulation, but not the complete inactivation, of several key genes.

The seed development of DP seeds was accompanied by an increase of anthocyanin content and the expression of *TaTCL2* and genes related to the flavonoid pathway.

**Phylogenetic tree and domain**

Evolutionary studies have revealed that of the genes involved in the anthocyanin pathway, structural genes evolve more slowly than regulatory genes.47 The evolutionary relationship of *TaTCL2* showed closely relationship from monocot plant in *Poaceae* (Supplemental Figure S2). It has been shown that *TaMpc1* and *HvMpcs* had a high connection between gene expression and the appearance of anthocyanin pigmentation, and the lineages of wheat and barley diverged 16 to 19 million years age (MYA).48 However, a previous study of the anthocyanin biosynthesis pathway showed that the anthocyanin genes are species specific and are not dependent on the class of the plant species.49

*TaTCL2* is a functional MYB transcription factor gene regulating anthocyanin biosynthesis. Upon examining the relationship between *TaTCL2* and MYB TFs in other plants, *TaTCL2* was revealed to be closest to *AtTCL2, PbODO1, SCAN2like, AtMYB, FuMYB, PbPH4, VvMYB5b*, and *VvMYBCS1* (Figure 6a). The dendrograms indicated that *TaTCL2* belonged to a distinct cluster of MYB proteins. We searched for domains in our sequence using Pfarm database, and the results showed that it contained a SANT (Swi3, Ada2, N-Cor, and TFIIIB) which has been reported to be an important domain for the anthocyanin biosynthesis.43,50 The SANT domain was found in the genes that *TaTCL2* was clustered with in our phylogenetic tree (Figure 6b). Moreover, the transcript level of *TaTCL2* was substantially higher at the early stages (10 and 20 DAP) in DP seed relative to Ye seeds with low anthocyanin content.
Interaction analysis

Our results showed that our cloned gene, TaTCL2, is likely involved in the anthocyanin biosynthesis pathway for wheat seed coat color deposition. R2R3 MYB TFs have been shown to interact closely with bHLH TFs31,52 of the IIIf subfamily of bHLH for Arabidopsis, which have been demonstrated to be involved in flavonoid biosynthesis and trichome formation.29,53,54 TT8 (Transparent testa), GL3 (Glabra 3), and EGL3 (Enhancer of Glabra 3) from the IIIf subfamily along with TTG1 (Transparent testa glabra 1-WD40) interact to form MBW complexes, which function in regulating the flavonoid biosynthetic genes and influencing seed coat color formation55 and anthocyanin biosynthesis for Arabidopsis, tomato, and strawberry.56-60 Our network interaction analysis of TaTCL2 based on the orthologous genes of Arabidopsis showed strong interactions with GL2, EGL3, and TTG1, which could form an MBW complex related to anthocyanin biosynthesis (Figure 7a).

We performed qRT-PCR of genes showing direct interactions in interaction network analysis. It is shown that there was a higher expression in GL2, TTG1, and EGL3 at the early stages (10 and 20 DAP) compared to the late ones (30 and 40 DAP), which is in accordance to the color formation in the seeds as seen in Figure 1. These allows us to conclude that there is a high possibility that an MBW complex can be formed and that TaTCL2 is involved in seed coat color deposition in Triticum aestivum L.

Conclusions

This study analyzed the networks underlying pigment formation, investigated the variation in flavonoid accumulation during seed development in Ye and DP seeds, and monitored the differential expression of the main structural genes and TFs involved in the flavonoid biosynthetic pathway using transcriptome analysis. The findings of this study elucidate the mechanisms underlying color formation in Triticum aestivum L. seed coats. Furthermore, by cloning the novel gene TaTCL2, we uncovered evidence that this gene is important in seed color formation through an MBW complex that regulates anthocyanin formation in DP seeds. The findings of this study will lay a foundation for understanding the molecular mechanism underlying DP color formation for DP seed breeding in T. aestivum L.

Acknowledgements

PCF acknowledges CONACYT for the scholarship granted.

Author Contributions

PCF and YWS conceived and designed the experiments. PCF performed experiments, analyzed data, and wrote the manuscript with support from JSY, DYK, and YWS. JSY helped with gene expression analysis. DYK performed transcriptome analysis. YWS contributed with valuable discussions. All authors have discussed the results and approved the final manuscript.

ORCID iD

Yong Weon Seo https://orcid.org/0000-0002-6052-7491

Supplemental material

Supplemental material for this article is available online.

REFERENCES

1. Rebetzke GJ, Jimenez-Berni J, Fischer RA, Deery DM, Smith DJ. Review: high-throughput phenotyping to enhance the use of crop genetic resources. Plant Sci. 2019;282:40-48.
2. Roucou A, Violle C, Fort F, Roumet P, Ecarnot M, Vile D. Shifts in plant functional strategies over the course of wheat domestication. J Appl Ecol. 2018;55:365-375.
3. Wood SA, Karp DS, DeClerck F, Kremen C, Naeem S, Palm CA. Functional traits in agriculture: agrobiodiversity and ecosystem services. Trends Ecol Evol. 2015;30:531-539.
4. Zýkín PA, Andreeva EA, Lykholyub AN, Tsvetkova NV, Volykova AV. Anthocyanin composition and content in eye plants with different grain color. Molecules. 2018;23:948.
5. Moise JA, Han S, Guányin-Savich L, Johnson DA, Miki BLA. Seed coats: structure, development, composition, and biotechnology. In: Vira Cell Dev Biol Plant. 2005;41:620-644.
6. Calderon Flores P, Yoon JS, Kim DY, Seo YW. Effect of chilling acclimation on germination and seedlings response to cold in different seed coat colored wheat (Triticum aestivum L.). BMC Plant Biol. 2021;21:252.
7. Himi E, Noda K. Isolation and location of three homoeologous dihydroflavonol-4-reductase (DFR) genes of wheat and their tissue-dependent expression. J Exp Bot. 2004;55:365-375.
8. Deng B, Yang K, Zhang Y, Li Z. The effects of temperature on the germination behavior of white, yellow, red and purple maize plant seeds. Acta Physiol Plant. 2015;37:174.
9. Furuta S, Takahashi M, Takahata Y, et al. Radical-scavenging activities of soybean cultivars with black seed coats. Food Sci Technol Res. 2003;9:73-75.
10. Takahashi R, Ohmori R, Kiyosue Y, Momiyama Y, Ohsumi F, Kondo K. Antioxidant activities of black and yellow soybeans against low density lipoprotein oxidation. J Agric Food Chem. 2005;53:4578-4582.
11. Slavin M, Kenworthy W, Yu LL. Antioxidant properties, phytochemical composition, and antiproliferative activity of Maryland-grown soybeans with colored seed coats. J Agric Food Chem. 2009;57:11174-11185.
12. Washburn S, Burke GL, Morgan T, Anthony M. Effect of soy protein supplementation on serum lipoproteins, blood pressure, and menopausal symptoms in perimenopausal women. Menopause. 1999;6:7-13.
13. Q C, Fu F, Lu K, et al. Differential accumulation of phenolic compounds and expression of related genes in black- and yellow-seeded Brassica napus. J Exp Bot. 2013;64:2885-2898.
14. Rahman MM, Lee KE, Kang SG. Studies on the effects of pericarp pigmentation on grain development and yield of black rice. Indian J Genet Plant Breed. 2015;75:426-433.
15. Trojan V, Musilova M, Vynhanek T, Klejdus B, Hanáček P, Havel L. Chalcone synthase expression and pigments deposition in wheat with purple and blue colored caryopsis. J Cereal Sci. 2014;59:48-55.
16. Eeau K. Anatomy of Seed Plants. John Wiley and Sons; 1977.
17. Knievel DC, Abdul-Aal ES, Rabalski I, Nakamura T, Huel P. Grain color development and the inheritance of high anthocyanin blue aleurone and purple pericarp in spring wheat (Triticum aestivum L.). J Cereal Sci. 2009;50:113-120.
18. Liu MS, Wang F, Dong YX, Zhang XS. Expression analysis of dihydroflavonol-4-Reductase genes involved in anthocyanin biosynthesis in purple grains of wheat. J Integr Plant Biol. 2005;47:1107-1114.
19. Xie DY, Sharma SB, Paiva NL, Ferreira D, Dixon RA. Role of anthocyanidin reductase, encoded by BANULYS in plant flavonoid biosynthesis. Science. 2003;299:396-399.
20. Zhang K, Lu K, Q C, et al. Gene silencing of BnTT10 family genes causes retarded pigmentation and lignin reduction in the seed coat of Brassica napus. PLoS One. 2013;8:e61247.
21. Hong M, Hu K, Tian T, et al. Transcriptomic analysis of seed coats in yellow-seeded Brassica napus reveals novel genes that influence proanthocyanidin biosynthesis. Front Plant Sci. 2017;8:1674.
22. Lepiniec L, Debeaujon I, Routaboul JM, et al. Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol. 2006;57:405-430.
23. Saito K, Yonekura-Sakakibara K, Nakabayashi R, et al. The flavonoid biosynthetic pathway in Arabidopsis: structural and genetic diversity. Plant Physiol Biochem. 2013;72:21-34.

24. Appelhagen I, Thiebig K, Nordholt N, et al. Update on transparent testa mutants from Arabidopsis thaliana: characterisation of new alleles from an isogenic collection. New Phytol. 2014;202:132-144.

25. Xu W, Grain D, Bobet S, et al. Complexity and robustness of the flavonoid transcriptional regulatory network revealed by comprehensive analyses of MYB-bHLH-WDR protein complexes and their targets in Arabidopsis seed. New Phytol. 2014;204:686-691.

26. Shin DH, Choi MG, Kang CS, Park CS, Choi SB, Park YI. A wheat R2R3-MYB protein PURPLE PLANT1 (TaPL1) functions as a positive regulator of anthocyanin biosynthesis. Biochem Biophys Res Commun. 2016;469:668-691.

27. Ramosy NA, Glover BJ. MYB-bHLH-WD40 protein complexes and the evolution of cellular diversity. Trends Plant Sci. 2005;10:63-70.

28. Páz-Ares J, Ghosal D, Wientand U, Peterson PA, Saedler H. The regulatory cl locus of Zea mays encodes a protein with homology to myb proto-oncogene products and with structural similarities to transcriptional activators. EMBO J. 1987;6:3553-3558.

29. González A, Zhao M, Leavitt JM, Lloyd AM. Regulation of the anthocyanin biosynthetic pathway by the TTT1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. Plant J. 2008;53:814-827.

30. Solano R, Nieto C, Avila J, Calas L, Daza I, Páz-Ares J. Dual DNA binding specificity of a petals epidermis-specific MYB transcription factor (MYB.PE3) from Petunia hybrida. EMBO J. 1995;14:1773-1784.

31. Appelhagen I, Nordholt N, Seidel T, et al. TRANSPARENT TESTA 13 is a tonoplast P3A-ATPase required for vacuolar deposition of proanthocyanidins in Arabidopsis thaliana seeds. Plant J. 2015;82:840-849.

32. Shin OH, Kim DY, Seo YW. Effects of different depth of grain colour on antioxidant capacity during water imbibition in wheat (Triticum aestivum L.). J Sci Food Agric. 2017;97:2750-2758.

33. Galicia L, Miranda A, Gutiérrez MG, et al. Laboratorio de calidad nutricional de maíz y análisis de tejido vegetal: protocolos de laboratorio. 2012. CMMYT, CIMMYT.

34. Hong MJ, Kim DY, Aho JW, Kang SY, Seo YW, Kim JB. Comparison of radiation sensitivity response to acute and chronic gamma irradiation in colored wheat. Genet Mol Biol. 2018;41:611-623.

35. Warren CR. Rapid measurement of chlorophylls with a microplate reader. J Plant Nutr. 2008;31:1331-1332.

36. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2 − ΔΔCT method. Methods. 2001;25:402-408.

37. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114-2120.

38. Kim D, Langmead B, Salberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357-360.

39. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166-169.

40. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with deseq2. Genome Biol. 2014;15:50.

41. Desper R, Gascuel O. Theoretical foundation of the balanced minimum evolution method of phylogenetic inference and its relationship to weighted least-squares tree fitting. Mol Biol Evol. 2004;21:587-598.

42. Zhao F, Li G, Hu P, et al. Identification of basic/helix-loop-helix transcription factors reveals candidate genes involved in anthocyanin biosynthesis from the strawberry white-flesh mutant. Sci Rep. 2018;8:2721.

43. Zong Y, Zhu X, Liu Z, et al. Functional MYB transcription factor encoding gene AN2 is associated with anthocyanin biosynthesis in Lycium ruthenicum Murray. BMC Plant Biol. 2019;19:169.

44. Aasland R, Stewart AF, Gibson T. The SANT domain: a putative DNA-binding domain in the SWI-SNF and ADA complexes, the transcriptional co-repressor N-cor and TFIIB. Trends Biochem Sci. 1996;21:87-88.

45. Yang GH, Li B, Gao JW, et al. Cloning and expression of two chalcone synthase and a flavonoid 3′′-5′′-hydroxylase 3′′-end CDNA from developing seeds of blue-grained wheat involved in anthocyanin biosynthetic pathway. Acta Bot Sin. 2004;46:588-594.

46. Yang GH, Zhao XQ, Li B, et al. Molecular cloning and characterization of a DFR from developing seeds of blue-grained wheat in anthocyanin biosynthetic pathway. Mol Biol Evol. 1999;16:266-274.

47. Rausher MD, Miller RE, Trifin P. Patterns of evolutionary rate variation among genes of the anthocyanin biosynthetic pathway. Mol Biol Evol. 2004;21:525-537.

48. Strigini K, Khaledstina EK. Structural and functional divergence of the Mpc1 genes in wheat and barley. BMC Evol Biol. 2019;19:45.

49. Shoeva OY, Glagoleva AY, Khaledstina EK. The factors affecting the evolution of the anthocyanin biosynthesis pathway genes in monocot and dicot plant species. BMC Plant Biol. 2017;17:236.

50. Oglesby L, Aranga A, Obuya J, Ochicago J, Cebert E, Trolova V. Anthocyanin accumulation in muscadine berry skins is influenced by the expression of the MYB transcription factors, MybA1, and MYBCS1. Antioxidants. 2016;5:35.

51. Mol J, Jenkins G, Schäfer E, Weiss D, Walbot V. Signal perception, transduction and gene expression involved in anthocyanin biosynthesis. CRC Crit Rev Plant Sci. 1996;15:525-557.

52. Winkel-Shirley B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. Plant Physiol. 2001;126:485-493.

53. Toledo-Ortiz G, Huq E, Quail PH. The Arabidopsis basic/helix-loop-helix transcription factor family. Plant Cell. 2000;12:1523-1534.

54. Pires N, Dolan L. Origin and diversification of the basic/helix-loop-helix transcription factor family. Plant Cell. 2003;15:1749-1770.

55. Pires N, Dolan L. Origin and diversification of the basic/helix-loop-helix transcription factors reveals candidate genes involved in anthocyanin biosynthesis from the strawberry white-flesh mutant. Plant Biol. 2008;10:133-144.

56. Feller A, Machemer K, Braun EL, Grotewold E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. Plant Biol. 2011;13:661-673.