Transcriptome Analysis Identifies Key Genes Responsible for Red Coleoptiles in *Triticum Monococcum*

Dong Cao [1,2,†]**, Jiequn Fan** [3,†]**, Xingyuan Xi** [1]**, Yuan Zong** [2]**, Dongxia Wang** [2]**, Huaigang Zhang** [1] **and Baolong Liu** [1,*]

[1] Qinghai Provincial Key Laboratory of Crop Molecular Breeding, Xining 810008, Qinghai, China; caodong@nwipb.cas.cn (D.C.); xixy@nwipb.cas.cn (X.X.); hgzhang@nwipb.cas.cn (H.Z.)
[2] State Key Laboratory of Plateau Ecology and Agriculture, Qinghai University, Xining 800010, Qinghai, China; laughing1898@icloud.com (Y.Z.); wangdx1127@163.com (D.W.)
[3] Shanghai Academy of Agricultural Sciences, Shanghai 201403, China; caod.08@163.com
[*] Correspondence: blliu@nwipb.cas.cn; Tel.: +86-971-6143-610; Fax: +86-971-6143-282
[†] These authors contributed equally to this paper.

check for updates

Abstract: Red coleoptiles can help crops to cope with adversity and the key genes that are responsible for this trait have previously been isolated from *Triticum aestivum*, *Triticum urartu*, and *Aegilops tauschii*. This report describes the use of transcriptome analysis to determine the candidate gene that controls the trait for white coleoptiles in *T. monococcum* by screening three cultivars with white coleoptiles and two with red coleoptiles. Fifteen structural genes and two transcription factors that are involved in anthocyanin biosynthesis were identified from the assembled UniGene database through BLAST analysis and their transcript levels were then compared in white and red coleoptiles. The majority of the structural genes reflected lower transcript levels in the white than in the red coleoptiles, which implied that transcription factors related to anthocyanin biosynthesis could be candidate genes. The transcript levels of *MYC* transcription factor *TmMYC-A1* were not significantly different between the white and red coleoptiles and all of the TmMYC-A1s contained complete functional domains. The deduced amino acid sequence of the MYB transcription factor TmMYB-A1 in red coleoptiles was homologous to TuMYB-A1, TaMYB-A1, TaMYB-B1, and TaMYB-D1, which control coleoptile color in corresponding species and contained the complete R2R3 MYB domain and the transactivation domain. TmMYB-a1 lost its two functional domains in white coleoptiles due to a single nucleotide deletion that caused premature termination at 13 bp after the initiation codon. Therefore, TmMYB-A1 is likely to be the candidate gene for the control of the red coleoptile trait, and its loss-of-function mutation leads to the white phenotype in *T. monococcum*.

Keywords: *T. monococcum*; coleoptile; anthocyanin biosynthesis; *MYB* transcription factor

1. Introduction

Red coleoptiles widely exist in wild relatives of cultivated wheat, such as *Triticum urartu*, *Aegilops tauschii*, and *T. monococcum*. Red coleoptiles, which are colored due to the accumulation of anthocyanin, can protect the emerging shoot from strong light, drought, and cold [1,2]. In addition, the coleoptile color trait is easy to observe and it has been employed as a useful characteristic in describing wheat varieties.

The anthocyanin biosynthetic pathway is relatively clear and the structural and regulatory genes have been isolated in several model plants [3–6]. There are some main structural genes in the anthocyanin biosynthesis pathway, including phenylalanine ammonia-lyase (PAL), cinnamate-4-hydroxylase (C4H),

4-coumarate: CoA ligase (4CL), shikimate O-hydroxycinnamoyltransferase (HCT), chalcone synthase (CHS), chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid-3′-hydroxylase (F3′H), flavonoid-3′,5′-hydroxylase (F3′5′H), flavonol synthase (FLS), dihydrofavonol-4-reductase (DFR), leucoanthocyanidin dioxygenase (LDOX), flavonoid-3-O-glucosyltransferase (UFGT), anthocyanidin reductase (ANR), and leucoanthocyanidin reductase (LAR) [6,7]. A complex of transcription factors MYB and MYC with WD40 proteins regulate these structural genes in the main [8]. The MYB protein includes two incomplete repeats, R2-MYB and R3-MYB, and a transactivation domain, whereas the MYC protein contains bHLH-MYC_N, HLH, and ACT-like domains. The conserved R3 repeat of the MYB protein is involved in an interaction with MYC, which forms the complex that enables their transcriptional functionality [9–11]. Both structural and regulatory genes are crucial in anthocyanin biosynthesis, but allelic variation in the *MYB* and *MYC* genes causes color variations in different plant tissues [12–16]. The genes for MYC-type transcription factors that control anthocyanin-related traits are *TT8* in *Arabidopsis* [17], *AN1* in *Petunia* [18], *ThMYC4E* and *TaMYC1* in *T. aestivum* [15,19,20], and *AetMYC1* in *Ae. tauschii* [16]. A similar phenomenon has been observed for the genes of MYB-type transcription factors, including *MYB75/PAP1* and *AtMYB90/PAP2* in *Arabidopsis* [21], *AN2* in *Petunia* [22], *MYB1* in sweet potato [23], *LAP1* in legumes [24], *MdMYBA* and *MdMYB1* in apple [25,26], and *EsMYBA1* in *Epimedium sagittatum* [27].

In common wheat, the red coleoptile trait is controlled by *Rc1*, *Rc2*, and *Rc3*, which are located on the short arm of homeologous chromosomes 7A, 7B, and 7D [28]. *Rc1*, *Rc2*, and *Rc3* have recently been isolated and found to encode three MYB transcription factors—*TaMYB-A1*, *TaMYB-B1*, and *TaMYB-D1*—respectively [14,29,30]. The loss-of-function mutation of *TuMYB-A1*, which is homologous to *Rc1*, *Rc2*, and *Rc3*, resulted in the white coleoptile trait in *T. urartu* [31,32]. In *Ae. tauschii*, the allelic variation of the *MYC* gene *AetMYC1* caused coleoptile color differentiation [16]. Color variation also exists in the coleoptiles of *T. monococcum*, but the genetic mechanism has not been identified. High-throughput RNA sequencing (RNA-Seq) provides an economic and convenient approach for detecting novel transcripts, single nucleotide polymorphisms, small RNAs, and alternate splicing products, as well as sense and antisense transcripts [32,33]. The technique offers the advantage of measuring the gene expression levels and obtaining corresponding nucleotide sequences without reference sequences [34–36]. In this manuscript, transcriptomic analysis was conducted to identify the key gene that is associated with coleoptile color variation in *T. monococcum* by comparing the nucleotide sequence variation with the expression level of the structural genes and transcription factors that are associated with anthocyanin biosynthesis in red and white coleoptiles.

2. Results

2.1. High-Throughput RNA Sequencing

To more comprehensively understand the coleoptile colour phenotype, the coleoptile colours of 143 *T. monococcum* cultivars were recorded. A total of 100 *T. monococcum* cultivars carried red coleoptiles and 43 carried white coleoptiles. The white coleoptile cultivars—PI190939 (W1), CItr13961 (W2), and CItr13962 (W3)—were selected to identify the key gene that is associated with the red coleoptile trait (Table S1; Figure 1), while two cultivars PI554495 (R1) and PI554500 (R2), carrying red coleoptiles, were chosen as the controls. Comparative transcriptome analysis was performed between the *T. monococcum* cultivars with red coleoptiles (R1 and R2) and white coleoptiles (W1, W2, and W3). After filtering, we obtained 53–60 million clean reads for each sample (Table S2). Trinity software was used to link the valid data to 294,658 UniGenes totaling 257,478,102 bp. The average length of each gene was 873 bp and the length of the N50 was 1466 bp.

BLASTX software was employed to annotate the assembled UniGenes using the non-redundant (Nr) (E_value $\leqq$ 0.01), Swissprot (E_value $\leqq$ 0.01), Kyoto Encyclopedia of Genes and Genomes (KEGG) (E_value $\leqq$ 1e^{-5}), and Clusters of Orthologous Groups (COG) (E_value $\leqq$ 1e^{-5}) databases. Totals of 171,946 (58.35%), 85,516 (29.02%), 48,893 (16.60), and 64,605 (21.93%) UniGenes could be annotated to

the Nr, Swiss-Prot, KEGG, and COG databases, respectively (Figure 2; Table S3). A sum of 176,630 proteins was determined following the removal of repeated proteins.

Figure 1. The coleoptiles of PI190939 (W1), CItr13961 (W2), CItr13962 (W3), PI554495 (R1), and PI554500 (R2).
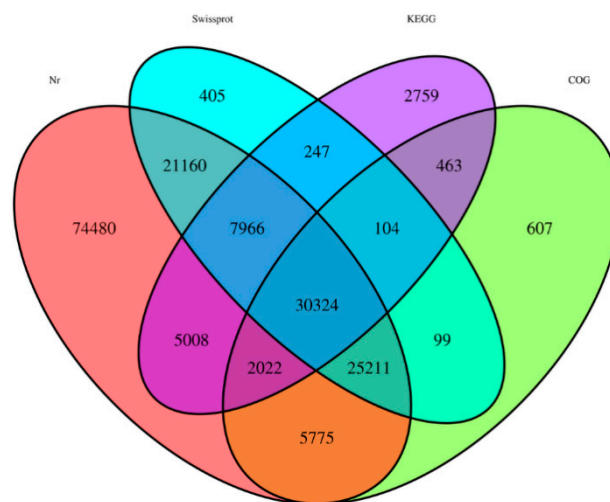

Figure 2. Outcome of the homology search of *T. monococcum* UniGenes.

2.2. Expression Profiles of Structural Genes Associated with Anthocyanin Biosynthesis

Coleoptile color variation is associated with the biosynthesis of anthocyanin and 15 structural genes that are related to anthocyanin biosynthesis derived from *Zea mays* were selected for a BLAST search of the assembled UniGene database and NCBI database. A total of 36 UniGenes were homologous to *PAL*, *C4H*, *4CL*, *HCT*, *CHS*, *CHI*, *F3H*, *FLS*, *F3′H*, *F3′5′H*, *DFR*, *LAR*, *LDOX*, *ANR*, and *UFGT* (Figure 3). The only UniGene that could not be identified was that homologous to *C3′H*. The phylogenetic trees were constructed using the deduced amino acid sequences of 36 unigenes in *T. monococcum* and structural genes associated with anthocyanin biosynthesis in other species (Figures S1–S15). The same structural genes were assigned to the same branch (Figures S1–S15). The fragments per kilobase of transcript per million mapped reads (FPKM) values of these genes were determined in order to evaluate their relative expression levels. The anthocyanin biosynthesis pathway was exhibited in *T. monococcum*, as described by Holton and Cornish [37] and Winkel-Shirley [6] (Figure 3). All of the structural genes exhibited higher expression levels in red coleoptiles (Figure 3; Table S4). The relative transcript level is between 1.06 and 1698.45 times higher in red coleoptiles than white coleoptiles, which implied that the transcription factor might be the key gene for the white coleoptile trait in *T. monococcum*.
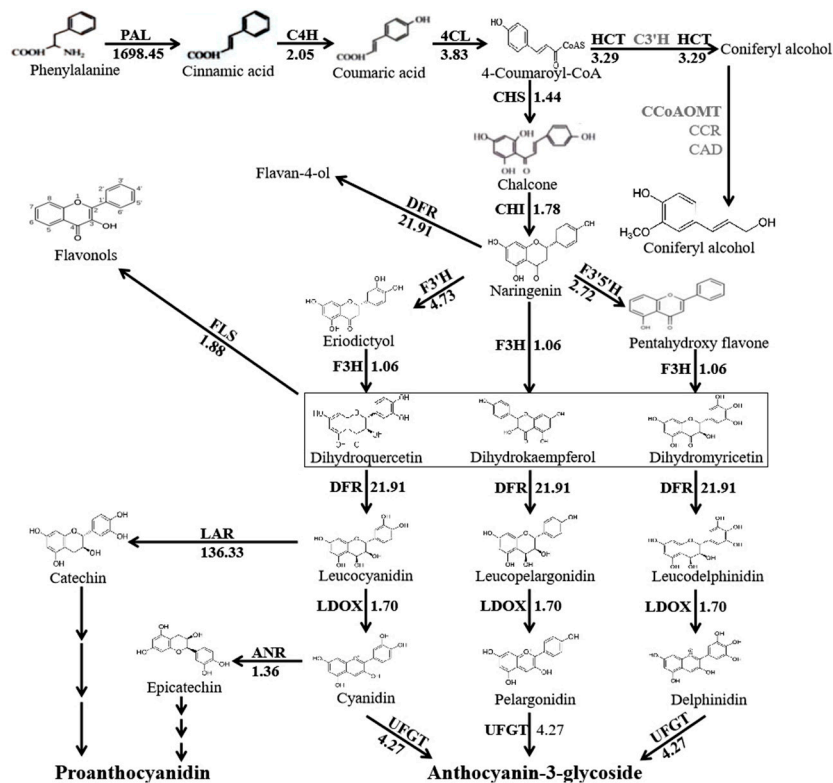
Figure 3. The expression level differences of anthocyanin biosynthesis-related structural genes in red and white coleoptiles. Arrows show the metabolic stream; the light color indicates that these genes are in the assembled UniGenes; the numbers represent the increased level of expression in red coleoptiles compared with white coleoptiles.

2.3. Sequence Characteristics and Expression Profiles of Regulatory Genes Associated with Anthocyanin Biosynthesis

TuMYB-A1, *TaMYB-A1*, *TaMYB-B1*, and *TaMYB-D1* control thecoleoptile color variation in *T. urartu* and *T. aestivum*, respectively [29–31], while *AetMYC1* is associated with variation of this trait in *Ae. Tauschii* [16]. Therefore, *TuMYB-A1* and *AetMYC1* were used in a BLAST search to identify homologous genes from the assembled UniGene database. One MYB transcription factor UniGenes (*R1.29805_c0_g6_i1*) and three MYC transcription factors (W2.26064_c0_g1_i6, W2.26064_c0_g1_i7, and W2.26064_c0_g1_ i12) UniGenes were obtained with scores >100. Each MYC contained between one and eight transcripts. A phylogenetic tree was constructed with the neighbor-joining method using the deduced amino acid sequences of the three MYC transcription factors and other MYCs that are involved in anthocyanin biosynthesis (Figure S16). W2.26064_c0_g1_i6, W2.26064_c0_g1_i7, and W2.26064_c0_g1_ i12 were assigned to the same branch as AetMYC1 (Figure S1), the key gene that is related to red coleoptiles in *T. aestivum* and *Ae. tauschii* [16,29]. W2.26064_c0_g1_i6, W2.26064_c0_g1_i7, and W2.26064_c0_g1_ i12 were splicing transcripts from the same gene, which was named *TmMYC1*. When compared with white coleoptiles, the average transcript levels of *TmMYC1* were lower in red coleoptiles (Table S4). The deduced amino acid sequences of TmMYC-A1 from white and red coleoptiles were the same length and they contained the same primary function domains, including the integrated bHLH-MYC_N domain, which is essential for protein–protein interaction, the HLH domain, which can facilitate protein-DNA binding, and the ACT-like domain, which initiates transcription (Figure S17).

The phylogenetic tree was also constructed by using the deduced amino acid sequences of the UniGenes *R1.29805_c0_g6_i1* and other MYBs that were involved in anthocyanin biosynthesis (Figure S18). The MYB transcription factor UniGene R1.29805_c0_g6_i1 was assigned to the same

branch as TuMYB-A1, TaMYB-A1, TaMYB-B1, and TaMYB-D1 (Figure S3), which control coleoptile color in *T. urartu* [31] and *T. aestivum* [29,30,38]. *R1.29805_c0_g6_i1* was named *TmMYB-A1*. The average transcript levels of *TmMYB-A1* in red coleoptiles were 13.31 times greater than in white coleoptiles. All *TmMYB-a1* from white coleoptiles exhibited one base deletion (G) at 13 bp when compared with *TmMYB-A1* from red coleoptiles. The deduced amino acid sequences of TmMYB-A1 from the red coleoptiles contained the primary functional structure of the MYB protein R2R3 MYB and transcript activator domains (Figure 4), while TmMYB-a1 lost the R2R3 MYB and transcript activator domains, because the frame shift mutation caused premature termination of translation for the encoded protein (Figure 4).
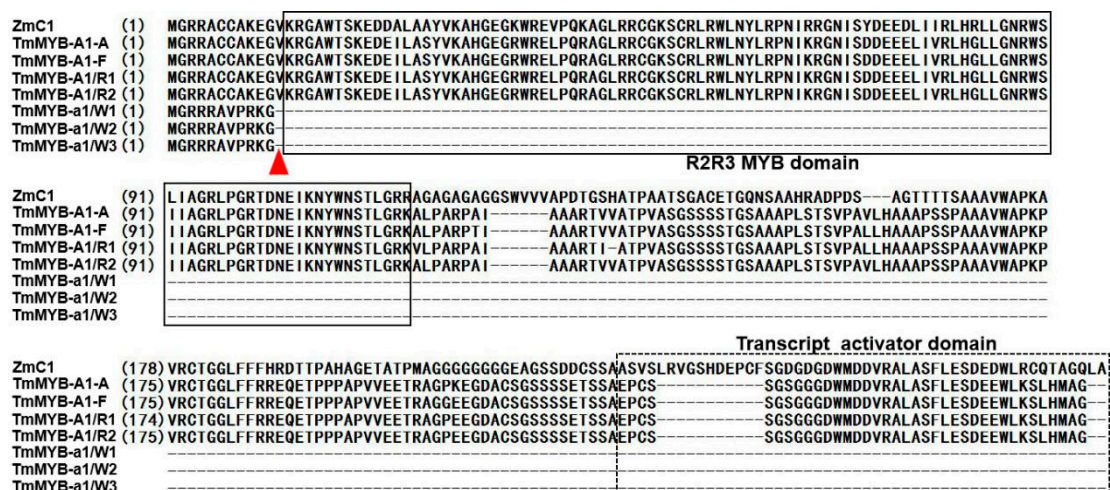

Figure 4. The deduced amino acid sequence alignment of TmMYB-A1 from R1, R2, W1, W2, W3, and MYBs, which regulate anthocyanin biosynthesis in other species. The solid frame represents the R2R3 MYB domain and the dotted frame represents the transcript activator domain. The red triangle represents the location of the stop codon.

3. Discussion

The comparative transcriptome analysis demonstrated that the structural genes involved in anthocyanin biosynthesis showed lower transcript abundance in white coleoptiles than the red coleoptiles, and these structural genes are primarily regulated by the MYB-bHLH-WD40 complex [39]. Loss of function or low expression of a *MYB* or *MYC* transcription factors could impede anthocyanin biosynthesis in coleoptiles [16,40–42] and, hence, the transcription factor is likely to be the candidate gene for coleoptile color variation in *T. monococcum*.

The MYC transcription factor *TmMYC1* was homologous to *AeMYC1*, which controls color variation in *Ae. tauschii*. However, the transcript levels of *TmMYC1* showed no significant difference between the white and red coleoptiles. All of the deduced amino acid sequences of *TmMYC1* from R1, R2, W1, W2, and W3 contained the integrated bHLH-MYC_N domain, HLH domain, and ACT-like domain. This suggested the *MYC* transcription factor could not be the key gene for the white coleoptile trait in *T. monococcum*. With regard to the MYB transcription factor, *TmMYB-A1* was on the same branch as TuMYB-A1, TaMYB-A1, TaMYB-B1, and TaMYB-D1 (Figure S18), which control coleoptile color in *T. urartu* [31] and *T. aestivum* [29,30,38] and the transcript levels of *TmMYB-A1* in red coleoptiles was significantly higher than in the white coleoptiles. The deduced amino acid sequence of TmMYB-A1 was homologous to TuMYB-A1 and it contained the complete R2R3 MYB domain and transcript activator domain, while the deduced amino acid sequences of TmMYB-a1 had lost the R2R3 MYB domain and transcript activator domain, because a single nucleotide deletion caused premature termination. Functional loss of its homologous genes, TuMYB-A1, TaMTB7A, TaMTB7B, and TaMTB7D, produced the white coleoptile in *T. urartu* [31] and *T. aestivum* [29,30,38] and, therefore, *TmMYB-a1* appeared to

be the key gene that is responsible for the white coleoptile trait in *T. monococcum* accessions W1, W2, and W3, as signified by low transcript levels and the loss of function.

4. Materials and Methods

4.1. Plant Materials

A total of 143 *T. monococcum* cultivars were obtained from the USDA-ARS National Small Grains Collection [43] and the Institute of Genetics and Development Biology, Chinese Academy of Sciences (Table S4). Ten seeds were germinated to survey coleoptile color [30].

4.2. Preparation of Total RNA and cDNA

Total RNA was isolated from the coleoptiles of five-day-old seedlings using the TIANGEN RNAprep Pure Plant Kit (Tiangen Company, Beijing, China). The total RNA of each sample was quantified and qualified with an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, USA), NanoDrop (Thermo Fisher Scientific, Wilmington, DE, USA), and 1% agrose gel. cDNA was obtained from the total RNA using the Thermo RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher Scientific, St. Louis, MO, USA).

4.3. Transcriptome Analysis

Ten cDNA library preparations were constructed (two replicates per accession) according to the manufacturer's protocol for mRNA-Seq sample preparation (Illumina, Inc., San Diego, CA, USA). Ten cDNA library products were sequenced by Illumina paired-end sequencing technology, with read lengths of 150 bp on the Illumina HiSeq X instrument by Genewiz Biotechnology Co. Ltd. (Suzhou, China). The raw sequence reads were stored in the National Center for Biotechnology Information (NCBI) SRA database, with the accession number SUB5053734. Before assembly, the raw paired-end reads were filtered in order to obtain high quality clean reads. Low quality sequences were removed, including the sequences with ambiguous bases (denoted by more than 10% "N" in the sequence trace), low quality reads (the rate of reads in which a quality value ≤ 10 is greater than 20%), and reads with adaptors. After filtering, the high quality reads were assembled by Trinity, with default parameters, to construct unique consensus sequences that are based on the sequences from the *T. monococcum* accessions R1, R2, W1, W2, and W3 genotypes [44]. UniGenes that were differentially expressed between red and white coleoptile genotypes were analyzed using chi-square tests with IDEG6 software (University of Padua, Padua, Italy) [45]. The UniGene expression level was calculated while using the FPKM values. The false discovery rate (FDR) method was introduced to determine the threshold p-value at FDR ≤ 0.05, and the absolute value of $|\log 2\text{Ratio}| \geq 2$ was used as the threshold to determine the significance of the differential expression of the UniGenes. The genes that are related to anthocyanin biosynthesis in KEGG pathways [46] were collected and then aligned to the UniGenes from a transcriptome mixture of red and white coleoptiles using the BLASTX algorithm with an E-value of $< 1.0 \times 10^{-5}$.

4.4. Bioinformatic Analysis

The sequence alignments were conducted using the Vector NTI 10 software package (Thermo Fisher Scientific). MEGA5 software was used to construct the phylogenetic tree (Tokyo Metropolitan University, Hachioji, Tokyo, Japan) using the neighbor-joining method [47].

5. Conclusions

In this study, transcriptome analysis demonstrated that structural genes showed lower transcript levels in the white than the red coleoptiles in *T. monococcum*, which implied that a transcription factor was likely to be the key gene for determining white-colored coleoptiles. The transcript levels of *MYC* transcription factor *TmMYC-A1* were not significantly different between the white and red coleoptiles

and all TmMYC-A1s contained the complete functional domains in white and red coleoptiles. The MYB transcription factor TmMYB-A1, homologous to genes controlling coleoptile color in common wheat, contained the complete functional domain in the red coleoptiles, while TmMYB-a1 had lost the functional domains in white coleoptiles due to a frameshift mutation. Moreover, the transcript levels of *TmMYB-A1* in red coleoptiles were significantly higher than in white coleoptiles. Therefore, *TmMYB-a1* is likely to be the key gene that is responsible for the white coleoptile trait in *T. monococcum,* as suggested by the low levels of transcription and functional loss.

Supplementary Materials: The following are available online: Figure S1. Phylogenetic relationships of the deduced amino acid sequences of one PAL and other PAL involved in anthocyanin biosynthesis; Figure S2. Phylogenetic relationships of the deduced amino acid sequences of three 4CL and other 4CL involved in anthocyanin biosynthesis; Figure S3. Phylogenetic relationships of the deduced amino acid sequences of four ANR and other ANR involved in anthocyanin biosynthesis; Figure S4. Phylogenetic relationships of the deduced amino acid sequences of two C4H and other C4H involved in anthocyanin biosynthesis; Figure S5. Phylogenetic relationships of the deduced amino acid sequences of three CHI and other CHI involved in anthocyanin biosynthesis; Figure S6. Phylogenetic relationships of the deduced amino acid sequences of two CHS and other CHS involved in anthocyanin biosynthesis; Figure S7. Phylogenetic relationships of the deduced amino acid sequences of four DFR and other DFR involved in anthocyanin biosynthesis; Figure S8. Phylogenetic relationships of the deduced amino acid sequences of one F3′5′H and other F3′5′H involved in anthocyanin biosynthesis; Figure S9. Phylogenetic relationships of the deduced amino acid sequences of two F3H and other F3H involved in anthocyanin biosynthesis; Figure S10. Phylogenetic relationships of the deduced amino acid sequences of one F3′H and other F3′H involved in anthocyanin biosynthesis; Figure S11. Phylogenetic relationships of the deduced amino acid sequences of one HCT and other HCT involved in anthocyanin biosynthesis; Figure S12. Phylogenetic relationships of the deduced amino acid sequences of two LAR and other LAR involved in anthocyanin biosynthesis; Figure S13. Phylogenetic relationships of the deduced amino acid sequences of two LDOX and other LDOX involved in anthocyanin biosynthesis; Figure S14. Phylogenetic relationships of the deduced amino acid sequences of three UFGT and other UFGT involved in anthocyanin biosynthesis; Figure S15. Phylogenetic relationships of the deduced amino acid sequences of two FLS and other FLS involved in anthocyanin biosynthesis; Figure S16. Phylogenetic relationships of the deduced amino acid sequences of 3 MYCs and other MYCs involved in anthocyanin biosynthesis; Figure S17. Amino acid sequence alignment of MYC transcription factors from R1, R2, W1, W2, W3, and other species; Figure S18. Phylogenetic relationships between MYBs from R1, R2, W1, W2, W3, and other species; Table S1. The information of *T. monococcum* used in this study; Table S2. Summary of sequencing data; Table S3. The annotation summary of predicted protein numbers from various databases; Table S4. Information on unigenes associated with anthocyanin biosynthesis in coleoptiles.

Author Contributions: B.L. and H.Z. conceived and designed the experiments. D.C. and J.F. coordinated and performed the experiments. X.X., Y.Z., and D.W. analyzed the data. D.C. and B.L. wrote the paper. All authors read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shoeva, O.Y.; Gordeeva, E.I.; Arbuzova, V.S.; Khlestkina, E.K. Anthocyanins participate in protection of wheat seedlings from osmotic stress. *Cereal Res. Commun.* **2017**, *45*, 1–10. [CrossRef]

2. Chalker-Scott, L. Environmental significance of anthocyanins in plant stress responses. *Photochem. Photobiol.* **1999**, *70*, 1–9. [CrossRef]

3. Sarma, A.D.; Sharma, R. Anthocyanin-DNA copigmentation complex:mutual protection against oxidative damage. *Phytochemistry* **1999**, *52*, 1313–1318. [CrossRef]

4. Petroni, K.; Tonelli, C. Recent advances on the regulation ofanthocyanin synthesis in reproductive organs. *Plant Sci.* **2011**, *181*, 219–229. [CrossRef] [PubMed]

5. Bartłomiej, S.; Justyna, R.K.; Ewa, N. Bioactive compounds in cerealgrains—Occurrence, structure, technological significance and nutritionalbenefits—A review. *Food Sci. Tech. Int.* **2012**, *18*, 559–568. [CrossRef] [PubMed]

6. Winkel-Shirley, B. Flavonoid biosynthesis. A colorful model for genetics, biochemistry, cell biology, and biotechnology. *Plant Physiol.* **2001**, *126*, 485–493. [CrossRef] [PubMed]

7. Bhaskar, P.B.; Raasch, J.A.; Kramer, L.C.; Neumann, P.; Wielgus, S.M.; Austin-Phillips, S.; Jiang, J. *Sgt1*, but not *Rar1*, is essential for the RB-mediated broad-spectrum resistance to potato late blight. *BMC Plant Biol.* **2008**, *8*, 8. [CrossRef] [PubMed]

8. Koes, R.; Verweij, W.; Quattrocchio, F. Flavonoids: A colorful modelfor the regulation and evolution of biochemical pathways. *Trends Plant Sci.* **2005**, *10*, 236–242. [CrossRef] [PubMed]

9. Hichri, I.; Barrieu, F.; Bogs, J.; Kappel, C.; Delrot, S.; Lauvergeat, V. Recent advances in the transcriptional regulation of the flavonoidbiosynthetic pathway. *J. Exp. Bot.* **2011**, *62*, 2465–2483. [CrossRef] [PubMed]

10. Styles, E.D.; Ceska, O.; Seah, K.T. Developmental differences in action of *R* and *B* alleles in Maize. *Can. J. Genet. Cytal.* **1973**, *15*, 59–72. [CrossRef]

11. Atchley, W.R.; Wollenberg, K.R.; Fitch, W.M.; Terhalle, W.; Dress, A.W. Correlations among amino acid sites in bHLH protein domains: An information theoretic analysis. *Mol. Biol. Evolut.* **2000**, *17*, 164–178. [CrossRef] [PubMed]

12. Zhang, Y.; Butelli, E.; Martin, C. Engineering anthocyanin biosynthesis in plants. *Curr. Opin. Plant Biol.* **2014**, *19*, 81–90. [CrossRef] [PubMed]

13. Cheng, J.; Liao, L.; Zhou, H.; Gu, C.; Wang, L.; Han, Y. A small indelmutation in an anthocyanin transporter causes variegated colouration ofpeach flowers. *J. Exp. Bot.* **2015**, *66*, 7227–7239. [CrossRef] [PubMed]

14. Wang, Y.Q.; Hou, X.J.; Zhang, B.; Chen, W.J.; Liu, D.C.; Liu, B.L.; Zhang, H.G. Identification of a candidate gene for *Rc-D1*, a locus controlling red coleoptile colour in wheat. *Cereal Res. Commun.* **2016**, *44*, 35–46. [CrossRef]

15. Jiang, W.; Liu, T.; Nan, W.; Jeewani, D.C.; Niu, Y.; Li, C.; Wang, Y.; Shi, X.; Wang, C.; Wang, J.; et al. Two transcription factors tappm1 and tappb1 co-regulate the anthocyanin biosynthesis in purple pericarp of wheat. *J. Exp. Bot.* **2018**, *65*, 2555–2567. [CrossRef] [PubMed]

16. Cao, D.; Ye, G.J.; Zong, Y.; Zhang, B.; Chen, W.J.; Liu, B.L.; Zhang, H.G. *AetMYC1*, the candidate gene controlling the red coleoptile trait in *Aegilops tauschii* coss. accession as77. *Molecules* **2017**, *22*, 2259. [CrossRef] [PubMed]

17. Nesi, N.; Debeaujon, I.; Jond, C.; Pelletier, G.; Caboche, M.; Lepiniec, L. The TT8 gene encodes a basic helix-loop-helix domain protein required for expression of dfr and ban genes in Arabidopsis siliques. *Plant Cell* **2000**, *12*, 1863–1878. [CrossRef] [PubMed]

18. Spelt, C.; Quattrocchio, F.; Mol, J.N.; Koes, R. Anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. *Plant Cell* **2000**, *12*, 1619–1632. [CrossRef] [PubMed]

19. Li, N.; Li, S.M.; Zhang, K.P.; Chen, W.J.; Zhang, B.; Wang, D.W.; Liu, B.L.; Zhang, H.G. ThMYC4E, candidate blue aleurone 1 gene controlling the associated trait in Triticum aestivum. *PLoS ONE* **2017**, *12*, e0181116. [CrossRef] [PubMed]

20. Shoeva, O.; Gordeeva, E.; Khlestkina, E. The regulation of anthocyanin synthesis in the wheat pericarp. *Molecules* **2014**, *19*, 20266–20279. [CrossRef] [PubMed]

21. Borevitz, J.O.; Xia, Y.; Blount, J.; Dixon, R.A.; Lamb, C. Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* **2000**, *12*, 2383–2394. [CrossRef] [PubMed]

22. Quattrocchio, F.; Wing, J.; Woude, K.V.D.; Souer, E.; Vetten, N.D.; Mol, J.; Koes, R. Molecular analysis of the anthocyanin2 gene of petunia and its role in the evolution of flower color. *Plant Cell* **1999**, *11*, 1433–1444. [CrossRef] [PubMed]

23. Mano, H.; Ogasawara, F.; Sato, K.; Higo, H.; Minobe, Y. Isolation of a regulatory gene of anthocyanin biosynthesis in tuberous roots of purple-fleshed sweet potato. *Plant Physiol.* **2007**, *143*, 1252–1268. [CrossRef] [PubMed]

24. Peel, G.J.; Pang, Y.Z.; Modolo, L.V.; Dixon, R.A. The lap1 myb transcription factor orchestrates anthocyanidin biosynthesis and glycosylation in medicago. *Plant J.* **2010**, *59*, 136–149. [CrossRef] [PubMed]

25. Ban, Y.; Honda, C.; Hatsuyama, Y.; Igarashi, M.; Bessho, H.; Moriguchi, T. Isolation and functional analysis of a MYB transcription factor gene that is a key regulator for the development of red coloration in apple skin. *Plant Cell Physiol.* **2007**, *48*, 958–970. [CrossRef] [PubMed]

26. Takos, A.M.; Jaffe, F.W.; Jacob, S.R.; Bogs, J.; Robinson, S.P.; Walker, A.R. Light-induced expression of a MYB gene regulates anthocyanin biosynthesis in red apples. *Plant Physiol.* **2006**, *142*, 1216–1232. [CrossRef] [PubMed]

27. Huang, W.J.; Sun, W.; Lv, H.Y.; Luo, M.; Zeng, S.H.; Pattanaik, S.; Yuan, L.; Wang, Y. A R2R3-MYB transcription factor from epimedium sagittatum regulates the flavonoid biosynthetic pathway. *PLoS ONE* **2013**, *8*, e70778. [CrossRef] [PubMed]

28. Khlestkina, E.K.; Pestsova, E.G.; Roder, M.S.; Borner, A. Molecular mapping, phenotypic expression and geographical distribution of genes determining anthocyanin pigmentation of coleoptiles in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* **2002**, *104*, 632–637. [CrossRef] [PubMed]

29. Himi, E.; Taketa, S. Isolation of candidate genes for the barley *Ant1* and wheat *Rc* genes controlling anthocyanin pigmentation in different vegetative tissues. *Mol. Genet. Genom.* **2015**, *290*, 1287–1298. [CrossRef] [PubMed]

30. Ye, G.J.; Wei, L.; Chen, W.J.; Zhang, B.; Liu, B.L.; Zhang, H.G. Frame-shift mutation causes the function loss of *TaMYB-A1* regulating anthocyanin biosynthesis in Triticum aestivum. *Cereal Res. Commun.* **2017**, *45*, 35–46. [CrossRef]

31. Cao, D.; Ye, G.J.; Zong, Y.; Zhang, B.; Chen, W.J.; Liu, B.L.; Zhang, H.G. Diversity and distribution of *Rc1* alleles in *Triticum urartu* from the Middle East. *Genet. Resour. Crop. Evolut.* **2017**, *35*, 637–644. [CrossRef]

32. Kumar, R.; Ichihashi, Y.; Kimura, S.; Chitwood, D.H.; Headland, L.R.; Peng, J.; Maloof, J.N.; Sinha, N.R. A high-throughput method for Illumina RNA-seq library preparation. *Front. Plant Sci.* **2012**, *3*, 202. [CrossRef] [PubMed]

33. Ozsolak, F.; Milos, P.M. RNA sequencing: Advances, challenges and opportunities. *Nat. Rev. Genet.* **2011**, *12*, 87–98. [CrossRef] [PubMed]

34. Liu, D.; Li, S.M.; Chen, W.J.; Zhang, B.; Liu, D.C.; Liu, B.L.; Zhang, H.G. Transcriptome analysis of purple pericarps in common wheat (*Triticum aestivum* L.). *PLoS ONE* **2016**, *11*, e0155428. [CrossRef] [PubMed]

35. Luo, J.R.; Duan, J.J.; Huo, D.; Shi, Q.Q.; Niu, L.X.; Zhang, Y.L. Transcriptomic analysis reveals transcription factors related to leaf anthocyanin biosynthesis in paeonia qiui. *Molecules* **2017**, *22*, 2186. [CrossRef] [PubMed]

36. Li, Y.Z.; Luo, X.; Wu, C.Y.; Cao, S.Y.; Zhou, Y.F.; Jie, B.; Cao, Y.L.; Meng, H.J.; Wu, G.L. Comparative transcriptome analysis of genes involved in anthocyanin biosynthesis in red and green walnut (*juglans regia* L.). *Molecules* **2017**, *23*, 25. [CrossRef] [PubMed]

37. Holton, T.A.; Cornish, E.C. Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* **1995**, *7*, 1071–1083. [CrossRef] [PubMed]

38. Himi, E.; Maekawa, M.; Miura, H.; Noda, K. Development of pcr markers for TaMYB10 related tor-1, red grain color gene in wheat. *Theor. Appl. Genet.* **2011**, *122*, 1561–1576. [CrossRef] [PubMed]

39. Strygina, K.V.; Khlestkina, E.K. *MYC* gene family in cereals: Transformations during evolution of hexaploid bread wheat and its relatives. *Mol. Biol.* **2017**, *51*, 772–779. [CrossRef]

40. Wang, C.; Shu, Q. Fine mapping and candidate gene analysis of purple pericarp gene *Pb* in rice (*Oryza sativa* L.). *Chin. Sci. Bull.* **2007**, *52*, 3097–3104. [CrossRef]

41. Qiu, Z.; Wang, X.; Gao, J.; Guo, Y.; Huang, Z.; Du, Y. The *tomatohoffman's anthocyaninless* gene encodes a bHLH transcription factor involved in anthocyanin biosynthesis that is developmentally regulated and induced by low temperatures. *PLoS ONE* **2016**, *11*, e0151067. [CrossRef] [PubMed]

42. Ogbonnaya, F.C.; Abdalla, O.; Mujeeb-Kazi, A.; Kazi, A.G.; Xu, S.S.; Gosman, N.; Lagudah, E.S.; Bonnett, D.; Sorrells, M.E.; Tsujimoto, H. Synthetic hexaploids: Harnessing species of the primary gene pool for wheat improvement. *J. Plant Breed. Rev.* **2013**, *37*, 35–122.

43. United States Department of Agriculture Agricultural Research Service. Available online: https://www.ars-grin.gov/npgs/sitelist.html (accessed on 18 September 2015).

44. Grabherr, M.G.; Haas, B.J.; Yassour, M.; Levin, J.Z.; Thompson, D.A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q. Trinity: Reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* **2011**, *29*, 644–652. [CrossRef] [PubMed]

45. Romualdi, C.; Bortoluzzi, S.; D'Alessi, F.; Danieli, G.A. Ideg6: A web tool for detection of differentially expressed genes in multiple tag sampling experiments. *Physiol. Genom.* **2003**, *12*, 159–162. [CrossRef] [PubMed]

46. Kyoto Encyclopedia of Genes and Genomes. Available online: http://www.genome.jp/kegg/ (accessed on 3 December 2018).

47. Tamura, K.; Peterson, D.; Peterson, N.; Stecher, G.; Nei, M.; Kumar, S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evolut.* **2011**, *28*, 2731–2739. [CrossRef] [PubMed]

Sample Availability: Samples of DNA of all *Triticum monococcum* accessions are available from the authors.