

RESEARCH ARTICLE

Open Access



RNA sequencing and anthocyanin synthesis-related genes expression analyses in white-fruited *Vaccinium uliginosum*

Yang Yang, Baihui Cui, Zhiwen Tan, Bingxue Song, Hounan Cao and Chengwen Zong* 

Abstract

Background: *Vaccinium uliginosum* (Ericaceae) is an important wild berry having high economic value. The white-fruited *V. uliginosum* variety found in the wild lacks anthocyanin and bears silvery white fruits. Hence, it is a good resource for investigating the mechanism of fruit color development. This study aimed to verify the differences in the expression levels of some structural genes and transcription factors affecting the anthocyanin biosynthesis pathway by conducting high-throughput transcriptome sequencing and real-time PCR analysis by using the ripening fruits of *V. uliginosum* and the white-fruited variety.

Results: We annotated 42,837 unigenes. Of the 325 differentially expressed genes, 41 were up-regulated and 284 were down-regulated. Further, 11 structural genes of the flavonoid pathway were up-regulated, whereas two were down-regulated. Of the seven genes encoding transcription factors, five were up-regulated and two were down-regulated. The structural genes *VuCHS*, *VuF3'H*, *VuFHT*, *VuDFR*, *VuANS*, *VuANR*, and *VuUFGT* and the transcription factors *VubHLH92*, *VuMYB6*, *VuMYBPA1*, *VuMYB11*, and *VuMYB12* were significantly down-regulated. However, the expression of only *VuMYB6* and *VuMYBPA1* rapidly increased during the last two stages of *V. uliginosum* when the fruit was ripening, consistent with anthocyanin accumulation.

Conclusions: *VuMYB6* was annotated as *MYB1* by the BLAST tool. Thus, the white fruit color in the *V. uliginosum* variant can be attributed to the down-regulation of transcription factors *VuMYB1* and *VuMYBPA1*, which leads to the down-regulation of structural genes associated with the anthocyanin synthesis pathway.

Keywords: *Vaccinium uliginosum*, White-fruited variant, RNA-seq, qRT-PCR

Background

Vaccinium uliginosum is a perennial deciduous shrub of the Ericaceae family. It is distributed in the northeast regions of China, including Xiao Xing'an Mountains, Da Xing'an Mountains, Inner Mongolia, and Changbai Mountain forest area, it also occurs in Sphagnum swamp meadow at an elevation of more than 700 m, along with a large community of other plant species such as *Carex*, *Larix gmelinii*, *Betula ovalifolia*, and *Ledum palustre*. This species can be found at temperatures as low as -40°C to 50°C [1]. The berries are fragrant, delicious, and nutritious and can be consumed either raw or after processing [2]. Although *V. uliginosum* is a very valuable species, it has not been intensively investigated. *V. uliginosum* berries

contain abundant amino acids, trace elements, anthocyanins, procyanidins and other polyphenols [3]. They have various beneficial health effects such as blood vessels softening, disease prevention and health care.

A *V. uliginosum* variety with white berries has been reported, but no further studies on this species were performed [4, 5]. During resource investigation, we found a white-fruited *V. uliginosum* variety in Wangqing County of Jilin Province in the Lanjia forest farm. They were sporadically distributed in the wild community. The berries were round or oblong, and the fruit weight was approximately 1–1.5 g, the ripening berries were silvery white and translucent. Preliminary investigation revealed that the berries lacked anthocyanins, and their total phenolic and flavonoid content was lower [6], whereas the vitamin C and titratable acid contents were higher than those of the wild-type *V. uliginosum*. During the

* Correspondence: zongchengwen@aliyun.com
Agriculture College of Yanbian University, Yanji, Jilin 133002, China



survey, we also found a chimera with blue and white berries on the same plant during the survey (Fig. 1). This phenotypic variation might have been caused by somatic mutation. This resource has a potential value in theoretical research and application for elucidating the mechanism of the synthesis and regulation of anthocyanidin. The white-fruited *V. uliginosum* variety produces tastier and sweeter berries with low tannin content than those of the wild type. This might reduce the processing and marketing cost of the berries [5].

Anthocyanin is a flavonoid produced via the flavonoid biosynthesis pathway during plant metabolism. It is an important water-soluble, nature, edible pigment, which exists widely in the vacuoles of plant epidermal cells and imparts orange and red to blue color to flowers, fruits, stems, leaves, and roots [7]. It is safe, non-toxic, and abundant, with important nutritional and pharmacological effects, such as antioxidant capacity, and ability to scavenge free radicals, prevent cardiovascular diseases, tumors, mutation, and radiation, regulate the activity of platelets, prevent platelet condensation, and induce immunomodulatory activity. In addition, it helps in improving the cold and drought resistance ability of plants [8].

The anthocyanin biosynthetic pathway is most extensively studied secondary metabolic pathway in plants studied most extensively, particularly in *Arabidopsis*, *Zea* and *Petunia*. Anthocyanin biosynthesis can be divided into two stages. First, phenylalanine transforms into 4-coumaryl: CoA referred to as the phenylpropanoid metabolic pathway. Subsequently, 4-coumaryl: CoA transforms into all kinds of flavonoid compounds, which is referred to as the anthocyanin biosynthetic pathway. Anthocyanin biosynthesis involves various structural and regulatory genes [9].

The genes involved in anthocyanin biosynthesis can be classified into two major groups. The first group consists of structural genes that directly encode the key enzymes in anthocyanin biosynthesis. The other group is transcriptional factors. Transcriptional regulation is an important aspect of

the regulation of gene expression in the anthocyanin biosynthetic pathway of plant. Its mechanism is very complicated. At present, three main types of regulatory factors have been identified: MYB, bHLH and WD40 transcription factors [10]. Anthocyanin biosynthesis in most of the species is regulated by the transcription factors that form a protein complex that binds to the promoter of the structural genes. The MADS-box transcription factor *VmTDR4* in *V. myrtillus* was suggested to play an important role in the synthesis of anthocyanins by the direct or indirect regulation of MYB transcription factors [11]. Moreover, the microRNA miR156 and its target gene *SPL3* decrease anthocyanin biosynthesis in *Arabidopsis thaliana* [12].

Fruit color is an important factor affecting the appearance and quality of the fruit. The study of the mechanism and regulation of fruit coloration is very important for elucidating somatic cell mutations in the peel. Mutations affecting fruit coloration, have been investigated in the model plants such as grape [13, 14], *V. myrtillus* [15], *Duchesnea indica* [16], and *Syzygium malaccense* [17]. The variations in the regulatory or structural genes in the anthocyanin biosynthesis pathway of these species are responsible for the mutations.

However, few studies have investigated the variations in the molecular mechanism of anthocyanin biosynthesis in *V. uliginosum*. This study aimed to analyze the differences among wild *V. uliginosum* and fruit color varieties at the molecular level by using transcriptome sequencing analysis. In addition we conducted a bioinformatics analysis and quantitative polymerase chain reaction (q-PCR) to identify the genes related to anthocyanin biosynthesis pathway and glucose metabolism. Our findings might reveal the main reason for the lack of anthocyanins in the *V. uliginosum* mutant.

Results

Transcriptome sequencing and assembly

In order to investigate the mutation mechanism of the white-fruited variant, we repeated the transcriptome



Fig. 1 The ripening white *Vaccinium uliginosum* variety and the chimera at the experimental site. The site is in Wangqing County of Jilin Province Lanjia forest farm

sequencing analysis thrice. We obtained 34.62Gb of Clean Data, which reached the 5.3Gb threshold for each sample and Q30 base percentage was $\geq 88.73\%$. The Read Number was more than 21,044,015, and the base number for each sample was not less than 5,303,019,780. The *V. uliginosum* GC content was in the range of 46.69–47.17%, whereas that of the white-fruited variant was in the range of 47.17–47.46% (Table 1).

The assembled sequences showed 225,777,919 transcripts items and 89,725 unigenes. The N50 value of the transcripts and unigenes were 1058.35 and 1094, respectively. This showed that the assembly had high integrity (Table 2). The number of contigs with 200–300 nt was the highest (27,838,434; 99.62%) and that of contigs with length of >2000 nt was the lowest (6441; 0.02%). In addition, the highest number of transcripts had length in the range of 200–300 nt, and the lowest number of transcripts had 1000–2000 nt (27,320; 12.81%). However, no unigenes had length in the range of 200–300 nt and the highest number of unigenes had length of 300–500 nt (42,605 accounting for 47.48% of the total).

Analysis by sequence alignment revealed approximately 18,355,782 mapped reads in *V. uliginosum*, accounting for 78.77% of the total clean reads. The number of mapped reads of the white-fruited variant was approximately 17,688,233 (78.58%). The mapped reads were not significantly different between the two samples (Table 3).

Functional annotation of unigenes

A total of 89,725 unigenes were annotated from different databases, with a total of 42,837 unigenes with functional annotations. Unigenes with a length of 300–1000 nt and ≥ 1000 nt were 26,734 (62.41%) and 16,103 (37.59%), respectively. Analysis by sequence alignment revealed that the maximum number of unigenes were annotated in NR (40,121, accounting for 93.66% of all the annotations). Whereas, the least number were annotated in COG (12,991, accounting for 30.33% of the total). The remaining unigenes were annotated in GO, KEGG, KOG, Pfam, and Swiss-prot databases (Table 4).

Table 1 Statistical evaluation of sequencing data

Samples	Read Number	Base Number	GC Content	Q30 (\geq %)
A1	21,137,847	5,326,737,444	46.78%	88.73%
A2	23,116,662	5,825,398,824	46.69%	90.89%
A3	25,635,645	6,460,182,540	47.17%	90.61%
B1	24,328,728	6,130,839,456	47.36%	90.85%
B2	22,137,140	5,578,559,280	47.46%	90.61%
B3	21,044,015	5,303,091,780	47.17%	90.97%

Note: A1, A2, A3 is the sample of *Vaccinium uliginosum*. B1, B2, B3 is the sample of its white-fruited variant

Table 2 Assembly results

Length Range	Contig	Transcript	Unigene
200–300	27,732,934(99.62%)*	68,779(32.24%)	0%
300–500	52,388(0.19%)	65,549(30.73%)	42,605(47.48%)
500–1000	32,032(0.12%)	51,682(24.23%)	26,647(29.70%)
1000–2000	14,639(0.05%)	27,320(12.81%)	13,722(15.29%)
2000+	6441(0.02%)	213,330	6751(7.52%)
Total Number	27,838,434	225,777,919	89,725
Total Length	1,139,232,020	1516	73,630,818
N50 Length	43	1058.35	1094
Mean Length	40.92		820.63

Screening of differentially expressed genes

For differential expression analysis, we corrected the significance value of the original hypothesis of the accepted effective Benjamini–Hochberg method. Finally, we adopted the corrected *P* value named false discovery rate (FDR) as the key indicator of differential gene expression, which can help to reduce the false positive rate by considering the expression levels of numerous of genes in an independent statistical hypothesis test.

A total of 325 differentially expressed genes (DEGs) were obtained, among which 41 were up-regulated and 284 genes were down-regulated (Fig. 2). A total of 118,804 genes showed no significant differences in expression level.

Of the 41 genes of up-regulated genes, 28 and 13 DEGs annotated and unannotated, respectively. Unannotated genes might be the new transcripts of *V. uliginosum*. The expression level of *F3'5'H* (c101538) and *DFR* (c130571) genes of the flavonoid biosynthesis pathway was high (6.1665 and 2.5403, respectively). The c74288 gene expressed in the white-fruited variant, but not in *V. uliginosum*; this could be attributed to the difference in aromatic compound metabolism (Additional file 1: Table S1).

Of the 284 down-regulated genes, 148 and 136 DEGs were annotated and unannotated, respectively. Moreover, 71 genes were expressed in *V. uliginosum*, but not in the white-fruited variant. The c17593 gene related to transcription factor B3, c114838 gene related to MADs-box transcription factor, and c86331 gene related to transcription factor *MYB3* were down-regulated in the white-fruited

Table 3 Comparison of the results of sequencing and assembly

Number	Clean Reads	Mapped Reads	Mapped Ratio
A1	21,137,847	16,607,615	78.57%
A2	23,116,662	18,181,365	78.65%
A3	25,635,645	20,278,365	79.10%
B1	24,328,728	19,199,505	78.92%
B2	22,137,140	17,420,598	78.69%
B3	21,044,015	16,444,595	78.14%

Table 4 Unigene annotation in different databases

Anno_Database	Annotated_Number	300 ≤ length ≤ 1000	Length ≥ 1000
COG_Annotation	12,991	6985	6006
GO_Annotation	24,469	15,061	9408
KEGG_Annotation	17,732	11,297	6435
KOG_Annotation	25,080	15,162	9918
Pfam_Annotation	29,299	15,609	13,690
Swissprot_Annotation	27,560	16,170	11,390
nr_Annotation	40,121	25,211	14,910
All_Annotated	42,837	26,734	16,103

variant, with expression fold values of -5.1825, -4.7228, and -6.6772, respectively. The gene c89580 related to sugar metabolism was not expressed in the white-fruited variant, but was expressed in *V. uliginosum*. In addition, the c116690 gene was down-regulated in the white-fruited

variant and its expression fold value was -2.9947 (Additional file 2: Table S2).

Functional annotation of DEGs

In all, 176 annotated DEGs were different between *V. uliginosum* and its white-fruit variant among the different databases. In the NR database, a maximum of 159 DEGs were annotated. In addition, a minimum of 39 DEGs were annotated in the COG database (Table 5).

GO function enrichment of DEGs

GO function enrichment revealed 82 annotated DEGs (Fig. 3). They participate in three major aspects including the biological process, molecular function and cellular components. In all, 88 GO terms were enriched. Of them, 47 were significantly enriched ($P \leq 0.05$). Moreover GO: 0009773 was the most significantly associated ($P = 0.0002$) with the light system I photoelectron transfer. The GO: 0009058 node was related to the biosynthesis process, and two DEGs were associated with this node one is sucrose synthase (c122231), the FPKM value of which was four fold lower in *V. uliginosum* and the other is c134490 of the glycosyl transferase group which was not expressed in the white-fruited variant. However, its FPKM value was 10.7933 in *V. uliginosum* (Additional file 3: Table S3).

Enrichment analysis of DEGs in the KEGG pathway

In living organisms, different gene products coordinate to perform different biological functions. Pathway annotation analysis of DEGs is helpful for further understanding the function of genes. KEGG database is the main public database on pathways. We analyzed of whether the appearance of DEGs in a pathway reflected pathway enrichment. KEGG pathway enrichment analysis revealed 42 annotated DEGs and 15 pathways. These

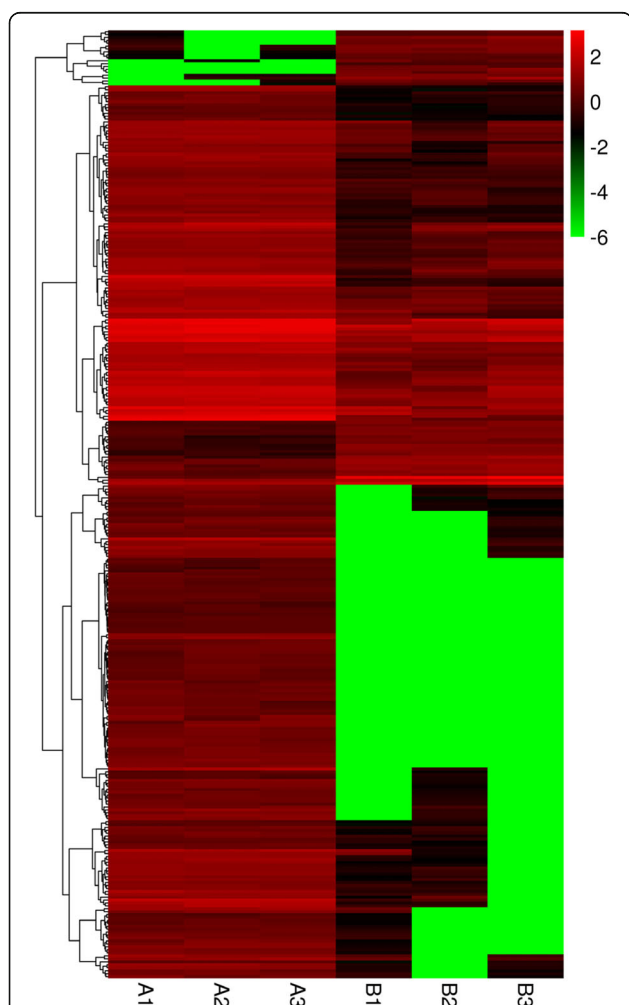


Fig. 2 Differential expression gene expression pattern clustering map. A1, A2, A3: sample of *V. uliginosum* ripening fruits. B1, B2, B3: sample of white-fruited variant ripening fruits

Table 5 The number of annotated differential expression genes

DEG Set	Annotated	COG	GO	KEGG	KOG	Pfam	Swiss-Prot	nr
A vs B	176	39	82	50	94	145	114	159

pathways involve five aspects, including cellular processes, genetic information processing, human diseases, metabolism, and organismal systems.

Among them, the number of unigenes and DEGs involved in the phenylpropanoid biosynthesis pathway was 240 and 5, respectively. The flavonoid biosynthesis pathway had 80 unigenes, and three DEGs. In addition, the numbers of unigenes and DEGs in starch and sucrose metabolism pathways were 390 and two, respectively (Additional file 4: Table S4).

Anthocyanin synthesis-related genes expression

The color of fruit peel is known to be correlated with the expression of anthocyanin synthesis related genes. Our expression analysis results showed that the expression of structural genes was mostly down-regulated in the white-fruited variant. However the expression of transcription factors in the anthocyanin biosynthesis pathway and genes in the sugar metabolism pathway were mostly up-regulated (Fig. 4). The structural genes peroxidase (*VuPOD*, c132702), cinnamoyl-CoA reductase (*VuCCR*,

c104371), and *VuCHI* (c126113) related to anthocyanin biosynthesis were significantly up-regulated in the white-fruited variety. In contrast, the genes *VuCHS* (c127976), flavanone-3 β -hydroxylase (*VuFHT*), flavonoid 3'-hydroxylase (*VuF3'H*, c123712), *VuDFR* (c130571), *VuANS* (c122374), anthocyanidin reductase (*VuANR*, c128846), and *VuUFGT* (c127140) were significantly down-regulated. However, the expression of flavonoid 3', 5'-hydroxylase (*VuF3'5'H*, c132630) and leucoanthocyanins reductase (*VuLAR*, c124332) was not significantly different between the white-fruited variant and wild-type *V. uliginosum*. The transcription factors were in the most up-regulated in the white-fruited variant. The genes *VubHLH63* (c80718), and *VuTDR4* and transcription factors *VubHLH130* (c113885), *VuMADS-box* (c114838) and transcription factor B3 (c117593) were significantly up-regulated in the white-fruited variant. However the expression of *VubHLH93* (c90489) was not significantly different from that of *V. uliginosum*. In the white-fruited variant, the *VubHLH92* (c112979) gene was significantly down-regulated. Among the MYB transcription factors,

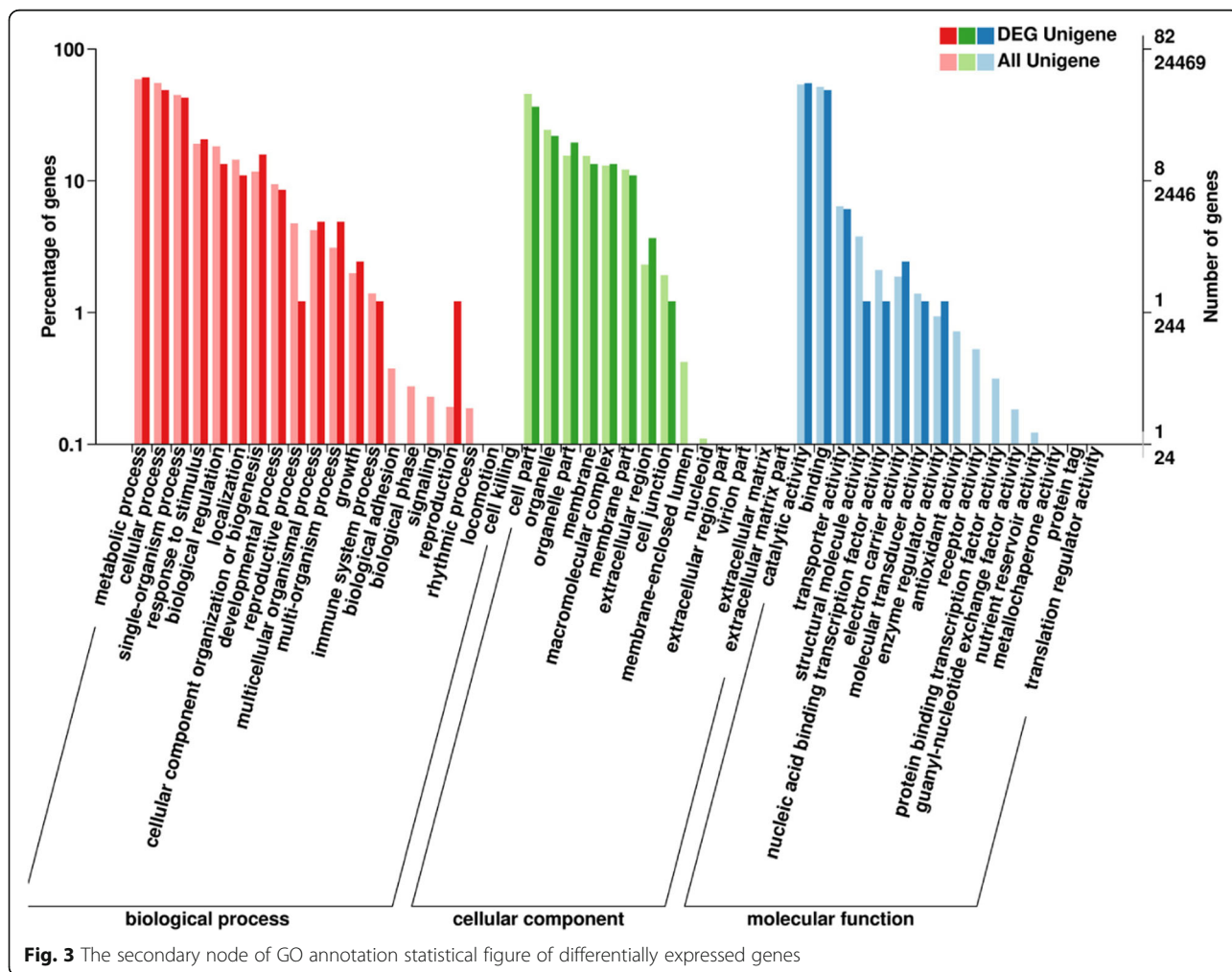


Fig. 3 The secondary node of GO annotation statistical figure of differentially expressed genes

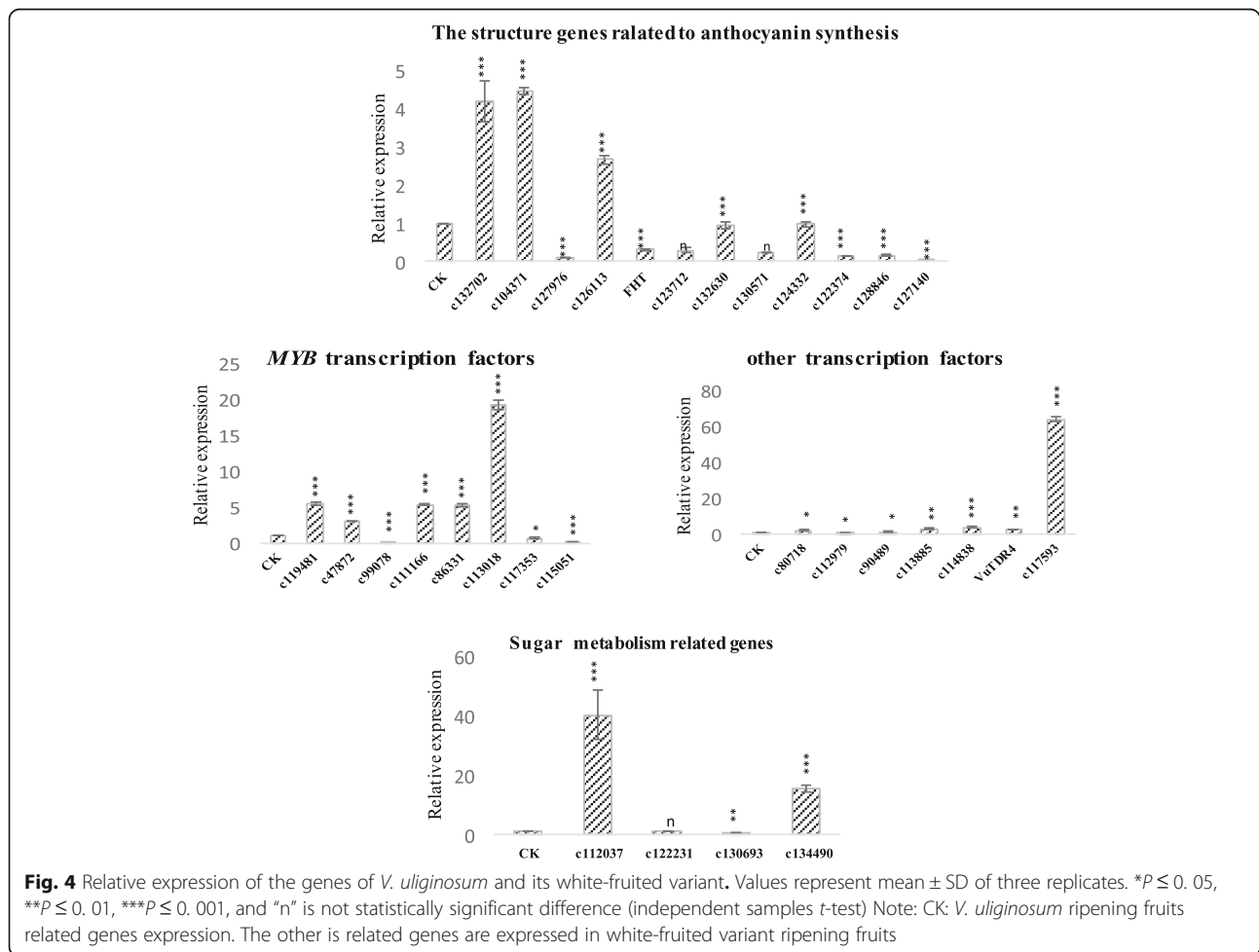
VuMYB2 (c119481), *VuMYB4* (c47872), *VuMYB7* (c111166), *VuMYB8* (c86331), and *VuMYB10* (c113018) were significantly up-regulated. The transcription factors *VuMYB6* (c99078) and *VuMYBPA1* (c115051) were significantly down-regulated. Moreover, the *VuMYB12* (c117353) transcription factor was significantly down-regulated. In the white-fruited variant sugar metabolism, beta-glucosidase (c112037) and glycosyl transferases group 1 (c134490) were significantly up-regulated, and the sugar (and other) transporter (c130693) was significantly down-regulated. The sucrose synthase (c122231) expression was not significantly different from that of *V. uliginosum*.

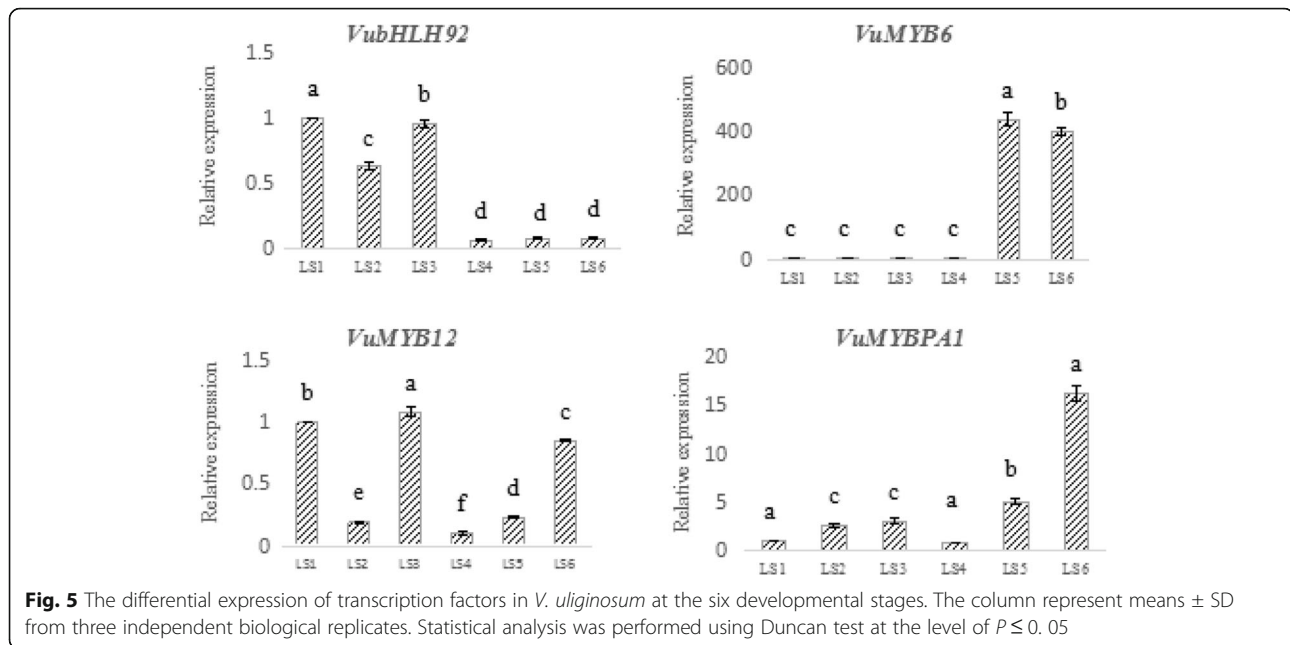
The expression of differentially expressed transcription factors at different developmental stages in *V. uliginosum*

Four transcription factors, *VubHLH92* (c112979), *VuMYB6* (c99078), *VuMYBPA1* (c115051), and *VuMYB12* (c117353), were expressed in the white-fruited variant, but significantly down-regulated in *V. uliginosum*.

Analysis of the expression of these four transcription factors at different developmental stages of *V.*

uliginosum showed that the expression of *VubHLH92* was first decreased and then increased (Fig. 5). When the fruit was about to ripen, almost no *VubHLH92* expression was noted. The expression of the *VuMYB6* was almost the same during the first four stages, and then the expression suddenly and rapidly increased when the fruit was ripening, to approximately 400 times of that during the first period, consistent with the accumulation pattern of anthocyanin. Evidently, this transcription factor is related to the synthesis of anthocyanin in *V. uliginosum*. The *VuMYB12* transcription factor expression was first decreased and then increased, followed by a decline. The expression began to increase when the fruit was ripening, but its expression was lower than that during the first and third periods. The expression of *VuMYBPA1* transcription factor first increased gradually and decreased in the fourth period, and increased rapidly in the last two periods. The expression ratio of the last period was approximately 16 times higher than that for the first period, consistent with the accumulation of anthocyanin.





Discussion

As early as 1986, Zhang et al. [5] identified a variant of *V. uliginosum* in Heihe City, Heilongjiang Province. They found that the sugar content of the white-fruited variety was higher than that of the blue-fruited wild-type one. Moreover, Ma et al. [4] reported that the white-fruited variant was slightly sweeter than the original variety from the Da Xing'an Mountains of China. These findings might explain why most of the genes related to sugar metabolism were up-regulated in the white-fruited variant. The expression of beta-glucosidase (c112037) was significantly up-regulated, probably because glucose was the principle sugar in the fruits of the white-fruited variant. Glycosyltransferase is the key enzyme associated with the catalytic synthesis of sugar chains; thus, the extremely significant up-regulation of glycosyl transferases group 1 (c134490) was reasonable. Moreover the sucrose is not primarily responsible for the sweetness of the white-fruited variety; thus, sucrose synthase (c122231) expression was not significantly different from that in *V. uliginosum*. After sugar formation in ripe fruits, the expression of sugar (and other) transporter (c130693) was significantly down-regulated in the white-fruited variant.

According to our survey, the morphological characteristics of the white-fruited variant did not differ from those of *V. uliginosum*, except that the ripe fruit was silver-white and lacked anthocyanin. This is not caused by the external environment, as the silver-white coloration does not change even in winter [4]. The formation of fruit peel color is closely related to the synthesis of anthocyanins, and its mechanism is complex. The expression of structural genes in the anthocyanin biosynthetic pathway is

generally required for upstream transcription factor binding to initiate gene transcription and expression, to finally synthesize anthocyanins and impart color to the fruit and peel [18].

The expression of structural genes *VuF3' 5' H* (c132630) and *VuLAR* (c124332) of the anthocyanin biosynthesis pathway in the white-fruited variant did not differ significantly from that in *V. uliginosum*. These genes do not direct effects on anthocyanin synthesis. The expression of *VuCHS* (c127976), *VuFHT*, *VuF3'H* (c123712), *VuDFR* (c130571), *VuANS* (c122374), *VuANR* (c128846) and *VuUFGT* (c127140) genes was extremely significantly down-regulated in the white-fruited variety their expression is closely linked to anthocyanin synthesis. This is probably because of the inhibition of transcription factor expression, which can affect the accumulation of anthocyanin.

Transcription factor B3 is unique to plants; it is a super-family existing mainly in gymnosperms, mosses, and algae and other plants. Luo et al. found that this transcription factor B3 plays an important role in the plant stress response and in plant growth and development [19]. Zhang et al. [20] found that transcription factor B3 plays a key role in the development and maturation of the embryo in *Theobroma cacao*. In *Arabidopsis*, the excessive expression of the transcription factor B3 resulted in the loss of petals [21]. At present, no studies have shown a relationship between transcription factor B3 and coloring. The expression of transcription factor B3 in the white-fruited variant was extremely significantly up-regulated; therefore, it might play a role in other aspects of development.

The *MADS-box* and *TDR4* genes encode transcription factors, and the *TDR4* gene is a member of the *MADS*

gene family [11]. Previous studies have showed that *MADS* and *TDR4* genes were involved in plant anthocyanin synthesis. Another study showed that *IbMADS10* was involved in the synthesis of anthocyanin in *Ipomoea batatas* [22]. In the *V. myrtillus* mutant, down-regulated expression of the *VmTDR4* gene directly or indirectly inhibited the expression of the *VmMYB2* gene, resulting in the down-regulation of the structural genes of the anthocyanin biosynthesis pathway, inhibiting the synthesis of anthocyanin [11]. The MADS super family has many members, and each member has its own function. We verified that the expression of the *VuMADS-box* (c114838) gene was extremely significantly up-regulated and that of the *VuTDR4* gene was significantly up-regulated in the white-fruited variety. Therefore, they might have no relation with anthocyanin synthesis.

At present, studies on anthocyanin synthesis in plants suggest that the known transcription factors can be mainly divided into three categories, namely the R2R3-MYB type, bHLH and WD40 complex protein transcription factors [10]. Some of these are single gene regulatory structural genes, some are two genes whose interaction controls anthocyanin synthesis, and some are MBW (MYB bHLH WD40) complexes regulating anthocyanin accumulation. Park et al. [23] revealed that the *bHLH2* gene contained a DNA transposon resulting in a change in the color of the petals of *Pharbitis purpurea* flowers. A single gene mutation can also affect anthocyanin accumulation. Qian et al. [24] found that *PyMYB10* regulated the synthesis and expression of anthocyanin in *Pyrus pyrifolia*. Azuma et al. [25] confirmed that *ViMYBA1-3* is a key gene that influences anthocyanin biosynthesis in grape peel. Aharoni et al. [26] reported that yellow strawberry *FaMYB1* inhibited the accumulation of anthocyanin and flavonoids in transgenic tobacco. Primetta et al. [27] showed that the homologous *MYBPA1* and *MYB2* are the members of the R2R3MYB family, and their down-regulation inhibited the expression of *CHS*, *DFR*, *LAR*, *ANR*, *ANS* and *UFGT* structural genes, affecting the accumulation of anthocyanin. Single-gene transcription factor can also directly regulate anthocyanin synthesis. Wada et al. [28] reported that *Arabidopsis* MYB (*CPC*) and bHLH (*GL3*) interact with each other to influence bHLH anthocyanin pigment synthesis in *Lycopersicon esculentum*. Schwinn et al. [29] reported that MYB and bHLH transcription factors increased the intensity of the color of *Petunia* and *Platycodon*, and strongly enhanced the phenotype of *Petunia*. Liu et al. [30] reported that *MrbHLH1* and *MrMYB1* affected anthocyanin synthesis in tobacco and *Myrica rubra*. Two transcription factors can also interact to regulate the accumulation of anthocyanin. Xie et al. [31] showed that *MdbHLH3* regulated *MdMYB1* expression to mediate the low-temperature-induced anthocyanin accumulation and coloration of apples. A transcription factor can also be

regulated by another transcription factor to affect the biosynthesis of anthocyanin. Albert et al. [32] confirmed that the MBW complex transcription factor activated a single transcription factor and then the R2R3-MYB type transcription factors *TrMYB133* and *RrMYB134* regulated the biosynthesis of anthocyanin and procyanidin in *Trifolium repens*. Thus, a single transcription factor can thus be regulated by a complex of transcription factors to induce the synthesis of anthocyanin.

Primetta et al. [27] used *V. uliginosum* and the white-fruited variant from Finland as study material and indicated that the expression of structural genes *CHS*, *DFR*, *LAR*, *ANR*, *ANS* and *UFGT*, related to the anthocyanin synthesis pathway, was significantly down-regulated, and the expression of *CHI* and *F3'5'H* genes did not differ significantly from those in *V. uliginosum*. However, in our study, the expression of *VuCHS*, *VuFHT*, *VuF3'H*, *VuDFR*, *VuANS*, and *VuUFGT* was significantly down-regulated in the white-fruited variety, and the expression of the *VuF3'5'H* and *VuLAR* genes did not differ significantly from that in *V. uliginosum*. Nevertheless, the expression of *VuCHI* was significantly up-regulated. The results of our study and those of Primetta et al. differ to some extent owing to the fact that mutants are formed differently because of the difference geographical regions.

In this study, high-throughput transcriptome sequencing was used to identify DEGs, the expression of which was verified to differ significantly between wild-type *V. uliginosum* and the white-fruited variant using qRT-PCR. Moreover, the expression of four transcription factors related to anthocyanin biosynthesis was significantly down-regulated, namely *VubHLH92* (c112979), *VuMYB6* (c99078), *VuMYBPA1* (c115051), and *VuMYB12* (c117353). The expression of only *VuMYB6* and *VuMYBPA1* increased rapidly in ripening fruit, among the six different developmental stages of *V. uliginosum*, consistent with the accumulation of anthocyanin, whereas the expression of other transcription factors did not increase during the anthocyanin accumulation period, indicating they had no direct relationship with anthocyanin synthesis. *VuMYBB6* was annotated to *MYB1* by BLAST sequence alignment analysis, and *MYB1* and *MYBPA1* were found to both be R2R3-MYB type transcription factors. The significant down-regulation of the expression of *VuMYB1* and *VuMYBPA1* inhibited the expression of structural genes *VuCHS*, *VuFHT*, *VuF3'H*, *VuDFR*, *VuANS*, and *VuUFGT*, affecting the anthocyanin accumulation in the white-fruited variant. This study provided another possibility for the mechanism of mutation in *V. uliginosum*. We intend to verify the specific regulatory mechanism in the future.

Conclusions

According to the transcriptome analysis of DEGs related to anthocyanin synthesis pathway in the ripening fruits of

V. uliginosum and white-fruited variant, the down-regulated expression of transcription factors *VubHLH92*, *VuMYB6*, *VuMYB12*, and *VuMYBPA1* might lead to the down-regulation of the structural genes related to the anthocyanin synthesis pathway.

The results of the expression of these four transcription factors in different developmental stages of *V. uliginosum* showed that only the expression of *VuMYB6* and *VuMYBPA1* is related to the accumulation of anthocyanins.

VuMYB6 was annotated as *MYB1* by the BLAST tool. Therefore, the down-regulation of the transcription factors *VuMYB1* and *VuMYBPA1* leads to the down-regulation of the structural genes associated with the anthocyanin synthesis pathway, which is the main reason for the lack of anthocyanins in the white berries.

Materials and methods

Plant materials

The ripening fruits (60 days after full-bloom stage) of *V. uliginosum* and its variant for RNA-seq were collected from Lanjia forest farm, Wangqing Contry, Jilin Province, China, and stored at -80°C .

Different developmental stages of *V. uliginosum* materials for q-PCR were also collected: blooming flowers (LS1), fruits with prominent ovarian enlargement 15 days after full-bloom stage (LS2), 0.5–0.7 cm fruits 30 days after full-bloom stage (LS3), 0.7–1.0 cm fruits 40 days after full-bloom stage (LS4), fruits at the stage when color changes 50 days after full-bloom stage (LS5), and the ripening fruits 60 days after full-bloom stage (LS6).

Methods

Extraction and detection of total RNA

The total RNA was isolated from the ripening fruits of *V. uliginosum* and its variant by using the RNA rapid extraction kit (Hua Yueyang, Beijing, China). Spectrophotometry method was used to detect the purity, concentration and integrity of RNA samples to ensure that the high-quality samples were used for RNA-seq.

RNA was also isolated from the materials collected at different developmental stages of *V. uliginosum* by using the RNA rapid extraction kit. The first chain of cDNA was synthesized using a reverse transcription Kit (Takara, Dalian, China) for q-PCR.

cDNA library, deep transcriptome sequencing and assembly

Enrichment of *V. uliginosum* and its variant mRNA with magnetic beads with Oligo (dT), randomly interrupt mRNA by adding Fragmentation Buffer. With the mRNA from the fruits *V. uliginosum* and its variant as template, the first cDNA strand was synthesized using random hexamers, then a second cDNA strand was synthesized by adding buffer, dNTPs, RNase H and DNA polymerase I, and cDNA was

purified using AMPure XP beads. The end was trimmed and a poly (A) tail was added at the sequencing joint. The AMPure XP beads were used to select the fragment size, and the cDNA library was obtained using PCR enrichment.

After the library was constructed, the concentration and insert size of the library were detected using Qubit2.0 and Agilent 2100 respectively, and the effective concentration of the library was accurately quantified by q-PCR method to ensure the library quality. Based on Sequencing By Synthesis technology, deep transcriptome sequencing was performed using Illumina HiSeq 2500 (Illumina, USA) and the read length was PE125. The Raw Data were filtered by removing the connector sequence and low-quality reads to achieve high-quality clean Data. Sequence alignment was conducted between the clean data for each sample and assembled transcript or unigene library. The obtained transcript and unigene reads were called Mapped Reads and were used for subsequent analysis. The two samples of *V. uliginosum* and its white-fruited variant were sequenced three times in order to ensure the accuracy of the experiment.

Functional annotation of unigenes

The BLAST tool [33] was used to compare the unigene sequences with those in NR [34], Swiss-Prot [35], GO [36], COG [37], KOG [38], and KEGG [39] databases. The results of unigene orthology in KEGG were obtained using KOBAS 2.0 [40]. After the amino acid sequences of the unigenes were predicted, the unigenes were annotated by comparing with HMMER [41] by using the Pfam [42] database.

Expression calculation of unigenes

The reads of each sequenced sample were compared with the unigene database by using Bowtie [43]. Based on this comparison we estimated the expression levels by using RSEM [44]. We used the value of FPKM to indicate the expression abundance of unigenes.

FPKM can eliminate the effect of gene length and sequencing on the calculation of gene expression. The calculated gene expression can be directly used to compare the differences in gene expression between different samples. The formula is as follows:

$$\text{FPKM} = \frac{\text{cDNA Fragment}}{\text{Mapped Fragment(Millions)} \times \text{Transcript Length(kb)}}$$

Differential gene expression and anthocyanin synthesis-related gene analysis

The FDR of less than 0.01 and the difference in expression level (fold change; FC) of ≥ 2 were used as a screening criterion. FC represents the ratio of the expression levels for two samples (groups). Therefore, $\text{FC} > 2$ indicates differences in the expression level of the two samples.

The DEGs and anthocyanin synthesis pathway and glucose metabolism-related genes expression in the ripening fruits of *V. uliginosum* and its white-fruited variant were analyzed using quantitative real-time polymerase chain reaction (qRT-PCR). Specific primers were designed and synthesized according to the gene sequence obtained using high throughput sequencing. Further, specific primers of anthocyanin biosynthesis-related genes designed by Michael Zifkin [45] for *Vaccinium corymbosum* were used to verify the expression levels of the genes by using q-PCR (Additional file 5: Table S5). The *VuGAPDH* gene [27] (GenBank Accession No. KP218509) was as reference gene. The SYBR FAST qPCR Kit Master Mix (2×) and the universal dye method (KAPA Biosystems, USA) were used to validate the expression of the genes by using ABI7900HT real time quantitative PCR instrument (ABI Company, USA). The PCR condition was as follows: pre-incubation at 95 °C for 5 min; amplification at 95 °C for 3 s, 60 °C for 20 s and 95 °C for 15 s. The melting curves were measured at 60 °C for 15 s, 95 °C for 15 s. The genes expression levels of the ripening fruits of the white-fruited variant were compared with those of the blue ripening fruits of *V. uliginosum*. Data analysis was performed using the $2^{-\Delta\Delta CT}$ method. Statistical analysis was performed using *t*-test at the level of $P \leq 0.05$.

Differential expression of transcription factors at the different developmental stages of *V. uliginosum*

The expression of significantly down-regulated transcription factors at the different developmental stages of the white-fruited variety of *V. uliginosum* was detected by performing q-PCR by using the $2^{-\Delta\Delta CT}$ method. The PCR condition was the same as mentioned above. Moreover, the gene expression level and anthocyanin accumulation in the samples were assessed using *VuGAPDH* as the reference gene. The expression of transcription factors at the different developmental stages of *V. uliginosum* was detected and compared with that at the full-bloom stage (LS1). Statistical analysis is performed using Duncan test at the level of $P \leq 0.05$. Each gene from each sample was analyzed three times to ensure the accuracy of the experimental results.

Additional files

Additional file 1: Table S1. The up-regulated expression genes from the annotated DEGs (XLS 48 kb)

Additional file 2: Table S2. The down-regulated expression genes from the annotated DEGs (XLS 182 kb)

Additional file 3: Table S3. Differentially expressed genes identified using GO function enrichment (XLS 37 kb)

Additional file 4: Table S4. Differentially expressed genes from KEGG enrichment (XLSX 13 kb)

Additional file 5: Table S5. Specific primers used for anthocyanin biosynthesis pathway related genes (XLSX 13 kb)

Abbreviations

ANR: Anthocyanidin reductase; ANS: Anthocyanidin synthase; CCR: Cinnamoyl-CoA reductase; CHI: Chalcone isomerase; CHS: Chalcone synthase; DEG: Differentially expressed genes; DFR: Dihydroflavonol 4-reductase; F3'5'H: Flavonoid 3', 5'-hydroxylase; F3'H: Flavonoid 3'-hydroxylase; F3H: Flavanone 3-hydroxylase; FC: Fold change; FDR: False discovery Rate; FHT: Flavanone-3β-hydroxylase; GAPDH: Glyceraldehyde-3-phosphate dehydrogenase; LAR: Leucoanthocyanins reductase; MBW: MYB bHLH WD40; OPC: Procyanidolicoligomer; PAs: Proanthocyanidins; POD: Peroxidase; PPC: Procyanidolic polymer; q-PCR: quantitative polymerase chain reaction; qRT-PCR: quantitative real-time polymerase chain reaction; UFGT: UDP glucose-flavonoid 3-O-glucosyltransferase

Acknowledgements

The authors thank the staff of Lanjia Forest Farm for their help and support for the collection of experimental materials for this study.

Funding

The work was carried out under the auspices of the National Science Foundation of China(31460504).

Availability of data and materials

All data, generated or analyzed during this study are included in this published article and its additional information files. In addition, raw data for this project have been submitted to the Sequence Read Archive (SRA) database of the National Center for Biotechnology Information (accession number SRP157240), the SRA records will be accessible with the following link after the release date of 11 September 2019: <https://www.ncbi.nlm.nih.gov/sra/SRP157240>.

Authors' contributions

The research project was initiated and conceptualized by ZCW who also performed the experiments and was involved in the drafting, revision, and improvement of the manuscript. YY helped with the experiments and in drafting the manuscripts. CBH participated in the collection of experimental materials, RNA extraction, and reverse transcription analysis. TZW participated in transcriptome data collection and analysis. SBX participated in gene sequence analysis and in designing the primers. CHN helped in designing this study and drafting and revised the manuscript. All authors reviewed and approved the final submission.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing of interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 30 June 2018 Accepted: 4 December 2018

Published online: 13 December 2018

References

- Zong CL, Deng M, Zong CW, Cao HN, Li WJ. The research progress of *Vaccinium uliginosum*. *Chinese Journal of Northern Horticulture*. 2011;12:173–6.
- Masuoka C, Yokoi K, Komatsu H, Kinjo J, Nohara T, Ono M. Two novel antioxidant Ortho-benzoyloxyphenyl acetic acid derivatives from the fruit of *Vaccinium uliginosum*. *Food Sci Technol Res*. 2007;13(3):215–20.
- Yousef GG, Brown AF, Funakoshi Y, Mbeunkui F, Grace MH, Ballington JR, et al. Efficient quantification of the health-relevant anthocyanin and phenolic acid profiles in commercial cultivars and breeding selections of blueberries (*Vaccinium* spp.). *J Agric Food Chem*. 2013;61(20):4806–15.
- Ma JY and Zhang Y. A new variant of *Vaccinium uliginosum*. *Chinese Journal of Plant Research*. 2002;22(01):8.
- Zhang XD. A new variant of *vaccinium uliginosum* - white-fruited *uliginosum*. *Chinese Journal of Northern Horticulture*. 1990;21:61.

6. Tan ZW. Comparison of the composition and transcription group of flavonoids in *Vaccinium uliginosum* and white-fruited varieties. Yanbian University; 2016.
7. Sadilova E, Stintzing FC, Carle R. Anthocyanins, colour and antioxidant properties of eggplant (*Solanum melongena* L.) and violet pepper (*Capsicum annuum* L.) peel extracts. *Zeitschrift Fur Naturforschung C*. 2006; 61(7–8):527–35.
8. Pascual-Teresa SD, Sanchez-Ballesta MT. Anthocyanins: from plant to health. *Phytochem Rev*. 2008;7(2):281–99.
9. Gao L, Yang H, Liu H, Yang J, Hu Y. Extensive Transcriptome changes underlying the flower color intensity variation in *Paeonia ostii*. *Front Plant Sci*. 2016;6(80):1205.
10. Lei Z, Gao L, HongXue W, XiaoTian C, YunSheng W, Hua Y, et al. The R2R3-MYB, bHLH, WD40, and related transcription factors in flavonoid biosynthesis. *Functional and Integrative Genomics*. 2013;13(1):75–98.
11. Jaakola L, Poole M, Jones Mo Kämäräinen-Karppinen T, Koskimäki JJ, Hohtola A, et al. A SQUAMOSA MADS box gene involved in the regulation of anthocyanin accumulation in bilberry fruits. *Plant Physiol*. 2010;153(4):1619–29.
12. Wu G, Poethig RS. Temporal regulation of shoot development in *Arabidopsis thaliana* by miR156 and its target SPL3. *Development*. 2006; 133(13):3539–47.
13. Yakushiji H, Kobayashi S, Gotoyamamoto N, Jeong ST, Sueta T, Mitani N, et al. A skin color mutation of grapevine, from black-skinned pinot noir to white-skinned pinot blanc, is caused by deletion of the functional allele. *Biosci Biotechnol Biochem*. 2006;70(6):1506–8.
14. Kobayashi S, Goto-Yamamoto N, Hirochika H. Retrotransposon-induced mutations in grape skin color. *Science*. 2004;304(5673):982.
15. Jaakola L, Määttä K, Pirttilä AM, Törrönen R, Kärenlampi S, Hohtola A. Expression of genes involved in anthocyanin biosynthesis in relation to anthocyanin, proanthocyanidin, and flavonol levels during bilberry fruit development. *Plant Physiol*. 2002;130(2):729–39.
16. Debes MA, Arias ME, Grellet-Bournonville CF, Wulff AF, Martínez-Zamora MG, Castagnaro AP, et al. White-fruited *Duchesnea indica* (Rosaceae) is impaired in ANS gene expression. *Am J Bot*. 2011;98(12):2077–83.
17. Panumas K, Saichol K, Wouterg VD. A white mutant of Malay apple fruit (*Syzygium malaccense*) lacks transcript expression and activity for the last enzyme of anthocyanin synthesis, and the normal expression of a MYB transcription factor. *Funct Plant Biol*. 2011;38(1):681–91.
18. Gonzalez DH. Chapter 1—introduction to transcription factor structure and function. *Plant Transcription Factors*. 2016:3–11.
19. Luo G, Lingfei YE, Chen XB. Research progress of *Arabidopsis* B3 transcription factor gene superfamily. *Chemistry of Life*. 2013;33(3):287–93.
20. Zhang YF, Clemens A, Maximova SN, Gultinan MJ. The Theobroma cacao B3 domain transcription factor TcLEC2 plays a dual role in control of embryo development and maturation. *BMC Plant Biol*. 2014;14(1):1–16.
21. Shao J, Liu X, Wang R, Zhang G, Yu F. The over-expression of an *Arabidopsis* B3 transcription factor, ABS2/NGAL1, leads to the loss of flower petals. *PLoS One*. 2012;7(11):e49861–e61.
22. Lalusin AG, Nishita K, Kim SH, Ohta M, Fujimura T. A new MADS-box gene (IbMADS10) from sweet potato (*Ipomoea batatas* (L.) lam) is involved in the accumulation of anthocyanin. *Molecular genetics. Genomics*. 2006;275(1):44–54.
23. Park KI, Ishikawa N, Morita Y, Choi JD, Hoshino A, Iida S. A bHLH regulatory gene in the common morning glory, *Ipomoea purpurea*, controls anthocyanin biosynthesis in flowers, proanthocyanidin and phytomelanin pigmentation in seeds, and seed trichome formation. *Plant journal for cell. Mol Biol*. 2007;49(4):641–54.
24. Feng S, Wang Y, Yang S, Xu Y, Chen X. Anthocyanin biosynthesis in pears is regulated by a R2R3-MYB transcription factor PpMYB10. *Planta*. 2010;232(1):245–55.
25. Azuma A, Kobayashi S, Mitani N, Shiraiishi M, Yamada M, Ueno T, et al. Genomic and genetic analysis of Myb-related genes that regulate anthocyanin biosynthesis in grape berry skin. *Theor Appl Genet*. 2008;117(6):1009–19.
26. Aharoni A, De Vos CHR, Wein M, Sun Z, Greco R, Kroon A, et al. The strawberry FaMYB1 transcription factor suppresses anthocyanin and flavonol accumulation in transgenic tobacco. *Plant J*. 2001;28(3):319–32.
27. Primetta AK, Karppinen K, Riihinen KR, Jaakola L. Metabolic and molecular analyses of white mutant *Vaccinium* berries show down-regulation of MYBPA1-type R2R3 MYB regulatory factor. *Planta*. 2015;242(3):631–43.
28. Wada T, Kunihiro A, Tominagawada R. *Arabidopsis* CAPRICE (MYB) and GLABRA3 (bHLH) control tomato (*Solanum lycopersicum*) anthocyanin biosynthesis. *PLoS One*. 2013;9(9):e109093–e93.
29. Schwinn KE, Boase MR, Bradley JM, Lewis DH, Deroles SC, Martin CR, et al. MYB and bHLH transcription factor transgenes increase anthocyanin pigmentation in petunia and lisianthus plants, and the petunia phenotypes are strongly enhanced under field conditions. *Front Plant Sci*. 2014;5:1–11.
30. Liu XF, Yin XR, Allan AC, Wang KL, Shi YN, Huang YJ, et al. The role of MrbHLH1 and MrMYB1 in regulating anthocyanin biosynthetic genes in tobacco and Chinese bayberry (*Myrica rubra*) during anthocyanin biosynthesis. *Plant Cell Tissue and Organ Culture*. 2013;115(3):285–98.
31. Xie XB, Shen LI, Zhang RF, Zhao J, Chen YC, Zhao Q, et al. The bHLH transcription factor MdbHLH3 promotes anthocyanin accumulation and fruit colouration in response to low temperature in apples. *Plant Cell Environ*. 2012;35(11):1884–97.
32. Albert NW. Subspecialization of R2R3-MYB repressors for anthocyanin and Proanthocyanidin regulation in forage legumes. *Front Plant Sci*. 2015; 6(1165):1165.
33. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25(17):3389–402.
34. Deng YY, Li JQ, Wu SF, Zhu YP, Chen YW, He FC. Integrated nr database in protein annotation system and its localization. *Comput Eng*. 2006;32(5):71–2.
35. Consortium UP. UniProt: the universal protein knowledgebase. *Nucleic Acids Res*. 2004;32(Database issue):115–9.
36. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium Nature Genetics. 2000;25(1):25–9.
37. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res*. 2000;28(1):33–6.
38. Koonin EV, Fedorova ND, Jackson JD, Jacobs AR, Krylov DM, Makarova KS, et al. A comprehensive evolutionary classification of proteins encoded in complete eukaryotic genomes. *Genome Biol*. 2004;5(2):R7.
39. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. *Nucleic Acids Res*. 2004;32(Database issue):277–80.
40. Xie C, Mao XZ, Huang JJ, Ding Y, Wu JM, Dong S, et al. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res*. 2011;39(Web Server issue):316–22.
41. Eddy SR. Profile hidden Markov models. *Bioinformatics*. 1998;14(9):1055–70.
42. Finn RD. Pfam: the protein families database. *Nucleic Acids Res*. 2014; 42(Database issue):D222–30.
43. Langmead BC, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10(3):1–10.
44. Li B, Dewey CN. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12(1):93–9.
45. Zifkin M, Jin A, Ozga JA, Zaharia LI, Scherthaner JP, Gesell A. Gene Expression and metabolite profiling of developing highbush blueberry fruit indicates transcriptional regulation of flavonoid metabolism and activation of abscisic acid metabolism. *Plant Physiol*. 2012;158(1):200.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

