RESEARCH ARTICLE https://doi.org/10.7235/HORT.20220007

Molecular Cloning and Functional Analysis of the Gene and Promoter of *IbbHLH2* from Purple-Fleshed Sweet Potato

Danwen Fu†, Yake Hui†, Haihang Li, Shaohua Yang, YaHui Chen, and Feng Gao*

Guangdong Provincial Key Lab of Biotechnology for Plant Development, School of Life Sciences, South China Normal University, Guangzhou 510631, China

*Corresponding author: 593197289@qq.com

†These authors contributed equally to the work.

Abstract

Transcription factors basic helix-loop-helix (bHLH) have been identified as regulators of flavonoid synthesis in flowers, seeds and fruits. Transcription factors *IbbHLH2* regulate the expression of structural genes that participated in anthocyanin synthesize of purple-fleshed sweet potato. In this study, *IbbHLH2* had been cloned, the expression pattern and function were investigated in purple-fleshed sweet potato. Real-time quantitative PCR was used to determine the expression levels of transcription factors (*IbbHLH2*, *IbMYB1*, *IbWD40*) in different colour and developmental stages of sweet potato. The results showed that the expression of *IbbHLH2* was completely synchronized with the accumulation of anthocyanins in different colour and developmental stages of sweet potato. The *IbbHLH2* was introduced into wild-type *Arabidopsis thaliana* by *Agrobacterium*-mediated transformation. The cotyledon and epicotyl of the transgenic plants became purple in color, the anthocyanin content was increased and the structural genes (*CHS*, *CHI*, *F3H*, *DFR*, *ANS* and *3GT*) were upregulated. In addition, the *IbbHLH2* promoter sequence of 1020 bp was cloned by hi-TAIL PCR and was shown using a histochemical assay to initiate the expression of the *GUS* gene. Bioinformatic analysis indicates that the *IbbHLH2* promoter contains a TATA-box and a CAAT-box, MYB-binding elements and elements governing responses to light, gibberellic acid (GA), abscisic acid (ABA) and salicylic acid (SA). The expression of *IbbHLH2* was regulated by MYB transcription factors and environmental factors.

Additional key words: anthocyanin, environmental factors, expression pattern, function, transcription factors

Received: September 27, 2021
Revised: November 17, 2021
Accepted: December 6, 2021

HORTICULTURAL SCIENCE and TECHNOLOGY
40(1):63-74, 2022
URL: http://www.hst-j.org

pISSN : 1226-8763
eISSN : 2465-8588

This work was funded by the Key-Area Research and Development Program of Guangdong Province (2020B020219001).

Author Contributions

Feng Gao designed the research, coordinated the study and interpreted the results. Danwen Fu and Yake Hui drafted the manuscript and conceived the study. Haihang Li, Shaohua Yang and YaHui Chen revised the manuscript. Yake Hui and Danwen Fu revised and conceived the article.

Introduction

Sweet potato [*Ipomoea batatas* (L.) Lam.] is the sixth largest crop in the world with an annual yield of more than 100 million tons (Loebenstein et al., 2009). Purple-fleshed sweet potato are rich in anthocyanins which can antitumor, antibacterial, antioxidant activities and so on (Springob et al., 2003; Philpott et al., 2004; Yang et al., 2006; Pu and Fu, 2010). Anthocyanins extracted from purple-fleshed sweet potato have a bright color, excellent thermal and light stabilities that can be used to produce high quality anthocyanins.

The synthesis of anthocyanins was regulated by the combination of transcription factors R2R3-MYB, bHLH and WD40 in different plant tissues and developmental stages (Koes et al., 2005; Ramsay and Glover 2005; Lepiniec et al., 2006; Xu et al., 2018). The transcription complex regulated the expression of anthocyanin synthetic genes that control the pattern of accumulation of anthocyanins in plants (Atchley et al., 1999). In most species, the co-expression of R2R3-MYB, bHLH and WD40 strongly induces anthocyanin production (Naing et al., 2018). Six MYB10-like and three bHLH-like TFs were identified as candidates to be the regulators of the anthocyanin accumulation in peach (Rahim et al., 2014).

The transcription factors basic helix-loop-helix (bHLH) are named after their conserved structural domain and typically regulate gene transcription by forming homodimers or heterodimers that bind to target gene promoters (Heim et al., 2003; Holton and Cornish 1995). The best-characterized regulatory genes involved in anthocyanin production were the maize (*Zea mays* L.) *leaf color* (*Lc*) and *colorless* (*CL*) genes, which play a role in regulating structural gene expression in maize. The *Lc* gene strongly activated *dihydroflavonol reductase* (*DFR*), *flavonoid-3'5'-hydroxylase* (*F3'5'H*), *anthocyanidin synthase* (*ANS*) and *UDPglucose flavoniod-3-glucosyltransferase* (*UF3GT*) and plays an important role in the anthocyanin synthesis pathway (Bradley et al., 1998). Following the cloning of the maize *Lc* gene, which encodes a bHLH protein, corresponding genes were found in other plants, such as the *Delila* gene in snapdragon (*Antirrhinummajus*) (Naing et al., 2018). In petunia (*Petunia hybrida*), two related bHLH gene, *AN1* and *B-Peru* regulated the color of the corolla gene (Spelt et al., 2000; Trinh et al., 2016). Glabrous 3(*SmGL3)* was associated with anthocyanin synthesis in eggplant roots, stems, leaves, flowers, peels and pulp, while the *transparent testa (SmTT8*) regulated key enzymes downstream of the anthocyanin synthesis pathway (Park et al., 2007). In the Convolvulaceae plant *Ipomoea purpurea*, the transcriptional regulatory gene *IpIVS* (*IpbHLH2*) regulated the expression of related genes in the anthocyanin synthesis pathway, in particular *DFR-B* and *ANS*. Transcript levels of *DFR-B* and *ANS* were almost zero in the seed coat of plants that did not express *IpIVS* (Mano et al., 2007). bHLH proteins have also been found to be another key anthocyanin regulator in berry skins (Liu et al., 2015).

The tuberous roots of purple-fleshed sweet potato are rich in anthocyanins. The transcription factors and their regulatory targets involved in the regulation of anthocyanin synthesis in tuberous roots were not well understood. The R2R3 MYB transcription factor gene *IbMYB1* had been cloned in purple-fleshed sweet potato. It can regulate the expression of the *chalcone synthase* (*CHS*), *chalcone isomerase* (*CHI*), *F3H*, *DFR*, *ANS* and *3-glucosyltransferase* (*3GT*) involved in anthocyanin synthesis in purple-fleshed sweet potato (Liu and Chen, 2007). However, it was not clear whether bHLH transcription factor involved in the regulation of anthocyanin synthesis in purple-fleshed sweet potato. In this work, molecular cloning and functional analysis of the *IbbHLH2* (GenBank ID: JF508437) to regulation of anthocyanin synthesis in purple-fleshed sweet potato were investigated. The results showed that the color of the cotyledon and epicotyl of *IbbHLH2* transgenic plants were purple, their anthocyanin content was increased and the expression levels of anthocyanin-related structural genes (*CHS*, *CHI*, *F3H*, *DFR*, *ANS* and *3GT*) were up-regulated.

Materials and Methods

Materials and Growth Conditions

Purple-fleshed sweet potato *cv.* 'Yamakawamurasaki' , 'A5' , 'A6' and white-fleshed sweet potato *cv.* 'Yubeibai' were grown in the Guangdong Key Lab of Biotechnology for Plant Development. Arabidopsis (*Arabidopsis thaliana*) 'Columbia'

was grown in a greenhouse with day/night cycle of 16h/8h at 20 ± 2°C. Wild type tobacco (*Nicotiana tabacum*) 'NC89' was grown in a greenhouse under day/night cycle of 16h/8h photo-period at 28 ± 2°C.

Extraction and Determination of Anthocyanin

Plant material (0.3 g) was added to 1.0 ml of methanol containing 1% hydrochloric acid (w/v), and extracted at 21°C for 18 h with shaking at 150 rpm. The extraction solution was obtained after centrifugation at 21500 × g for 3 min and analyzed spectrophotometrically at 530 nm and 657 nm. The anthocyanin content was calculated using the formula: $Q_{Anthocyanin} = (A_{530} - 0.25 \times A_{657}) \times M^{-1}$, $Q_{Anthocyanin}$ represents the relative content of anthocyanins, A_{530} and A_{657} represent absorbance at 530 nm and 657 nm respectively, and M represents the fresh weight of plant material. All quantitative tests were repeated three times and the data shown are the average of the three repeats ± standard deviation.

Extraction of RNA and Genomic DNA

Total RNA was extracted using the Rapid Extraction Kit from Beijing Baitek Biotechnology Co. Ltd. (Beijing, China) following the manufacturer's instructions. The RNA concentration was determined using a Nano Drop ND1000 microassay spectrophotometer, and its purity was analyzed by agarose gel electrophoresis. DNA was extracted using a Plant Genomic DNA Extraction Kit (Tiangen Biotech, China).

Isolation and Structure Analysis of the Gene and Promoter of *IbbHLH2*

Approximately 2 mg of total RNA was reverse transcribed with oligo (dT) as primer using M-MLV reverse transcriptase (Takara, Japan) according to the manufacturer's recommendations. The *IbbHLH2* conservative fragment sequence was extracted using the PrimeScript® One Step RT-PCR Kit Ver.2 (Takara, Japan) following the manufacturer's instructions. The 3'-cDNA and 5'-cDNA of IbbHLH2 sequence was extracted by the SMART RACE cDNA amplification instructions. PCR products were analysed on 1% agarose gels, for each reaction product, a single fragment was recovered from gels and purifified using a DNA purifification Kit (Magen, Beijing, China). The fragment was ligated into the plasmid, transformed into Escherichia coli DH5α competent cells (Weidi Shanghai, China), and then sequenced (Sangong, Shanghai, China). The primers used for PCR were listed in Suppl. Table 1.

The *IbbHLH2* promoter sequence was cloned by hiTAIL-PCR. The reaction mixture (total volume 20 μL) contained 2.0 μL PCR buffer, 2.0 μL dNTP mix, 0.1 μL Ex Taq, 1 μL tuber DNA, 0.6 μL pIbbHLH2R1, 2.0 μL any one of the LAD primers, and water to 20 μL. Each 25 μL reaction mixture contained 2.5 μL PCR buffer, 2 μL dNTP mix, l μL *AC1*, 1 μL *pIbbHLH2R2/R3*, 0.1 μL Ex Taq, and 1.0 μL 10-fold-diluted primary Tail-PCR product, plus water to 25 μL. PCR products were analysed on 1% agarose gels, for each reaction product, a single fragment was recovered from gels and purifified using a DNA purifification Kit (Magen, Beijing, China). The fragment was ligated into the plasmid, transformed into Escherichia coli DH5α competent cells (Weidi Shanghai, China), and then sequenced (Sangong, Shanghai, China). The primers used for hiTAIL-PCR were listed in Suppl. Table 1.

The *IbbHLH2* ORF and nucleotide sequences was translated by ORF Finder tool in NCBI (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). IbbHLH2 biochemical properties was predicted by ProtParam tool and secondary structure was analysis by SOPMA software in Expasy (http://www.expasy.ch/tools/). Three-dimensional Model of IbbHLH2 was predicted and

analysis by Swiss-Model software (http://swissmodel.expasy.org/) and Weblab viewlite 4.0 software. The structure analysis of *IbbHLH2* promoter was pridicted by PLACE (http://www.dna.affrc.go.jp/PLACE/signalscan.html) and Plant-CARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/).

Construction of Expression Vectors

To identify the function of the *IbbHLH2* gene of purple-fleshed sweet potato *cv.* 'A5' in the regulation of anthocyanin synthesis, the plant overexpression vector *35S*::*IbbHLH2* gene was constructed and transformed into the wild type *arabidopsis thaliana* under the mediation of *agrobacterium* EHA10. Seeds of T1 generation were harvested and planted on an MS plate containing Hyg antibiotic. DNA was extracted from the growing seedlings and PCR was performed to confirm that the recombinant plasmid had been transferred to *arabidopsis thaliana*.

The pCambia1300 vector was used for plant expression constructs. The PCR product was purified and recovered using AxyPrep DNA Gel Recovery Kit (Axygen) according to the manufacturer's instructions. The amplified products were analyzed on 0.8% agarose gels, and single fragments were recovered from the gels. pCambia1300 was digested with KpnI and NcoI, the resulting fragment was ligated into the *bHLH2* DNA fragment and then transformed into *E. coli* DH5 α. Constructs were confirmed by sequencing. pCambia1300 containing *IbbHLH2* was transformed into *A. tumefaciens* strain EHA105.

To examine the functions of the *IbbHLH2* promoter, *pIbbH2* was fused with the *GUS* reporter gene in the pBI121 binary vector to generate the plant expression vector *pIbbHLH2::GUS*. The *pIbbHLH2* sequence was amplified from mountain purple genomic DNA using the specific primers *pIbbHLH2F* and *pIbbHLH2R*. The recombinant vector was then transformed into *E. coli* DH5 α and identified by PCR. After agarose gel electrophoresis, a target band of about 1000 bp was observed, indicating that the *pIbbHLH2::GUS* plant overexpression vector had been successfully constructed. The recombinant plasmid *pIbbH2::GUS* was transformed into *Agrobacterium* EHA105 and its presence confirmed by PCR. The *pIbbHLH2::GUS* plasmid was introduced into tobacco leaves by *Agrobacterium*-mediated transformation, and GUS activity was assessed in infected leaves transiently expressing either *pIbbHLH2::GUS* or *35S::GUS*.

pIbbHLH2 was recombined into the pBI121 vector enzymed by *HindIII* and *BamHI,* replacing the 35S promoter sequence. The *IbbHLH2* promoter was truncated for construct chimeric genes consisting of GUS driven by various *IbbHLH2* promoter regions. Full-length of *pIbbHLH2* were fused to a beta-glucuronidase (GUS) reporter gene to investigate the activity of *IbbHLH2* promoter. The PCR product was subcloned into the pBI121 vector by replacing the CaMV35S promoter in PBI121. Positive clones were verified by sequencing. The pBI121-bearing was transferred into *A. tumefaciens* strain EHA105. The primers used for constructing vectors were listed in Suppl. Table 1.

Stable Expression and GUS Assays

Transgenic Arabidopsis plants were generated by the standard floral-dip protocol as described elsewhere (Schoenbohm et al., 2000). Transformants were selected by growing on MS medium containing 50 $mg \cdot L^{-1}$ kanamycin. All hygromycin-resistant T3 plants were analyzed by PCR. Beta-glucuronidase (GUS) histochemical assays were performed according to the method described as our reported previously (Fu et al., 2020).

Real-time Quantitative PCR

The expression of transcription factors (IbMYB1, IbbHLH2 and IbWD40) in fibrous roots, thick roots and storage roots of purple-fleshed sweet potato cv. A5 and white-fleshed sweet potato cv. Yubeibai were analyzed using real-time quantitative PCR. First-strand cDNA was synthesized from total RNA using Prime Script™ RT Master Mix (Takara). RT-qPCR was conducted using SYBR® Premix Ex Taq™ II (Takara) in a total reaction volume of 20 μL consisting of 100 ng of template cDNA, each primer at 0.5 μM and 10 μL of SYBR® Premix Ex Taq™ II, and the amplification program was as follows: 1 cycle of 95 °C for 10 s followed by 40 cycles of 95 °C for 5 s and 60 °C for 30 s with a Bio-Rad CFX96 Real-Time PCR system (BIO-RAD, USA) according to the manufacturers' instructions. IbG14 was used as internal controls and calculated using the comparative Ct analysis method. The primers used for RT-qPCR were listed in Suppl. Table 1.

Results

Isolation and Functional Analysis of the Gene Sequence of *IbbHLH2*

In this study, the ORFs of the *IbbHLH2* genes of 'Yamakawamurasaki', 'A5' and 'Yubeibai' were cloned and systematically analyzed. Three fragments with a uniform size of about 2000 bp were obtained from cDNA of the three variety, which were named IbbHLH2-Ya (Yamakawamurasaki), IbbHLH2-A5 and IbbHLH2-Yu (Yubeibai). Sequence analysis showed that the three fragments have 94.8% cDNA sequence identity and 92.0% amino acid sequence identity. The amino acid sequence of the 'Yubeibai' had seven amino acid changes compared with the 'Yamakawamurasaki' and 'A5', but no differences were found in the key functional bHLH domain (Fig. 1).

The genomic DNA sequences of the four ORF regions of the *IbbHLH2* gene were obtained by PCR amplification using the segmented cloning strategy from the same three variety. After splicing, the complete coding sequences of three variety of *IbbHLH2* genomic DNA were obtained and named *gIbbHLH2-Ya*, *gIbbHLH2-A5* and *gIbbHLH2-Yu* (Fig. 2). *gIbbHLH2-Ya* was 6819 bp in length and contained eight exons and seven introns and the intron boundary sequences were in line with GT-AG rules. *gIbbHLH2-A5* was 6755 bp in length and also contained eight exons and canonical exon boundary sequences. *gIbbHLH2-Yu* was 4975 bp in length containing seven exons with conventional boundary sequences. It can be seen from the above results that the *IbbHLH2* genes of the two purple-fleshed sweet potato have an extra exon and intron compared with the white-fleshed sweet potato. Further analysis revealed that the lengths of the extra intron in the purple-fleshed sweet potato were 1815 bp in *gIbbHLH2-Ya* and 1747 bp in *gIbbHLH2-A5*, and the insertion position was in the middle of the sixth exon of the white-fleshed gene, located 323 bp upstream of the bHLH functional domain (Suppl. Fig. 1). Further analysis of the extra intron sequences in *gIbbHLH2-Ya* and *gIbbHLH2-A5* revealed that they were very similar (86.2%), both containing MYB binding sites, multiple photoresponsive elements and hormone (ABA, GA and SA) response elements. Therefore, the results indicated that this intron may regulate the expression of the *IbbHLH2* gene in purple-fleshed sweet potato.

To identify the function of the *IbbHLH2* gene in the regulation of anthocyanin synthesis, the plant overexpression vector *35S*::*IbbHLH2* gene was constructed and transformed into the wild type *arabidopsis thaliana* under the mediation of *agrobacterium* EHA10. Seeds of T1 generation were harvested and planted on an MS plate containing Hyg antibiotic.

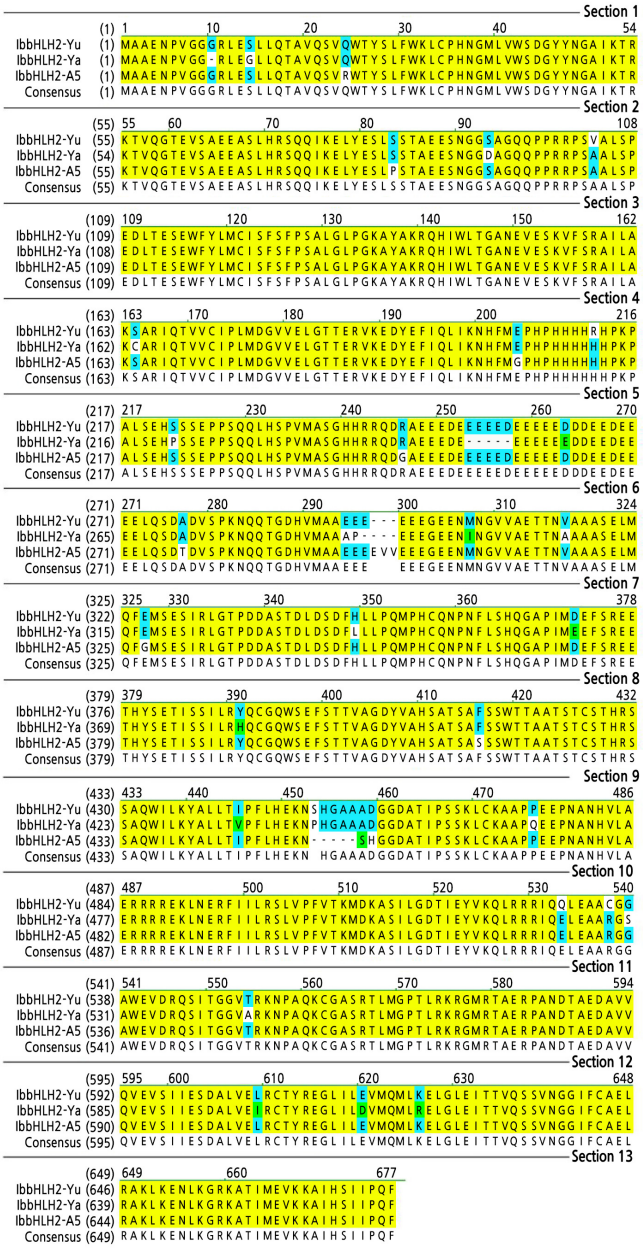


Fig. 1. Comparison of the amino acid sequences encoded by the *IbbHLH2* gene in the purple-fleshed and white-fleshed sweet potato. Ya and A5 was purple-fleshed sweet potato *cv.* 'Yamakawamurasaki' and 'A5', Yu was white-fleshed sweet potato cv. 'Yubeibai'. The green color on the sequence represent the different of the amino acid sequences encoded by the IbbHLH2 gene in the purple-fleshed and white-fleshed sweet potato.

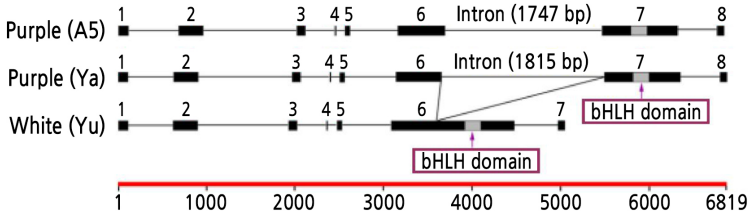


Fig. 2. Comparison of *IbbHLH2* genomic DNA structure in purple-fleshed and white-fleshed sweet potato. Ya and A5 was purple-fleshed sweet potato *cv.* 'Yamakawamurasaki' and 'A5', Yu was white-fleshed sweet potato *cv.* 'Yubeibai'.

DNA was extracted from the growing seedlings and PCR was performed to confirm that the recombinant plasmid had been transferred to *arabidopsis thaliana*. The T3 generation transgenic homozygous plants were obtained through twice on Hyg resistance screening. Genomic DNA and RNA were extracted from homozygous transgenic lines, and the *IbbHLH2* gene sequence of about 2000 bp was obtained by genomic PCR amplification, indicating that the gene *IbbHLH2* has been successfully inserted into the wild-type *Arabidopsis* genome (Fig. 3). Semi-quantitative RT-PCR showed that the *IbbHLH2* gene was expressed at high levels in transgenic plants (Fig. 3).

The color traits, anthocyanin content and related structural gene expression were investigated in transgenic *A. thaliana* (Suppl. Fig. 1). The result showed that the cotyledons and epicotyls of the seedlings of transgenic plants emerge purple color and the anthocyanin content was higher than in controls. Furthermore, the expression levels of anthocyanin synthesis-related structural genes (*CHS*, *CHI*, *F3H*, *DFR*, *ANS* and *3GT*) were all higher in transgenic plants. This results indicated that the *IbbHLH2* gene cloned in this paper was a functional sequence and that its introduction into the *Arabidopsis* genome can increase the content of anthocyanins in the resultant transgenic plants by up-regulating the expression level of related bio-synthetic genes.

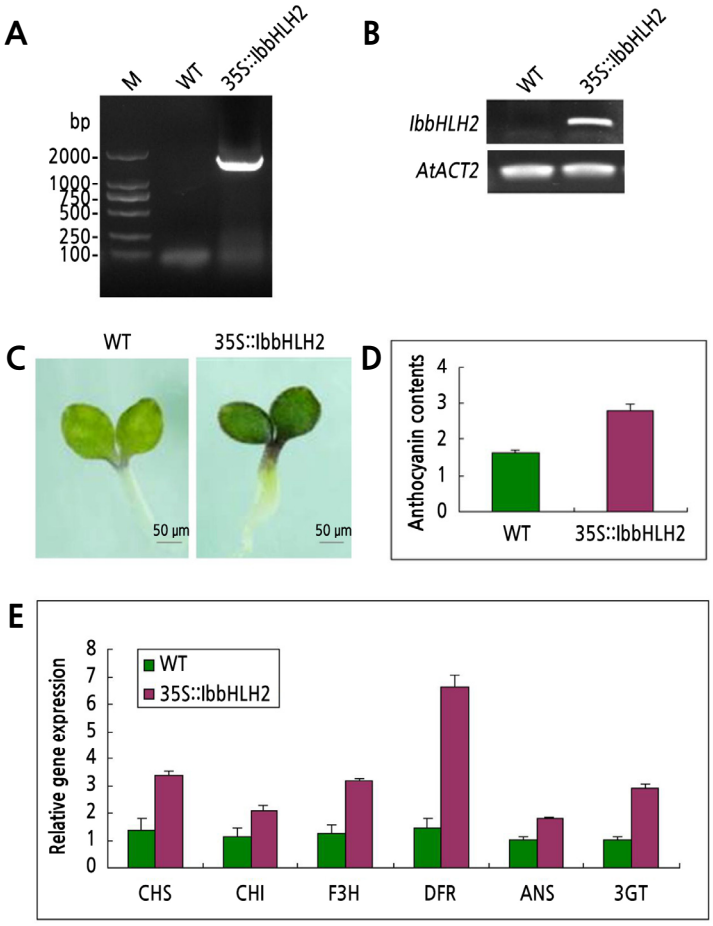


Fig. 3. Functional verification of *IbbHLH2* in transgenic *Arabidopsis*. A: PCR detection of *IbbHLH2* gene insertion in transgenic *Arabidopsis* genome, B: RT-PCR detection of *IbbHLH2* gene expression in transgenic *Arabidopsis*, C: wild-type and transgenic seedlings, D: content of anthocyanin in flowers of wild-type and transgenic *Arabidopsis* seedlings, E: expression level of anthocyanin synthesis-related structural genes in wild-type and transgenic *Arabidopsis* seedlings. The unit of anthocyanin content is g^{-1}FW.

Isolation and Functional Analysis of the Promoter of *IbbHLH2*

The genomic DNAs of 'Yamakawamurasaki', 'A5' and 'Yubeibai' were used as templates to clone the promoter sequences, which were named *pIbbH2-Ya* (Yamakawamurasaki), *pIbbH2-A5* (A5) and *pIbbH2-Yu* (Yubeibai). Sequence analysis showed that the above three sequences were very similar with an identity of 99.1% and no base mutations in the positions of the key *cis*-acting elements (Suppl. Fig. 2).

The promoter of *IbbHLH2* (*pIbbH2*) was obtained from purple-fleshed sweet potato 'Yamakawamurasaki' by hi-TAIL PCR and shown by sequencing to be 1020 bp in length (Suppl. Fig. 3). Analysis of the sequence revealed that *pIbbH2* contains a number of *cis*-acting transcription elements, MYB binding elements, a TATA-box, a CAAT-box, a MYC binding site, an abscisic acid (ABA) response element (ABRE), two binding sites for the transcription factor WRK710S associated with the gibberellin (GA) response and WBOX elements associated with the salicylic acid (SA) response. In addition, multiple photo-responsive structural elements were found in *pIbbHLH2*.

To examine the functions of the *IbbHLH2* promoter, *pIbbH2* was fused with the *GUS* reporter gene in the pBI121 binary vector to generate the plant expression vector *pIbbHLH2::GUS*. The *pIbbHLH2* sequence was amplified from mountain purple genomic DNA using the specific primers *pIbbHLH2F* and *pIbbHLH2R*. *pIbbH2* was recombined into the pBI121 vector enzymed by *HindIII* and *BamHI,* replacing the 35S promoter sequence. The recombinant vector was then transformed into *E. coli* DH5 α and identified by PCR. After agarose gel electrophoresis, a target band of about 1000 bp was observed, indicating that the *pIbbHLH2::GUS* plant overexpression vector had been successfully constructed. The recombinant plasmid *pIbbH2::GUS* was transformed into *Agrobacterium* EHA105 and its presence confirmed by PCR. The *pIbbHLH2::GUS* plasmid was introduced into tobacco leaves by *Agrobacterium*-mediated transformation, and GUS activity was assessed in infected leaves transiently expressing either *pIbbHLH2::GUS* or *35S::GUS*. The results indicate that *pIbbH2* has a promoter activity which can initiate the expression of the *GUS* gene (Fig. 4).

The Expression Characteristics of *IbbHLH2*

To explore the relationship between the expression pattern of the *IbbHLH2* and anthocyanin biosynthesis in the root of sweet potato, three purple-fleshed sweet potato and one white sweet potato were researched (Fig. 5). The anthocyanin content and the expression levels of *IbbHLH2* in different tissues at different developmental stages of the roots were measured. The results showed that the expression of *IbbHLH2* was completely synchronized with the accumulation of

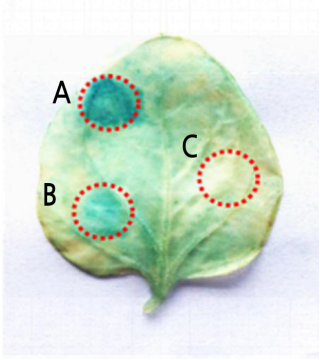


Fig. 4. Histochemical GUS staining analyses. GUS expression in tobacco leaves transiently expressing *pIbbHLH2::GUS* and (B) *35S::GUS*; no GUS expression was detected in leaf areas treated with *Agrobacterium EHA105* only (C).

anthocyanins in different variety and developmental stages.

The transcription levels of three transcription factor genes (*IbbHLH2*, *IbMYB1* and *IbWD40*) involved in anthocyanin synthesis were assessed by qRT-PCR in roots of the four sweet potato cultivars. The expression levels of *IbbHLH2* and *IbMYB1* across the four sweet potato were consistent with anthocyanin content and the expression levels of key anthocyanin biosynthesis pathway genes. *IbWD40* represented an exception to this general trend, since its expression level was highest

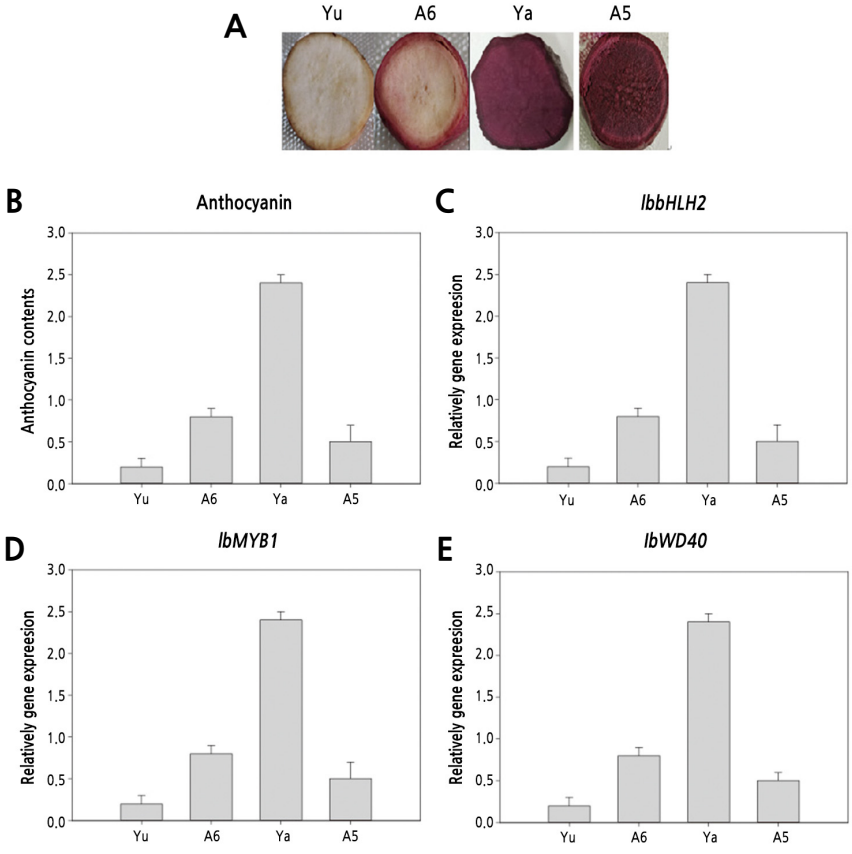


Fig. 5. Anthocyanin content and expression of transcription factors in different sweet potato. A: Cross-section of sweet potato roots; B: Anthocyanin content in different sweet potato; C, D and E: Expression of transcription factors in different sweet potato. Yu ('Yubeibai'), Ya ('Yamakawamurasaki'), A5 ('A5'), A6 ('A6'). The unit of anthocyanin content is $g^{-1}FW$.

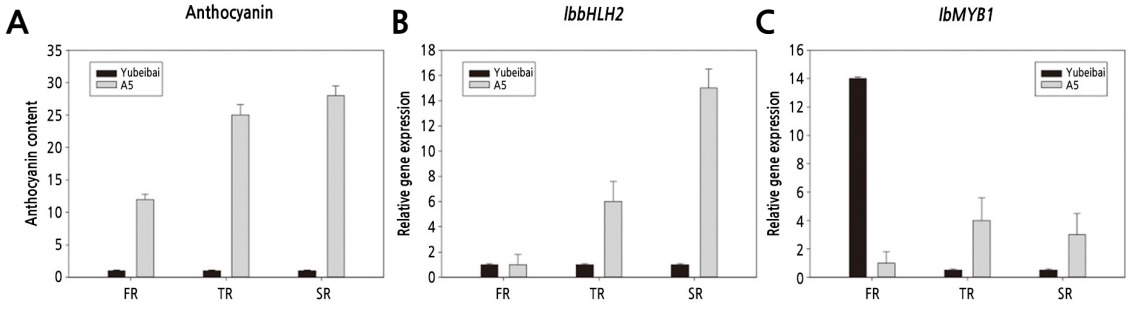


Fig. 6. Anthocyanin content and expression of transcription factors in different developmental stages of purple-fleshed and white-fleshed sweet potato. FR: Fibrous roots (diameter < 2 mm), TR: Thick roots (2 mm < diameter < 5 mm), SR: Storage roots (diameter > 5 mm). The unit of anthocyanin content is $g^{-1}FW$.

in 'Yamakawamurasak' rather than in 'A5' (Fig. 6). The result indicated that *IbbHLH2* and *IbMYB1* were the key transcription factors that regulate the synthesis of anthocyanin in rhizomes of purple sweet potato while *IbWD40* may be less important.

The anthocyanin content and the expression level of structure genes and transcription factors were measured in purple-fleshed strain 'A5' with the highest anthocyanin content and white-fleshed variety 'Yubeibai' with the lowest anthocyanin content at different developmental stages. Three developmental stages as foollows: fibrous roots (diameter < 2 mm), thick roots (2 mm < diameter < 5 mm) and storage roots (diameter > 5 mm) .

In the purple-fleshed sweet potato, the accumulation of anthocyanin content gradually increased as the roots developed and was highest in mature roots, the expression pattern of the *IbbHLH2* gene was also consistent with the trend in anthocyanin content, namely, fibrous roots < thick roots < storage roots. *IbMYB1* showed the highest expression in storage roots, and the lowest level in fibrous roots, which was inconsistent with the trend in anthocyanin content.

In white-fleshed sweet potato, the anthocyanin content was decreased gradually as roots matured (Fig. 4), the expression levels of all the tested genes were consistent with the trend of anthocyanin content, namely, fibrous roots > thick roots > storage roots. The expression levels of all the genes tested and the anthocyanin content in thick roots and mature storage roots of the purple-fleshed was higher in the white-fleshed sweet potato. *IbbHLH2* was expressed at a higher level in purple-fleshed roots while *IbMYB1* was higher in the white variety at the fibrous root. These results indicated that both *IbbHLH2* and *IbMYB1* involved in the regulation of anthocyanin synthesis in the thick roots and mature storage roots, while in the fibrous root, *IbbHLH2* play a key role in transcriptional regulation at the synthesis of anthocyanins.

In summary, the expression of the *IbbHLH2* gene was completely synchronized with the synthesis and accumulation of anthocyanins in different sweet potato variety and their different root developmental stages, indicating that *IbbHLH2* is the key transcriptional regulatory gene involved in the synthesis and metabolism of anthocyanins in purple-fleshed sweet potato.

Discussion

The *Lc* (*leaf colour*) gene, a member of the maize R gene family, was the first reported plant gene to encode a transcription factor with the bHLH domain. It is crucial for the synthesis of maize anthocyanins. Further studies revealed that *Lc* interacts with the maize MYB transcription factor *C1* to regulate the synthesis of anthocyanins, and that *Lc* interacts with the structural gene promoter sequence 5'-CAGGTG-3', which is found in genes encoding enzymes of the anthocyanin synthesis pathway, including the *CHS* 5'-CACGTA-3', *CHI* 5'-CACGTG-3'and *DFR* 5'-CACGTG-3' sites (Li et al., 2007). Transformation of the *Lc* gene into petunia produces a significant increase in expression of *PAL*, *DFR*, *F3'H*, *F3H*, *ANS* and *UFGT* (Bradley et al., 1998). Studies of *Lc* gene-transformed apples have also shown that *PAL*, *CHS*, *F3'H*, *DFR* and *ANS* are upregulated (Aboudi et al., 2012). In this research, the cognate *IbbHLH2* gene is expressed coordinately with accumulation of anthocyanin, indicating that *IbbHLH2* is the key transcriptional regulatory gene in the synthesis and metabolism of anthocyanins in purple-fleshed sweet potato. The low level of expression of *IbbHLH2* in white-fleshed sweet potato is likely one of the main reasons for the lack of color in roots of this variety.

Although the comparison of *IbHLH2* ORF sequences were changed in seven amino acid residues, there were no differences in the functional domain. This indicated that the loss of pigment synthesis in the white-fleshed cultivar was not caused by a change in the structure of *IbHLH2* protein, and there may be no difference in function between the

purple-fleshed and white-fleshed versions of IbbHLH2 protein. However, the genomic DNA sequences of *IbbHLH2* in purple and white-fleshed variety had significant differences. An extra intron, located 323 bp upstream of the sequence encoding the bHLH domain, had been inserted into the purple-fleshed sweet potato *IbbHLH2* gene. Further preliminary examination of this intron revealed that it contained a large number of *cis*-acting elements, including MYB binding elements, as well as light and hormone (ABA, GA and SA) response elements. Thus, this intron sequence may had a regulatory effect on the expression of the *IbHLH2* gene.

In recent researches, there had been a number of reports of the regulation of gene transcription by introns. In general, the presence of introns in yeast, nematodes, insects, mammals and plants has a positive effect on gene expression (Rose 2002). In plants, a variety of introns had been found to enhance gene expression. For example, *A. thaliana UBQ10* intron 1 can increase gene expression levels 15-fold, inclusion of maize *shl* gene subunit 1 can increase the expression level of a reporter gene by 40-fold, the intron of the *SeFAD2* gene of sesame also had an enhanced expression effect (Kim et al., 2006). The regulation of gene expression by introns occurs not only during pre-mRNA splicing, but can also had broader effects on the regulation of gene expression (Maquat and Carmichael 2001; Maniatis and Reed 2002; Orphanides and Reinberg 2002; Dong et al., 2014; Xu et al., 2014). Therefore, the extra intron in *IbbHLH2* may enhance the expression of its cognate gene by interacting with transcription factors. The details of this regulatory mechanism will be the focus of further studies. The identification of *IbbHLH2* as a key transcriptional regulatory gene in the anabolism of anthocyanin in purple sweet potato lays the foundations for detailed molecular studies of the synthesis and accumulation of anthocyanins in purple sweet potato tuber, and may stimulate new approaches for molecular-oriented breeding of sweet potato pigment traits (anthocyanin content). In this research, the study was based on an indirect functional test in the model system, although the genetic transformation system of *I. batatas Lam*. was already established. However, there needs more researchs to evidence the main idea in the future.

Literature Cited

Aboudi MM, Neha A, Belal EK, Athar A (2012) Role for gene looping in intron-mediated enhancement of transcription. PNAS 109:8505-8510. doi:10.1073/pnas.1112400109

Atchley WR, Terhalle W, Dress A (1999) Positional dependence, cliques, and predictive motifs in the bHLH protein domain. J Mol Evol 48:501-516. doi:10.1007/PL00006494

Bradley JM, Davies KM, Deroles SC, Bloor SJ, Lewis DH (1998) The maize *Lc* regulatory gene up-regulates the flavonoid biosynthetic pathway of *Petunia*. Plant J 13:381-392. doi:10.1046/j.1365-313X.1998.00031.x

Dong W, You YX, Niu LL, Gao F (2014) Isolation and analysis of the promoter of an anthocyanin synthase gene from purple-fleshed sweet potato tubers. Acta Physiol Plant 36:2637-2649. doi:10.1007/s11738-014-1635-4

Fu DW, Zhang JC, Wang DC, Gao F (2020) Molecular cloning, functional characterization and transcriptional regulation of the promoter of dihydroflavonal-4-reductase gene from purple-fleshed sweet potato. S Afr J Bot 131:206-213. doi:10.1016/j.sajb.2020.02.031

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

Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. Plant Cell 7:1071-1083. doi:10.2307/3870058

Kim MJ, Kim H, Shin JS, Chung CH, Ohlrogge JB, Suh MC (2006) Seeding-specific expression of sesame microsomal oleic acid desaturase is controlled by combinatorial properties between negative cis-regulatory elements in the *SeFAD2* promoter and enhancers in the 5'-UTR intron. Mol Genet Genomics 276:351-368. doi:10.1007/s00438-006-0148-2

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

Lepiniec L, Debeaujon I, Routaboul JM, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol 57:405-430. doi:10.1146/annurev.arplant.57.032905.105252

Li H, Flachowsky H, Fischer TC, Hanke MV, Forkmann G, Treutter D, Schwab W, Hoffmann T, Szankowski I (2007) Maize *Lc* transcription

factor enhances biosynthesis of anthocyanins, distinct proanthocyanidins and phenylpropanoids in apple (*Malus domestica Borkh.*). Planta 226:1243-1254. doi:10.1007/s00425-007-0573-4

Liu MY, Song CZ, Chi M, Wang TM, Zuo LL, Li XH, Zhang ZW, Xi ZM (2015) The effects of light and ethylene and their interaction on the regulation of proanthocyanidin and anthocyanin synthesis in the skins of *Vitis vinifera* berries. Plant Growth Regul 79:377-390. doi:10.1007/s10725-015-0141-z

Liu YG, Chen YL (2007) High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. Biotechniques 43:649-656. doi:10.2144/000112601

Loebenstein G, Thottappilly G, Fuentes S, Cohen J (2009) Virus and Phytoplasma Diseases. The Sweet Potato 10:105-134. doi:10.1007/978-1-4020-9475-0_8

Maniatis T, Reed R (2002) An extensive network of coupling among gene expression machines. Nature 416:499-506. doi:10.1038/416499a

Mano H, Ogasawara F, Sato K, Higo H, Minobe Y (2007) Isolation of a regulatory gene of anthocyanin biosynthesis in tuberous roots of purple-fleshed Sweet Potato. Plant Physiol 143:1252-1268. doi:10.1104/pp.106.094425

Maquat LE, Carmichael GG (2001) Quality control of mRNA function. Cell 104:173-176. doi:10.1016/S0092-8674(01)00202-1

Naing AH, Kin CK (2018) Roles of R2R3-MYB transcription factors in transcriptional regulation of anthocyanin biosynthesis in horticultural plants. Plant Mol Biol 98:1-18. doi:10.1007/s11103-018-0771-4

Orphanides G, Reinberg D (2002) A unified theory of gene expression. Cell 108:439-451. doi:10.1016/S0092-8674(02)00655-4

Park KI, Ishikawa N, Morita Y, Choi JD, Hoshino A, Iida S (2007) 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 J 49:641-654. doi:10.1111/j.1365-313X.2006.02988.x

Philpott M, Gould KS, Lim C (2004) In situ and in vitro antioxidant activity of sweet potato anthocyanins. J Agr Food Chem 52:1511-1513. doi:10.1021/jf034593j

Proudfoot NJ, Furger A, Dye MJ (2002) Integrating mRNA processing with transcription. Cell 108:501-512. doi:10.1016/S0092-8674(02)00617-7

Pu Y, Fu Y (2010) Health care function and utilization of purple meat sweet potato. Sci Tech Inform 13:237-240

Rahim MA, Busatto N, Trainotti L (2014) Regulation of anthocyanin biosynthesis in peach fruits. Planta 240:913-929. doi:10.1007/s00425-014-2078-2

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

Rose AB (2002) Requirements for intron-mediated enhancement of gene expression in *Arabidopsis*. RNA 8:1444-1453. doi:10.1017/S1355838202020551

Schoenbohm C, Martens S, Eder C, Forkmann G, Weisshaar B (2000) Identification of the *Arabidopsis thaliana* flavonoid 3'-hydroxylase gene and functional expression of the encoded P450 enzyme. Biol Chem 381:749-53. doi:10.1515/BC.2000.095

Spelt C, Quattrocchio F, Mol JN, Koes R (2000) Anthocyanin1 of petunia encodes a basic helix-loop-helix protein that directly activates transcription of structural anthocyanin genes. Plant Cell 12:1619-1632. doi:10.1105/tpc.12.9.1619

Springob K, Nakajima J, Yamazaki M, Saito K (2003) Recent advances in the biosynthesis and accumulation of anthocyanins. Nat Prod Rep 20:288-303. doi:10.1039/b109542k

Trinh NA, Aung HN, Muthukrishnan A, Sun HL, Chang KK (2016) Sucrose-induced anthocyanin accumulation in vegetative tissue of Petunia plants requires anthocyanin regulatory transcription factors. Plant Sci 252:144-150. doi:10.1016/j.plantsci.2016.06.021

Xu KD, Huang XH, Wu MM, Wang Y, Chang K, Liu J, Zhang Y, Zhang FL, Zhang LM, et al (2014) Rapid, highly efficient and economical method of Agrobacterium-mediated in planta transient transformation in living onion epidermis. PLoS ONE 9:e83556. doi:10.1371/journal.pone.0083556

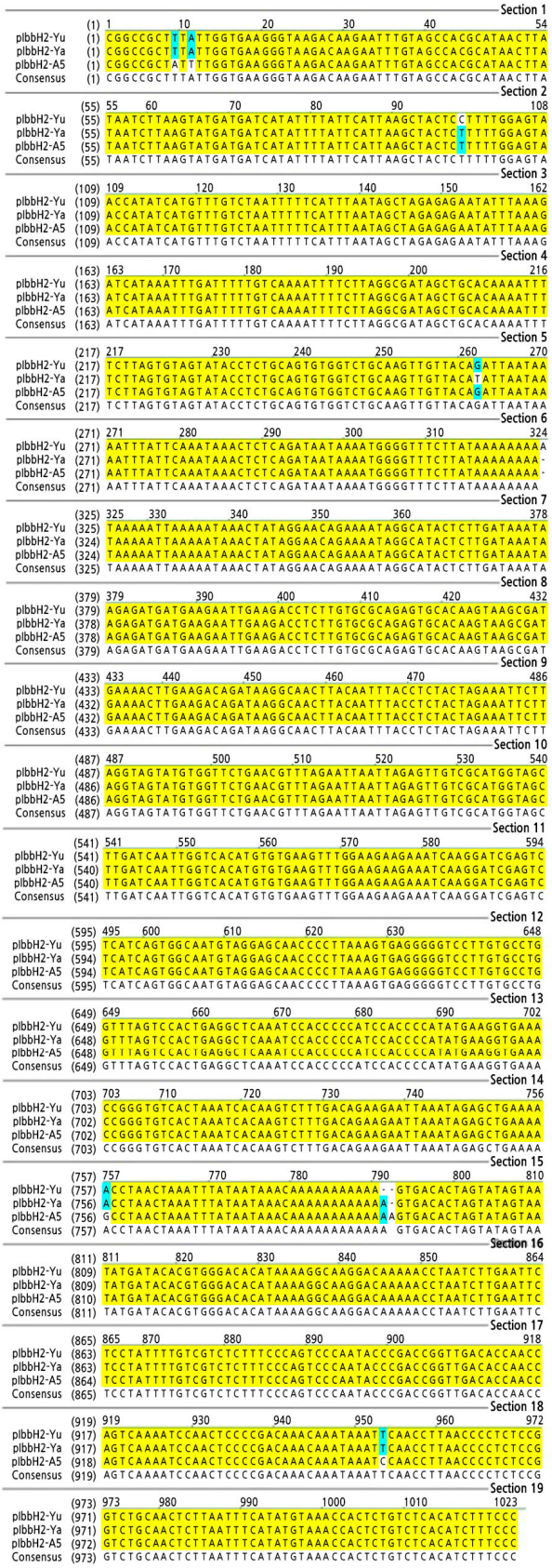
Xu PB, Zawora C, Li Y, Wu J, Liu LC, Liu ZC, Cai R, Lian HL (2018) Transcriptome sequencing reveals role of light in promoting anthocyanin accumulation of strawberry fruit. Plant Growth Regul 86:121-132. doi:10.1007/s10725-018-0415-3

Yang XS, Yang JM, Gao F (2006) Advances in purple sweet potato pigment. Chinese Agr Sci Bull 04:94-98

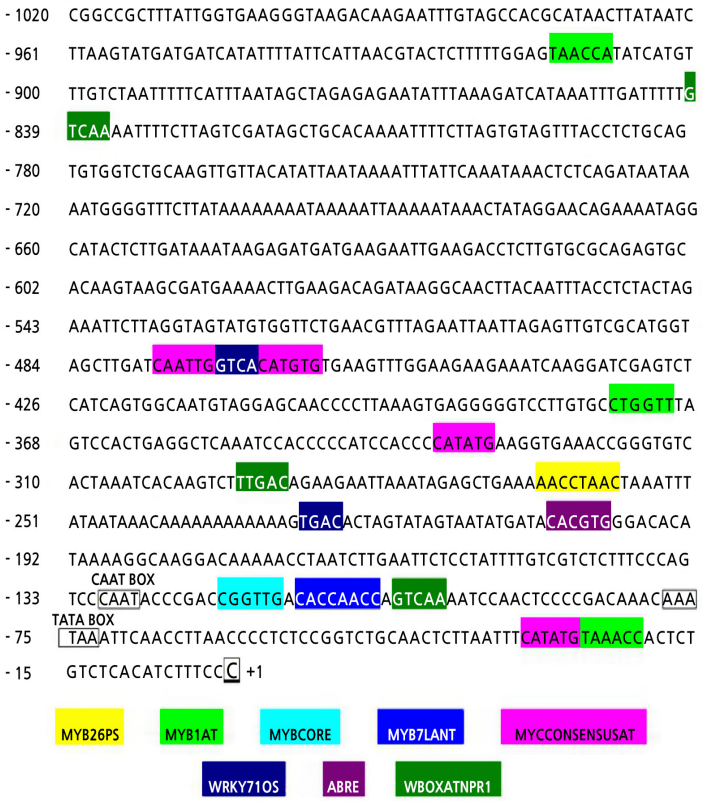
Supplemental Table 1. The primer sets used in this research

Primer name	Primer sequence (5'-3')
Isolation of the gene *IbbHLH2*	
IbbHLH2F-core	GTGGTTCTACCTCATGTGCATCTCC
IbbHLH2R-core	CGTACTCAATGGTGTCGCCGA
UPM	long:CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT short:CTAATACGACTCACTATAGGGC
NUPM	AAGCAGTGGTATCAACGCAGAGT
IbbHLH2F3-1	TAGAGCAATTCTCGCCAAGTGTG
IbbHLH2F3-2	CCGAAACAATATCGAGCATCC
IbbHLH2R5-1	CGTACTCAATGGTGTCGCCGA
IbbHLH2R5-2	GTTCCCAGTTCAACCACGCC
Isolation of the *pIbbHLH2*	
LAD1-1	ACGATGGACTCCAGAGCGGCCGCVNVNNNGGAA
LAD1-2	ACGATGGACTCCAGAGCGGCCGCBNBNNNGGTT
LAD1-3	ACGATGGACTCCAGAGCGGCCGCVVNVNNNCCAA
LAD1-4	ACGATGGACTCCAGAGCGGCCGCBBNBNNNCGGT
AC1	ACGATGGACTCCAGAG
pIbbHLH2R1	GAGGTAGAACCACTCAGACTCCGTCAA
pIbbHLH2R2	CTACCACCACAAATACGCATAAACACCC
pIbbHLH2R3	GTCCCTTGACTATAATATCCCATCCGAG
Construction of expression vectors	
IbbHLH2GF	TTGGTACCGATGGCGGCGGAAAACCCTGT
IbbHLH2GR	GGCCCGGGCAAACTGAGGAATTATACTATGAATTGC
pIbbHLH2F:	GCAAGCTTCGGCCGCTTTATTGGTGAAGGGTAAGAC
pIbbHLH2R:	CCGGATCCTACTATATATAGATTACAGAGAAGCTAATTAAGGCGGCGATGC
Real-time quantitative PCR	
QbHLH2-F	GCTTAGGGAGCTTGGACTAGAAATC
QbHLH2-R	CTTTTTCACTTCCATGATTGTTGCT
QMYB1-F	CCGACGACTCTATTGACCG
QMYB1-R	TGCTGAAGGCAAAGTGGTT
QWD40-F	GGGGTTTTCGCTTCAGTGTCG
QWD40-R	ATCTGAGCAAAGGGGTGTCCG
QCHS-F	TCACTAATAGCGAGCATAAAACCG
QCHS-R	CGCCATGTATTCACAAAAGCT
QCHI-F	GAAAAGTTGGGAGTGAGAATGATGC
QCHI-R	GCACACAACACGCAATAAAACATAG
QF3H-F	CTGCTGGAGGTTTTATCAGAGGC
QF3H-R	TAGGTCAGGCTGGGGG CATTTTG
QDFR-F	GCTTTATCGGCTCCTGGTTGGTCAT
QDFR-R	ATTCGTGTCCGCTTTCGGTAGTTC
QANS-F	GACACGCCCAAACCTGATGAA
QANS-R	ATACCAAACGGACTCCACAAAGC
Q3GT-F	CTTGCTCTATTTCTGCTCACTTGTA
Q3GT-R	TTCTGACATTCCTGGGATTACTTT
QACT-F	GCGGATAGAATGAGCAAGG
QACT-R	AGCCTCCAATCCAGACA CT
G14-F	ATGTCGGACAAGTGCGGAAACTGCG
G14-R	TTAGTGGCCACAGGTGCGGTCGGTA

Supplemental Fig. 1. The sequence and motifs of *IbbHLH2* in purple-fleshed sweet potato. The red color of the box on the sequence represent the domain of CAAT-box. The yellow color of the box on the sequence represent the domain of ABA. The blue color of the box on the sequence represent the domain of TATA-box. The green color of the box on the sequence represent the domain of MYB binding site. The orange color of the box on the sequence represent the domain of light responsive element.



Supplemental Fig. 2. Alignment of the promoter squences of the *IbbHLH2* genes from purple-fleshed and white-fleshed sweet potato.



Supplemental Fig. 3. The sequence and motifs of *pIbbHLH2* in purple-fleshed sweet potato.