

Identification and characterization of seed-specific transcription factors regulating anthocyanin biosynthesis in black rice

C.-K Kim · M.-A. Cho · Y.-H. Choi · J.-A. Kim ·
Y.-H. Kim · Y.-K. Kim · S.-H. Park

Received: 18 May 2010 / Revised: 16 November 2010 / Accepted: 8 December 2010 / Published online: 1 February 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Black rice is rich in anthocyanin and is expected to have more healthful dietary potential than white rice. We assessed expression of anthocyanin in black rice cultivars using a newly designed 135 K *Oryza sativa* microarray. A total of 12,673 genes exhibited greater than 2.0-fold up- or down-regulation in comparisons between three rice cultivars and three seed developmental stages. The 137 transcription factor genes found to be associated with production of anthocyanin pigment were classified into 10 groups. In

addition, 17 unknown and hypothetical genes were identified from comparisons between the rice cultivars. Finally, 15 out of the 17 candidate genes were verified by RT-PCR analysis. Among the genes, nine were up-regulated and six exhibited down-regulation. These genes likely play either a regulatory role in anthocyanin biosynthesis or are related to anthocyanin metabolism during flavonoid biosynthesis. While these genes require further validation, the results here underline the potential use of the new microarray and provide valuable insight into anthocyanin pigment production in rice.

C.-K. Kim · S.-H. Park (✉)
Genomics Division,
National Academy of Agricultural Science (NAAS),
Suwon 441-707, Korea
e-mail: shp1790@rda.go.kr

M.-A. Cho
Fruit Research Division,
National Institute of Horticultural & Herbal Science (NIHHS),
Suwon 440-706, Korea

Y.-H. Choi
Bioenergy Crop Research Center,
National Institute of Crop Science (NICS),
Suwon 441-706, Korea

J.-A. Kim
Bio-crop Development Division,
National Academy of Agricultural Science (NAAS),
Suwon 441-707, Korea

Y.-H. Kim
Policy Development Office, Korea Institute of Planning &
Evaluation for Technology in Food, Agriculture,
Forestry & Fisheries,
Anyang 431-060, Korea

Y.-K. Kim
School of Electrical Information Communication Engineering,
Wonkwang University,
Iksan, Korea

Keywords Anthocyanin · Black rice · Pigmentation ·
Rice microarray · Transcription factor

Introduction

Rice, one of the most economically important cereal crops in the world, is a significant staple for feeding much of the world's population. Although white rice is most commonly consumed, several rice cultivars contain color pigments, such as black and red rice. Black rice has a high anthocyanin content located in the pericarp layers, which gives it a dark purple color (Ryu et al. 1998; Takashi et al. 2001).

Anthocyanins, a group of reddish-purple water-soluble flavonoids, are the primary pigments in the red and black grains, and are responsible for the attractive red, purple, and blue colors of many flowers, fruits, and vegetables. Based on several reviews, it is estimated that more than 400 naturally occurring anthocyanins have been found (Kong et al. 2003). The pigment composes a group of natural colorants belonging to the flavonoid family and is produced by tissues in response to developmental and environmental signals (Quattrocchio et al. 1993). The regulatory elements that confer tissue-specific accumulation of anthocyanin have

been characterized in several plant species (Holton and Cornish 1995). The main pigments comprising anthocyanin of dark purple rice are cyanidin (cyanidin-3-*O*- β -glucoside) and peonidin (peonidin-3-*O*- β -glucoside) (Ryu et al. 1998; Abdel-Aal et al. 2006).

Recently, anthocyanins have been recognized as health-promoting food ingredients due to their antioxidant activity (Nam et al. 2006; Philpott et al. 2006), and anticancer (Hyun and Chung 2004), hypoglycemic (Tsuda et al. 2003), and anti-inflammatory effects (Tsuda et al. 2002). For example, pigment-supplemented diets of black rice reduced oxidative stress in mice (Xia et al. 2003) and its pigment fraction may have antiatherogenic activity (Xia et al. 2006). A recent report showed that black rice is a good source of fiber, minerals, and several important amino acids (Zhang et al. 2005), and there is increased interest in alternative sources of anthocyanins due to a rising demand for economical sources of natural and stable pigments (Hu et al. 2003; Zhang et al. 2004).

The examined cv. Heugjinju has the richest anthocyanin content among dark purple rice cultivars, with the highest cyanidin-3-glucoside (C3G) pigment content of these cultivars (Ryu et al. 2000; Kim et al. 2007). C3G is known to have high antioxidative function and is the most powerful among the 14 main anthocyanins (Ding et al. 2006).

Orthology assignment in pigmentation biosynthesis is a critical prerequisite of numerous comparative genomics procedures (Dessimoz et al. 2006). The clusters of orthologous groups (COGs) analysis is delineated by comparing protein sequences encoded in complete genomes, representing major phylogenetic lineages. It is important that identification of orthologs is derived from the COGs methodology (Remm et al. 2001). The Gene Ontology (GO) is a major bioinformatics initiative with the aim of standardizing the representation of genes involved in the identification of orthologs (Ashburner et al. 2000). Recently, a number of studies have demonstrated that microarray analyses can be used to examine changes in genome-wide gene expression. Microarray experiments have been used to analyze gene expression changes in a number of crop species, including rice (Kim et al. 2009; Kim et al. 2010; Troester et al. 2009).

Here, we investigated anthocyanin gene expression used by black rice to gain insight into the causes of pigment production. Our results will greatly facilitate future breeding of hybrid varieties of rice rich in anthocyanin.

Materials and methods

Rice

The white rice used was the cv. Dongjin. Black rice was a Heugjinju and Heugseol cultivar. Three rice cultivars were

cultivated under controlled conditions in an experimental field of the National Academy of Agricultural Science (NAAS). Heugjinju represents the first breeding cultivar of black rice in Korea (Moon et al. 1998). Heugseol is a new black pericarp cultivar for high anthocyanin, which was developed by the National Institute of Crop Science in 2008 (Choi et al. 2008). It was derived as a crossbreed between Seolgaeng/Heugjinju.

Experimental design

This experiment was designed to assess three factors (i.e., one white and two black rice cultivars) and three treatments (i.e., three seed developmental stages, heading+7 days, +14 days, and +21 days), in triplicate. The samples were harvested from research plots in the experimental fields in 2009. They were manually hulled and ground to obtain a fine powder using a cyclone mixer mill (HMF-590; Hanil, Seoul, Korea) and a mortar and pestle. The milled rice powders were kept at -80°C before RNA extraction.

The results of anthocyanin analysis showed total anthocyanin content of Heugjinju cultivar to be 295 mg/100 g and Heugseol to be 417 mg/100 g. The major components for both consisted of cyanidin-3-*O*-glucoside (C3G) and peonidin-3-*O*-glucoside. The distribution ratio of individual anthocyanins showed Heugjinju to have 95.4% C3G and Heugseol, 94.7% C3G. In contrast, the Dongjin cultivar (white seed) showed no detectable anthocyanins.

Therefore, the experiments were performed with Dongjin (white color seed of no anthocyanin), Heugjinju (black color seed of lower anthocyanin) and Heugseol (black color seed of higher anthocyanin). We assumed that two black cultivars would be useful for comparing the gene expression of anthocyanins because they have a similar genetic background but very different anthocyanin content.

RNA extraction

Total RNA was extracted from seeds of two black rice cultivars, Dongjin and Heugjinju, and the wild type Dongjin cultivar. We performed at least three replicates for each treatment. Frozen samples were homogenized with a mortar and pestle in liquid N_2 . The ground powder was transferred to empty falcon tubes on liquid N_2 until the homogenization procedure was ready to be performed, and 0.5 mL of RLC buffer (Qiagen, Hilden, Germany) was added. Because of the high starch content of rice seed, the homogenates were shaken with a vortex for only 10 s, and then plant debris was pelleted by centrifugation. Supernatants were extracted using an RNeasy Kit (Qiagen, Valencia, CA, USA) according to manufacturer's instructions. The RNA samples were further purified using

phenol-chloroform-isoamylalcohol (25:24:1) and an RNeasy mini plant kit (Qiagen). The quantity of total RNA was determined by measuring absorbance at 260 nm and 280 nm by using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, NC, USA). In addition, the level of protein contamination in the RNA was determined based on the A260/A280 ratio. Only RNA samples with ratios of 2.0–2.2 were used for these experiments.

Anthocyanin analysis

The aleurone layer, about 10% of the outer layer of black rice, was extracted with 10 mL of 1% HCL in methanol for 24 h in the dark. Quantification of anthocyanin was performed using a Waters CapLC XE pump (Millford, MA, USA) with Waters 2996 photodiode array detection system. A Symmetry C₁₈, 5 μm (0.32×150 mm) column, operated at 25°C, from Waters was used for separation. The mobile phase consisted of 0.1% trifluoroacetic acid (TFA) in water (eluant A) and 0.1% TFA in 95% acetonitrile (eluant B). The injection volume for all samples was 0.2 μL. Spectra were recorded from 200 to 600 nm at a flow-rate of 0.5 μL min⁻¹. The electrospray ionization mass spectrometry (ESI-MS) was carried out with a Micromass electrospray interface ZMD 4000 (Micromass, Manchester, UK). Nitrogen was used as the nebulizing gas. Source blocks and desolvation temperatures were 110°C and 20°C, respectively. Prior to analysis, all samples were filtered through a 0.45-μm membrane filter. The standard anthocyanins (i.e., cyanidin-3-*O*-glucoside and peonidin-3-*O*-glucoside) were obtained from Extrasynthese (R&D Chemicals, Genay, France), and apigenin, kaempferol and quercetin were purchased from Sigma Chemical Co. (St. Louis, MO, USA).

Microarray analysis

Microarray analysis was performed using newly designed 135 K *Oryza sativa* microarrays. These microarrays contained probes to assess 31,439 genes deposited at the International Rice Genome Sequencing Project (IRGSP, <http://rgp.dna.affrc.go.jp/E/IRGSP/>) and the Rice Annotation Project version 2 (RAP2, <http://rapdb.dna.affrc.go.jp/>). Four 60-nt probes were designed from each gene starting at 60 bp upstream of the stop codon and with a 30-bp shift such that four probes could be used to cover a 150-bp region at the 3' end of the gene. In addition, 200 genes were included from 123 chloroplasts, 74 mitochondria, and 3 selection markers. Therefore, the microarray had a total of 125,956 probes. The microarray was scanned using the GenePix 4000B (Molecular Devices, Inc., Sunnyvale, CA, USA), and signals were digitized and analyzed by

Nimblescan (Nimblegen, Madison, WI, USA). To improve the sensitivity and reproducibility of microarray analysis (Irizarry et al. 2003), normalization was processed using the cubic spline normalization method (Workman et al. 2002), and probe-level summarization was processed by Robust Multichip Analysis (RMA) using a median polish algorithm (Sadlier et al. 2004).

RT-PCR analysis

Reverse transcription-polymerase chain reaction (RT-PCR) analysis was carried out to verify the microarray hybridization data. To reduce experimental variation, a total of 27 RNA samples were extracted from the Dongjin, Heuginju, and Heugseol cultivars at 7, 14, and 21 days after the heading stage, in triplicate. Each RNA was independently isolated for RT-PCR, apart from those used for the microarray analysis.

After treatment with RNase-free DNase followed by phenol extraction and ethanol precipitation, 5 μg of total RNA from each sample was used to synthesize each pool of cDNA using the SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA). For PCR amplification, 1 μL of the resulting cDNA reaction was used as a template. The PCR reactions were carried out in 50-μL volumes with 1.25 unit of *Taq* DNA polymerase and 20 pmol of each primer pair. The primers were designed based on sequence information available from the Rice Annotation Project Database (RAP, <http://rapdb.dna.affrc.go.jp>) and the NCBI/Primer-BLAST (http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi?LINK_LOC=BlastHome). The PCR program was as follows: 3 min at 94°C, 24 cycles of 30 sec at 94°C, 30 sec at the optimal annealing temperature (58 to 65°C), and 1 min at 72°C, followed by 5 min at 72°C. To validate our RT-PCR results, we performed each experiment three times. The PCR reaction with 24 cycles of amplification was analyzed by gel electrophoresis. Actin mRNA was used as a loading control.

Data collection and statistical analysis

All experiments were performed using only Cy3 to eliminate the dye-swap error value. Data-based background subtraction using a local background estimator was performed to improve fold change estimates on arrays with high background. Spot intensity was calculated as the median value of the spot compared to the background median value. The functional classification categories assessed during COGs analysis using the NCBI/COGs database (<http://www.ncbi.nlm.nih.gov/COG/>). To determine the processes affected significantly by anthocyanin production, GO analysis was performed using GoMiner

software (Zeeberg et al. 2003). The statistical significance of these matches was assessed by calculating p-values using the one-sided Fisher exact test for the number of categorized GO terms in the total analysis. False discovery rate (FDR) values were collected from less than 0.05. The selected genes were analyzed by average linkage clustering to group genes with similar function using SAS/STAT (SAS Institute Inc., Cary, NC, USA, <http://www.sas.com>). Transcription factors were evaluated and clustered using the hyper-geometric analysis method. Candidate genes in the hyper-geometric analyses were defined as those with p-values less than 0.05 according to ANOVA assessment using GeneSpring GX 11 software (Agilent Technologies Inc., Santa Clara, CA, USA, <http://www.chem.agilent.com/>).

Results and Discussion

To evaluate anthocyanin biosynthesis in rice cultivars, we performed a four-stage screen. First, we tested the hypothesis that comparison between cv. Dongjin and high anthocyanin cultivars, such as Dongjin/Heugjinju, Dongjin/Heugseol, and Heugjinju/Heugseol, express the same genes at three time points following the heading stage. Second, COG and GO analysis were performed to classify the genes functionally and identify orthologous genes involved in anthocyanin biosynthesis. Third, transcription factors involved in anthocyanin pigmentation were evaluated and clustered by the hyper-geometric distribution analysis method. Finally, selected unknown and predicted genes related to anthocyanin biosynthesis were verified by reverse transcription-polymerase chain reaction (RT-PCR).

Data collection and p-value analysis

Microarray analysis was performed to assess three factors (i.e., a white cultivar and two black rice cultivars) and three treatments (i.e., seed developmental stages) in triplicate. From this experiment, a total of 852,876 intensity items was yielded. We screened 12,673 genes to identify those with greater than a 2-fold up- or down-regulation in each combination of the three rice cultivars and three seed developmental stages. The number of genes significantly upregulated in each cultivar during the three seed developmental stages ranged from 528 to 3,695. The number of down-regulated genes ranged between 295 and 1,782. The number of significant up-regulated genes for each developmental stage was higher than that of down-regulated genes. In the Dongjin/Heugjinju combination, 1,089 upregulated genes were significantly expressed in the heading+7 days stage, and this number increased to 3,695 genes at the heading+21 days stage. In the Dongjin/Heugseol combina-

tion, the expression of 928 upregulated genes increased to 2,305 genes at the heading+21 days stage. In the Heugjinju/Heugseol combination, the expression of 528 upregulated genes increased to 1,682 genes at the heading+21 days stage.

Finally, 365 genes were upregulated in the two black rice cultivars at all developmental stages and showed a significant correlation ($r=0.6503$, $p=0.0325$), suggesting that these genes may be associated with pigmentation metabolism. We also identified 57 genes that demonstrated downregulation and showed a significant correlation ($r=0.4391$, $p=0.0453$). We assumed these genes were related to inhibition of pigmentation production or anthocyanin metabolism.

Clusters of orthologous groups (COGs) analysis

COGs analysis was performed on the 12,673 candidate genes using the NCBI/COGs database (<http://www.ncbi.nlm.nih.gov/COG/>). The upregulated genes were classified as general function (21.6%), signal transduction (12.3%), and posttranslational modification, protein, and chaperones (8.2%). Of the regulated genes, 3,271 genes (25.8%) were poorly characterized, and 617 of those genes (4.8%) had an unknown function. The COGs related to anthocyanin function were assumed due to the presence of highly conserved genes between two specific categories: post-translational modification, protein turnover, and chaperones (8.2%) and carbohydrate transport and metabolism (4.7%).

GO analysis associated with anthocyanin production

The processes that were most affected during anthocyanin production were determined by GO analysis. First, GoMiner categorizes each gene according to its GO terms and mode of gene expression. Modes of expression are denoted as under, over, and change to represent genes that were downregulated, upregulated, or both, respectively. With this analysis, 1,558 and 295 genes were identified as orthologous to *Arabidopsis* genes and up- or downregulated, respectively. GoMiner analysis demonstrated that these genes were involved in 25, 19, and 12 biological processes, cellular components, and molecular functions, respectively. Figure 1 shows the composition ratio of genes in the each category.

Candidate genes associated with anthocyanin pigmentation

Genes identified through COGs and GO analysis were compared to the rice genome database at <http://rgp.dna.affrc.go.jp/E/IRGSP/> and the rice genome system supported by NAAS website <http://nabic.naas.go.kr/>. Comparison with these databases enabled elimination of genes related

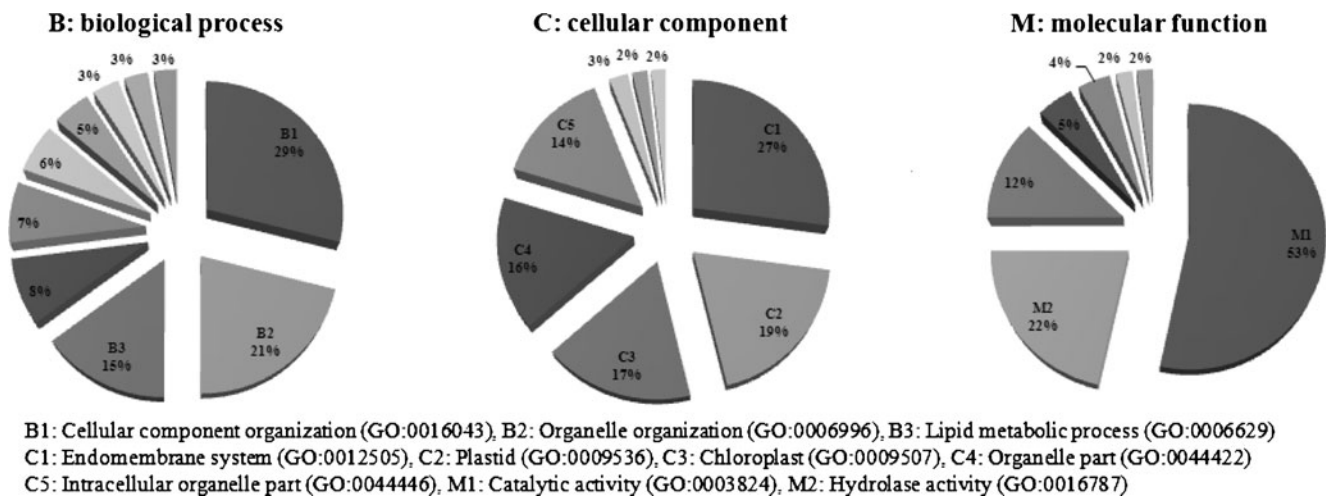


Fig. 1 GO analysis of anthocyanin production biosynthesis. The three process categories show the composition ratio of genes

to pigments other than anthocyanin from the candidate gene list. Ultimately, 1,289 candidate genes were predicted to be related to anthocyanin biosynthesis and production.

Transcription factors associated with anthocyanin

We characterized the transcription factors associated with anthocyanin pigmentation using the hyper-geometric analysis method. Table 1 shows the transcription factor groups predicted from the 1,289 candidate genes.

We identified 10 groups that exhibit functionally diverse transcription factor activity involved in anthocyanin biosynthesis. Of the 1,289 candidate genes, 137 (10.6%) were identified as putative transcription factors. The MYB and GT transcription factor families showed the highest expression levels with these groups accounting for approximately 37% of the predicted transcription factors. Other putative groups of anthocyanin production included the NAC families, basic helix-loop-helix (bHLH) and peroxisome proliferator-activated receptor binding protein (PBP).

Involvement of the Myb and GT families, NAC, bHLH and PBP transcription factors in colored tissues has been reported previously (Martin and Paz-Ares 1997; Yujie et al. 2010; Reddy et al. 1998; Borevitz et al. 2000; Sheng et al. 2005). These results illustrate the functional diversity of the transcription factor families and that these factors may be highly activated during anthocyanin pigmentation.

MYB family members activate flavonoid pigmentation, are involved in specific steps of the anthocyanin pathway, and have evolved specific functions with different biochemical properties. The GT factors, GT-1 and GT-2, have been shown to interact with multiple sequences within the promoter of the *Tdc* gene and express a *trans*-acting factor possessing a GT-motif. GT-1-containing genes encoded factors such as the nuclear protein of light-responsive in rice. NAC factors play an important role in regulating the expression of flavonoid biosynthesis-related genes (Morishita et al. 2009). The bHLH factors function in determining red seed color in anthocyanin biosynthesis (Gonzalez et al. 2008). PBP factors may be closely related to the flower-specific Myb305 factor

Table 1 Transcription factor groups (TFs) predicted by the cumulative hyper-genomic distribution analysis method

TFs	Description	Gene	P-value	Ratio
MYB	A life cycle activating related to metabolism of cell	23	0.0413	16.79
MYB-related	Flavonoid -biosynthesis related to Myb-like protein	10	0.0319	7.30
GT family	Nuclear factors related to trans-acting	18	0.0325	13.14
WRKY	To biotic, abiotic stresses and development process	17	0.0287	12.41
NAC	Genes related to flavonoid biosynthetic pathway	16	0.0317	11.68
bHLH	Anthocyanin biosynthesis and signaling	12	0.0329	8.76
PBF	The regulates cereal storage protein gene	10	0.0108	7.30
MADS	Determination of floral meristem and organ identity	9	0.0272	6.57
AP2	Ethylene responsive element binding factor	8	0.0305	5.84
PBP	Flower-specific gene related to Myb305	3	0.0496	2.19
Etc.	STF, GRF, bZIP, RAV and C2H2 factors	11	0.1258	8.03

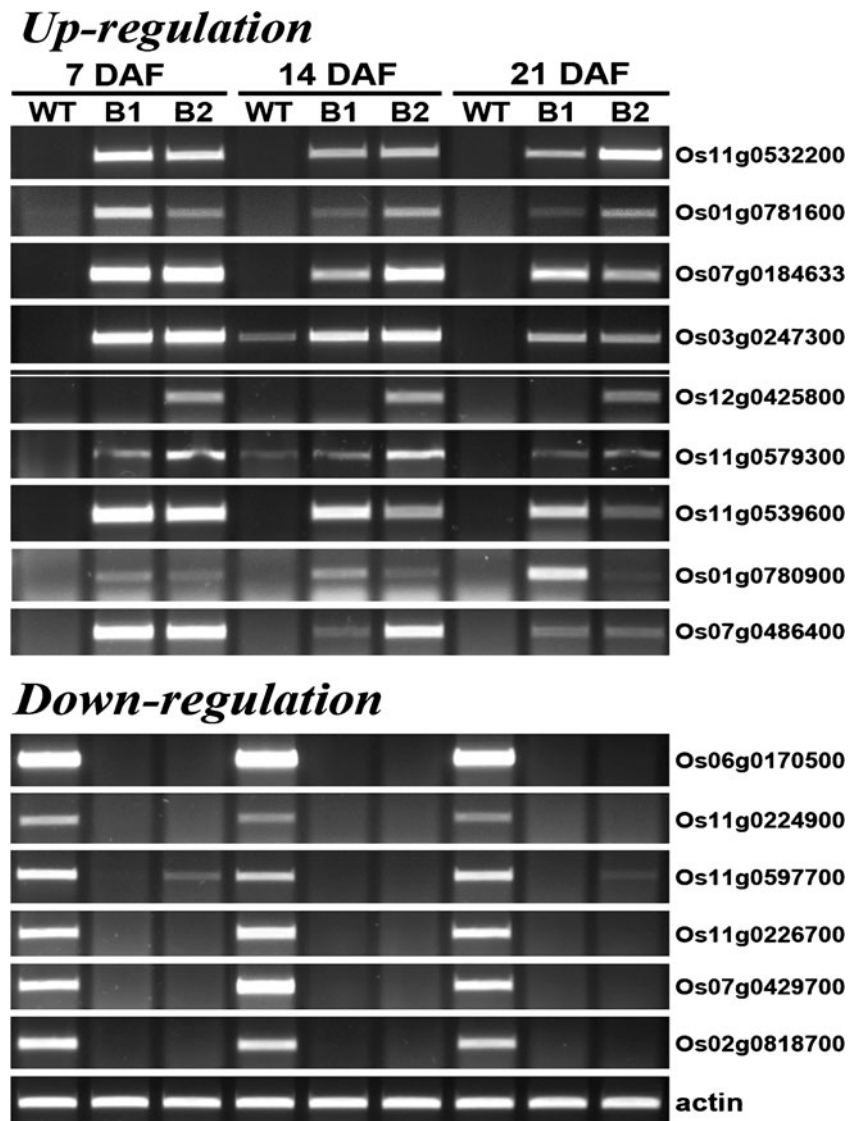
and MYB-like protein (Reddy et al. 1998). The WRKY factors encode the seed coat development gene of *Arabidopsis* (Johnson et al. 2002), and several MADS box transcription factors are expressed preferentially in flowers and cause early flowering when ectopically expressed in rice plants (Jeon et al. 2000). Papillomavirus binding factor (PBF) regulates expression of cereal storage protein, which interacts with Opaque-2 and Related to ABI3/VP1 1 (RAV) factors, to act as negative regulators of growth. APETALA2 (AP2) is related to seed development, ethylene-responsive expression, and floral morphogenesis (Jofuku et al. 1994; Shukla et al. 2006).

Unknown genes involved in anthocyanin biosynthesis

Among the 137 transcription factor genes identified above, the expression of 24 unknown and hypothetical genes was

found at each seed developmental stage. Among these, 14 showed a pattern of upregulation, and 10 genes were down-regulated. Therefore, this method is able to efficiently screen the unknown genes associated with specific biological process. Finally, 17 unknown and hypothetical genes differed between the *Donjin* cultivar and the two black cultivars at all seed developmental stages. Of the 17 candidate genes, 10 showed a pattern of upregulation and 7 appeared to be downregulated genes. These genes most likely play either a regulatory role in the anthocyanin production process or are related to anthocyanin metabolism during flavonoid biosynthesis. Additional function studies of these genes are necessary to further characterize possible regulatory factors of the anthocyanin pathway and pigmentation metabolism. Nevertheless, these 17 transcription factor genes play a potential role in anthocyanin production. This study provides valuable insight into anthocyanin pigmentation production

Fig. 2 RT-PCR analysis result of 15 candidate genes, representing nine up-regulated and six down-regulated genes. A total 27 RNA samples were isolated from three cultivars and three developmental stages, in triplicate. Actin was used as a control. DAF = Day after heading, WT = Dongjin (white color seed of no anthocyanin), B1 = Heugjinju (black color seed of lower anthocyanin), B2 = Heugseol (black color seed of higher anthocyanin)



and will greatly facilitate the future breeding of anthocyanin-rich hybrid rice varieties.

RT-PCR analysis of selected genes

RT-PCR analysis was performed on the 17 unknown and hypothetical genes using 27 RNA samples representing each cultivar. Among the 17 genes, nine showed an up-regulated pattern similar to white rice at all seed developmental stages tested, while six genes exhibited a down-regulated pattern. The remaining two genes were both up- and down-regulated, and were omitted as a result of small differences with control rice (Fig. 2).

Most of the up- and down-regulated genes appear to be related to major biological changes induced in the anthocyanin biosynthesis pathway. The results of RT-PCR

analysis coincided largely with the expression profiles obtained through the microarray hybridization experiments. Interestingly, four unknown genes (Os07g0184633, Os03g0247300, Os11g0539600, and Os07g0486400) were highly induced during the early heading stage in the two black rice cultivars. This result suggested that these four unknown genes may play a role in anthocyanin production in the early rice heading stage. However, the Os01g0780900 gene is highly induced during late heading stage. Gene Os12g0425800 demonstrates a difference between Heugjinju and Heugseol. This gene was not expressed in the cv. Dongjin without anthocyanin, was weakly induced in the Heugjinju cultivar with low anthocyanin, and was highly expressed in the Heugseol cultivar with high anthocyanin. Therefore, this gene may play a role in controlling anthocyanin levels.

Table 2 RT-PCR primer list of the 15 candidate genes consisting of nine up-regulated and six down-regulated genes

Locus ID	Primer	Sequence (5' → 3')	Product size/bp
Up-regulation			
Os11g0532200	U1 F	GCCACTGTTACGTCGCCGGA	619
	U1 R	CCCACGGCAACCACGTCTCC	
Os01g0781600	U2 F	GGCTCCAGTGCGGCAGGAGA	304
	U2 R	TCACGGTCATGTCCACGCC	
Os07g0184633	U3 F	TCCGATCCCTGGGTCGCTCG	397
	U3 R	GGGGCAGCTCAAGGCAAGGT	
Os03g0247300	U4 F	TTCGTCGCCATGGGGTTGCC	604
	U4 R	GCCATCAACCAGCCCGGCAT	
Os12g0425800	U5 F	GACGGGATGGGTCATGCCGGC	378
	U5 R	GAGGACATCAGCCGCAAGATTC	
Os11g0579300	U6 F	CAGCAGCTGCGGAAGGTTGC	622
	U6 R	ACACGCTCACGCACAGCAGA	
Os11g0539600	U7 F	TGCGTGCGCAAGTGTAGGCT	537
	U7 R	AAGGCGCAGGCTCCATGCAA	
Os01g0780900	U8 F	GACGGGAGGGAGAGGGGTCCG	850
	U8 R	TCCCCACGGGGCTCATCAGG	
Os07g0486400	U9 F	GCCGAGTACCCGCAAGGTGG	607
	U9 R	GTTGGAAGCCACGCGCCTCT	
Down-regulation			
Os06g0170500	D1 F	GGCAAACCCGGTCCACAGCA	541
	D1 R	ACAGGCTCCAAGCGTTCGGC	
Os11g0224900	D2 F	TGCTCACGAGGGTTGCCAGC	443
	D2 R	CAAGCGGCAAGCCCCACAT	
Os11g0597700	D3 F	TGTCAAGGCCCAACGCCGTG	788
	D3 R	GCTCCGCTGGCGACAAACCT	
Os11g0226700	D4 F	GAGCCATGCAGCGCTGGGAA	476
	D4 R	TGCTGCACCGACCAAGGTA	
Os07g0429700	D5 F	CATCAGCAGCGGGAGGCAATG	455
	D5 R	GGGACCAACAGATGGATCGAC	
Os02g0818700	D6 F	GCTCCATCGGCTCTGTCCACCG	520
	D6 R	CTTGACGCCATAGCCCCGCA	

It was assumed that the down-regulated genes were related to inhibition of anthocyanin metabolism or had a negative effect on the anthocyanin biosynthesis pathway. Among the six down-regulated genes, Os06g0170500 showed levels of expression in all developmental stages of the Dongjin cultivar. Table 2 shows the primer list of the 15 candidate genes. Two genes were omitted due to small differences with control rice. Each primer used to amplify the candidate genes.

Conclusions

There has been increased interest in the use of black rice due to the numerous health benefits associated with the anthocyanin pigment. In this study, we identified putative transcription factors involved in anthocyanin biosynthesis in black rice cultivars using a newly designed 135 K *Oryza sativa* microarray. We performed a four-stage analysis to evaluate anthocyanin biosynthesis in black rice: first, we tested for the presence of commonly expressed genes between different rice cultivars, second, functional classification categories were identified by COGs and GO analysis, third, transcription factors were identified by the hyper-geometric analysis method, and finally, selected unknown and hypothetical genes were verified by RT-PCR. Among the 137 transcription factor genes identified, 24 genes were unknown and hypothetical genes. Among these, 17 genes differed between the white cultivar and the two black rice cultivars at all the seed developmental stages. The 17 candidate genes consist of 10 and 7 genes were up- and down-regulated gene, respectively. These genes may play a regulatory role in anthocyanin production, or are related to anthocyanin metabolism during flavonoid biosynthesis.

Results of RT-PCR analysis largely coincided with the expression profiles obtained through the microarray experiments. While these genes still require further investigation and validation, these results demonstrate the potential of this method using newly designed microarrays. This study provides valuable insight into pigmentation production and will greatly facilitate future breeding of anthocyanin-rich hybrid varieties of rice.

Acknowledgements This work was supported in part by the National Academy of Agricultural Science, Bio-green 21 Program of Rural Development Administration, and with a grant from the Crop Functional Genomics Center of the 21st Century Frontier Research Program funded by the Ministry of Science and Technology, Seoul, Korea.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Abdel-Aal E-SM, Young JC, Rabalski I (2006) Anthocyanin composition in black, blue, pink, purple, and red cereal grains. *J Agric Food Chem* 54:4696–4704
- Gonzalez A, Zhao M, Leavitt JM, Lloyd AM (2008) Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in Arabidopsis seedlings. *Plant J* 53:814–827
- Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 25:25–29
- Borevitz JO, Xia Y, Blount J, Dixon RA, Lamb C (2000) Activation tagging identifies a conserved MYB regulator of phenylpropanoid biosynthesis. *Plant Cell* 12:2383–2393
- Choi YH, Yang YI, Lee SB, Lee KS, Lee JH, Hong HC et al. (2008) A new black pericarp and processing soft rice cultivar 'Heugseol'. *Korean Crop Research* 9:27–36
- Dessimoz C, Boeckmann B, Roth AC, Gonnet GH (2006) Detecting non-orthology in the COGs database and other approaches grouping orthologs using genome-specific best hits. *Nucleic Acids Res* 34:3309–3316
- Ding M, Feng R, Wang SY, Bowman L, Lu Y, Qian Y, Castranova V, Jiang BH, Shi X (2006) Cyanidin-3-glucoside, a natural product derived from blackberry, exhibits chemopreventive and chemotherapeutic activity. *J Biol Chem* 281:17359–17368
- Holton TA, Cornish EC (1995) Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* 7:1071–1083
- Hu C, Zawistowski J, Ling WH, Kitts DD (2003) Black rice (*Oryza sativa* L. *indica*) pigmented fraction suppresses both reactive oxygen species and nitric oxide in chemical and biological model systems. *J. Agric. Food Chem* 51:5271–5277
- Hyun JW, Chung HS (2004) Cyanidin and malvidin from *Oryza sativa* cv. Heugjinjubyeo mediate cytotoxicity against human monocytic leukemia cells by arrest of G2/M phase and induction of apoptosis. *J Agric Food Chem* 52:2213–2217
- Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 31:e15
- Jeon JS, Jang S, Lee S, Nam J, Kim C, Lee HS, Chung YY, Kim SR et al. (2000) *Leafy hull sterile1 (lhs1)* is a homeotic mutation in a rice MADS box gene affecting rice flower development. *Plant Cell* 12:871–884
- Jofuku KD, den Boer BG, Van Montagu M, Okamoto JK (1994) Control of Arabidopsis flower and seed development by the homeotic gene *APETALA2*. *Plant Cell* 6:1211–1225
- Johnson CS, Kolevski B, Smyth DR (2002) TRANSPARENT TESTA GLABRA2, a trichome and seed coat development gene of Arabidopsis, encodes a WRKY transcription factor. *Plant Cell* 14:1359–1375
- Kim BG, Kim JH, Min SY, Shin KH, Kim JH, Kim HI, Ryu SN, Ahn JH (2007) Anthocyanin content in rice is related to expression level of anthocyanin biosynthetic genes. *J Plant Biol* 50:156–160
- Kim CK, Kikuchi S, Satoh K, Kim JA, Kim DH, Kim YH, Park SH, Yoon YH (2009) Genetic analysis of seed-specific gene expression for pigmentation in colored rice. *BioChip J* 3:125–129
- Kim CK, Kikuchi S, Kim YK et al. (2010) Computational identification of seed-specific transcription factors involved in anthocyanin production in black rice. *BioChip J* 3:247–255
- Kong JM, Chia LS, Goh NK, Chia TF, Brouillard R (2003) Analysis and biological activities of anthocyanins. *Phytochemistry* 64 (5):923–933
- Martin C, Paz-Ares J (1997) MYB transcription factors in plants. *Trends Genet* 13:67–73

- Moon HP, Choi YG, Lee JH, Jung KH, Cho SY, Hwang HG et al. (1998) A new early maturing, anthocyanin pigmented rice variety 'Heugjinjubyeo'. *J Breed Sci* 30(4):383–383
- Morishita T, Kojima Y, Maruta T, Nishizawa-Yokoi A, Yabuta Y, Shigeoka S (2009) Arabidopsis NAC transcription factor, ANAC078, regulates flavonoid biosynthesis under high-light. *Plant Cell Physiol* 50:2210–2222
- Nam SH, Choi SP, Kang MY, Koh HJ, Kozukue N, Friedman M (2006) Antioxidative activities of bran extracts from twenty one pigmented rice cultivars. *Food Chem* 94:613–620
- Philpott M, Gould KS, Lim C, Ferguson LR (2006) In situ and in vitro antioxidant activity of sweetpotato anthocyanins. *J Agric Food Chem* 54:1710–1715
- Quattrocchio F, Wing JF, Leppen H, Mol J, Koes RE (1993) Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *Plant Cell* 5:1497–1512
- Reddy VS, Scheffler BE, Wienand U, Wessler SR, Reddy AR (1998) Cloning and characterization of the rice homologue of the maize C1 anthocyanin regulatory gene. *Plant Mol Biol* 36:497–498
- Remm M, Storm C, Sonnhammer E (2001) Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J Mol Biol* 314:1041–1052
- Ryu SN, Park SZ, Ho CT (1998) High performance liquid chromatographic determination of anthocyanin pigments in some varieties of black rice. *J Food Drug Anal* 6:729–736
- Ryu SN, Han SJ, Park SZ, Kim HY (2000) Antioxidative activity and varietal difference of cyanidin 3-glucoside and peonidin 3-glucoside contents in pigmented rice. *Korean J Crop Sci* 45(4):257–260
- Sadlier DM, Connolly SB, Kieran NE, Roxburgh S, Brazil DP, Kairaitis L, Wang Y, Harris DC, Doran P, Brady HR (2004) Sequential extracellular matrix-focused and baited- global cluster analysis of serial transcriptomic profiles identifies candidate modulators of renal tubulointerstitial fibrosis in murine adriamycin-induced nephropathy. *J Biol Chem* 279:29670–29680
- Sheng T, Joost K, Maarten K, Sjef S (2005) Sucrose-Specific Induction of Anthocyanin Biosynthesis in Arabidopsis requires the MYB75/PAP1 gene. *Plant Physiol* 139:1840–1852
- Shukla RK, Raha S, Tripathi V, Chattopadhyay D (2006) Expression of CAP2, an APETALA2-Family Transcription Factor from Chickpea, Enhances Growth and Tolerance to Dehydration and Salt Stress in Transgenic Tobacco. *Plant Physiol* 142:113–123
- Takashi I, Bing X, Yoichi Y, Masaharu N, Tetsuya K (2001) Antioxidant Activity of Anthocyanin Extract from Purple Black Rice. *J Med Food* 4:211–218
- Troester MA, Millikan RC, Perou CM (2009) Microarrays and epidemiology: ensuring the impact and accessibility of research findings. *Cancer Epidemiol Biomark Prev* 18(1):1–4
- Tsuda T, Horio F, Osawa T (2002) Cyanidin 3-O- β -glucoside suppresses nitric oxide production during a zymosan treatment in rats. *J Nutr Sci Vitaminol* 48:305–310
- Tsuda T, Horio F, Uchida K, Aoki H, Osawa T (2003) Dietary cyanidin 3-O- β -D-glucoside-rich purple corn color prevents obesity and ameliorates hyperglycemia. *J Nutr* 133:2125–2130
- Workman C, Jensen LJ, Jarmer H, Berka R, Gautier L, Nielsen HB, Saxild HH, Nielsen C, Brunak S, Knudsen S (2002) A new non-linear normalization method for reducing variability in DNA microarray experiments. *Genome Biol* 3(0048):1–0048, 16
- Xia M, Ling WH, Ma J, Kitts DD, Zawistowski J (2003) Supplementation of diets with black rice pigment fraction attenuates atherosclerotic plaque formation in apolipoprotein E deficient mice. *J Nutr* 133:744–751
- Xia X, Ling W, Ma J, Xia M, Hou M, Wang Q, Zhu H, Tang Z (2006) An anthocyanin-rich extract from black rice enhances atherosclerotic plaque stabilization in apolipoprotein E-deficient mice. *J Nutr* 136:2220–2225
- Yujie F, Kabin X, Xin H, Honghong H, Lizhong X (2010) Systematic analysis of GT factor family of rice reveals a novel subfamily involved in stress responses. *Mol Genet Genomics* 283:157–169
- Zeeberg BR, Feng W, Wang G, Wang MD, Fojo AT, Sunshine M, Narasimhan S, Kane DW, Reinhold WC, Lababidi S, Bussey KJ, Riss J, Barrett JC, Weinstein JN (2003) GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* 4:R28
- Zhang Z, Kou X, Fugal K, McLaughlin J (2004) Comparison of HPLC methods for determination of anthocyanins and anthocyanidins in bilberry extracts. *J Agric Food Chem* 52:688–691
- Zhang MW, Guo BJ, Peng ZM (2005) Genetic effects on grain characteristics of indica black rice and their uses on indirect selections for some mineral element contents in grains. *Genet Resour Crop Ev* 52:1121–1128