

Genetic Analysis of Seed-specific Gene Expression for Pigmentation in Colored Rice

ChangKug Kim¹, Shoshi Kikuchi², Kouji Satoh²,
JinA Kim³, DongHern Kim⁴, YongHwan Kim¹,
SungHan Park¹, JeomHo Lee⁵ & UngHan Yoon¹

¹Genomics Division, ³Functional Bio-Material Division and ⁴Bio-crop Development Division, National Academy of Agricultural Science (NAAS), RDA, Suwon 441-707, Korea

²Division of Genome and Biodiversity Research, National Institute of Agrobiological Sciences (NIAS), Tsukuba, Ibaraki 305-8602, Japan

⁵On-Farm Research Division, Rural Development Administration (RDA), Korea

Correspondence and requests for materials should be addressed to U.H. Yoon (uhyoon@rda.go.kr)

Accepted 4 May 2009

Abstract

Anthocyanin is a major class of flavonoids that tissues produce in response to environmental signals. The pigmentation of colored rice was analyzed by oligo microarray based on two factors, cultivar and developmental stage. As many as 250 and 350 genes were identified to be significantly up- and down-regulated. Using hypergeometric analysis for transcription factor function, the Myb and GT families, PBP, PBF, RAV, and STF factors were identified have potential anthocyanin-specific functions. We also obtained seventeen unknown genes which display novel functions. Among the genes within the GT-1 and Myb1 groups, three unknown genes were found to be up-regulated significantly in both combinations of cultivar and developmental stage. These results showed functional diversity of TF families and that the biological functions of the particular TFs may be activated in a pigmentation pathway in rice.

Keywords: Pigmentation, Colored rice, Microarray, Anthocyanin, Transcription factor

Introduction

Rice, one of the most important cereal crops in the world, is a significant staple for feeding much of the world's population. Although white rice is mostly consumed, there are many special cultivars that contain color pigments, such as black and red rice. Anthocya-

nin pigmentation has been genetically studied in rice for several decades. This pigment composes a group of natural colorants belonging to the flavonoid family, and is produced by tissues in response to developmental and environmental signals¹. The genes regulating anthocyanin, which exhibit significant homology to the *R* and *C1* families in maize, have been identified²⁻⁴. Moreover, the regulatory elements that confer tissue-specific accumulation of anthocyanin have been characterized in several plant species⁵. The anthocyanins present in black rice are chrisanthemin, keracyanin, and peonidin 3-glucoside, while those found in red rice are catechin and catechatannin⁶. In addition, pigmented rice contains an acetylated procyanidin, a common anthocyanin with free radical activity⁷. Classical genetic studies identified the red seed color to be encoded by basic helix-loop-helix (bHLH)-containing proteins and dihydroflavonol 4-reductase gene products^{8,9}. Red pigmentation, which is prevalent amongst ancestors of *Oryza Sativa*, is closely associated with seed shattering and dormancy. One gene found to be involved in the production of this red color is *Rc*, which encodes a bHLH protein¹⁰.

Many eukaryotic transcription factors (TFs) consist of at least two discrete domains, a DNA binding domain and an activation or repression domain¹¹. Genes regulated by Myb factors play important roles in many developmental processes and defense responses in plants¹². For instance, many TFs harboring the R2R3-Myb domain have been identified in various plant species as regulators of flavonoid biosynthesis^{13,14}. To assess the functional significance of regulatory elements, a computational approaches, such as cumulative hypergeometric distribution analysis, have been performed^{15,16}.

A recent report showed that diet supplementation with pigment from black rice markedly reduced oxidative stress in mice¹⁷. Data has also suggested that an anthocyanin-rich extract from black rice improves the performance of certain metabolic pathways associated with diets high in fructose¹⁸. These marked health benefits have been attributed to the antioxidant property of anthocyanin¹⁹. Due to these and other recent reports stating the various potential health benefits of anthocyanin²⁰, there has been increased interest on the use of this pigment in functional food research using the new microarray technology as nanotechnology^{21,22}. Therefore, this study was performed to identify by

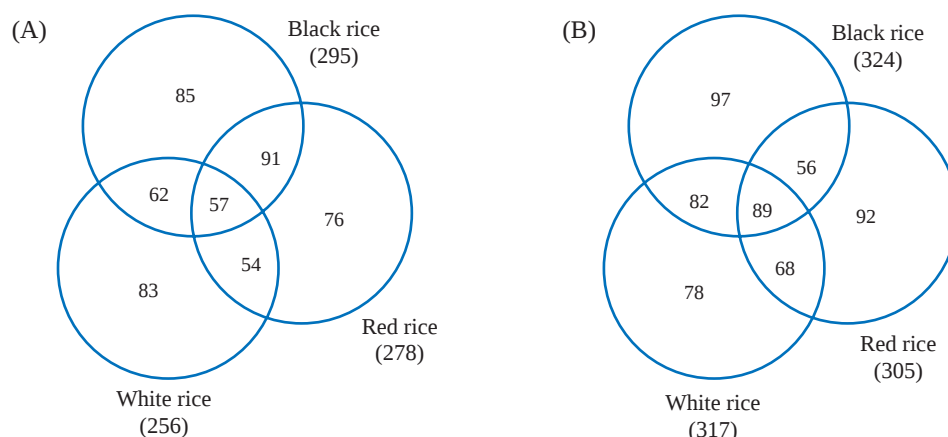


Figure 1. The number of up- and downregulated genes in the three cultivars. The diagrams show the number of overlapping and non-overlapping genes that were (A) upregulated or (B) downregulated significantly during the different developmental stages in black, red, and white rice. The numbers in parentheses indicate the total number of genes with a greater than 2.5-average change in expression in the three cultivars.

microarray novel genes that consistently associate with anthocyanin pigmentation.

Results and Discussion

An Evaluation of Selected Genes

To evaluate the pigmentation of colored rice, we performed a two-stage screen. First, we performed hypothesis testing to find a common gene from various combinations of colored rice, such as black/white, red/white, and black/red, and verified the significance of the data. Then, we evaluated the transcription factors (TFs) involved in pigmentation by average linkage clustering and the hypergeometric analysis method. From this second step, we retained those genes whose expression was related to color production without differences in developmental stage and cultivar. The data was screened to select genes with a greater than 2.5 average upregulation (or downregulation) using three cultivars across three developmental stages. To verify the hypothesis that a common gene for pigmentation exists among colored rice, we separated the samples into two groups according to cultivar and developmental stage. The cultivar, ANOVA was used to analyze the variation from all three cultivars. The developmental stages were used to analyze the variation to compare the three different developmental stages. Both groups were significant in regards to cultivar (P -value=0.0012) and developmental stage (P -value=0.0001).

The number of genes expressed significantly in each cultivar was between 250 and 350. Fifty-seven genes were expressed in all three cultivars (Figure 1A). In

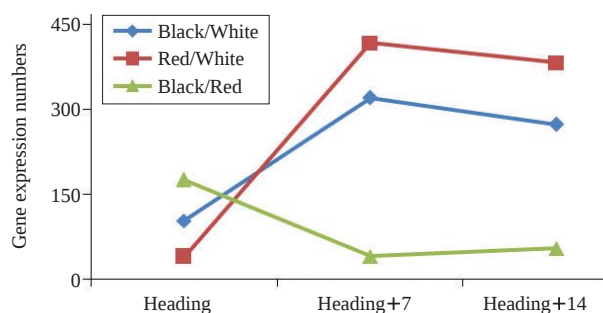


Figure 2. Changes in gene expression compared to black, red, and white rice combination. The X-axis represents time of rice sampling and developmental stage, i.e., heading, after 7 days, and heading after 14 days.

black/red rice, ninety-one genes common to both showed a significant correlation ($R=0.7612$, $P=0.0315$), suggesting that these genes may be associated with pigmentation metabolism. In contrast, eighty-nine genes, of which most coded for storage-induced proteins, were downregulated amongst the three cultivars (Figure 1B).

Next, we compared the effects of variation in developmental stage tested (e.g., heading, heading +7 days, heading +14 days) to the combination of cultivars (e.g. black/white, red/white, and black/red). In the black/white rice combination, the expression of 103 genes increased significantly in the heading stage, and this number increased to 320 genes by the heading +7 days stage. In the black/red rice combination, 175 genes were increased significantly in the heading stage, but this number was reduced in subsequent develop-

mental stages (Figure 2).

The selected genes were compared to the International Rice Genome Sequencing Project (IRGSP, <http://rgp.dna.affrc.go.jp/E/IRGSP/>) database and annotated by the rice genome system supported by National Academy of Agricultural Science (NAAS). During comparison to the IRGSP database, genes related to tannin were removed. Ultimately, 304 candidates were found to be significant genes related to anthocyanin.

Potential Transcription Factors Associated with Pigmentation

The selected 304 genes were analyzed by average linkage clustering to group genes with similar function. Clustering analysis was performed according to the median intensity measured from the array. To prevent a complicated model, we assigned a value of 0.35 as the average distance between clusters. From this, seven clusters were identified (Table 1).

Among the 304 genes, 79% were related to anthocyanin and Myb-induced proteins. Most genes within the first cluster were tissue-specific and related to Myb protein activation involved in flavonoid biosynthesis.

Table 1. The attribution of gene groups by clustering analysis.

Clusters	Attribution	Numbers
1	A flower-specific Myb-like protein family	112
2	Nuclear factors GT-1 interact within the Tdc-promoter	95
3	Cereal storage protein and carbohydrate gene expression	39
4	A trans-acting factor that binds to a GT-motif	24
5	Closely be related to the flower specific in Myb-like protein	13
6	Negative growth regulator in a signal pathway	12
7	Auxin-responsive gene and a novel binding factor	9

These genes were triggered at the very beginning of rice coloration. Clusters 1 and 2 contained genes potentially involved in particular steps of the anthocyanin pathway. The 39 genes belonging to cluster 3 were expressed at the ripening stage and their expression levels increased continuously from the heading stage to 14 days thereafter. Cluster 5 grouped genes that may be closely related to the flower-specific Myb-like protein. Finally, *Oryza sativa* growth-regulating factor and auxin response factor were categorized into clusters 6 and 7, respectively.

To identify the best candidate genes, transcription factor analysis was performed on genes consistently found to be associated with rice pigmentation. Table 2 shows the transcription factor groups predicted by the cumulative hypergeometric distribution analysis method utilized in this study.

Involvement of the Myb and GT families, Peroxisome proliferator-activated receptor binding protein (PBP), and Papillomavirus binding factor (PBF) have already been reported in colored tissues^{2,4}. Most of these genes are related to the production of anthocyanins, as well as other flavonoid compounds such as flavonols and proanthocyanidins. It is likely that Myb genes are present in all higher plants and have evolved specific functions with different biochemical properties²³. The Myb genes most likely code for proteins involved in specific steps of the anthocyanin pathway such as that catalyzed by the putative leucoanthocyanidin dioxygenase (AK110790, AK068517, and AK069008). 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. The GT-1-containing genes encoded factors such as the nuclear protein of light-responsive in rice (AK106249 and AK108243). The PBP factor may be closely related to the flower-specific Myb305 factor and Myb-like protein (AK073606, AK111740, and AK111807). The PBF factor regulates cereal storage protein gene expression, which interacts with Opaque-2 and Related to ABI3/VP1 1 (RAV) factors to act as negative grow-

Table 2. Results of transcription factor (TF) analysis performed on selected gene expression data associated with colored rice.

TFs	Description	Gene	P-value	TF ratio
MYB1	A flower-specific related to Myb protein	69	0.0323	48.25
MYB26	A Myb-like protein of flowers	16	0.0095	11.18
MYB305	Gene related to flavonoid-biosynthesis	30	0.0298	20.98
GT-1	Nuclear factors related to trans-acting	91	0.0180	63.63
GT-2	A trans-acting factor in GT-motif	27	0.0217	18.89
PBF	The regulates cereal storage protein gene	31	0.0319	21.67
PBP	Flower-specific gene related to Myb305	15	0.0087	10.48
RAV	Negative growth regulator gene	8	0.0273	5.59
STF	Auxin-responsive gene	6	0.0173	4.19

th regulators during growth stages (AK103508 and AK66400).

There were seventeen unknown genes among the 304 genes selected and categorized. Within the GT-1 and Myb1 groups, 6 and 5 unknown genes were identified in the black/red rice combination, respectively. Among these two groups, the expression of three unknown genes (AK059529, AK105918, and AK060724) differed dramatically between the black/red rice and developmental stages with great statistical significance (P -value=0.0016). These three genes may potentially play either a regulatory role in the developmental process or be related to anthocyanin metabolism. Functional studies are necessary to clarify the nature of these three new candidates as possible regulatory factors of the anthocyanin pathway in colored rice.

In conclusion, our results identified nine groups of genes that exhibit functionally diverse TF activity involved in rice pigmentation. Seventeen unknown genes were also revealed to potentially play a role in this biological process. These particular TFs may be highly activated in a pigmentation pathway in colored rice.

Materials and Methods

Rice

The samples used represent three different genotypes, namely white rice (Manmibyeo: Milyang162), red rice (Jinjubyeo: Milyang194), and black rice (Hukcholbyeo: Suwon451). The colored rice was harvested at different developmental stages (heading, heading +7 days, heading +14 days) during the fall season in 2006 in Suwon, Korea. The harvested seeds were immediately frozen in liquid nitrogen at the field and stored at -80°C .

Statistical Analysis and Scanning of the Oligo Microarray

The oligo microarray experiments were performed using Agilent Rice 60-mer 22K arrays which were designed and validated by the National Institute of Agrobiological Sciences (NIAS, <http://www.nias.affrc.go.jp/>) in Japan. The microarray includes oligonucleotide probes against 21,495 genes from the genome of *Oryza sativa* L. ssp japonica (cultivar Nipponbare). All experiments were run using only Cy3 in order to eliminate the dye-swap error value. Spot intensity was calculated as the median value of the spot compared to the background median value. Samples were separated into two groups, cultivar and developmental stage, in order to test our hypothesis. For gene expression clustering, eigen values of the data were first

generated with SAS Enterprise software and then statistically analyzed with TM4 software developed by the J. Craig Venter Institute. Data were subjected to a t-test ($\alpha=0.01$) where experiments were assigned to white, black, or red groups. Transcription factor analysis was performed by the cumulative hypergeometric distribution analysis method¹⁵.

References

1. Quattrocchio, F., Wing, J.F., Leppen, H.T.C., Mol, J.N.M. & Koes, R.E. Regulatory genes controlling anthocyanin pigmentation are functionally conserved among plant species and have distinct sets of target genes. *Plant Cell* **5**, 1497-1512 (1993).
2. Takahashi, M. Gene analysis and its related problems. *J. Fac. Agr. Hokkaido University* **61**, 91-142 (1982).
3. Hu, J., Reddy, V.S. & Wessler, S.R. The rice R gene family: two distinct subfamilies inverted repeat transposable elements. *Plant Mol. Biology* **42**, 667-678 (2000).
4. Reddy, V.S., Scheffler, B.E., Wienand, U., Wessler, S.R. & Reddy, A.R. Cloning and characterization of the rice homologue of the maize C1 anthocyanin regulatory gene. *Plant Mol. Biology* **36**, 497-498 (1998).
5. Holton, T.A. & Cornish, E.C. Genetics and biochemistry of anthocyanin biosynthesis. *Plant Cell* **7**, 1071-1083 (1995).
6. Clifford, M.N. Anthocyanins: Nature, occurrence and dietary burden. *Journal of Science Food Agriculture* **80**, 1063-1072 (2000).
7. Oki, T. *et al.* Polymeric procyanidins as radical-scavenging components in red-hulled rice. *J. Agric. Food Chemistry* **50**, 7524-7529 (2002).
8. Sweeney, M.T., Thomson, M.J., Pfeil, B.E. & McCouch, S. Caught red-handed: Rc encodes a basic helix-loop-helix protein conditioning red pericarp in rice. *Plant Cell* **18**, 283-294 (2006).
9. Furukawa, T. *et al.* The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp. *Plant Journal* **4**, 90-102 (2007).
10. Megan, T. *et al.* Caught red-handed: Rc encodes a bHLH protein conditioning red pericarp in rice. *Plant Cell* **18**, 288-294 (2006).
11. Ptashne, M. How eukaryotic transcriptional activators work. *Nature* **335**, 683-689 (1988).
12. Li, J. *et al.* A subgroup of MYB transcription factor genes undergoes conserved alternative splicing in Arabidopsis and rice. *Journal Experimental Botany* **57**, 1263-1273 (2006).
13. Hironori, M., Fumiaki, O., Kazuhito, S., Hiromi, H. & Yuzo, M. Isolation of a regulatory gene of anthocyanin. *Plant Physiology* **143**, 1252-1268 (2007).
14. Wunschiers, R., Axelson, R., Vellguth, M. & Lindblad, P. Experimental and bioinformatic approaches for analyzing and visualizing cyanobacterial nitrogen and

- hydrogen metabolism. *Electronic Journal of Biotechnology [online]* **10**, DOI: 10.2225/vol10-issue4-full-text-2 (2007).
15. Priya, S., Yitzhak, P. & Church, M.C. Genome-wide Co-occurrence of promoter elements reveals a cis-regulatory cassette of rRNA transcription. *Genome Research* **12**, 1273-1731 (2002).
 16. Jensen, L.J. & Knudsen, S. Automatic discovery of regulatory elements patterns in promoters regions based on whole cell expression data and functional annotation. *Bioinformatics* **16**, 326-333 (2000).
 17. Xia, M., Ling, W.H., Ma, J., Kitts, D.D. & Zawistowski, J. Supplementation of diets with black rice pigment fraction attenuates atherosclerotic plaque formation in apolipoprotein E deficient mice. *Journal Nutr.* **133**, 744-751 (2003).
 18. Xia, X. *et al.* An anthocyanin-rich extract from black rice enhances atherosclerotic plaque stabilization in apolipoprotein e-deficient mice. *Journal Nutr.* **136**, 2220-2225 (2006).
 19. Wu, X. & Prior, R.L. Systematic identification and characterization of anthocyanins by HPLC-ESI-MS/MS in common foods in the United States: fruits and berries. *Journal of Agricultural and Food Chemistry* **53**, 2589-2599 (2005).
 20. Zhang, Z., Kou, X., Fugal, K. & McLaughlin, J. Comparison of HPLC methods for determination of anthocyanins and anthocyanidins in bilberry extracts. *Journal of Agricultural and Food Chemistry* **52**, 688-691 (2004).
 21. Kim, S.-K. *et al.* SELEX: Combinatorial library methodology and applications. *BioChip J.* **2**, 97-102 (2008).
 22. Cho, Y.-J., Kim, C.-J., Kim, N., Kim, C.-T. & Park, B. Some cases in applications of nanotechnology to food and agricultural systems. *BioChip J.* **2**, 183-185 (2008).
 23. Kobayashi, S., Ishimaru, M., Hiraoka, K. & Honda, C. Myb-related genes of the Kyoho grape regulate anthocyanin biosynthesis. *Planta* **215**, 924-933 (2002).