Identification of novel genes involved in anthocyanin and proanthocyanidin pigments accumulation in rice tissues using genome-wide association study (GWAS)

Haghi R.[1], Fazeli A.[1*], Ahmadikhah A.[2*] and Shariatiat V.[3]

[1] **Dept. of Plant Breeding, Ilam University, Ilam, I.R. of Iran.**

[2] **Faculty of Life Sciences and Biotechnology, Shahid Beheshty University, Tehran, I.R. of Iran.**

[3] **National Institute of Genetic Engineering and Biotechnology, Tehran, I.R. of Iran.**

Receive Date: 2019/10/16 Revise Date: 2019/10/28 Accept Date: 2019/12/08

Abstract

As natural mutants, some rice varieties have brown pericarp and/or purple stems and leaves. The brown and purple colors are caused by the accumulation of proanthocyanidin and anthocyanin pigments, respectively, belong to flavonoids. The anthocyanin and proanthocyanidin pigments are biosynthesized by the phenylpropanoid pathway and are modulated by several structural and regulatory gene families. In this study chromosome loci involved in the accumulation of anthocyanidin pigments in purple stems and leaves and the accumulation of proanthocyanidin pigments in brown pericarp were identified using 44K SNP array in 282 rice varieties based on association mapping method. Based on genes annotation database of rice and significant single nucleotide polymorphism (SNP) signals obtained from association mapping analysis, 30 chromosome positions including Chalcone isomerase, Chalcone synthase, Leucoanthocyanidin reductase, bHLH, MYB, WDR, and B-box coding genes were identified as candidate influencer genes in proanthocyanidin pigmentation in the rice grains pericarp. According to the mentioned method, 39 genes were introduced as likely involved genes in purple stem appearance including, bHLH, MYB, WDR, B-box, F3H, ChI, Glucosyltransferase and also 23 genes were detected for leaf pigmentation, including bHLH, Myb, WD40, B-box and Glucosyltransferase coding genes. Considering this fact that progenitors and domesticating processes for all the rice cultivars are not the same, and many structural and regulatory genes are introduced for controlling anthocyanin pigment accumulation, it seems that in different cultivars, one or some of these introduced genes can be the influencer gene for the accumulation of these pigment.

Key words: Anthocyanin, Association mapping, Proanthocyanidin, Rice, SNP

*Corresponding author: a.fazeli@ilam.ac.ir, a_ahmadikhah@sbu.ac.ir

Introduction

Rice as one of the most important crop plants has a critical role in human feeding in the world. Green aerial parts and white grains are common in cultivated varieties, but many wild rice ecotypes and some cultivated varieties have a purple stem (PS) and purple leaves (PL) and/or brown grain pericarp (3, 14). The purple and brown colors in rice plant organs are caused by the accumulation of anthocyanin and proanthocyanidin pigments, respectively (3, 14). Intensity and extension of the purple and brown pigments have a different pattern in various varieties and landraces, indicating genomic control of this trait is quantitative. In some varieties, purple pigments accumulate in the bottom part of the stem, but in some of them, they are spread over the leaf blade (fig.1-a). The intensity of brown color in grain pericarp of different cultivars is different, from light brown to heavy brown (fig.1-b).

Figure 1- a: Rice plants with purple and green stems b: Rice grains with white to brown pericarp

Accumulation of anthocyanin pigments in plants has several advantage for plant and its consumers, for example, they are involved in plant development, plant responses to biotic and abiotic stresses, and attraction of pollinator insects. Also, they play important role in antioxidant activity, anticancer, hypoglycemic, and anti-inflammatory effects for their consumers (2, 4, 5, 6). Anthocyanin and proanthocyanidin pigments belong to flavonoids, a major class of secondary metabolites that are synthesized from phenylalanine by phenylpropanoid pathway. In this pathway, several functional and regulatory genes are involved.

Different variants of anthocyanins are composed of an anthocyanidin backbone combined with sugar and acyl groups. Anthocyanidins are composed of two aromatic benzene rings conjunct by an oxygenated heterocycle (fig.2). Six variants of anthocyanidin, pelargonidin, cyanidin, delphinidin, Peonidin, petunidin and malvidin, are prevalent in plants. In addition to different anthocyanidin backbone, the difference in the structure, quantity and position of sugar and acyl conjugates led to diversity in anthocyanins. Biosynthesis of anthocyanins branched from general flavonoid pathway by the activity of Chalcone Synthase (CHS), converts 4-Coumaroyl-CoA and Malonyl-CoA to Nargenine Chalcone. Then Chalcone isomerase (CHI) converts Nargenine Chalcon to Nargenine. In the next step, three variants of dihydroflavonol can be synthesized by hydroxylase enzymes (flavanone 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3′H), flavonoid 3′,5′-hydroxylase(F3′5′H)). Next, dihydroflavonol 4-reductase (DFR) converts dihydroflavonol to a Leucoanthocyanidin variant, based on dihydroflavonol variant. Then, Anthocyanidin Synthase (ANS) synthesize different variants of a colored anthocyanidin from Leucoanthocyanidins. Finally, attachment of Glucose (sugar) by glucosyltransferase enzymes such as flavonoid 3-O-glucosyltransferase (UFGT) and then Acyl group by acyltransferase led to anthocyanin synthesis (2, 4, 5, 6) (fig.3).

Proanthocyanidin can be produced from Leucoanthocyanidins and Leucoanthocyanidins derivatives by a reductase enzyme (Anthocyanidin reductase (ANR), Leucoanthocyanidin reductase (LAR)) (2, 5, 6) (fig.3).

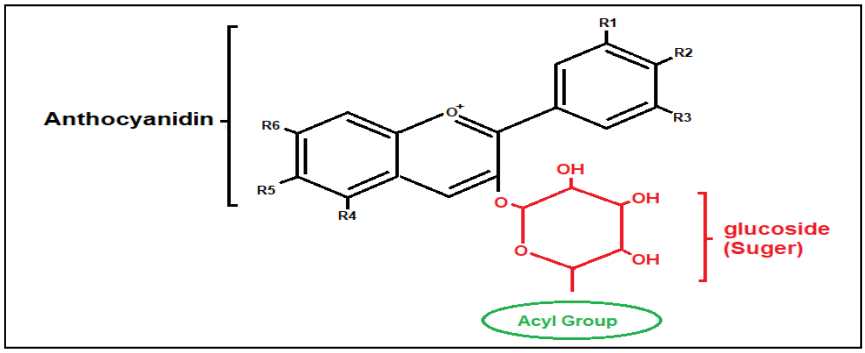

Figure 2- The general structure of an anthocyanin

Biosynthesis of CHS, CHI and F3H are common for most of the flavonoid compounds. In the biosynthesis of colored derivatives of flavonoids including anthocyanin and proanthocyanidin, regulatory genes modulate the expression of downstream structural genes consisting F3□H, F3□5□H, DFR, ANS, and UFGT, cause tissue-specific pattern and different intensity of the pigmentation (2,6). In previous studies, bHLH domain-containing proteins and Myb DNA binding domains containing proteins were introduced as transcription factors involved in the biosynthesis of anthocyanin and proanthocyanidin in rice (14, 15, 18). Recently, a regulatory complex called MBW including proteins containing Myb DNA banding domains, basic Helix-Loop-Helix (bHLH) domain-containing protein and a WD repeat, was introduced for the regulation of structural genes for the accumulation of anthocyanin pigments in rice and some other plants (20). The WD repeats regulate the specific activity and quantity of the complex (20). Also, a C-S-A system has been proposed by Sun et al, (2018) (17) for pigmentation of rice hull, while C is an MYB transcription factor, S is a bHLH domain-containing protein and A gene which encode a structural protein in anthocyanin biosynthesis pathway. So that, S interact with C for activating A gene.

Furukawa et al (2007) (5) introduced the RC gene, encoding a bHLH protein and a positive regulator for proanthocyanidin biosynthesis in brown pericarp and seed coat of rice. The RC gene activates CHS, DFR, LAR genes in brown grains. In white rice, 14 nucleotides have been deleted from the RC gene promotor. Recently, a B-box domain-containing protein was identified by Kim et al (2018) (8) as an inducer of anthocyanin biosynthesis in rice. This protein is the coactivator of HY5 protein in photomorphogenesis signaling. HY5 directly activate MYB transcription factors in the anthocyanin biosynthesis pathway (8). Xia et al (2016) isolated an artificial mutant line for brown hull and molecular analysis showed that a mutation occurred in a single recessive gene, OsFBX310^{bh6}. This gene mapped on the long arm of chromosome 9, coding an F-box domain-containing protein (19).

Assessment of the expression of transcripts in both black and white rice cultivars using the 135K *Oryza sativa* microarray identified 57 candidate genes involved in anthocyanin pigmentation in black rice. The genes are structural genes and transcription factors that distributed in 11 chromosomes (11).

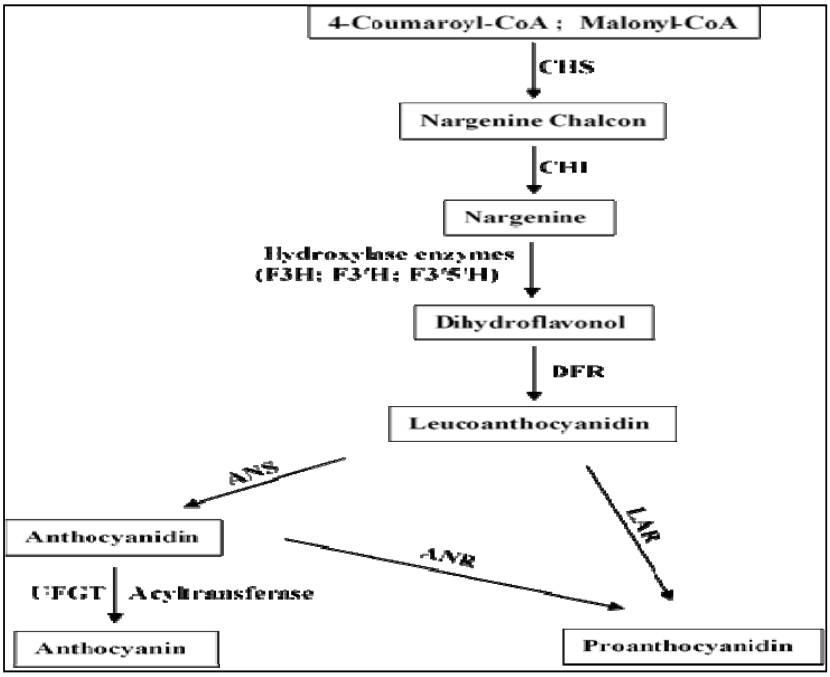

Figure 3- The Anthocyanin and Proanthocyanidin biosynthesis pathway, Chalcone Synthase (CHS), Chalcone isomerase (CHI), flavanone 3-hydroxylase (F3H), flavonoid 3′-hydroxylase (F3′H), flavonoid 3′,5′-hydroxylase (F3′5′H), dihydroflavonol 4-reductase (DFR), Anthocyanidin Synthase (ANS), flavonoid 3-O-glucosyltransferase (UFGT), anthocyanidin reductase (ANR), leucoanthocyanidin reductase (LAR).

Park et al, (2016) analyzed cDNA colons of two F3′H genes variants from white, red and black Korean rice varieties and revealed one or two amino acid substitution in the allelic variant of each gene (12). Heterologous expression of colons in yeast indicated different catalytic efficiency of an allelic variant of white, red and black varieties. Also expression level of F3′H genes was dominantly higher in black than red and white grains (12).

Linkage mapping and linkage disequilibrium mapping (association mapping) are two applied methods to find genomic regions involved in the control of quantitative traits. In linkage mapping strategy, recombination events between molecular markers and the genomic regions are measured in a sister-brotherhood population derived from a crossing between two or more parents, for example, F_2, backcross, doubled haploid and recombinant inbred lines populations (21). In this method, because of the low crossing-over rate, the resolution of the map is low and also the breeding of an appropriate population is very time-consuming and costly (7). In contrast, association mapping counts the correlation between marker alleles and the given trait in a natural prevalent population. In this method, a large and diverse collection of cultivars and lines is randomly collected and the relevant QTLs are mapped based on linkage disequilibrium (10). In comparison to the linkage mapping, association mapping is very fast, inexpensive and accurate. The natural population experienced a high rate of meiotic events and recombinations through the evolution, led to the high resolution of genetic maps and increases the accuracy of QTL mapping (10, 22). The aim of association mapping is the identification of a single polymorphism within a DNA locus linked to the given phenotype in a quantitative trait (7). One of

the most appropriate genetic markers for association analysis is SNP. SNP genotyping has the advantage of presenting whole-genome data without prohibitive cost, providing high-resolution mapping with a large number of SNPs entire the genomes (9). The population structure and genetic relationship between different varieties of the population can create false marker-trait associations. With the inclusion of population structure effect (Q coefficients) and kinship effects (K coefficients) to statistic model, the effect of population structure and kinship on the marker-trait association can be eliminated (13). Accordingly, association mapping with a large number of varieties and complete genome coverage with 40 K SNP array can identify all probable genes involved in the accumulation of anthocyanin pigments in rice. The objective of this research was to identify genomic regions involved in anthocyanin accumulation in rice and identification of candidate genes using association mapping with 40 K SNP array.

Material and Methods

The grains of a population including 282 rice varieties was kindly given by International Rice Research Institute (IRRI) and cultured at the normal condition in research farm of Shahid Beheshti University, Iran. The 44 K SNP array data for all the 282 varieties downloaded from GRAMINE website (www.gramene.org). For measurement of intensity and saturation of pericarp color, the hull of grains was removed and only the pericarp part of seeds coat remained. Then, the de-hulled grains of all the 282 varieties were scanned using a Samsung scanner and brown color intensity was recorded (between 8-58) using RGB color option in Photoshop CS5 software. Phenotypic data for stem and leaf color was recorded in field-grown plants for all 282 varieties in well grown plants before flowering stage and converted to 0 and 1 data for purple and green tissues, respectively. The 282 rice varieties were clustered using STRUCTURE software Version 2.3 based on SNP data for extracting Q values to be utilized in a mixed linear model (MLM) as covariates (13). Then association analysis between SNPs and phenotypic data was done using TASSEL 3.0 software (1). The MLM model is as follows:

$$P = \mu + M + Q + K + E$$

Where P is the phenotype, M and Q are the genotypic fixed effects and population structure, respectively. To split the effect of the genetic relationship between varieties, K matrix is added to the model as the kinship coefficient of the varieties, and E is the residual effects. After association analysis, the results were presented as Manhattan plots based on the negative $\log_{10}$ transformed observed p-values for each SNP-trait association. We used a threshold of 1e-04 to declare a significant association. The chromosome loci with high –Log (p-value) were selected as genomic regions carrying a candidate gene. Based on gene annotation database of rice (*https://rapdb.dna.affrc.go.jp*) and probable function of known genes in ±200kb closed to (16) selected SNPs, candidate genes for the purple stem and leaf and also brown pericarp were identified.

Results

Result of association analysis nominated different genomic regions as influencer loci in the accumulation of proanthocyanidin and anthocyanin pigments in rice organs, even identified loci for the accumulation of anthocyanin in stem and leaf were different. Based on Manhattan plots, the minimum considerable –Log (p-value) of SNP signals was determined about 4 for stem and pericarp color and 10 for leaf color.

Proanthocyanidin pigmentation

Most of the appointed SNPs with remarkable p-values for proanthocyanidin accumulation in brown pericarp located on chromosomes 1, 7, and 10. Other chromosomes had few considerable signals (fig. 4). Concerning previous studies in rice and other plant species, 11 structural genes and transcription factor families including bHLH, MYB, WDR, B-box DFR, F3H, PAL, CHS, CHI, ANR, LAR were determined as probably responsible genes for the accumulation of proanthocyanidin pigments in rice plants. Based on the position of SNPs with significant signals for pericarp pigmentation and the gene annotation database of rice, 30 chromosome positions including 1 Chalcone isomerase, 1 Chalcone synthase, 3 Leucoanthocyanidin reductase, 12 bHLH, 4 MYB, 8 WDR, and 1 B-box coding genes were identified as candidate influencer genes in proanthocyanidin pigmentation in the rice grains pericarp (Table 1). As expected, most of these identified loci are the regulatory genes, 25 regulatory genes in contrast to 5 structural genes. The SNP with the highest –Log(p-value) was positioned in around a bHLH transcription factor in the short arm of chromosome 7. This transcription factor was introduced by Furukawa et al (2007) (5) as a responsible gene for pigmentation of the brown pericarp of rice.

Anthocyanin pigmentation

Major striking SNP for anthocyanin accumulation in purple tissues was located on chromosomes 2, 4, and 6 for purple stem (fig. 5) and chromosomes 4, 6, 7 and 11 for purple leaves (fig. 6). Based on numerous studies, bHLH, MYB, WDR, B-box were considered as probably regulatory genes, similar to proanthocyanidin pigmentation. In addition to DFR, PAL, CHS, CHI, three other genes including ANS, UDP, and GST were also selected as likely structural genes involved in the accumulation of anthocyanins in stem and leaves. Four structural gene families including Chalcone isomerase, flavanone 3-hydroxylase, and UDP-glucosyl transferase and all the intended regulatory gene families including bHLH, MYB, WDR and B-box coding genes were detected close to significant SNP signals for anthocyanin accumulation both in stem and leaf (Tables 2 and 3). 39 genes were identified as likely involved genes in purple stem appearance including, 15 bHLH, 13 MYB, 5 WDR, 1 B-box, 2 F3H, 2 ChI, 1 Glucosyltransferase (Table 2). For stem pigmentation, four SNPs with a strong signal were detected in a restricted region of the short arm of chromosome 6. In this chromosomal region four annotated genes, including bHLH, MYB, ChI, and Glucosyltransferase, were identified.

–log10 (p-value) of SNP signals for leaf pigmentation was remarkably higher than stem and pericarp, especially on chromosome 6 with –log10 (p-value) about 22. Close to the significant SNP signals for leaf pigmentation, 5 bHLH, 3 Myb, 6 WD40, 1 B-box and 8 Glucosyltransferase coding genes were detected. Glucosyltransferase was the only structural genes identified as probably influencer gene in leaf pigmentation. 45 SNPs with significantly high –log10 (p-value) were detected in a small region of the short arm of chromosome 6 (close to the mentioned striking region for stem color) and interestingly all the intended regulatory genes were found in that area.

Discussion and Conclusion

The result of this study implies that genomic controlling of anthocyanin pigmentation is very complex and several structural and regulatory genes are involved in the control of this trait. In spite of almost the same biosynthesis pathway, there were

not any commonly identified chromosome regions responsible for the accumulation of anthocyanin and proanthocyanidin pigments and also, co-segregation of purple tissues and the brown pericarp is not occurring forever. Considering the fact that all the rice plants with purple leaves have purple stems and no common involved chromosomal region was identified for stem and leaf pigmentation, the involved genes in leaf pigmentation are likely regulatory gens for the tissue specific accumulation of the metabolite or are involved in the transportation of the metabolites from stem to leaves.

In previous studies two systems have been proposed, MBW regulatory complex (20) and C-S-A system (17), as genomic control of anthocyanin accumulation in rice. MBW regulatory complex has been introduced by Xu et al (2015) (20), wherein M is an MYB domain-containing protein, B is a bHLH and W is a WD repeat-containing protein. All these regulatory genes were identified in all three analyses for pericarp, stem, and leaf pigmentation. In C-S-A system proposed by Sun et al (2018) (17), C and S are MYB and bHLH domain-containing proteins, respectively, and interact together for activating A gene which encodes a structural gene in anthocyanin biosynthesis pathway.

Recently Kim et al (2018) (8) introduced a B-box protein as the final tuner of anthocyanin biosynthesis. In our study, we found B-box coding genes in all three analyses, on the long arm of chromosome 6, long arm of chromosome 2, and the short arm of chromosome 6 for pericarp, stem, and leaf, respectively. The number of identified B-box genes (one gene per trait) was much lower than other candidate regulatory genes.

Xia et al (2016) (19) have introduced an F-box domain-containing protein as responsible for rice brown hull pigmentation and the coding gene mapped on the long arm of chromosome 9. Our association analysis did not find any significant SNP signal on chromosome 9 for pigmentation of brown pericarp and purple stem and purple leaf. So it seems that the identified F-box protein just contributes to hull pigmentation.

According to C-S-A system and also result of Oh et al (2015) (11) and Park et al (2016) (12), in addition to regulatory genes, the mutation in structural genes in some varieties led to change in anthocyanin pigments accumulation in various rice organs. Also in this study, some structural genes (tables 1, 2, and 3) were identified for both proanthocyanidin and anthocyanin pigmentation. Based on this result and considering this fact that anthocyanin pigmentation is more common in wild progenitors of rice than cultivated varieties and most of common varieties and cultivars with green aerial parts and white grains domesticated from wild progenitors with purple aerial parts and black or brown grains, mutation in different structural and regulatory genes in the pathway could have occurred for each variety during domestication process. So in various varieties, mutation in different genes can be responsible for the change in organs color. Therefore, we cannot introduce universal gene(s) as responsible genes for anthocyanin pigmentation in all the varieties. In this study, an association mapping analysis by using a wide range of varieties (282 varieties) carrying various genome structure together with well genome coverage with 40 K SNP array and accurate phenotyping data have been performed. Result of this study has presented a reliable and comprehensive list of likely involved genes in accumulation of anthocyanin and proanthocyanidin pigments in stem-leaf and pericarp of the rice plant, respectively. One or more of these genes can be the responsible for pigmentation in a special variety. The candidate genes identified in our research

need to be confirmed by molecular approaches such as qRT-PCR etc. The results of this research add new information on the complexity of genetic control of anthocyanin accumulation in rice.

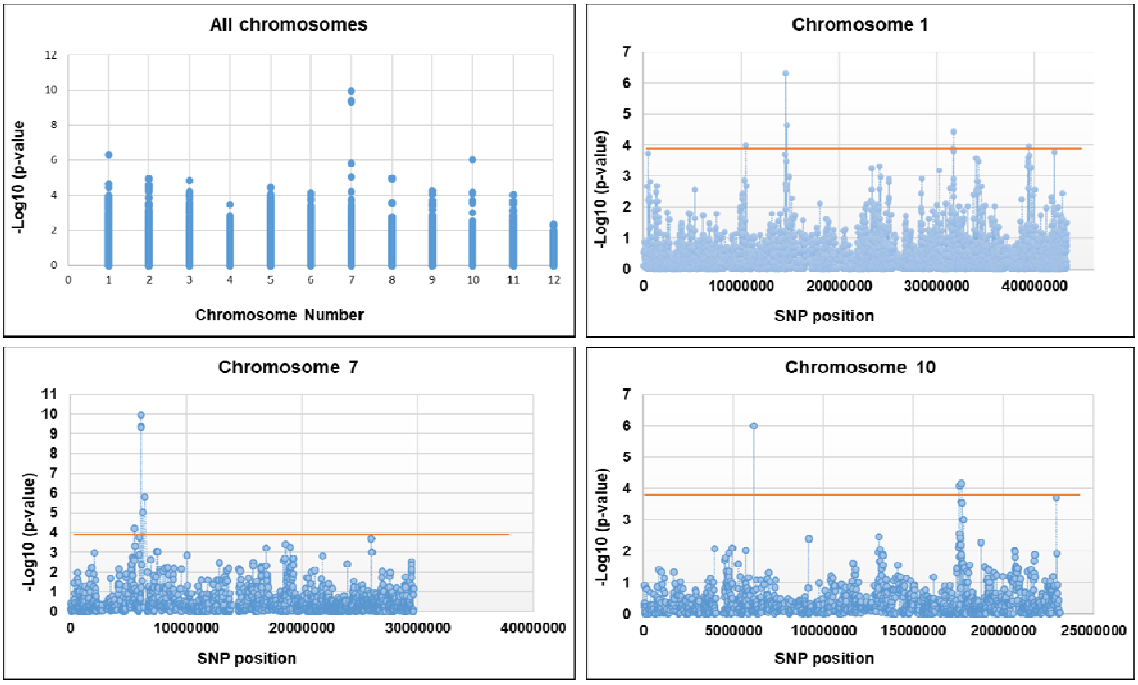

Figure 4- Association analysis result for genes involved in accumulation of proanthocyanidin pigments in brown pericarps of rice grains. Chromosomes with higher association signals are shown.

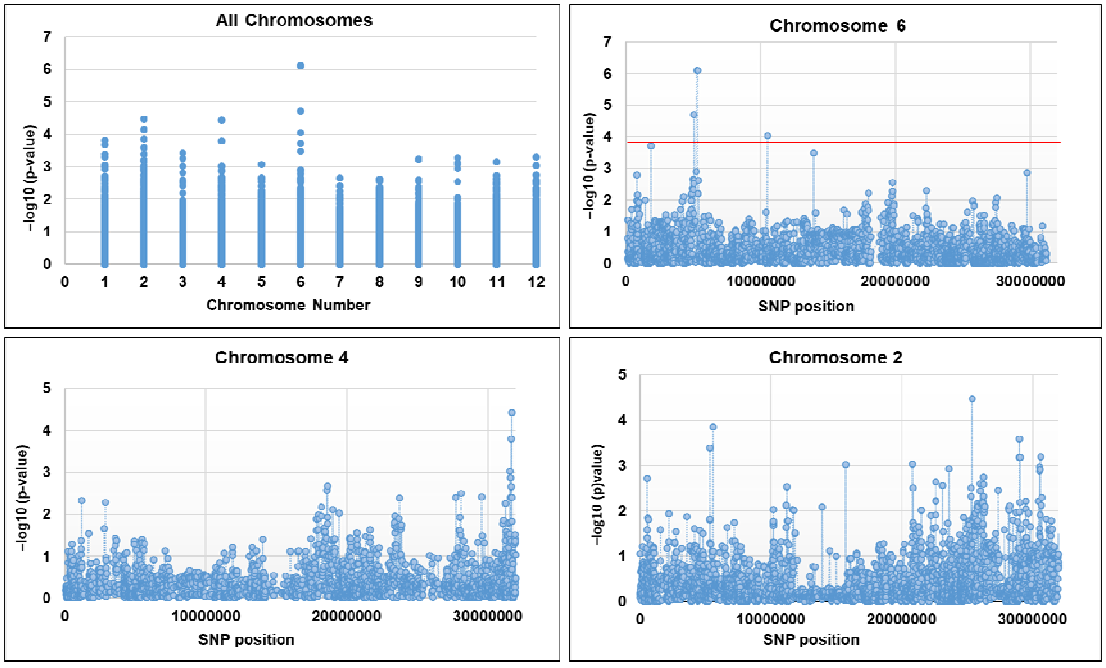

Figure 5-Association analysis results for genes involved in the accumulation of anthocyanin pigments in purple stems of rice plants. Chromosomes with higher association signals were shown.

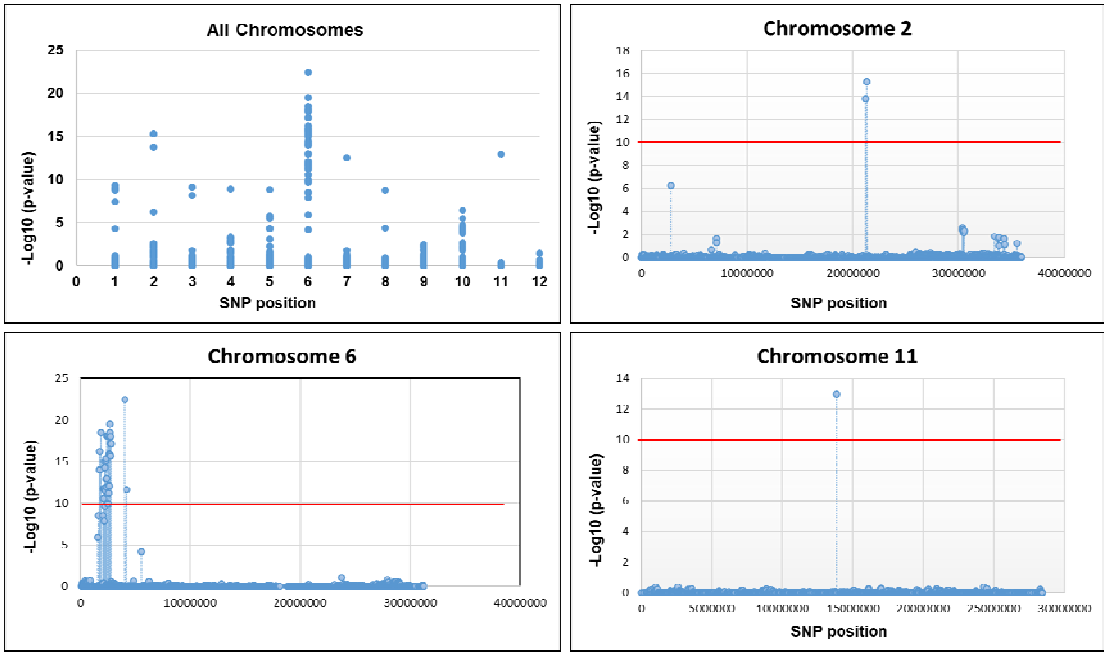

Figure 6: Association analysis result for genes involved in the accumulation of anthocyanin pigments in purple leaves of rice plants. Chromosomes with higher association signals were shown.

Table 1. Identified candidate genes for proanthocyanidin pigmentation in pericarps of brown rice.

Brown Pericarp					
Chromosome No.	**Chromosome region**	**SNP position**	**-log10 (p) of SNP**	**Candidate Genes**	**Candidate Gene Position**
Chr.1	1-1	10493521	4	Myb transcription factor	10217782-10219691
				bHLH transcription factor	10271157-10274304
				bHLH transcription factor	10664792-10666903
				MYB transcription factor	10934895-10935992
	1-2	39370281	4	bHLH transcription factor	39211307-39216467
				WD40 repeat domain-containing protein	40463040-40467753
Chr.2	2-1	6685191	4.41	bHLH transcription factor	6744958-6749884
				WD40 repeat domain-containing protein	5924727-5930096
	2-2	20350830	4.94	WD40 repeat domain-containing protein	20026633-20030067
		20351790	4.56	WD40 repeat domain-containing protein	20182213-20186136
				bHLH transcription factor	20562601-20565664
		20592789	4.23	bHLH transcription factor	21436299-21440664
Chr.3	3-1	34469629	4.21	MYB transcription factor	34350604-34352915
				Chalcone isomerase	34394508-34395638
Chr.5	5-1	22517298	4.02	bHLH transcription factor	22371227-22373519

Chr.6	6-1	4048543	4.12	WD40 repeat domain containing protein	3632670-3639453
				leucoanthocyanidin dioxygenase	3889190-3890392
				leucoanthocyanidin dioxygenase	3892063-3892359
				leucoanthocyanidin dioxygenase	3898331-3898775
				bHLH transcription factor	4192207-4194418
				bHLH transcription factor	4709744-4715552
	6-2	26717868	4.01	WD40 repeat domain-containing protein	26529305-26538020
				bHLH transcription factor	26756214-26757596
				WD40 repeat domain-containing protein	26779067-26781799
				B-box-containing protein	26843118-26843680
Chr.7	7-1	5524672	4.20	bHLH transcription factor	5081853-5084921
		6068267	9.31	bHLH transcription factor	6062889-6069304
		6090274	9.95		
		6256569	5.02	Chalcone synthase	6294479-6296305
		6416076	5.80	WD40 repeat domain-containing protein	6949934-6951925
Chr.10	10-1	6118562	6	MYB transcription factor	6024884-6028718

Table 2. Identified candidate genes for anthocyanin pigmentation in rice purple stems

Purple Stem					
Chromosome No.	**Chromosome region**	**SNP position**	**-log10 (p) of SNP**	**Candidate Genes**	**Candidate Gene Position**
Chr.1	1-1	18260021	3.8	MYB transcription factor	18755303-18756817
				WD40 repeat domain-containing protein	20717034-20722254
				bHLH transcription factor	21687437-21687759
				bHLH transcription factor	21687799-21689199
				bHLH transcription factor	22138044-22142395
				bHLH transcription factor	22236858-22238067
				bHLH transcription factor	22325472-22325974
Chr.2	2-1	5590334	3.84	MYB transcription factor	4058194-4060315
				bHLH transcription factor	4350872-4351973
				MYB transcription factor	4867738-4868831
				WD40 repeat domain-containing protein	4913036-4913371
				MYB transcription factor	4966920-4967015
				WD40 repeat domain-containing protein	5924727-5930096
	2-2	25401253	4.47	bHLH transcription factor	23643455-23651476
				MYB transcription factor	24580457-24582527
				MYB transcription factor	24878777-24879932
				MYB transcription factor	25759841-25762800
				MYB transcription factor	25778113-25780795
				B-box-containing protein	26027785-26029488

	2-3	33022850	4.14	flavanone 3-hydroxylase	32310799-32312184
				WD40 repeat domain-containing protein	32469587-32472658
				Chalcone isomerase	32957456-32961345
				WD40 repeat domain-containing protein	33629573-33633614
				MYB transcription factor	34829856-34831322
Chr.4	4-1	31715352	4.43	bHLH transcription factor	28369370-28371848
				flavanone 3-hydroxylase	29331563-29338260
				bHLH transcription factor	29880299-29881660
				MYB transcription factor	29995755-29997261
				MYB transcription factor	30028840-30030685
				bHLH transcription factor	30235422-30238854
				bHLH transcription factor	31416378-31420095
				bHLH transcription factor	32166724-32169575
				bHLH transcription factor	32657307-32658183
				bHLH transcription factor	33676818-33683540
Chr.6	6-1	5058807	4.70	bHLH transcription factor	4709744-4715552
		5295675	6.10	Chalcone isomerase	5227416-5231610
		5295818	6.10	MYB transcription factor	5315163-5316640
		5351422	6.10	UDP-glucosyl transferase	5908095-5909128
				MYB transcription factor	6239975-6241172

Table 3. Identified candidate genes for anthocyanin pigmentation in rice purple leaves

Purple Leaf					
Chromosome No.	**Chromosome region**	**SNP position**	**-log10 (p) of SNP**	**Candidate Genes**	**Candidate Gene Position**
Chr.2	2-1	21218858	13.79	bHLH transcription factor 5	20562601-20565664
				WD40 repeat domain-containing protein 6	20182213-20186136
		21309432	15.27	bHLH transcription factor	21436299-21440664
Chr.6	6-1	45 SNPs between: 1763642-4229437	10.0 to 22.46	WD40 repeat domain-containing protein	1504211-1509611
				WD40 repeat domain-containing protein	1648595-1654151
				UDP-glucosyl transferase8	2018053-2021997
				MYB transcription factor3	2187658-2187867
				WD40 repeat domain-containing protein	2320837-2329651
				B-box-containing protein	2695460-2699468
				UDP-glucosyl transferase	2902268-2907302
				bHLH transcription factor	3273390-3276888
				WD40 repeat domain-containing protein	3632670-3639453
				UDP-glucosyl transferase	3840732-3844710
				UDP-glucosyl transferase	3878788-3882766
				bHLH transcription factor	4192207-4194418
				UDP-glucosyl transferase	4648131-4649391

				UDP-glucosyl transferase	4653460-4655652
Chr.7	7-1	26602971	12.53	MYB transcription factor	26075493-26077145
				MYB transcription factor	27043501-27046097
				WD40 repeat domain-containing protein	27656036-27665871
Chr.11	11-1	13841323	12.97	UDP-glucosyl transferase	14509845-14516320
				bHLH transcription factor	14569673-14571090
				UDP-glucosyl transferase	14672877-14675385

References

1- Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ramdoss, Y., Buckler, E., 2007, TASSEL: software for association mapping of complex traits in diverse samples, bioinformatics, 2633–2635. https://doi.org/10.1093/bioinformatics/btm308

2- Chen, C., 2015, Pigments in Fruits and Vegetables, Genomics and Dietetics, Springer.

3- Chin, H.S., Wu, Y.P., Hour, A.L., Hong, C.Y., Lin, Y.R., 2016, Genetic and Evolutionary Analysis of Purple Leaf Sheath in Rice, Rice, 9: 8. https://doi.org/10.1186/s12284-016-0080-y

4- Chunthaburee, S., Sakuanrungsirikul, S., Wongwarat, T., Sanitchon, J., Pattanagul, W., Theerakulpisut, P., 2016, Changes in Anthocyanin Content and Expression of Anthocyanin Synthesis Genes in Seedlings of Black Glutinous Rice in Response to Salt Stress, Asian Journal of Plant Sciences, 15: 56-65. https://doi.org/10.3923/ajps.2016.56.65

5- Furukawa, T., Maekawa, M., Oki, T., Suda, I., Iida, S., Shimada, H., Takamure, I., Kadowaki, K.I., 2006, The Rc and Rd genes are involved in proanthocyanidin synthesis in rice pericarp, The Plant Journal, 49: 91-102. https://doi.org/10.1111/j.1365-313X.2006.02958.x

6- Gould, K., Davies, K., Winefield, C., 2009, Anthocyanins: Biosynthesis Functions and Applications, Plant Sciences, Springer.

7- Gupt, P., Rustgi, S., Kulwal, P., 2005, Linkage disequilibrium and association studies in higher plants: present status and future prospects, Plant Molecular Biology, 57: 461-485. https://doi.org/10.1093/bioinformatics/btm308

8- Kim, D., Park, S., Lee, J.Y., Ha, S.H., Lee, J.G, Lim, S.H., 2018, A Rice B-Box Protein, OsBBX14, Finely Regulates Anthocyanin Biosynthesis in Rice, International Journal of molecular Science, 19: 2190. https://doi.org/10.3390/ijms19082190

9- McCouch, S., Zhao, K., Wright, M., Tung, C.W., Ebana, K., Thomson, M., Reynolds, A., Wang, D., DeClerck, G., Ali, M.L., McClung, A., Eizenga, G., Bustamante, C., 2010, Development of genome-wide SNP assays for rice, Breed Science, 60: 524-535. https://doi.org/10.1270/jsbbs.60.524

10- Moose, S., Mumm, R., 2008, Molecular plant breeding as the foundation for 21 century crop improvement, Plant Physiology, 147: 969-977. https://doi.org/10.1104/pp.108.118232

11- Oh, J.H., Won, S.Y., Lee, J.H., Shin, D.H., Kim, C.K., 2015, Computational Identification of Anthocyanin-Related Genes UsingTranscriptome Data from Black Rice Plants, Cell & Developmental Biology, 133–141. https://doi.org/10.4137/EBO.S6077

12- Park, S., Choi, M.J., Lee, J.Y., Kim, J.K., Ha, S.H., Lim, S.H., 2016, Molecular and Biochemical Analysis of Two Rice Flavonoid 3'-Hydroxylase to Evaluate Their Roles in Flavonoid Biosynthesis in Rice Grain, International Journal of molecular Science, 17, E1549. https://doi.org/10.3390/ijms17091549

13- Pritchard, J.K., Stephens, M., Donnelly, P., 2000, Inference of population structure using multilocus genotype data, Genetics, 55: 945-959.

14- Saitoh, K., Onishi, K., Mikami, I., Thidar, K., Sano, Y., 2004, Allelic diversification at the C (OsC1) locus of wild and cultivated rice: nucleotide changes associated with phenotypes, Genetics, 168: 997–1007. https://doi.org/10.1534/genetics.103.018390

15- Sakamoto, W., Ohmori, T., Kageyama, K., Miyazaki, C., Saito, A., Murata, M., Noda, K., Maekawa, M., 2001, The Purple leaf (Pl) Locus of Rice: The Plw Allele has a Complex Organization and Includes Two Genes Encoding Basic Helix-Loop-Helix Proteins Involved in Anthocyanin Biosynthesis, Plant Cell Physiol. 42: 982–991. https://doi.org/10.1093/pcp/pce128

16- Song, J., Li, J., Sun, J., Hu, T., Wu, A., Liu, S., Wang, W., Ma, D., Hui, Z.M., 2018, Genome-wide association mapping for cold tolerance in a core collection of rice (*Oryza sativa* L.) landraces by using high-density single nucleotide polymorphism markers from

specific-locus amplified fragment sequencing. Frontiers in Plant Science. 9: 875-890. https://doi.org/10.3389/fpls.2018.00875

17- Sun, X., Zhang, Z., Chen, C., Wu, W., Ren, N., Jiang, C., Yu, J., Zhao, Y., Zheng, X., Yang, Q., Zhang, H., Li, J., Li, Z., 2018, The C–S–A gene system regulates hull pigmentation and reveals evolution of anthocyanin biosynthesis pathway in rice, Journal of Experimental Botany, 69; 1485–1498. https://doi.org/10.1093/jxb/ery001

18- Sweeney, M., Thomson, M., Pfeil, B., McCouch, S., 2006, Caught red-handed: Rc encodes a basic Helix-Loop-Helix protein conditioning red pericarp in rice, Plant Cell, 18: 283–294. https://doi.org/10.1105/tpc.105.038430

19- Xia, X., Xiao-bo, Z., Yong-feng, S., Hui-mei, W., Bao-hua, F., Xiao-hong, L., Qi-na, H., Li-xin, S., Dan, G., Yan, H., Jian-li, W., 2016, A Point Mutation in an F-Box Domain-Containing Protein Is Responsible for Brown Hull Phenotype in Rice, Rice Science, 23: 1−8. https://doi.org/10.1016/j.rsci.2016.01.001

20- Xu, W., Dubos, C., Lepiniec, L., 2015, Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes, Trends Plant Sci., 20: 176–185. https://doi.org/10.1016/j.tplants.2014.12.001

21- Xu, Y., Li, P., Yang, Z., Xu, C., 2017, Genetic mapping of quantitative trait loci in crops, The Crop Journal, 5: 175-184. https://doi.org/10.1016/j.cj.2016.06.003

22- Zhang, P., Zhong, K., Shahid, M.Q., Tong, H., 2016, Association analysis in rice: From application to utilization, Frontiers in Plant Science, 7: 1202-1218. https://doi.org/10.3389/fpls.2016.01202

شناسایی ژن‌های دخیل در تجمع رنگیزه‌های آنتوسیانین و پروآنتوسیانیدین در بافت‌های مختلف ۲۸۲ رقم برنج با استفاده از نقشه‌یابی ارتباطی در مقیاس ژنوم

رضا حقی[۱]، آرش فاضلی[۱*]، اسدالله احمدی خواه[۲*] و وحید شریعتی[۳]

[۱] **ایران، ایلام، دانشگاه ایلام، گروه زراعت و اصلاح نباتات**

[۲] **ایران، تهران، دانشگاه شهید بهشتی، دانشکده علوم زیستی و بیوتکنولوژی**

[۳] **ایران، تهران، پژوهشگاه ملی مهندسی ژنتیک و بیوتکنولوژی**

چکیده

بعنوان یک موتانت طبیعی، برخی از ارقام برنج دارای رنگ دانه قهوه‌ای و/یا رنگ ساقه و برگ بنفش می‌باشند. رنگ قهوه‌ای برونبر به دلیل تجمع رنگیزه پروآنتوسیانیدین و رنگ بنفش ساقه و برگ بواسطه‌ی تجمع آنتوسیانین ایجاد می‌شود. این رنگیزه‌ها که متعلق به دسته فلاونوئیدها بوده از طریق مسیر فنیل‌پروپانوئید ساخته می‌شوند و ژن‌های ساختاری و تنظیمی زیادی در بیوسنتز آنها نقش دارند. در این پژوهش، جایگاه‌های کروموزومی دخیل در ایجاد رنگ قهوه‌ای پریکارپ دانه و رنگ بنفش ساقه و برگ گیاه برنج با روش نقشه‌یابی ارتباطی با استفاده از ۴۴۰۰۰ نشانگر SNP در ۲۸۲ رقم شناسایی شد. براساس اطلاعات پایگاه داده ژن‌های آنوتیت شده برنج وهمچنین نتایج نقشه‌یابی ارتباطی، ۳۰ جایگاه کروموزومی شامل ژن‌های کدکننده‌ی آنزیم‌های چالکون سنتتاز، چالکون ایزومراز، لوکوآنتوسیانیدین ردوکتاز، bHLH، MYB، WDR، و B-box بعنوان ژن‌های احتمالی دخیل در ایجاد رنگ قهوه‌ای برونبر دانه برنج شناسایی شدند. با استفاده از همین روش، ۳۹ جایگاه کروموزومی شامل: ژن‌های کدکننده‌ی آنزیم‌های bHLH ، MYB، WDR، B-box، F3H، چالکون ایزومراز و گلیکوزیل ترانسفراز برای تجمع رنگیزه ارغوانی در ساقه و ۲۳ جایگاه کروموزومی شامل: ژن‌های کدکننده‌ی آنزیمهایbHLH ، Myb، WD40، B-boxو گلیکوزیل ترانسفراز برای تجمع رنگیزه ارغوانی در برگ‌های برنج معرفی شدند. با توجه به این واقعیت که اجداد وحشی و فرآیند اهلی سازی برای همه ارقام امروزی برنج مشابه نبوده است و همچنین با توجه به اینکه ژن‌های ساختاری و تنظیمی بسیاری در مسیر بیوسنتز این رنگیزه‌ها نقش دارند، به نظر میرسد که در هر کدام از ارقام موجود یک یا تعدادی از این ژن‌های معرفی شده می‌توانند ژن کلیدی در تجمع این رنگیزه‌ها در آن رقم باشند.

واژه‌های کلیدی: آنتوسیانین، پروآنتوسیانیدین، نقشه‌یابی ارتباطی، برنج، SNP

* نویسندگان مسئول، تلفن: ۰۸۴۵۹۲۴۱۷۰۰ ، پست الکترونیکی: a.fazeli@ilam.ac.ir, a_ahmadikhah@sbu.ac.ir