INTERNATIONAL JOURNAL OF AGRICULTURE & BIOLOGY
ISSN Print: 1560–8530; ISSN Online: 1814–9596
17–0599/2018/20–1–149–156
DOI: 10.17957/IJAB/15.0468
http://www.fspublishers.org

Full Length Article

DNA Methylation for Transcriptome Changes during Starch Biosynthesis in Rice

Jae-Hyeon Oh[1], Hwang-Weon Jeong[1], Yuna Kang[1], Gang-Seob Lee[2], Yong-Hwan Kim[3], Yong-Kab Kim[4] and Chang-Kug Kim[1*]

[1]*Genomics Division, National Institute of Agricultural Sciences, RDA, Jeonju 54874, Korea*
[2]*Biosafety Division, National Institute of Agricultural Sciences, RDA, Jeonju 54874, Korea*
[3]*Department of Crop Science and Biotechnology, Dankook University, Cheonan 31116, Korea*
[4]*School of Electrical Information Communication Engineering, Wonkwang University, Iksan 54538, Korea*
[*]For correspondence: chang@korea.kr

Abstract

DNA methylation plays an important role in gene regulation for the biosynthesis of amylose and other starches in rice. In starch biosynthesis, methylation variation has been largely influenced by environmental factors. Here, we describe the DNA methylation patterns in four rice accessions with variable amylose content, and integrate these data with DNA resequencing and transcriptomics analysis to determine how methylation may relate to single nucleotide polymorphism (SNP) variation and gene expression. We discovered numerous differentially-methylated regions (DMRs) through our genome-wide analysis of the four rice accessions at two distinct growth stages. In total, 4,525,209 SNPs were identified, and 2,501 (17.1%) common genes were methylated in all accessions at both time-points. Three candidate genes were selected based on comparison of the SNP variation, gene expression, methylation patterns, and RT-PCR validation. Our results indicate that DNA methylation is associated with amylose and amylopectin biosynthesis, and expression of our candidate genes may be influenced by increased/decreased methylation. Although methylation is greatly influenced by environmental factors, our results suggest that gene methylation may play an important role starch biosynthesis in rice and therefore should considered in breeding strategies for the development of improved varieties.

Keywords: Amylose; DNA methylation; SNP; Starch biosynthesis

Introduction

DNA methylation is an epigenetic modification that is critical for gene regulation in response to environmental conditions (Garg *et al*., 2015). Methylation systems are mosaics of conserved features, suggesting methylation is an ancient property of eukaryotic genomes (Zemach *et al*., 2010). Although DNA methylation likely has a conserved role in gene silencing, the levels and patterns of DNA methylation appear to vary drastically among different organisms (Feng *et al*., 2010). A large number of methods are available to measure DNA methylation (Cokus *et al*., 2008; Harrison and Parle McDermott, 2011).

Critically, advances in high-throughput sequencing have enabled the measurement of whole-genome DNA methylation (Adusumalli *et al*., 2015), and methyl-CpG-binding domain (MBD) patterns (Decock *et al*., 2016). DNA methylation affects the transcriptome, and DNA sequence divergence is thought to be the major determinant of methylation differences at the whole genome level (Chodavarapu *et al*., 2012). In the promoter regions, methylation represses gene expression, whereas methylation in gene bodies is positively correlated with enhanced expression (Wang *et al*., 2014). Additionally, methylation in transcriptional termination regions (TTRs) can significantly repress gene expression, and the effect is even stronger than that of promoter methylation (Li *et al*., 2012).

Starch is the most abundant storage carbohydrate in rice (Zeeman *et al*., 2010; Farooq *et al*., 2011). It is an insoluble glucan composed of two polymers of glucose, amylopectin, and amylose (Wani *et al*., 2012). The amylose biosynthesis is synthesized by granule bound starch synthase (Hanashiro *et al*., 2008) and amylopectin biosynthesis is highly active in the developing endosperm through the coordinated activity of many factors (Ohdan *et al*., 2005; Fujita *et al*., 2006). Whole-genome sequencing has revealed many isozymes of each class, indicating that starch biosynthesis is a highly-regulated process, involving a variety of metabolic pathways (Bahaji *et al*., 2014), transcription factor genes (Fu and Xue, 2010), and starch-synthesizing enzymes (Pandey *et al*., 2012).

In rice, relationship between starch biosynthesis and DNA methylation, genes regulated through methylation increase starch synthesis during endosperm development

To cite this paper: Oh, J-H., H-W. Jeong, Y. Kang, G-S. Lee, Y-H. Kim, Y-K. Kim and C-K. Kim, 2018. DNA methylation for transcriptome changes during starch biosynthesis in rice. *Int. J. Agric. Biol*. 20: 149–156

(Xing *et al*., 2015). Especially, OsSRT1-mediated histone deacetylation has been reported to be involved in starch accumulation (Zhang *et al*., 2016). Recently, a genome-wide DNA methylation analysis reveals that numerous differentially-methylated regions (DMRs) have been associated with gene expression in response to environmental factors and stress conditions (Garg *et al*., 2015). Further, in this organism, most DNA methyltransferases, histone methyltransferases, and DNA demethylases are differentially regulated (Lu *et al*., 2016), and both DNA methylation patterns and the expression of these enzymes are modulated in response to organic toxins, such as atrazine (Zhang *et al*., 2014). Although the OsSRT1 gene has been reported, a number of starch-related genes expressed through methylation are not well known. It is assumed that methylation change of starch biosynthesis is largely due to its environmental factors (Thitisaksakul *et al*., 2012). Here, we aimed to determine the role of amylose-specific DNA methylation in the regulation of starch and sucrose metabolism.

Materials and Methods

Rice Accessions

We obtained samples of four rice (*Oryza sativa* ssp. *japonica*) accessions from the National Agrobiodiversity Center (http://genebank.rda.go.kr/): Suweon 384 (IT191829), Suweon 460 (IT215182), Suweon 355 (IT191781), and Milyang 168 (IT213236). These contain amylose ratios of 0.0%, 9.1%, 16.2%, and 28.0%, respectively. Suweon 460 is a mutant of Suweon 355, and Milyang 168 was generated from a three-way cross-breeding strategy that included Suweon 355. We hypothesized that these accessions would be useful for comparing expression and methylation of genes involved in starch biosynthesis, as they exhibit variable amylose content, but otherwise have similar genetic backgrounds.

Analysis Details

Genome resequencing: Paired-end sequencing libraries were constructed using the TruSeq TM DNA Sample Preparation Kit v2 (Illumina Inc., San Diego, CA, USA), according to manufacturer protocols. Whole-genome resequencing was performed using the Illumina HiSeq 2000 sequencing platform (Caporaso *et al*., 2012). Raw reads were subjected to forced trimming, ambiguity filtering, quality check, and mapping (Patel and Jain, 2012; Kawahara *et al*., 2013).

SNP detection: The SAMtools/BCFtools was used for calling SNP and analysis was processed on the rice reference sequence (Os-Nipponbare-Reference-IRGSP-1.0) using the SnpEff program (http://snpeff.sourceforge.net/). Visual inspection of SNP distribution was performed with the Genomic Ranges (http://bioconductor.org/). We also determined whether non-synonymous SNPs mediate substitutions of amino acid with altered characteristics (e.g., hydrophobic to basic or amino acid to stop codon) manually.

Gene expression and positions: We performed 12 RNA seq experiments, comparing gene expression in four rice accessions at three stages (ear emergence +7, +14, and +21 days). Library construction and sequencing was performed using the Illumina HiSeq 2000 platform (Hayward, CA, USA). The fragments per kilobase of transcript per million mapped reads (FPKM) scores and gene expression levels were determined using Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/). The heat map clustering was drawn using the gplots (https://www.rdocumentation.org/packages/gplots/). We identified the most likely chromosomal positions of each gene using the FSTVAL program (GGBio Inc., Yongin, Korea) and located the best mapping position with BLAST/NCBI, employing an e-value cut-off $\leq 1.0 \times 10^{-5}$.

Methylation and comparison: A single library was constructed for MBD sequencing analysison the four accessions at two stages of growth (ear emergence +7 days and +21 days). Each sample was sequenced to yield >90% with at least 10X coverage and Q30-score> 90% (Decock *et al*., 2016). Mapping was performed using BWA software, with options (i.e., -m 1, -v 3) (Li and Durbin, 2009). DMR peaks were detected using the MACS (http://liulab.dfci.harvard.edu/MACS/) with options (i.e., -nomodel, -shiftsize = 180 "for 350 bp-fragment library, 50 bp reads, 120 bp adaptor"). Because DNA hypermethylation related to increase in the epigenetic methylation and occurred specifically for the majority of the genes, we annotated DMR-associated transcripts based on the hypermethylation (Garg *et al*., 2015; Lu *et al*., 2016). These transcripts were assigned using two methods such as promoter and gene-body models. The promoter-based method identifies DMRs overlapping gene promoter (between 2 kb upstream and 500 bp downstream from transcription start site [TSS]) within-house scripts. The gene body-based method identifies DMRs overlapping gene-body regions (between the TSS and transcription end site [TES]) with in-house scripts (Fig. 1). The screened transcripts were finally identified compared to NCBI/NR database. Assigned genes in the promoter/gene-body methods were screened using two cut-off values, such as $P<10^{-5}$ and false discovery rate (FDR) <5%.The SNP variation, gene expression, and methylation and results were represented as a circos diagram (http://circos.ca/).

RT-PCR: Total RNA was extracted from ground leaf tissues using the RN easy Plant Mini Kit (QIAGEN, Valencia, CA, USA), according to manufacturer instructions. RNA was quantified using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific Inc., Wilmington, NC, USA), and reverse transcription polymerase chain reaction (RT-PCR) was performed in triplicate using the SYBR Premix Ex Taq reagent (Takara Bio, Inc., Otsu, Shiga and Japan). Reaction parameters were as follows: 1 cycle at 94°C for 5 min; 25 cycles at 94°C for

30 sec, 58°C for 30 sec, and 72°C for 1 min; and a 10-min final extension at 72°C. All results were normalized to expression of *OsActin*.

Accessions

Resequencing data have been deposited in EMBL-EBI (http://www.ebi.ac.uk) under the accession numbers: ERP014478, ERP014479, ERP014495, and ERP014496. The RNAseq transcriptome dataset has been deposited in EMBL-EBI with accession numbers, ERP014497-ERP014504, and ERP017036-ERP017039. The MBD-seq methylation dataset has been deposited in EMBL-EBI with accession numbers, ERP014506-ERP014513.

Results

Genetic Variation among the Genomes

Our sequencing yielded an average of 1.13×10^8 reads with an average 45.8X depth. A Q30 score was obtained for 89.9-92.1% of bases, and average 43.2% GC ratio. For all samples, 93.1% of pre-processed reads were uniquely mapped, and average alignment depth was 40.5X, suggesting our resequencing data is sufficient for subsequent analysis (Table 1). In total, we identified 4,525,209 SNPs in our four rice accessions. Of these, 83% were homozygous, and 17% were heterozygous SNPs. In coding sequence (CDS) regions, we identified average 14,370 SNPs, including 6,536 synonymous and 7,834 non-synonymous SNPs. The promoter region sequences are most highly enriched for SNPs than other regions (Table 2).

Identification of Differentially-Methylated Genes

We next performed whole-genome methylation mapping. Total methylated sequence was 7.2 GB with 95.4% Q30 score. When methylated sequences were aligned, the ratio of aligned reads was 55.1% and that of un-aligned reads was 5.4%. The remaining 39.5% was comprised of multi-aligned reads which occurs due to similar sequences on the reference genome (Table 3). We annotated DMR-associated transcripts based on the hypermethylation, and these genes were assigned based on the promoter and gene-body regions according to two cut-off values (Table 4). We screened transcripts based on stages, accessions, and cut-off value in the two models, and transcripts that appeared more than once among them were finally selected. The 1,395 transcripts of promoter and 1,106 transcripts of gene-body were anchored. Overall, we identified total 2,501 transcripts that were differentially methylated except for the gene that appeared only oncein two models.

Gene Expression and Genomic Localization of Starch-Related Biosynthesis

To assess the gene expression, we performed 12 non-replicated RNA seq experiments of each rice accession at three stages (ear emergence +7, +14 and +21 days).

Table 1: Resequencing alignment statistics after pre-processing of raw resequencing reads

Sample	Raw reads				Mapped reads	
	Number	Depth	GC (%)	Q30 (%)	Ratio (%)	Depth
Suweon384	93,267,836	37.7	43.18	92.11	92.64	33.5
Suweon460	143,676,668	58.1	43.05	90.01	92.74	50.8
Suweon355	126,279,094	51.1	43.23	89.86	93.5	45.5
Milyang168	89,169,212	36.1	43.38	90.41	93.51	32.2
Average	113,098,203	45.8	43.21	90.60	93.09	40.5

Table 2: Distribution of SNPs within various genetic regions from four rice accessions

Sample	Total	Intron	CDS		Promoter	UTR		Inter Region
			SY[a]	NS[b]		5'-UTR	3'-UTR	
Suweon384	1,094,425	36,058	6,934	8,447	775,828	5,667	8,602	252,889
Suweon460	1,401,352	43,774	7,551	8,949	971,903	6,815	10,567	351,793
Suweon355	802,287	27,484	5,104	5,924	573,506	4,351	6,930	178,988
Milyang168	1,227,145	39,066	6,554	8,017	847,300	5,610	8,897	311,701
Average	1,131,302	36,596	6,536	7,834	792,134	5,611	8,749	273,843

[a]SY: synonymous SNPs in the coding sequence (CDS), [b]NS: non-synonymous SNPs

Table 3: Statistics of methyl binding domain (MBD) sequencing on the reference genome

Sample	Yield (Mb)	Input reads	Q-Score[a]	M-Score[b]	Mapped reads (%)		
					Align	Un-align	Multi-align
Suweon384	1,850	36,261,403	95.5	37	53.7	2.9	43.4
Suweon460	1,570	30,787,484	96.3	37.5	58.0	5.1	36.9
Suweon355	1,602	31,412,746	95.6	37.2	53.6	3.3	43.1
Milyang168	2,193	42,994,881	94.1	36.8	55.0	10.1	34.9
Average	1,804	35,364,129	95.4	37.1	55.1	5.4	39.5

[a]Q-score: % Bases Q ≥30 (pf, pass filter), [b]M-score: mean quality score (pf)

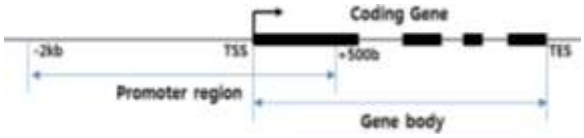

Fig. 1: Diagram of promoter and gene body regions for annotation of differentially-methylated region (DMR)-associated genes. Promoter contains 2 kb of upstream sequence and 500 bp of sequence downstream from the transcription start site (TSS) in protein-coding genes. TES, transcription end site

We detected the 42,528 unique transcripts using identified 31,648,792 reads, with an average 50.4% GC ratio and 90.2% Q30 score. We then utilized amulti-stepped analysis to identify genes. First, we identified 5,267 genes that exhibited a >2.0-fold up- or down-regulation. Second, we screened 18 candidate genes for regulation of starch metabolism using the KEGG database (Fig. 2 and Table 5). Finally, we identified 28 transcripts from the 18 candidate genes and generated clusters using heat map analysis. A composite tree illustrates that these transcripts into three subgroups (I, II and III). Group I contains eight amylose-

Table 4: The numbers of differentially-methylated regions (DMRs) identified by the promoter and gene-body annotation models, with two cut-off options

Compared samples			Hyper-methylation			NCBI/NR database[c]		Cut-off
Control	S[a]	Accessions[b]	Control	Accessions	Both	Promoter	Gene-body	
S384	7	S460	8,452	4,920	13,372	1,788	1,359	P<10-5
		S355	2,110	2,160	4,270	664	424	
		M168	2,922	3,017	5,939	927	606	
		S460	5,937	2,801	8,738	1,858	2,610	FDR
		S355	1,008	1,035	2,043	483	343	<5%
		M168	1,136	1,033	2,169	2,472	3,428	
	21	S460	17,774	3,518	21,292	967	645	P<10-5
		S355	1,810	1,654	3,464	334	211	
		M168	29,932	3,724	33,656	323	174	
		S460	13,912	2,727	16,639	1,411	2,104	FDR
		S355	715	599	1,314	156	104	<5%
		M168	17,242	3,315	20,557	2,235	3,148	

[a]S: Time stages (ear emergence +7 days, +21 days)
[b]Accessions: S (Suweon), M (Milyang)
[c]Numbers of DMRs in the accessions

Table 5: Mapping matrix for 18 candidate genes that were identified from plant secondary-metabolite pathways for starch and sucrose metabolism

Gene_ID	Strand	Transcipt[a]	TranscriptID[b]	Protein	Protein ID[c]
Os01g0633100	+	Os01t0633100-01	AK071497	*AGPL2*	Q7G065
Os01g0851700	-	Os01t0851700-01	AK103367	*PYG*	B8ACF5
Os02g0528200	-	Os02t0528200-01	AK108535	*SBE3*	Q40663
Os02g0744700	-	02t744700-01/02	AK066446	*SSIIB*	D0TZP8
Os03g0735000	-	03t735000-01/02	AK069296	*AGPL1*	B8AR31
Os03g0758100	-	Os03t0758100-01	AK063766	*PYG*	Q9ATK9
Os05g0580000	-	Os05t0580000-01	AK100910	*AGPL3*	A2Y7W1
Os06g0133000	+	Os06t0133000-01	AK070431	*GBSSI*	D3U2H9
Os06g0160700	-	Os06t0160700-01	AK109458	*SSI*	D3U2H8
Os06g0229800	+	Os06t0229800-01	AK101978	*SSIIA*	Q0DDE3
Os06g0367100	+	06t0367100-01/02	AK066930	*SBE*	B0FLE9
Os06g0726400	-	6400-01/ 02/ 03/ 04	AK119436	*SBE1*	D0TZI4
Os07g0243200	-	Os07t0243200-01	AK121036	*AGPL4*	B9FWD3
Os07g0412100	-	Os07t0412100-01	AK102058	*GBSSII*	D0TZD9
Os08g0345800	-	5800-01/ 02/ 03/ 04	AK103906	*AGPS2*	P15280-2
Os08g0520900	-	Os08t0520900-00	AB015615	*ISA1*	D0TZF9
Os09g0298200	-	09t0298200-01/02	AK073146	*AGPS1*	D3U2H7
Os10g0437600	+	Os10t0437600-00	AF383878	*GBSSII*	D0TZN5

[a]Transcript name is partially omitted
[b]ID is referenced by DDBJ (http://www.ddbj.nig.ac.jp/)
[c]ID is referenced by UniProt (http://www.uniprot.org/) and RefSeq (http://www.ncbi.nlm.nih.gov/)

related transcripts, Group II contains 12 amylopectin-related transcripts, and Group III contains eight ADP-glucose-related transcripts (Fig. 3a). To reveal the their chromosomal positions, we mapped the positions of these 18 genes on the 12 rice chromosomes using the FSTVAL program and grouped them into four function alcategories: amylose, amylopectin, ADP-glucose, and starch phosphorylase (Fig. 3b). However, we did not find a significant relationship between the three groups of the heat map and the 18 genes on the chromosomal positions.

Comparison of SNP Structure, Gene Expression and Methylation

We investigate the relationship among SNPs, gene expression, and methylation using 14,370 SNP sand 28 transcripts of starch-related biosynthesis. In the relationship between SNPs and gene expression, we analyzed 6,536 synonymous and 7,834 non-synonymous SNPs. The 14,370 SNP sand FPKM values of identified 28 transcripts were compared. However, we did not find significant sequence variation between the transcriptome and SNPs, except Os06T0133000-01 transcript. In the relationship between SNPs and methylation reads, we compared the two factors in the entire locus region of transcript. However, we did not find significant correlation between the methylation and SNP variation.

In the methylation difference between amylose-variable rice, we compared the four accessions in both the promoter (-2 kb upstream) and the TTR (+2 kb downstream). Transcripts were selected based on Suweon384 (amylose ratio 0.0%), which showed high methylation reads in the high-amylose accessions.

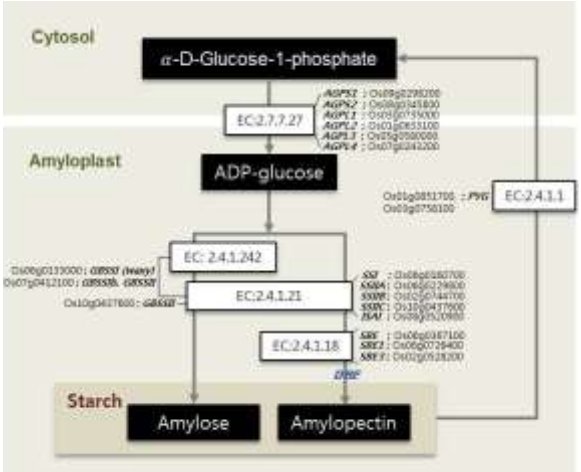

Fig. 2: Simplified model of the starch biosynthetic pathway in a rice plants. Legend is as follows: ADP-Glucose, ADP-Glucose pyrophosphorylase; AGPS, Glucose-1-phosphate adenylyltransferase large subunit; AGPL, Glucose-1-phosphate adenylyltransferase; GBSSI, Granule bound starch synthase; SS, Starch synthase; ISA,I soamylase; SBE, Starch-branching enzyme

We detected four transcripts (Os06T0133000-01, Os03T0735000-01, Os01T0633100-01, andOs07T0412100-01) with increase methylation in the promoter (Fig. 4a) and two transcripts (Os06T0229800-01 and Os06T0367100-01) with increased methylation in both promoter and TTR region (Fig. 4b). Therefore, this suggests that differences of amylose contents mainly occur due to more epigenetic modifications (i.e., methylation) than SNP variation.

RT-PCR Validation of Selected Genes

We identified the six DMR-associated genes (i.e., *Os06g0133000*, *Os03g0735000*, *Os06g0229800*, *Os07g0412100*, *Os01g0633100* and *Os06g0367100*) from comparison of SNP distribution, gene expression, and differential methylation. To validate the results, we measured expression of six candidate genes by RT-PCR. Three of the six genes were up regulated (i.e., *Os06g0133000*, *Os03g0735000*, and *Os06g0229800)* during all of the developmental stages tested (Fig. 5). It is likely that the three up-regulated genes either play regulatory roles in amylose production or are related to starch biosynthesis.

Discussion

In this study, we identified amylose- and starch metabolism-related genes based on whole-genome DNA methylation patterns and transcriptome profiling in rice. We measured distinct methylation patterns in four rice accessions, which have similar genetic backgrounds, but exhibit different amylose content. This is consistent with previous studies showing whole-genome DNA methylation status is related to starch biosynthesis (Ohdan *et al*., 2005; Li *et al*., 2012; Garg *et al*., 2015). We further annotated 2,501 genes that were differentially methylated in all samples using our promoter and gene-body annotation models. However, expression of only three was verified by RT-PCR. Therefore, we suggest environmental factors other than methylation are more significant for amylose biosynthesis, although multiple methylated regions were observed in coding regions and/or promoters.

In both rice and *Arabidopsis*, DNA methylation in promoters was found to repress gene expression, whereas methylation in gene bodies was associated with enhanced expression and active transcription (Takuno and Gaut, 2013). We identified 2,501 (17.1%) methylated genes in the rice genome, which is similar to the 16% predicted for this organism. The methylation pattern within and around protein-coding genes is also consistent with that observed in previous studies (Chen and Zhou, 2013; Garg *et al*., 2015).

Previous studies have demonstrated that DNA methylation in rice can be influenced by environmental stimuli and starch metabolism (Li *et al*., 2012; Garg *et al*., 2015). Other reports suggest DNA methylation is directly responsible for the increased starch biosynthesis observed during endosperm development (Xing *et al*., 2015). Starch is comprised of two carbohydrate polymers, a linear amylose and a branched amylopectin. Their synthesis is inhibited by environmental stress such as water deficit (Farooq *et al*., 2009). Our data reveal different patterns of DNA methylation in amylose and amylopectin genes. For example, glucose-1-phosphate adenylyltransferase (*AGPL1*, *Os03g073500*), which generates ADP-glucose in starch metabolism, was expressed at lower levels in Suweon384 (amylose ratio 0.0%) than in high-amylose accessions. Similarly, in high-amylose varieties, GBSS genes that function in amylose biosynthesis, *GBSSI (waxy)* and *GBSSIb*, showed higher expression, whereas levels of *SSIIA* and *SBE* were lower. This is consistent with the observation that amylopectin content decreases with increasing amylose.

Based on the starch and sucrose metabolic pathway (KEGG pathway: map00500), amylose production can be enhanced by either increasing the precursor or decreasing the reductase. We investigated the effect of methylation on expression of 18 candidate genes predicted from the KEGG database. These genes encode 28 representative transcripts, and each was confirmed by PubMed Medline (https://www.ncbi.nlm.nih.gov/pubmed/). We found six transcripts (Os06T0133000-01, Os03T0735000-01, Os06T0229800-01, Os07T0412100-01, Os01T0633100-01, and Os06T0367100-01) with increased or decreased methylation, when comparing the SNP variation and methylation patterns between promoter and TTR regions (Fig. 4). To gain insight into the relationship between SNPs distribution, gene expression, and methylation, we compared expression of these six transcripts in the four rice accessions (Fig. 6).

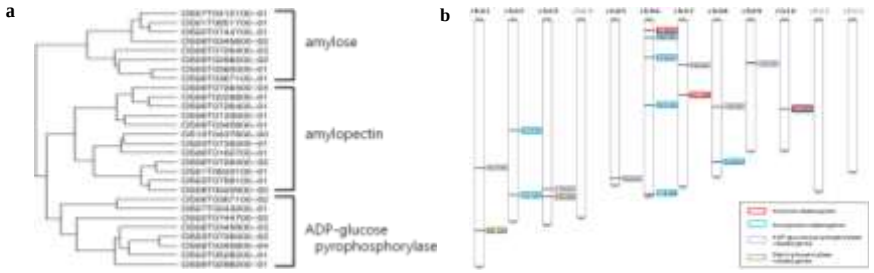

Fig. 3: Dendrogram of 18 starch-related genes based on expression analysis and their genomic localizations. a) A composite tree for 28 starch-related biosynthetic transcript as obtained from a heat map based on gene expression profiles reveals three gene clusters. b) Genomic localization showing the position of 18 starch-related biosynthetic genes on the 12 rice chromosomes. Some of each gene identifier has been omitted, and square color indicates starch-related gene type: red, amylose; blue, amylopectin; purple, ADP-glucose; dark green, starch phosphorylase

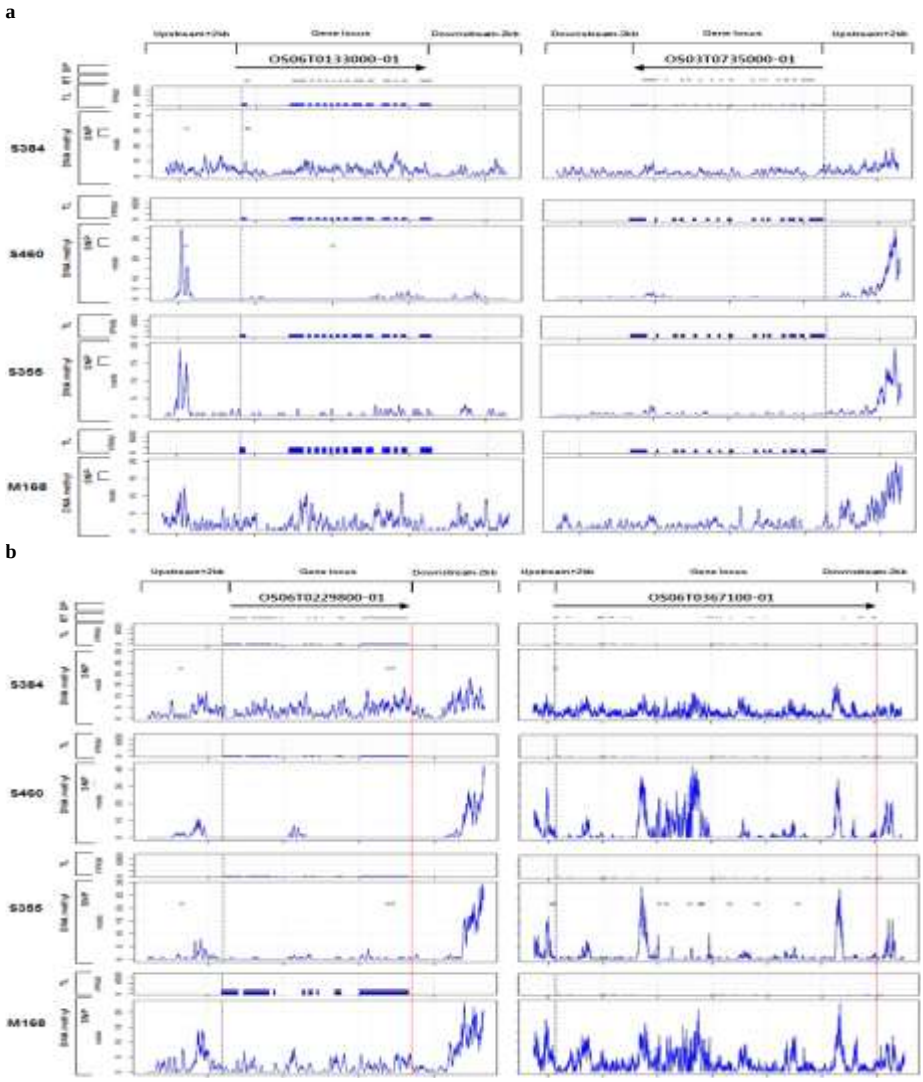

Fig. 4: Methylation patternsin promoter (-2 kb upstream) and TTR (+2 kb downstream) regions with SNP variation and FPKM gene expression values. Transcripts were selected based on S384 (amylose ratio 0.0%). SNPs are shown in green circles, and DNA methylation is expressed as the peak of the read count. Legend is as follows: S, Suweon; M, Milyang; DP, direction of progress; RT, representative transcript; TL, transcript level for FPKM; SNP,SNP variation. a) Four transcripts with increased methylation in the promoter region. Two transcripts have been omitted due to similar pattern.b) Two transcripts with increased methylation in both promoter and TTR region

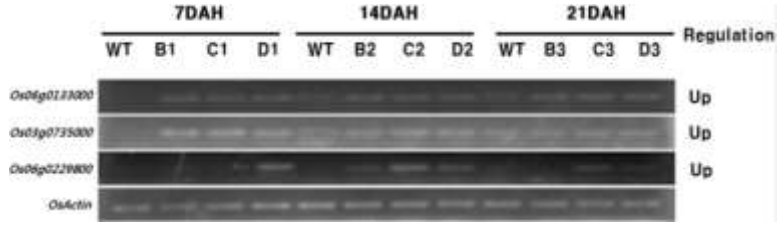

Fig. 5: Reverse transcription polymerase chain reaction (RT-PCR) analysis of the starch-related genes identified in this study. Expression of three candidate genes was assessed during three growth stages. The 7DAH, 14DAH, and 21DAH columns represent respective days after ear emergence. Abbreviations are as follows: WT, Suweon384; B, Suweon460; C, Suweon355; D, Milyang168

Fig. 6: The interrelationship among SNPs variation, gene expression of transcriptomeand methylation in each rice accession using Circos diagram. Three factors were showed by peak graph between promoter (up-steam: -2 kb) and TTR (down-stream: +2 kb) region. The legend is as follows. MET1: methylation of ear emergence + 7days, MET3: methylation of ear emergence + 21 days, LT1: gene expression on the transcript level of ear emergence + 7days, LT3: gene expression on the transcript level of ear emergence + 21 days

Using a Circos diagram, we visualized the SNP variation, gene expression, and methylation patterns at each growth stage in promoter (-2 kb upstream) and TTR (+2 kb downstream) regions. Although the Os06T0133000-01 transcript displayed significant differences in gene expression, no correlation was found between the three factors examined and amylose content in our rice accessions. Finally, we validated expression of four of the six DMR-associated genes by RT-PCR. Although there is no direct evidence that the three genes are directly involved in amylose biosynthesis, these candidate genes play a crucial role in shaping the DNA methylation landscape in rice.

Conclusion

Our results suggest a potential role for DNA methylation in the regulation of starch and sucrose metabolism. The data further suggest that DNA methylation plays a more important role in amylose biosynthesis than SNP variation. DMR-associated candidate genes may provide the basis for future selection strategies, potentially impacting rice breeding for the development of improved varieties. In future studies, it will also be interesting to assess changes in methylation that occur in rice cultivars grown with variable amylose.

Acknowledgements

This work was supported by the "Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ01045502)" Rural Development Administration, Republic of Korea.

References

Adusumalli, S., M.F.M. Omar, R. Soong and T. Benoukraf, 2015. Methodological aspects of whole-genome bisulfite sequencing analysis. *Brief Bioinform*., 16: 369–379

Bahaji, A., J. Li, Á.M. Sánchez López, E. Baroja Fernández, F.J. Muñoz, M. Ovecka, G. Almagro, M. Montero, I. Ezquer and E. Etxeberria and J. Pozueta Romero, 2014. Starch biosynthesis, its regulation and biotechnological approaches to improve crop yields. *Biotechnol. Adv.*, 32: 87–106

Caporaso, J.G., C.L. Lauber, W.A. Walters, D. Berg Lyons, J. Huntley, N. Fierer, S.M. Owens, J. Betley, L. Fraser and M. Bauer and N. Gormley, 2012. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*., 6: 1621–1624

Chen, X. and D.X. Zhou, 2013. Rice epigenomics and epigenetics: challenges and opportunities. *Curr. Opin. Plant Biol*., 16: 164–169

Chodavarapu, R.K., S. Feng, B. Ding, S.A. Simon, D. Lopez, Y. Jia, G.L. Wang, B.C. Meyers, S.E. Jacobsen and M. Pellegrini, 2012. Transcriptome and methylome interactions in rice hybrids. *Proc. Natl. Acad. Sci. U.S.A*., 109: 12040–12045

Cokus, S.J., S. Feng, X. Zhang, Z. Chen, B. Merriman, C.D. Haudenschild, S. Pradhan, S.F. Nelson, M. Pellegrini and S.E. Jacobsen, 2008. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. *Nature*, 452: 215–219

Decock, A., M. Ongenaert, R. Cannoodt, K. Verniers, B. De Wilde, G. Laureys, N. Van Roy, A.P. Berbegall, J. Bienertova Vasku and N. Bown and N. Clement. 2016. Methyl-CpG-binding domain sequencing reveals a prognostic methylation signature in neuroblastoma. *Oncotarget*, 7: 1960–1972

Farooq, M., A. Wahid, N. Kobayashi, D. Fujita and S.M.A. Basra, 2009. Plant drought stress: effects, mechanisms and management. *Agron. Sustain. Dev.*, 29: 185–212

Farooq, M., K.H.M. Siddique, H. Rehman, T. Aziz, D.J. Lee and A. Wahid, 2011. Rice direct seeding: Experiences, challenges and opportunities. *Soil Till. Res.*, 111: 87–98

Feng, S., S.J. Cokus, X. Zhang, P.Y. Chen, M. Bostick, M.G. Goll, J. Hetzel, J. Jain, S.H. Strauss and M.E. Halpern and C. Ukomadu, 2010. Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U.S.A*., 107: 8689–8694

Fu, F.F. and H.W. Xue, 2010. Coexpression analysis identifies Rice Starch Regulator1, a riceAP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol*., 154: 927–938

Fujita, N., M. Yoshida,N. Asakura, T. Ohdan, A. Miyao, H. Hirochika and Y. Nakamura, 2006. Function and characterization of starch synthase I using mutants in rice. *Plant Physiol.*, 140: 1070–1084

Garg, R., V.N. Chevala, R. Shankar and M. Jain, 2015. Divergent DNA methylation patterns associated with gene expression in rice cultivars with contrasting drought and salinity stress response. *Sci. Rep*., 5: 14922

Hanashiro, I., K. Itoh, Y. Kuratomi, M. Yamazaki, T. Igarashi, J.I. Matsugasako and Y. Takeda, 2008. Granule-bound starch synthase I is responsible for biosynthesis of extra-long unit chains of amylopectin in rice. *Plant Cell Physiol*., 49: 925–933

Harrison, A. and A. Parle McDermott, 2011. DNA methylation: a timeline of methods and applications. *Front. Genet.*, 2: 74

Kawahara, Y., M. de la Bastide, J.P. Hamilton, H. Kanamori, W.R. McCombie, S. Ouyang, D.C. Schwartz, T. Tanaka, J. Wu and S. Zhou and K.L. Childs, 2013. Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice*, 6: 4

Li, H. and R. Durbin, 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25: 1754–1760

Li, X., J. Zhu, F. Hu, S. Ge, M. Ye, H. Xiang, G. Zhang, X. Zheng, H. Zhang and S. Zhang and Q. Li, 2012. Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genom.*, 13: 300

Lu, Y.C., S.J. Feng, J.J. Zhang, F. Luo, S. Zhang and H. Yang, 2016. Genome-wide identification of DNA methylation provides insights into the association of gene expression in rice exposed to pesticide atrazine. *Sci. Rep*., 6: 18985

Ohdan, T., P.B. Francisco, T. Sawada, T. Hirose, T. Terao, H. Satoh and Y. Nakamura, 2005. Expression profiling of genes involved in starch synthesis in sink and source organs of rice. *J. Exp. Bot.*, 56: 3229–3244

Pandey, M.K., N.S. Rani, M.S. Madhav, R.M. Sundaram, G.S. Varaprasad, A.K. Sivaranjani, A. Bohra, G.R. Kumar and A. Kumar, 2012. Different isoforms of starch-synthesizing enzymes controlling amylose and amylopectin content in rice (*Oryza sativa* L.). *Biotechnol. Adv.*, 30: 1697–1706

Patel, R.K. and M. Jain, 2012. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. *PLOS ONE*, 7: e30619

Takuno, S. and B.S. Gaut, 2013. Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. *Proc. Natl. Acad. Sci. U.S.A*., 110: 1797–1802

Thitisaksakul, M., R.C. Jimenez, M.C. Arias and D.M. Beckles, 2012. Effects of environmental factors on cereal starch biosynthesis and composition. *J. Cereal Sci.*, 56: 67–80

Wang, J., N.C. Marowsky and C. Fan, 2014. Divergence of gene body DNA methylation and evolution of plant duplicate genes. *PLOS ONE*, 9: e110357

Wani, A.A., P. Singh, M.A. Shah, U. Schweiggert Weisz, K. Gul and I.A. Wani, 2012. Rice starch diversity: Effects on structural, morphological, thermal, and physicochemical properties—A review. *Comprehensive Rev. Food Sci. Food Saf.*, 11: 417–436

Xing, M.Q., Y. Zhang, S.R. Zhou, W.Y. Hu, X.T. Wu, Y.J. Ye, Y.P. Xiao, X. Li and H.W. Xue, 2015. Global analysis reveals the crucial roles of DNA methylation during rice seed development. *Plant Physiol*., 168: 1417–1432

Zeeman, S.C., J. Kossmann and A.M. Smith, 2010. Starch: its metabolism, evolution, and biotechnological modification in plants. *Annu. Rev. Plant Biol.*, 61: 209–234

Zemach, A., I.E. McDaniel, P. Silva and D. Zilberman, 2010. Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, 328: 916–919

Zhang, J.J., Y.C. Lu, J.J. Zhang, L.R. Tan and H. Yang, 2014. Accumulation and toxicological response of atrazine in rice crops. *Ecotoxicol. Environ. Saf.*, 102: 105–112

Zhang, H., Y. Lu, Y. Zhao and D.X. Zhou, 2016. OsSRT1 is involved in rice seed development through regulation of starch metabolism gene expression. *Plant Sci*., 248: 28–36

(Received 02 June 2017; Accepted 17 August 2017)