International Journal of Agricultural and Environmental Information Systems
Volume 14 • Issue 1

Expression Characteristics and Sequence Variation Analysis of Rice Starch Regulator 1 Gene in Japonica Rice With Transgressive Variation

Haiying Liu, Heilongjiang Academy of Agricultural Sciences, China
Yongcai Lai, Heilongjiang Academy of Agricultural Sciences, China*
Zhenhua Xu, Biotechnogogy Research Institute, Heilongjiang Academy of Agricultural Sciences, China
Zhonliang Yang, Biotechnogogy Research Institute, Heilongjiang Academy of Agricultural Sciences, China
Yanmin Yu, Biotechnogogy Research Institute, Heilongjiang Academy of Agricultural Sciences, China
Ping Yan, Biotechnogogy Research Institute, Heilongjiang Academy of Agricultural Sciences, China

ABSTRACT

The parents and transgressive variation lines of hybrids with significant difference in amylose content were selected to compare and analyze the accumulation characteristics of amylose and the change of OsRSR1 expression in grains in the process of grain filling, and the PCR technology was used to clone the OsRSR1 gene base sequence of four varieties. The results showed that the amylose content in grains increased gradually with grain filling process, the amylose content of offspring and parents with high amylose content were higher than the offspring and parents with low amylose content, hybrids could obtain the transgressive variation lines through the continuous directional selection of amylose content in grain, and the accumulation of amylose content in grain was closely related to genotypes. The expression quantity of OsRSR1 gene in grain was increasing during the grain filling process, the amylose content of grain was closely related to the activity of OsRSR1 gene, and the expression of grain OsRSR1 gene could also produce transgressive variation.

KEYWORDS

Japonica Rice, RSR1 Gene Expression, Sequence Variation, Transgressive Variation

INTRODUCTION

Transcription factors are also called trans-acting factors. Typical transcription factors consist of four functional domains. Transcription factors regulate transcription and expression of genes through interaction between functional areas and promoters cis-acting elements or functional areas of other transcription factors (Liu et al., 2001; Hou, 2014; Xie, 2019). For *OsRSR1,* promoter elements may produce a marked effect in response to ABA, ethylene, and abiotic stress signals,

DOI: 10.4018/IJAEIS.317417 *Corresponding Author

thereby regulating the expression of downstream stress resistance related genes (Li, et al., 2015). The *OsRSR1* gene is affected by low temperature, drought, and NaCl stress in adversity expression. *OsRSR1* is an AP2/EREBP family transcription factor that regulates starch synthesis in rice. Its expression is negatively correlated with the expression of the genes responsible for the synthesis of starch in the endosperm in seeds and sink tissues (Fu & Hue, 2010). The study of its physiological function showed that the lack of *OsRSR1* could lead to the increase of the expression of genes related to starch synthesis in the seeds. In the deletion mutant of the *OsRSR1* gene, the expression of a class of starch synthase genes was unregulated, the amylose content in rice seed was improved, and the amylopectin structure in rice seeds was changed. Therefore, the final morphology of starch granules in rice seeds was changed, the gelatinization temperature of starch was reduced, and the microstructure of amylopectin was changed. *OsRSR1* is a transcription factor of the APETALA2/ethylene response element binding protein family. There are two AP2/ERF domains in the structure of *OsRSR1* protein, which are the typical characteristics of the AP2 family transcription factors. There is a DNA binding area of two conserved sequence blocks: YRG element and RADY element in the AP2/ERF binding domain.

The occurrence of transgressive inheritance variation in quantitative characters is a common phenomenon in sexual hybrids. The sexual hybridization between rice varieties is still one of the main ways to cultivate new rice varieties at present and in the future. However, there are few studies on the relationship between the sequence variation of the transcription factor base sequence and its transcriptional expression quantity in the sexual hybrids among the related varieties. Therefore, this study selected two rice varieties with a significant difference in the amylose content in the grains as parents. It then took the amylose as the selection index and continuously and directionally selected the transgressive variants with a significant difference in the amylose content to compare and analyze the base sequence and protein domain of the grain *OsRSR1* gene between the parent and the transgressive variant. This was to provide a theoretical basis for elucidating the relationship between the base sequence variations, the amylose content of the *OsRSR1* gene, and the molecular mechanism of transgressive inheritance variation in the grain amylose content of the hybrid progeny.

MATERIALS AND METHODS

Test Materials and Test Methods

Two japonica rice varieties with significant difference in grain amylose content were selected, i.e., parent Xixuan 1 (18.48%), Tong 769 (15.81%), and transgressive offspring stable strains Dongnong 1101 (19.71%) and Dongnong 1124 (7.40%). In 2015–2016, the pot experiment was carried out in the Agricultural College of Northeast Agricultural University. The length of the pot was 100 cm, the width was 40 cm, and the height was 60 cm. From April 1 to 15, according to the growth period of the test materials, sowing was carried out in stages to ensure the heading stage was as consistent as possible. Furthermore, plug-seedlings in trays in a greenhouse, equidistantly dibble sowing of a single sprouting seed, and dry rice-nursery management were selected. On May 15, rice seedlings with consistent growth potential were selected for transplantation. Each variety was inserted in three pots and 24 seedlings were equidistantly planted in every pot. After seedling survival, 12 seedlings were planted and treated with normal fertilizer.

At heading stage, the rice ears of the same size and extension at the same time were selected and marked with a sign. On the 10th, 20th, and 30th day after heading, eight rice ears with signs were selected, respectively. Then, 20 grains in the middle of the ear, with the same grain filling, were selected. With shells and embryos removed at a low temperature, they were put into a sterile tube, which was then quickly placed into liquid nitrogen. Finally, they were stored at -80ºC to be used for quantitative analysis of fluorescence.

RNA Extraction and qRT-PCR Analysis

The total RNA of grains was extracted with TRIzol (Kang et al., 2013), and digested with RNase-free DNase I to eliminate the pollution of genomic DNA. The Oligo (dT) was used as the primer to synthesize the first chain of cDNA. According to the sequence of OsRSR1(AY685117.1) and actin (NM_197297) cDNA in rice, polymerase chain reaction (PCR) primers were designed by Premier 5.0 (Table 1). The real-time quantitative reverse transcription PCR (qRT-PCR) reaction system (20 μL) included SYBR Premix Ex-Taq (2×) 10 μL, upstream and downstream primers (10 μm) 0.5 μL, template 2 μL, and sterile water. There were three repetitions for each sample.

The qRT-PCR reaction procedure is: pre-denaturation at 95ºC for 30 seconds, denaturation at 95ºC for 10 seconds, annealing at 60ºC for 30 seconds, extending at 72ºC for 30 seconds, 40 cycles, and extending at 72ºC for 5 minutes. Referring to the ΔΔCT method of genes (Livak & Schmittgen, 2001), the relative expression quantity of the target gene relative to the reference gene (actin) was calculated using the $2^{-\Delta\Delta Ct}$ method. The relative expression quantity of the target gene equaled $2^{-\Delta\Delta Ct}$. Among them, ΔΔCt = (Ct target gene - Ct reference gene) experimental group, and (Ct target gene - Ct reference gene) control group.

Cloning and Full Sequence Analysis of Rice Starch Regulatory Factor (*OsRSR1*) Gene cDNA

Plasmids and Bacterial Strains

The Escherichia coli strain was DH5α, and pMDT-18 was a plasmid vector for gene cloning and construction.

PCR Amplifies the Full-Length cDNA of Rice Starch Regulatory Factor (OsRSR1) Gene

Referring to the nucleotide sequence of rice starch regulatory factor (*OsRSR1*) published on The National Center for Biotechnology Information (NCBI), primers were designed and synthesized (Table 1). Through the homologous comparison function of BLAST in NCBI, the designed primer specificity was evaluated. The full-length cDNA fragment of the rice starch regulatory factor (*OsRSR1*) gene was amplified by the PCR method from the cDNA library of the rice grain. PCR conditions are as follows: pre-denaturation at 94ºC for 10 minutes, denaturation at 94ºC for 1 minute, annealing at 61ºC for 1 minute, extending at 72ºC for 2 minutes, 32 cycles, and extending at 72ºC for 2 minutes.

Cloning and Identification of PCR Products

After the PCR products were separated by 0.7% agarose gel electrophoresis, the target fragments were recovered by NucleoTrap® gel extraction kit (Clontech). Then, they were connected to pMDT-18 vector to transform the Escherichia coli DH5α competent cells. On the Luria broth (LB) medium plate of 100 mg/L antimicrobial peptide (AMP) antibiotics containing IPTG (isopropyl beta-D-1-thiogalactopyranoside) and XG-gal (5-Bromo-4-Chloro-3-Indolyl-beta-D-Galactoside), sterilized

Table 1. PCR primers

	Gene Name	Primer Name	Primer Sequence Number	Primer Sequence (5¢–3¢)
qRT-PCR Primer	Actin	Actin-F	NM_197297	TTATGGTTGGGATGGGACA
		Actin-R		AGCACGGCTTGAATAGCG
	OsRSR1	q-RSR1-F	AY685117.1	TGCGGCAAGCAAGTCTACCT
		q-RSR1-R		AAGCTCTTCATCTGCCTCATGTC
PCR Primer	OsRSR1	Q-OsRSR1-F	AY685117.1	ATGGAGTTGGATCTGAACAACGTGGCGGAA
		Q-OsRSR1-R		TCAATGGTGGTGGTGATGGCGGCTTGACGA

toothpicks were used to pick out the white colonies. Through blue-white selection and a small amount of culture, the plasmid DNA was extracted using the alkaline lysis method (Xianguang et al., 2003) and identified by enzyme digestion.

cDNA Sequence Analysis

The cloning of rice starch regulatory factor gene cDNA was carried out in pMDT-18 vector. Shanghai Invitrogen Company was commissioned to determine the complete nucleotide sequence, and a bidirectional repeatable determination was carried out. The nucleotide sequence and its deduced amino acid sequence were analyzed on the website of NCBI, and the sequence comparative analysis was carried out based on the DNAMAN database.

RESULTS AND DISCUSSION

Comparison of Amylose Accumulation Characteristics in Parent and Transgressive Variation Line Grains at Different Grain Filling Stages

The results of multiple comparisons of the amylose content in grains of parents and transgressive variation lines at different grain filling stages are listed in Table 2.

Table 2 shows that the amylose content of grain increases with the grain filling process. On 10 days, 20 days, and 30 days after heading, the amylose content of the grains of the transgressive offspring Dongnong 1124 was significantly lower than that of parent Tong 769 with low amylose content. However, the amylose content of the grains of the transgressive offspring Dongnong 1101 were significantly higher than that of parent Xixuan 1, with a high amylose content, and the amylose content of Dongnong 1124 was significantly lower than that of Dongnong 1101. In the grain filling period, the amylose content of the offspring and parents with high amylose contents are higher than those of the offspring and parents with low amylose content. This indicates that the sexual hybrids between varieties with no major genetic differences can obtain the transgressive variation lines with significantly high and low grain amylose content through the continuous orientation selection of amylose content in grains. The accumulation of amylose content in grains is closely related to genotype, and it is the quantitative genetic trait controlled by genotypes.

The Change and Comparison of Osrsr1 Gene Expression in Grains of Parents and Transgressive Variation Lines During the Grain Filling Process

The change of the *OsRSR1* gene expression in grains of the parents and transgressive variation lines at different stages of the grain filling is shown in Table 3.

Table 3 shows that the variation trend of the *OsRSR1* gene expression in the grains of parents and transgressive variation lines is basically consistent during the grain filling process, and all of them are on the rise. The grain *OsRSR1* gene expression quantity gradually increases with the grain

Table 2. Comparison of amylose content in rice grain filling (%)

Variety	Days After Heading/d			Polished Rice
	10 days	20 days	30 days	
Xixuan 1	9.16b	15.18b	18.47a	18.74b
Tong 769	6.59c	12.45c	15.05b	15.78c
Dongnong 1101	10.45a	16.68a	18.93a	19.71a
Dongnong 1124	2.14d	5.23d	7.33c	7.40d

Note: Figures followed by the same lowercase letter mean significant difference at 0.05 level. The same as below.

Table 3. Relative expression levels of OsRSR1 genes parents and hybrid progenies after heading

Parents and Progenies	Days After Heading/d		
	10 days	20 days	30 days
Xixuan 1	1.000a	0.676a	1.341b
Tong 769	0.401c	0.659a	1.251c
Dongnong 1101	0.432b	0.668a	1.553a
Dongnong 1124	0.396c	0.479b	1.328b

filling process, and the expression quantity is small at the early stage of grain filling. The expression quantity increases rapidly in the middle and late filling stages, indicating that the expression of the *OsRSR1* gene is mainly in the middle and late stages of grain filling.

It can also be seen from Table 3 that there is no significant difference between the grain *OsRSR1* gene expression quantity of the transgressive variation line Dongnong 1124 with low amylose content and the parents on the 10th day after heading. However, on the 20th day after heading, they are all lower than the parent. There is a significant difference between the grain *OsRSR1* gene expression quantity of the transgressive variation line with high amylose content and parents on the 10th day and 30th day after heading, and they are significantly lower than the parents with high amylose content, and on the 20th day after heading, there is no significant difference in expression quantity. This shows that the difference in the expression quantity of the *OsRSR1* gene in the grains of the parents and transgressive variation lines, with significant difference in amylose content, varies with the grain filling period, and the expression quantity is closely related to the change of amylose content, and he expression quantity can show the transgressive feature.

Cloning and Similarity Comparison of the OsRSR1 Gene of Parents and Transgressive Variants Grains

According to the results of gene sequencing, the cDNA full sequence length of the *RSR1* gene of the parents and transgressive variation line grains is 1539 bp, which encodes 512 amino acids, respectively, and contains all the coding region sequences, which is exactly the same with the expected results. The cDNA full-length nucleotide sequence and the deduced amino acid sequence of the *OsRSR1* gene of the parent Xixuan 1 grain are shown in Figures 1 and 2, respectively.

The cDNA complete sequence of the tested materials, i.e., *OsRSR1*-Xixuan 1, *OsRSR1*-Tong 769, *OsRSR1*-Dongnong 1101, and *OsRSR1*-Dongnong 1124 is compared with the nucleotide sequence of *RSR1* published in GenBank, and the results are shown in Figure 2.

It can be seen in Figure 2(a) that compared with Nipponbare, the homology of the *OsRSR1* gene cDNA sequence of the parent Xixuan 1, Tong 769, and the transgressive variation lines Dongnong 1101 and Dongnong 1124 is 99.94%, 99.94%, 99.87%, and 99.94%, respectively. This proves that this study has successfully cloned the cDNA of *OsRSR1*, indicating that there is no significant cDNA sequence difference of starch synthesis regulatory factor *OsRSR1* gene between parents and transgressive variation lines with significant difference in grain amylose content, and the sequence has a very high conservatism. It can be seen in Figure 2(b) that compared with Nipponbare, the homology of the amino acid sequence derived from the *OsRSR1* gene of the parent Xixuan 1, Tong 769, and the transgressive variation lines Dongnong 1101 and Dongnong 1124 is 99.9%, indicating that there is no significant amino acid sequence difference of the rice starch synthesis regulatory factor between parents and transgressive variation lines with significant difference in grain amylose content, and the sequence has a very high conservatism.

Figure 1. Deduced cDNA sequence (top) and amino acid sequence (bottom) of gene in Xixuan 1

```
   1  ATGGAGTTGG ATCTGAACAA CGTGGCGGAA GGGGTGGTGG AGAAGCATGA GACGGCGGCG AGGAGCGACT
  71  CCGGCACGTC GGAGTCGTCG GTGCTCAACG GGGAGGCGTC TGGCGCGGCC ATCGCGCCGG CGGAAGAGGG
 141  GTCGAGCTCG ACGCCGCCGT CGCCGCCGCC GCCTCCCGCG AGGAGGAGGA GGCCACCCCC TCGCCGCCGC
 211  TCGGCGTCGG CGTCGGGCGA GAACGACGCC GACGACGACG AGGAGGAGGA GGCCACCCCC TCGCCGCCGC
 281  CGCACCACCA ACACCAGCAG CTGCTCGTCA CCCGGGAGCT ATTCCCTTCC GCCGCTCCCT CGCCGCAGCA
 351  TTGGGCGGAG CTCGGCTTCC TCCGCCCCGA CCCACCGCGC CCACACCCAG ACATCAGAAT CCTCGCCCAC
 421  GCGCCTCCCC CGGCGCCACC GCCGCCGCCG CCGCAGCCGC AGCCTCAGGC GGCCAAGAAA AGCCGCCGCG
 491  GGCCGCGCTC TCGCAGCTCG CAATACCGCG GCGTCACCTT CTACCGCCGC ACCGGCCGCT GGGAATCCCA
 561  CATCTGGGAT TGCGGCAAGC AAGTCTACCT AGGTGGATTC CACACTGCTC ACGCAGCTGC AAGGGCGTAC
 631  GACAGGGCGG CGATCAAGTT CAGGGGAGTA GAGGCTGACA TCAACTTCAA CCTGAGCGAC TACGAGGAGG
 701  ACATGAGGCA GATGAAGAGC TTGTCCAAGG AGGAGTTCGT GCACGTTCTC CGGCGACAGA GCACCGGCTT
 771  CTCCCGCGGC AGCTCAAAGT ACAGGGGTGT CACCCTCCAC AAGTGCGGCC GCTGGGAGGC TCGCATGGGC
 841  CAATTCCTTG GCAAGAAGTA CATATATCTT GGGCTATTCG ACAGCGAAGT AGAGGCTGCA AGGGCTTATG
 911  ATAAGGCTGC GATCAAATGC AATGGCAGAG AAGCCGTCAC CAACTTCGAG CCCAGCACAT ATGATGGTGA
 981  GCTGCCTACT GATGCTGCTG CTCAAGGAGC CGATGTGGAT CTGAACCTGA GAATATCTCA GCCTGCAGCC
1051  TCACAGCAGA GCCCCAAGAG GGATAGCGGC TCCCTTGGCC TGCAAATCCA CCATGGATCA TTTGAAGGTT
1121  CTGAATTCAA GAGAGCAAAG AATGATGCAG CTCCCTCTGA ACTTGCTAGC CGCCCTCATC GGTTCCCTCT
1191  TCTGACCGAG CATCCGCCAA TCTGGACTGC CCAACCTCAT CCCCTATTCC CAAATAATGA GGATGCATCC
1261  AGATCATCGG ATCAGAAGAG GAAGCCATCA GAGGGGGTAG CTGTTCCAAG CTGGGCATGG AAGCAGGTGA
1331  GCCATCATCA CCCTGCTCCT CCTCACACGC TGCCATTGCC CTTCTTCTCC TCCTCCTCGT CGTCGCCGTC
1401  GTCGTCCTCC GCTGCAGCAT CATCAGGATT CCCCAAAGCC GCCACGACAG CAGCTGCTGC CCAACACACT
1471  GCCACCCTCC GGTTCGACCC GACGGCGCCG TCGTCTTCGT CGTCAAGCCG CCATCACCAC CACCATTGA
```

```
   1 MELDLNNVAE GVVEKHETAA RSDSGTSESS VLNGEASGAA IAPAEEGSSS TPPSPPPPPA
  61 AVLEFSILRS SASASGENDA DDDEEEEATP SPPPHHQHQQ LLVTRELFPS AAPSPQHWAE
 121 LGFLRPDPPR PHPDIRILAH APPPAPPPPP PQPQPQAAKK SRRGPRSRSS QYRGVTFYRR
 181 TGRWESHIWD CGKQVYLGGF DTAHAAARAY DRAAIKFRGV EADINFNLSD YEEDMRQMKS
 241 LSKEEFVHVL RRQSTGFSRG SSKYRGVTLH KCGRWEARMG QFLGKKYIYL GLFDSEVEAA
 301 RAYDKAAIKC NGREAVTNFE PSTYDGELPT DAAAQGADVD LNLRISQPAA SQQSPKRDSG
 361 SLGLQIHHGS FEGSEFKRAK NDAAPSELAS RPHRFPLLTE HPPIWTAQPH PLFPNNEDAS
 421 RSSDQKRKPS EGVAVPSWAW KQVSHHHPAP PHTLPLPFFS SSSSSPSSSS AAASSGFPKA
 481 ATTAAAAQHT ATLRFDPTAP SSSSSSRHHH HH-
```

Figure 2. Genetic distance OsRSR1 gene sequences and deducing amino acids for the tested materials

Distance matrix a

Homology matrix

	1	2	3	4	5
1	100%	0.004	0.004	0.004	0.002
2	99.8%	100%	0.004	0.004	0.002
3	99.8%	99.6%	100%	0.004	0.002
4	99.8%	99.6%	99.6%	100%	0.002
5	99.8%	99.6%	99.6%	99.6%	100%

Distance matrix b

Homology matrix

	1	2	3	4	5
1	100%	0.001	0.001	0.001	0.001
2	99.9%	100%	0.002	0.001	0.002
3	99.9%	99.8%	100%	0.002	0.001
4	99.9%	99.9%	99.8%	100%	0.001
5	99.9%	99.9%	99.8%	99.9%	100%

Note: 1-Nipponbare; 2-Xixuan 1; 3- Dongnong 1101; 4-Dongnong 1124; 5-Tong 769

Full Length cDNA Sequence Alignment of OsRSR1 Gene in Parent and Transgressive Variation Line Grains

The triplet code variable loci in the results of the *OsRSR1* gene cDNA complete sequence base alignment of the Nipponbare, parents, and transgressive variation line grains are shown in Table 4.

Table 4 shows that the bases of three loci of the transgressive variation line Dongnong 1101 change compared with the female parent Xixuan 1. Among them, the base T of 160 locus is converted into C, the triplet code changes from TCG to CCG, and the amino acid changes from serine (Ser)

Table 4. Comparison of triplets for gene cDNA sequence of OsRSR1 Nipponbare, tested parent, and their derived progenies

Variety	Base Site				
	157	160	395	834	1432
Nipponbare	CCG	TCG	CAC	CGC	TCA
Xixuan 1	CCG	TCG	CAC	CGC	CCA
Dongnong 1101	CCG	CCG	CAC	CGT	TCA
Dongnong 1124	CCG	TCG	CGC	CGC	TCA
Tong 769	TCG	TCG	CAC	CGC	TCA

to proline (Pro). The base C of 834 locus is converted into T, the triplet code changes from CGC to CGT, and there is no change in amino acids. The base C of 1432 locus is converted into T, the triplet code changes from CCA to TCA, and the amino acid changes from proline (Pro) to serine (Ser). Compared with the male parent, the bases of the three loci change. Among them, the base C of 157 locus and 160 locus is converted into T, the triplet code changes from CCG to TCG, and the amino acid changes from proline (Pro) to serine (Ser). For 834 locus, the triplet code changes from CGT to CGT, and there is no change in amino acids. Compared with the parents, the bases of four loci in the transgressive variation line Dongnong 1101 are different from the parents. Among them, the change of the bases of 157 locus, 160 locus, and 1432 locus, and the change of the triplet code cause the change of the amino acid, and the change of the base of the 834 locus does not cause the change in amino acids.

The bases of two loci of the transgressive variation line Dongnong 1124 change compared with the female parent Xixuan 1. Among them, the base G of 395 locus is converted into A, the triplet code changes from CGC to CAC, and the amino acid changes from arginine (Arg) to histidine (His). The base T of 1432 locus is converted into C, the triplet code changes from TCA to CCA, and the amino acid changes from serine (Ser) to proline (Pro). Compared with the male parent Tong 769, the bases of two loci change. Among them, the base C of 157 locus is converted into T, the triplet code changes from CCG to TCG, and the amino acid changes from proline (Pro) to serine (Ser). The base G of 395 locus is converted into A, the triplet code changes from CGC to CAC, and the amino acid changes from arginine (Arg) to histidine (His). Compared with the parents, the bases of three loci in the transgressive variation line Dongnong 1124 are different from the parents. Among them, the change of the bases of 157 locus, 395 locus, and 1432 locus causes the change of the amino acid.

The comparison of the transgressive variation line Dongnong 1101 and Dongnong 1124 showed that the bases of three loci and the triplet code were different between the two variants. Among them, the change of the triplet codes of 160 locus, and 395 locus, caused the change of the amino acid, the change of the triplet code of the 834 locus did not cause the change in amino acids.

A comparison of the above base sequences showed that the base after mutation could form a same sequence or a different sequence as the female or male parent, or a sequence that is different from both parents. Moreover, it could form a synonymous triple code or a mutational triplet code. The offspring produced by hybridization could form a polymorphic gene with individual different bases and a mutant gene that changed the protein amino acid sequence in the process of gene separation and stability through the base change of certain site.

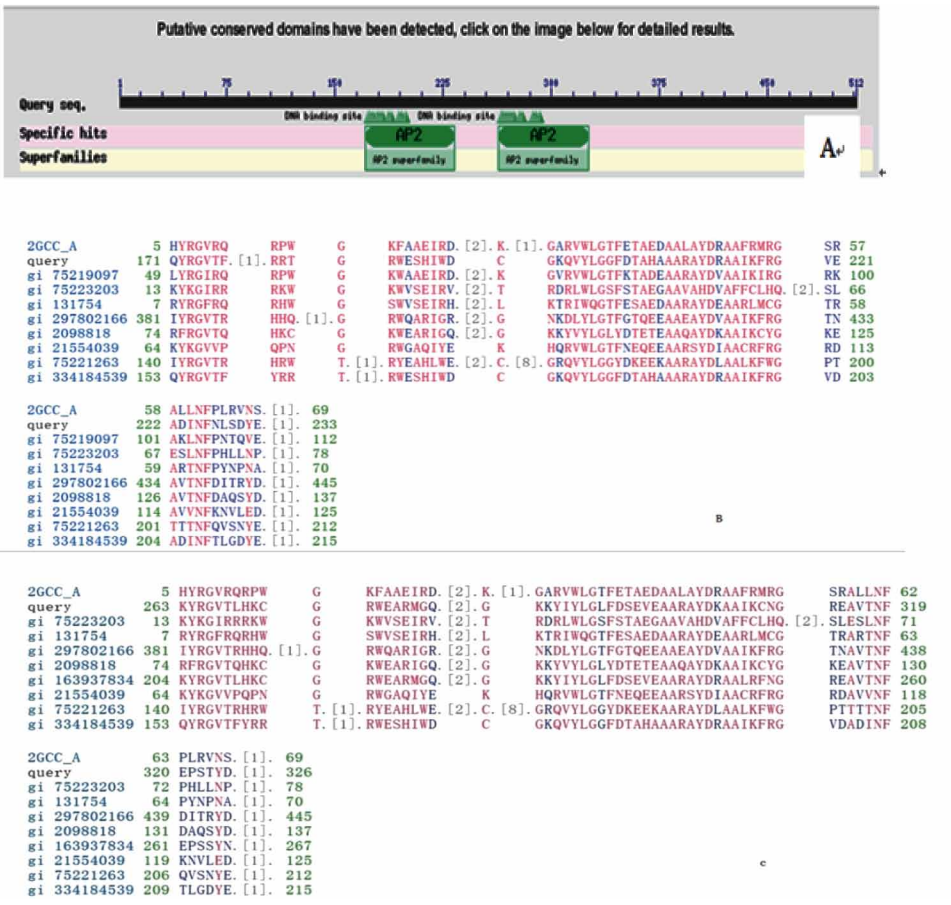
Sequence Structure Analysis of OsRSR1 Gene in Rice

The *OsRSR1* nucleotide sequence analysis between the parent and transgressive variation showed that the four cloned cDNA sequences all encoded the protein open reading frame consisting of 512 amino acids. The estimated molecular weight is 55.6 kDa, which belongs to the AP2 subfamily of the AP2/EREBP transcription factor family, and there are two AP2-DNA domains. The AP2-DNA

domains are respectively located in 171–233 and 263–326, as shown in Figure 3(A, B, C). Each AP2/EREBP binding domain has two conservative sequence modules (block): the YRG element and the RADY element. The YRG element is composed of 19–22 amino acid residues, which is extremely alkaline. It contains a conservative YRG amino acid motif and is located at the N end of AP2 domain, which is beneficial to the combination with DNA. The RAYD element consists of 42–43 amino acid residues, and there is a core sequence with high conservatism composed of 18 amino acid residues in the amino acid residues. This sequence can form an amphipathic alpha-helix (Chen et al., 2011). These two elements may participate in the interaction between AP2/EREBP transcription factors and other transcription factors or DNA, and mainly regulate the development process of the flowers, meristem, ovule, and seeds.

Four key enzymes were involved in the synthesis and accumulation of rice grain starch. The gene expression process of starch synthase is regulated by some transcription factors, and *OsRSR1* is negatively correlated with the gene of starch synthase (Fu & Hue 2010). *flo*2 can encode a structural protein with RPT, and the structural protein can regulate the expression features of the key enzyme gene of starch synthesis (She et al., 2010). Zinc finger protein *OsbZIP58* can activate the synthesis of rice starch, affect the quantity of starch particles, and change the proportion of short-chain and medium-long-chain (Wang et al., 2013). Albani et al., (1997)research results indicate that some transcription factors in the rice bZIP family can regulate the specific expression characteristics of the

Figure 3. Analysis of conserved AP2/EREBP DNA binding domain among cDNA of OsRSR1-Xixuan1 (A, B, C)



OsGBSSI gene in seeds by binding the GCN4 gene sequence in GBSSI promoter region endosperm box. The starch regulatory factor is negatively correlated with the activity of starch synthetase in rice (Fu & Hue, 2010). The results showed that for the parents and transgressive variation lines, the amylose content of grains increased with the grain filling process from the perspective of the accumulation characteristics of amylose in grains during the grain filling process. During the grain filling process, the variation trend of the *OsRSR1* gene expression of grains was basically the same, showing a rising curve. The relative expression quantity of the *OsRSR1* gene increased gradually along with the grain filling process, which was consistent with the amylose accumulation trend. There was a difference in the expression of the *OsRSR1* gene in grains of the parent and transgressive variation lines during grain filling, and the expression quantity was closely related to the change of amylose content, which could express the transgressive feature. Therefore, the amylose content of rice grain was closely related to the expression quantity of the *OsRSR1* gene.

Studies have pointed out that in nature there is a lot of allele expression (variation) in the plant body. The mutation of these gene sequences directly leads to the phenotypic variation of plant traits (Yong-Sheng et al., 2009). The changes in bases on different loci in the biological genes are the intrinsic reasons for the formation of multiple alleles and genetic polymorphisms. The study of Branlard (1987) showed that when the wheat grain protein content varied between 10% and 15%, the Glu-1 allelic variation had a significant effect on bread processing quality. However, when the wheat grain protein content was more than 15%, it had no effect on bread processing quality. The study of Wang et al. (2014) showed that the insertion or deletion of base fragments in the intron region was closely related to the variation of phenotypic characteristics. Due to the deletion and insertion of base fragments in the intron region, the transcriptional level of Vrn-B1c increased, and the heading of the allelic variation Vrn-B1c is earlier than Vrn-B1a (Shcherban et al., 2013). The results of this study show that although the homology of the base sequence in the *OsRSR1* gene coding region among parents and transgressive variation lines with significant difference in grain amylose content was high, the amino acid sequence was somewhat conservative. The similarity of the gene sequence after base mutation was over 99%, and the base sequence of the *OsRSR1* gene and the protein amino acid sequence among different varieties and offspring with significant difference in amylose content were not completely consistent. Compared with the parent, the grain amylose content transgressive variation lines, through directional selection of sexual hybridization, still changed at some base loci. These results indicated that the sexual hybrids between varieties generated alleles and isoforms with individual different bases and amino acids in the process of gene isolation and recombination through base conversion or transversion, which was an intrinsic molecular mechanism and an important way for biologic genes to produce polymorphism, phenotypic traits, and genetic variation. Previous studies have also shown that the single nucleotide polymorphisms (SNPs) base insertion loss and structural changes were the most variations in most biological genomes (Kruglyak, 1997). The results of this study also show that the change in the sequence of gene bases is the intrinsic basis for alleles to produce the polymorphism between different varieties. However, because of the degeneracy of the triplet code, no changes in each base could lead to changes in the amino acid sequence and the changes in the function of the protein.

The transcription factor is a class of nuclear proteins that recognize and combine specific DNA regulatory sequences and activate or inhibit transcription. Its quantity or activity abnormality can cause the abnormal expression of genes that play a key role in cell growth and differentiation. The base mutation of the transcription factor may occur in the structural functional areas and unstructured functional areas of the transcription factors. The study of Ying et al. (2014) indicated that the GATA-4 gene was amplified in patients with conotruncal heart defects, and in healthy people without basic congenital heart disease. The human genes in patients with conotruncal heart defects identified that the base of one missense mutation 799 locus changed from G to A, thus the valine (V) at 267 locus of the amino acid sequence mutated to methionine (M). The mutation was located between two zinc finger domains, and the mutation was near the protein kinase A phosphorylation locus. Therefore, it

was presumed that this mutation was likely to directly affect the DNA binding and transcription of the zinc finger domain. 0.1% to 0.5% of total nucleic acid variation in the human body causes many diseases, including cancer (Yang et al., 2020), and the mutation of these gene sequences directly leads to the phenotypic variation of plant traits (Yong-Sheng et al., 2009). However, the study by Gao et al. (2021) indicated the mutations of these heading date alleles might not be the cause of the heading date changes. The results of this study show that the variation of the individual locus of the base sequence of the parent and the transgressive variation lines lead to the variation of the amino acid sequence. The amino acid sequences of parents and transgressive variation lines had two loci variations compared with the Nipponbare. These variations did not occur in the DNA binding area or the conservative structural functional area of the model area of EAR (with LxLxL and $^{L}/_{F}DLN^{L}/_{F}$ (X) P types) in C end, and the protein structure of the functional area of the transcription factor OsRSR1 did not change. However, the amylose content of the transgressive lines is significantly higher or lower than that of parents, and rice quality of ultra-low parent variation lines is better than parents. The new findings of this study were a difference in the expression of the *OsRSR1* gene in the grains of the parent and transgressive variation lines, and the variation of the individual locus of the base sequence of the parent and the transgressive variation lines. The amylose content of the rice grain was closely related to the expression quantity of the *OsRSR1* gene, and the transgressive variation of transgressive hybrids was not induced by the structural function of the transcription factor *OsRSR1*. The synthesis and the transgressive phenomenon of grain starch was another way, such as the change of the *OsRSR1* gene expression, or the results of the common interaction of *OsRSR1* transcription factors, with other transcription factors or DNA caused by the core sequence contained by the DNA binding region of the transcription factor *OsRSR1* (Liu et al., 2001). We can use the whole genome sequencing method to obtain high-throughput SNP genotypes by means of simplified genome sequencing or gene chip, analyze the structural differences between the genomes of different individuals of parents and transgressive variation lines, and combine the phenotypic traits of parents and transgressive variation lines. The main genes controlling individual traits are obtained, and the genetic mechanism of traits related to rice eating quality is analyzed.

CONCLUSION

The variation trend of the *OsRSR1* gene expression of grains is basically the same, showing a rising curve. The results show that the relative expression quantity of the *OsRSR1* gene increases gradually along with the grain filling process, which is consistent with the amylose accumulation trend. Although the homology of the base sequence in the *OsRSR1* gene-coding region among parents and transgressive variation lines with significant difference in grain amylose content is high, the amino acid sequence is rather conservative. Compared with the parent, the grain amylose content transgressive variation lines, through directional selection of sexual hybridization, still changed at some base loci. The variation of the individual locus of the base sequence of the parent and the transgressive variation lines leads to the variation of the amino acid sequence. These variations do not occur in functional areas of the model area, and the protein structure of the functional area of the transcription factor OsRSR1 does not change.

AUTHOR NOTE

We have no conflicts of interest to disclose. The authors of this manuscript have covered open Access Funding for this article.

This study was supported by the project of scientific research fund of Heilongjiang Scientific Research Institute (CZKYF2022-1-C042) and China Rice Industrial Technology System (CARS-01-59).

REFERENCES

Albani, D., Hammond-Kosack, M. C., Smith, C., Conlan, S., Colot, V., Holdsworth, M., & Bevan, M. W. (1997). The wheat transcriptional activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *The Plant Cell*, *9*(2), 171–184. doi:10.1105/tpc.9.2.171 PMID:9061949

Branlard, G. (1987, May). Prediction of the bread wheat quality from HMW glutenins and gliadins. In *3. International Workshop on Gluten Proteins*. World Scientific Publishing Co Pte Ltd.

Chen, L., Fang, C., Wang, Y., Li, Y., Jiang, C., & Liang, M. (2011). Cloning and expression analysis of stress-resistant ERF genes forn tea plant [Camellia sinensis (L.) O. Kuntze]. *Journal of Tea Science, 31*(1), 53–58. doi: 10.13305/j.cnki.jts.2011.01.009

Fu, F. F., & Xue, H. W. (2010). Coexpression analysis identifies Rice Starch Regulator1, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiology, 154*(2), 927–938. https://doi.org/10.1104/pp.110.159517

Gao, R., Zhang, H., Xuan, N., Liu, X., Zhang, H., Zhang, M., & Yao, F. (2021). Variation analysis of heading date alleles using rice single segment substitution lines. *Huabei Nongxuebao*, *36*(5), 29–34. doi:10.7668/hbnxb.20192258

Hou, S. Y., Sun, Z. M., Guo, B., Wang, Y. G., Li, G. Q., & Han, Y. (2014). Cloning and expression analysis of two C2H2 transcription factors in soybean. *Plant Physionlogy Journal*, *50*(5), 665–674. doi:10.13592/j.cnki.ppj.2013.0447

Kang, G. Z., Xu, W., Liu, G. Q., & Guo, T. C. (2013). Comprehensive analysis of the transcription of starch synthesis genes and the transcription factor RSR1 in wheat (Triticum aestivum) endosperm. *Genome*, *56*(2), 115–122. doi:10.1139/gen-2012-0146 PMID:23517321

Kruglyak, L. (1997). The use of a genetic map of biallelic markers in linkage studies. *Nature Genetics, 17*(1), 21–24. doi: 10.1038/ng0997-21. PMID: 9288093

Li, Z., Zhang, S., & Huang, R. (2015). The expression of rice OsSHAT1 and OsRSR1 in response to hormones and abiotic stresses. *Shengwu Jishu Tongbao*, *31*(3), 88. doi:10.13560/j.cnki.biotech.bull.1985.2015.04.012

Liu, Q., Zhang, G., & Chen, S. (2001). Structure and regulatory function of plant transcription factors. *Chinese Science Bulletin*, *46*(4), 271–278. doi:10.1007/BF03187184

Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods (San Diego, Calif.)*, *25*(4), 402–408. doi:10.1006/meth.2001.1262 PMID:11846609

Shcherban, A. B., Khlestkina, E. K., Efremova, T. T., & Salina, E. A. (2013). The effect of two differentially expressed wheat Vrn-B1 alleles on the heading time is associated with structural variation in the first intron. *Genetica*, *141*(4), 133–141. doi:10.1007/s10709-013-9712-y PMID:23564445

She, K. C., Kusano, H., Koizumi, K., Yamakawa, H., Hakata, M., Imamura, T., Fukuda, M., Naito, N., Tsurumaki, Y., Yaeshima, M., Tsuge, T., Matsumoto, K., Kudoh, M., Itoh, E., Kikuchi, S., Kishimoto, N., Yazaki, J., Ando, T., Yano, M., & Shimada, H. et al. (2010). A novel factor FLOURY ENDOSPERM2 is involved in regulation of rice grain size and starch quality. *The Plant Cell*, *22*(10), 3280–3294. doi:10.1105/tpc.109.070821 PMID:20889913

Wang, J. C., Xu, H., Zhu, Y., Liu, Q. Q., & Cai, X. L. (2013). OsbZIP58, a basic leucine zipper transcription factor, regulates starch biosynthesis in rice endosperm. *Journal of Experimental Botany*, *64*(11), 3453–3466. doi:10.1093/jxb/ert187 PMID:23846875

Wang, X., Ju, L., Liu, F., Zhang, Y., Zhang, F., Fu, X., & Zhang, X. (2014). Vernalization effects of dominant alleles Vrn-B1a and Vrn-B1b and their distributions in cultivars from Yellow and Huai River Valleys Facultative Winter Wheat Zone. *Zuo Wu Xue Bao*, *40*(3), 439–446. doi:10.3724/SP.J.1006.2014.00439

Xianguang, Y., Zhiguang, Q., Baocun, Z., & Wenshi, M. (2003). The compare study on the method of extracting plasmid DNA. *Shengwu Jishu Tongbao*, *6*, 24–26. doi:10.13560/j.cnki.biotech.bull.1985.2003.06.006

Xie, L., Zhang, Y. F., & Wu, C. L. (2019). Microphthalmia family of transcription factors associated renal cell carcinoma. *Asian Journal of Urology*, *6*(4), 312–320. doi:10.1016/j.ajur.2019.04.003 PMID:31768316

Yang, D., Sun, Y., Chang, F., Tian, H., Liu, C., & Li, Z. (2020). Highly specific quantification of mRNA mutation in single cells based on RNase H cleavage-assisted reverse transcription (RT)-PCR. *Chinese Chemical Letters*, *31*(5), 1095–1098. doi:10.1016/j.cclet.2019.09.015

Ying, T., Wang, X., & Zhang, Y. (2014). A novel GATA-4 mutation identified in children with cardiac conotruncal defects. *Hainan Yixueyuan Xuebao*, *20*(8), 1022–1025. doi:10.13210/j.cnki.jhmu.20140630.002

Yong-Sheng, W. U., Xin-Hai, L. I., & Zhuan-Fang, H. A. O. (2009). Genomic DNA sequence, gene structure, conserved domains, and natural alleles of Gln1-4 gene in maize. *Zuo Wu Xue Bao*, *35*(6), 983–999. doi:10.3724/SP.J.1006.2009.00983

Liu Haiying, female, was born in 1978. She works in the Biotechnology Research Institute of Heilongjiang Academy of Agricultural Sciences, and studies in the postdoctoral mobile station, Heilongjiang Academy of Agricultural Sciences. My research direction is rice genetics and breeding work.

Lai Yongcai, male, Ph. D., was born in 1964. He works at the Heilongjiang Academy of Agricultural Sciences.

Zhenhua Xu, male, master, was born in 1987, His research direction is rice genetics and breeding work. He works in the Biotechnology Research Institute of Heilongjiang Academy of Agricultural Sciences.

Zhongliang Yang, male, master, was born in 1973, His research direction is rice genetics and breeding work. He works in the Biotechnology Research Institute of Heilongjiang Academy of Agricultural Sciences.

Yu Yanmin, male, Ph.D., was born in 1981, His research direction is rice genetics and breeding work. He works in the Biotechnology Research Institute of Heilongjiang Academy of Agricultural Sciences.

Ping Yan, male, master, was born in 1967, His research direction is rice genetics and breeding work. He works in the Biotechnology Research Institute of Heilongjiang Academy of Agricultural Sciences.