UDC 575.630
https://doi.org/10.2298/GENSR2003991B
Original scientific article

GENETIC DIVERSITY, POPULATION STRUCTURE AND ALLELE MINING OF GENES GOVERNING GRAIN SIZE RELATED TRAITS IN RICE (*Oryza sativa* L.)

Swarajyalakshmi N. BOLLINENI[1], Bukya J. NAIK[1], Suresh E. NAIK[1], Gopalakrishnamurthy KADAMBARI[1], G. Eswar REDDY[1], S. MALATHI[6], Eswar N.P. REDDY[1], Srividhya AKKAREDDY[2], Reddi M. SEKHAR[3], Bahben TANTI[4], B. Krishna VENI[5], B. Vijaya LAKSHMI[5] and Lakshminarayana R. VEMIREDDY[1,3*]

[1]Department of Molecular Biology and Biotechnology, S.V Agricultural College, Acharya NG Ranga Agricultural University (ANGRAU), Tirupati-517502, Andhra Pradesh, India
[2]Regional Agricultural Research Station, ANGRAU, Tirupati-517502, Andhra Pradesh, India
[3]Department of Genetics and Plant Breeding, S.V Agricultural College, Acharya NG Ranga Agricultural University (ANGRAU), Tirupati-517502, Andhra Pradesh, India
[4] Department of Botany, Gauhati University, Guwahati, Assam, India
[5]Rice Research Unit, ANGRAU, Bapatla-522101, Andhra Pradesh, India
[6]Indian Institute of Rice Research, Hyderabad-500030, India

Bollineni N.S., B.J. Naik, S. E. Naik, G. Kadambari, G. E. Reddy, S. Malathi, E.N.P. Reddy, S. Akkareddy, R. M. Sekhar, B. Tanti, B. K. Veni, B. V. Lakshmi, L. R. Vemireddy (2020). *Genetic diversity, population structure and allele mining of genes governing grain size related traits in rice (Oryza sativa L.)*- Genetika, Vol 52, No.3, 991-1007.

The present study was undertaken for allele mining of genes governing grain dimensions viz., grain length, grain width, grain thickness, grain size and grain weight. The molecular markers linked to the reported genes for these traits were used to screen 124 diverse rice genotypes. Thirty-two molecular markers used in this study produced a total of 86 alleles among 124 rice genotypes. The number of alleles ranged from 2 to 4 with an average of 2.58 alleles per locus. A dendrogram consisting of 124 rice genotypes revealed that all the genotypes can be divided into two groups. An analysis of the model-based population structure using simple sequence repeats (SSRs) covering all 12 chromosomes provided evidence of a significant population structure in the rice genotypes. The novel alleles identified in the study could be of great value for development of consumer-targeted rice varieties.

Corresponding author: Lakshminarayana R. Vemireddy, Department of Genetics and Plant Breeding, S.V Agricultural College, Acharya NG Ranga Agricultural University (ANGRAU), Tirupati-517502, Andhra Pradesh, India. Phone: +91 9440507965, E-mail: drvlnreddy@gmail.com

Keywords: Rice, Grain size, Allele mining, Population structure, and Marker-trait associations.

INTRODUCTION

Rice is one of the most important cereal crops of the world nourishing more than 50% of the world population. Rice is also an excellent model crop plant owing to its small genome size (*Indica*- 390 Mb and *Japonica*- 430 Mb), availability of a high-quality reference genome and innumerable transcriptomic and proteomic resources. Although rice yields have increased tremendously after the Green Revolution in the 1960's and the success of hybrid rice technology in the 1970's, the current day yield levels have almost reached a plateau. As per the predictions, 50% more yield is to be produced to feed the anticipated 9 billion human population by 2050.

Rice yield, a complex trait, is governed by numerous major and minor factors such as the number of productive tillers, panicle length, number of panicles per plant, grain weight, grain filling rate and grain size. Among them, the grain size is determined by the grain length, width and thickness, which in combination affect the grain weight. The physical appearance of rice grain is one of the main components of rice grain quality and major factor defining the market value of the crop. Preference for rice grain shape, however, varies among consumer groups. For instance, long and slender grain varieties are preferred by consumers in the USA and West Europe, and in most Asian countries including China, India, Pakistan and Thailand; in contrast, consumers in Japan, South Korea and Sri Lanka prefer short and bold grain cultivars (UNNEVEHR *et al*., 1992; JULIANO and VILLAREAL, 1993). Therefore, a comprehensive understanding of grain shape is warranted for designing varieties according to the market preference.

The complete whole genome sequence of rice has largely facilitated the discovery of several QTLs (Quantitative Trait Loci) in rice. To date, over 8500 QTLs governing different agronomically important traits of rice, including grain size have been mapped using various segregating populations generated from diverse parents (www.gramene.org). In recent years, several major QTLs affecting the rice grain size have been cloned and characterized. Some of the major genes/QTLs governing grain size include *GS3, GW2, qSW5/GW5, GS5, qGL3/qGL3.1, GW8, TGW6, GLW7, Gn1a, DEP1,* and *OsSPL14*. Among them, the major QTL, *GS*3 for grain length, which encodes a protein containing a putative PEPB-like domain, a transmembrane region, a putative TNFR/NGFR family cysteine-rich domain and a VWFC module (MAO *et al*., 2010). For grain width, GW2 was identified, which encodes a RING-type E3 ubiquitin ligase and loss of function leads to an increase in grain width and weight (XIAN-JUN *et al*., 2007). Another important QTL for grain width, *GW*5 encodes a novel nuclear protein of 144 amino acids that is localized to the nucleus and likely acts in the ubiquitin-proteasome pathway to regulate cell division during seed development with a 1212-bp deletion associated with the increased grain width (WENG *et al*., 2008). Among the other cloned genes for grain size, *qSW*5 (SHOMURA *et al*., 2008), *qGL*3.1 (QI *et al*., 2012), *GIF*1 (WANG *et al*., 2008), *TGW*6 (ISHIMARU *et al*., 2013), *GS*5 (LI *et al*., 2011) and *GW*8 (WANG *et al*., 2012 were reported as major regulators of grain size. From the recently cloned and characterized genes, it is revealed that multiple signaling pathways such as ubiquitination-mediated proteasomal degradation, phytohormones, and G-protein signaling pathways are involved in the determination of grain length (ZUO *et al*., 2014). However, the molecular mechanism underlying the control of grain size is still unclear.

During rice domestication and improvement, several valuable alleles in the germplasm were remained unused. Uncovering of the superior alleles for grain size facilitates the development of consumer-specific rice varieties. Hence, the present study was undertaken to identify novel or superior alleles for grain size traits in rice. Here, we evaluated 124 rice genotypes with diverse grain size traits to find out the allelic variation of major grain size regulating genes viz., *GS*3, *GS*5, *GLW*7, *GW*5/*qSW*5, *qGRL7.1*, *GS*2 and *qsgw7* through allele mining. In addition, we also studied their allelic diversity in the rice germplasm.

MATERIALS AND METHODS

Plant material

A total of 124 rice genotypes with different grain shapes and sizes were used in this experiment (Supplementary Table 1). The seeds were sown in the nursery and transplanted in the field with three replications in a randomized block design with 20x15cm spacing at wetland farm, S.V. Agricultural College, Acharya NG Ranga Agricultural University (ANGRAU), Tirupati, Andhra Pradesh, India.

Phenotyping of grain- size-related traits

At the stage of plant maturity, the seeds were collected and the grain length (GL) and grain width (GW) were recorded (in mm) as average from 10 completely filled and matured grains using a vernier caliper. The grain length-width ratio (GLWR) was estimated by dividing the grain length by the grain width of each genotype. The 1000 grain weight (TGW) was measured using an electronic digital weighing balance by taking 100 filled grains from each sample and the values were multiplied by 10 factors to get 1000 grain weight in grams. The frequency distribution for GL, GW, GLWR and TGW was analyzed using Microsoft-excel (2007). The genotypes were classified as extra long (>7.50 mm), long (6.61 to 7.50 mm), medium (5.51 to 6.60 mm) and short (<5.50 mm) based on the grain length of the respective genotype.

Genomic DNA isolation

The genomic DNA was isolated from young leaf tissues of each genotype 60 days after sowing using the cetyltrimethyl ammonium bromide (CTAB) method. The isolated DNA was tested for quantity and quality on 0.8% agarose gel electrophoresis and NanoDrop-1000 Spectrophotometer (ND-1000). The DNA samples were diluted with nuclease-free water to a working concentration of 100 ng/µl.

Genotyping

The rice genotypes were screened using eight known gene-specific markers for grain size. The grain size related genes included in the present study were *GS*3, *GS*5, *GLW*7, *GW*5/*qSW*5, *qGRL7.1*, *GS*2 and *qsgw7* (Table 1). PCR reaction was carried out in a 10 µL reaction volume containing 1X buffer, 0.5 µL of 1mM dNTPs, 5 pico moles of each forward and reverses primers, 0.1 µL of Taq DNA polymerase (5U/uL), 5.4 µL of sterile, nuclease free Millipore water. The PCR conditions were set up as follows: 94°C for 5 min for initial

denaturation followed by 34 cycles of 94°C for 45 s, primers annealing for 45 s set at varied temperatures (55-60°C) and elongation for 1 min at 72°C, followed by a final elongation at 72°C for 10 min. The PCR products were analyzed by electrophoresis on 4% agarose or 3% metaphore agarose gels stained with ethidium bromide together with a 50 bp DNA ladder. After electrophoresis, the gels were documented using a gel documentation system (Alpha Imager, USA). The PCR reactions were repeated for the markers that were not amplified at least two times to cross-check the scoring data.

Table 1. Gene specific markers used in this study for grain size traits

S. No	Gene	Primer	Primer sequence (5'-3')	Type	Reference
1	*GS*5	GS5INDEL1	F-CTAACTCCCATGGAATTACTAG R-GGAAAGCGAAACTGATTGACA	Indel	Li et al 2011
2	*GS*3	GS3RGS1	F-TCCACCTGCAGATTTCTTCC R-GCTGGTCTTGCACATCTCTCT	Indel	Wang et al 2011
3	*GLW*7	GLW7	F-CCCGCCTTTATATCCCTTTC R-CTGGAGGAGGGTGGAGAGAG	Indel	Si et al 2016
4	*GW*5/qSW5	RMw513	F-GTATTTGTTTGTCGCATTC R-TAGGACCATAGATGTGAGTTA	SSR	Wan et al. 2008
5	*qGRL*7.1	RM505	F-AGAGTTATGAGCCGGGTGTG R-GATTTGGCGATCTTAGCAGC	SSR	Anand et al 2012
6	*GS*2	RM3212	F-GACCCGAGCTAGACGACAAACACC R-ACGCAGCCTCGCACTGTACTCG	SSR	Zhang et al 2013
7	*qsgw*7	RM22015	F-TTTGGTCTGGACCATCATAAGG R-TGACTTTCTCAAGGACATCCTACC	SSR	Bian et al 2013
8	*SLG*7	RM21945	F-CTACACAAGTGAACGCCATCAGG R-GTTCTAGGGTGTCCTTTCATGAGC	SSR	Zhao et al 2015

Statistical analysis

The markers were scored as present (1) or absent (0) to generate the binary matrix for each individual and used to infer the assessment of genetic distance. An unweighted neighbor joining tree was constructed using NTSYS pc -2.02 software.

The association between selected markers for the grain size genes (GL, GW, GLWR and TGW) was analyzed using the general linear model (GLM) in TASSEL3 software (BRADBURY *et al*., 2007). To study the presence of genetic structure for grain size, population structure analysis was performed using the program STRUCTURE version 2.3.4 (PRITCHARD *et al*., 2000). The model was run based on an admixture model with correlated allele frequencies and the number of sub- groups (K) in the clusters was determined by simulating different K-values (K=1 to 10) with 5 independent runs and a run length of 100,000 burn-in period and 100,000 MCMC. The optimal K-value was determined through the ΔK method using Structure Harvester ver.0.6.193 application (EARL, 2012).

RESULTS AND DISCUSSION

Rice yield is often determined by three important agronomic traits viz., grain number per panicle, number of panicles and grain weight. Among them, grain weight is governed by grain size and has drawn great attention in recent days due to its importance as yield as well as quality trait; hence, it is a key trait for both farmers and consumers. Like any other yield component traits, grain size is also governed by innumerable genes/QTLs located on different chromosomes. The key genes among them are *GS*3, *qSW*5, *GS*5, *GW*2, *GW*8, *TGW*6 and *GLW*7. Although a number of QTL/genes were cloned and characterized for grain size, the effects of different allelic combinations from different genes to determine the final grain shape and size are still unclear. Uncovering of various alleles of these genes and their effect in different genetic backgrounds is the need for the hour for development of high-yielding varieties suitable for diverse consumer preferences. Therefore, in the present investigation, we aimed to identify novel alleles from targeted grain-size related genes in diverse rice genotypes.

Phenotypic variations and correlation coefficient analysis for grain size traits

The phenotypic variation in grain-size traits was determined in 124 rice genotypes. The rice genotypes included in the present study exhibited substantial variation in grain-size traits (Fig.1). The phenotypic data for grain length (GL), grain width (GW), length-width ratio (LWR) and 1000-grain weight (TGW) are given in Supplementary Table 1. Estimates for range, mean, standard deviation and coefficient of variation for the rice genotypes are provided in Table 2. The mean of the grain length (GL) ranged from 5.6 mm (Shobini) to 12.05 mm (Pusa1121), grain width (GW) ranged from 1.8 mm (WGL11427) to 3.1 mm (MGD101, MGD103 and Leoit), 1000-grain weight (TGW) ranged from 8 g (Sona masoori) to 31.2 g (Koshihikari) and LWR ranged from 2.38 (Leoit) to 5.35 (Pusa 1121). All four traits *i.e*., grain length, grain width, grain weight and length to width ratio exhibited normal distribution in all the rice genotypes, indicating quantitative inheritance of these traits (Fig. 2).

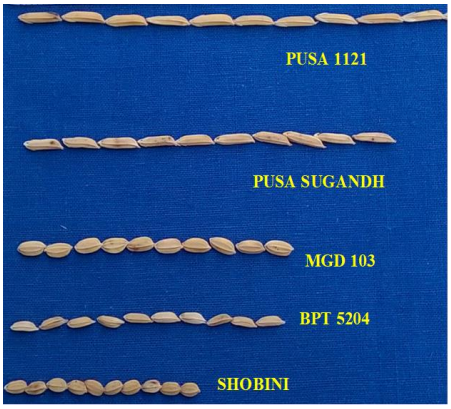

Figure 1 Variation in grain length in representative rice germplasm

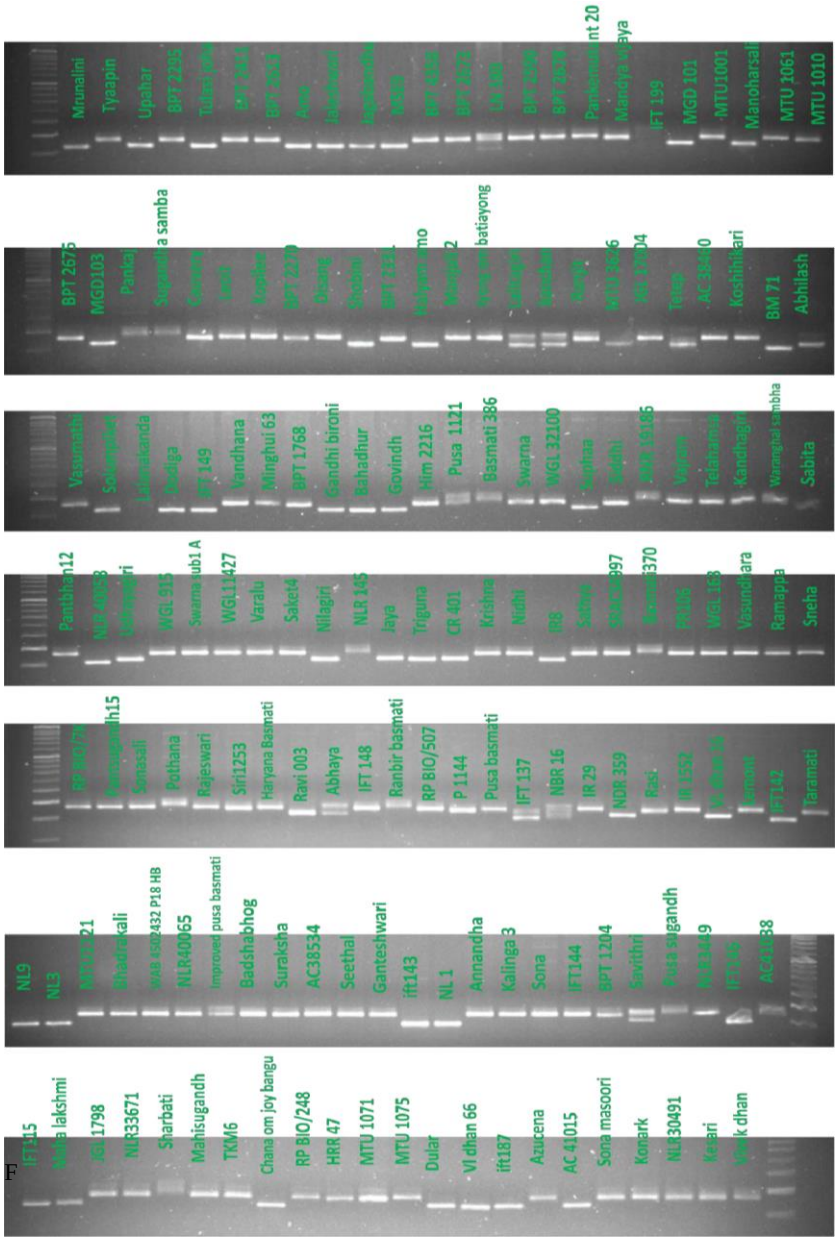

Fig. 2. Variation in alleles in rice cultivars using RMw 513 primer

Table 2. Descriptive statistics for grain size traits in rice genotypes.

Trait	Range	Mean	Std. deviation	Coefficient of variation (%)
TGW(g)	8.0-31.8	19.20	3.95	20.57
GL(mm)	5.6-12.05	8.57	0.98	11.44
GW(mm)	1.8-3.1	2.35	0.31	13.43
GLWR	2.38-5.35	3.71	0.65	17.57

TGW-Thousand grain weight, GL-Grain length, GW-Grain width, GLWR-Grain length-width ratio

The correlation coefficients were estimated for pairwise analysis among the GL, GW, GLWR and TGW traits (Table 3) using 124 diverse genotypes. The traits grain length and grain width showed a negative correlation, though non-significant. Such negative correlation has also been reported by the previous studies among the grain size traits (KOUTROUBAS *et al*., 2004; DALLA *et al*., 2010; KITAGAWA *et al*., 2010). However, the other traits GW and TGW showed strong and highly significant positive correlation (0.367^{**}). A highly significant and positive correlation was observed between GL and GLWR ($r = 0.698^{**}$), GL and TGW ($r = 0.420^{**}$) while negative correlation was observed between GW and GLWR ($r = -0.757^{**}$). These results suggest that there is a strong association among the four grain traits that play important role in determining the final rice grain size.

Table 3. Correlation coefficients among four grain size component traits in 124 rice genotypes

	Grain weight	Grain length	Grain width	Length/width ratio
Grain weight	1.000	-	-	-
Grain length	0.420^{**}	1.000	-	-
Grain width	0.367^{**}	-0.078^{NS}	1.000	-
Length/width ratio	0.010^{NS}	0.698^{**}	-0.757^{**}	1.000

**Significant at 0.01 NS- nonsignificant

Allele mining of grain-size related traits in rice germplasm

In total, 32 molecular markers comprising of 8 gene-specific markers for grain size and 24 SSR markers covering all 12 chromosomes were used in this study and all markers showed polymorphism and produced a total of 86 alleles among the 124 rice varieties (Fig.3 and Table 4). All 32 molecular markers used in this study were polymorphic and they produced a total of 86 alleles among the 124 rice varieties. Number of alleles ranged from 2 to 4 with an average of 2.58 alleles per locus. The mean polymorphism information content (PIC) value was 0.34, with values ranging from 0.70 (RM 252) to 0.03 (RM 502) (Table 4). Eleven markers were highly informative (PIC>0.5), seven markers were moderately informative (0.5 > PIC > 0.25) and fourteen markers were slightly informative (PIC < 0.25). DANG *et al*. (2015) reported an average PIC value of 0.70, with values ranging from 0.0736 to 0.9394, using 258 SSR markers in 532 rice accessions collected from East and Southeast Asia. EDZESI *et al*. (2016) in their study reported an average PIC value of 0.7365, with values ranging from 0.2425 to 0.9373, using 262 SSR markers in 628 rice accessions.

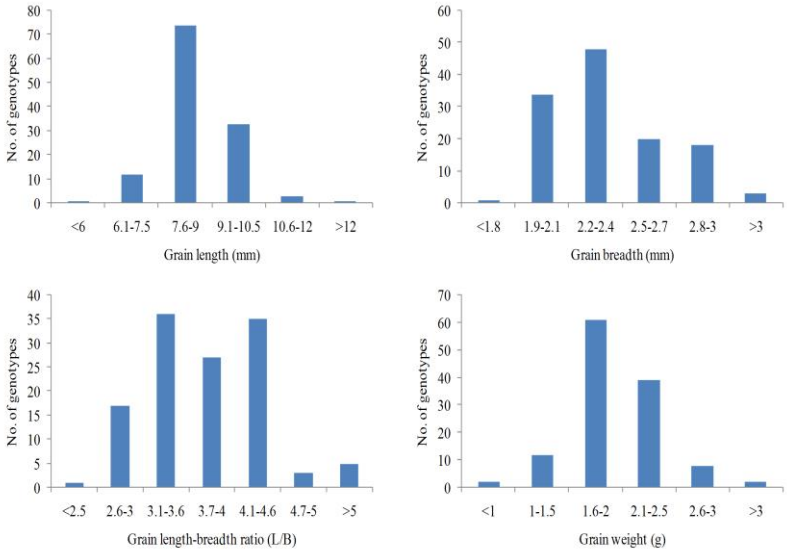

Fig. 3. Frequency distribution of grain size traits in rice genotypes

Table 4. SSR markers used for population structure analysis

S.No	Primer	Chromosome	No. of alleles	PIC
1	RM580	1	2	0.032
2	RM431	1	2	0.078
3	RM555	2	3	0.545
4	RM240	2	4	0.701
5	RM523	3	2	0.063
6	RM520	3	2	0.473
7	RM551	4	3	0.334
8	RM252	4	4	0.705
9	RM574	5	2	0.123
10	RM440	5	3	0.601
11	RM589	6	3	0.555
12	RM340	6	4	0.702
13	RM533	7	2	0.516
14	RM505	7	2	0.107
15	RM502	8	2	0.032
16	RM447	8	2	0.531
17	RM316	9	2	0.191
18	RM464	9	3	0.334
19	RM258	10	2	0.501
20	RM228	10	2	0.474
21	RM552	11	3	0.244
22	RM254	11	2	0.093
23	RM511	12	2	0.047
24	RM313	12	4	0.665

PIC- Polymorphism information content

In order to discover the novel alleles of the grain size genes from 124 rice genotypes, the reported gene-specific primers have been chosen. For each gene the allele number was given from low molecular weight to high molecular weight. The gene wise results of novel alleles have been discussed hereunder.

***GS*3:** The *GS*3 gene on chromosome 3 has been reported as the most important gene controlling grain size (grain length), which encodes a protein with a putative phosphatidyl ethanol amine-binding protein (PEBP)-like domain, a transmembrane domain, a putative tumor necrosis factor receptor (TNFR)/ nerve growth factor receptor (NGFR) family domain and von Willebrand factor type C (VWFC) module (FAN *et al*., 2006).

Screening rice genotypes with *GS*3-specific marker revealed three alleles, with a considerable variation. Out of all tested rice genotypes, 53 were detected with Allele 1, of which 4 were extra- long grains, 10 were long, 31 were medium long and 8 were short grain length, which clearly suggest that this allele is common to medium long grain varieties. The Allele 2 was observed in 74 genotypes, out of which 15 were extra-long, 35 were long, 21 were medium long and 3 were short length grains. The Allele 3 was observed in only four rice genotypes (Supplementary Table 2). TAKANO-KAI *et al*. (2013) evaluated the seed size of 282 diverse rice strains and identified three novel GS3 alleles with independent deletions that all map to the fifth exon of the gene (320-bp, 13-bp and 1 + 3-bp deletion). LU *et al*. (2013) also genotyped the *GS*3 specific CAPS marker SF28 and observed functional (*GS*3-C) and nonfunctional (*GS*3-A) GS3 alleles in 43 and 36 varieties of *indica* population and 23 and 24 varieties of japonica population, respectively.

***GS*5:** The gene *GS5 (Grain Size 5)* positively regulates grain width, grain filling, and weight by promoting cell division and, to a lesser extent, cell elongation of the palea and lemma. Identified by positional cloning, the *GS*5 encodes a putative serine carboxy peptidase and acts as a positive regulator of a subset of the G1-to-S transition genes of the cell cycle (LI *et al*., 2011).

Screening of rice genotypes with *GS*5 specific marker revealed two alleles, with almost equal frequencies. There were 61 rice genotypes with Allele 1, of which 10 were extra-long grains, 19 were long, 25 were medium long and 7 were short grain length indicating that this allele is common in long to medium grain type genotypes. Whereas the Allele 2 was observed in 62 genotypes, of which 9 were extra-long, 25 were long, 25 were medium long and 3 were short length grains, suggesting that this allele is common to long slender grain rice varieties (Supplementary Table 3). These results obviously suggest that these two alleles are largely specific to genotypes with medium to long slender grain type. LI *et al*. (2011) reported that the deletion in *GS*5 gene causes wide grain and also identified two major *GS*5 haplotypes (*GS*5-1 and *GS*5-3), which showed significant differences in grain length and grain width.

Grain width 5/Seed width 5 (*GW*5/*qSW*5): The *qSW*5 gene is also considered as major gene controlling grain width and minor grain length. The gene *GW*5/*qSW*5 encodes a novel nuclear protein that is localized in the nucleus and physically interacts with polyubiquitin, suggesting that *GW*5 acts in the proteasomal degradation pathway to regulate grain size. It negatively regulates grain width (SHOMURA *et al*., 2008; WENG *et al*., 2008). The deletion in the

*GW*5 gene has been reported to play an important role in the selection of increasing grain size from artificial and natural crossing during rice domestication (SHOMURA *et al*., 2008).

Screening of rice genotypes with *qSW*5 specific marker revealed four alleles, with a substantial variation in the number of genotypes. The rice number of genotypes with Allele 1 is 22, of which 3 were extra- long grains, 8 were long, 11 were medium long and no short grain length genotypes. The genotypes with Allele 2 were 61, of which 8 were extra-long, 20 were long, 28 were medium long and 5 were short length grains. The genotypes with Allele 3 were 34, out of which 6 were extra-long, 13 were long, 10 were medium long and 5 were short length grains. The genotypes with Allele 4 were 6, out of which one was extra-long, three were long, and two were medium long and no short length grains **(**Supplementary Table 4). Out of the six genotypes, all six genotypes were medium slender grains. These results clearly suggest that there are many alleles common to medium slender genotypes. LU *et al*. (2013) reported 77 SNPs and two large indel regions in the *qSW*5 locus when they sequenced 127 varieties.

qGRL7.1: This QTL affects GL, GW, and GLWR. Genome-wide association mapping and haplotype analysis of Wang et al (2017) revealed five candidates predicted underlying this QTL, including Os07g0563700 (IKI3 family protein), Os07g0563800 (a GTPase-activating protein), Os07g0564000 (Conserved hypothetical protein), Os07g0564100 (a UDPglucuronosyl / UDP-glucosyltransferase family protein) and Os07g0564150 (a hypothetical gene).

Screening of the rice genotypes with *qGRL7.1* specific marker, revealed two alleles in all the rice genotypes. The number of rice genotypes with Allele 1 was 119, of which 16 were extra- long grains, 43 were long, 50 were medium long and 10 were short grain length genotypes (Supplementary Table 5). The number of genotypes with Allele 2 was only five, out of which 3 were extra-long, one was long, and one was medium long. In this QTL, Allele 1 was found to be the predominant one in long to medium slender grain varieties, while Allele 2 was the rarely observed in the rice germplasm, indicating that Allele 2 might be a novel allele and recently evolved in Indian rice germplasm.

GS2: It is a dominant gene responsible for grain length and width in rice. The cloning and characterization of the *GS*2 QTL revealed that it encodes *OsGRF4* (Growth-Regulating Factor 4), a regulator of transcription (HU *et al*., 2015). Furthermore, it was reported that a mutation in *GS*2 gene at the OsmiR396c target site resulted in elevated transcript levels of *GS*2 and the accumulation of *GS*2 led to enlarged cell size and increased cell number, which in turn resulted in enhanced grain weight and yield.

Screening of rice genotypes with the *GS*2 linked marker revealed two alleles. Allele 1 was found in 113 rice genotypes, of which 17 were extra- long grains, 42 were long, 45 were medium long and 9 were short grain length genotypes (Supplementary Table 6). The number of genotypes with Allele 2 was 11, of which 2 were extra-long, 2 were long, and 6 were medium long and one short length grains. It appears that the Allele 1 is the most prevalent in the rice genotypes studied while Allele 2 is rare. The sequence analysis of *GS*2 gene in 66 accessions revealed that SNP (TC487-488AA) was only found in the BDL accession and concluded that it was a rare mutation of *GS*2 and might not have been selected by breeders (HU *et al*., 2015).

qsgw7: It a single major QTL affecting the 1000-grain weight of rice was identified on the short arm of chromosome 7 using a segregating population derived from sgw (low grain weight) and cultivar 9311 (indica, high grain weight) (BIAN *et al*., 2013).

Screening of the rice genotypes with *qsgw*7 linked marker revealed two alleles, with a great amount of variation. The number of rice genotypes with Allele 1 was 121, of which 16 were extra- long grains, 44 were long, 51 were medium long and 10 were short grain length genotypes (Supplementary Table 7). The number of genotypes with Allele 2 was only three with extra-long slender grains. Allele 1 appears to be most prevalent in the medium slender rice genotypes, whereas Allele 2 is appears to be a rarest and specific to extra-long grain varieties.

***SLG*7 (Slender grain on chromosome 7):** This gene is responsible for grain shape that encodes a protein homologous to LONGIFOLIA 1 and LONGIFOLIA 2, both of which increase organ length in Arabidopsis. The *SLG*7 is constitutively expressed in various tissues in rice, and the *SLG7* protein is located in the plasma membrane. Morphological and cellular analyses suggested that *SLG*7 produces slender grains by longitudinally increasing cell length, while transversely decreasing cell width, which is independent of cell division (ZHOU *et al*., 2015).

Screening of the rice genotypes with *SLG*7 specific marker unfolded two alleles, with plenty of variation within the germplasm. Allele 1 was present in 119 genotypes, which comprising 15 extra- long grains, 43 long, 51 medium long and 10 short grain length genotypes. The number of genotypes with Allele 2 is 5, of which 4 are extra-long grains and one was a long grain type. All five genotypes belonged to the slender grain type. Allele 1 seems to be the most abundant type of *SLG*7 gene in the genotypes considered in the present study (Supplementary Table 8). The sequence analysis by ZHOU *et al*. (2015) in 50 rice germplasm from a wide geographic range in Asia and America suggested that the *SLG*7 alleles were classified into five types, namely Azucena, 9311, Nipponbare, Guangluai 4, and Dular. Nipponbare and Guangluai, with four haplotypes, were exclusively found in japonica and indica subspecies, respectively, while the other three haplotypes were detected in both subspecies. The Azucena *SLG*7 allele was identified in 13 genotypes from the United States, the Philippines, Indonesia, and southern China, suggesting that SLG7 is not a rare allele and has been collected during rice breeding. In addition, the varieties carrying the Azucena-type *SLG*7 allele were found to be significantly higher than those with the other four *SLG*7 alleles in grain length.

***GLW 7* (Grain length and weight on chromosome 7):** The *GLW*7 is a major QTL producing larger grains and panicles and eventually improves grain yield in cultivated rice. *GLW*7 encodes the plant-specific transcription factor *Os*SPL13 and positively regulates cell size in the grain hull, resulting in enhanced rice grain length and yield (SI *et al*., 2016).

Screening of the rice genotypes with the *GLW7* specific marker revealed two alleles, with a substantial variation in the genotypes studied. Out of all, 110 genotypes had Allele 1, of them 17 were extra- long grains, 43 were long, 48 were medium long and 2 were short grain length genotypes (Supplementary Table 9). The genotypes with Allele 2 were 15, of which 4 were extra-long grains and 5 were long grain type, 4 were medium length and 2 were short length grains. Allele 1 was appears to be the most predominant one in the long slender genotypes studied. Sequencing of the *GLW*7 gene in 26 small-grain and 21 large-grain japonica varieties

revealed that 16 of 29 polymorphisms were tightly associated with the lead SNP (rs19784266 on chromosome 7): 6 SNPs in the promoter region, 3 polymorphisms in the 5′ UTR (g.19763436G>A (C>T variant at position −172), g.19763399GAAGTG[1] (one copy of the repeat sequence beginning at −135), g.19763279G>C (C>G variant at position −15)), 1 synonymous polymorphism in exon 1, 2 SNPs in the intron, and 1 SNP and 3 indels in the 3′ UTR (SI *et al*., 2016).

The plausible reason for the appearance of the common alleles in majority of the rice genotypes could be due to the severe selection pressure imposed whereas the appearance of novel or rare alleles is due to their recent evolutionary past or least preferred during the domestication and crop improvement. Even though, grain size types, long slender, medium and short varieties are consumed in India, the most preferred type for the North Indians is long slender while it is medium slender in the South India. In the present study, none of the alleles of the targeted genes found to be unique to different types of grain sizes, which clearly reinforces that fact that the genetics of grain size is much more complex than we assumed. Hence, it is warranted to uncover more QTLs/genes governing grain size for complete understanding of the trait.

Allelic diversity of rice genotypes

A dendrogram consisting of 124 rice genotypes was drawn using the unweighted pair group method using arithmetic averages (UPGMA) based on genotyping data using NTSYS pc - 2.02 software (Fig. 4). The dendrogram revealed that the total of 124 genotypes could be divided into two groups: A and B. Group A included the extra-long grain length basmati genotypes. Previously, many reports also grouped the basmati varieties as a separate group (GLASZMANN, 1987; NARASHIMULU *et al*., 2011; NAGARAJU *et al*., 2002). Group B consisted of all classes of grain size. However, group B could again be divided into two groups *i.e.,* B1 and B2. Group B1 includes mostly long grain genotypes. The B2 group consisted of classes of genotypes of all grain lengths and sizes. However, falling of different classes of grain size groups into a single group suggests that the gene-specific markers used in the present study are not sufficient enough to classify all the genotypes as per their trait. These results also reinforce the fact that grain size is one of the complex traits governed by many genes.

Population structure analysis

An analysis of the model-based population structure using SSRs covering all 12 chromosomes provided evidence of a significant population structure in the 124 rice accessions and identified the highest likelihood value at K= 4 for all five replicates (five runs for each K). All the rice genotypes could be divided into four clusters/subpopulations, *viz.*, POP1 to POP4 (Figure 5).

The first subpopulation, POP1 included a total of 47 genotypes, of which 20 were extra-long grains, 10 were long grain, 16 were medium long and only one genotype had short grain length. From this, it was revealed that the POP1 subpopulation was grouped under extra-long grain type. The second subpopulation, POP2 included a total of 25 genotypes, of which 8 were extra-long, 14 were long and 3 were medium long. From this, it was revealed that POP2 was grouped under long grain type. The third subpopulation, POP3 includes a total of 24 genotypes,

of which 10 were extra-long, 9 were long, 4 were medium and 1 was short. The fourth subpopulation, POP4 includes a total of 28 genotypes, of which 8 were extra-long, 7 were long, 12 were medium and 1 was short.

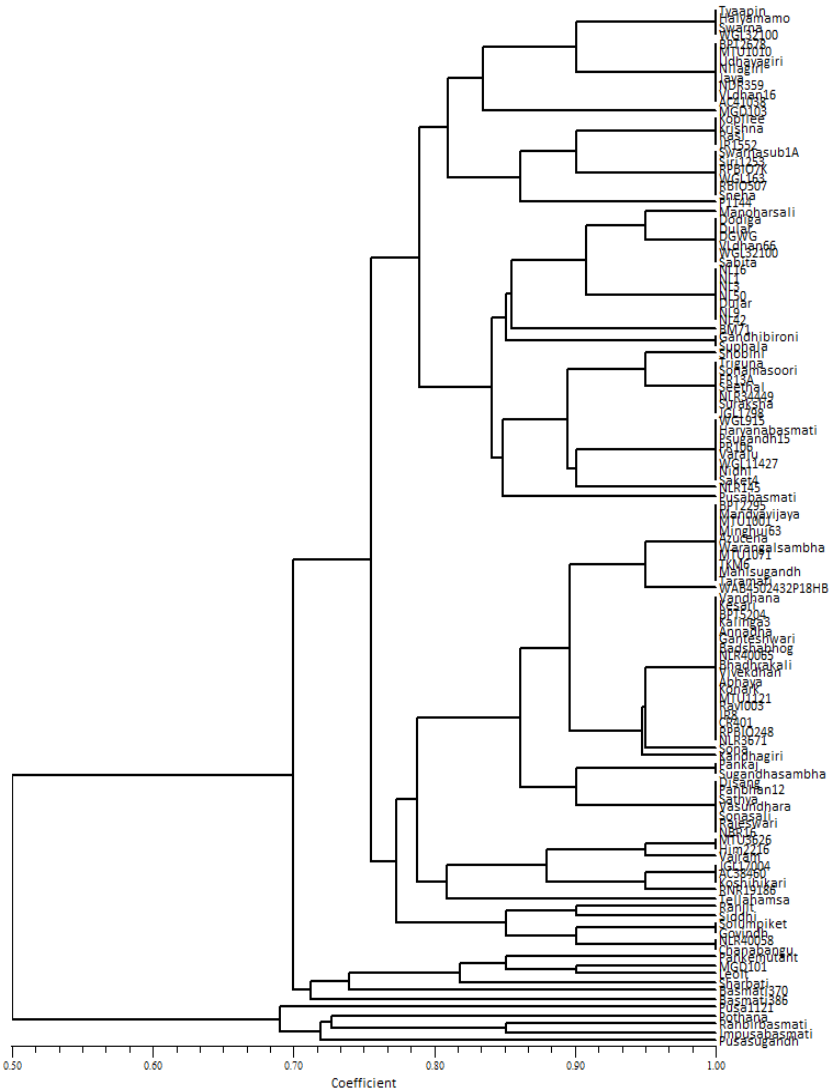

Fig.4 Dendrogram of 124 rice genotypes drawn based on genotyping data of gene-specific markers

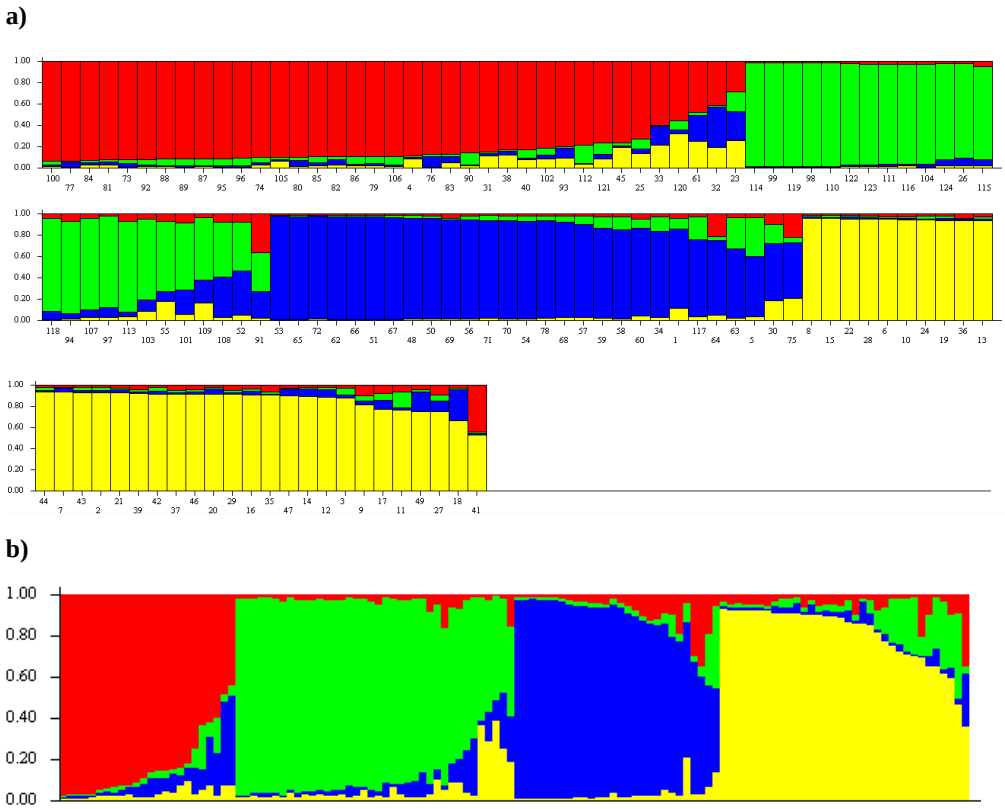

Fig. 5. a) Population structure of all 124 rice genotypes defined by STURCTURE software using 24 SSR markers. Each genotype is represented by a vertical bar. b) Identified subpopulations are POP1 (red colour), POP2 (green colour), POP3 (blue colour) and POP4 (yellow colour).

CONCLUSION

Elucidation of the complete mechanism of the grain size traits is imperative to develop consumer-targeted rice varieties. The present investigation sheds more light on finding the types of allele existed in different rice genotypes for grain size traits. The alleles identified could be useful to maneuver the grain size trait for consumer's interest and also to increase the grain yield. The gene-specific markers used in the study can also have the potential to be used as foreground markers in marker-assisted breeding programs. The markers that are associated with grain size are highly informative and can be used in the selection of parental lines for the development of new rice breeding populations.

ACKNOWLEDGMENTS

This research was supported by the Department of Biotechnology (DBT), Govt. of India (Sanction Order: No. BT/PR5493/AGIII/103/848/2012 Date: 08.02.2016) awarded to LRV. We

thank ANGRAU for providing research lab facilities and Bayer crop science for providing fellowship to SNB.

FINANCIAL SUPPORT

Department of Biotechnology (DBT), Govt. of India (Sanction Order: No. BT/PR5493/AGIII/103/848/2012 Date: 08.02.2016).

Received, December 09th, 2019
Accepted July 18th, 2020

REFERENCES

BIAN, H.H., C.J., HE, H., LI, C.L., SHI, X.S., ZHU, J.R., PENG, X.P., FU, X.R., HE, L.F., CHEN, L.J., HU OUYANG (2013): Identification and validation of a new grain weight QTL in rice. Genetics and Molecular Research, *12* (4): 5623-5633.

BRADBURY, P.J., Z., ZHANG, D.E., KROON, T.M., CASSTEVENS, Y., RAMDOSSM, E.S., BUCKLER (2007): TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics, *23*(19): 2633–2635.

DALLA CORTE, A., V., MODA-CIRINO, C.A., ARIAS, J.F., TOLEDO, D., DESTRO (2010): Genetic analysis of seed morphological traits and its correlations with Grain yield in common bean. Braz. Arch. Biol. Technol., *53*(1): 27–34.

EARL, D.A. (2012): STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources, *4*(2): 359–61.

EDZESI, W.M., X., DANG, L., LIANG, E., LIU, D., HONG (2016): Genetic diversity and elite allele mining for grain traits in rice (O. sativa L.) by association mapping. Frontiers in Plant Science, *7*:787.

FAN, C., Y., XING, H., MAO, T., LU, B., HAN, C., XU, X.H., LI, Q., ZHANG (2006): GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. TAG, *112*(6): 1164–71.

GLASZMANN, J.C. (1987): Isozymes and classification of Asian rice varieties. TAG, *42*: 233–240.

HU, J., Y., WANG, Y., FANG, X., ZENG, X., Y., JIE HAIPING, S., ZHENYUAN, P., JIANGJIE, Z., DONG, K., SHUJING, Z., LI, D., GUOJUN, G., LONGBIAO, Z., DALI, Z., GUANGHENG, X., LIHONG, X., GUOSHENG, L., JIAYANG, Q., QIAN (2015): A Rare Allele of *GS*2 Enhances Grain Size and Grain Yield in Rice. Mol. Plant., *8* (10):1455-1465.

ISHIMARU, K., N., HIROTSU, Y., MADOKA, N., MURAKAMI, N., HARA, H., ONODERA (2013): Loss of function of the IAA-glucose hydrolase gene *TGW*6 enhances rice grain weight and increases yield. Nature Genetics, *45*(6): 707–11.

JULIANO, B. and C.P., VILLAREAL (1993): *Grain quality evaluation of world rice*. International Rice Research Institute, Manila, Philippines.

KITAGAWA, K., S., KURINAMI, K., OKI, Y., ABE, T., ANDO, I., KONO (2010): A novel kinesin 13 protein regulating rice seed length. Plant and Cell Physiology, *51*(8): 1315–29.

KOUTROUBAS, S.D., F., MAZZINI, B., PONS, D.A., NTANOS (2004): Grain quality variation and relationships with morpho-physiological traits in rice (*Oryza sativa* L.) genetic resources in Europe. Field Crops Research, *86*(2): 115–30.

LI, Y., C., FAN, Y., XING, Y., JIANG, L., LUO, L., SUN, D., SHAO, C., XU, X., LI, J., XIAO, Y., HE, Q., ZHANG (2011): Natural variation in *GS*5 plays an important role in regulating grain size and yield in rice. Nature Genetics, *43*:1266–1269.

LU, L., D., SHAO, X., QIU, L., SUN, W., YAN, X., ZHOU, L.,YANG, Y., HE, S., YU, Y., XING (2013): Natural variation and artificial selection in four genes determine grain shape in rice. The New phytologist, *200*(4):1269-80.

MAO, H., S., SUN, J., YAO, C., WANG, S., YU, C., XU (2010): Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. Proc. Nat. Acad. Sci.USA, *107*(45): 19579–84.

NAGARAJU, J., M., KATHIRVEL, R., RAMESH KUMAR, E.A., SIDDIQ, S.E., HASNAIN (2002): Genetic analysis of traditional and evolved Basmati and non-Basmati rice varieties by using fluorescence-based ISSR-PCR and SSR markers. Proc. Nat. Acad. Sci. USA, *99*: 5836-5841.

NARSHIMULU, G., M., JAMALODDIN, L.R.,VEMIREDDY, G., ANURADHA, E.A., SIDDIQ (2011): Potentiality of evenly distributed hypervariable microsatellite markers in marker-assisted breeding of rice. Plant Breeding, *130*(3): 314-320.

PRITCHARD, J.K., M., STEPHENS, P., DONNELLY (2000): Inference of population structure using multi locus genotype data. Genetics, *155*(2): 945–959.

QI, P., Y.S., LIN, X.J., SONG, J.B., SHEN, W., HUANG, J.X., SHAN (2012): The novel quantitative trait locus GL3. 1 controls rice grain size and yield by regulating CyclinT1; 3. Cell Response, *22*(12): 1666.

SHOMURA, A., T., IZAWA, K., EBANA, T., EBITANI, H., KANEGAE, S., KONISHI, M., YANO (2008): Deletion in a gene associated with grain size increased yields during rice domestication. Nature Genetics, *40*:1023–1028.

SI, L., J., CHEN, X., HUANG, H., GONG, J, LUO, Q., HOU, T., ZHOU, T., LU, J., ZHU, Y., SHANGGUAN, E., CHEN, C., GONG, Q., ZHAO, Y., JING, Y., ZHAO, Y., LI, L., CUI, D., FAN, Y., LU, Q., WENG, Y., WANG, Q., ZHAN, K., LIU, X., WEI, K., AN, G., AN, B., HAN (2016): *Os*SPL13 controls grain size in cultivated rice. Nature genetics, *48*(4):447-56.

TAKANO-KAI, N., H., JIANG, A., POWELL, S., MCCOUCH, I., TAKAMURE (2013): Multiple and independent origins of short seeded alleles of *GS*3 in rice. Breeding Science, *63*:77–85.

UNNEVEHR, L.J,. B., DUFF, B.O., JULIANO (1992): Consumer demand for rice grain quality. International Rice Research Institute, Manila, Philippines and International Development Research Centre, Ottawa, Canada.

WANG, E., J., WANG, X., ZHU, W., HAO, L., WANG, Q., LI (2008): Control of rice grain-filling and yield by a gene with a potential signature of domestication. Nat. Genet., *40*(11): 1370–4.

WANG, C., S., CHEN, S., YU (2011): Functional markers developed from multiple loci in GS3 for fine marker-assisted selection of grain length in rice. TAG, *122*(5): 905–13.

WANG, S., K., WU, Q.,YUAN, X., LIU, Z., LIU, X., LIN (2012): Control of grain size, shape and quality by OsSPL16 in rice. Nat Genet., *44*(8): 950-4.

WENG, J., S., GU, X., WAN, H., GAO, T., GUO, N., SU, C., LEI, X., ZHANG, Z., CHENG, X., GUO, J., WANG, L., JIANG, H., ZHAI, J., WAN (2008): Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. Cell Research, *18*:1199-1209.

XIAN-JUN, S., W., HUANG, M., SHI, M.Z., ZHU, L., HONG-XUAN (2007): A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nature Genetics, *39*(5): 623-30.

XU, F., L., JIN, Y., HUANG, C., TONG, Y., CHEN, J., BAO (2016): Association mapping of quantitative trait loci for yield-related agronomic traits in rice (*Oryza sativa* L.). Journal of Integrative Agriculture, *15*(10): 2192-2202.

ZHANG, P.S., H., QIANG, S., FU, W., JIE, D., HUAFENG (2013): Fine mapping of *GS*2, a dominant gene for big grain rice. The Crop J., *1*: 160-165.

ZHOU, J., M., JUN, G., HAIYONG, P., XIURONG, L., MAMOTSHEWA, Y., FUHAI, G., HOUWEN, G., YUN, T., YAJUN, Z., JINYAN, G., ZHIYUN, Y., CHUANDENG, G., MINGHONG, Y., ZEFENG, L., GUOHUA (2015) Natural Variations in *SLG7* Regulate Grain Shape in Rice. Genetics, *201*:1591-1599.

ZUO, J. and J., LI (2014): Molecular genetic dissection of quantitative trait loci regulating rice grain size. Annual Review of Genetics, *48*: 99–118.

GENETIČKI DIVERZITET, STRUKTURA POPULACIJE I GENSKI ALELI KOJI SU POVEZANI SA OSOBINAMA VELIČINE ZRNA KOD PIRINČA (*Oryza sativa* L.)

Swarajyalakshmi N. BOLLINENI[1], Bukya J. NAIK[1], Suresh E. NAIK[1], Gopalakrishnamurthy KADAMBARI[1], G. Eswar REDDY[1], S. MALATHI[6], Eswar N.P. REDDY[1], Srividhya AKKAREDDY[2], Reddi M. SEKHAR[3], Bahben TANTI[4], B. Krishna VENI[5], B. Vijaya LAKSHMI[5] and Lakshminarayana R. VEMIREDDY[3*]

[1]Departman za molekularnu biologiju i biotehnologiju, S.V. Poljoprivredni koledž, Acharya NG Ranga Poljoprivredni Univerzitet (ANGRAU), Tirupati-517502, Indija
[2]Regionalna poljoprivredna istraživačka stanica, ANGRAU, Tirupati-517502, Indija
[3]Departman za genetiku i oplemenjivanje biljaka, S.V. Poljoprivredni koledž, Acharya NG Ranga Poljoprivredni Univerzitet (ANGRAU), Tirupati-517502, Indija
[4] Departman za botaniku, Gauhati Univerzitet, Guwahati, Asam, Indija
[5]Istraživačka jedinica za pirinač, ANGRAU, Bapatla-522101, Indija
[6]Indijski Institute za istraživanje pirinča, Hyderabad-500030, Indija

Izvod

Ova studija je sprovedena za traženje alela gena koji regulišu dimenzije zrna, odnosno dužinu zrna, širinu zrna, debljinu zrna, veličinu zrna i težinu zrna. Molekularni markeri povezani sa prijavljenim genima za ove osobine korišćeni su za pregled 124 različitih genotipova pirinča. Trideset dva molekularna markera korišćena u ovoj studiji proizvela su ukupno 86 alela među 124 genotipa pirinča. Broj alela kretao se od 2 do 4 sa prosekom od 2,58 alela po lokusu. Dendrogram koji se sastojao od 124 genotipa pirinča dobijen je metodom neponderisanih parnih grupa koristeći aritmetički prosek (UPGMA), otkrivajući da se svi genotipovi mogu podeliti u dve grupe. Analiza populacione strukture zasnovane na modelu korišćenjem SSR-ova koja pokrivaju svih 12 hromozoma pružila je dokaze o značajnoj strukturi populacije u genotipovima pirinča. Novi aleli identifikovani u studiji mogli bi biti od velike vrednosti za razvoj sorti pirinča usmerenih ka potrebama potrošača.

Primljeno 09.XII.2019.
Odobreno 22. VII. 2020.