



OPEN

Superior haplotypes towards development of low glycemic index rice with preferred grain and cooking quality

Ramchander Selvaraj^{1,4}, Arun Kumar Singh^{1,4}, Vikas Kumar Singh^{1,4}, Ragavendran Abbai³, Sonali Vijay Habde², Uma Maheshwar Singh² & Arvind Kumar^{1,2}✉

Increasing trends in the occurrence of diabetes underline the need to develop low glycemic index (GI) rice with preferred grain quality. In the current study, a diverse set of 3 K sub-panel of rice consisting of 150 accessions was evaluated for resistant starch and predicted glycemic index, including nine other quality traits under transplanted situation. Significant variations were noticed among the accessions for the traits evaluated. Trait associations had shown that amylose content possess significant positive and negative association with resistant starch and predicted glycemic index. Genome-wide association studies with 500 K SNPs based on MLM model resulted in a total of 41 marker-trait associations (MTAs), which were further confirmed and validated with mrMLM multi-locus model. We have also determined the allelic effect of identified MTAs for 11 targeted traits and found favorable SNPs for 8 traits. A total of 11 genes were selected for haplo-pheno analysis to identify the superior haplotypes for the target traits where haplotypes ranges from 2 (*Os10g0469000-GC*) to 15 (*Os06g18720-AC*). Superior haplotypes for RS and PGI, the candidate gene *Os06g11100* (H4-3.28% for high RS) and *Os08g12590* (H13-62.52 as intermediate PGI). The identified superior donors possessing superior haplotype combinations may be utilized in Haplotype-based breeding to developing next-generation tailor-made high quality healthier rice varieties suiting consumer preference and market demand.

Abbreviations

AC	Amylose content
KL	Kernel length
KB	Kernel breadth
LBR	Length/ Breadth ratio
ASV	Alkali spreading value
KLAC	Kernel length after cooking
KBAC	Kernel breadth after cooking
LER	Linear elongation ratio
GC	Gel consistency
RS	Resistant starch
PGI	Predicted glycemic index
TPR	Transplanted rice
MTAs	Marker trait associations
GWAS	Genome wide association studies

Rice (*Oryza sativa* L.) is a major source for carbohydrate of more than 50% of the global population. In recent times, an increase in living standards has created awareness among the peoples all over the world for consumption of superior quality rice to improve their health status¹. Thus, the presence of premium grain quality in rice

¹IRRI South Asia Hub (IRRI-SAH), ICRISAT Campus, Patancheru, Hyderabad, India. ²South-Asia Regional Centre (SARC), International Rice Research Institute (IRRI), Varanasi, India. ³Leibniz Institute of Plant Genetics and Crop Plant Research (IPK), Gatersleben, Germany. ⁴These authors contributed equally: Ramchander Selvaraj, Arun Kumar Singh and Vikas Kumar Singh. ✉email: arvind.kumar@cgiar.org

is one of the crucial determinant factors for all the stakeholders including breeders, producers and consumers². Currently, more efforts have been put forth to improve the rice quality traits by utilizing the key allelic variants from the germplasm variation to meet quality rice demand of consumers and potential markets. But, what are the key grain quality determinants? They are broadly classified in to, milling, appearance, cooking/eating parameters and nutritional quality. Among these, the later two are associated with the end-users. The physicochemical characters which determine the cooking and eating quality are amylose content (AC), gel consistency (GC) and gelatinization temperature (GT) or alkali spreading value (ASV). Generally, rice with intermediate amylose, GT and soft to medium gel consistency are preferred by consumers in South Asia³. Apart from quality, these days people have more concerns about their health, and prefer low glycemic index rice varieties which could be safe for diabetics and obesity. Hence, grain quality improvement of rice together with the lower glycemic index are of vital importance for the rice breeders in the current scenario of increasing diabetic population all over the world, especially in the Asia. Low GI diets effectively prevent type II diabetes^{4,5} and consumption of high amylose class rice varieties with soft-textured on cooking might be an alternative for intermediate amylose rice varieties especially for those who are suffering from type II diabetes⁶. The most critical factor responsible for breeding low glycemic index (GI) rice is resistant starch (RS). It is the starch portion resistant to enzyme hydrolysis that escapes the small intestine and enters the large intestine where it gets fermented and slowly releases glucose to the bloodstream. Generally, it has been reported that increased level of RS content in rice grains lowers the GI value ensuring a negative correlation with amylose content⁷. Recently low-to-intermediate RS phenotypic variations in rice panel were identified, resulting novel RS associations to numerous genes associated with amylopectin biosynthesis and degradation⁷.

In any crop species, genetic diversity plays important role in breeding elite varieties. Identification of favorable alleles⁸ and its superior haplotype of the various genes associated with traits linking to cooking and eating quality of rice are pre-requisite of breeding to develop healthier rice. Nowadays, with the improvement of high-throughput sequencing technologies with reduced cost makes genome wide association studies (GWAS) as one of the prominent techniques to identify marker-trait associations (MTAs). Research on the glycemic index identified a novel association of candidate loci *Os05g03600* reporting intermediate to high GI variations⁹. Another hotspot on chromosome 6 was found to include *GBSSI*, hydrolase genes and genes involved in signaling and chromatin modification with differential methylation patterns in *GL6.1* region. Alternative splicing of *GBSSI* promoter region resulted in intermediate to high GI variations. Novel SNP associations on chromosomes namely 2, 5, 6 and 11 has been reported and these SNPs influences the final viscosity variations but no significant association with GI⁹. Besides, the predicted glycemic index (PGI) is estimated by *in-vitro* enzymatic action of starch digestion leads to hydrolysis and the glycemic index (GI) estimated *in-vivo*, which requires human clinical evaluation of the two to three hours of blood glucose response after food intake. Thus, it is more time consuming and resource demanding. The *in-vitro* digestion methods have been developed to measure the starch hydrolysis index and it is used to calculate PGI using the formula developed¹⁰. Significant positive association ($r=0.946$) between pGI and GI of rice samples by using bread as the reference¹¹.

Recently, 11 candidate loci controlling grain quality traits has been identified that are involved in the starch biosynthesis¹². Nine MTAs were identified for seven QTLs namely *GS3*, *TUD*, *qGRL7.1*, *qPGWC7*, *qGL3.4*, *qGW1.1* and *qGW10.2* controlling quality traits¹³. Numerous genes/QTLs were identified for grain number and grain length viz., *GW2*, *GIF1*, *qSW5*, *GS3*, *GS5*, *qGL3*, *GW8*, *GS6*, *GS2*, *GL7/GW7*, *OsMA PK6*, *GLW7* and *GAD1*^{14–26}. QTLs controlling grain size have also been detected via genetic mapping and association studies^{24,27–32}. Starch is made up of two components namely amylose and amylopectin. Amylose, an important parameter affecting cooking and eating quality of rice directly involved in changing the grain texture by absorbing water on cooking. However, the difference in the amylopectin structure of similar AC rice explained variations in the textural qualities of cooking. It is a well-established fact that the waxy (*Wx*) gene, *granule-bound starch synthase* (*GBSSI*) located on chromosome 6 and genomic regions surrounding the gene are highly diverse^{9,33}. The *Alk* locus carrying *SSIIa* (*starch synthase IIa*) mapped on chromosome 6 confers alkali spreading to determine the gelatinization temperature of milled rice³⁴. There are several other starch synthesis-related genes involved namely, starch debranching enzymes, starch branching enzyme soluble starch synthase with a predominant role in controlling eating and cooking properties of rice³⁵. Two QTLs for resistant starch in rice *qRS7-1* and *qRS7-2* (chromosome 7), explaining phenotypic variance from 7.6 to 17.3%, have been reported³⁶.

Rice grain quality is one of the major determinants in selecting parents for any breeding program and the genetic constitution of the genotypes determine the effect of prevailing environmental conditions on the developmental process involved in seed formation and maturation. The good quality grain is an important factor influencing its acceptance by consumers and thus, is one of the major traits in rice breeding to withstand marketability in trade and commerce. However, one of the significant challenges to date has been the lack of extensive knowledge of the genetic and molecular basis of several grain quality traits including GI. The availability of 3000 rice genomes (3 K-RGP) offers opportunities for harnessing haplotype diversity for GI along with other critical quality-related genes and enhances the possibility for the identification of ideal haplotypes for deployment in rice breeding. The current study was undertaken with the objective to identify the candidate genes and their superior haplotypes associated with RS, PGI and nine major grain quality traits across the 3 K-RGP subset under transplanted (TPR) situation (Fig. 1) and explore the possibility of developing healthy rice varieties with preferred grain quality by assembling superior haplotypes via *haplotype-based breeding*.

Results

Trait variation for resistant starch, glycemic index and grain and cooking quality. The subset of 3 K RGP was extensively phenotyped to study the variations for RS, PGI along with nine other grain quality-related traits. The results revealed significant variations among the accessions (Table S1). The RS content ranged

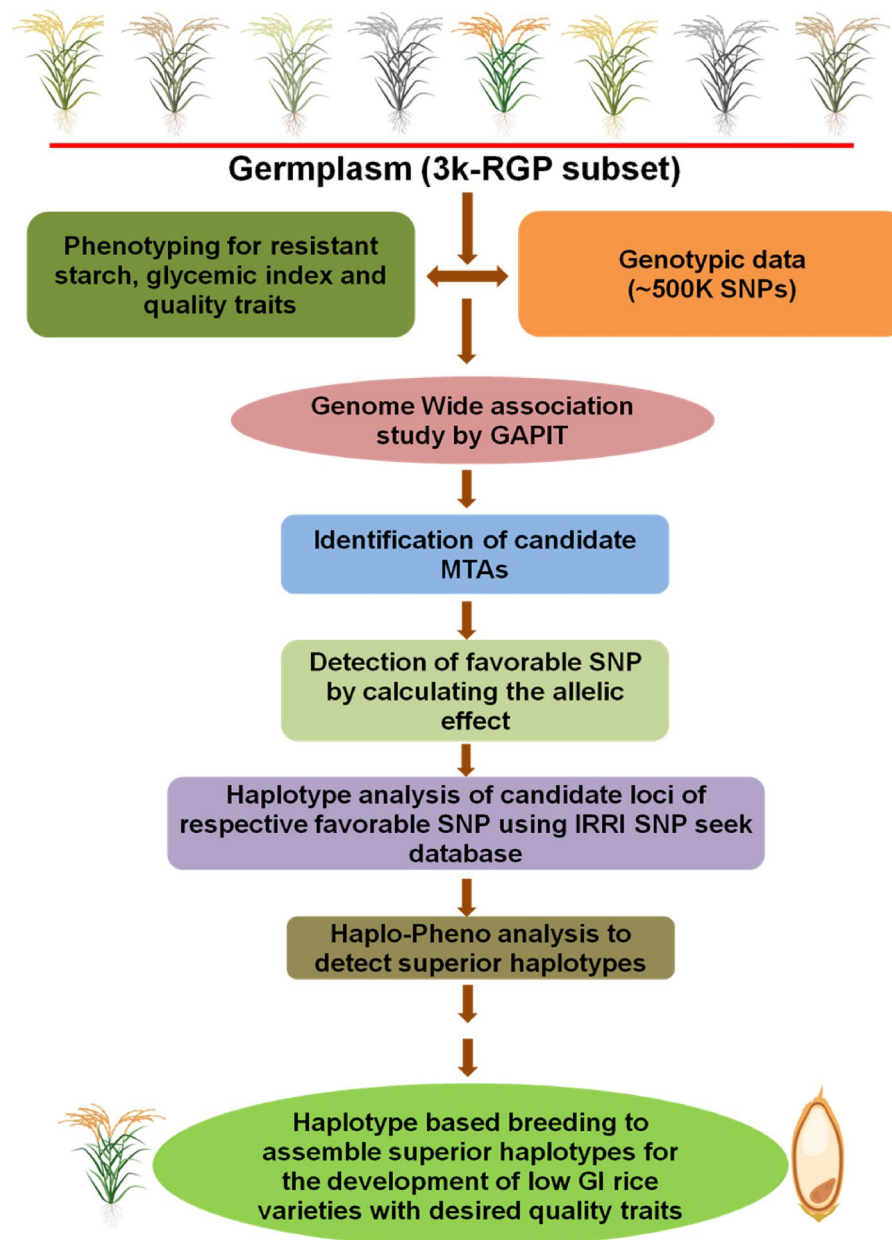


Figure 1. Overall methodology flow of the study to unravel the SNP associations, identifying favorable SNPs, locating the candidate genes and superior haplotype identification in transplanted rice (TPR) by haplotype based breeding to assemble superior haplotypes to design the rice varieties suitable for diabetics with improved grain quality parameters.

from 0.57 to 10.00%, whereas, the PGI ranged from 52.91 to 99.94. Significant variation was also observed for the other quality traits-like AC, LBR, LER, ASV and GC that play an important role in the selection of desirable quality of rice varieties in the view of consumer preference. In our study, AC ranged from 12 to 33%, KL in the range of 4.60 to 7.7 mm, KB ranged from 1.80 to 3.00 mm, LBR in the range of 1.58 to 3.60, ASV from 1 to 7 score, LER ranged between 1.02 to 1.96 and GC in the range of 14.50 to 100 mm.

Correlation among grain quality traits. Association among the traits was observed. RS and PGI showed significant negative correlations. The important quality trait AC showed a significant positive ($r=0.19^*$, p -value-0.048) and negative correlation ($r=-0.16^*$, p -value-0.049) with RS and PGI, respectively whereas; it showed a negative significant relationship with ASV and GC. KL had a significant positive correlation with LBR and KLAC whereas it was negatively correlated with KB (Fig. 2).

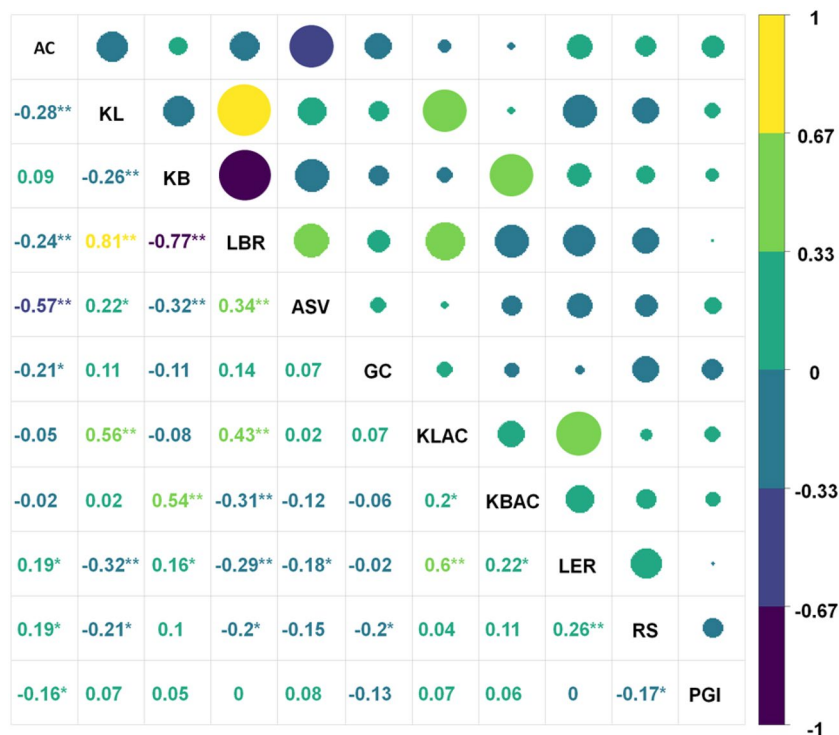


Figure 2. Investigation on correlation coefficients explaining associations among predicted glycemic index and resistant starch with rice quality parameters under transplanted rice system. AC—Amylose content, KL—Kernel length, KB—Kernel breadth, LBR—Length/ Breadth ratio, ASV—Alkali spreading value, KLAC—Kernel length after cooking, KBAC—Kernel breadth after cooking, LER—Linear elongation ratio, GC—Gel consistency, RS—Resistant starch, PGI—Predicted glycemic index, TPR—Transplanted rice. * Significance at p -value < 0.05 , ** Significance at p -value < 0.01 .

Genome-wide association study (GWAS) of targeted traits. Single locus GWAS (MLM) was performed using the phenotypic data with approximately more than 0.5 million (500 K) SNPs of concerned accessions and the peak associations were predicted based on considering the value of PVE (phenotypic variance explained) with suggestive significant p -value < 0.00001 ($-\log_{10}(p) > 5$). In general, modification of bonferroni correction significant threshold value to control the false positive rate in a single-locus GWAS are so conservative that some associated SNPs may be eliminated. To avoid this problem multi-locus GWAS of mrMLM was also conducted to confirm the MTAs with the LOD score of greater than 3. A set of 41 causatives MTAs were detected for all the investigated quality traits in MLM (Table 1). In MLM, single MTA was identified for KL, KBAC, L/B ratio and LER, and up to nine MTAs for AC, RS and PGI. We identified nine MTAs associated with RS located on chromosomes 1, 2, 6, 8 and 11 with PVE ranging between 15.92 to 19.31% (Fig. 3). For PGI, nine MTAs were identified on chromosomes 1, 3, 7, 8, 9 and 11 with PVE ranging between 23.47 to 17.77% (Fig. 3). For AC, nine MTAs were detected on chromosomes 2, 6, 8, 10 and 11, respectively with the PVE ranging between 14.38 to 17.96% (Fig. S1). For KL, one MTA was identified on chromosome 11 with the PVE of 11.87% (Fig. S2). For KB, 4 MTAs were identified on chromosome 2, 3 and 5 with a PVE of 14.48, 14.79, 15.35 and 21.48%, respectively (Fig. S3). LBR recorded an MTA on chromosome 5 with the PVE of 16.99% (Fig. S4). For ASV, an indicator for GT, one peak was detected on chromosome 6 with a PVE of 25.53% (Fig. S5). Three MTAs on chromosome 1, 5 and 12 were detected for KLAC with PVE of 23.55, 24.17 and 18.05%, respectively (Fig. S6). For KBAC only one association was detected on chromosome 8 with the PVE of 18.35 (Fig. S7). For LER, one MTA was identified on chromosome 9 explaining PVE of 16.69%. Two MTAs for GC on chromosome 10 and 7 with the PVE of 18.39 and 18.11% were detected in TPR (Fig. S8).

To confirm the real associations, multi-locus GWAS (mrMLM) was also conducted with three models mrMLM³⁷, FastmrMLM³⁸ and FASTmrEMMA³⁹ by adopting the critical threshold of significance for SNP-trait association was set at a LOD score of greater than 3. On comparing both MLM and mrMLM, we found that all the MTAs identified by MLM were also detected by any one method of mrMLM namely mrMLM, FastmrMLM and FASTmrEMMA with significant LOD score (Table 1, Table S2a). In mrMLM GWAS, S6_6896749 and S11_28809145 exhibited with AC with a high PVE range of 11.52 to 21.83 and 9.26 to 11.50%, S5_5361877 for KB and LBR with PVE of 1.01 to 5.29% and 1.01 to 12.26%, S6_6711302 for ASV with the PVE of 10.55 to 27.19%, S12_17888167 for KLAC with 16.62–17.19% of PVE, S9_17873147 for LER with 1.00–16.07% PVE, S7_24356119 for GC with 37.55% PVE and S11_5737145 for PGI with PVE of 5.71 to 23.55%. As per the model of mrMLM GWAS, the common MTAs identified shown higher values of PVE in the range of 5.29% (S5_5361877

Trait	SNP ID	MLM		mrMLM		Gene ID (RAP)	Gene function
		P value	PVE (%)	LOD Score	PVE (%)		
AC	S8_10622940	1.21E-06	17.96	4.64–5.63	1.85–4.44	<i>Os08g0276000</i>	transmembrane 9 superfamily member, putative, expressed
	S6_6896749	6.05E-06	15.43	4.74–5.82	11.52–21.83	<i>Os06g0232700</i>	transposon protein, putative, unclassified, expressed
	S8_10166593	7.33E-06	15.13	3.88–4.65	1.64–3.61	<i>Os08g0266700</i>	Rad21 / Rec8 like protein, putative, expressed
						<i>Os08g0267050</i>	expressed protein
	S11_28809145	8.24E-06	14.95	4.56–10.66	9.26–11.50	<i>Os11g0704000</i>	selT/selW/selH selenoprotein domain containing protein, expressed
	S2_1904604	9.63E-06	14.71	4.52–6.34	3.61–11.47	<i>Os02g0135800</i>	WD domain, G-beta repeat domain containing protein, expressed
	S10_5085990	1.02E-05	14.61	3.17–4.35	3.94–13.37	<i>Os10g09430</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	S6_10612428	1.12E-05	14.47	5.12–8.70	7.70–25.81	<i>Os06g18710</i>	retrotransposon protein, putative, unclassified
						<i>Os06g18720</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	S8_12286113	1.19E-05	14.39	3.56–4.30	1.70–3.86	<i>Os08g0299700</i>	OsFBO20—F-box and other domain containing protein, expressed
	S2_2474176	1.20E-05	14.38	4.65–6.75	4.20–13.68	<i>Os02g05180</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
KL	S11_28909306	6.64E-06	11.87	3.55–3.95	1.05–1.22	<i>Os11g0706100</i>	alpha-hemolysin, putative, expressed
						<i>Os11g0706500</i>	expressed protein
KB	S5_5361877	1.19E-07	21.48	5.32–8.12	1.01–5.29	<i>Os05g09510</i>	hypothetical protein
						<i>Os05g0187500</i>	IQ calmodulin-binding motif family protein, expressed
	S3_13543674	5.46E-06	15.35	3.68–4.18	1.00–1.01	<i>Os03g23860</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	S3_19893628	7.87E-06	14.79	4.17–5.04	1.01–1.04	<i>Os03g35870</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
	S2_13138765	9.63E-06	14.48	33.55–4.38	1.01–1.06	<i>Os02g0326500</i>	expressed protein
LBR	S5_5361877	8.33E-08	16.99	6.30–9.12	1.01–12.26	<i>Os05g09510</i>	hypothetical protein
						<i>Os05g0187500</i>	IQ calmodulin-binding motif family protein, expressed
ASV	S6_6711302	6.31E-08	25.53	7.87–18.86	10.55–27.19	<i>Os06g0229000</i>	OsFtsH6 FtsH protease, homologue of AtFtsH6, expressed
						<i>Os06g12380</i>	expressed protein
KLAC	S5_28169331	5.88E-07	24.17	12.02	33.68	<i>Os05g49110</i>	retrotransposon protein, putative, Ty1-copia subclass, expressed
						<i>Os05g0566200</i>	NLI interacting factor-like phosphatase, putative, expressed
	S1_9126931	7.90E-07	23.55	8.03	10.41	<i>Os01g16160</i>	retrotransposon protein, putative, unclassified, expressed
	S12_17888167	1.22E-05	18.03	5.55–9.61	16.62–17.19	<i>Os12g29920</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
					<i>Os12g0484375</i>	conserved hypothetical protein	
KBAC	S8_6811	5.78E-06	18.35	4.29–5.25	1.02–1.08	<i>Os08g01010</i>	hypothetical protein
LER	S9_17873147	8.72E-06	16.69	4.19–6.44	1.00–16.07	<i>Os09g0469400</i>	glycogen operon protein glgX, putative, expressed
						<i>Os09g0469900</i>	queuine tRNA-ribosyltransferase, putative, expressed
GC	S10_17333120	9.24E-06	18.39	3.91	8.82	<i>Os10g0469000</i>	leucine-rich repeat receptor protein kinase EXS precursor, putative, expressed
	S7_24356119	1.07E-05	18.11	5.62	37.55	<i>Os07g0597400</i>	expressed protein
RS	S2_22999187	1.83E-06	19.31	4.20–5.53	1.60–3.37	<i>Os02g0594100</i>	protein kinase domain containing protein, expressed
	S8_15000020	1.85E-06	19.29	4.82–6.57	1.87–4.47	<i>Os08g0335500</i>	protein kinase, putative, expressed
	S10_4535302	3.52E-06	18.15	5.64–7.50	1.80–4.21	<i>Os10g08370</i>	transposon protein, putative, CACTA, En/Spm sub-class, expressed
	S2_5050167	4.86E-06	17.58			<i>Os02g0191000</i>	disease resistance protein RPM1, putative, expressed
	S1_8845858	6.13E-06	17.17	4.33–7.14	2.70–7.73	<i>Os01g15720</i>	tobamovirus movement protein containing protein, expressed
	S6_5819355	8.84E-06	16.53	5.37–6.93	1.57–3.27	<i>Os06g11100</i>	retrotransposon protein, putative, unclassified, expressed
	S6_9155098	9.38E-06	16.43	3.59–4.59	1.45–2.80	<i>Os06g16080</i>	transposon protein, putative, CACTA, En/Spm sub-class, expressed
	S11_11627020	1.00E-05	16.32	3.82–5.12	1.69–3.79	<i>Os11g20120</i>	retrotransposon protein, putative, unclassified, expressed
						<i>Os11g20130</i>	retrotransposon protein, putative, unclassified, expressed
	S2_9683531	1.26E-05	15.92	4.53–6.49	2.11–5.46	<i>Os02g16970</i>	retrotransposon protein, putative, unclassified, expressed
PGI	S11_5737145	4.88E-07	23.47	6.23–7.83	5.71–23.55	<i>Os11g10500</i>	retrotransposon protein, putative, Ty1-copia subclass, expressed
	S11_12661155	6.93E-07	22.77	4.42–6.15	1.46–5.55	<i>Os11g22020</i>	retrotransposon protein, putative, unclassified, expressed
						<i>Os11g22030</i>	retrotransposon protein, putative, unclassified, expressed
	S1_20415780	1.99E-06	20.71	5.09–15.20	4.49–17.64	<i>Os01g0548000</i>	expressed protein
	S9_17671895	2.33E-06	20.41	4.84–14.32	4.27–16.76	<i>Os09g29090</i>	retrotransposon protein, putative, unclassified, expressed
	S7_28567910	3.61E-06	19.57	6.17–19.09	5.41–21.35	<i>Os07g0675200</i>	expressed protein
						<i>Os07g0675300</i>	expressed protein
	S8_25864590	5.29E-06	18.84	5.66–9.74	2.77–10.77	<i>Os08g0520300</i>	RNA recognition motif containing protein, putative, expressed
						<i>Os08g0520400</i>	spermidine synthase-related, putative, expressed
	S9_7528951	6.83E-06	18.35	4.00–11.40	3.51–13.73	<i>Os09g13060</i>	retrotransposon protein, putative, Ty3-gypsy subclass, expressed
S3_21132448	8.81E-06	17.87	4.53	3.51	<i>Os03g38070</i>	retrotransposon protein, putative, unclassified, expressed	
S8_7447826	9.30E-06	17.77	4.18–4.71	1.33–5.03	<i>Os08g12580</i>	retrotransposon protein, putative, unclassified, expressed	
					<i>Os08g12590</i>	retrotransposon, putative, centromere-specific	

Table 1. GWAS for grain quality traits, RS content and predicted glycemic index by single locus (MLM) and multi-locus model (mrMLM). AC—Amylose content, KL—Kernel length, KB—Kernel breadth, LBR—Length/Breadth ratio, ASV—Alkali spreading value, KLAC—Kernel length after cooking, KBAC—Kernel breadth after cooking, LER—Linear elongation ratio, GC—Gel consistency, RS—Resistant starch, PGI—Predicted glycemic index, TPR—Transplanted rice and significant associations were detected with the P value < 0.00001 in MLM and $\text{LOD} > 3$ in mrMLM. Bolded MTAs shown strong association confirmed by MLM and mrMLM with high value of PVE.

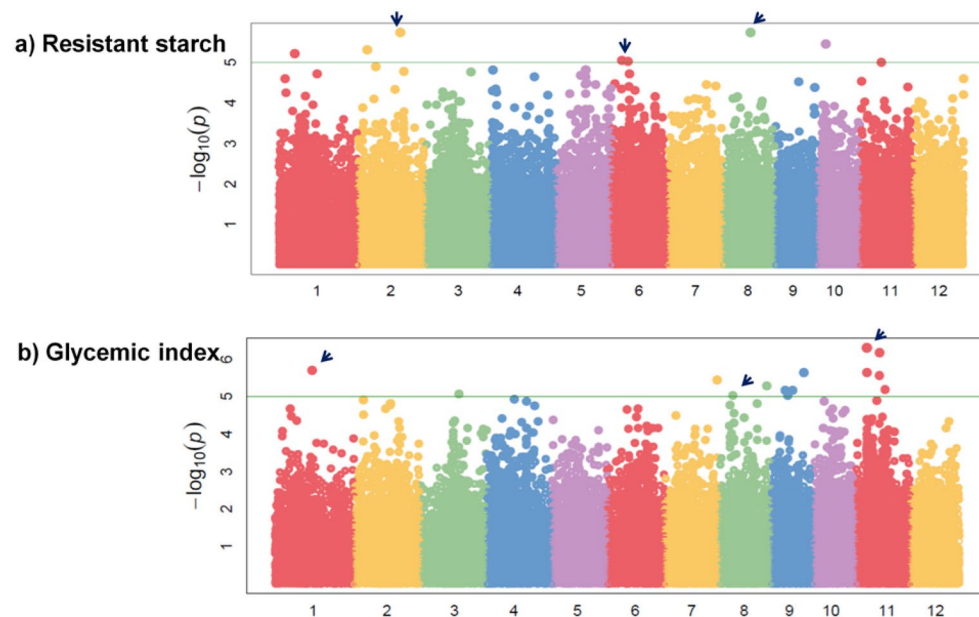


Figure 3. Genome wide association studies—Manhattan plot showing peak associations for resistant starch and predicted glycemic index in (a) Peak association for resistant starch on chromosome 2, 6 and 8 in transplanted rice, (b) Peak association for PGI on chromosome 1, 8 and 11 in transplanted rice.

for KB) to 33.68% (S5_28169331 for KLAC). The strong association MTAs identified by MLM with high PVE or R^2 value were also confirmed by mrMLM models.

Detection of favorable SNP alleles associated with a trait of interest. Associated MTAs of RS, PGI and nine quality traits were subjected for the identification of favorable SNP alleles (Table S2b). In this investigation, the positive effect of candidate SNP alleles that led to increase in AC, KL, L/B ratio, ASV, KLAC, LER, GC, and RS or decrease in KB, KBAC and PGI were defined as “favorable alleles”, and those that resulted in decrease of AC, KL, L/B ratio, ASV, KLAC, LER, GC, and RS or an increase in KB, KBAC and PGI were defined as “unfavorable alleles”. As a result, one MTA for each trait like AC, ASV, GC and RS had an increased phenotypic effect whereas KB, KBAC and PGI revealed decreased effect in their phenotype are designated as favorable alleles. S6_6896749 possessing ‘G’ allele had strong phenotypic effect (0.88%) on AC, S5_5361877 possessing ‘A’ allele with 0.18 of L/B ratio, S6_6711302 with ‘T’ allele had 1.96 phenotypic effects on ASV, S6_5819355 with ‘A’ allele had 1.74% increase in phenotypic effect in RS and S8_7447826 with ‘A’ allele had resulted in decrease of -2.91 in PGI (Table S2b). Findings of this study indicated that the favorable SNP significantly increased or decreased the phenotypic effect of the trait in the genotypes they were present. Identification of superior haplotypes of these mined favorable alleles would be beneficial to develop elite entries with preferred grain quality and lower GI value.

Haplotype analysis of identified candidate loci. All the MTAs associated with the quality-related traits were subjected to the identification of candidate loci using RAP database (Table S2b). Out of the 41 MTAs detected in this investigation, we found 27 MTAs were within the candidate loci and 14 MTAs were found to be flanked by two candidate loci. Of these, nine loci for PGI, nine loci for AC, nine loci for RS, four loci each for KB, KLAC, two loci for GC, one loci each for KL, LBR, ASV, LER and KBAC were identified in this study.

The identified candidate loci were subjected to haplotype analysis by rice SNP seek database to estimate the number of haplotypes present in the sub-set of 3 K-RGP for all the investigated quality-related traits and to use superior haplotypes to breed quality rice varieties. Haplotype analysis reported a minimum of one haplotype to the maximum of 15 haplotypes for the identified loci. The loci *Os08g0276000*, *Os02g0135800* and *Os10g09430*

for AC, *Os05g09510* for KB and LBR, *Os05g0566200* and *Os12g0484375* for KLAC, *Os09g0469400* for LER, *Os02g0594100*, *Os02g0191000*, *Os11g20120* and *Os11g20130* for RS and *Os08g0520400* for PGI had registered one haplotype, respectively. The loci namely, *Os06g18710* and *Os06g18720* for AC, *Os05g0187500* for KB and LBR, *Os06g12380* for ASV, *Os08g01010* for KBAC, *Os06g16080* and *Os02g16970* for RS and *Os07g0675300* for PGI were found to be recorded 15 haplotypes explaining maximum diversity. On the other hand, 14 loci had 2 haplotypes, 7 loci had 3 haplotypes, 4 loci had 4 haplotypes, 3 loci had 5, 1 locus had 8, 10 and 14 haplotypes, respectively (Table S2b).

Haplotype frequency of identified causative genes. The major loci *Os06g0232700* associated with AC had five haplotypes where H5 representing the higher frequency of 57.33%. The other loci on chromosome 6, *Os06g18710* representing maximum haplotype diversity of 15 haplotypes where H4 recorded a higher frequency of 52.66% (Table S3). The candidate SNP for KL centered between two loci namely *Os11g0706100* and *Os11g0706500* exhibited a higher frequency of H5 with 58.66% and H3 with 42.66%, respectively. S5_5361877 is the major allele for KB and LBR falls within candidate genes namely *Os05g09510* and *Os05g0187500*, where *Os05g0187500* possessing 15 haplotypes with the higher haplotype frequency of 48.66% in H14. The trait ASV had an association with S6_6711302 within two loci, in which *Os06g12380* had 15 haplotypes where H5 recorded the maximum frequency of 41.33%.

Haplo-Pheno analysis for the identification of superior haplotypes. For a better understanding of phenotypic performance of the accessions carrying specific haplotype, trait mean was worked out for each haplotype separately. Here, candidate loci possessing more than 2 haplotypes were statistically analyzed by Duncan's test to identify the superiority of the haplotype for the identified MTAs associated with RS, PGI and other quality traits (Table 2). Once MTAs with favorable alleles identified, it was subjected to haplotype analysis with the help IRRI SNP seek database. The candidate loci associated with AC on chromosome 6 namely *Os06g0232700* (H4- 26.50% for high amylose; H5-24.00% for intermediate amylose), *Os06g18710* (H3-25.50%; H6-23.30% for intermediate amylose) and *Os06g18720* (H4-26.58% for high amylose; H14-23.36% for intermediate amylose) had shown positive allelic effect with the superior haplotype of increasing in amylose content. For RS and PGI, the candidate loci *Os06g11100* (H4-3.28% for high RS; H2-2.42% for intermediate RS) and *Os08g12590* (H13-62.52 as intermediate PGI) (Fig. 4) possesses positive and negative allelic effect of 1.74 and -2.91, respectively. For grain size (LBR), H9-2.81 of *Os05g0187500* had the positive allelic effect of 0.18. For ASV and GC, H1-3.43 and H1-56.06 as the desirable haplotype for the candidates of *Os06g0229000* and *Os10g046900* (Table 2).

For KL, *Os11g0706100* had H8 (6.71 mm) and H7 (5.67 mm), *Os11g0706500* had H2 (6.41 mm) as a desirable haplotype for development of long and medium slender grain rice varieties. *Os05g0187500* and *Os03g35870* were associated with KB had H7 (2.37 mm) and H1 (2.42 mm) as superior haplotypes for the corresponding loci. For LBR, *Os05g0187500* had H9 as desirable haplotype having genotypes recorded higher length-breadth ratio (2.89) possessing slender grain rice lines. *Os06g0229000* and *Os06g12380* associated with ASV showed H4 (3.67 of mean score) and H5 (4.48 of mean score) as anticipated one since intermediate ASV is desirable to breed good quality rice varieties. The other loci for RS *Os06g16080* registered H11 (5.37%) for high and H7 (2.95%) as desirable haplotypes for intermediate RS content, respectively. Other loci, *Os02g16970* linked with RS reported H14 (4.49%) as a superior haplotype for improving RS content (Table S4).

Discussion

Rice, one of the essential staple food crop species is consumed by half of the global population especially in Asian continents. The ever-increasing population of the world demands more than 50% additional rice production to be attained by 2030 to meet future needs⁴⁰. Dissecting the complexity of grain quality is possible by means of genome wide association mapping using a diverse set of population with a higher precision approach to map the causative alleles than the bi-parental mapping population since GWAS population consist higher evolutionary recombination events in their genomes⁴¹. During the past three decades, tremendous efforts have been put in by rice breeders and geneticists to detect the causative QTLs/genes responsible for quality improvement in rice and many researchers came up with remarkable results. Grain quality includes cooking and eating quality of rice grains and these are complex traits governed by multiple genes with various gene expression and regulatory network pathways and several QTLs/genes associated with grain quality were identified and cloned in rice⁴². Starch biosynthesis in rice grains during grain formation and development is one of the critical factors in improving the quality aspects of rice varieties and regulation of this complex metabolic pathway controlled by a network of various genes and gene combinations are poorly understood. *Waxy* gene (*GBSSI*) plays a key role in amylose biosynthesis and the formation of long-chain amylopectin⁴³. Another gene *Starch synthase IIa* (*SSIIa*) modifies amylopectin structure and improving grain quality^{44,45}. Likewise, several other genes namely *SSI*, *SSIIIa*, *isoamylase1* (*ISA1*), *pullulanase* (*PUL*), branching enzyme (*BEI* and *BEIIb*)⁴⁶⁻⁵⁰, have also been reported to control starch structure formation and modifies the physicochemical properties of rice grains. Therefore, amylopectin content and structure alteration have a significant impression on the modification of starch granules and affect cooking properties in rice.

In recent times, non-infectious diseases cause several health-related issues like cardiovascular problems, obesity and diabetics due to the high consumption of calories and improper physical exercises. Recent study reported that high amylose starch-rich RS decreases the glycemic value of starch, insulin response in human metabolism and reduce the risk of type II diabetes mellitus and other health-related issues⁵¹. Currently, breeders and geneticists are making efforts to develop high RS rice varieties. Cultivars rich in amylose content of rice, maize and barley developed by either mutational breeding or biotechnological approaches have been found to possess high RS content⁵²⁻⁵⁴.

Trait	Gene ID (RAP)	Allelic effect	Gene function	Average performance of superior haplotype	Average performance of other haplotype	Lines with superior haplotype	Disrtibution of Superior Haplotype		
							Indica	Aus	Admixture
AC	<i>Os06g0232700</i>	0.88	transposon protein, putative, unclassified, expressed	H4-26.50 ^a	H1-24.25 ^b , H2-24.44 ^b , H3-24.03 ^b , H5-24.00 ^b	4	4	0	0
	<i>Os06g18710</i>	0.28	retrotransposon protein, putative, unclassified	H3-25.50 ^a	H1-24.50 ^b , H4-24.12 ^b , H6-23.30 ^c , H7-24.28 ^b , H9-22.11 ^d and H15-24.37 ^b	2	1	1	0
	<i>Os06g18720</i>	0.28	retrotransposon protein, putative, Ty3-gypsy subclass, expressed	H4-26.58 ^a	H1-20.00 ^g , H2-24.45 ^d , H3-25.00 ^g , H5-23.80 ^g , H6-24.64 ^d , H7-24.9 ^{cd} , H8-25.75 ^b , H10-23.11 ^f , H12-19.5 ^h , H13-24.75 ^{cd} , H14-23.36 ^{ef}	3	3	0	0
KB	<i>Os05g0187500</i>	-0.1	IQ calmodulin-binding motif family protein, expressed	H7-2.37 ^a	H1-2.4 ^b , H2-2.42 ^c , H3-2.55 ^b , H4-2.4 ^b , H5-2.56 ^f , H6-2.46 ^{de} , H9-2.4 ^b , H11-2.47 ^g , H12-2.58 ^g , H13-2.45 ^d , H14-2.46 ^{de} , H15-2.46 ^e	6	4	2	0
LBR	<i>Os05g0187500</i>	0.18	IQ calmodulin-binding motif family protein, expressed	H9-2.81 ^a	H1-2.71 ^b , H2-2.69 ^c , H3-2.14 ⁱ , H4-2.66 ^{de} , H5-2.38 ^g , H6-2.68 ^{cd} , H7-2.66 ^c , H11-2.59 ^g , H12-2.33 ^k , H13-2.61 ^f , H14-2.57 ^h , H15-2.51 ⁱ	73	72	0	1
ASV	<i>Os06g0229000</i>	1.96	OsFtsH6 FtsH protease, homologue of AtFtsH6, expressed	H1-3.43 ^c	H2-3.47 ^b , H3-3.00 ^d , H4-3.57 ^a	9	8	1	0
	<i>Os06g12380</i>	1.96	expressed protein	H11-2.75 ^c	H5-4.48 ^h , H6-2.92 ^d , H8-3.67 ^g , H9-3.25 ^f , H10-3.00 ^c , H13-2.58 ^b , H15-2.60 ^b	8	8	0	0
KBAC	<i>Os08g01010</i>	-1.16	hypothetical protein	H11-3.17 ^a	H1-3.40 ^f , H3-3.52 ^g , H4-3.31 ^d , H6-3.18 ^b , H10-3.29 ^c , H13-3.35 ^e , H14-3.53 ^g , H15-3.60 ^h	8	8	0	0
GC	<i>Os10g0469000</i>	5.01	leucine-rich repeat receptor protein kinase EXS precursor, putative, expressed	H1-56.06 ^a	H2-47.73 ^b	6	5	1	0
RS	<i>Os06g11100</i>	1.74	retrotransposon protein, putative, unclassified, expressed	H4-3.28 ^a	H1-3.06 ^a , H2-2.42 ^c , H3-2.85 ^{ab}	48	48	0	0
PGI	<i>Os08g12590</i>	-2.91	retrotransposon, putative, centromere-specific	H13-62.52 ^{ab}	H1-69.34 ^f , H2-67.64 ^{ef} , H5-63.36 ^b , H6-66.68 ^{de} , H9-65.41 ^c , H11-65.53 ^c , H14-71.67 ^g	5	5	0	0

Table 2. Haplo-pheno analysis of favourable candidate loci for grain quality traits, RS content and predicted glycemic index in rice. AC—Amylose content, KL—Kernel length, KB—Kernel breadth, LBR—Length/Breadth ratio, ASV—Alkali spreading value, KLAC—Kernel length after cooking, KBAC—Kernel breadth after cooking, LER—Linear elongation ratio, GC—Gel consistency, RS—Resistant starch, PGI—Predicted glycemic index, TPR—Transplanted rice.

In the present investigation, diverse panel of rice accessions (3 K-RGP sub-set) were not only phenotyped for grain quality-related traits but also phenotyped for the most important health/well-being related traits such as RS and PGI to map candidate SNPs, functional annotations and to unravel the haplotype diversity of identified candidate loci across the 3 K-RGP. Then, the *haplo-pheno* analysis was carried out to ascertain superior or appropriate haplotype for the improvement of grain quality in rice breeding programs.

RS, a valuable trait was found to have a significant positive correlation with amylose content and was negatively correlated with glycemic index as suggested by several previous reports and thus a suitable parameter for rice consumption to improve the health status of those who are suffering from diabetics, colon cancer, and obesity⁷. We identified nine strong MTAs for RS with significant PVE on chromosomes 1, 2, 6, 8 and 11 (Table 1). Primarily, previous reports found that *sbe3-rs* on chromosome 2^{55,56} and *SSIIa* (*Os08g0191433*) on chromosome 8⁴⁵ involved in the production of RS content and it may be associated with expression of *Waxy* gene as well. In our study, we report two MTAs in proximity to the candidate regions mentioned above of chromosome 2 (*Os02g0594100*) and chromosome 8 (*Os08g0335500*) with not much significant haplotype diversity but significant variations in the trait value was observed among the accessions⁵⁷. We also identified a locus on chromosome 6 (*Os06g11100*) near (1 Mb away) to the position of *starch synthase IIa* (*SSIIa: Os06g0229800*, 6,748,398–6,753,302) gene, significantly affecting GC, degree of starch crystallinity by increasing the fractional

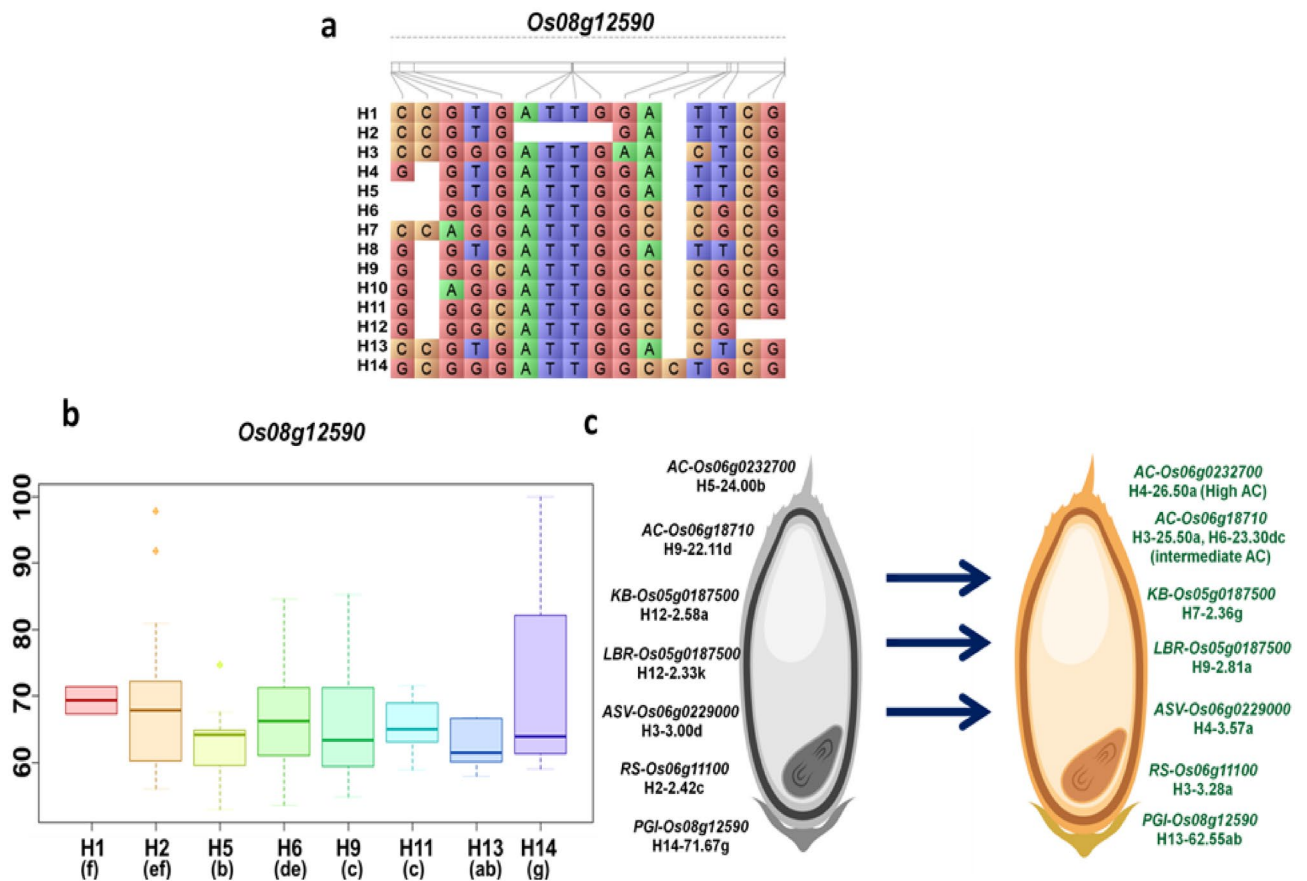


Figure 4. (a) Haplotype variations observed for *Os08g12590* associated with predicted glycemic index in TPR situation, (b) Boxplot shown haplotype diversity and variations exists between haplotypes of *Os08g12590* linked with PGI with the significance level of p value < 0.05 and different alphabets denotes significant differences between haplotypes, (c) Assembly of superior haplotypes (Haplotype based breeding) of RS, PGI and other important quality traits with favorable allelic effect to design diabetic free healthy rice with preferred grain quality and genetic gains.

amount of A chains in amylopectin^{8,44,58}. All the identified MTAs and their corresponding candidate loci exhibit a significant number of variable haplotypes for the traits studied. In recent past, the discovery of mutation/variants each in *GBSSI*, *SSI*, *SSIIa*, and *SSIIIa* with a potential to increase RS content and hydrolysis index in rice has been reported^{59,60}. For PGI, we identified a total of 9 associations on chromosome 1 (*Os01g0548000*), 3 (*Os03g38070*), 7 (*Os07g0675200*), 8 (*Os08g12580* and *Os08g0520300*), 9 (*Os09g13060* and *Os09g29090*) and 11 (*Os11g10500* and *Os11g22020*) (Table 1). Similar kind of results were reported⁹ and we also found significant MTAs influencing PGI in close proximity on chromosome 1 (snp_01_30302075 and snp_01_36980087), Chr 3 (snp_03_22422723), Chr 9 (snp_09_21456056 and snp_09_21523496) and Chr 11 (snp_11_334055 and snp_11_28758733). Similar is the case for GC; our results match with “*qAC7*” associated with RM8261 in the physical position Chr7: 25,866,581⁶¹. Here, we identified a candidate *Os07g0597400* on chromosome 7 associated with GC. A QTL *qRS7-2* on chromosome 7 associated with RS content between RM3404 (chr7:20,105,832) and RM478 (chr7:25,949,521) was identified and reported³⁶.

Over several decades, quality traits like AC, GC, ASV, LER, etc., of rice grains have been extensively studied and it has been reported that AC is governed by waxy gene⁶², by several other loci⁶³ and even by unidentified non-waxy genes⁶⁴. Our study reported nine associations for AC (Fig. S1) and the major SNP S6_6896749 located with the candidate loci of *Os06g0232700* on chromosome 6 found adjacent to *ALK* gene (6,726,252)^{65–67}. Several other loci were identified on Chromosome 2, *Os02g0135800* encoding for WD domain, G-beta repeat domain-containing protein, expressed and *Os02g05180* (2,474,176) encrypting for retrotransposon protein, putative Ty3-gypsy subclass. *Starch branching enzyme 3 (SBEIIIb- Os02g32660)* is essential for amylopectin synthesis in the endosperm^{68,69}. Several reports found that high amylose rice and maize reveals in amylose extender (ae) phenotype due to the inactivation of *SBEIIIb*⁴⁶. These results clearly confirmed that a network of genes are involved in the biosynthesis of amylose. Our association analysis study provides a stronger picture of how genetic network of starch synthesizing genes involved in cooking and eating quality improvement in rice. We also identified an association on chromosome 11, S11_28809145 on *Os11g0704000* with the function of selT/selW/selH seleno-protein domain-containing protein, expressed slightly closer to *Os11g31330*⁷⁰.

Kernel length, kernel breadth, and LBR are the important traits to assess the grain size and dimensions for selecting appropriate breeding lines for their consumer’s preference and market acceptability. Medium slender

rice is more preferable than bold rice varieties. For kernel length, an MTA was identified on chromosome 11 (S11_28909306) with 9 haplotypes. For kernel breadth, one MTA was observed on chromosome 5, two on chromosome 3 and one on chromosome 2. LBR determines the grain size and is controlled by a complex network of genes and also affected by environmental factors. Our study reported similar findings of¹³, who reported quite a lot of QTLs for grain length, grain width and grain length–width ratio on chromosome 3, 5 and 7 by utilizing 3 k panel sub-set. To date, several hundred QTLs associated with grain size have been identified and functionally characterized viz., *GS3*, *GW2*, *qSW5/GW5*, *GS5* and *qGL3/qGL3.1*^{14,16,17,19,71–74}. Among the above said QTLs *GS3* and *qSW5* showed consistent and sturdy effect on grain size variations reported in our investigation also^{71,75–77}. For instance, the locus *Os05g0187500* identified having large haplotype diversity responsible for the putative function of IQ calmodulin-binding motif family protein exhibits peak association close to the SNP position of S5_5361877 reported for KB and LBR^{27,31,41,78}.

ASV is an indicator for gelatinization temperature, we identified a strong and peak association on chromosome 6 at the position of S6_6711302 (between *Os06g0229000* and *Os06g12380*) with worthy haplotype diversity in the range of 4 to 15 haplotypes, respectively. The same QTL on chromosome 6 at the position of 6,726,252 associated with *ALK* gene was reported^{65–67,78}. A study conducted with 258 accessions of 3 k panel also reported the same genic location (S6_6752888) for GT and a candidate of *ALK* gene encoding *SSIla* affecting the chain-length distribution of amylopectin causing alkali disintegration of rice grains^{13,65,79}. For gel consistency, we identified two candidates on chromosome 7 (*Os07g0597400*) and 10 (*Os10g0469000*) with minimum haplotype variations of 8 and 2. Previously, a *qGC6* was reported between close proximity to *WAXY* gene¹³. More importantly, a new novel association for GC is identified on chromosome 7 had near vicinity and it was not previously reported. Linear elongation ratio is an important cooking quality trait and we identified an associations for LER on chromosomes 9. A QTL for LER was reported on chromosome 4 with the marker interval of C933 to C946 using RIL population of *indica* x *indica* hybrids⁸⁰.

On a whole, amylose and amylopectin production is a complex phenomenon governed by several networks of genes and their coordinating expression during grain filling and seed maturation affecting the grain textural properties and even prevailing environmental conditions also. Noteworthy, it is not an easy breeding protocol to improve the quality aspects of rice varieties by selecting a notable gene that may be either superior or inferior for a specific trait and introgression by markers assisted pyramiding and selection. In this, we identified several superior haplotypes by studying haplotype diversity and linking haplotype variations with the trait value. Assembling superior haplotype combinations for most of the grain quality traits viz., AC, ASV, GC, LBR and RS coupled with GI in to one background to maximize the textural properties of grains, yield and most of morpho-agronomic traits to develop tailor-made new generation rice with enhanced genetic gain compared to available popular mega rice varieties^{81,82}. For example, AC (S6_10612428) had maximum haplotype diversity with the superior haplotype of H3 (*Os06g18710* – 25.50%) possessing high amylose class, LBR (*Os05g0187500*) owns H9 (2.81) as superior one among 15 haplotypes, RS (*Os06g11100*) holds H4 (3.28%) as topmost one from all the 4 haplotypes.

In this investigation we also identified better donors for RS, PGI and other quality-related traits which can be utilized in the *haplotype-based breeding* program to develop elite lines with low glycemic index value with desirable quality traits by assembling superior haplotypes suited for different situations (Table S5). Our results were compared with previous reports of glycemic index, we found UQUIHUA::IRGC 117,037–1 had registered intermediate GI of 66.88 with the RS content of 3.23% whereas⁹ reported as low GI (< 55). Apart from this we also selected different classes of glycemic index lines from the 3 K sub-panel grown in 2019WS and different quality-related traits were phenotyped and result showed that four lines namely BAIANG 6::IRGC 6129–1 from Indonesia, MAKRO::IRGC 74,763–1 from India, AUS 329::IRGC 29,116–1 from Bangladesh and KOTTEYARAN::IRGC 47,383–1 from Srilanka possesses medium slender grains with good elongation ratio and medium to soft gel consistency after cooking (Table S6). The line MAKRO::IRGC 74,763–1 from India having intermediate RS content and low PGI of 52.91.

But when considering ASV for GT, the identified entries for the low glycemic index showed the score of 1 to 2 which represents high GT requires more water, cooking time and poor in texture not suitable for cooking and eating (Table S6). Amylose content alone does not describe the cooking and eating quality of rice grains, as varieties with the same range of AC possess variable variations in the cooking and eating quality⁵⁵. In our study, intermediate ASV of 4 to 5 score possesses high PGI values. Superior haplotypes associated with the low glycemic index, intermediate ASV and soft GC in the preferred combinations can be utilized for the development of low glycemic rice varieties with desirable grain quality. Hence, the approach of *haplotype-based breeding* is anticipated to assist in the development of -premium quality rice varieties with low GI to meet the increasing demands of the rice consuming population.

Conclusion

Significant variations were observed for the grain quality traits, especially for AC, ASV, GC and RS. We captured several novel and significant associations of SNPs for the target traits and studied haplotype differences of identified candidate genes. Newly identified candidate genes might be useful for further functional characterization and pathway elucidation for grain quality traits in rice. Here we proposed an exclusive strategy on use of superior haplotype-possessing elite donors and incorporating superior/appropriate haplotypes for the majority of quality traits in a genetic background as a prominent way to develop high yielding and quality rich rice varieties suitable for the consumption of diabetics, obese population and as preferred by consumer's needs and demands.

Material and methods

Genetic material and field trial. The genetic material used in this study comprised of 150 accessions having different duration groups of 3 K RGP evaluated under transplanted situation (TPR) during dry season 2017 (DS2017) and Wet season 2018 (WS2018) (Table S7). These accessions were evaluated for RS, PGI and 9 quality parameters including physico-chemical properties, cooking and eating properties. The analysis was carried out at the rice quality lab at IRRI SA Hub, ICRISAT campus. The agronomic management of the transplanted experiments including was carried as per the procedure detailed⁸³.

Phenotyping of physicochemical parameters and cooking and eating quality of rice. Traits namely, KL, KB and LBR were measured with 10 kernels using Standard Evaluation System of Rice (SES)⁸⁴ and the mean was calculated. Amylose content of the grains was estimated by the rapid protocol of cut grain dip method⁸⁵. For cooking-related parameters, 20 kernels from each entry were soaked in 5 ml of distilled water in 15 ml test tubes for 20 min and cooked in a boiling water bath at 100 °C intended for the time period of 8 min to determine the KLAC, KBAC and LER. Gel consistency was determined by the method formulated by⁸⁶, where known amount of rice flour was placed in the culture tubes and wetted with 0.2 ml 95% ethanol containing 0.03% thymol blue and 2 ml 0.2 N KOH added and mixed with a Vortex Genie mixer set at specified speed. Tubes were covered with glass marbles and heated in a vigorously boiling water bath for 8 min. Then, the tubes were removed from the water bath and kept at room temperature for 5 min, cooled in an ice-water bath for 20 min, and laid flat on a laboratory table and length of the blue-colored gel was measured in millimeters. The method separated different classes of GC into soft (61–100 mm), medium (41–60 mm) and hard gel consistency (25–40 mm) based on SES, 2013. GT was estimated based on alkali spreading value (ASV) of milled rice by⁸⁷. Six whole-grain, milled rice samples were placed in duplicate Petri plates containing 10 ml of 1.7% KOH. The Petri plates were covered and incubated for 23 h at 30 °C. The appearance and disintegration of grains were visually observed after incubation based on the scale given by SES, 2013.

Phenotyping for RS and predicted GI. *Measurement of RS.* RS content of the rice samples was determined by using a resistant starch assay kit (K-RSTAR, Megazyme, Irishtown, Ireland) by ten-fold downscaling of sample and reagents with slight modifications⁸⁸. Fine powder of polished rice flour (10 mg) was taken in 2 ml eppendorf tube and it was digested with 400 µl of enzymatic mixture-containing 10 mg/ml of pancreatic α-amylase and 3.0U/ml of amyloglucosidase (AMG) and the mixture was incubated at 37°C for 16 h with continuous shaking (200 strokes/minute). After incubation, 400 µl of 99% ethanol was added to the mixture and vortexed to stop the reaction. Then, the samples were centrifuged for 10 min at 12,000 rpm and the supernatant was collected and the pellets were washed repeatedly with 200 µl (vortexed) and 600 µl of 50% ethanol. Washing with 50% ethanol was made twice to remove all the non-resistant starch in the samples and supernatants were pooled for the measurement of non-resistant starch. Then, the pellets were air-dried for 30 min to remove moisture content in the residue. Around, 200 µl of 2 M KOH was added to the residue and mixed properly to avoid the formation of clumps and incubated at 5°C for 1 h in a shaker (200 strokes/minute). Then, 800 µl of 1.2 M sodium acetate buffer and 10 µl of AMG (3,300 U/ml), vortexed and tubes were placed in a water bath for 30 min at 50°C with intermittent mixing with the help of vortex mixer for every 10 min. After incubation, samples were cooled and centrifuged at 12,000 rpm for 10 min. Then, 100 µl of sample aliquot was taken in a fresh test tube (15 ml), 3 ml of GOPOD reagent (D-Glucose assay kit, Megazyme, Irishtown, Ireland) and incubated in a water bath at 50°C for 30 min. The absorbance was determined with the help of UV-1800 (Shimadzu Corporation, Japan) at 510 nm against a blank containing buffer and GOPOD reagent.

Estimation of predicted GI (PGI). The glycemic index was determined by the protocol given by¹⁰ with little modifications. Fine powder of polished rice flour (50 mg) was cooked in 5 ml of water for 30 min and 10 ml of HCl-KCl buffer (pH = 1.5) was added to the sample. Then, 0.2 ml of a solution containing 1 g of pepsin in 10 ml of HCl-KCl buffer was added to each sample and incubated at 40 °C for 1 h in a shaking water bath. After incubation, the volume was adjusted to 25 ml with Tris maleate buffer (pH = 6.9). Then, 5 ml of α-amylase (2.6 units) in Tris maleate buffer (pH 6.9) was added and samples were incubated at 37 °C in a shaking water bath. After incubation, 1 ml of the sample aliquot was collected serially at intervals of 30 min for 3hrs (30, 60, 90, 120 and 180 min). The enzyme activity in the aliquot was inactivated by heating at 100 °C for 5 min and refrigerated until the end of the incubation period. 3 ml of 0.4 M Sodium acetate buffer (pH 4.75) and 60 µl amyloglucosidase was added to hydrolyze the digested starch to glucose. The samples were incubated at 60 °C for 45mins in a water bath and glucose content in each aliquot was estimated using the GOPOD kit (D-Glucose assay kit, Megazyme, Irishtown, Ireland). Glucose was converted into starch by multiplying with 0.9. Kinetics of starch digestion was estimated by non-linear first-order equation; $C = C_{\infty} (1 - e^{-kt})$. Where, C = Concentration of starch hydrolyzed at the time (t), C_{∞} = equilibrium concentration (i.e., % of starch hydrolyzed after 180 min which is the glucose content after 120 min divided by the total starch) and k = kinetic constant. The area under the hydrolysis curve (AUC) was calculated using the equation; $AUC = C_{\infty} / k (1 - e^{-k(t_f - t_0)})$. Where, C_{∞} corresponds to the concentration at equilibrium (t_{180}), t_f is the final time (180 min), t_0 is the initial time (0 min) and k is the kinetic constant. A hydrolysis index (HI) was calculated by comparison with the AUC of a reference food (white bread). The predicted glycemic index (PGI) was estimated using the formula $PGI = 39.7 + 0.548 (HI)$.

GWAS, Haplotype analysis and Haplo-Pheno analysis. Mean data was used for GWAS analyses using approximately more than 500 K SNP data. The analysis was carried out by using GAPIT (MLM model) (<https://CRAN.R-project.org/>), R based approach considering both kinship values (k-values) and population structure (Q-matrix)⁸⁹. Of the several methods suggested to correct false positive in association analysis even

keeping stringent p-value benchmark, the most stringent correction method called “Bonferroni Correction” was used in the present analysis. The bonferroni threshold calculated by the formula $1/m$ where ‘m’ is the number of test performed which resulted in $-\log_{10}(1/m) = 5.69$. As it is too conservative and suggestive value of $p = 0.00001$ ($-\log_{10}(p) > 5$)^{31,90} was used for the identification of peak associations (MTAs) with the targeted grain quality traits. Multi-locus GWAS (mrMLM) was also conducted to validate the MTAs identified by single locus model of MLM by adopting the significant LOD score of 3^{91} . The phenotypic allele effect (ai) was determined by the formula given by²⁹ and the favorable alleles of each trait were subsequently identified according to the breeding objective. Strongly associated SNPs with the targeted traits were utilized to find the causative QTLs/genes by the RAP database. In-built tool of SNP seek database⁸⁹ was utilized to conduct haplotype analysis for entire candidate locus, by employing default parameters with Calinski criteria for k-group determination. Nipponbare was used as the reference genome. All of the 150 lines belonging to 8 sub-populations namely, aro, aus, admix, ind1A, ind1B, ind2, ind3, and indx was considered for the haplotype analysis. We utilized the ‘3kfiltered’ SNP set present in the SNP seek database for the entire analysis. The filtered was obtained from the Base SNP set by applying the following filtering criteria: (1) alternative allele frequency at least 0.01, (2) proportion of missing calls per SNP at most 0.2⁹² (http://snp-seek.irri.org/_download.zul) and this SNP set was already available in the SNP seek database which was directly utilized in this study. Haplotype analysis for candidate locus have been carried out considering only the nonsynonymous SNPs and indels in the exon region that results in amino-acid change. The information regarding haplotype and their diversity was obtained from the SNP-seek database (http://snp-seek.irri.org/_download.zul) to detect the superior haplotype by categorizing haplotypes by using phenotyping data of concerned genotype trait means for the associated genes.

Statistical analysis. Mean, range and standard deviation were calculated with the help of a standard excel program. Duncan’s multiple range tests were carried out using XL Stat 2019. Correlation analysis among all the component traits related to quality was conducted by SPSS ver. 20. Boxplot and correlogram were constructed with the help of R-package (<https://CRAN.R-project.org/>). Significant difference among the haplotypes (only the haplotypes validated in at least two lines were considered) was studied with the help of F-statistic and Duncan’s test with the significance level of $p < 0.05$.

Received: 7 October 2020; Accepted: 19 March 2021

Published online: 12 May 2021

References

- Hao, W., Zhu, M., Gao, J., Sun, S. & Lin, H. Identification of quantitative trait loci for rice quality in a population of chromosome segment substitution lines. *J. Integr. Plant Biol.* **51**, 500–512 (2009).
- Guo, X., Lei, C., Wang, J., Jiang, L. & Wan, J. Identification of a stable quantitative trait locus for percentage grains with white chalkiness in rice (*Oryza sativa*). *J. Integr. Plant Biol.* **53**, 598–607 (2011).
- Khush, G.S., Paule, C.M. and Cruz, N.M. Rice grain quality evaluation and improvement at IRRI. In Proceedings of the Workshop on Chemical Aspects of Rice Grain Quality, IRRI, Los Banos, Philippines, pp. 21–31 (1979).
- Barclay, A. W. *et al.* Glycemic index, glycemic load, and chronic disease risk—a meta-analysis of observational studies. *Am. J. Clin. Nutr.* **87**, 627–637 (2008).
- Gnagnarella, P., Gandini, S., La Vecchia, C. & Maisonneuve, P. Glycemic index, glycemic load, and cancer risk: a meta-analysis. *Am. J. Clin. Nutr.* **87**, 1793–1801 (2008).
- Fitzgerald, M. A. *et al.* Identification of a major genetic determinant of glycaemic index in rice. *Rice* **4**, 66–74 (2011).
- Parween, S. *et al.* Balancing the double-edged sword effect of increased resistant starch content and its impact on rice texture: the genetics and molecular physiological mechanisms. *Plant Biotechnol. J.* <https://doi.org/10.1111/pbi.13339> (2020).
- Abbai, R., Singh, V. K., Snowdon, R. J., Kumar, A. & Schnurbusch, T. Seeking crops with balanced parts for the ideal whole. *Trends Plant Sci.* **25**(12), 1189–1193 (2020).
- Anacleto, R. *et al.* Integrating a genome-wide association study with a large-scale transcriptome analysis to predict genetic regions influencing the glycaemic index and texture in rice. *Plant Biotechnol. J.* **17**, 1261–1275 (2019).
- Goni, I., Garcia-Alonso, A. & Saura-Calixto, F. A starch hydrolysis procedure to estimate glycemic index. *Nutrition Res.* **17**, 427–437 (1997).
- Lai, M. H., Liu, K. L., Chen, P. & Y., Ke N.J., Chen J., Sung J., Wu Y. L., Lin S.D. . Predicted glycemic index and glycemic index of rice varieties grown in Taiwan. *Cereal Chem. J.* **93**(10), 150–155 (2015).
- Huggins, T. D. *et al.* Association analysis of three diverse rice (*Oryza sativa* L.) germplasm collections for loci regulating grain quality traits. *Plant Genome* **12**, 170085 (2019).
- Wang, X. *et al.* New candidate genes affecting rice grain appearance and milling quality detected by genome-wide and gene-based association analyses. *Front. Plant Sci.* **7**, 1998 (2017).
- Song, X. J., Huang, W., Shi, M., Zhu, M. Z. & Lin, H. X. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nature Genet.* **39**, 623–630 (2007).
- Wang, E. *et al.* Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nature Genet.* **40**, 1370–1374 (2008).
- Shomura, A. *et al.* Deletion in a gene associated with grain size increased yields during rice domestication. *Nature Genet.* **40**, 1023–1028 (2008).
- Mao, H. L. *et al.* Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl Acad. Sci. USA* **107**(45), 19579–19584 (2010).
- Li, Q. F. *et al.* Characterization of expression of the OsPUL gene encoding a pullulanase-type debranching enzyme during seed development and germination in rice. *Plant Physiol. Biochem.* **47**, 351–358 (2009).
- Zhang, X. *et al.* Rare allele of OsPPKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc. Natl Acad. Sci. USA* **109**(52), 21534–21539 (2012).
- Wang, S. *et al.* Control of grain size, shape and quality by OsSPL16 in rice. *Nature Genet.* **44**, 950–954 (2012).
- Sun, L. *et al.* GS6, a member of the GRAS gene family, negatively regulates grain size in rice. *J. Integr. Plant Biol.* **55**, 938–949 (2013).
- Hu, J. *et al.* A rare allele of GS2 enhances grain size and grain yield in rice. *Mol. Plant* **8**(10), 1455–1465 (2015).

23. Wang, S. *et al.* The *OsSPL16GW7* regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nature Genet.* **47**, 949–954 (2015).
24. Liu, S. *et al.* *OsMAPK6*, a mitogen-activated protein kinase, influences rice grain size and biomass production. *Plant J.* **84**, 672–681 (2015).
25. Si, L. *et al.* *OsSPL13* controls grain size in cultivated rice. *Nature Genet.* **48**, 447–456 (2016).
26. Jin, J. *et al.* *Plant Cell* **28**, 2453–2463 (2016).
27. Zhao, K. *et al.* Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* **2**, 467 (2011).
28. Singh, R., Singh, A., Sharma, T., Singh, A. & Singh, N. Fine mapping of grain length QTLs on chromosomes 1 and 7 in Basmati rice (*Oryza sativa* L.). *J. Plant Biochem. Biotechnol.* **21**, 157–166 (2012).
29. Zhang, T. *et al.* Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in China. *PLoS ONE* **8**, e57220 (2013).
30. Qiu, X. *et al.* Genome-wide association study of grain appearance and milling quality in a worldwide collection of Indica rice germplasm. *PLoS ONE* **10**, e0145577 (2015).
31. Feng, Y. *et al.* Genome wide association mapping for grain shape traits in *indica* rice. *Planta* **244**(4), 819–830 (2016).
32. Edzesi, W. M. *et al.* Genetic diversity and elite allele mining for grain traits in rice (*oryza sativa* L.) by association mapping. *Front. Plant Sci.* **7**, 787 (2016).
33. Wang, J., Hu, P. & Lin, L. Gradually decreasing starch branching enzyme expression is responsible for the formation of heterogeneous starch granules. *Plant Physiol.* **176**, 582–595 (2018).
34. Zhang, G. *et al.* Double repression of soluble starch synthase genes *SSIIa* and *SSIIIa* in rice (*Oryza sativa* L.) uncovers interactive effects on the physicochemical properties of starch. *Genome* **54**(6), 448–459 (2011).
35. Pandey, M. K. *et al.* Different isoforms of starch-synthesizing enzymes controlling amylose and amylopectin content in rice (*Oryza sativa* L.). *Biotechnol. Adv.* **30**, 1697–1706 (2012).
36. Zeng, Y. W. *et al.* Identification of QTLs for resistant starch and total alkaloid content in brown and polished rice. *Genet. Mol. Res.* **15**(3), 15037268 (2015).
37. Wang, S. B. *et al.* Improving power and accuracy of genome-wide association studies via a multi-locus mixed linear model methodology. *Sci. Rep.* **6**, 19444. <https://doi.org/10.1038/srep19444> (2016).
38. Wen, Y. J. *et al.* Methodological implementation of mixed linear models in multi-locus genome-wide association studies. *Brief. Bioinform.* **19**, 700–712 (2018).
39. Zhang, Y.W., Tamba, C.L., Wen, Y.J., Li, P., Ren, W.L., Ni, Y.L., Gao, J., Zhang, Y.M. mrMLM v4.0: An R platform for multi-locus genome-wide association studies. *Genomics, Proteomics & Bioinformatics*, online, DOI: <https://doi.org/10.1016/j.gpb.2020.06.006> (2020).
40. Khush, G. S. What it will take to feed 5.0 billion rice consumers in 2030. *Plant Mol. Biol.* **59**, 1–6 (2005).
41. Wang, J., Hu, P., Chen, Z., Liu, Q. & Wei, C. Progress in high-amylose cereal crops through inactivation of starch branching enzymes. *Front. Plant Sci.* **8**, 469 (2017).
42. Zhang, Y. *et al.* Genetic dissection of yield traits in super hybrid rice Xieyou9308 using both unconditional and conditional genome-wide association mapping. *Sci. Rep.* **7**, 824 (2017).
43. Hanashiro, I. *et al.* Granule-bound starch synthase I is responsible for biosynthesis of extra-long unit chains of amylopectin in rice. *Plant Cell Physiol.* **49**(6), 925–933 (2008).
44. Nakamura, Y. *et al.* Essential amino acids of starch synthase IIa differentiate amylopectin structure and starch quality between japonica and indica rice varieties. *Plant Mol. Biol.* **58**, 213–227 (2005).
45. Zhou, H. *et al.* Critical roles of soluble starch synthase *SSIIIa* and granule-bound starch synthase *waxy* in synthesizing resistant starch in rice. *Proc. Natl Acad. Sci. USA* **113**, 12844–12849 (2016).
46. Nishi, A., Nakamura, Y., Tanaka, N. & Satoh, H. Biochemical and genetic analysis of the effects of amylose extender mutation in rice endosperm. *Plant Physiol.* **127**, 459–472 (2001).
47. Fujita, N. *et al.* Antisense inhibition of isoamylase alters the structure of amylopectin and the physicochemical properties of starch in rice endosperm. *Plant Cell Physiol.* **44**, 607–618 (2003).
48. Satoh, M. *et al.* Nonhydrostatic icosahedral atmospheric model (NICAM) for global cloud resolving simulations. *J. Comput. Phys.* **227**, 3486–3514 (2008).
49. Fujita, N. *et al.* Function and characterization of starch synthase I using mutants in rice. *Plant Physiol.* **140**, 1070–1084 (2006).
50. Fujita, N. *et al.* Characterization of *SSIIIa*-deficient mutants of rice: the function of *SSIIIa* and pleiotropic effects by *SSIIIa* deficiency in the rice endosperm. *Plant Physiol.* **144**, 2009–2023 (2007).
51. Zhu, L. *et al.* High-amylose rice improves indices of animal health in normal and diabetic rats. *Plant Biotechnol. J.* **10**, 353–362 (2012).
52. Bird, A. R. *et al.* Wholegrain foods made from a novel high-amylose barley variety (Himalaya 292) improve indices of bowel health in human subjects. *Br. J. Nutr.* **99**, 1032–1040 (2008).
53. Jiang, H., Campbell, M., Blanco, M. & Jane, J. Characterization of maize amylose-extender (ae) mutant starches. Part II: Structures and properties of starch residues remaining after enzyme hydrolysis at boiling-water temperature. *Carbohydr. Polym.* **80**, 1–12 (2010).
54. Hallstrom, E., Sestili, F., Lafandra, D., Bjorck, I. & Ostman, E. A novel wheat variety with elevated content of amylose increases resistant starch formation and may beneficially influence glycaemia in healthy subjects. *Food Nutr. Res.* **55**, 7074 (2011).
55. Yang, R. *et al.* A putative gene *sbe3-rs* for resistant starch mutated from *SBE3* for starch branching enzyme in rice (*Oryza sativa* L.). *PLoS ONE* **7**(8), e43026 (2012).
56. Biselli, C. *et al.* GWAS for starch-related parameters in japonica rice (*Oryza sativa* L.). *Plants* **8**(8), pii: E292 (2019).
57. Bao, J., Zhou, X., Xu, F., He, Q. & Park, Y. J. Genome-wide association study of the resistant starch content in rice grains. *Starch/Stärke* **69**, 1600343 (2017).
58. Kong, X., Kasapis, S. & Bao, J. Viscoelastic properties of starches and flours from two novel rice mutants induced by gamma irradiation. *LWT - Food Sci. Technol.* **60**(1), 578–582 (2015).
59. Raja, R. B. *et al.* EcoTILLING by sequencing reveals polymorphisms in genes encoding starch synthases that are associated with low glycemic response in rice. *BMC Plant Biol.* **17**, 13 (2017).
60. Selvakumar, G., Ramadoss, M., Nayak, Kalagatur, Bapu., Mohan., Alqarawi., Hashem and Abd_Allah. Single Nucleotide Polymorphisms (SNPs) in starch biosynthetic genes associated with increased resistant starch concentration in rice mutant. *Front. Plant Sci.* doi: <https://doi.org/10.3389/fgene.2019.00946> (2019).
61. Bruno, E., Choi, Y. S., Chung, I. K. & Kim, K. M. QTLs and analysis of the candidate gene for amylose, protein, and moisture content in rice (*Oryza sativa* L.). *Biotech* **7**, 40 (2017).
62. Mikami, I. *et al.* Allelic diversification at the *wx* locus in landraces of Asian rice. *Theor. Appl. Genet.* **116**, 979–989 (2008).
63. Aluko, G. *et al.* QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* x *O. glaberrima*. *Theor. Appl. Genet.* **109**, 630–639 (2004).
64. Li, J. *et al.* QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. *Genome* **47**(4), 697–704 (2004).

65. Gao, Z. *et al.* *ALK*, the key gene for gelatinization temperature, is a modifier gene for gel consistency in rice. *J. Integr. Plant Biol.* **53**, 756–765 (2011).
66. Bao, J. S., Corke, H. & Sun, M. Microsatellites, single nucleotide polymorphisms and a sequence tagged site in starch-synthesizing genes in relation to starch physico-chemical properties in non-waxy rice (*Oryza sativa* L.). *Theor. Appl. Genet.* **113**, 1185–1196 (2006).
67. Mogga, M., Sibiya, J., Shimelis, H., Lamo, J. & Yao, N. Diversity analysis and genomewide association studies of grain shape and eating quality traits in rice (*Oryza sativa* L.) using DArT markers. *PLoS ONE* **13**, e0198012 (2018).
68. Jeon, J. S., Ryoo, N., Hahn, T. R., Walia, H. & Nakamura, Y. Starch biosynthesis in cereal endosperm. *Plant Physiol. Biochem.* **48**, 383–392 (2010).
69. Lu, F. H. *et al.* An SNP downstream of the *OsBELLb* gene is significantly associated with amylose content and viscosity properties in rice (*Oryza sativa* L.). *J. Cereal Sci.* **56**, 706–712 (2012).
70. Wambugu, P., Ndjiondjop, M. N., Furtado, A. & Henry, R. Sequencing of bulks of segregants allows dissection of genetic control of amylose content in rice. *Plant Biotechnol. J.* **16**, 100–110 (2018).
71. Fan, C. *et al.* *GS3*, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* **112**, 1164–1171 (2006).
72. Weng, J. *et al.* Isolation and initial characterization of *GW5*, a major QTL associated with rice grain width and weight. *Cell Res.* **18**, 1199–1209 (2008).
73. Li, Y. *et al.* Natural variation in *GS5* plays an important role in regulating grain size and yield in rice. *Nature Genet.* **43**(12), 1266–1269 (2011).
74. Qi, P. *et al.* The novel quantitative trait locus *GL31* controls rice grain size and yield by regulating Cyclin-T13. *Cell Res.* **22**, 1666–1680 (2012).
75. Duan, P. *et al.* Natural variation in the promoter of *GSE5* contributes to grain size diversity in rice. *Mol. Plant* **10**, 685–694 (2017).
76. Liu, Q. *et al.* G-protein β subunits determine grain size through interaction with MADS-domain transcription factors in rice. *Nature Commun.* **9**(1), 852 (2018).
77. Misra, G. *et al.* Whole genome sequencing-based association study to unravel genetic architecture of cooked grain width and length traits in rice. *Sci. Rep.* **7**, 12478 (2017).
78. Huang, X. *et al.* Genome-wide association studies of 14 agronomic traits in rice landraces. *Nature Genet.* **42**, 961–967 (2010).
79. Umemoto, T., Yano, M., Satoh, H., Shomura, A. & Nakamura, Y. Mapping of a gene responsible for the difference in amylopectin structure between *japonica*-type and *indica*-type rice varieties. *Theor. Appl. Genet.* **104**, 1–8 (2002).
80. Liu, L. L. *et al.* Identification of stably expressed quantitative trait loci for cooked rice elongation in non-Basmati varieties. *Genome* **51**(2), 104–112 (2008).
81. Bevan, M. W. *et al.* Genomic innovation for crop improvement. *Nature* **543**, 346–354 (2017).
82. Abbai, R. *et al.* Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains. *Plant Biotechnol. J.* **17**, 1612–1622 (2019).
83. ICAR-Indian Institute of Rice Research (IIRR), 52th All India Co-Ordinated Programme: Agronomy Technical Programme, Rajendranagar, Hyderabad, (2016–2017).
84. Standard Evaluation System for Rice. International rice research institute (2013).
85. Agasimani, S., Selvakumar, G., Joel, A. J. & Ganesh Ram, S. A simple and rapid single kernel screening method to estimate amylose content in rice grains. *Phytochem. Anal.* **24**(6), 569–573 (2013).
86. Cagampang, G. B., Perez, C. M. & Juliano, B. O. A gel consistency test for eating quality of rice. *J. Sci. Food Agric.* **24**, 1589–1594 (1973).
87. Little, R. R., Hilder, G. B. & Dawson, E. H. Differential effect of dilute alkali on 25 varieties of milled white rice. *Cereal chem.* **35**, 111–126 (1958).
88. Butardo, V. M. *et al.* Impact of down-regulation of starch branching enzyme IIb in rice by artificial microRNA- and hairpin RNA-mediated RNA silencing. *J. Exp. Bot.* **62**, 4927–4941 (2011).
89. Yu, J. *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat. Genet.* **38**, 203–208 (2006).
90. Misra, G. *et al.* Dissecting the genome-wide genetic variants of milling and appearance quality traits in rice. *J. Exp. Bot.* **70**(19), 5115–5130 (2019).
91. Zhang, Y. M., Jia, Z. & Dunwel, J. M. The applications of new multi-locus GWAS methodologies in the genetic dissection of complex traits. *Front. Plant Sci.* **10**, 100. <https://doi.org/10.3389/fpls.2019.00100> (2019).
92. Mansueto, L. *et al.* Rice SNP-seek database update: new SNPs, indels, and queries. *Nucleic Acids Res.* **45**, 1075–1081 (2016).

Acknowledgements

Authors are grateful to the Department of Science and Technology (DST), Science and Engineering Research Board (SERB), Government of India and Department of Biotechnology, Government of India for providing financial support.

Author contributions

A.K. and V.K.S. conceived and designed the experiments. R.S., A.K.S., S.V.H. and U.M.S. performed the phenotyping of quality and related traits. R.S., A.K.S. and R.A. analysed and interpreted the result. R.S. and A.K.S. prepared the manuscript. A.K. and V.K.S. revised and finalized the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-87964-8>.

Correspondence and requests for materials should be addressed to A.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021