Analysis of genomic regions for crude protein and fractions of protein using a recombinant inbred population in Rice (Oryza sativa L.)

Sajid Fiaz a,b, Zhonghua Sheng a, Aqib Zeb a, Hirendra Nath Barman a, Tahmina Shar a, Umed Ali a and Shaoqing Tang a

a State Key Laboratory of Rice Biology, China National Centre for Rice Improvement, China National Rice Research Institute, Hangzhou, People’s Republic of China; b Department of Plant Breeding and Genetics, University of Haripur, Haripur, Pakistan

ABSTRACT

The improvement for grain seed storage proteins (SSPs) is integral for rice breeding to achieve superior nutritional quality. In present investigation, a sum of 44 QTLs for PC, GLU, GLO, ALB and PRO were detected on all 12 chromosomes, with single QTL explaining 3.70–12.47% of phenotypic variation. The majority of detected QTLs located in the region of RM7158-RM3414 including Waxy (Wx) gene along with qPC1. The highly positive significant correlation showed traits under investigation are related to each other. Three QTLs (qPC6, qPC7 and qGLU6) had shown relatively higher rate of phenotypic contribution suggesting influence of environmental conditions. Whereas, a pair of epistatic QTLs, GLO and PRO were also detected meanwhile, the M-QTLs were the primary genetic basis for PC, GLU and ALB, respectively. The outcome of present study will help to unearth the genetic foundation of protein and fractions of protein for future grain nutritional quality improvement programs.

ARTICLE HISTORY

Received 20 November 2020 Revised 27 May 2021 Accepted 8 July 2021

KEYWORDS

QTLs mapping; Waxy; marker-assisted selection; nutritional quality

1. Introduction

Rice, being the important source of caloric intake, ensures food security to billions of people around the globe [1]. The quality improvement with yield is a hotspot for rice breeders worldwide, especially in Asian rice-growing countries [2]. The grain quality is complex and the sum of different attributes includes cooking and eating, nutritional, milling and appearance quality [3]. Healthy quality improvement ultimately uplifts human nutrition and health; however, approximately one-third of humans on earth are suffering from an inadequate supply of protein. The inferior protein quality, unavailability of essential vitamins and micronutrients may cause some diseases [4]. The rice grain is composed of more than 90% protein and starch [5,6]. In addition, protein and fractions of proteins, known as seed storage proteins (SSPs), are key elements, explaining the nutritional quality aspects [7]. Rice holds lower protein content than other cereals, but net utilization of protein is higher owing to the most significant consumers [8]. Similarly, the high nutritional quality rice demand has been estimated to upsurge over time [9].

The rough endoplasmic reticulum (ER) is the key place for the synthesis of SSPs, translocated to ER lumen and further transferred to discrete intracellular compartments of the plant’s endomembrane system [10]. Rice contains a relatively balanced amino acid composition and the SSPs are fractioned into albumins (ALB), globulins (GLO), prolamins (PRO) and glutelins (GLU), according to the differences in solubility [11]. Rice SSPs have the second-highest lysine content, the limiting amino acid after oats [12]. The GLU, an alkali-soluble protein, makes 80% of SSPs, found in the milled fraction [13]. It is accumulated in irregularly shaped protein bodies II (PB-II) derived from the protein storage vacuole and globulin, as 57 kDa precursor [14]. The GLUs are further divided into subfamilies: GluA, GluB, GluC, and GluD, depending on resemblance in amino acid [15], with high nutrition value for human’s diet [16]. GLU is a major SSP of the rice grain. Any modification in it may cause a significant influence on grain quality. The PRO, an equally distributed alcohol-soluble protein, makes lower than 5% of SSPs and is stored in spherical PB-I. The PRO is categorized into three types based on their molecular mass: 10 kDa prolamin (RP10), 13 kDa prolamin (RM1, RM2, RM4, and RM9), and 16 kDa prolamin (RP16) [15]. Water-soluble ALB and salt-soluble GLO are primarily concentrated in the embryo and outer aleurone layer of the endosperm. We lose the major portion of these protein fractions during the polishing process [17]. Moreover, classes of proteins, i.e. ALB and GLO, have also been known for allergenic proteins. The rice hypoallergenic protein may be helpful for patients being allergic to these proteins. In addition, a
low glutelin diet is recommended for patients suffering from diabetes and kidney stones [18]. Therefore, more consideration must be given to protein quality and its concentration in any rice breeding programme.

The rice SSPs are polygenic and a mixture of highly polymorphic polypeptides, mutation in some the structural genes may have less influence on grain protein contents [19]. From the past decades, more intensive efforts have been undertaken to understand the genetic base of crude protein content (PC) of rice [20,21]. These studies have shown significant influencing QTLs located on chromosomes 1 and 6, respectively. A significant gene qPC1 encoding a putative amino acid transporter OsAAP6, functioning as a positive regulator of PC was cloned and functionally characterized [21]. GLU is encoded by 15 genes, and the past studies have revealed the mutants of 57H have high concentration of 57 kDa pro-glutelin with floury/opaque endosperm phenotype [22]. Several genes have been isolated from 57H mutants; however, only gpa3, Osvpe1, and OsRab5a have been successfully cloned [23]. The fraction PRO is encoded by 34 gene copies, whereas, only a few loci have been successfully cloned and functionally characterized [24]. Zhang et al. [16] conducted a study and reported 16 QTLs for PC and other fractions of protein.

The SSPs are responsible for determining the nutritional value of rice grains. However, it is also well known that PC influences the cooking quality via protein-starch interaction and can hamper starch gelatinization; therefore, any disorder in the protein’s structure increases the viscosity of rice meal during cooking [25]. Thus, the present study investigated physio-chemical properties of a recombinant inbred lines (RILs) population of 193 lines, derived from inter-sub-specific cross between Japonica rice Nipponbare (NIP) and Indica super rice YK 17, to identify QTLs associated with crude protein and fractions of protein. The detected QTLs were analysed further to understand epistatic effects and their interaction with the environment. The results of the present study displayed the genetic architecture of SSPs in rice, which may help breeders further improve the nutritional quality by marker-assisted selection.

2. Materials and methods

2.1. Plant materials and field experiments

A recombinant inbred lines (RILs) plant population, containing 193 lines, was developed by the single-seed-descent method (SSD) from a cross between genomic sequenced Japonica rice NIP with Indica super rice YK17 as parents (Figure 1). Field experiments were conducted in Hangzhou (HZ) (30° N latitude) in 2017 and 2018, and Hainan province (HN) (18.4° N latitude) in 2017. In HZ 2017 and HZ 2018 field trials, seeds were sown in May, and seedlings were transplanted in June, while in the HN 2017 field trial, seeds were sown in November, and seedlings were transplanted in December. Each plot of plant population consisted of three rows of 21 plants at a spacing pattern of 25 cm (between rows) by 20 cm (within rows). The field trials were arranged in a randomized complete block design (RCBD) with three replications. Irrigation, fertilizer application and other management measures followed standard field production practices. At maturity, each plot of RILs was harvested in bulk and dried naturally. The dried rice grains were stored for three months at room temperature before the evaluation of physio-chemical properties.

2.2. Quality trait evaluation

From each individual line, filled grains were utilized to evaluate grain quality. Hulls were removed from 125 g of grains using a Satake testing husker (THU-35A Satake Engineering, Japan) and de-branned with a McGill number 2 mill (seedburo Equipment, U.S.A.). Milled rice flour samples were obtained by grinding milled rice grains to pass through a 0.42 mm screen on a Udy cyclone mill (Cyclotec 1093 sample mill, Tecator, Sweden). The milled flour samples were sieved through a 100 mesh sieve to get a uniform granule size. The following standard procedures were followed for the traits under investigation.

2.3. Crude protein and fractions of protein

The PC was measured using the micro-Kjeldahl pre-treatment method [26]. Rice protein fractions were...
prepared from rice flour following the process of [27] with minor modifications. The milled rice flour of 1.5 g with three repeats of each line was weighted for all fractions of protein separately; 0.1 M NaOH was used as extraction buffer for GLU. The samples were stirred with 10 ml of extraction buffer at ice water bath for 2 hr; 0.5 M NaCl was used as extraction buffer for GLO; ddH2O was used as extraction buffer for ALB, and 70% n-propanol was used as extraction buffer for PRO protein. The samples of GLO, ALB and PRO were stirred with 10 ml of extraction buffer at room temperature for 4 hr. The procedure was repeated three times. The extracts were stored in the freezer for further analysis. Extracts were separated from residues by centrifugation (10,000 crf, 10 min). The PC of extracts was measured by the semi-micro-Kjeldahl method. The PC and fractions of protein were converted to ammonium nitrogen by sulphuric acid digestion and the absorbance value of the blue production of reaction with natrium salicylicum and hypochlorous acid at the wavelength of 660 nm was checked. The nitrogen contents were measured by employing Rapid Flow Auto Analyzer (AA3, SEAL, Germany). A conversion factor of 5.95 was used to extract PC from the calculated nitrogen content of milled rice flour.

2.4. Linkage map construction

A previously constructed linkage map of RIL population was used in the present investigation. The DNA was extracted through the sodium dodecyl sulphate method from the young leaves of rice seedlings. During linkage map construction, 163 single sequence repeats (SSRs) markers were used with excellent polymorphisms between Nipponbare and YK17. The markers were selected from the public database. MAPMAKER/Exp V 3.0 [28] was used to construct the linkage map, and the recombination rate was converted into the genetic distance (cM) using the Kosambi function. MapDraw V2.1 [29] was used to draw the linkage map based on the obtained linkage data. It spanned a total of 1479.40 cM on all 12 chromosomes with an average interval of 9.08 cM between adjacent markers (Table 1).

2.5. QTL analysis

QTLs controlling PC and fractions of proteins were mapped using Windows QTL Cartographer Version 2.5 (WinQTLCart 2.5) [30] with the composite interval mapping (CIM), and a LOD value of 2.5 was set as the threshold for the detection of putative QTLs. QTLs with epistatic effects and QTL-by-environment interaction (QEs) effects were analysed using QTL Network-2.1 [31] with the mixed-model-based composite interval mapping (MCIM).

2.6. Statistical analysis

All data were analysed using SPSS 22.0 and Excel 2016.

3. Results

3.1. Phenotypic variation in parents and the RIL lines for ECQs

Over three rice growing seasons, the variation in parents and RIL population was significant and broad. Normal frequency distribution of phenotypic data was observed for PC, GLU, GLO, ALB and PRO, indicating the influence of several genes in controlling traits under observation (Table 2 and Figure 2).

3.2. Correlation analysis of traits

The correlation coefficient analysis showed a correlation among all traits under investigation ranged from −0.0793 to 0.5526 (Table 3). A significant to highly significant correlation was observed between PC and fractions of protein except a non-significant correlation between PC to ALB and PC to PRO. A highly significant correlation was observed between PC and GLU in all three (HZ 2017, HN 2017 and HZ 2018) environmental conditions.

| Chr | Total length (Mb) | Physical distance (Mb) | Genetic distance (cM) | No. markers | Coverage (%) | Average interval (cM) |
|-----|-------------------|------------------------|-----------------------|------------|--------------|----------------------|
| 1   | 45.05             | 43.20                  | 177.30                | 18         | 95.89        | 9.85                 |
| 2   | 36.78             | 35.71                  | 134.80                | 18         | 97.09        | 7.49                 |
| 3   | 37.37             | 36.02                  | 190.20                | 22         | 96.39        | 8.65                 |
| 4   | 36.15             | 34.92                  | 125.00                | 14         | 96.60        | 9.45                 |
| 5   | 30.00             | 29.43                  | 160.00                | 16         | 98.10        | 10.00                |
| 6   | 31.60             | 30.77                  | 84.60                 | 11         | 97.37        | 7.69                 |
| 7   | 30.28             | 28.91                  | 113.40                | 12         | 95.48        | 9.07                 |
| 8   | 28.57             | 27.67                  | 127.00                | 10         | 96.85        | 9.07                 |
| 9   | 30.53             | 22.29                  | 78.00                 | 9          | 73.01        | 8.67                 |
| 10  | 23.96             | 22.75                  | 89.00                 | 10         | 94.95        | 8.90                 |
| 11  | 30.76             | 28.02                  | 95.40                 | 9          | 91.09        | 10.60                |
| 12  | 27.77             | 25.22                  | 104.70                | 10         | 90.82        | 10.47                |
| Total| 388.82            | 364.91                 | 1479.40               | 163        | 93.64        | 9.08                 |

Note: Chr = chromosome, Mb = Megabite, cM = Centimorgan.
Figure 2. Phenotypic distribution of PC, GLU, GLO, ALB and PRO in the japonica rice NIP × YK17 RIL population across the three growing conditions. Mean value NIP and YK17 from three environments were shown above with an arrow.

Table 2. Performance of protein and fractions of protein-related traits in parents and their RILs in three cropping seasons.

| Year      | Traits | PC (%) | YK17 (%) | P value | Range | Mean ± SD | Skewness | Kurtosis |
|-----------|--------|--------|----------|---------|-------|-----------|----------|----------|
| 2017 (Hangzhou) | PC (%) | 11.17 ± 0.24 | 7.44 ± 0.09 | 0.0003 | 12.21–5.02 | 9.26 ± 1.35 | 0.17      | 0.05     |
|  | Glu (%) | 5.10 ± 0.02 | 4.53 ± 0.04 | 0.005  | 3.63–9.99  | 6.64 ± 1.47 | 0.34      | 0.71     |
|  | GLO (%) | 0.51 ± 0.02 | 0.39 ± 0.01  | 0.007  | 0.29–1.06  | 0.64 ± 0.18 | 0.17      | 0.05     |
|  | ALB (%) | 0.27 ± 0.021 | 0.14 ± 0.005 | 0.015  | 1.06–0.12  | 0.44 ± 0.21 | 0.48      | 0.61     |
|  | PRO (%) | 0.52 ± 0.02 | 0.31 ± 0.06  | 0.05   | 0.24–1.01  | 0.54 ± 0.14 | 0.38      | 0.17     |
| 2017 (Hainan) | PC (%) | 11.17 ± 0.19 | 7.72 ± 0.20 | 0.002  | 11.76–5.50 | 8.94 ± 1.30 | 0.11      | 0.06     |
|  | Glu (%) | 5.15 ± 0.06 | 4.23 ± 0.04  | 0.008  | 3.12–9.96  | 5.40 ± 1.48 | 0.79      | 0.18     |
|  | GLO (%) | 0.70 ± 0.01 | 0.46 ± 0.003 | 0.003  | 0.20–1.07  | 0.63 ± 0.20 | 0.11      | 0.52     |
|  | ALB (%) | 0.26 ± 0.007 | 0.15 ± 0.014 | 0.01   | 0.90–0.10  | 0.34 ± 0.17 | 0.37      | 0.59     |
|  | PRO (%) | 0.42 ± 0.005 | 0.30 ± 0.014 | 0.0008 | 1.04–0.17  | 0.36 ± 0.14 | 0.43      | 0.61     |
| 2018 (Hangzhou) | PC (%) | 11.08 ± 0.09 | 7.06 ± 0.03 | 0.0003 | 12.93–6.56 | 9.92 ± 1.59 | 0.16      | 0.71     |
|  | Glu (%) | 5.15 ± 0.05 | 4.03 ± 0.04  | 0.008  | 3.31–10.94 | 6.95 ± 1.92 | 0.06      | 0.87     |
|  | GLO (%) | 0.73 ± 0.16 | 0.57 ± 0.01  | 0.005  | 0.36–1.19  | 0.67 ± 0.13 | 0.27      | 0.68     |
|  | ALB (%) | 0.37 ± 0.014 | 0.19 ± 0.007 | 0.004  | 0.65–0.10  | 0.21 ± 0.09 | 0.36      | 0.41     |
|  | PRO (%) | 0.50 ± 0.005 | 0.27 ± 0.005 | 0.0005 | 0.72–0.16  | 0.35 ± 0.11 | 0.96      | 0.55     |

Notes: Data are presented as the mean ± standard deviation (SD). PC = protein content, GLU = Glutelin, GLO = Globulin, ALB = Albumin, PRO = Prolamin.

conditions. The GLU exhibited a highly positive significant to significant positive correlation with ALB, GLO and PRO under all three population growing seasons. Similarly, ALB showed a significant to highly positive significant correlation with GLO and PRO. Meanwhile, GLO showed a highly positive significant to significant
correlation with PC in HZ 2017 and HZ 2018 whereas, it was non-significant in HN 2017. GLO exhibited a highly positive significant to significant positive correlation with PRO under all three (HZ 2017, HN 2017 and HN 2018) population growing seasons.

3.3. Mapping of QTLs

A total of 44 QTLs related to PC, GLU, GLO, ALB and PRO were detected by employing WinQTLCart 2.5 based on CIM with the phenotypic performance of the RIL population (Table 4). The detected QTLs were distributed on all 12 chromosomes, with a single QTL explaining −0.02% to 0.43% of phenotypic variation. Among the 44 detected QTLs, 3 QTLs for GLOB were detected on chromosome 1, whereas 3 QTLs for PC, 3 QTLs for GLU, 3 QTLs for ALB were detected on chromosome 6 under all three population growing seasons (Figure 3).

3.4. QTLs for protein content (PC)

The QTLs for PC, qPC6 were detected under all three environment conditions (HZ 2017, HN 2017 and HZ 2018) with the explained phenotypic variations of 5.24%, 6.63% and 6.50%, respectively. The allele of qPC6 came from Indica YK17 and reduced PC, −0.33%, −0.46% and −0.45% under all three population growing seasons, respectively. The QTL qPC1 harboured with marker interval RM8236–RM5536 in HN 2017 and RM7405–RM128 in HZ 2018 with a phenotypic variation of 6.27% and 5.60%, respectively. The allele from Indica YK17 reduced PC, −0.39% in HN 2017 and −0.40% in HZ 2018, respectively. Moreover, qPC7 was detected under two population growing seasons, harboured with marker interval RM3859–RM11 in HZ 2017 and RM21242-RM5875 in HN 2017 with the explained phenotypic variations of 5.78% and 3.70%, respectively. The allele of Japonica NIP increases qPC7, 0.34% in HZ 2017 and 0.36% in HN 2017, respectively. The qPC6 peak of the LOD score covered the Wx gene region, predicting qPC6 might correspond to the Wx gene. Moreover, qPC1, detected under HN 2017 and HZ 2018, encodes a putative amino acid transporter OsAAP6, functioning as a positive regulator of PC. The QTL, qPC7 was detected under two population growing seasons HZ2017 and HN 2017 with the explained phenotypic variances of 5.78% and 3.70%, respectively. The positive additive effect of qPC7 was contributed by the allele of Japonica NIP under both environmental conditions. The other two QTLs, qPC5 and qPC11, were detected under HN 2017 and HZ 2018 with the explained phenotypic variances of 6.61% and 7.26%, respectively. The negative additive effect of both of these QTLs was contributed by the allele of Indica YK17.

3.5. QTLs for glutelin content (GLU)

Eight QTLs for GLU were identified under all three growing seasons located on chromosomes 1, 4, 5, 6, 8 and 10. The explained phenotypic variation contributed by individual QTLs ranged from 5.36% to 9.79%. Among the detected QTLs, qGLU6 was significant under all three (HZ 2017, HN 2017 and HZ 2018) population growing seasons with the phenotypic variations of 7.82%, 3.70% and 5.47%, respectively. The allele of qGLU6 came from Indica YK 17 and reduced GLU −0.49%, −0.40% and −0.50% under all three population growing seasons, respectively. The other QTLs qGLU4 detected in HZ 2017, qGLU8 detected in HN 2017, whereas qGLU5 and qGLU10 were only significant in HZ 2018. The QTLs, qGLU4, qGLU8, qGLUS and qGLU10 explained 6.14%, 9.79%, 7.93% and 5.36% phenotypic variance, respectively. The positive additive effect of qGLU4 and qGLU10 was contributed by the allele of Japonica NIP, whereas qGLU8 and qGLUS were decreased by the allele of Indica YK17.

3.6. QTLs for globulin content (GLO)

Ten QTLs were detected for GLO under all three studied population growing seasons. The explained phenotypic variation contributed by individual QTLs ranged from 5.10% to 12.47%. The QTL, qGLO1 was detected under all three (HZ 2017, HN 2017 and HZ 2018) population growing seasons harboured with markers interval RM6738–RM3738 with the explained phenotypic variations of 12.47%, 6.18% and 6.18%, respectively. The positive additive effect of qGLO1 was contributed by the allele of Japonica NIP. Another QTL, qGLO5 was detected under HZ 2017 and HZ 2018 environmental conditions with the explained phenotypic variations of 11.55% and 5.95%, respectively. The positive additive effect was contributed by the Japonica NIP in HZ 2017, whereas the allele from Indica YK17 contributed negative allele under HZ 2018 environmental condition. The QTLs,

| Table 3. Coefficients of pairwise correlation for protein and fractions of protein from an RIL population derived from the cross of NIP × YK17. |
|-----------------|-----------------|-----------------|-----------------|
| PC              | GLU             | ALB             | GLO             |
| GLU             | 0.5526**        | 0.3429**        | 0.1593**        |
| ALB             | 0.1337          | 0.3873**        | 0.0585          |
| GLO             | 0.2300**        | 0.2029**        | 0.1379**        |
| PRO             | 0.1408          | 0.3642**        | 0.2066**        |

Notes: a, b, and c represent (HZ 2017), (HN 2017) and (HZ 2018), respectively. * and ** are significant at 0.05 and 0.01 levels, respectively.
Table 4. QTLs for protein content and fractions of protein in the RIL population of NIP × YK17.

| Trait | QTL | Chr. | Interval | LOD | Var.% | Add. |
|-------|-----|------|----------|-----|------|------|
| HZ 2017 |     |      |          |     |      |      |
| PC    | qPC6| 6    | RM7158–RM3414 | 2.50 | 5.24 | −0.33 |
| GLU   | qGLU4| 4    | RM17303–RM17377 | 2.81 | 6.14 | 0.43 |
| GLO   | qGLO6| 6    | RM1758–RM3414 | 3.60 | 7.82 | −0.49 |
| ALB   | qALB3| 3    | RM282–RM6446 | 3.73 | 10.11 | −0.03 |
| PRO   | qPRO1| 1    | RM600–RM3341 | 3.11 | 7.80 | −0.04 |
| HN 2017 |     |      |          |     |      |      |
| PC    | qPC1| 1    | RM8236–RM5536 | 2.80 | 6.27 | −0.39 |
| GLU   | qGLU5| 5    | RM17852–RM165 | 3.15 | 6.61 | −0.40 |
| GLO   | qGLO8| 8    | RM3414–RM276 | 2.91 | 5.70 | −0.40 |
| ALB   | qALB5| 5    | RM529–RM18065 | 3.30 | 8.31 | −0.05 |
| PRO   | qPRO4| 4    | RM7585–RM6659 | 2.88 | 6.30 | 0.05 |
| HZ 2018 |     |      |          |     |      |      |
| PC    | qPC1| 1    | RM7405–RM128  | 2.84 | 5.60 | −0.40 |
| GLU   | qGLU5| 5    | RM2699–RM27186 | 3.59 | 7.26 | −0.46 |
| GLO   | qGLO10| 10  | RM1024–RM1805 | 3.92 | 7.95 | −0.54 |
| ALB   | qALB3| 3    | RM590–RM5907  | 3.13 | 5.95 | 0.03 |
| PRO   | qPRO2| 2    | RM5812–LM3763 | 4.18 | 8.84 | −0.04 |
| ALB   | qALB6| 6    | RM3414–RM5963 | 4.86 | 11.40 | −0.05 |
| ALB   | qALB8| 8    | RM310–RM22694 | 3.90 | 7.60 | 0.04 |

$qGLO12$ was identified under HZ 2017 and HN 2017 population growing seasons with the explained phenotypic variations of 5.10% and 7.25%, respectively. The positive additive effect was derived from $Japonica$ NIP for both of these QTLs. The three QTLs, $qGLO11$ detected in HN 2017 and $qGLO6$, $qGLO10$ detected in HZ 2018, with the phenotypic variations of 11.52%, 10.00% and 6.03%, respectively. The allele from $Indica$ YK17 decreased GLO, −0.05% in $qGLO6$ and −0.06% in $qGLO11$, whereas the additive effect from $Japonica$ NIP increases 0.03% GLO in $qGLO10$.

3.7. QTLs for albumin content (ALB)

A total of nine QTLs associated with ALB were detected under all three population growing seasons, with the explained phenotypic variation by individual QTL ranged from 5.53% to 10.11%. The QTL, $qALB6$ was detected under all three population growing seasons (HZ 2017, HN 2017 and HZ 2018) harboured with marker interval RM7158–RM3414 with the explained phenotypic variations of 6.19%, 7.23% and 5.53%, respectively. The allele of $qALB6$ came from $Indica$ YK17 and reduced ALB, −0.49%, −0.40% and −0.50% under all three population growing seasons, respectively. The QTL, $qALB3$ was detected under HZ 2017 and HZ 2018 with the phenotypic variations of 10.11% and 6.73%, respectively. The allele from $Indica$ YK17 decreased ALB from −0.03% and −0.01% separately in HZ 2017 and HN 2018. The other four QTLs, $qALB10$ in HZ 2017, $qALB5$ and $qALB9$ in HN 2017, and $qALB4$ in HZ 2018 explained phenotypic variation of 9.91%, 8.31%, 6.50% and 6.19%, respectively. The positive additive effect of $qALB9$ and $qALB4$ was contributed from the allele of $Japonica$ NIP, whereas the allele for $qALB10$ and $qALB5$ came from $Indica$ YK17.

3.8. QTLs for prolamin content (PRO)

Nine QTLs were detected under all three (HZ 2017, HN 2017 and HZ 2018) population growing seasons, with
the explained phenotypic variation by individual QTL ranged from 5.19% to 11.40%. The QTL, qPRO1 was detected flanked with marker interval RM600–RM3341 under HZ 2017 and RM8068–RM3746 under HN 2017, with the contributed phenotypic variation of 7.80% and 5.19%, respectively. The allele from Indica YK17 reduced qPRO1, −0.04% and −0.03%, respectively. The QTL, qPRO6 was detected under HZ 2017 and HZ 2018 with the explained phenotypic variation of 6.85% and 11.40%, respectively. The allele from Indica YK17 reduced qPRO6, −0.04% and −0.05%, respectively. Moreover, QTL, qPRO8 was detected under HN
2017 and HZ 2018 with the explained phenotypic variation of 5.76% and 7.60%, respectively. The allele from Indica YK17 came in HN 2017, whereas, the positive additive effect was contributed by Japonica NIP in HZ 2018. Other QTLs, qPRO4 and qPRO12 in HZ 2017 and qPRO8 in HZ 2018 contributed phenotypic variation of 7.80%, 7.85% and 8.84%, respectively. The allele came from Indica YK17 for all these three QTLs.

3.9. Detection of QTLs with additive × environment and epistasis interactions

To understand the genetic architecture of protein content and fractions of protein attributes, the digenic epistatic effects of PC, GLU, GLO, ALB and PRO were estimated. Three QTLs (qPC6, qPC7 and qGLU6) were detected by the joint analysis of PC, GLU, GLO, ALB and PRO under all three population growing seasons, and under non-significant additive × environment interaction, respectively (Table 5). The phenotypic contribution rate was relatively higher. This suggested that the QTL expressions of PC and GLU were influenced by the environmental conditions. The epistatic interaction also played an important role in determining rice grain quality. So, to elaborate understanding for genetic components of these attributes, five locus bi-allelic epistatic interactions were estimated (Table 6). The attributes GLOB and PRO showed two pairs of epistatic loci; however, PC, GLU, ALB exhibited non-significant epistasis effect indicating M-QTLs controlling PC, GLU, ALB. GLUB revealed one pair of epistatic loci, which elucidated 1.66% of phenotypic variation. One pair of epistatic loci was detected for PRO accounted for 7.41% of the phenotypic variation.

4. Discussion

Rice provides a large caloric and nutrition demand to approximately half of the world population [32]. The improvement in rice grain protein quality is possible only by regulating SSP contents. The genetic architecture of SSPs is much complexed, multigenic and with a mixture of highly polymorphic polypeptides. The mutation in some structural genes has little or no effect on the overall grain’s protein composition [17,19]. Similarly, the significance of major and minor genes influenced by epistatic and environmental interaction are also noteworthy. Therefore, the combined utilization of conventional and molecular breeding techniques, e.g. marker-assisted selection (MAS), may be the effective and reliable technique in harnessing the SSP concentration in rice grain [16,33]. Moreover, the contemporary studies designed to identify QTLs for rice grain protein content have elaborated the genetic pillars of protein content [34]. The interrelationship among fractions of SSPs is complicated requiring comprehensive research. In the present investigation, PC, GLU, GLO, ALB and PRO were analysed to detect M-QTLs, epistatic QTLs and QEs association under three population growing seasons.

PC is an integral component for improving nutritional quality and palatability of cooked rice [25]. The synergistic relationship of amylose with PC determines that the rice with less than 7% protein is good in taste [35]. The QTL qPC6 detected repeatedly on the marker position near the Wx locus located on chromosome 6 confirmed the results derived from several other mapping populations [20,36]. Similarly, the qPC1, a major QTL for PC on chromosome 1 was detected under HN 2017 and HZ 2018 population growing seasons, suggesting the influence of qPC1 for rice grain protein contents. However, the qPC7 noticed under two population (HZ 2017 and HN 2017) growing seasons needs further investigation, and it can be a potential gene controlling PC along with qPC1 and qPC6. GLU is the highest in concentration and contains more essential amino acids, especially lysine, for human consumption compared to other fractions of protein. A QTL, qGLU6 was

### Table 5. QTLs with additive × environment interaction for PC, GLU, GLO, ALB and PRO detected in the RIL population under three trials.

| Trait/QTL | Marker Interval | Marker (Position, cM) | Range | A² | AE1 | AE2 | AE3 | R² (%) |
|-----------|-----------------|-----------------------|-------|----|-----|-----|-----|--------|
| **PC**    |                 |                       |       |    |     |     |     |        |
| qPC6      | RM7158–RM190    | 0.0                   | 0.0–4.0|−0.3137|0.0000|0.0000|−0.0002|4.20    |
| qPC7²      | RM3859–RM214    | 59.4                  | 54.3–65.8|0.3464|0.0000|−0.0003|0.0003|5.12    |
| qGLU6      | RM3414–RM4276   | 11.9                  | 7.2–15.9|−0.4992|0.0001|0.0000|−0.0001|7.36    |

Notes: A², Significant additive effects contributed by QTL mapped in the environments. AE1, AE2, and AE3 represent the additive effects QTL × environment interaction in three locations (HZ 2017, HN 2017 and HZ 2018).

²QTLs were also detected in the CIM analysis.

### Table 6. Epistatic interactions for PC, GLU, GLO, ALB and PRO detected in the RIL population under three trials.

| Trait | Chr | Marker Interval | Chr | Marker Interval | a₁a₂ | R² (%) |
|-------|-----|-----------------|-----|-----------------|------|--------|
| GLO   | 3   | RM14574–RM7197  | 5   | RM3790–RM3664   | −0.023|1.66    |
| PRO   | 5   | RM18053–RM165  | 5   | RM18349–RM18457 | −0.047|7.41    |

Notes: a₁a₂, Additive × additive epistatic effect, a positive value indicates parental type > recombinant type; a negative value indicates parental type < recombinant type.
detected repeatedly along the same genomic region as qPC6, with a negative additive effect contributed from Indica YK17. Moreover, a highly significant correlation \((r = 0.5526***)\) between PC and GLU was observed in the RIL population. These results indicated that PC and GLU might share the exact genetic mechanism. GLO is concentrated in the embryo and the outer aleurone layer of the endosperm, and a major portion of this protein is removed during milling. Only a few studies were conducted to clone and characterize the genes controlling GLO [24]. It is noteworthy that qGLO1 was detected as a major gene for GLO under all three population growing seasons with the same marker interval. It can be assumed that the qGLO1 globulin gene might be under a regulatory mechanism different from those of GLU, PRO and ALB, due to its unique genetic organization. The ALB protein is associated with the allergen proteins similar to GLO. The QTL, qALB6 was found as a major determinant for ALB under all three population growing seasons with same marker interval of qPC6 and qGLU6. Chen et al. [34] reported the influence of Wx gene for controlling ALB. They found 3.3 kb Wx pre-mRNA is positively correlated with ALB, providing new insights into the genetic basis of rice grain quality. In addition, qPRO6 QTLs were detected for the alcohol-soluble PRO protein. But, qPRO6 shares the same genomic region as qPC6, qGLU6 and qALB6, indicating the influence of Wx gene region to control PRO. These results indicated the simultaneous influence of major genes to control the fraction of protein, whereas the highly positive significant correlation of fraction of protein (GLU, GLO, ALB and PRO) with each other indicated the partially common genetic mechanism, which is consistent with the fact that these traits are related [37]. The non-significant correlations of ALB and PRO for PC contradict with the findings reported by [16]. In other studies, it was found that ALB was negatively correlated with PC, and there was no correlation between GLO and PC [16,38]. The discrepancy of present results may be due to the differences in germplasm evaluated, experiment location, and environmental influence. But, the QTLs for PC, GLO and PRO were observed at different chromosomes; it could be helpful to fine map underlying alleles controlling these attributes. These identified loci for PC, GLO and PRO can further introgress into elite cultivars via marker-assisted selection (MAS) to enhance the nutritional and other specific objectives. In the present investigation, the genetic basis for protein and fraction of proteins (GLU, GLO, ALB and PRO) was obtained by analysing QEs and epistatic QTLs. The QEs interactions had shown phenotypic contribution rates are relatively higher, suggesting that the QTLs expressions of PC and GLU were influenced by the environmental conditions except for GLO, ALB and PRO, and showed the influence of M-QTL was dominant over QEs. For epistatic interaction, GLO and PRO were observed for epistatic interaction excluding PC, GLU and ALB. The sum effect of detected epistatic QTLs for GLO and PRO was higher than M-QTLs, suggesting the influence of the epistatic effect of one QTL over another QTL.

In conclusion, a better grain quality could be achieved by regulating the SSP content. The present investigation revealed the significance of minor QTLs, epistatic QTLs and QE interactions and M-QTLs on rice grain quality improvements programmes. The genetic mechanism of quantitative regulation for PC and fractions of protein in rice is inter-correlated, which is evident from the co-localization of QTLs responsible for these SSPs. Future genetics based studies, such as fine mapping of novel QTLs, interaction studies of genes and searching for non-environment-specific QTL, are required to elucidate the genetic mechanism of quantitative regulation and obtain DNA markers tightly linked to the desirable QTL to facilitate MAS in rice breeding for high nutritional quality.

Acknowledgements

ST conceived and designed the experiment. ZS prepared materials. SF and HNB harvested germplasm. SF performed the experiments. SF, AZ, TS and UA analyzed the data. ST critically revised the article. All authors have read and approved the final manuscript.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Natural Science Foundation of Zhejiang province [LR20C13002], the special support plan for high level talents in Zhejiang Province [2019R52032].

ORCID

Sajid Fiaz https://orcid.org/0000-0001-9097-4359
Zhonghua Sheng https://orcid.org/0000-0002-5960-8307
Hirenda Nath Barman https://orcid.org/0000-0003-2601-3048
Umed Ali https://orcid.org/0000-0003-1748-784X
Shaoqing Tang https://orcid.org/0000-0001-5532-9784

References

[1] Anis G, Zhang Y, Xu X, et al. QTL analysis for rice seedlings under nitrogen deficiency using chromosomal segment substitution lines. Pak J Bot. 2018;50(2):537–544.
[2] Fiaz S, Ahmad S, Riaz A, et al. Applications of the CRISPR/Cas9 system for rice grain quality improvement: perspectives and opportunities. Int J Mol Sci. 2019;20(4):888.
[3] Fiaz S, Jiao G, Sheng Z, et al. Analysis of genomic regions affecting cooking and eating quality traits of a recombinant inbred population in rice. Inter J Agri Biol. 2019;22:611–619.
[4] Balyan HS, Gupta PK, Kumar S, et al. Genetic improvement of grain protein content and other health-related constituents of wheat grain. Plant Breed. 2013;132(5):446–457.
[5] Shar T, Sheng Z, Ali U, et al. Mapping quantitative trait loci associated with paste viscosity attributes in double haploid population of rice (Oryza sativa L.). J Integr Agr. 2019;18(0):2–14.

[6] Jiao G, Wei X, Tang S, et al. Stirring affects starch granule morphology and the functional properties of rice flour. Starch-Stärke. 2018;70:1700247.

[7] Ufaz S, Galli G. Improving the content of essential amino acids in crop plants: goals and opportunities. Plant Physiol. 2008;147:954–961.

[8] Kim Y, Lee J, Lee T, et al. The suppression of the glutelin storage protein gene in transgenic rice seeds results in a higher yield of recombinant protein. Plant Biotechnol Rep. 2012;6(4):347–353.

[9] Fitzgerald MA, McCouch SR, Hall RD. Not just a grain of rice: the quest for quality. Trends Plant Sci. 2009;14(3):133–139.

[10] Li X, Wu Y, Zhang DZ, et al. Rice prolamine protein body biogenesis: a BIP-mediated process. Science. 1993;262(5136):1054–1056.

[11] Shewry P. Plant storage proteins. Biol Rev. 1995;70:375–426.

[12] Mckevith B. Nutritional aspects of cereals. Nutr Bull. 2004;29(8):111–142.

[13] Yamagata H, Sugimoto T, Tanaka K, et al. Biosynthesis of storage proteins in developing rice seeds. Plant Physiol. 1982;70:1094–1100.

[14] Kim Y, Ahn Y, Lim Y, et al. Daily nutritional dose supplementation with antioxidant nutrients and phytochemicals improves DNA and LDL stability: a double-blind, randomized, and placebo-controlled trial. Nutrients. 2013;5(12):5218–5232.

[15] Kawakatsu T, Yamamoto MP, Hirose S, et al. Characterization of a new rice glutelin gene GluD-1 expressed in the starchy endosperm. J Exp Bot. 2008;59(15):4233–4245.

[16] Zhang W, Bi J, Chen L, et al. QTL mapping for crude protein and protein fraction contents in rice (Oryza sativa L.). J Cereal Sci. 2008;48(2):539–547.

[17] Shewry PR. Improving the protein content and composition of cereal grain. J Cereal Sci. 2007;46(3):239–250.

[18] Morita R, Kusaba M, Iida S, et al. Development of PCR markers to detect the glb1 and Lgc1 mutations for the production of low easy-to-digest protein rice varieties. Plant Physiol. 2010;154(5):812–824.

[19] Kawakatsu T, Hirose S, Yasuda H, et al. Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. Plant Physiol. 2010;151(1):164343.

[20] Lou J, Chen L, Yue G, et al. QTL mapping of grain quality traits in rice. J Cereal Sci. 2009;50(2):145–151.

[21] Peng B, Kong H, Li Y, et al. OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. Nat Commun. 2014;5:4847.

[22] Ren Y, Wang Y, Liu F, et al. GLUTELIN PRECURSOR ACCUMULATION3 encodes a regulator of post-Golgi vesicular traffic essential for vacuolar protein sorting in rice endosperm. Plant Cell. 2014;113:121376.

[23] Wang Y, Ren Y, Liu X, et al. Osrab5a regulates endomembrane organization and storage protein trafficking in rice endosperm cells. Plant J. 2010;64(5):812–824.

[24] Bhullar NK, Gruissem W. Nutritional enhancement of rice for human health: the contribution of biotechnology. Biotechnol Adv. 2013;31(1):50–57.

[25] Yu TQ, Jiang W, Ham T, et al. Comparison of grain quality traits between japonica rice cultivars from Korea and Yunnan province of China. J Crop Sci Biotechnol. 2008;11:135–140.

[26] CBS (Chinese Bureau of Standardization). Method for determination of crude protein in cereals and bean seeds (semi-micro-Kjeldahl method). GB 2905e32, UDC 633.1.8. revealed by genome-wide association analysis. Front Plant Sci. 1982;9:593–601.

[27] Kumamaru T, Satoh H, Iwata N, et al. Mutants for rice storage proteins: 1. Screening of mutants for rice storage proteins of protein bodies in the starchy endosperm. Theor Appl Genet. 1988;76(1):11–16.

[28] Lincoln SE, Daly MJ, Lander ES. (1993). Constructing genetic linkage maps with MAPMAKER/EXP version 3.0: a tutorial and reference manual. A whitehead institute for biomedical research technical report, 3.

[29] Liu RH, Mep JL. Mapdraw: a micro soft excel macro for drawing genetic linkage maps based on given genetic linkage data. Hereditas. 2003;25:317–321.

[30] Wang S. (2007). Windows QTL cartogapher 2.5. Available from: http://statgen.ncsu.edu/qtlcart/WQTLCart.html.

[31] Yang J, Hu C, Hu H, et al. QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. Bioinformatics. 2008;24(5):721–723.

[32] Barman HN, Sheng Z, Fiaz S, et al. Generation of a new thermo-sensitive genic male sterile rice line by targeted mutagenesis of TMS5 gene through CRISPR/Cas9 system. BMC Plant Biol. 2019;19:109.

[33] Sheng Z, Fiaz S, Li Q, et al. Molecular breeding of fragrant early-season hybrid rice using the BADH2 gene. Pak J Bot. 2019;51:2089–2095.

[34] Chen P, Shen Z, Ming L, et al. Genetic basis of variation in rice seed storage protein (albumin, globulin, prolamin, and glutelin) content revealed by genome-wide association analysis. Front Plant Sci. 2018;9:612.

[35] Lee GH, Yun BW, Kim KM. Analysis of QTLs associated with the rice quality related gene by double haploid populations. Int J Genomics. 2014;2014:781832.

[36] Yu YH, Li G, Fan YY, et al. Genetic relationship between grain yield and the contents of protein and fat in a recombinant inbred population of rice. J Cereal Sci. 2009;50(1):121–125.

[37] Nakase M, Hotta H, Adachi T, et al. Cloning of the rice seed α-globulin-encoding gene: sequence similarity of the 5′-flanking region to those of the genes encoding wheat high-molecular-weight glutenin and barley D hordein. Gene. 1996;170(2):223–226.

[38] Sturgis FE. (1952). Protein in rice as influenced by variety and fertilizer levels. LSU Agricultural Experiment Station Reports. 94.