

## RESEARCH ARTICLE

## Mapping QTLs underpin nutrition components in aromatic rice germplasm

M. Z. Islam<sup>1\*</sup>, M. Arifuzzaman<sup>2</sup>, S. Banik<sup>3</sup>, M. A. Hossain<sup>4</sup>, J. Ferdous<sup>5</sup>, M. Khalequzzaman<sup>1</sup>, B. R. Pittendrigh<sup>6</sup>, M. Tomita<sup>7\*</sup>, M. P. Ali<sup>8\*</sup>

**1** Genetic Resources and Seed Division, Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh, **2** Department of Genetics and Plant Breeding, Hajee Mohammad Danesh Science and Technology University, Dinajpur, Bangladesh, **3** Grain Quality and Nutrition Division, Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh, **4** Regional Station, Barisal, Bangladesh Rice Research Institute, Gazipur, Bangladesh, **5** Biotechnology Division, Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh, **6** Department of Entomology, Michigan State University, East Lansing, Michigan, United States of America, **7** Genetics and Genome Engineering Laboratory, Green Biology Research Division, Research Institute of Green Science and Technology, Shizuoka University, Shizuoka City, Shizuoka, Japan, **8** Entomology Division, Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh

\* [zahid.grs@gmail.com](mailto:zahid.grs@gmail.com) (MZI); [panna\\_ali@yahoo.com](mailto:panna_ali@yahoo.com) (MPA); [tomita.motonori@shizuoka.ac.jp](mailto:tomita.motonori@shizuoka.ac.jp) (MT)



## OPEN ACCESS

**Citation:** Islam MZ, Arifuzzaman M, Banik S, Hossain MA, Ferdous J, Khalequzzaman M, et al. (2020) Mapping QTLs underpin nutrition components in aromatic rice germplasm. PLoS ONE 15(6): e0234395. <https://doi.org/10.1371/journal.pone.0234395>

**Editor:** Evangelia V. Avramidou, Institute of Mediterranean Forest Ecosystems of Athens, GREECE

**Received:** November 27, 2019

**Accepted:** May 22, 2020

**Published:** June 11, 2020

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0234395>

**Copyright:** This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

**Data Availability Statement:** The data underlying the results presented in the study are available within manuscript.

## Abstract

As rice is an important staple food globally, research for development and enhancement of its nutritional value it is an imperative task. Identification of nutrient enriched rice germplasm and exploiting them for breeding programme is the easiest way to develop better quality rice. In this study, we analyzed 113 aromatic rice germplasm in order to identify quantitative trait loci (QTL) underpinning nutrition components and determined by measuring the normal frequency distribution for Fe, Zn, amylose, and protein content in those rice germplasm. Comparatively, the germplasm Radhuni pagal, Kalobakri, Thakurbhog (26.6 ppm) and Hatisail exhibited the highest mean values for Fe (16.9 ppm), Zn (34.1 ppm), amylose (26.6 ppm) and protein content (11.0 ppm), respectively. Moreover, a significant linear relationship ( $R^2 = 0.693$ ) was observed between Fe and Zn contents. Cluster analysis based on Mahalanobis  $D^2$  distances revealed four major clusters of 113 rice germplasm, with cluster III containing a maximum 37 germplasm and a maximum inter-cluster distance between clusters III and IV. The 45 polymorphic SSRs and four trait associations exhibited eight significant quantitative trait loci (QTL) located on eight different chromosomes using composite interval mapping (CIM). The highly significant QTL (variance 7.89%, LOD 2.02) for protein content (*QTL.pro.1*) was observed on chromosome 1 at 94.9cM position. Also, four QTLs for amylose content were observed with the highly significant *QTL.amy.8* located on chromosome 8 exhibiting 7.2% variance with LOD 1.83. Only one QTL (*QTL.Fe.9*) for Fe content was located on chromosome 9 (LOD 1.24), and two (*QTL.Zn.4* and *QTL.Zn.5*) for Zn on chromosome 4 (LOD 1.71) and 5 (LOD 1.18), respectively. Overall, germplasm from clusters III and IV might offer higher heterotic response with the identified QTLs playing a significant role in any rice biofortification breeding program and released with development of new varieties.

**Funding:** National Agricultural Technology Project (NATP): Phase-1 in Bangladesh. Financed by Bangladesh Agricultural Research Council (BARC), Farmgate, Dhaka. Grant No. 12. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.”. Lead author MZ Islam received fellowship from this grant. M. Khalequzzaman received project award”. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## 1. Introduction

While nutritious crop varieties (including those enriched with micronutrients) represent a crucial demand for healthy growth and development of human beings, few micronutrient-enriched food crop varieties exist among global staple cultivated crop varieties. In particular, iron (Fe) is an essential element for hemoglobin with zinc (Zn) also providing a key co-factor for more than three hundred enzymes involved in principal biological activities [1, 2]. Deficiencies of Fe and Zn cause a wide range of health problems, including but not limited to anemia, reduced growth, poor cognitive development, stunting, reduced immunity, diarrhea, lesions on eyes and skin, delayed healing of wounds, and mental lethargy [3, 4]. Currently, more than two billion people in the world—especially children and pregnant or lactating women—suffer from Fe and Zn deficiency diseases [5]. And while micronutrient enriched supplementary drugs are available to remediate these problems, such drugs are typically prescribed by doctors only after micronutrient deficiency symptoms appear and are often prohibitively expensive [6].

According to FAO, more than 90% of Asian people depend on rice as food, delivering micronutrients via rice would afford these missing micronutrient benefits in a more cost-effective and wider-reaching way [7, 8]. “Biofortification”—or, the hybridization of food crops with rich vitamin and mineral densities—represents one powerful and cost-effective tool for providing these micronutrients and for reducing or eliminating micronutrient-related malnutrition [7–9]. In general, although consumers prefer milled rice, which unfortunately provides an only limited source of Fe and Zn micronutrients, research to identify rice germplasm with higher micronutrient content can make rice biofortification breeding programs possible [10, 11]. A biofortification breeding program was succeeded for Zn-enriched rice varieties in Bangladesh, India, and Philippines [12]. However, amylose content in rice is also an important determinant for maintaining the eating and cooking quality of rice, as determined by three principal physicochemical properties: amylose content (AC), gel consistency (GC), and gelatinization temperature (GT). In general, rice growers prefer high amylose content rice, such that amylose-categorized rice germplasm would be required for any rice biofortification program to match market viability and grower and consumer preferences. Protein is an essential component of diet which provides basic function in nutrition and supply adequate amounts of needed amino acids in the body. Therefore, enhancing protein content in rice is an important criterion for consumer preference for ensuring healthy life [13].

Micronutrient Fe, Zn, amylose, and protein content varies widely among rice germplasm, landraces, cultivars, breeding lines, and wild rice. Among these, aromatic rice germplasms constitute a small but important group of rice genotypes that are popular (and increasingly in demand) in many countries of the world for their aroma and/or super-fine grain quality [14]. As such, aromatic rice represents a strategic choice for biofortification and is also readily available and popular in Bangladesh, where it comprises short and medium bold types with a mild to strong aroma [15, 16]. Although genetic diversity and population structure analysis based on phenotypic and genotypic characteristics of aromatic rice have been reported [17, 18], the micronutrient content for Fe, Zn, amylose, and protein content have not yet been analyzed.

The rice varieties enriched with micronutrients like iron (Fe), zinc (Zn), and/or other vitamins are considered better quality rice. Recently, demands for this type of rice has increased as a way to offset micronutrient deficiencies both in children and older people. While Fe- and Zn-rich rice varieties have been developed through conventional breeding approaches and cultivation in rice-growing countries, several issues—like complex genetic backgrounds, sophisticated phenotyping, lack of knowledge about molecular markers, and interactions between germplasm and environmental [19, 20]—make such an approach time-consuming to meet

growing demand and climate change effects. In contrast, identification of quantitative trait loci (QTLs) and genes for high Fe, Zn, amylose, and protein content could lead to more results more quickly. Furthermore, these genes or loci could be introgressed into cultivar backgrounds precisely through molecular marker assisted breeding (MAB) and genetic engineering approaches to enable early release of rice varieties [21].

QTL is a locus, or a part, of DNA estimated from quantitative trait and molecular marker data in a population. To map QTLs or genes, bi-parental mapping and genome-wide association studies (GWAS) have been applied to a number of rice traits [20–24]. However, GWAS have several merits over bi-parental mapping, including high resolution and scanning for numerous and rare alleles [25]. GWAS have been utilized to identify QTLs for Fe, Zn, and several other mineral elements in rice seeds [20, 24]. In addition, 31 putative QTLs have been identified for Fe and Zn (as well as Mn, Cu, Ca, Mg, P, and K) contents [26]. In one study, 14 QTLs for Fe and Zn and candidate genes *OsYSL1* and *OsMTP1* for Fe and *OsARD2*, *OsIRT1*, *OsNAS1*, and *OsNAS2* for Zn were identified in rice seeds [27].

While mapping of QTLs for Fe and Zn concentrations in non-aromatic rice has been performed, limited studies have investigated Fe and Zn concentrations (and amylose and protein content) in aromatic rice. The purpose of this study, therefore, was: (1) to identify high yield genotypes for Fe, Zn, amylose, and protein content in 113 aromatic rice germplasm, (2) to cluster the genotypes based on the studied traits, and (3) to identify QTLs for Fe, Zn, amylose, and protein content using molecular marker-trait associations studies.

## 2. Results

### 2.1 Mean comparison and frequency distribution

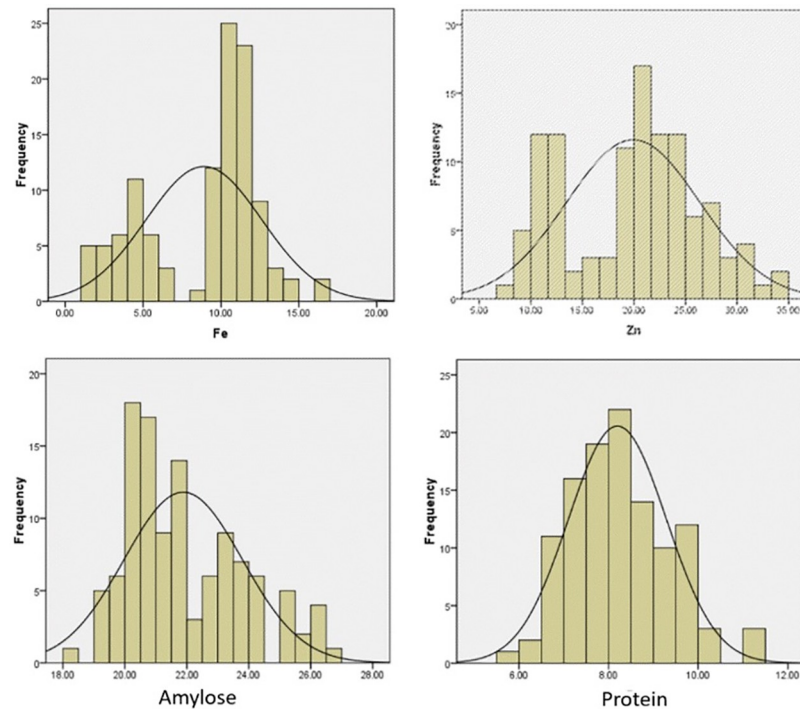
Fig 1 and Table 1 present frequency distributions and means comparison of 113 aromatic germplasm for four phenotypic traits (Fe, Zn, amylose, and protein content). Fig 1 depicts a normal frequency distribution, along with a wider phenotypic variation, for grain-Fe content ( $8.90 \pm 3.72$  ppm, range 1.10–16.90). The highest number of genotypes (25) occupied the 10–11 range followed by 11–12. The Zn content also showed a normal frequency distribution with an average  $19.93 \pm 6.48$  ppm, with the highest number of genotypes (22) occupying the 20–21.66 range followed by 10.00–13.33 and 21.67–25. Overall, the Radhuni pagal genotype exhibited the maximum Zn content and the Luina genotype the minimum.

Amylose content exhibited a wider range, from 18.20–26.60, with an average for the 113 genotypes of  $21.88 \pm 1.91$ . The highest number of genotypes (17) occupied the 20.00–21.25 range followed by 21.25–22.5, with the genotype Thakurbhog exhibited the maximum amylose content and the genotype Jirabuti the minimum. The least significant difference (LSD) (5%) was estimated at 0.35 considering all genotypes for the trait.

Protein content also showed a normal frequency distribution, with an average for the 113 genotypes of  $8.19 \pm 0.10$  ppm. The highest number of genotypes (22) occupied the 8.00–8.5 range, with genotype Hatisail exhibiting the maximum protein content (11.00) and Sadagura (SI-104) the minimum (6.30). The Zn content ranged from 7.30 to 34.10 with an average of  $19.93 \pm 0.61$  ppm, and the genotype Kalobakri exhibiting the maximum.

### 2.2 Regression analysis

The response function for the relationship between Zn concentration and Fe concentration (shown in Fig 2) exhibited a strong, highly significant linear relationship ( $R^2 = 0.693$ ) between both. Zn content (ppm) was positively affected by Fe concentration and showed a strong relationship with Fe ( $F = 250.73$ ,  $p < 0.001$ ), e.g., with increasing Zn content, the Fe content increases and vice-versa.



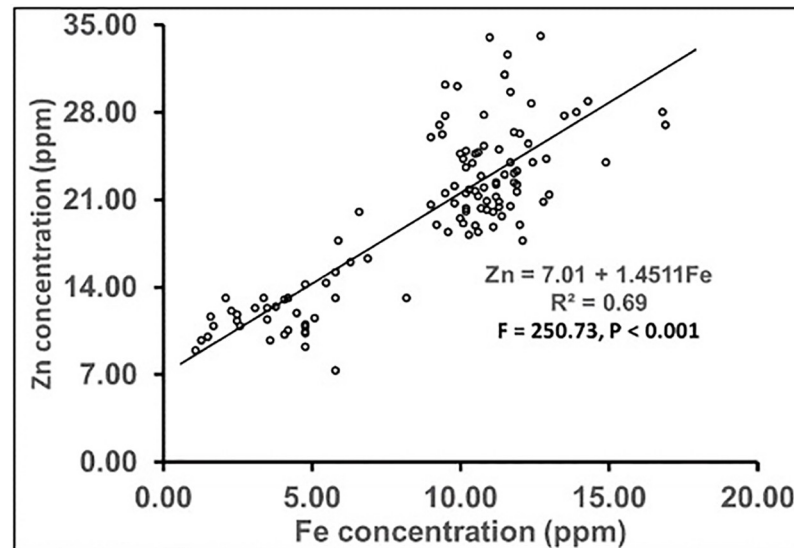
**Fig 1. Frequency distribution of 113 aromatic rice germplasm for four traits.**

<https://doi.org/10.1371/journal.pone.0234395.g001>

**Table 1. Mean values of selected superior and inferior lines for studied traits.**

Variety/ Line	Amylose	Protein	Zn	Fe
<b>Selected seven superior genotypes</b>				
Kalobakri	23.80	8.22	34.10	12.70
Sakkorkhana	20.70	9.50	34.00	11.00
Hatisail	20.40	11.00	32.60	11.60
Rajbut	20.60	9.10	27.80	10.80
Sakkor khora	20.30	8.00	27.70	13.50
Radhuni pagal	21.00	9.00	27.00	16.90
Thakurbhog	26.60	7.20	19.00	12.00
<b>Selected seven inferior genotypes</b>				
Lal Soru	20.10	7.60	7.30	5.80
Luina	22.50	8.10	8.90	1.10
Sadagura (SI -104)	24.30	6.30	18.20	10.30
Gobindhabhog (SI -110)	20.80	6.40	20.60	9.00
Jirabuti	18.20	8.40	25.30	10.80
Begunmala	19.10	8.10	26.40	11.80
Chinniguri	19.10	8.30	20.30	10.20
<b>Calculations</b>				
Max.	26.60	11.00	34.10	16.90
Min.	18.20	6.30	7.30	1.10
Average	21.88	8.19	19.93	8.90
Std	1.91	1.03	6.48	3.72
SE	0.18	0.10	0.61	0.35
CV	8.72	12.54	32.50	41.74
LSD (5%)	0.35	0.19	1.19	0.69

<https://doi.org/10.1371/journal.pone.0234395.t001>



**Fig 2. Linear regression showing Zn concentration as a dependent variable on Fe concentration.**

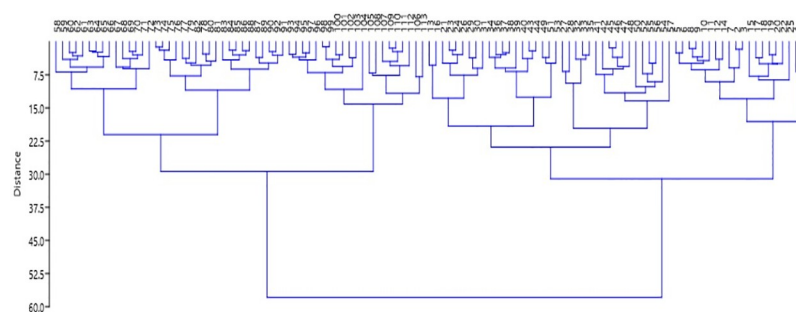
<https://doi.org/10.1371/journal.pone.0234395.g002>

### 2.3 Cluster analysis

The 113 germplasm were grouped into four major clusters (I, II, III and IV) at a 26% level of genetic similarity coefficient (Fig 3). Cluster I comprised 35 germplasm and was further classified into two sub-clusters IA and IB with 15 and 20 germplasm, respectively. Similarly, Cluster II comprised 21 genotypes and was further subdivided, with sub-clusters IIA and IIB having 12 and 9 germplasm, respectively, as well as higher diversity. Cluster III (with 37 genotypes, divided into IIIA and IIIB with 20 and 17, respectively) had a higher number of genotypes exhibit variable genetic distances among the germplasm. Cluster IV (with 19 genotypes, and 17 and 2 sub-clusters in IVA and IVB, respectively) was the smallest; in IVB, the two genotypes were Sakor and Chini Sagar. Higher inter-cluster was revealed between clusters III and IV, followed by clusters I and II (see Fig 3).

### 2.4 Molecular marker-trait associations

In the study, 45 polymorphic SSR markers were used in 113 rice genotypes for four phenotypic traits. Table 2 shows results of significant marker-trait associations. A total of eight significant



**Fig 3. A UPGMA cluster dendrogram of 113 aromatic rice germplasm based on 4 physio-chemical traits using Ward's method.**

<https://doi.org/10.1371/journal.pone.0234395.g003>

**Table 2. Significant marker-trait associations indicating quantitative trait loci for Fe, Zn, amylose and protein content in 113 rice genotypes.**

Traits	QTL name	Marker	Chromosome	Position (cM)	LOD	R <sup>2</sup> (%)	p-value
Fe	<i>QTL.Fe.9</i>	RM215	9	1.8	1.24	3.65	0.0424
Zn	<i>QTL.Zn.4</i>	RM551	4	8.5	1.71	5.89	0.0095
	<i>QTL.5n.5</i>	RM413	5	26.7	1.18	4.54	0.0234
Amylose	<i>QTL.amy.6</i>	RM190	6	7.4	1.11	4.41	0.0256
	<i>QTL.amy.7</i>	RM125	7	63.5	1.45	5.7	0.0109
	<i>QTL.amy.8</i>	RM284	8	83.7	1.83	7.2	0.0041
	<i>QTL.amy.11</i>	RM144	11	68.6	1.49	5.9	0.0095
Protein	<i>QTL.pro.1</i>	RM5	1	94.9	2.02	7.89	0.0026

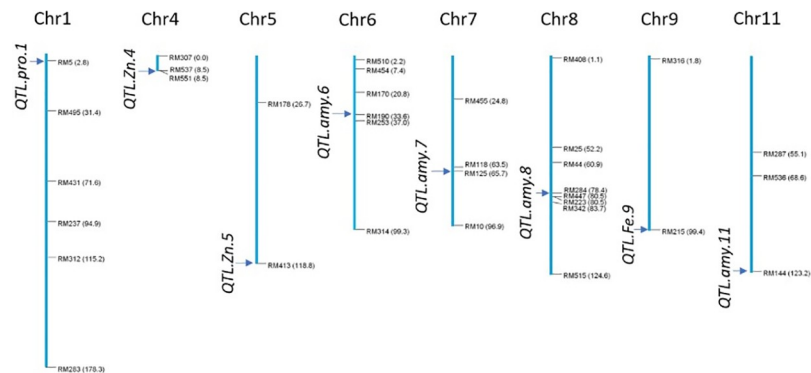
<https://doi.org/10.1371/journal.pone.0234395.t002>

markers were identified for four nutritional traits including high Fe, Zn, amylose, and protein, with each significant marker designated a quantitative trait loci (QTL): four QTLs for protein, two QTLs for Zn, one QTL for Fe, and one QTL for protein. These QTLs were mapped using Mapdisto Version 2 (see Fig 4).

For Fe content, its single QTL was located on chromosome 9 at position 1.8cM, with 3.65% variance ( $p < 0.05$ , LOD 1.24). For Zn content, two QTLs were located on chromosomes 4 and 5 at positions 8.5 and 26.7cM, respectively, with *QTL.Zn.4* contributing higher variance (5.89%,  $p < 0.01$  and LOD 1.71). For protein content, one QTL, *QTL.pro.1*, located on chromosome 1 at position 94.9 cM contributed a maximum 7.89% variance ( $p < 0.01$ , LOD 2.02). For amylose, 4 QTLs (*QTL.amy.6*, *QTL.amy.7*, *QTL.amy.8* and *QTL.amy.11*) were identified. In particular, *QTL.amy.8* revealed on chromosome 8 exhibited 7.2% variance ( $p < 0.01$ , LOD 1.83), followed by *QTL.amy.11* (LOD 1.49), *QTL.amy.7* (LOD 1.45) and *QTL.amy.6* (LOD 1.11).

### 3. Discussion

Plant breeding programs for biofortification of staple food crops (like rice or wheat) require screening of germplasm and varieties and/or elite lines having Fe, Zn, protein, and amylose dense grains to be used as donor parents [28]. An increase in concentration of these elements in grain is a high-priority research area. For example, maximum micronutrients are frequently present in some landraces and /or genetically distant wild varieties [29]. As such, exploitation of large genetic variation for Fe, Zn, amylose, and protein existing in cereal germplasm



**Fig 4. QTL map showing eight significant markers for four phenotypic traits in 113 aromatic rice germplasm.**

<https://doi.org/10.1371/journal.pone.0234395.g004>



represents an important strategy for minimizing the extent of Fe and Zn deficiencies in developing world.

In our study, the frequency distribution was normal for Fe, Zn, amylose, and protein content with maxima found in the Radhuni pagal, Kalobakri, Hatisail, and Thakurbhog genotypes, respectively. While the reasons for variably high-content varieties remain to be determined exactly, similar findings have been reported in other varieties of rice [30, 31]. One study obtained normal frequency distribution for Zn content and skewed distribution for Fe content in a recombinant inbred lines (RILs) population obtained from Madhukar×Swarna [27]. That study identified higher concentrations of Fe and Zn in Swarna (22.5ppm) and Madhukar (53.7 ppm), respectively. Moreover, pH, organic matter content, inherent Fe/Zn levels of native soil, environment, genotype, and genotype×environment interaction have all shown significant effects on rice grain Fe and Zn content [30, 31].

For genetic improvement of aromatic rice accessions in Bangladesh, a broader genetic base is required. Cluster analysis, which indicates the diversity and distances among experimental germplasm, offers a key tool in this effort. In this study, the 113 germplasms were grouped into four distinct clusters based on Mahalanobis  $D^2$  distances. Previous research clustered 113 aromatic and fine rice genotypes into ten groups based on 19 quantitative traits [17], ten clusters in rice genotypes [32, 33], and clusters of 58 rice varieties in groups based on 18 morphological traits with a genetic distance of approximately 0.75 [34].

In this study, cluster III consisted of 37 genotypes, with a higher number of genotypes exhibiting variable genetic distances. While the other clusters also exhibited variable genetic distances, the highest inter-cluster was observed between III and IV. Genotypes having distant clusters offer the best opportunity for hybridization in order to achieve a higher heterotic response.

The eight significant marker trait associations (MTA) were assayed using a general linear model (GLM). Most of these loci were similar to the major loci reported for these traits [35–42]. In the genotypic analysis, seven out of the 119 SSR markers used were linked to known major-effect drought grain yield QTLs identified at International Rice Research Institute (IRRI) [38]. Here, we identified several novel loci for Fe, Zn, amylose, and protein content in rice.

The most significant QTL protein content was *QTL.pro.1* located on chromosome 1 in position 94.9 cM with LOD 2.02 and 7.89% variance. This QTL is useful for marker assisted breeding (MAB) for higher protein content in rice. Several consistent minor-effect QTLs were also identified for amylose content, with the most significant QTL located on chromosome 8. QTLs are considered significant when above the threshold LOD score 3.0 [43]. Other research reported two QTLs on chromosomes 1 (markers RM8111-RM14323, LOD 2.03) and 9 (markers RM219-RM23914, LOD 2.88) for amylose content [44], as well three QTLs on chromosomes 8 (RM506-RM1235, LOD 2.57), 9 (RM219-RM23914, LOD 2.66) and 10 (RM24934-RM25128, LOD 6.13) for protein content in 120 doubled haploid (DH) rice population. For nutrient content itself, three QTLs were identified (one for Fe and two for Zn content) [44]. In particular, the highly significant *QTL.Zn.4* located on short arm of chromosome 4 with LOD 1.71 contributed a considerable amount of variability ( $R^2 = 5.89\%$ ) [44].

Other studies obtained 42 and 3 SSR markers association with grain Fe and Zn content, respectively. Using a GLM model with  $r^2 > 0.10$  and  $p \leq 0.05$  filters, they reported novel QTLs *qFe3.3* and *qFe7.3* for grain Fe and *qZn2.2*, *qZn8.3* and *qZn12.3* for Zn in 485 germplasm lines of milled rice [45]. Also, one study reported QTL *qSDW3*, associated with stem dry weight and a significant LOD score of 10.7, that explained approximately 7.5% of the phenotypic variation [46], as well as another two QTLs located on chromosomes 3 and 4 with LOD scores 5.3 and 4.8 explaining 4.5% and 12.5% variance. In the same study, another locus between

RM119-RM518 was also observed on chromosome 4 that shared about 2.45% in phenotypic variance [46]

Overall, these results demonstrate that association mapping represents a feasible option for identifying major-effect QTLs for traits studied. Indeed, genome-wide association analysis (GWAS) for complex traits using molecular markers like simple sequence repeats (SSRs) are becoming a more efficient tool for identifying loci or genes for a particular trait in crop plants. In the present study, using this technique, we have identified potential candidate QTLs for bio-fortification of rice for four key micronutrient traits.

## 4. Materials and methods

### 4.1 Plant materials

In this study, 113 aromatic rice germplasm representing landraces, fine rice genotypes, elite cultivars, and exotic genotypes preserved in the genebank of Bangladesh Rice Research Institute (BRRI), Gazipur, Bangladesh were used. Names for the 113 aromatic rice germplasm, along with quantitative phenotypic traits, have been previously described [17]. Amylose and protein content were measured at the BRRI, Grain Quality and Nutrition Division, while micronutrient Fe and Zn were determined at the BRRI, Plant Breeding Division.

### 4.2 Measurement of protein content

The Micro-Kjeldahl method was used for the determination of nitrogen [47]. Fifty milligrams of powdered sample were introduced into a 30 ml Kjeldahl flask. The catalyst mixture ( $K_2SO_4$ ,  $CuSO_4$  and Selenium powder) of 1.95 g was added followed by 2.3 ml concentrated  $H_2SO_4$ . Digestion was continued until the mixture became clear. After digestion, the flask was connected to a distillation set up. An Erlenmeyer flask (125 ml) containing 10 ml boric acid solution plus one drop of mixed indicator was placed under the condenser with the tip of the condenser extending below the surface of the solution. Then 9 ml of  $NaOH-Na_2S_2O_3$  solution was added slowly to the digest. The flask was connected to the steam source and distilled until about 30 ml distillate was collected. The distillate was immediately titrated against a standard HCl solution to the first appearance of a violet or reddish color. A blank determination was made simultaneously.

$$\% N = \frac{ml\ HCl\ for\ sample - mL\ HCl\ for\ bank}{weight\ of\ sample\ (g)} \times N\ HCl \times 0.014 \times 100$$

$$Crude\ protein\ (\%) = \% N \times 5.95$$

### 4.3 Measurement of amylose content

Juliano's method was used for the analysis of amylose content [48]. A milled rice sample of 100 mg was placed in a 100 ml volumetric flask. One ml of 95% ethanol and 9 ml of 1N NaOH were added to it. The contents were heated on a boiling water bath to gelatinize the starch. After cooling for one-hour, distilled water was added and the contents were mixed well. Five ml of the starch solution was put in a 100 ml volumetric flask. One ml of 1N acetic acid and 2 ml of iodine solution were added. Then distilled water was added to bring the volume up to the mark. The sample solution was set at 620 nm with a spectrophotometer. Absorbance values were plotted and a calculation was made with the help of a standard curve, available for rice samples of predetermined amylose content. Results were expressed as per cent amylose content in milled rice weight.



#### 4.4 Measurement of Fe and Zn content (mg/kg)

Polished grain was analyzed by Inductively Coupled Plasma (ICP) from the Plant Breeding Division Laboratory at BRRI. For each treatment, 0.6 g of powder were weighed accurately and placed in high-pressure digestion vessels (125 ml Erlenmeyer flask). A 10 ml mixture of  $\text{HClO}_4:\text{HNO}_3$  (1:10) was added to each flask and covered loosely with a plastic film label “perchloric” and kept overnight or longer in the fume hood. Pre-digested samples were mixed well by swirling and placed onto a hotplate and the temperature increased gradually to 225 °C. Then 2 ml of 2:5 mixture of  $\text{HClO}_4:\text{HNO}_3$  were added to each sample. When digests were clear to light yellow, the temperature was increased up to 240 °C and heated until 1 ml of digest (sample) was left. After cooling, the samples were diluted in 10 ml of 0.5N HCl. The digested samples were used for the quantification of Fe and Zn by inductively coupled plasma-optical emission spectrophotometer (ICP-OES) using a modified procedure [49]. The concentrations of Fe and Zn were determined and presented as mg/kg (ppm). The seed-Fe and Zn contents of germplasm were estimated, and statistical analysis done using their average values.

#### 4.5 Molecular characterization

We used 52 well-distributed SSRs for the molecular characterization (S1 Data). The cM positions, repeat motifs, and chromosomal positions for the SSR markers can be found in the rice genome database [50]. Out of these 52 SSR markers, 45 were polymorphic, while 7 were monomorphic. The 45 polymorphic markers selected for analysis were distributed across the 12 chromosomes, from those 3 linked to aromatic traits, 4 related to cooking and eating quality traits, and 31 were listed in the panel of 50 standard SSR markers used for marker-trait analysis; the remainder of SSRs were selected randomly.

DNA was extracted from the young leaves of 21-day-old plants using the miniscale method [51]. Each PCR was carried out in a 20  $\mu\text{l}$  reaction volume containing 1  $\mu\text{l}$  of  $\text{MgCl}_2$  free  $10\times$  PCR buffer with  $(\text{NH}_4)_2\text{SO}_4$ , 1.2  $\mu\text{l}$  of 25 mM of  $\text{MgCl}_2$ , 0.2  $\mu\text{l}$  of 10 mM of dNTPs, 0.2  $\mu\text{l}$  of 5 U/ $\mu\text{l}$  Taq DNA polymerase, 0.5  $\mu\text{l}$  of 10  $\mu\text{M}$  forward and reverse primers, and 3  $\mu\text{l}$  (10 ng) of DNA using a 96-well thermal cycler. An additional 10  $\mu\text{l}$  of mineral oil was added in each well to prevent evaporation. Amplification was carried out using a G-storm PCR machine (Gene Technologies Ltd., England). Amplification conditions were one cycle at 94 °C for 5 minutes (initial denaturation) followed by 35 cycles at 94 °C for 1 minute (denaturation), 55 °C for 1 minute (annealing), 72 °C for 2 minutes (extension), with a final extension for 7 minutes at 72 °C at the end of 35 cycles. After mixing with the loading dye, PCR products were run through polyacrylamide gels. A 50 bp DNA ladder was used to determine the amplicon size. Three 4  $\mu\text{l}$  PCR products were resolved by running gel in 1X TB buffer for 1.5 to 2.5 h depending upon the allele size at approximately 90 volts and 500 mA electricity. Gels were then stained with 1  $\mu\text{g}/\text{mL}$  of ethidium bromide and documented using a Molecular Imager gel documentation unit (XR System, BIO-RAD, Korea).

#### 4.6 Statistical analysis

**Phenotypic analysis.** All data used to analyse and presented in this paper are provided as supplementary files (S1 Fig and S1 Table). The diverse statistical parameters, including mean, standard deviation, coefficient of variation (CV), analysis of variance (ANOVA), frequency distribution and Pearson’s correlation coefficient, regression co-efficient of seed Fe, Zn, amylose, and protein contents among accessions were measured using SPSSv17.0. A UPGMA cluster dendrogram of 113 aromatic rice germplasm based on 4 physio-chemical traits was constructed using PAST software package [52] following Ward’s hierarchical clustering method.

**Molecular marker-trait association analysis.** The size of the band for each 45 polymorphic markers was scored by AlphaEaseFC 4.0 software. The summary statistics, including the number of alleles, major allele size, and frequency was determined using PowerMarker version 3.25 [53]. The allele frequency data from PowerMarker was used to export the data in binary format (allele presence = “1” and allele absence = “0”) for analysis with NTSYS-pc version 2.2 [54]. Association mapping was conducted for four quantitative traits viz., Fe, Zn, amylose, and protein content in rice grain. A general linear model (GLM) was used for assessment of marker-trait associations (MTA) in R version 5.2.2 [55]. The molecular marker-trait associations were calculated using the binary data. To detect the association of Fe, Zn, amylose, and protein content in the rice genome, logarithm of odds (LOD) thresholds were calculated using 1000 permutations with a significance threshold of  $p = 0.05$  used as a criteria for QTL analysis of markers with the trait [56]. Here, Composite Interval Mapping (CIM), one of the most frequently used quantitative trait loci (QTL) analysis method [57], was used. Using CIM, putative QTLs were identified using the phenotypic and SSR marker by R/qtl package [58] in R. The QTL map was constructed using MAPdisto version 2.0 [59]. During this analysis, the derived genetic map, missing phenotypes and frequency of other phenotypic data presented as [S1 Fig](#).

## Supporting information

### **S1 Data.**

(XLSX)

### **S1 Fig. Genetic map, frequency distribution pattern of phenotypic data.**

(DOCX)

### **S1 Table. List of 52 SSR markers used in this study.**

(DOC)

## Author Contributions

**Conceptualization:** M. Khalequzzaman.

**Data curation:** M. Z. Islam, S. Banik, M. A. Hossain, J. Ferdous.

**Formal analysis:** M. Z. Islam, M. Arifuzzaman, S. Banik, M. P. Ali.

**Investigation:** M. Z. Islam, M. A. Hossain, M. Khalequzzaman.

**Methodology:** M. Z. Islam.

**Project administration:** M. Khalequzzaman.

**Resources:** M. Khalequzzaman.

**Software:** M. Z. Islam.

**Validation:** B. R. Pittendrigh, M. Tomita, M. P. Ali.

**Visualization:** M. Arifuzzaman, M. P. Ali.

**Writing – original draft:** M. Z. Islam, M. Arifuzzaman, B. R. Pittendrigh, M. Tomita, M. P. Ali.

**Writing – review & editing:** M. Arifuzzaman, B. R. Pittendrigh, M. Tomita, M. P. Ali.

## References

1. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, De Onis M, et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*. 2013; 382(9890):427–51. [https://doi.org/10.1016/S0140-6736\(13\)60937-X](https://doi.org/10.1016/S0140-6736(13)60937-X)
2. Stevens GA, Finucane MM, De-Regil LM, Paciorek CJ, Flaxman SR, Branca F, et al. Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995–2011: a systematic analysis of population-representative data. *The Lancet Global Health*. 2013; 1(1):e16–e25. [https://doi.org/10.1016/S2214-109X\(13\)70001-9](https://doi.org/10.1016/S2214-109X(13)70001-9) PMID: 25103581
3. Hotz C, Brown KH. Assessment of the risk of zinc deficiency in populations and options for its control. *Food and Nutrition Bulletin*. 2004; 25(1, Supplement 2):S94–S202.
4. Prasad AS. Zinc deficiency: Its characterization and treatment. In: Sigel A, editor. *Met Ions Biol Syst*. 41. Boca Raton, FL: CRC Press; 2004. p. 103–38. PMID: 15206115
5. Wessells KR, Brown KH. Estimating the global prevalence of zinc deficiency: results based on zinc availability in national food supplies and the prevalence of stunting. *PLoS One*. 2012; 7(11):e50568. <https://doi.org/10.1371/journal.pone.0050568> PMID: 23209782
6. Richard DS. The Historical Evolution of Thought Regarding Multiple Micronutrient Nutrition. *J Nutr*. 2012; 142(1):143S–56S. <https://doi.org/10.3945/jn.110.137745> PMID: 22157539
7. Swamy BM, Rahman MA, Inabangan-Asilo MA, Amparado A, Manito C, Chadha-Mohanty P, et al. Advances in breeding for high grain zinc in rice. *Rice*. 2016; 9(1):49. <https://doi.org/10.1186/s12284-016-0122-5> PMID: 27671163
8. Bouis HE, Welch RM. Biofortification—A sustainable agricultural strategy for reducing micronutrient malnutrition in the global south. *Crop Science*. 2010; 50(Supplement\_1):S-20–S-32. <https://doi.org/10.2135/cropsci2009.09.0531>
9. Bouis HE. Micronutrient fortification of plants through plant breeding: Can it improve nutrition in man at low cost? *Proc Nutr Soc*. 2003; 62(2):403–11. <https://doi.org/10.1079/pns2003262> PMID: 14506888
10. White PJ, Broadley MR. Biofortifying crops with essential mineral elements. *Trends in Plant Science*. 2005; 10(12):586–93. <https://doi.org/10.1016/j.tplants.2005.10.001> PMID: 16271501
11. Zhang M, Guo B, Peng Z. Genetic effects on Fe, Zn, Mn and P contents in *Indica* black pericarp rice and their genetic correlations with grain characteristics. *Euphytica*. 2004; 135(3):315–23. <https://doi.org/10.1023/B:EUPH.0000013340.98344.60>
12. HarvestPlus. Biofortification Progress Briefs 2014 [cited 2019 29 October]. [https://www.harvestplus.org/sites/default/files/Biofortification\\_Progress\\_Briefs\\_August2014\\_WEB\\_0.pdf](https://www.harvestplus.org/sites/default/files/Biofortification_Progress_Briefs_August2014_WEB_0.pdf).
13. Wang X, Pang Y, Zhang J, Wu Z, Chen K, Ali J, et al. Genome-wide and gene-based association mapping for rice eating and cooking characteristics and protein content. *Sci Rep*. 2017; 7(1):17203. <https://doi.org/10.1038/s41598-017-17347-5> PMID: 29222496
14. Singh R, Gautam P, Saxena S, Singh S. Scented rice germplasm: Conservation, evaluation and utilization. In: Singh R, Singh U, Khush GS, editors. *Aromatic rices*. New Delhi, IN: Oxord & IBH Publishing Co. Pvt. Ltd.; 2000. p. 107–33.
15. Islam M, Banik S, Haque A, Siddique M, Khalequzzaman M, Mian M. Physico-chemical and cooking properties of local aromatic rice germplasm in Bangladesh. *Eco-friendly Agriculture Journal*. 2013; 6:243–8.
16. Shahidullah S, Hanafi M, Ashrafuzzaman M, Ismail MR, Khair A. Genetic diversity in grain quality and nutrition of aromatic rices. *African Journal of Biotechnology*. 2009; 8(7):1238–46.
17. Islam M, Khalequzzaman M, Bashar M, Ivy N, Haque M, Mian M. Variability Assessment of Aromatic and Fine Rice Germplasm in Bangladesh Based on Quantitative Traits. *The Scientific World Journal*. 2016; 2016:Article ID 2796720. <https://doi.org/10.1155/2016/2796720> PMID: 27127800
18. Islam M, Khalequzzaman M, Bashar M, Ivy N, Mian M, Pittendrigh BR, et al. Variability Assessment of Aromatic Rice Germplasm by Pheno-Genomic traits and Population Structure Analysis. *Sci Rep*. 2018; 8(1):Art. 9911.
19. Gande NK, Kundur PJ, Soman R, Ambati R, Ashwathanarayana R, Bekele BD, et al. Identification of putative candidate gene markers for grain zinc content using recombinant inbred lines (RIL) population of IRRI38 X Jeerigesanna. *African Journal of Biotechnology*. 2014; 13(5):657–63. <https://doi.org/10.5897/AJB2013.13334>
20. Zhang M, Pinson SR, Tarpley L, Huang X-Y, Lahner B, Yakubova E, et al. Mapping and validation of quantitative trait loci associated with concentrations of 16 elements in unmilled rice grain. *Theoretical and Applied Genetics*. 2014; 127(1):137–65. <https://doi.org/10.1007/s00122-013-2207-5> PMID: 24231918

21. Collard BC, Mackill DJ. Marker-assisted selection: An approach for precision plant breeding in the twenty-first century. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2007; 363 (1491):557–72. <https://doi.org/10.1098/rstb.2007.2170> PMID: 17715053
22. Swamy BM, Kumar A. Sustainable rice yield in water-short drought-prone environments: conventional and molecular approaches. In: Lee TS, editor. *Irrigation Systems and Practices in Challenging Environments*. Rijeka, Croatia: InTech; 2012. p. 149–68.
23. Nawaz Z, Kakar KU, Li X-b, Li S, Zhang B, Shou H-x, et al. Genome-wide association mapping of quantitative trait loci (QTLs) for contents of eight elements in brown rice (*Oryza sativa* L.). *J Agric Food Chem*. 2015; 63(36):8008–16. <https://doi.org/10.1021/acs.jafc.5b01191> PMID: 26317332
24. Norton GJ, Douglas A, Lahner B, Yakubova E, Guerinot ML, Pinson SR, et al. Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (*Oryza sativa* L.) grown at four international field sites. *PLoS One*. 2014; 9(2):e89685. <https://doi.org/10.1371/journal.pone.0089685> PMID: 24586963
25. Nordborg M, Tavaré S. Linkage disequilibrium: What history has to tell us. *Trends Genet*. 2002; 18 (2):83–90. [https://doi.org/10.1016/s0168-9525\(02\)02557-x](https://doi.org/10.1016/s0168-9525(02)02557-x) PMID: 11818140
26. Garcia-Oliveira AL, Tan L, Fu Y, Sun C. Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *J Integr Plant Biol*. 2009; 51(1):84–92. <https://doi.org/10.1111/j.1744-7909.2008.00730.x> PMID: 19166498
27. Anuradha K, Agarwal S, Rao YV, Rao K, Viraktamath B, Sarla N. Mapping QTLs and candidate genes for iron and zinc concentrations in unpolished rice of Madhukar × Swarna RILs. *Gene*. 2012; 508 (2):233–40. <https://doi.org/10.1016/j.gene.2012.07.054> PMID: 22964359
28. Stangoulis J. Technical aspects of zinc and iron analysis in biofortification of the staple food crops, wheat and rice. 19th World Congress of Soil Science, Soil Solutions for a Changing World; 1–6 August 2010; Brisbane, Australia 2010.
29. Brar B, Jain S, Singh R, Jain R. Genetic diversity for iron and zinc contents in a collection of 220 rice (*Oryza sativa* L.) genotypes. *Indian Journal of Genetics and Plant Breeding*. 2011; 71(1):67–73.
30. Suwanto N. Genotype × environment interaction for iron concentration of rice in central Java of Indonesia. *Rice Science*. 2011; 18(1):75–8.
31. Pandian S, Robin S, Vinod K, Rajeswari S, Manonmani S, Subramanian K, et al. Influence of intrinsic soil factors on genotype-by-environment interactions governing micronutrient content of milled rice grains. *Australian Journal of Crop Science*. 2011; 5(13):1737–44.
32. Singh S, Singh R, Maurya D, Verma O. Genetic divergence among lowland rice cultivars. *The Indian Journal of Genetics and Plant Breeding*. 1987; 47(1):11–4.
33. Rao A, Prasad A, Krishna TS, Seshu D, Srinivasan T. Genetic divergence among some brown planthopper resistant rice varieties. *Indian Journal of Genetics and Plant Breeding*. 1981; 41(2):179–85.
34. Ahmadikhah A, Nasrollanejad S, Alishah O. Quantitative studies for investigating variation and its effect on heterosis of rice. *International Journal of Plant Production*. 2008; 2(4):297–308.
35. Bernier J, Kumar A, Ramaiah V, Spaner D, Atlin G. A large-effect QTL for grain yield under reproductive-stage drought stress in upland rice. *Crop Science*. 2007; 47(2):507–16. <https://doi.org/10.2135/cropsci2006.07.0495>
36. Courtois B, Audebert A, Dardou A, Roques S, Ghneim-Herrera T, Droc G, et al. Genome-wide association mapping of root traits in a japonica rice panel. *PLoS One*. 2013; 8(11):e78037. <https://doi.org/10.1371/journal.pone.0078037> PMID: 24223758
37. Ghimire KH, Quiatchon LA, Vikram P, Swamy BM, Dixit S, Ahmed H, et al. Identification and mapping of a QTL (qDTY1. 1) with a consistent effect on grain yield under drought. *Field Crops Research*. 2012; 131:88–96. <https://doi.org/10.1016/j.fcr.2012.02.028>
38. Kumar A, Dixit S, Ram T, Yadaw R, Mishra K, Mandal N. Breeding high-yielding drought-tolerant rice: Genetic variations and conventional and molecular approaches. *Journal of Experimental Botany*. 2014; 65(21):6265–78. <https://doi.org/10.1093/jxb/eru363> PMID: 25205576
39. Palanog AD, Swamy BM, Shamsudin NAA, Dixit S, Hernandez JE, Boromeo TH, et al. Grain yield QTLs with consistent-effect under reproductive-stage drought stress in rice. *Field Crops Research*. 2014; 161:46–54. <https://doi.org/10.1016/j.fcr.2014.01.004>
40. Swamy BM, Kaladhar K, Ramesha M, Viraktamath B, Sarla N. Molecular mapping of QTLs for yield and yield-related traits in *Oryza sativa* cv Swarna × O. nivara (IRGC81848) backcross population. *Rice Science*. 2011; 18(3):178–86. [https://doi.org/10.1016/S1672-6308\(11\)60025-5](https://doi.org/10.1016/S1672-6308(11)60025-5)
41. Vikram P, Swamy BM, Dixit S, Ahmed HU, Cruz MTS, Singh AK, et al. qDTY 1.1, a major QTL for rice grain yield under reproductive-stage drought stress with a consistent effect in multiple elite genetic backgrounds. *BMC Genet*. 2011; 12(1):89. <https://doi.org/10.1186/1471-2156-12-89> PMID: 22008150

42. Venuprasad R, Bool M, Dalid C, Bernier J, Kumar A, Atlin G. Genetic loci responding to two cycles of divergent selection for grain yield under drought stress in a rice breeding population. *Euphytica*. 2009; 167(2):261–9. <https://doi.org/10.1007/s10681-009-9898-3>
43. Zobaer Akond M, Alam J, Hasan MN, Uddin MS, Alam M, Mollah MNH. A Comparison on Some Interval Mapping Approaches for QTL Detection. *Bioinformatics*. 2019; 15(2):90–4. <https://doi.org/10.6026/97320630015090> PMID: 31435154
44. Yun B-W, Kim M-G, Handoyo T, Kim K-M. Analysis of rice grain quality-associated quantitative trait loci by using genetic mapping. *American Journal of Plant Sciences*. 2014; 2014(5):9. <https://doi.org/10.4236/ajps.2014.59125>
45. Pradhan S, Pandit E, Pawar S, Naveenkumar R, Barik S, Mohanty S, et al. Linkage disequilibrium mapping for grain Fe and Zn enhancing QTLs useful for nutrient dense rice breeding. *BMC Plant Biol*. 2020; 20(1):57. <https://doi.org/10.1186/s12870-020-2262-4> PMID: 32019504
46. Sabar M, Shabir G, Shah SM, Aslam K, Naveed SA, Arif M. Identification and mapping of QTLs associated with drought tolerance traits in rice by a cross between Super Basmati and IR55419-04. *Breeding science*. 2019; 69(1):169–78. <https://doi.org/10.1270/jsbbs.18068> PMID: 31086495
47. AOAC. *Methods of analysis*. Washington, DC: Association of Official Agricultural Chemists; 1995.
48. Juliano B. A simplified assay for milled rice amylose. *Cereal Science Today*. 1971; 16:334–60.
49. Boss CB, Fredeen KJ. *Concepts, instrumentation and techniques in inductively coupled plasma optical emission spectrometry*. Waltham, MA: Perkin Elmer Corporation; 1997.
50. Gramene. *Gramene Portals 2017* [cited 2017 22 October]. <http://www.gramene.org/>.
51. Zheng K, Huang N, Bennett J, Khush GS. PCR-based marker-assisted selection in rice breeding (IRRI Discussion Paper Series, Philippines). Los Baños, Philippines: IRRI; 1995.
52. Hammer Ø, Harper D, Ryan P. *Paleontological statistics software: Package for education and data analysis*. *Palaeontologia Electronica*. 2001; 4(1):1–9.
53. Liu K, Muse SV. PowerMarker: An integrated analysis environment for genetic marker analysis. *Bioinformatics*. 2005; 21(9):2128–9. <https://doi.org/10.1093/bioinformatics/bti282> PMID: 15705655
54. Rohlf FJ. *NTSYS-pc. Numerical taxonomy and multivariate analysis system version 2.1*. New York, USA: Exeter Software; 2002.
55. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2019. <https://www.R-project.org/>.
56. Churchill GA, Doerge RW. Empirical threshold values for quantitative trait mapping. *Genetics*. 1994; 138(3):963–71. PMID: 7851788
57. Zeng Z-B. Precision mapping of quantitative trait loci. *Genetics*. 1994; 136(4):1457–68. PMID: 8013918
58. Broman KW, Wu H, Sen S, Churchill GA. R/qtl: QTL mapping in experimental crosses. *Bioinformatics*. 2003; 19(7):889–90. <https://doi.org/10.1093/bioinformatics/btg112> PMID: 12724300
59. Heffelfinger C, Fragoso CA, Lorieux M. Constructing linkage maps in the genomics era with MapDisto 2.0. *Bioinformatics*. 2017; 33(14):2224–5. <https://doi.org/10.1093/bioinformatics/btx177> PMID: 28369214