Original Research Article
ISSN: **3041-5357**
Journal homepage: https://diversity.researchfloor.org/

Journal of Diversity Studies

Evaluation of Scented Non-Basmati Rice (*Oryza sativa* L.) Accessions for Seed Storage Proteins and SSR Markers Linked to Major QTLs of Protein

Yasmin M Shaikh, Vitthal T Barvkar* and Altafhusain B Nadaf*

Department of Botany, Savitribai Phule Pune University, Pune-411007, India

ABSTRACT

Protein in rice (Oryza sativa L.) grains is considered as on of the fundamental nutrients after carbohydrates. Understanding the genetic background underlying protein content is pivotal for improving the nutritional quality in rice. The present study aimed to analyse variation for SSP content and to find marker trait association in rice accessions. The total protein content along with protein fraction: albumin (0.11-0.78), globulin (0.11-1.43), glutelin (1.44-9.23), and prolamin (0.05 – 0.36) mg/100 mg in the 101 rice accessions was determined. A correlation analysis demonstrated that globulin was positively correlated with total protein content and albumin. Glutelin content display positive correlation with total protein content, while no significant correlation was observed between glutelin and either albumin or globulin. Genotyping with 8 protein fraction-linked simple sequence repeats (SSR) markers in 50 rice accessions identified 42 polymorphic alleles, and the mean polymorphism information content value for each marker was 0.61 (0.2-0.75), suggesting that the studied markers were informative and helpful for diversity analysis. All evaluated accessions were divided into two subpopulations. Further marker-trait association analysis identified 3 SSR marker alleles, including two novel ones associated with highest glutelin (RM472_260) and lower prolamin content (RM472_330), will be useful for marker-assisted breeding. This study provide first insight into the SSR based association analysis for SSP fraction in landraces of scented non basmati rice suggesting their utility in molecular breeding targeting nutritional quality improvement.

Keywords: *Rice; SSPs; SSR ; Association Analysis.*

Citation: Yasmin M Shaikh, Vitthal T Barvkar and Altafhusain B Nadaf [2026]. Evaluation of Scented Non-Basmati Rice (*Oryza sativa* L.) Accessions for Seed Storage Proteins and SSR Markers Linked to Major QTLs of Protein. *Journal of Diversity Studies.*
DOI: https://doi.org/10.51470/JOD.2026.5.1.180

Corresponding Author: *Vitthal T Barvkar and Altafhusain B Nadaf*

E-mail Address: *bvitthal@unipune.ac.in, abnadaf@unipune.ac.in*

Article History: *Received 20 January 2026 | Revised 17 February 2026 | Accepted 21 March 2026 | Available Online April 27, 2026*

Introduction

Rice is the primary staple source of food for half of the world, as it gives necessary nutrients for healthy living. Protein represents the second abundant macronutrient in rice grains, following carbohydrates. Its concentration varies among different rice varieties, approximately 5% to 16%. Indica rice cultivars exhibit higher protein content (2%-3%) than Japonica [1]. Based on differential protein solubility, SSPs categorised into four types: glutelin, albumin, globulin and prolamin, and accounting glutelins, 80 % or higher of the total seed protein [2]. "The remaining 20% consists of 1-5% albumin, 4-15% globulin, and 2-8% prolamins" [3]. Compared to other major cereal crops, rice (*Oryza sativa* L.) exhibits low SSPs. Yet, rice protein is recognized for its high nutritional quality, because of the presence of glutelin a major protein fraction compared to prolamins [4]. Furthermore, it is considered hypoallergenic, making it a suitable ingredient in the formulation of infant weaning foods and specialized dietary products for individuals needing a gluten-free diet. These attributes underscore the functional and nutritional advantages of high-quality rice protein.

To assess genetic variations, SSRs or rice microsatellite (RM) are most preferentially and widely used in assessing genetic diversity since they are cost effective, reliable, rapid, and easy to score [5, 6]. By combining SSPs evaluation with molecular data offers valuable insights in marker-assisted breeding programme for developing high-quality rice. Several recent studies have been identified the quantitative trait loci (QTLs) linked to protein content in rice furnishing essential information useful for the improvement of rice protein quality [7, 8, 9, 10]. Most recent studies have only focused on total protein, but little is known about SSP fraction and its linked SSR markers in rice [11], which will be used in association analysis in landraces. Scented rice is one of the most popularly accepted rice due to its pleasant aroma. Besides basmati, India cultivates numerous indigenous scented non-basmati varieties. However, comprehensive information regarding their nutritional quality remains scarce. The study aimed to evaluate the local non-basmati rice accessions for SSP and to study the association based on reported protein-linked SSR markers. However, a small set of 8 SSR markers was used, as the objective was to validate the association of already reported markers in the selected accessions rather than marker discovery hence, a selected 8 marker set was appropriate.

Materials and Methods

Plant material

A total of 101 rice accessions (80- scented non basmati, 12-non scented, 8-basmati and 1-jasmine) were used (Table S1). The majority of accessions were available at Savitribai Phule Pune University [12], additional were collected from the Rice Diversity Center, Kirugavalu, Karnataka. Accessions were grown at Savitribai Phule Pune University using a Randomized Complete Block Design (RCBD) and preserved for further phenotypic evaluation.

Quantification of SSP fractions

The rice SSPs namely glutelin, albumin, prolamin and globulin, were determined following [13] with slight modifications. A 100 mg powder of each rice sample was centrifuged for protein fraction: albumin, prolamin, globulin, and glutelin extraction with 1.0 mL of their respective extraction solutions, namely 10 mM Tris–HCl buffer (pH7.5), 60 % n-propanol containing 1 mM EDTA-2Na, 1 M NaCl and 0.05 M NaOH. The mixture was stirred for 2 h at room temperature. The extracts were centrifuged at 12,000 rpm for 15 min at 4 °C and the supernatant was collected as a protein fraction extract. The Coomassie brilliant blue G-250 dye-binding method [14] was used for protein fraction quantification.

Correlation analysis between SSP fractions

RStudio software version 2023.06.2 was used for the correlation analysis between protein fraction content and to generate corrplot (Fig 2).

DNA isolation and PCR

A set of 50 diverse rice accessions was selected for SSR based genotyping based on the phenotypic dataset of 101 rice accessions, ensuring representation across key traits of interest. Genomic DNA from leaf was isolated using the CTAB method [15] with some modifications. Total 8 SSR markers (Table S2) associated with major qtls of total protein were selected for genotyping [16, 17, 18, 19, 20]. PCRs were carried out under standard conditions, and the amplified products were subsequently separated on 2.5% agarose gels. Finally, the PCR amplicons were resolved on 8% denaturing polyacrylamide gel electrophoresis. SSR marker data scoring was performed visually for allele size and presence or absence of band.

Genetic diversity, Population structure and Phylogenetic analysis

Genetic diversity and multivariate analysis (AMOVA) were conducted with the help of GenALEx version 6.5 software [21]. Principal Coordinate Analysis (PcoA) was performed based on Nei's unbiased genetic distance pairwise population matrix following Nei and Li method [22]. Principal component analysis (PCA) was performed using TASSEL 3.0 software [23]. To determine population structure the STRUCTURE 2.3.4 software was used, with K values ranging from 1 to 10 with ten repeats. Furthermore, to determine the Delta K value by implementing the Evanno's method [24], an online software program StructureSelector was used [25]. A phylogenetic tree was constructed with MEGA 10 [26] using genetic distance matrix and unweighted pair group method with arithmetic mean (UPGMA) statistical method.

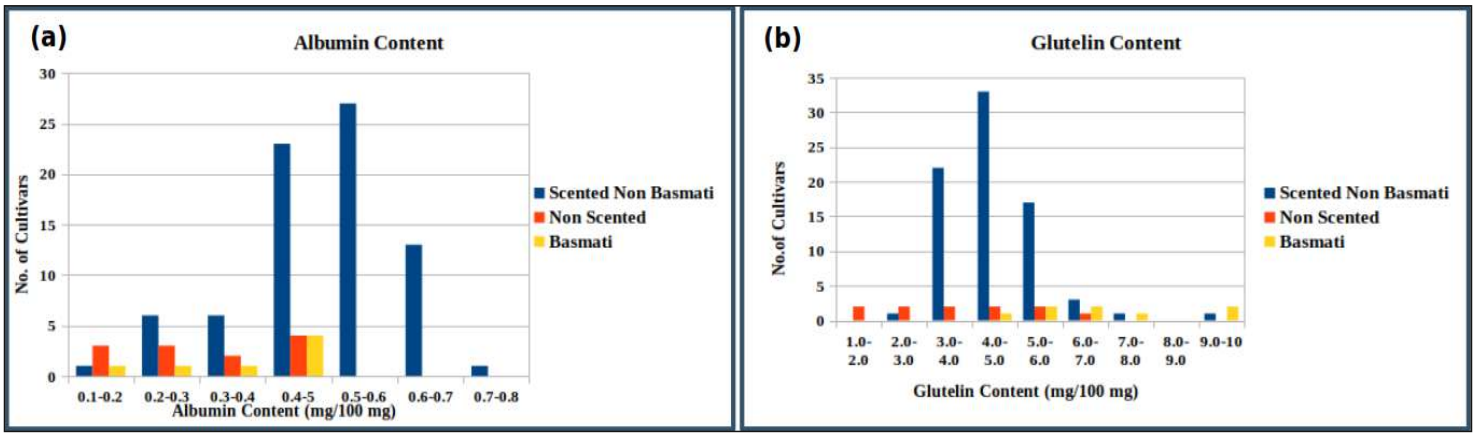
Marker trait association analysis

Associations between SSR and SSPs were analysed using TASSEL 3.0 software through a general linear model (GLM). Quantile-quantile (Q–Q) plots were constructed (Fig 6). Associations between SSPs and SSR markers were considered significant at a threshold of $P < 0.05$.

Results and Discussion

Quantitative and correlation analysis of SSP fractions

Total protein in the studied accessions ranged from 2.34 to 10 mg/100 mg grains with an average of 5.88, which is in line with the earlier reports of Kim and Jeong [27] for japonica cultivars, but comparatively lower than [28, 29]. The albumin ranged from 0.11 to 0.78 mg/100 mg with an average of 0.46. Globulin varied between 0.11 and 1.43 mg/100 mg with an average of 0.67 mg/100 mg. Glutelin ranged from 1.44 to 9.23 mg/100 mg, with a mean value of 4.58 mg/100 mg, while prolamin varied from 0.05 to 0.36 mg/100 mg, with a mean of 0.18 mg/100 mg (Fig 1 and Table S1). Majority of values are in agreement with the value reported by [30]. However, albumin and globulin were higher. A similar study was done by [31], reported lower globulin and higher prolamin, while albumin and glutelin levels were comparable. This deviation suggests a possible genotypic or environmental influence specifically affecting the accumulation of storage proteins like prolamin.



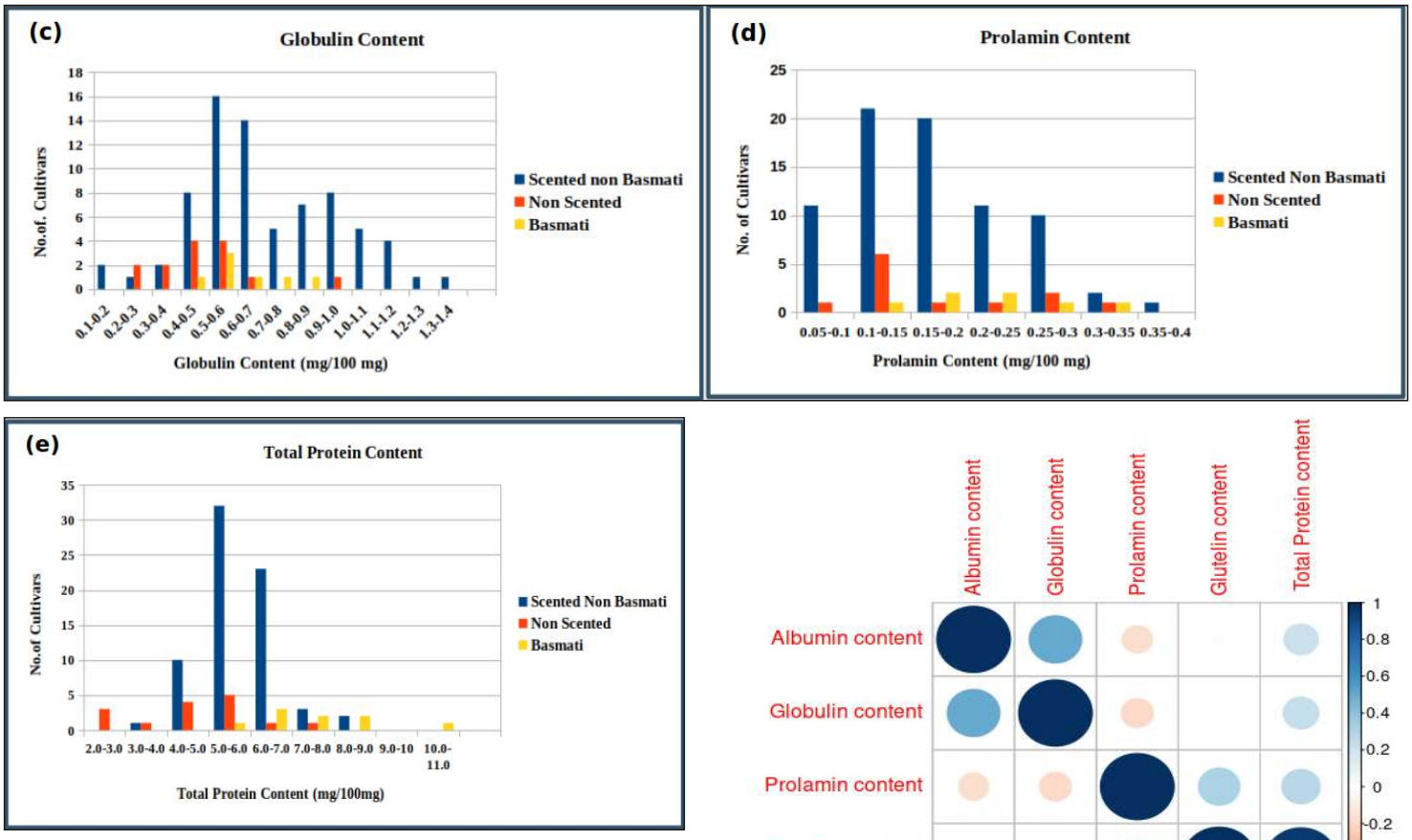


Fig 1: Histograms showing the distributions of Albumin (a), Glutelin (b), Globulin (c), Prolamin (d) and total protein (e) among the rice accessions.

Considerable varietal variation was observed in the contents of all SSPs, indicating good phenotypic diversity. The Royal Basmati accession was identified with the maximum level of total protein and glutelin. Whereas, Luchai Moti and Shrikant accessions were observed with the least total protein and globulin fraction respectively. The least glutelin was found in the Majeri, Ratibhog, and Navedhan genotypes, which could be suitable for kidney and diabetic patients. Furthermore, albumin, globulin, and prolamin were found to be maximum in Kamavatya, Geerige Sanna, and Kali Kumud, respectively. The least albumin, prolamin, and glutelin content were found in the Gham, Bhansphool-A, and Majeri, respectively.

A correlation analysis demonstrated that globulin was positively correlated with total protein content and albumin. Glutelin content display positive correlation with total protein content; no significant correlation was observed between glutelin and either albumin or globulin (Fig 2). Similar observations were reported by [1]. However, several studies [30,31,32] have documented positive correlation between glutelin and other three SSPs. The considerable phenotypic variations and significant correlation among the studied accessions supports the use of this panel for genotyping and association analysis.

Fig 2: Corr plot representing the correlation among protein fraction.

Genetic diversity analysis based on SSR markers

The 50 rice accessions exhibited notable genetic variations, revealed by protein content linked SSR markers (Table 1 and Fig 3). All SSR markers used in this study were polymorphic. A total of 42 polymorphic alleles were detected, and the number of alleles per marker was between the range of 3 and 8, with a mean of 5.25. The MAF, and Ho varied from 0.317 to 0.881 and 0.041–0.714, respectively. The 5.25 average numbers of alleles per locus were in agreement with 5.54 [33] and 5.49 [11], but higher than 3.11 [34], 2.5 [35] and 3 [36]. However, it was lower than 7.43 [37]. Collectively, this indicates there appears to be favourable allelic diversity, which is essential for the genetic diversity assessment. The mean PIC for each marker was 0.61 (range 0.2-0.75). Despite, using the small number of markers the majority of markers were found to be the most suitable markers to differentiate among the rice accessions, attributable to the PIC >0.5. However, further research is required to reveal the genetic architecture underlying the protein fraction content in rice and to develop more effective breeding strategies for these important traits.

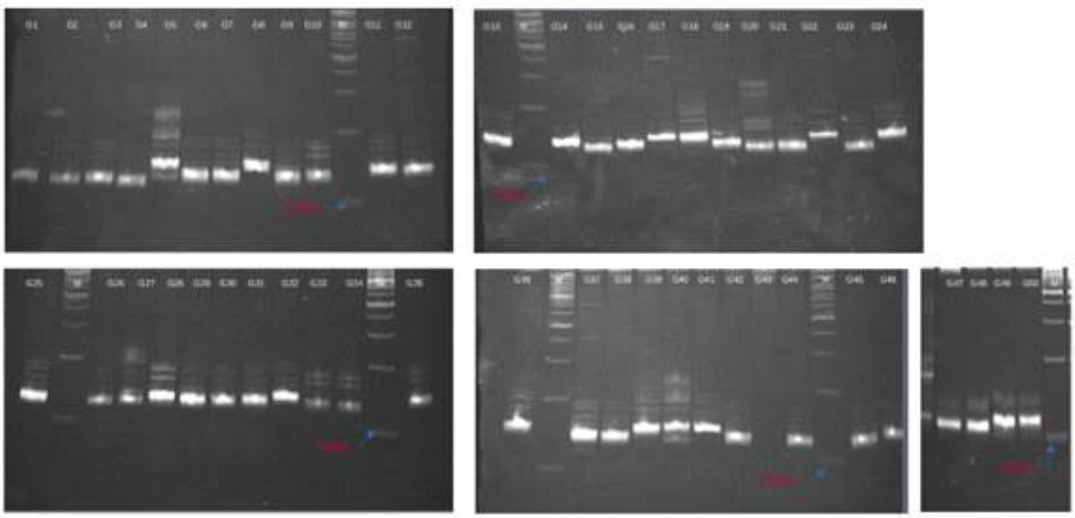


Fig 3: PCR amplification profile of RM125 markers of selected rice accession on 8% PAGE. (M-100 bp DNA ladder, accession codes (G1–G50) and their corresponding names are provided in Supplementary file S2)

Table 1: Genetic diversity indices revealed by SSR markers in a given rice accession. MAF; Major allele frequency, Na; No. Alleles, Ne; No. Effective Alleles, I; Information Index, Ho; Observed Heterozygosity, He; Expected Heterozygosity, uHe; Unbiased Expected Heterozygosity, and F; Fixation Index.

Sr. No.	Marker Name	MAF	Na	Ne	I	Ho	He	uHe	F	PIC
1	RM472	0.454	7	3.338	1.433	0.488	0.700	0.709	0.300	0.70
2	RM7124	0.364	6	3.681	1.445	0.682	0.728	0.737	0.064	0.73
3	RM6973	0.590	3	2.015	0.773	0.103	0.504	0.510	0.796	0.50
4	RM5672	0.489	5	2.792	1.201	0.205	0.642	0.649	0.681	0.74
5	RM3917	0.881	2	1.265	0.365	0.143	0.210	0.212	0.319	0.20
6	RM541	0.317	8	4.789	1.719	0.390	0.791	0.801	0.507	0.75
7	RM180	0.398	8	3.219	1.431	0.455	0.689	0.697	0.341	0.70
8	RM125	0.591	3	2.310	0.960	0.045	0.567	0.574	0.920	0.56

Population structure and phylogenetic analysis

Structure Harvester's delta K plot (Fig 4) demonstrated the occurrence of two optimal subpopulations,pattern was consistent with the previous studies [38-40]. The dendrogram generated by UPGMA showed 50 rice accession into 5 clusters (Fig 5). The distribution of the accessions into the five major groups was heterogeneous. In cluster I, the majority of accessions had low albumin and low glutelin content. In Cluster II, most accessions with 4.0-5.0 mg/100 mg glutelin and B-370 were separated from other accessions, indicating higher genetic distance. In cluster III, Kali Kumud, Lal Dodki, Ambemohar Tambda, and Ambemohar Ajra cluster together shows similar type of globulin (0.51-0.6 mg/100 mg) and albumin content (0.36-0.49 mg/100 mg). The local varieties Ambemohar Ajra and Ambemohar Tambda were grouped into separate sub group which shows high glutelin, medium total protein, and a unique allele for RM472_260. In cluster IV, non scented Burma Black and Kolamb were grouped into separate subgroup with similar albumin content (0.3 and 0.32). Kalsal, Kala Krishna, present in different branch revealing that they may have different genetic constitutions. Therefore, these accessions could be served as potential reservoirs of valuable novel genes in rice breeding. Phylogenetic grouping pattern in the current study showed the occurrence of considerable genetic diversity. However, the cluster patterns showed no correspondence with the pre defined population structure based on the protein-associated marker, which may be due to the polygenic nature of the protein.

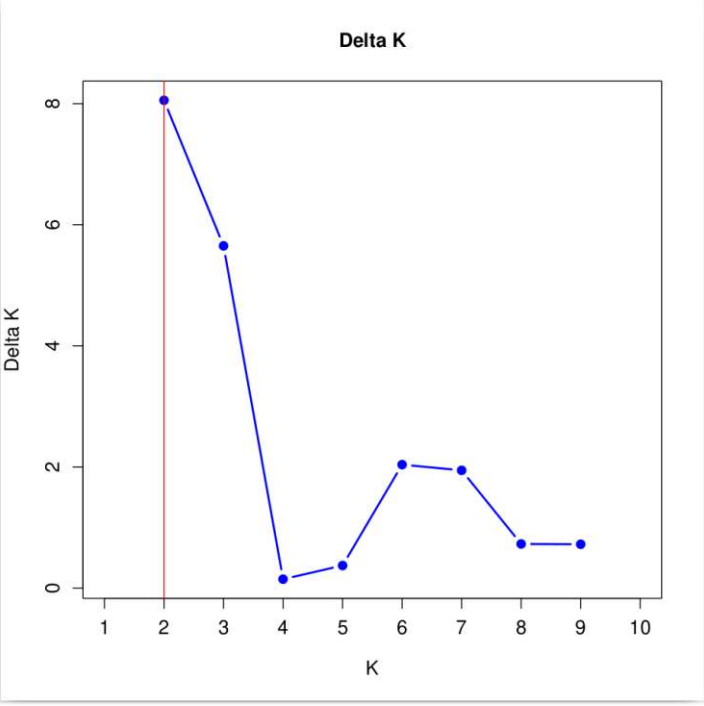


Fig 4. Population structure analysis of 50 rice accession; (a) Sharp peak of Delta K at K = 2 indicates two subpopulations

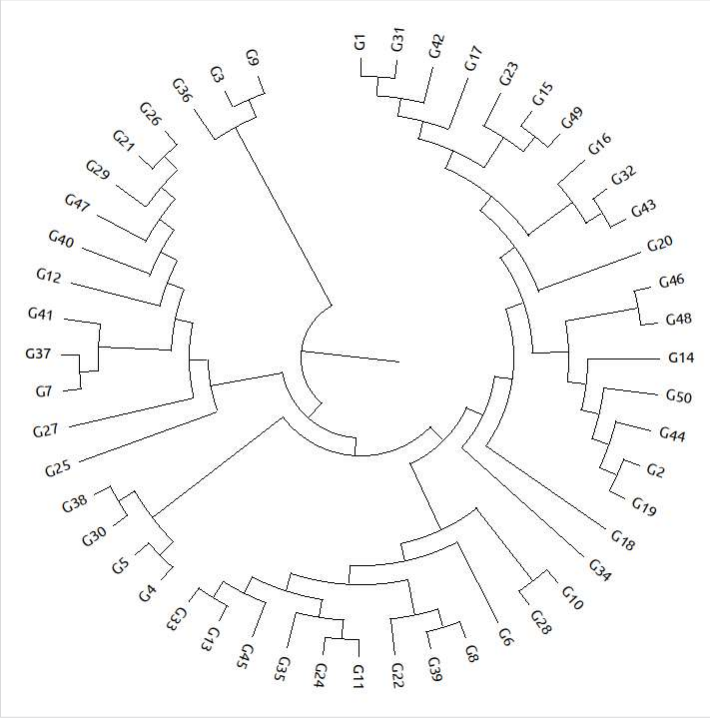


Fig 5. UPGMA dendrogram based on SSR marker in rice accession.

Association analysis between protein fractions and SSR markers

The GLM-QQ plots exhibited a close to perfect distribution of scores from the reference line (Fig 6). 3 SSR marker alleles associated with protein fraction with P<0.05 (Table 2) were detected. Two SSR marker alleles (RM472_330 and RM541_225) were significantly associated with prolamin, whereas one marker allele (RM472_260) was associated with glutelin. Two SSR RM472 and RM541 have previously been identified to be linked with grain protein [16] and prolamin [17] in rice, respectively. We identified two novel trait-specific alleles at this locus: the marker allele RM472_260 showed relatively highest glutelin content, while RM472_330 exhibited relatively lower prolamin in 3 and 6 accession, respectively. Despite at low frequency, these alleles showed significant and meaningful phenotypic variance, suggesting its potential utility in marker-assisted selection. To our knowledge, this appears to be the first report on association analysis of SSP fractions using SSR markers in scented non-basmati rice accessions. Therefore, these results merit further validation on large and diverse accessions to ensure their reliability for the development of biofortified rice.

Table 2: Association between SSR and SSP fraction with P<0.05

Sr. No.	SSR Locus	Trait	P-value	R^2 value (%)
1	RM472_260	Glutelin	0.022	10.7
2	RM472_330	Prolamin	0.046	8.2
3	RM541_225	Prolamin	0.044	8.7

Note: *SSR locus, SSR marker; the size of allele (bp).*

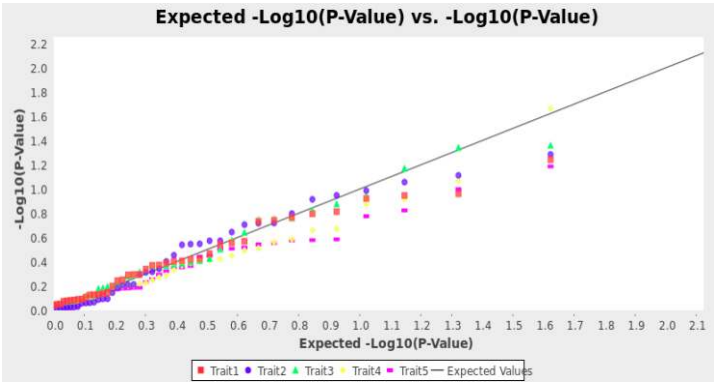


Fig 6: QQ plot of association analysis between SSR marker alleles and SSP in rice accessions under study: Trait1-Albumin, Trait2-Globulin, Trait3-Prolamin, Trait4-Glutelin, Trait5-Total protein content

Conclusions

Considerable variation for the rice SSP in a large collection of local scented non-basmati rice germplasm was observed, which might be useful for the selection of accession in future breeding. Genotyping analysis revealed that SSR markers used in the present study were valuable for assessing genetic diversity but might be effective for the differentiation of rice accessions with higher and lower glutelin content. Furthermore, by marker-trait association analysis, highest glutelin and lower prolamin content linked marker alleles were identified, which can aid marker-assisted breeding to develop biofortified rice, after independent validation on large population in a multi environment trail. Although a less number of SSR markers were used, the present study provided useful preliminary findings of genetic diversity and marker–trait associations, which can be further validated using large number of marker and population.

References

1. Yang, Y. et al. (2019). Natural variation of OsGluA2 is involved in grain protein content regulation in rice. Nature communications, 10(1), 1949.
2. Juliano, BO. (1972) Rice: Chemistry and Technology: The rice caryopsis and its composition. American Association of Cereal Chemists, Inc.
3. Houston, D. F., Iwasaki, T., Mohammad, A., & Chen, L. (1968). Radial distribution of protein by solubility classes in the milled rice kernel. Journal of Agricultural and Food Chemistry, 16(5), 720-724.
4. Friedman, M. (1996). Nutritional value of proteins from different food sources. A review. Journal of agricultural and food chemistry, 44(1), 6-29.
5. Gupta, P. K., & Varshney, R. K. (2000). The development and use of microsatellite markers for genetic analysis and plant breeding with emphasis on bread wheat. Euphytica, 113(3), 163-185.
6. Temnykh, S., DeClerck, G., Lukashova, A., Lipovich, L., Cartinhour, S., & McCouch, S. (2001). Computational and experimental analysis of microsatellites in rice (Oryza sativa L.): frequency, length variation, transposon associations, and genetic marker potential. Genome research, 11(8), 1441.

7. Chattopadhyay, K. et al. (2019). Detection of stable QTLs for grain protein content in rice (Oryza sativa L.) employing high throughput phenotyping and genotyping platforms. Scientific reports, 9(1), 3196.
8. Zhao, L. et al. (2022). Mapping QTLs for rice (Oryza sativa L.) grain protein content via chromosome segment substitution lines. Cereal Research Communications, 50(4), 699-708.
9. Lou, J., Chen, L., Yue, G., Lou, Q., Mei, H., Xiong, L., & Luo, L. (2009). QTL mapping of grain quality traits in rice. Journal of cereal science, 50(2), 145-151.
10. Chen, P. et al. (2023). The genetic basis of grain protein content in rice by genome-wide association analysis. Molecular Breeding, 43(1), 1.
11. Zhang, W. et al. (2019). Association analysis of four storage protein components using microsatellite markers in a japonica rice collection. Chilean journal of agricultural research, 79(1), 3-16.
12. Mathure, S., Jawali, N. and Nadaf A. (2010) Diversity analysis in selected non-basmati scented rice collection. Rice Science, 17, 35-42.
13. Kumamaru, T., Satoh, H., Iwata, N., Omura, T., Ogawa, M., & Tanaka, K. (1988). Mutants for rice storage proteins: 1. Screening of mutants for rice storage proteins of protein bodies in the starchy endosperm. Theoretical and Applied Genetics, 76(1), 11-16.
14. Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry 72:248-254.
15. Doyle, J. J., & Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical bulletin.
16. Peng, B. et al. (2014). OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. Nature communications, 5(1), 4847.
17. Bryant, R. J., Jackson, A. K., Yeater, K. M., Yan, W. G., McClung, A. M., & Fjellstrom, R. G. (2013). Genetic variation and association mapping of protein concentration in brown rice using a diverse rice germplasm collection. Cereal Chemistry, 90(5), 445-452.
18. Yang, Y. et al. (2015). Identification of quantitative trait loci responsible for rice grain protein content using chromosome segment substitution lines and fine mapping of qPC-1 in rice (Oryza sativa L.). Molecular Breeding, 35(6), 130.
19. Zhong, M. et al. (2011). Identification of QTL affecting protein and amino acid contents in rice. Rice Science, 18(3), 187-195.
20. Wang, L. et al. (2008). The QTL controlling amino acid content in grains of rice (Oryza sativa) are co-localized with the regions involved in the amino acid metabolism pathway. Molecular Breeding, 21(1), 127-137.
21. Peakall, R. O. D., & Smouse, P. E. (2006). GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular ecology notes, 6(1), 288-295.
22. Nei, M., & Li, W. H. (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of the National Academy of Sciences, 76(10), 5269-5273.
23. Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. Bioinformatics, 23(19), 2633-2635.
24. Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular ecology, 14(8), 2611-2620.
25. Li, Y. L., & Liu, J. X. (2018). StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods. Molecular ecology resources, 18(1), 176-177.
26. Kumar, S., Stecher, G., Li, M. et al. (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. Molecular Biology Evolution 35:1547-1549.
27. Kim, M., Jeong, Y. (2002). Extraction and electrophoretic characterization of rice proteins. Nutraceuticals Food 7:437-441.
28. Banerjee, S., Chandel, G., Mandal, N., Meena, B. M., & Saluja, T. (2011). Assessment of nutritive value in milled rice grain of some Indian rice landraces and their molecular characterization. Bangladesh Journal of Agricultural Research, 36(3), 369-380.
29. Gomez, KA. (1979). Effect of environment on protein and amylose content of rice. In Proceedings of the workshop on chemical aspects of rice grain quality. International Rice Research Institute Los Baños Laguna, Philippines. 59-68
30. Chen, P. et al. (2018). Genetic basis of variation in rice seed storage protein (albumin, globulin, prolamin, and glutelin) content revealed by genome-wide association analysis. Frontiers in plant science, 9, 612.
31. Zhang, W. et al. (2008). QTL mapping for crude protein and protein fraction contents in rice (Oryza sativa L.). Journal of Cereal Science, 48(2), 539-547.
32. Fiaz, S., Sheng, Z., Zeb, A., Barman, H. N., Shar, T., Ali, U., & Tang, S. (2021). Analysis of genomic regions for crude protein and fractions of protein using a recombinant inbred population in Rice (Oryza sativa L.). *Journal of Taibah University for Science*, *15*(1), 579-588.
33. Pratap, A., Bisen, P., Loitongbam, B., Rathi, S. R., Parmita, P., Singh, B. P., & Singh, P. K. (2020). Assessment of genetic diversity in rice germplasm (Oryza sativa L.) using SSR markers. Intenational Journal of Current. Microbiology, 9, 3346-3355.
34. Singh, N. et al. (2016). Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. BMC genetics, 17(1), 127.
35. Vanniarajan, C., Vinod, K. K., & Pereira, A. (2012). Molecular evaluation of genetic diversity and association studies in rice (Oryza sativa L.). Journal of genetics, 91(1), 9-19.
36. Nachimuthu, V. et al. (2015). Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in Oryza sativa. Rice, 8(1), 30.
37. Suvi, W. T., Shimelis, H., Laing, M., Mathew, I., & Shayanowako, A. I. T. (2020). Assessment of the genetic diversity and population structure of rice genotypes using SSR markers. Acta Agriculturae Scandinavica, Section B—Soil & Plant Science, 70(1), 76-86.
38. Fan, J., Zhang, X., Cheng, B., Zhang, X., Xu, X., Zeng, B., & Gong, J. (2026). Genetic diversity analysis of newly approved rice cultivars in southern China from 2014 to 2022. Genetic Resources and Crop Evolution, 73(4), 171.
39. Salem, K. F. et al. (2024). Analysis of genetic diversity, population structure and phylogenetic relationships of rice (Oryza sativa L.) cultivars using simple sequence repeat (SSR) markers. Genetic Resources and Crop Evolution, 71(5), 2213-2227.
40. Anandan, A., Anumalla, M., Pradhan, S. K., & Ali, J. (2016). Population structure, diversity and trait association analysis in rice (Oryza sativa L.) germplasm for early seedling vigor (ESV) using trait linked SSR markers. PloS one, 11(3), e0152406.