PLANT SCIENCE TODAY
Vol 13(1): 1-9
https://doi.org/10.14719/pst.11374

eISSN 2348-1900

RESEARCH ARTICLE

Molecular screening of parental lines and trait association for grain protein content in F_2 population of rice (*Oryza sativa* L.)

Gugulothu Prasanna[1,2*], Lavuri Krishna[3], Jukanti Aravind Kumar[3], Kuna Aparna[4], Sadu Narendar Reddy[5], Vishal Kumar Gupta[6] & Ajaz A Lone[7*]

[1]Department of Genetics and Plant Breeding, College of Agriculture, Professor Jayashankar Telangana Agricultural University (PJTAU), Rajendranagar 500 030, Hyderabad, India

[2]Department of Genetics and Plant Breeding, School of Agriculture, S R University, Warangal 506 371, Telangana, India

[3]Department of Genetics and Plant Breeding, Institute of Rice Research, Professor Jayashankar Telangana Agricultural University (PJTAU), Rajendranagar 500 030, Hyderabad, India

[4]Department of Food and Nutrition, Quality control lab, Professor Jayashankar Telangana Agricultural University (PJTAU), Rajendranagar 500 030, Hyderabad, India

[5]Department of Crop Physiology, College of Agriculture, Professor Jayashankar Telangana Agricultural University (PJTAU), Rajendranagar 500 030, Hyderabad, India

[6]Department of Genetics and Plant Breeding, M S Swaminathan School of Agriculture, Centurion University of Technology and Management, Odisha, India

[7]Dryland Agriculture Research Station (DARS), Rangreth, Srinagar 190 001, Jammu and Kashmir, India

*Correspondence email - prasannagricos@gmail.com; ajazlone@skuastkashmir.ac.in

Received: 20 August 2025; Accepted: 31 December 2025; Available online: Version 1.0: 31 January 2026; Version 2.0: 05 February 2026

Cite this article: Gugulothu P, Lavuri K, Jukanti AK, Kuna A, Sadu NR, Vishal KG, Ajaz AL. Molecular screening of parental lines and trait association for grain protein content in F_2 population of rice (*Oryza sativa* L.). Plant Science Today. 2026; 13(1): 1-9. https://doi.org/10.14719/pst.11374

Abstract

Rice grain protein is the second most abundant component of milled rice grain and has been extensively studied due to its significant role as a nutrient. There are very limited high grain protein varieties identified. Therefore, studying the genetics of grain protein, yield and quality traits is essential for developing a breeding program that will increase yield while maintaining rice quality. The key purpose of this research work was to identify lines with high protein content. Three diverse parents (RDR 1295, JAK 686 and JAK 685) were screened using 27 already reported grain protein markers in rice. Among these markers, 6 markers were linked with JAK 686 parent, while a single marker was linked to JAK 685 parent. So, these identified high-protein donors (JAK 686 and JAK 685) were then crossed with low-protein genotype (RDR 1295) during kharif 2020. In this study, F_2 segregating populations from 2 different cross combinations in rice RDR 1295 x JAK 686 (Cross-I) and RDR 1295 x JAK 685 (Cross-II) were studied. Character association studies revealed that plant height (0.108), kernel length (0.077), kernel width (0.025) and L/B ratio (0.045) have shown a clear positive correlation with grain protein content (GPC) for Cross-I and in Cross-II, panicle length (0.041), test weight (0.065), kernel length (0.138) and kernel width (0.101) have positive association with protein content. Path coefficient analysis further indicated that, in Cross-I, kernel length (1.247) exhibited the highest direct and positive effect on GPC. In Cross-II, the number of productive tillers per plant (0.183) showed highest direct and positive influence on GPC. The key traits identified include kernel length and kernel width, as both high correlation coefficients and direct effects displayed strong association with GPC.

Keywords: grain protein; molecular markers; protein malnutrition; segregating population

Introduction

Grain protein content (GPC) in rice is a vital nutritional trait, as rice serves as a staple food for over half the global population and enhancing its protein content can significantly improve dietary protein intake, especially in developing countries. Protein content of the rice grain plays an important role in human nutrition, as protein malnutrition can hinder normal growth and physiological development particularly in regions where rice is mostly consumed. Rice contributes significantly to the global dietary energy supply, surpassing maize and wheat in importance (1). However, despite providing 14 % of the world's protein, rice lacks essential amino acids. Genetic diversity in protein content presents an opportunity for breeding programs aimed at improving nutritional value. As the world population increases, the demand for nutrient-rich foods will rise. While rice is a staple food, it contains limited amounts of key micronutrients and essential compounds (2). Despite having less protein than many other crops, it remains an important part of the human diet (3). Main grain quality in rice, especially protein level, significantly affects consumer approval and economic value in the export markets. Although rice has an approximate protein value of 8.5 %, milled rice that is often consumed typically has an average protein content of 6–8 %. Protein content with < 8 % are defined as low protein, those with 8–10 % protein content as moderate and those with more than 10 % protein content as high.

Although molecular markers and quantitative trait loci associated with GPC have been extensively reported in rice, their effective integration with phenotypic trait relationships in segregating

populations remains limited. The novelty of the present study lies in the combined use of marker-based parental screening and comprehensive character association and path coefficient analyses in two F_2 populations sharing a common parent. This integrated approach enables the identification of high-protein genotypes while simultaneously elucidating the direct and indirect influence of yield and grain quality traits on protein content. The findings provide practical insights for developing nutritionally enhanced rice varieties without compromising agronomic performance.

Improving the protein content of rice has become a major focus in recent genetic studies, with QTL mapping being employed to identify loci linked to protein content (4–7). QTLs governing GPC have been observed on almost all chromosomes, with notable clusters found on chromosomes 1, 2, 6, 7, 10, 11 and 12 (8). This study aimed to identify high-protein genotypes by screening 3 parents with molecular markers associated with protein content. To further understand the genetic relationships, character association and path coefficient analysis were studied on 2 F_2 populations from Cross I (RDR 1295 x JAK 686) and Cross-II (RDR 1295 x JAK 685) to determine how yield and quality traits influence GPC. Greater emphasis has now been placed on identifying the research gap, particularly the need to integrate molecular marker-based screening with trait association and path coefficient analysis in segregating populations.

Materials and Methods

In this study, 27 SSR markers, previously mapped on chromosomes 1, 2, 3, 4, 7, 8, 10 and 12 (8), were screened with 3 parents (RDR 1295, JAK 686 and JAK 685). Owing to their high protein content JAK 686 and JAK 685 were crossed with low protein genotype (RDR 1295) in rabi, 2020 to develop experimental material. During kharif, 2021 F1's developed and F_1's was also selfed simultaneously to generate F_2. List of parents and their prominent characteristics are presented in (Table 1). The role of molecular markers has been clarified by explicitly linking parental marker screening with phenotypic performance in the F_2 populations. Although formal marker trait association or segregation analysis was not the primary objective of this study, the F_2 populations were extensively evaluated for GPC and related traits to assess the consistency of parental marker information.

CTAB technique was used to extract genomic DNA from young leaves (21 DAS). SSR markers were utilized subsequently following DNA quantification, PCR amplification and agarose gel electrophoresis. The Syngene Ingenius gel documentation system was used to record the amplified PCR products after they were separated on 3.0 % Seakem® LE agarose gel. List of markers used for screening parents for protein content are presented in Table 2. During the rabi, 2020-21 crossing was done between RDR 1295 with both the protein donors JAK 686 and JAK 685 to generate F_1's. During kharif 2021, the F_1 plants were selfed to generate the F_2 population. The F_2 seeds from these crosses were sown in nursery and then transplanted in the main field as single seedlings. The data was recorded on 12 traits in 3 replications: days to 50 % flowering, plant height, panicle length, number of productive tillers per plant, number of filled grains per panicle, 1000 grain weight, grain yield per plant, kernel length, kernel breadth, kernel L/B ratio, amylose content and GPC for 50 F_2 plants in each replication for yield and quality. A total of 150 plants provided sufficient degrees of freedom for reliable estimation of correlation coefficients and path analysis, as supported by standard quantitative genetic and biometric practices in segregating populations. The TNAU-STAT statistical package was used for determination of character association and path coefficient analysis (9). The experimental field conditions during the kharif season, including uniform soil fertility management and recommended nitrogen application practices followed across all plots. Since all genotypes were evaluated under the same environmental and management conditions, environmental variation was minimised, allowing genetic effects to be reliably assessed.

The nitrogen content was determined using the Micro-Kjeldahl method (10). For this, 2 g of powdered rice flour was taken in a 100 mL Micro-Kjeldahl flask, along with a 5:1 ratio of potassium sulphate and cupric sulphate to speed up the reaction. Then, 10 mL of concentrated H_2SO_4 was added to the mixture, which was allowed to digest for 3 hr. The mixture was transferred to a distillation apparatus after digestion. Into a 250 mL conical flask, 15 mL of a 4 % solution of boric acid was poured and finally a mixed indicator was added. To absorb released ammonia (NH_3), the condenser tip was submerged in the boric acid solution. The Kjeldahl flask was then automatically filled with 25 mL of 40 % NaOH. After starting the distillation process, it will take around 9 min to collect 10 mL of distillate in the conical flask. The distillate was then titrated using standard 0.02N H_2SO_4 until bluish green to pink colour.

To estimate GPC, the nitrogen content of the rice was multiplied by a protein conversion factor (11).

Protein (%) = Nitrogen (%) x 6.25

Results and Discussion

Three diverse parental lines RDR 1295, JAK 685 and JAK 686 were screened using 27 previously reported grain protein markers, including both milled and brown rice-specific markers. In this investigation, 27 SSR markers associated with GPC that had previously been mapped in separate mapping populations by different workers were used. Out of 27 SSR markers screened, 6 markers (RM297, RM493, RM562, RM12532, RM257 and RM209) were found to be linked with JAK 686 and single marker RM 309 was found to be associated with JAK 685. Gel pictures for identifying genetic markers reported in rice (Fig. 1).

Correlation analysis

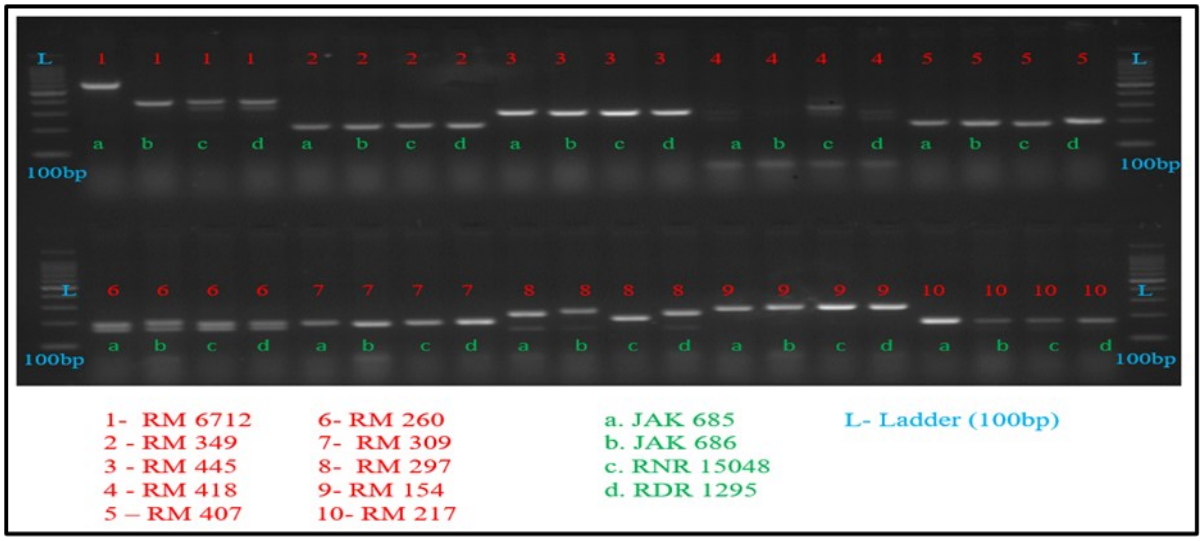
The results of character association reveal the relationships between various traits in 2 crosses RDR 1295 x JAK 686 and RDR 1295 x JAK 685 are presented in Table 3 and 4. In the Cross-I, days to 50 % flowering showed a significant positive correlation associated with number of filled grains per panicle. Traits such as number of productive tillers per plant, kernel width, kernel length/width ratio and GPC were observed to correlate negatively with days to 50 % flowering. Likewise, in Cross-II, days to 50 % flowering were highly

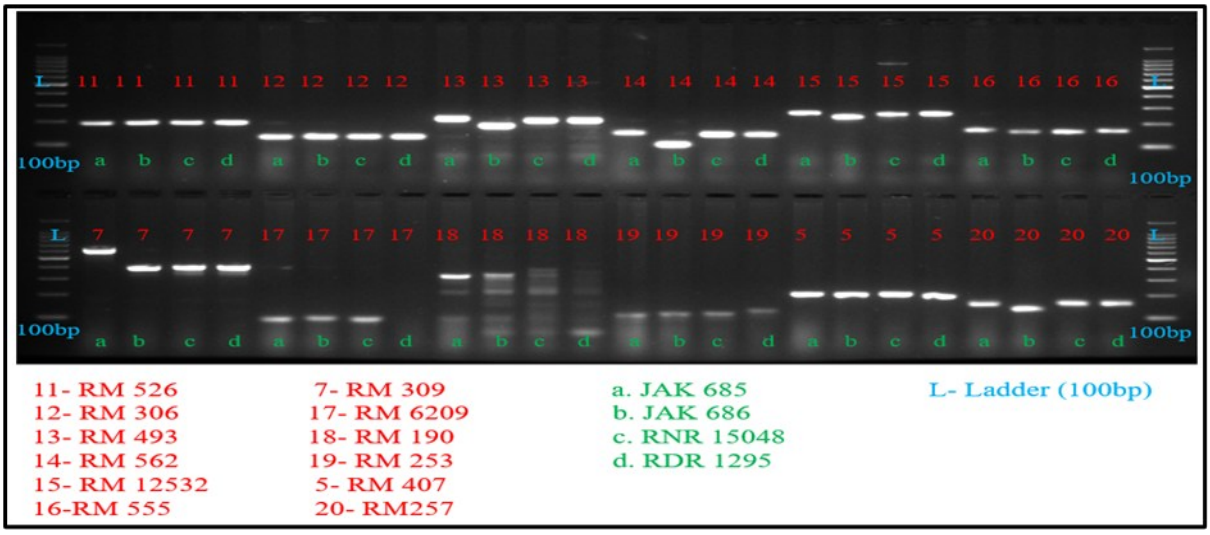
Table 1. List of parents and their prominent characteristics

Parents	Source	Characters
RNR 15048	IRR, ARI, Hyderabad	Short duration, blast resistant, short slender, low glycemic index and moderate protein content (8.0 %)
RDR 1295	RS & RRS, Rudrur	Medium duration, medium slender and low protein content (7.0 %)
JAK 686	ICAR-IIRR, Hyderabad	Early duration and high protein content (~12.5 %)
JAK 685	ICAR-IIRR, Hyderabad	Medium duration and high protein content (10.5 %)

Table 2. List of molecular markers used for screening of parents for grain protein content in rice

Sl. No	Marker name	Rice type	Forward sequence	Reverse sequence	Chromosome number	References
1	**RM 6712**	Milled rice	CCAGCATCATCATTGTCATCATCG	ATCCATCCAGCAGGAGAAACAGG	3	(12)
2	**RM 349**	Milled rice	GCTCGTCTTTCGTCTCTGTGTGC	AAGTACGCGCTGTCCATCATCC	4	(13)
3	**RM 445**	Milled rice	GCCTTGTCCCTAGCTAATCATTTCC	GGCTCGAATCTACGAACAACAGC	7	(13)
4	**RM 418**	Milled rice	CGATCGAGCATCAACACAACG	GACGTATCGCGTATCGTCATGC	7	(13)
5	**RM 407**	Milled rice	GACTACGAGACGAGTGATTTGAACC	GCGTGGGAAATGACTAGGAGTAGG	8	(12)
6	**RM 260**	Milled rice	GATAGAGGATTGGGTGCGTGTGC	TACGCCAACCAATTCCAAACACC	12	(12)
7	**RM 309**	Milled rice	CACGCACCTTTCTGGCTTTCAGC	AGCAACCTCCGACGGGAGAAGG	12	(12)
8	**RM 297**	Brown rice	ACAGGGCTATGCAGACACAGTGC	AGCAAGCGAAGGGAAGTGACC	1	(14)
9	**RM 154**	Brown rice	GACGGTGACGCACTTTATGAACC	CGATCTGCGAGAAACCCTCTCC	2	(13)
10	**RM 7217**	Brown rice	AGGATGACACGTGGCGACTTAGG	CAACGGACGGGATTTCAGTACC	10	(15)
11	**RM 526**	Milled rice	TACAGGCAGAAAGAGCAGTTCA	CAGCGTTCTTCATCATTTCATC	2	(16)
12	**RM 306**	Brown rice	GGACTCCGGCAGATCATCA	CTGGTTTCATCATGTGTGCCTA	2	(16)
13	**RM493**	Milled rice	GTACGTAAACGCGGAAGGTGACG	CGACGTACGAGATGCCGATCC	1	(13)
14	**RM562**	Milled rice	GGAAAGGAAGAATCAGACACAGAGC	GTACCGTTCCTTTCGTCACTTCC	1	(13)
15	**RM12532**	Milled rice	GCATGGAGACCTTAATATCCAACTCC	GATAGACGATCGAGTTGGGTTGC	2	(17)
16	**RM555**	Milled rice	TTGACATGCGAAATGGAGATGG	TTGGATCAGCCAAAGGAGACC	2	(17)
17	**RM6209**	Milled rice	GGCTTCGTCTTCCTCATCTCG	TCCATCCATAGACTTGTTGACTGC	5	(12)
18	**RM190**	Milled rice	GGAGTGGTCAAATAAGTTGCTTGC	GGCTCTTACTCGTCAATGAACTCC	6	(14)
19	**RM253**	Milled rice	CCATCTCTGCCTCTGACTCACC	TCCTTCAATGGTCGTATCTTCTCC	6	(14)
20	**RM257**	Milled rice	CCGTGCAACTTAAATCCAAACAGG	GGAATCCTATATGAGCCAGTGATGG	9	(13)
21	**RM184**	Brown rice	AACGAAGACGATCGAGAGGAAGC	CCATCTCCACCCAACCAAACC	10	(18)
22	**RM209**	Milled rice	ACAAAGGGAGTATGTCCTCATCC	GGAGGTAGCTCTATCGTTGTCG	11	(14)
23	**RM229**	Milled rice	ACGACTATCAACACAACACTGCAACC	CGCTCGCACATCTTATCCTCTCC	11	(14)
24	**RM226**	Brown rice	GAAGCTAAGGTCTGGGAGAAACC	AATGGCCTTAACCAAGTAGGATGG	1	(14)
25	**RM23914**	Brown rice	GAGGATCCTTACCATCAAACTTCG	CCAAGAACCTGCATTCTTCAAGG	9	(15)
26	**RM506**	Brown rice	CAGTATCCATGTCCTTGCTTAACG	AATAGATTGAGTGGTCGACTGAGG	8	(15)
27	**RM25128**	Brown rice	CCCAAACGAAATGTTTCAGACG	TATCGCATCCGATCTTACCTTCC	10	(15)





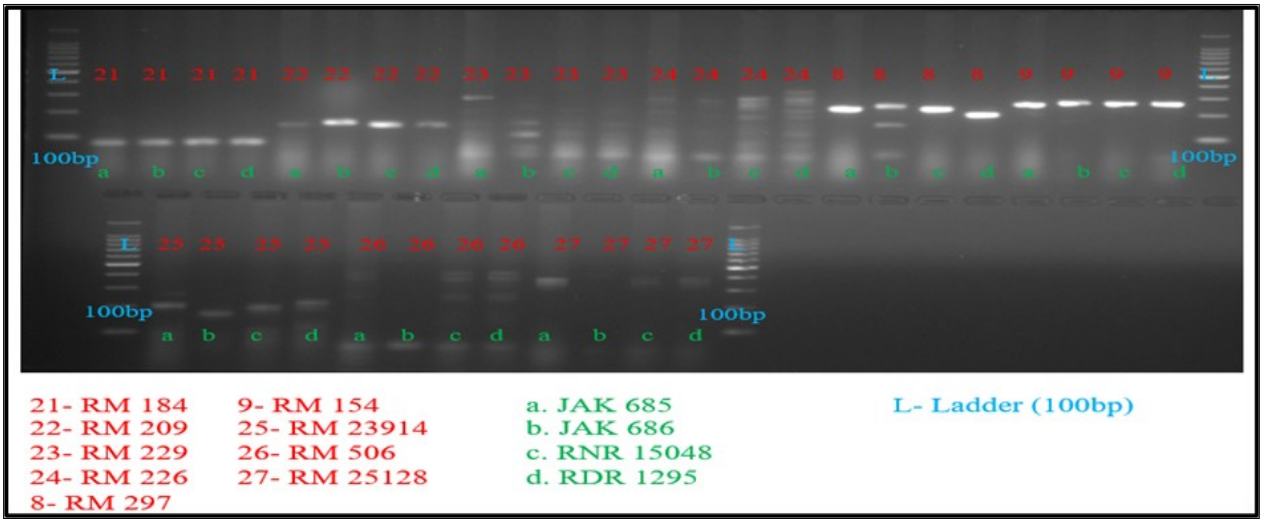


Fig. 1. Gel picture for screening of parents with reported protein markers in rice.

associated with plant height and amylose content. These findings confer similarity with similar studies on plant height and for single plant yield (19–21).

Association studies revealed that in Cross-I (RDR 1295 x JAK 686), plant height was strongly correlated with amylose content, test weight and panicle length, consistent with previous findings (22). In Cross-II (RDR 1295 x JAK 685), plant height exhibited a highly significant and positive correlation with kernel length. Additionally, in Cross-I (RDR 1295 x JAK 686), panicle length showed a negative non-significant association with both the number of filled grains per panicle and protein content. A negative non-significant correlation was also observed in Cross-II (RDR 1295 x JAK 685) between panicle length and amylose content. The differences are attributed to variation in the genetic background of the donor parents (JAK 686 and JAK 685), which likely harbor distinct alleles governing plant height, grain quality and protein related traits. These allelic differences may influence pleiotropic effects, resulting in cross-specific correlations.

The number of productive tillers per plant had significant and negative correlation with number of filled grains per panicle, test weight, kernel width and protein content and positive significant association with single plant yield in Cross-I. In contrast, this trait displayed a positive and non-significant association with both single plant yield and protein content in Cross-II.

In Cross-I, negative and non-significant correlations were observed for the number of filled grains per panicle with GPC, amylose content and the kernel width, but it was significantly and positively correlated with single plant yield. These findings were found to be in direct agreement with early works who studied single plant yield (23–26). In Cross-II, this trait was significantly and negatively correlated with protein content but positively correlated with amylose content and single plant yield.

The negative correlation of test weight with length- to-width ratio and protein content, whereas, it had positive and significant association with single plant yield, kernel length and kernel width in Cross-I. There was a positive and significant association of test weight with the single-plant yield, kernel width and kernel length in Cross-II.

The negative association of single-plant yield with the protein content and a significant positive association with kernel length was observed in Cross-I. The present results were similar to those of early works on protein content (20). In Cross-II, a negative non-significant association was recorded for single-plant yield versus protein content, while positive significant association exists with amylose content and kernel length.

In Cross-I, protein content had a positive non-significant association with kernel length, kernel width and the L/B ratio whereas in Cross-II had a positive significant association with number of productive tillers per plant and positive non-significant association with number of filled grains per panicle. The L/B ratio

Table 3. Correlation coefficient for grain protein content with yield, yield related and grain quality traits in F_2 population of cross-I (RDR 1295× JAK 686)

Traits	DFF	PH	PL	PT	GPP	TW	SPY	KL	KB	L/B	AC	PC
DFF	1.000	-0.646**	-0.329**	-0.065	0.408**	-0.100	0.08	-0.168*	-0.018	-0.134	-0.532**	-0.059
PH		1.000	0.510**	0.092	-0.251**	0.196*	0.021	0.130	0.066	0.066	0.465**	0.108
PL			1.000	0.018	-0.051	0.214**	0.111	0.070	0.007	0.051	0.215**	-0.018
PT				1.000	-0.245**	-0.105	0.239**	0.048	-0.013	0.055	0.068	-0.037
GPP					1.000	0.042	0.482**	0.077	-0.034	0.093	-0.156	-0.091
TW						1.000	0.261**	0.233**	0.307**	-0.024	0.153	-0.060
SPY							1.000	0.122*	0.078	0.055	0.144	-0.036
KL								1.000	0.198*	0.715**	0.149	0.077
KB									1.000	-0.541**	-0.087	0.025
L/B										1.000	0.187*	0.045
AC											1.000	-0.039
PC												1.000

DFF- Days to 50 % flowering, PH- Plant height, PL-Panicle length, PT-Number of productive tillers per plant, GPP-Number of grains per panicle, TW- 1000 grain weight, SPY- Single plant yield, KL- Kernel length, KB- Kernel breadth, L/B- Kernel L/B ratio, AC- Amylose content, PC-Protein content.

*Significant at 5 % level, ** Significant at 1 % level

Table 4. Correlation coefficient for grain protein content with yield, yield related and grain quality traits in F_2 population of cross-II (RDR 1295× JAK 685)

Traits	DFF	PH	PL	PT	GPP	TW	SPY	KL	KB	L/B	AC	PC
DFF	1.000	0.660**	-0.024	-0.103	0.001	-0.124	-0.206*	-0.513**	-0.149	-0.168*	0.174*	-0.074
PH		1.000	0.094	0.014	0.048	0.072	0.158	0.412**	0.148	0.112	-0.129	-0.018
PL			1.000	0.104	0.02	0.114	0.107	0.044	-0.016	0.033	-0.161*	0.041
PT				1.000	-0.31**	0.110	0.270**	0.129	0.002	0.070	-0.194	0.181*
GPP					1.000	-0.009	0.560**	0.020	-0.111	0.122	0.233**	-0.220**
TW						1.000	0.448**	0.257**	0.319**	-0.114	0.129	0.065
SPY							1.000	0.226**	0.116	0.049	0.173*	-0.090
KL								1.000	0.103	0.521**	-0.179*	0.138
KB									1.000	-0.790**	-0.068	0.101
L/B										1.000	-0.034	-0.006
AC											1.000	-0.083
PC												1.000

DFF- Days to 50 % flowering, PH- Plant height, PL-Panicle length, PT-Number of productive tillers per plant, GPP-Number of grains per panicle, TW- 1000 grain weight, SPY- Single plant yield, KL- Kernel length, KB- Kernel breadth, L/B- Kernel L/B ratio, AC- Amylose content, PC-Protein content.

*Significant at 5 % level, ** Significant at 1 % level

was positively and significantly correlated with amylose content in Cross-I , while in Cross-II , it revealed a negative non-significant correlation with both amylose and protein content.

Path analysis

The path analysis assists us in understanding the causes and effects different traits have on each other. In the present study, correlations were computed to estimate direct and indirect effects at a phenotypic level using the GPC as a dependent variable based on F_2 population data. Summaries of path analysis results are presented in (Tables 5 , Table 6).

Days to 50 % flowering in Cross-I exhibited a negative direct effect on protein content (-0.102), suggesting that early flowering is associated with higher GPC and positive indirect effects through panicle length, number of productive tillers per plant, number of filled grains per panicle, test weight, single plant yield, kernel width, L/B ratio and amylose content. Also, in Cross-II, showed a negative direct effect (-0.080) on protein content. However, positive indirect effects *via* plant height, length/breadth ratio, amylose content and single plant yield in Cross-II.

In Cross-I, plant height had a positive direct effect (0.220) on protein content, indicating that taller plants tend to have higher protein levels. Despite this, it exerted multiple negative indirect effects through panicle length, single plant yield, kernel width, test weight, L/B ratio, number of filled grains per panicle and amylose content. In Cross-II, contrarily, plant height had a negative direct effect (-0.125) on protein content. The negative effect was further reinforced by indirect effects *via* amylose content, single plant yield and L/B ratio. This suggests that plant height may be less desirable for improving protein due to its negative effect in Cross-II.

Panicle length in Cross-I showed a negative direct effect on protein content (-0.111), indicating that longer panicles may be associated with reduced protein content. In contrast, panicle length displayed a positive direct effect with protein content (0.048), suggesting a potentially supportive role in improving protein levels in this genetic background in Cross-II.

In Cross-I number of productive tillers per plant exhibited a negative direct effect on protein content (-0.036) and positive indirect effects through days to 50 % flowering, plant height, test weight, single plant yield, kernel length and kernel width. In Cross-II, it had a positive and direct effect (0.183) on GPC. However, negative indirect effects are mediated via single plant yield, length/breadth ratio, plant height and amylose content. This suggests that while productive tillers directly contribute to protein content in Cross-II, their interaction with other traits may reduce the overall gain. Number of productive tillers per plant had negative effect it will reduce GPC in Cross-I.

Number of filled grains per panicle displayed a negative direct effect on protein content (-0.137). Despite the negative direct effect, several other traits contributed positively through indirect effects such as single plant yield, kernel length, panicle length, kernel width, length/breadth ratio, number of productive tillers per plant and amylose content in Cross-I. These traits, while not always exerting direct effects, indirectly enhanced protein content. Protein content was negatively influenced by an overall direct effect (-0.056) in Cross-II. Positive indirect effects were observed through amylose content, plant height and panicle length. These results emphasise the compensatory nature of indirect effects, helping buffer the negative direct effect on protein content in Cross-II.

In Cross-I test weight had a negative direct effect on GPC (-0.115), indicating that increased grain weight may reduce GPC. In Cross-II (RDR 1295 x JAK 685), it exerted a positive direct effect on protein content (0.044). However, it also exhibited negative indirect effects through plant height and single plant yield. The contrast between the 2 crosses shows that the role of test weight is genotype-dependent, with both positive and negative interactions.

In Cross-I, single plant yield demonstrated a negative direct effect on GPC (-0.026), indicating that higher yield may reduce the protein content. In Cross-II, it showed a negative direct effect (-0.183) on GPC. Additionally, negative and indirect effects were observed via plant height, number of filled grains per panicle, length/breadth ratio (L/B) ratio in Cross-II. These results suggest a consistent negative association between yield and protein content in both crosses, highlighting the yield and quality substitution.

Kernel length in Cross-I had a strong positive direct effect on protein content (1.247), making it a major contributing trait to increased protein levels. The high direct effect of kernel length on GPC observed in Cross-I may be attributed to strong genetic linkage or pleiotropic effects between kernel size and protein accumulation

in this population. It also showed a moderate positive effect with protein content (0.158) in Cross-II, indicating a favourable but less pronounced effect. Kernel length is a beneficial trait for protein improvement, particularly in Cross-I.

In Cross-I, length/breadth ratio exerted negative direct effect on protein content (-1.375). Despite this, it contributed positive indirect effects through several traits number of filled grains per panicle, days to 50 % flowering, plant height, single plant yield, kernel length, kernel length, kernel breadth. This suggests that while L/B ratio directly reduces protein content, it indirectly improves protein content with yield and grain quality traits. In Cross-II protein content was found to have a negative direct effect on the L/B ratio (-0.018), indicating a reciprocal relationship between protein content and L/B ratio.

Amylose content had a negative direct effect on GPC (-0.144), suggesting that higher amylose levels may slightly reduce protein content in Cross-I and positive indirect effects through days to 50 % flowering, plant height, single plant yield, kernel length and kernel width. It had a positive direct effect on protein content (0.030) in cross-II and contributed positive indirect effects through plant height, test weight, length/breadth ratio. This indicates that amylose content plays positive role in Cross-II, both directly and indirectly, compared to its effect in Cross-I.

Conclusion

Out of 27 SSR markers screened, 6 markers (RM297, RM493, RM562, RM12532, RM257 and RM209) were found to be linked with JAK 686 and single marker RM 309 was found to be associated with JAK 685. It is concluded that both JAK 685 and JAK 686 have high GPC, indicating that they could be useful as donors in hybridisation programme. In Cross-I (RDR 1295 x JAK 686), kernel length, plant height, L/B ratio and the kernel width, all had positive effects on GPC. Meanwhile, in Cross-II (RDR 1295 x JAK 685) I, test weight, number of filled grains per panicle, number of productive tillers per plant, kernel width and kernel length were all positively correlated with protein content. Whereas amylose content had negative effect on protein content in cross-I and positive direct effect in Cross-II (RDR 1295 x JAK 685). The identification of JAK 685 and JAK 686 as high GPC donors, along with associated SSR markers, provides valuable resources for marker-assisted selection in rice breeding programme. Traits such as kernel length and kernel width, which exhibited strong positive correlations and direct effects on GPC, can serve as effective phenotypic selection criteria in early segregating generations. Furthermore, mapping populations derived from the RDR 1295 background offer a promising platform for validating and fine-mapping QTLs governing GPC, thereby facilitating the development of nutritionally enhanced rice varieties without compromising agronomic performance.

Table 5. Phenotypic path coefficients for grain protein content with yield and grain quality traits in F_2 population of cross-I (RDR 1295 × JAK 686)

Days to 50 % flowering vs Protein content	r = -0.059	**Productive tillers per plant vs Protein content**	r = -0.037
Direct effect	-0.102	Direct effect	-0.036
Indirect effects through PH	-0.142	Indirect effects through DFF	0.006
PL	0.036	PH	0.020
PT	0.002	PL	-0.002
GPP	0.056	GPP	-0.033
TW	0.011	TW	0.012
SPY	0.007	SPY	0.007
KL	-0.209	KL	0.060
KB	0.018	KB	0.013
L/B ratio	0.185	L/B ratio	-0.075
AC	0.077	AC	-0.009
Plant height vs Protein content	r = 0.108	**Filled grains per panicle vs Protein content**	r = -0.091
Direct effect	0.220	Direct effect	-0.137
Indirect effects through DFF	0.066	Indirect effects through DFF	-0.041
PL	-0.057	PH	-0.055
PT	0.003	PL	0.005
GPP	-0.034	PT	0.009
TW	-0.022	TW	-0.004
SPY	-0.001	SPY	0.017
KL	0.162	KL	0.096
KB	-0.063	KB	0.032
L/B ratio	-0.090	L/B ratio	-0.128
AC	-0.067	AC	0.022
Panicle length vs Protein content	r =-0.018	**Test weight vs Protein content**	r = -0.060
Direct effect	-0.111	Direct effect	-0.115
Indirect effects through DFF	0.033	Indirect effects through DFF	0.010
PH	0.112	PH	0.043
PT	-0.007	PL	-0.003
GPP	-0.007	PT	-0.010
TW	-0.024	GPP	0.092
SPY	0.008	SPY	-0.033
KL	0.088	KL	0.212
KB	-0.007	KB	-0.121
L/B ratio	-0.070	L/B ratio	-0.080
AC	-0.031	AC	-0.006

Single plant yield vs Protein content	r =-0.036	**L/B ratio vs Protein content**	r = 0.045
Direct effect	-0.026	Direct effect	-1.375
Indirect effects through DFF	-0.027	Indirect effects through DFF	0.013
PH	-0.011	PH	0.014
PL	-0.003	PL	-0.005
PT	-0.010	PT	-0.002
GPP	0.092	GPP	0.012
TW	-0.033	TW	0.002
KL	0.212	SPY	0.001
KB	-0.121	KL	0.892
L/B ratio	-0.080	KB	0.517
AC	-0.006	AC	-0.027
Kernel length vs Protein content	r = 0.077	**Amylose content vs Protein content**	r = -0.039
Direct effect	1.247	Direct effect	-0.144
Indirect effects through DFF	0.017	Indirect effects through DFF	0.054
PH	0.028	PH	0.102
PL	-0.007	PL	-0.024
PT	-0.001	PT	-0.002
GPP	0.010	GPP	-0.021
TW	-0.027	TW	-0.017
SPY	0.004	SPY	0.001
KB	-0.189	KL	0.186
L/B ratio	-0.983	KB	0.083
AC	-0.021	L/B ratio	-0.257
Kernel breadth vs Protein content	r =0.025		
Direct effect	-0.957		
Indirect effects through DFF	0.001		
PH	0.014		
PL	-0.008		
PT	0.005		
GPP	-0.004		
TW	-0.035		
SPY	0.003		
KL	0.247		
L/B ratio	0.744		
AC	0.012		

Table 6. Phenotypic path coefficients for grain protein content with yield and grain quality traits in F_2 population of Cross-II (RDR 1295 × JAK 685)

Days to 50 % flowering vs Protein content	r = -0.074	**Productive tillers per plant vs Protein content**	r = 0.181
Direct effect	-0.080	Direct effect	0.183
Indirect effects through PH	0.082	Indirect effects through DFF	0.008
PL	-0.001	PH	-0.001
PT	-0.019	PL	0.005
GPP	-0.005	GPP	0.017
TW	-0.005	TW	0.004
SPY	0.037	SPY	-0.049
KL	-0.081	KL	0.020
KB	-0.011	KB	0.002
L/B ratio	0.003	L/B ratio	-0.001
AC	0.005	AC	-0.006
Plant height vs Protein content	r = -0.018	**Filled grains per panicle vs Protein content**	r = -0.220
Direct effect	-0.125	Direct effect	-0.056
Indirect effects through DFF	0.053	Indirect effects through DFF	-0.007
PL	0.004	PH	0.002
PT	0.002	PL	0.001
GPP	0.001	PT	-0.056
TW	0.003	TW	-0.003
SPY	-0.029	SPY	-0.089
KL	0.065	KL	-0.003
KB	0.011	KB	-0.012
L/B ratio	-0.002	L/B ratio	-0.002
AC	-0.004	AC	0.007
Panicle length vs Protein content	r =-0.041	**Test weight vs Protein content**	r = 0.065
Direct effect	0.048	Direct effect	0.044
Indirect effects through DFF	0.001	Indirect effects through DFF	0.010
PH	-0.011	PH	-0.009
PT	0.019	PL	0.005
GPP	-0.001	PT	0.020
TW	0.005	GPP	0.004
SPY	-0.019	SPY	-0.082
KL	0.007	KL	0.040
KB	-0.001	KB	0.025
L/B ratio	-0.006	L/B ratio	0.002
AC	-0.004	AC	0.004

Single plant yield vs Protein content	r =-0.090	**L/B ratio vs Protein content**	r = -0.006
Direct effect	-0.183	Direct effect	-0.018
Indirect effects through DFF	0.016	Indirect effects through DFF	0.013
PH	-0.019	PH	-0.014
PL	0.005	PL	0.001
PT	0.049	PT	0.012
GPP	-0.027	GPP	-0.007
TW	0.020	TW	-0.005
KL	0.035	SPY	-0.009
KB	0.009	KL	0.082
L/B ratio	-0.009	KB	-0.062
AC	0.005	AC	-0.001
Kernel length vs Protein content	r = 0.138	**Amylose content vs Protein content**	r = -0.083
Direct effect	0.158	Direct effect	0.030
Indirect effects through DFF	0.041	Indirect effects through DFF	-0.014
PH	-0.051	PH	0.016
PL	0.002	PL	-0.007
PT	0.023	PT	-0.035
GPP	0.004	GPP	-0.013
TW	0.011	TW	0.005
SPY	-0.041	SPY	-0.031
KB	0.008	KL	-0.028
L/B ratio	-0.009	KB	-0.005
AC	-0.005	L/B ratio	0.006
Kernel breadth vs Protein content	r =0.101		
Direct effect	0.078		
Indirect effects through DFF	0.012		
PH	-0.018		
PL	-0.008		
PT	0.005		
GPP	0.008		
TW	0.014		
SPY	-0.021		
KL	0.016		
L/B ratio	-0.014		
AC	-0.002		

Acknowledgements

The present study has been carried at the Institute of Rice Research (IRR), Rajendranagar, Hyderabad and molecular work was carried at IIRR lab. We are thankful to the institutions for providing the required support during the research period.

Authors' contributions

JAK carried out molecular studies, grain quality parameters and provided materials required for research work. LK participated in the design of the study. GP performed statistical analysis. KA, SNR, VKG and AAL edited and revised the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors declare that they don't have any conflict of interest.

Ethical issues: None

References

1. Kennedy G, Burlingame B. Analysis of food composition data on rice from a plant genetic resources perspective. Food Chem. 2003;80 (4):589—96. https://doi.org/10.1016/S0308-8146(02)00507-1
2. Das P, Adak S, Lahiri Majumder A. Genetic manipulation for improved nutritional quality in rice. Front Genet. 2020;11:776. https://doi.org/10.3389/fgene.2020.00776
3. Juliano BO. Structure, chemistry and function of the rice grain and its fractions. Cereal Foods World. 1992;37:772–79.
4. Tan YF, Sun M, Xing YZ, Hua JP, Sun XL, Zhang QF, et al. Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. Theor Appl Genet. 2001;103 (6):1037–45. https://doi.org/10.1007/s001220100665
5. Wang L, Zhong M, Li X, Yuan D, Xu Y, Liu H, et al. 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. Mol Breed. 2008;21(1):127–37. https://doi.org/10.1007/s11032-007-9141-7
6. Zhang W, Bi J, Chen L, Zheng L, Ji S, Xia Y, et al. QTL mapping for crude protein and protein fraction contents in rice (*Oryza sativa* L.). J Cereal Sci. 2008;48(2):539–47. https://doi.org/10.1016/j.jcs.2007.11.010
7. Ye G, Liang S, Wan J. QTL mapping of protein content in rice using single chromosome segment substitution lines. Theor Appl Genet. 2010;121(4):741–50. https://doi.org/10.1007/s00122-010-1345-2
8. Peng B, Kong HL, Li YB, Wang LQ, Zhong M, Sun L, et al. Genetic mapping and genetic basis of grain protein content in rice (*Oryza sativa* L.). Theor Appl Genet. 2014;127(6):1211–23.
9. Manivannan N. TNAUSTAT – statistical package. 2014.
10. Piper CS. Soil and plant analysis. New York: Inter Science Publishers; 1966. p. 47–9.
11. Sadasivam S, Manickam A. Peroxidase. In: Methods in biochemistry. New Delhi: New Age International; 1996. p. 108–10.
12. Pradhan SK, Pandit E, Pawar S, Bharati B, Chatopadhyay K, Singh S, et al. Association mapping reveals multiple QTLs for grain protein content in rice useful for biofortification. Mol Genet Genomics. 2019;294(4):963–83. https://doi.org/10.1007/s00438-019-01556-w
13. Zhong M, Wang LQ, Yuan DJ, Luo LJ, Xu CG, He YQ. Identification of QTL affecting protein and amino acid contents in rice. Rice Sci. 2011;18(3):187–90. https://doi.org/10.1016/S1672-6308(11)60026-7

14. Aluko G, Martinez C, Tohme J, Castano C, Bergman C, Oard JH. QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* × *O. glaberrima*. Theor Appl Genet. 2004;109(3):630–9. https://doi.org/10.1007/s00122-004-1668-y

15. Yun BW, Kim MG, Handoyo T, Kim KM. Analysis of rice grain quality-associated quantitative trait loci by using genetic mapping. Am J Plant Sci. 2014;5(9):1125–32. https://doi.org/10.4236/ajps.2014.59125

16. Lee S, Park J, Lee J, Shin D, Marmagne A, Lim PO, et al. OsASN1 overexpression in rice increases grain protein content and yield under nitrogen-limiting conditions. Plant Cell Physiol. 2020;61 (7):1309–20. https://doi.org/10.1093/pcp/pcaa060

17. 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. https://doi.org/10.1155/2014/781832

18. Yu YH, Li G, Fan YY, Zhang KQ, Min J, Zhu ZW, 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):121–5. https://doi.org/10.1016/j.jcs.2009.03.008

19. Immanuel SC, Pothiraj N, Thiyagarajan K, Bharathi M, Rabindran R. Genetic parameters of variability, correlation and path coefficient studies for grain yield and other yield attributes among rice blast disease resistant genotypes of rice (*Oryza sativa* L.). Afr J Biotechnol. 2011;10(17):3322–34. https://doi.org/10.5897/AJB10.2575

20. Dhakal R, Deosarkar DB, Paudel DP. Studies of correlation and path analysis of grain yield with yield contributing characters, grain quality characters and iron chlorosis in aerobic (*Oryza sativa* L.) on vertisols. Int J Curr Res. 2017;9(6):52926–31.

21. Devi PS, Veni BK, Jyothula DP, Raja DS. Association studies for yield components, physico-chemical and nutritional quality parameters in coloured rice (*Oryza sativa* L.). Int J Curr Microbiol Appl Sci. 2019;8 (6):359–65. https://doi.org/10.20546/ijcmas.2019.806.040

22. Ratna M, Begum S, Husna A, Dey SR, Hossain MS. Correlation and path coefficients analyses in basmati rice. Bangladesh J Agric Res. 2015;40(1):153–61.

23. Kalaiselvan S, Subramanian A, Thirumurugan T, Rajanbabu V. Genetic variability and association studies in F2 population of rice under sodicity. Electron J Plant Breed. 2019;10(2):601–13. https://doi.org/10.5958/0975-928X.2019.00076.0

24. Singh SK, Habde S, Singh DK, Khaire A, Mounika K, Majhi PK. Studies on character association and path analysis studies for yield, grain quality and nutritional traits in F2 population of rice (*Oryza sativa* L.). Electron J Plant Breed. 2020;11(3):969–75. https://doi.org/10.37992/2020.1103.158

25. Lohiteswararao RV, Kumar P, Sravani P, Raju KK, Gogada R. Correlation and path analysis of grain yield and its attributes of traditional rice (*Oryza sativa* L.) genotypes. Pharma Innov J. 2021;10 (12):1055–8.

26. Manojkumar D, Srinivas T, Rao LS, Suneetha Y, Sundaram RM, Kumari VP. Study of genetic variability and trait associations in F2 population of YH3 × AKDRMS 21-54 intra-specific cross of rice. Pharma Innov J. 2022;11(9):1735–42.

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Copyright: © The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited (https://creativecommons.org/licenses/by/4.0/)

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.