



Review Article



Toward high-protein rice: breeding, genomics, and metabolic strategies for improving grain protein accumulation

Sheetanshu Gupta

Assistant Professor, Department of Food Safety and Quality Assurance, College of Food Technology, Central Agriculture University, Imphal, Manipur, India

*Corresponding author e-mail: anshus279@gmail.com

(Received: 30/01/2026; Revised: 25/04/2026; Accepted: 12/06/2026; Published: 20/06/2026)

ABSTRACT

Rice is the primary staple food for more than half of the world's population, yet its milled grain typically contains only modest levels of protein compared with other cereals and grain legumes. Improving grain protein content and quality in rice is therefore an important goal for nutritional security, especially in regions where rice supplies a large fraction of daily calories and access to diverse protein sources is limited. Over the past two decades, significant progress has been made in understanding the genetic, genomic, and physiological basis of protein synthesis, deposition, and storage in the rice endosperm. This has enabled breeders and biotechnologists to identify key quantitative trait loci (QTLs), regulatory genes, transporters, and metabolic pathways that control both total grain protein content and the composition of storage proteins such as glutelins, prolamins, albumins, and globulins. Conventional and marker-assisted breeding have been used to introgress high-protein alleles into elite backgrounds, while genomic selection offers new opportunities to accelerate genetic gain for protein traits without sacrificing yield. At the same time, functional genomics and transcriptomics have revealed critical roles for amino acid transporters, transcription factors, and endomembrane trafficking components in governing protein loading and accumulation in seeds. More recently, transgenic and genome-editing strategies, including targeted modification of amino acid transporters and regulators of nitrogen assimilation, have demonstrated the potential to elevate grain protein content and improve amino-acid balance with minimal penalties to agronomic performance. However, trade-offs among protein content, grain yield, eating quality, and processing characteristics remain a major challenge.

Keywords: High-protein rice; grain protein content; rice biofortification; quantitative trait loci (QTL); amino acid transporters; functional genomics; genome editing

INTRODUCTION

Rice (*Oryza sativa* L.) is one of the most important cereal crops in the world and serves as the primary staple food for billions of people, especially across Asia, where it contributes a major proportion of daily caloric intake (Mahender et al., 2023; Shrestha et al., 2022). Because of this extraordinary dietary dependence, even modest improvements in the nutritional composition of rice can have major public health implications. While rice is highly valued for its digestibility, adaptability, and cooking versatility, polished rice is not considered a protein-dense food, and this limitation becomes nutritionally significant in populations where access to pulses, dairy, meat, or other high-quality protein sources is restricted (Shrestha et al., 2022; Sreenivasulu et al., 2024).

The challenge is not simply that rice contains less protein than legumes; it is that rice-dependent diets may meet energy needs without adequately supplying total protein or essential amino acids. This creates a hidden nutritional gap, particularly in low-income populations whose diets

are dominated by cereals. Therefore, improving grain protein content and protein quality in rice is increasingly viewed as a strategic goal in crop biofortification and nutritional security research (Mahender et al., 2023; Sreenivasulu et al., 2024). In recent years, this goal has become more realistic because advances in plant breeding, genomics, and molecular biotechnology have begun to clarify the biological mechanisms that regulate protein accumulation in rice grain (Li et al., 2023; Yang et al., 2022).

Rice grain composition is dominated by starch, but protein is the second most important storage component and plays a major role in nutritional value as well as grain quality. The total protein fraction in rice is mainly composed of glutelins, prolamins, globulins, and albumins, with glutelins accounting for the largest share of storage protein in the endosperm (Shrestha et al., 2022). These proteins differ in digestibility, amino acid composition, and their effects on cooking and eating quality. Rice protein is often considered superior to the

proteins of several other cereals in terms of digestibility and hypoallergenic properties, and rice protein isolates are increasingly explored for food applications, including gluten-free products and nutraceutical formulations (Shrestha et al., 2022). However, despite these advantages, rice proteins remain limited in certain essential amino acids, particularly lysine, which reduces their biological value when rice is consumed as the dominant staple without adequate dietary complementation (Shrestha et al., 2022; Sreenivasulu et al., 2024). For this reason, efforts to improve rice protein must address both quantity and quality. A rice variety with moderately increased protein but a more balanced amino acid profile may be more nutritionally useful than one with a larger increase in total protein alone. This is especially relevant in the context of modern biofortification, where the objective is not only to raise nutrient concentration but also to improve bioavailability, functionality, and compatibility with consumer preferences (Mahender et al., 2023). The concept of high-protein rice, therefore, extends beyond simple grain composition and enters the broader framework of designing rice varieties that are nutritionally enhanced, agronomically stable, and acceptable in terms of taste, texture, and cooking behaviour (Li et al., 2023).

The biology of protein accumulation in rice grain is complex and tightly regulated. Storage-protein deposition in the developing endosperm depends on a network of physiological and molecular processes that include nitrogen uptake from soil, assimilation into amino acids, long-distance transport through the plant, remobilization from vegetative tissues, and deposition into storage organelles in the seed (Li et al., 2023). Each of these steps can influence final grain protein content. This means that protein accumulation is not governed by a single gene or pathway but by an integrated system linking plant nitrogen metabolism, source-sink dynamics, endosperm development, and intracellular protein trafficking (Li et al., 2023; Yang et al., 2022). Such complexity explains why grain protein content is considered a quantitative trait and why its improvement has historically been challenging.

Even so, natural variation in rice germplasm offers an important opportunity for improvement. Studies summarised in recent reviews show that grain protein content varies substantially across rice accessions, cultivars, and breeding populations, indicating that the trait is amenable to selection and genetic dissection (Li et al., 2023). In some germplasm sets, protein content has ranged from approximately 5% to nearly 15%, demonstrating that there is enough phenotypic diversity to support both conventional and molecular breeding approaches (Li et al., 2023). This variation reflects the combined influence of genotype, environment, agronomic management, and genotype-by-environment interaction. Among these, nitrogen fertilization, growth duration, grain filling conditions, and polishing intensity are particularly important in shaping the final protein

content measured in harvested grain (Yang et al., 2022; Li et al., 2023).

One of the principal obstacles in breeding for high-protein rice is the frequent trade-off between grain protein content and other desirable traits, especially yield and eating quality. In many cereal crops, higher protein concentration is associated with lower starch deposition or altered grain structure, and rice follows a similar trend. Increased protein can modify the physicochemical properties of the grain, influence texture after cooking, and reduce the softness and palatability that consumers often prefer (Li et al., 2023; Custodio et al., 2019). Since rice is marketed and consumed according to very specific quality preferences—such as aroma, stickiness, elongation, whiteness, and softness—protein improvement cannot be pursued in isolation. A successful high-protein rice variety must balance nutritional gain with grain quality and agronomic performance, which makes the problem both biologically and commercially complex (Custodio et al., 2019; Li et al., 2023). Despite these challenges, recent progress in rice genetics has substantially improved prospects for targeted breeding. Genome-wide association studies, quantitative trait locus mapping, and candidate-gene analyses have identified several loci associated with grain protein content in rice (Yang et al., 2022; Wang et al., 2025). Some of these loci are linked to nitrogen metabolism, plant architecture, heading date, and grain development, which illustrates how deeply grain protein is connected with broader physiological processes. Yang et al. (2022), for example, demonstrated that the genetic basis of rice grain protein content involves not only direct regulators of seed composition but also genes affecting agronomic performance and developmental timing. This reinforces the idea that grain protein is a systems trait rather than an isolated quality attribute.

The emergence of multi-omics approaches has accelerated this transition from descriptive breeding to mechanism-based improvement. Transcriptomics, metabolomics, proteomics, and genomic prediction tools are increasingly being used to identify pathways and biomarkers associated with high-protein grain types. A recent multi-omics study identified rice lines with both high protein content and low glycemic index and revealed genomic regions and candidate genes contributing to these traits in a high-yielding background (Sreenivasulu et al., 2024). This finding is particularly important because it suggests that protein enrichment in rice can be integrated with other nutritional targets rather than treated as a separate breeding objective. It also signals a broader shift in rice biotechnology toward simultaneous optimization of multiple health-related grain traits.

Recent reports have further strengthened interest in high-protein rice by showing that significant gains may be possible in elite breeding materials. According to work highlighted by the Max Planck Institute, researchers developed rice lines with unusually high protein content

along with favourable starch characteristics, indicating that enhanced nutritional performance can be combined with other desirable dietary properties (Max Planck Society, 2024). Although such developments require extensive field validation, quality testing, and regional adaptation before broad deployment, they show that high-protein rice is no longer merely a conceptual breeding goal. It is becoming a practical target supported by molecular evidence, functional trait analysis, and translational breeding pipelines (Max Planck Society, 2024; Sreenivasulu et al., 2024).

Within this context, biotechnology offers several routes for improving grain protein accumulation. Conventional breeding remains essential, especially when high-protein donor lines are available and when non-transgenic outcomes are preferred for regulatory or market reasons. Marker-assisted selection can improve efficiency by enabling breeders to track favourable alleles for grain protein and related physiological traits in segregating populations (Li et al., 2023). Genomic selection adds another layer of power by predicting breeding values for complex traits controlled by many loci, thereby accelerating gain in traits that are expensive or unstable to phenotype directly (Mahender et al., 2023). These breeding approaches can be combined with pre-breeding strategies to broaden the genetic base for protein improvement and reduce linkage drag from undesirable donor backgrounds (Mahender et al., 2023).

At the same time, functional genomics and metabolic engineering provide opportunities to directly manipulate the mechanisms of protein accumulation. Genes controlling amino acid transport, nitrogen assimilation, storage-protein synthesis, and endosperm development are increasingly being investigated as intervention points for nutritional enhancement (Li et al., 2023; Yang et al., 2022). Targeted manipulation of such genes can, in principle, increase the transport of nitrogenous compounds to the grain, modify storage-protein balance, or improve the abundance of essential amino acids. Genome editing is particularly attractive in this regard because it enables precise modification of endogenous loci with fewer unintended genomic effects than conventional transgenesis (Sreenivasulu et al., 2024). As tools such as CRISPR-based editing mature in rice, the possibility of producing protein-enriched varieties with improved compositional and agronomic stability becomes increasingly realistic.

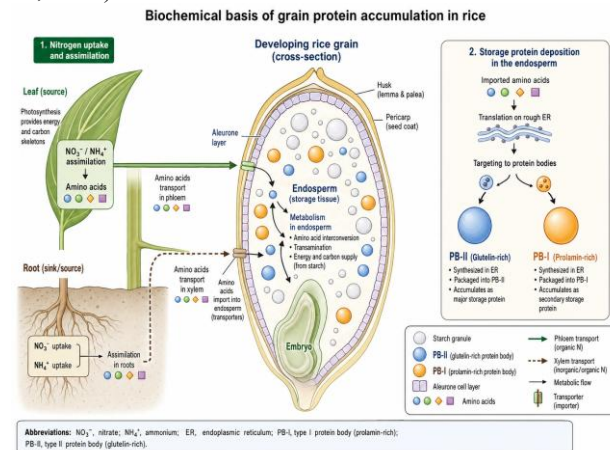
Another important dimension of this topic is the relationship between rice protein and broader grain quality attributes. Protein does not act independently within the grain matrix; it interacts with starch composition, amylose content, water absorption, and gelatinization behavior, all of which contribute to final cooking quality (Li et al., 2023; Custodio et al., 2019). Therefore, a review of high-protein rice must consider not only breeding targets and gene networks but also the practical implications for consumer acceptance and food use. This is especially important in countries where rice is selected and traded according to fine distinctions in

sensory quality, and where any deterioration in taste or texture may prevent even nutritionally superior varieties from being adopted widely. The topic is also timely because rice improvement is increasingly expected to address multiple nutritional and metabolic concerns at once. The same breeding pipelines that once focused on zinc, iron, and provitamin A biofortification are now being expanded toward protein quality, glycemic response, and functional food characteristics (Mahender et al., 2023; Sreenivasulu et al., 2024). This integrated perspective reflects a shift in crop improvement from yield-centered thinking toward nutrition-sensitive breeding. In that context, high-protein rice represents more than a single trait objective; it represents an effort to redesign one of the world's most important staple crops so that it contributes more effectively to human health.

Biochemical basis of grain protein accumulation in rice
The development of high-protein rice depends on a clear understanding of how proteins are synthesized, transported, deposited, and remodelled during grain filling. In rice, protein is the second most abundant storage component after starch, but its importance extends beyond simple percentage composition. Grain proteins influence nutritional value, milling behaviour, cooking quality, texture, and even the physiological response of the grain to environmental stress. As a result, the biochemical basis of protein accumulation is central to both rice nutrition research and biotechnology-driven crop improvement (Shrestha et al., 2022; Li et al., 2023). Rice grain proteins are generally classified into four major fractions: glutelins, prolamins, globulins, and albumins. Among these, glutelins are the dominant storage proteins and account for the largest share of total endosperm protein, while prolamins form the second major storage fraction. Albumins and globulins occur in smaller quantities but contribute to the overall nutritional and functional profile of the grain (Shrestha et al., 2022; Huang et al., 2019). The relative abundance of these fractions matters because each differs in digestibility, amino acid composition, and cellular localization. Glutelins are generally considered more nutritionally favourable than prolamins, which are relatively less digestible and contribute differently to grain texture and end-use quality (Shrestha et al., 2022; Huang et al., 2019).

A key feature of rice endosperm biology is that these storage proteins are not deposited randomly. Instead, they accumulate in distinct intracellular structures known as protein bodies. Prolamins are synthesized on the rough endoplasmic reticulum and remain within the lumen, where they form ER-derived protein bodies known as PB-I. By contrast, glutelin precursors are synthesized in the endoplasmic reticulum but are subsequently transported through the Golgi apparatus to protein storage vacuoles, where they are processed and deposited in vacuole-derived protein bodies known as PB-II. This separation reflects differences in trafficking pathways as well as biochemical behavior, and it has

major implications for attempts to engineer storage-protein composition in rice (Tosi et al., 2014; Shrestha et al., 2022).



The predominance of glutelin in rice grain is explained not only by its synthesis but also by its accumulation kinetics. Detailed work on grain development has shown that glutelin accumulates at a faster rate and for a longer duration than prolamin during grain filling. Huang et al. (2019) reported that final glutelin weight per grain was more than three times higher than final prolamin weight and that the higher glutelin content resulted from both a greater accumulation rate and a longer active deposition period. This finding is significant because it suggests that increasing total grain protein is not merely a matter of inducing more synthesis overall; it may also depend on extending the active accumulation period of nutritionally favorable fractions or improving their trafficking efficiency during seed development.

Protein accumulation in rice grain is tightly linked with nitrogen metabolism. Grain storage proteins are ultimately built from nitrogen assimilated by the plant, and the efficiency with which rice absorbs, remobilizes, and transports nitrogen to the developing grain has a direct influence on endosperm protein deposition. Nitrogen absorbed from soil is incorporated into amino acids, especially glutamine and asparagine, which are then transported to sink tissues. During grain filling, the plant reallocates nitrogen from leaves and stems toward the developing caryopsis, where amino acids are used for protein synthesis and storage-protein assembly (Li et al., 2023; Zhang et al., 2026). This means that the protein content of rice grain reflects whole-plant physiological coordination rather than a seed-specific process alone.

The relationship between nitrogen availability and grain protein accumulation has been documented extensively. Nitrogen fertilization often increases grain protein content, but the response is influenced by timing, cultivar, and environmental conditions. Beyond the simple increase in total protein, nitrogen supply can also alter the relative abundance of specific storage-protein fractions and affect grain quality traits such as chalkiness, hardness, and eating quality (Li et al., 2023; Tang et al., 2021). This dual role makes nitrogen a powerful but complex lever for enhancing protein

content. In breeding and biotechnology, it also highlights the importance of identifying genotypes that maintain efficient nitrogen use and favourable grain composition under practical field conditions. At the molecular level, protein accumulation in rice grain is governed by a coordinated network of genes involved in transcriptional regulation, amino acid transport, endomembrane trafficking, and post-translational processing. Storage-protein genes must be expressed in the correct temporal and spatial pattern during endosperm development, and their products must be folded, transported, and packaged efficiently into the proper storage organelles. Proteomic and cytological studies of developing rice endosperm have shown that protein synthesis is closely integrated with amino acid metabolism, redox regulation, cellular trafficking, and the metabolic changes that accompany starch deposition (Han et al., 2010). These findings reinforce the view that grain protein accumulation is part of a broader developmental program rather than an isolated biochemical event.

Transport processes are especially important in determining how much protein ultimately accumulates in the grain. Since amino acids must move from source organs into the developing endosperm, amino acid transporters represent critical control points in the pathway from plant nitrogen assimilation to seed protein deposition. Although the precise contribution of individual transporter families varies, current reviews emphasize that transport regulation can influence both total protein content and the balance of storage-protein fractions by determining the substrate supply available during grain filling (Li et al., 2023; Zhang et al., 2026). This is one reason why transporter genes have emerged as promising targets in high-protein rice research.

Another important component of the biochemical framework is protein-body formation and intracellular trafficking. The rice endosperm uses two distinct deposition systems, one for prolamins and another for glutelins, and perturbations in trafficking can alter both grain composition and seed ultrastructure. Experiments with transgenic rice expressing heterologous storage proteins have shown that changes in trafficking behaviour can produce abnormal protein bodies and disrupt the normal segregation of endosperm storage proteins (Tosi et al., 2014). Such findings are highly relevant to biotechnology because they indicate that attempts to enhance protein content by modifying storage-protein expression must also consider whether the cellular machinery can correctly process, sort, and store the altered protein load. The biochemical balance among storage proteins also affects amino acid composition. Reducing one storage-protein fraction can trigger compensatory accumulation of others, thereby changing the overall nutritional profile of the grain. Kawakatsu et al. (2010) demonstrated that suppression of certain seed storage proteins in transgenic rice led to compensatory synthesis of alternative proteins and altered amino acid distribution, including changes in total lysine content. This compensation phenomenon is

important because it suggests that the rice seed protein system is flexible and responsive, which may be advantageous for engineering improved protein quality. At the same time, it means that targeted interventions may produce indirect effects that must be evaluated carefully.

Environmental stress adds another layer of complexity to protein accumulation. Temperature during grain filling is particularly important, as elevated temperature can alter storage-substance deposition, grain quality formation, and the abundance of proteins involved in signaling, transport, and oxidative stress responses. Proteomic analyses under high-temperature conditions have identified stimulus-responsive proteins linked to auxin signaling, calcium signaling, protein modification, and reactive oxygen species regulation, suggesting that stress can reshape the network controlling storage-protein deposition (Wang et al., 2023). For high-protein rice biotechnology, this matters because a genotype with improved protein accumulation under normal conditions may behave differently under heat stress, which is increasingly relevant in the context of climate change.

The interaction between protein accumulation and eating quality also deserves attention. Protein in rice is not simply a nutritional component; it influences starch gelatinization, water absorption, grain hardness, and textural properties after cooking. Higher protein content is often associated with reduced palatability or changes in softness, and the specific contributions of different protein components may not be identical. Recent studies examining associations between rice protein fractions and eating quality indicate that not only total protein, but also the balance among components, can shape consumer-perceived grain quality (Li et al., 2023; Association between rice protein components and eating quality, 2023). This means that improving rice protein accumulation must be approached with biochemical precision rather than by pursuing total protein gain alone. Taken together, the biochemical basis of grain protein accumulation in rice reveals a tightly interconnected system involving storage-protein composition, endosperm organelle biology, nitrogen metabolism, amino acid transport, developmental regulation, and environmental response. Glutelin predominance, differential protein-body trafficking, nitrogen remobilization, and compensatory changes among storage fractions all contribute to the final protein content and quality of the rice grain. Understanding these processes is essential for designing breeding and biotechnology strategies that increase protein content without compromising yield, grain quality, or adaptation. Therefore, any serious effort to develop high-protein rice must begin with the biochemical logic of how the grain builds and stores its proteins during development.

Genetic architecture and QTLs for grain protein content in rice

Grain protein content in rice is a classic quantitative trait. It is controlled by multiple loci, influenced by

environment, and often interacts with other grain-quality and agronomic traits such as amylose content, gel consistency, chalkiness, heading date, and yield. Because of this complexity, progress in high-protein rice breeding has depended heavily on genetic mapping, genome-wide association studies, and the identification of functionally relevant candidate genes rather than on phenotype-based selection alone (Chen et al., 2022; Bharali et al., 2025). These approaches have shown that grain protein content is not regulated by one major pathway alone, but by a dispersed genetic architecture spread across the rice genome.

Early QTL studies revealed that loci associated with grain protein content occur on almost all 12 rice chromosomes, although some genomic regions appear repeatedly across populations and environments. More recent work has improved this picture by focusing specifically on milled rice rather than brown rice, which is important because consumers predominantly eat polished rice and because nutritional improvement in milled grain is the more practical breeding target (Bharali et al., 2025). In an F_2 population derived from the cross BPT 5204 $\times$ JAK-686, Bharali et al. (2025) identified four QTLs for grain protein content in milled rice: *qPC1.1*, *qPC1.2*, *qPC5.1*, and *qPC5.2*. Among these, *qPC1.2* on chromosome 1 was the major QTL, explaining 15.71% of phenotypic variation, while the favourable additive alleles for all four QTLs were contributed by the high-protein donor JAK-686. This result is important because it demonstrates that useful QTLs for grain protein enhancement can be identified in breeding-relevant material without completely compromising agronomic traits.

Genome-wide association analysis has expanded this understanding even further. Chen et al. (2022) evaluated 529 rice accessions from 87 countries and detected 135 significant loci associated with grain protein content across years and subpopulations. Several loci were repeatedly detected under different environments, suggesting that at least part of the genetic basis of grain protein content is stable enough to support marker-assisted improvement. Four minor but validated QTLs—*qPC2.1*, *qPC7.1*, *qPC7.2*, and *qPC1.1*—were confirmed using near-isogenic line F_2 populations, with some explaining surprisingly large proportions of phenotypic variation in specific validation backgrounds. In particular, *qPC7.1* and *qPC7.2* on chromosome 7 explained 43.4% and 29.2% of the variation, respectively, in their NIL-derived populations, demonstrating that certain genomic regions can exert strong effects depending on genetic background and allele combination (Chen et al., 2022).

An important lesson from these studies is that grain protein content in rice is shaped by both additive and epistatic effects. Bharali et al. (2025) not only detected main-effect QTLs but also reported a large number of digenic epistatic interactions influencing protein content and related quality traits. This means that breeders cannot assume that favorable loci will act independently

when combined in a new genetic background. Some alleles may reinforce each other, while others may interfere, reducing the expected phenotypic gain. For high-protein rice development, this makes validation across populations and environments essential before a QTL can be considered truly useful in breeding.

Another major advance has been the cloning and functional validation of genes underlying grain protein QTLs. Among the best characterized is *qPC1*, which corresponds to *OsAAP6*, a putative amino acid transporter. This gene positively regulates grain protein content by enhancing amino acid uptake and redistribution, thereby increasing the supply of nitrogenous substrates to the developing grain (Chen et al., 2022; Bharali et al., 2025). *OsAAP6* is particularly significant because it provides a mechanistic bridge between whole-plant nitrogen transport and endosperm protein accumulation. Rather than simply marking a statistical association, it points directly to a biological process that can be manipulated through breeding or biotechnology.

Another important candidate is *OsGluA2*, which has been validated as the gene underlying *qGPC-10*. This gene encodes a glutelin type-A2 precursor and acts as a positive regulator of grain protein content, especially by influencing glutelin accumulation. Since glutelin is the major and nutritionally favorable storage-protein fraction in rice, *OsGluA2* is of particular interest for improving not only total protein but also storage-protein composition (Chen et al., 2022; Bharali et al., 2025). Together, *OsAAP6* and *OsGluA2* illustrate two complementary genetic routes for improving grain protein: one through amino acid transport and the other through storage-protein synthesis. Recent work also suggests that some genes influencing starch metabolism and grain appearance may indirectly affect grain protein content. In the GWAS study of Chen et al. (2022), the gene *Flo5* emerged as a candidate affecting protein content as well as chalkiness. CRISPR/Cas9 knockout mutants of *Flo5* showed increased grain protein content together with increased chalkiness, indicating that starch biosynthesis genes may have pleiotropic effects on endosperm structure and compositional balance. This is highly relevant for rice biotechnology because it highlights a recurring theme in grain-quality genetics: protein traits are deeply connected with starch traits, and improving one may alter the other. A high-protein line with excessive chalkiness or poor eating quality may have limited practical value, regardless of its nutritional gain.

The relationship between QTLs for protein content and other quality traits is therefore a central concern. In the BPT 5204 × JAK-686 study, QTLs for grain protein content were mapped alongside loci for amylose content, gel consistency, kernel traits, and yield components, showing that protein improvement must be evaluated in a multi-trait context (Bharali et al., 2025). Correlation

analyses in that study suggested weak or negative associations between protein content and some quality parameters, while previous literature also indicates that higher protein often contributes to harder-cooked rice and reduced palatability. This does not mean that high-protein rice is inherently unacceptable, but it does mean that breeding strategies must avoid selecting for protein in isolation. Marker-assisted pyramiding will likely be most effective when protein QTLs are combined with favorable alleles for grain texture, appearance, and consumer acceptance.

The availability of stable QTLs and validated genes is now making marker-assisted selection more realistic for grain protein improvement. QTLs such as *qPC1.2*, *qPC7.1*, and *qPC7.2*, along with cloned genes like *OsAAP6* and *OsGluA2*, provide useful entry points for developing breeder-friendly markers. Once validated across target environments, such markers can help introgress high-protein alleles into elite backgrounds while minimizing linkage drag from undesirable donor traits (Bharali et al., 2025; Chen et al., 2022). This is especially useful in rice breeding, where farmers and consumers often demand the retention of specific varietal features such as maturity duration, grain shape, aroma, and cooking quality.

At the same time, the current evidence also shows that grain protein content has moderate heritability and remains environmentally sensitive. Bharali et al. (2025) reported moderate heritability and moderate genetic advance for grain protein content in their segregating population, indicating that environmental influence remains substantial even when genetic variability is present. This reinforces the need for multilocation testing, replicated phenotyping, and, increasingly, genomic prediction models that can capture both marker effects and genotype-by-environment interaction. For a trait like grain protein content, robust breeding progress will likely come from combining classical QTL mapping, GWAS, candidate-gene validation, and genomic selection rather than relying on any one approach alone. Overall, the genetic architecture of grain protein content in rice can now be described with much greater clarity than before. Multiple QTLs have been mapped, several stable genomic regions have been validated, and key functional genes such as *OsAAP6*, *OsGluA2*, and *Flo5* have begun to reveal the biological networks underlying protein accumulation. Yet the trait remains genetically complex, environmentally responsive, and tightly linked with grain quality traits that influence adoption. For this reason, the value of QTL research lies not only in identifying loci but in building a workable framework for breeding high-protein rice that is nutritionally improved, agronomically competitive, and acceptable to consumers.

Table 1. Major QTLs and candidate genes associated with grain protein content in rice

S. No.	QTL/gene	Chromosome	Reported effect	Key significance	Source
1	<i>qPC1.2</i>	1	Explained 15.71% phenotypic variation in milled rice	Major QTL detected in BPT 5204 × JAK-686; favorable allele from high-protein parent	Bharali et al. (2025)
2	<i>qPC1.1</i>	1	Detected in both biparental and GWAS-linked validation studies	Suggests repeated importance of chromosome 1 region for protein accumulation	Bharali et al. (2025); Chen et al. (2022)
3	<i>qPC5.1</i>	5	Minor QTL for grain protein content	Useful as part of pyramiding strategy	Bharali et al. (2025)
4	<i>qPC5.2</i>	5	Minor QTL for grain protein content	Indicates multiple useful loci on chromosome 5	Bharali et al. (2025)
5	<i>qPC2.1</i>	2	Validated in NIL-F ₂ ; explained 9.82% variation	Stable GWAS-supported locus	Chen et al. (2022)
6	<i>qPC7.1</i>	7	Validated in NIL-F ₂ ; explained 43.4% variation	Strong-effect locus in validation population	Chen et al. (2022)
7	<i>qPC7.2</i>	7	Validated in NIL-F ₂ ; explained 29.2% variation	Important chromosome 7 locus for grain protein regulation	Chen et al. (2022)
8	<i>OsAAP6</i> (<i>qPC1</i>)	1	Positive regulator of grain protein content via amino acid transport	Key cloned gene for marker-assisted and biotechnological improvement	Chen et al. (2022); Bharali et al. (2025)
9	<i>OsGluA2</i> (<i>qGPC-10</i>)	10	Positive regulator of protein, especially glutelin accumulation	Important for improving storage-protein composition	Chen et al. (2022); Bharali et al. (2025)
10	<i>Flo5</i>	8	Affects grain protein content and chalkiness	Illustrates pleiotropic linkage between starch and protein traits	Chen et al. (2022)

Breeding strategies for developing high-protein rice
Improving grain protein content in rice through breeding is both promising and challenging. The promise lies in the presence of useful natural variation, mapped QTLs, validated markers, and increasing genomic resources. The challenge lies in the fact that protein content is a polygenic and environmentally responsive trait that is often linked with yield, grain appearance, and cooking quality. For this reason, breeding high-protein rice requires an integrated strategy that combines conventional selection with molecular and genomic tools rather than relying on a single approach (Bharali et al., 2025; Chen et al., 2022).

Conventional breeding and donor selection

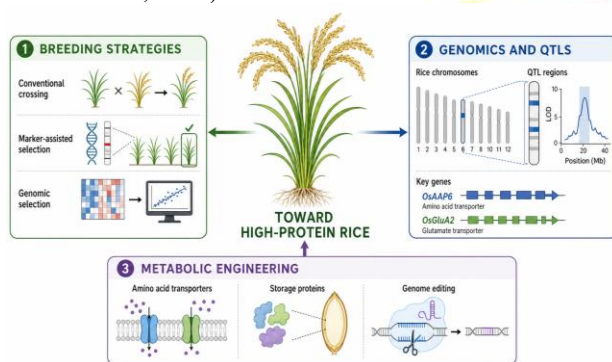
The foundation of any breeding program for high-protein rice is the identification of suitable donor germplasm. Natural variation for grain protein content exists across traditional varieties, landraces, breeding lines, and diverse germplasm collections, and this variation remains the starting point for improvement. Conventional hybridization followed by pedigree or population-based selection can be effective when donor lines possess clearly superior protein content and at least

acceptable agronomic traits. This approach remains especially relevant in breeding systems where transgenic or edited products face regulatory delay or low public acceptance (Li et al., 2023; Yang et al., 2022). However, donor choice is critical. A high-protein donor that carries poor grain type, low yield, late maturity, or undesirable cooking characteristics may introduce severe linkage drag into elite backgrounds. That is why breeders increasingly favour donors that combine elevated protein content with at least partial adaptation or acceptable grain quality. The recent mapping work using JAK-686 as a donor in a cross with BPT 5204 is a good example of this logic, because it identified useful protein QTLs in material directly relevant to rice improvement rather than in highly exotic germplasm alone (Bharali et al., 2025). In practical breeding, this reduces the gap between discovery and deployment. Conventional breeding remains useful not only for trait introgression but also for validating whether improved protein levels can be maintained across environments. Since protein content is sensitive to nitrogen availability, temperature, and grain filling conditions, field-based selection under target environments remains

indispensable. Even with strong marker support, breeders still need phenotypic confirmation to ensure that high-protein lines perform consistently and retain consumer-acceptable grain quality (Yang et al., 2022; Li et al., 2023).

Marker-assisted selection and QTL pyramiding

Marker-assisted selection has become one of the most practical tools for improving complex grain-quality traits in rice. Once a QTL or gene has been validated, linked DNA markers can be used to track favourable alleles in segregating populations without waiting for full phenotypic expression. This is particularly useful for grain protein content because phenotyping can be laborious, destructive, and influenced by the environment. Marker-assisted selection, therefore, increases efficiency, allows earlier selection, and helps breeders retain elite recurrent-parent backgrounds while introducing nutritionally favourable loci (IRRI, n.d.; Bharali et al., 2025).



The success of marker-assisted breeding depends on the reliability of the underlying loci. QTLs such as *qPC1.2*, *qPC7.1*, and *qPC7.2*, along with candidate genes such as *OsAAP6* and *OsGluA2*, offer promising targets because they have shown repeatability or functional relevance for grain protein improvement (Chen et al., 2022; Bharali et al., 2025). In breeding practice, these loci can be introduced through marker-assisted backcrossing or combined through QTL pyramiding. Pyramiding is especially attractive because grain protein is controlled by many loci of small to moderate effect, and combining favourable alleles may produce larger and more stable gains than introgressing a single QTL. Even so, marker-assisted selection for protein must be done carefully. Since some protein-associated loci are linked with grain chalkiness, cooking quality, or other end-use traits, breeders need flanking markers and strong background selection to avoid undesirable carryover. The broader experience of rice breeding shows that marker-assisted pyramiding works best when breeders simultaneously track target loci and key background regions, allowing rapid recovery of the elite parent while preserving the desired grain-quality profile (Singh et al., 2024; IRRI, n.d.). For high-protein rice, this means MAS should be used not as a stand-alone shortcut, but as part of a wider breeding design that includes rigorous phenotype confirmation.

Genomic selection and multi-trait prediction

While marker-assisted selection works well for a few major loci, it becomes less efficient when a trait is governed by many small-effect genes. Grain protein content in rice often fits this description, which makes genomic selection particularly relevant. Genomic selection uses genome-wide marker information to predict the breeding value of individuals without needing to identify all causal loci explicitly. This is useful for complex, quantitative traits because it captures additive effects distributed across the genome and can reduce breeding cycle time (Wang et al., 2021; Hasan et al., 2023).

For high-protein rice, genomic selection offers at least three clear advantages. First, it can improve selection efficiency for a trait that is moderately heritable and environmentally sensitive. Second, it can be used in early generations before labor-intensive grain-quality testing is feasible. Third, it can be combined with multi-trait models so that protein content is improved alongside yield, maturity, grain type, and other quality parameters rather than at their expense (Wang et al., 2021; Hasan et al., 2023). This is important because one of the greatest risks in protein breeding is selecting nutritionally improved lines that later fail in farmer fields or consumer markets. Recent rice breeding experience in South Asia has shown that genomic selection can accelerate gain when used together with trait-specific marker-assisted selection. Hasan et al. (2023) reported that integrating genomic selection with marker-based breeding improved efficiency in irrigated rice improvement programs by enabling better capture of breeding value and shortening selection cycles. Although that work was not focused only on protein, the principle applies directly to high-protein rice. A combined strategy in which major validated protein loci are fixed using markers and the remaining polygenic background is improved through genomic prediction may be one of the most effective breeding routes available today.

Haplotype-based breeding and allele mining

As rice genomic resources expand, breeding is moving beyond simple QTL presence-absence logic toward a more refined understanding of favourable haplotypes and allelic combinations. Haplotype-based breeding identifies superior versions of genes or linked genomic blocks and then combines them in elite backgrounds with greater precision than conventional marker breeding alone. This strategy is especially valuable for grain quality traits, where the phenotypic effect of a locus often depends on the exact allelic state and the wider genetic background (Rathour et al., 2021; Singh et al., 2024).

For high-protein rice, haplotype analysis may help distinguish truly favourable alleles of genes such as *OsAAP6* from neutral or less effective variants. It can also help breeders identify elite donors in landraces, regional germplasm, or underutilised collections without relying solely on extreme phenotypes. Allele mining further strengthens this process by searching for superior

natural variants in diverse germplasm pools. Together, these tools can enrich the breeding pipeline with novel sources of protein improvement that may not be visible through conventional phenotyping alone (Rathour et al., 2021; Li et al., 2023). Another advantage of haplotype-based breeding is that it supports better management of trade-offs. If one allele improves protein but harms eating quality, whereas a related haplotype offers moderate protein gain with less quality penalty, breeders can make more informed choices. In rice, where market acceptance depends heavily on grain quality, this fine-scale discrimination is extremely valuable.

Integrating protein improvement with quality and yield No breeding strategy for high-protein rice will succeed unless it accounts for the balance among protein content, yield, and consumer quality. Higher protein is desirable, but not at the cost of drastic yield loss, unacceptable grain appearance, or poor cooking behaviour. This is

why modern rice breeding increasingly emphasizes multi-trait selection. Grain protein must be evaluated together with amylose content, gel consistency, chalkiness, grain length and breadth, head rice recovery, and field performance under target ecologies (Bharali et al., 2025; Custodio et al., 2019). This integration also requires careful stage-wise selection. Early generations may focus on markers, broad agronomic suitability, and elimination of undesirable segregants. Intermediate stages should combine biochemical screening for protein content with grain-quality testing. Advanced testing must then confirm performance across environments, nitrogen levels, and farmer-relevant management conditions. Such a structured pipeline is more time-consuming than selecting for a single trait, but it is essential if high-protein rice is to move from research novelty to farmer adoption.

Table 2. Major breeding approaches for high-protein rice improvement

S. No.	Breeding approach	Main purpose	Strength	Limitation
1	Conventional hybridization and selection	Introgress protein-rich traits from donor parents	Simple, widely accepted, no special regulatory issues	Slow; strongly influenced by environment
2	Marker-assisted selection (MAS)	Track validated QTLs/genes for grain protein	Efficient early-generation selection; precise introgression	Limited when many small-effect loci are involved
3	Marker-assisted backcrossing	Transfer protein QTLs into elite cultivars	Helps recover recurrent parent background quickly	Needs tightly linked markers and strong phenotyping
4	QTL pyramiding	Combine multiple favorable protein loci	Can improve cumulative gain and stability	Epistasis may reduce expected benefits
5	Genomic selection	Predict breeding value using genome-wide markers	Useful for polygenic traits; shortens breeding cycle	Requires training population and strong statistical models
6	Haplotype-based breeding	Select superior allelic combinations	Greater precision than single-marker selection	Depends on good genomic data and allele characterization
7	Allele mining	Identify novel high-protein variants in germplasm	Expands breeding diversity and donor options	Useful alleles may carry linkage drag
8	Multi-trait selection	Improve protein while retaining yield and quality	Most realistic route for cultivar development	More complex and resource-intensive

Genomics, functional genes, and metabolic engineering for high-protein rice

The transition from conventional breeding to precision improvement has made genomics and metabolic engineering central to the development of high-protein rice. Mapping studies have already shown that grain protein content is controlled by multiple loci, but genomics helps move beyond statistical association to biological mechanism. It allows researchers to identify the genes that regulate amino acid uptake, nitrogen redistribution, storage-protein synthesis, endosperm development, and the balance between protein and starch deposition. This is particularly important because rice grain protein content is not a simple trait of seed composition; it is the final outcome of coordinated transport, metabolism, signaling, and developmental regulation across the whole plant (Chen et al., 2022; Sreenivasulu et al., 2024).

Functional genes regulating grain protein accumulation Among the known genes affecting grain protein content in rice, *OsAAP6* is one of the best characterized. It encodes an amino acid permease and has been shown to positively regulate grain protein accumulation by increasing amino acid uptake and translocation toward developing seeds. Functional analyses have shown that elevated expression of *OsAAP6* enhances amino acid import and promotes the expression of genes associated with storage-protein synthesis, ultimately increasing rice grain protein content (Chen et al., 2022; Ren et al., 2023). Because this gene connects source–sink amino acid movement with seed storage-protein formation, it has become one of the most promising targets for both molecular breeding and gene-based improvement. Another emerging regulator is *OsNAC74*, a transcription factor that influences grain protein content by regulating the expression of *OsAAP6* and other protein-related

genes. Gene-editing studies have shown that knockout of *OsNAC74* reduces *OsAAP6* expression and significantly lowers grain protein content, while also affecting traits such as chalkiness and gel consistency (Ren et al., 2023). This finding is important because it demonstrates that protein accumulation in rice is not controlled only by transporter genes or storage-protein genes directly, but also by upstream regulatory factors that coordinate broader developmental and metabolic pathways. Such transcriptional regulators may eventually provide more flexible intervention points than structural genes alone.

Storage-protein genes themselves also remain important targets. The identification of *OsGluA2* as a positive regulator of grain protein content highlights the importance of glutelin biosynthesis in improving both total protein and nutritionally favorable storage-protein fractions (Chen et al., 2022). Since glutelins are the predominant and relatively digestible storage proteins in rice, manipulating their abundance or deposition pattern offers a direct route to nutritional enhancement. However, the effect of altering storage-protein expression is often buffered by compensatory mechanisms within the seed, which means that direct changes in one protein class may trigger increases or decreases in others (Kawakatsu et al., 2010). This internal balance must be considered when engineering grain composition.

Multi-omics and pathway discovery

One of the most powerful recent advances in rice biotechnology is the use of multi-omics to connect phenotype with mechanism. Genomics identifies candidate loci, transcriptomics reveals gene-expression patterns, metabolomics captures pathway activity, and proteomics helps define the actual biochemical output of developing grains. When these layers are integrated, researchers can identify not only genes associated with high-protein phenotypes, but also the metabolic context in which those genes operate (Sreenivasulu et al., 2024). This systems-level view is crucial because high-protein rice is shaped by interactions among nitrogen assimilation, amino acid biosynthesis, endosperm filling, and starch metabolism. A recent multi-omics study of a rice population identified lines combining high amylose, low glycemic index, and high protein content and revealed that these superior lines were enriched in glycolytic and amino acid metabolic pathways (Sreenivasulu et al., 2024). This is a particularly valuable insight because it suggests that protein enrichment is not simply a matter of more nitrogen deposition but may also involve a broader rewiring of carbon–nitrogen metabolism. It also indicates that nutritionally improved rice can be developed through the coordinated selection of pathways affecting both grain composition and metabolic health properties, rather than through single-trait selection alone. Proteomics and systems reviews further support this integrated view. Recent analyses of grain proteins in rice and barley have emphasized that future improvement will require coordinated

manipulation of biosynthesis, subcellular transport, and amino acid-pool management. These studies argue that increasing total protein content may be most effective when combined with strategies that suppress lysine-poor proteins, redirect nitrogen toward nutritionally superior proteins, and expand the free amino acid pool available during grain filling (Akhter et al., 2025). Such insights are helping shift rice biotechnology from descriptive protein analysis toward predictive metabolic engineering.

Metabolic engineering and amino acid enhancement

Metabolic engineering aims to improve grain composition by redesigning the pathways that determine substrate flow, biosynthesis, storage, and end-product balance. In the context of high-protein rice, this means increasing the supply of amino acids to the grain, improving the accumulation of desirable storage proteins, and, where possible, enhancing essential amino acid composition. Since lysine is a limiting amino acid in rice, one promising strategy is to engineer metabolic sinks that favor lysine-rich proteins or increase lysine biosynthesis while reducing its degradation (Akhter et al., 2025). Although much of this work remains at the experimental stage, it provides a conceptual framework for improving protein quality at the same time as protein quantity.

Another metabolic strategy involves changing the relative abundance of storage-protein fractions. Rice endosperm has a built-in capacity to rebalance protein accumulation, and studies reducing specific storage proteins have shown that the seed can compensate by increasing alternative proteins and altering amino acid composition (Kawakatsu et al., 2010). This compensation can be used strategically. If lysine-poor protein classes are reduced and the plant is simultaneously engineered to increase amino acid supply, the result may be a more favourable overall protein profile. However, such interventions must be assessed carefully because protein-body organization, grain texture, and seed viability can all be affected. Metabolic engineering is also being linked increasingly with the regulation of starch–protein relationships. Recent reports on high-protein, low-glycemic rice suggest that modifying starch biosynthesis genes, such as *sbeIIb*, can alter overall grain metabolism in ways that coincide with higher protein content and improved health-related characteristics (Sreenivasulu et al., 2024; Max Planck Society, 2024). This does not mean starch genes should be treated as direct protein genes, but it does show that manipulating one major component of the endosperm can reshape the broader metabolic environment in which protein accumulates. Such cross-pathway effects may become highly useful in designing multi-benefit rice varieties.

Genome editing as a precision tool

Genome editing has rapidly become one of the most promising tools for rice trait improvement because it allows direct modification of endogenous genes with far greater precision than traditional transgenic approaches.

In rice, CRISPR/Cas systems are already being used widely to improve yield, stress tolerance, and grain quality, and their application to nutritional biofortification is growing quickly (Sukegawa & Toki, 2022; Li et al., 2024). For high-protein rice, genome editing is particularly attractive because many useful targets are already known, including transporters, transcription factors, and quality-related genes connected with protein deposition.

The appeal of genome editing lies in its flexibility. It can be used to knock out negative regulators, fine-tune promoter activity, alter coding sequence function, or create favorable alleles directly in elite cultivars. In the case of *OsNAC74*, gene-editing work has helped clarify its regulatory relationship with *OsAAP6* and grain protein accumulation (Ren et al., 2023). In the case of endosperm composition, edited alleles of *sbellb* have confirmed the feasibility of generating rice lines with altered grain composition and nutritionally improved profiles (Sreenivasulu et al., 2024). These examples show that genome editing serves not only as a breeding tool but also as a functional validation platform for gene discovery. At the same time, genome editing should not be viewed as a simple shortcut. The rice grain is a tightly integrated biological system, and editing one gene can have unintended effects on chalkiness, cooking quality, yield, or seed development. For this reason, edited lines must be evaluated for whole-plant performance and end-use quality, not only for target-trait success. Nonetheless, compared with conventional transgenics, genome editing offers a more precise and potentially more acceptable route for improving high-protein rice, especially in cases where no foreign DNA remains in the final product (Sukegawa & Toki, 2022; Li et al., 2024).

Trade-offs among protein content, yield, and eating quality

One of the most persistent challenges in developing high-protein rice is that protein enrichment is rarely an isolated improvement. In most cases, increasing grain protein content affects other important traits, especially yield, cooking quality, and consumer acceptance. This creates a classic breeding dilemma: the nutritional goal of raising protein must be balanced against the agronomic and market realities that determine whether a new variety will actually be adopted. For this reason, the development of high-protein rice cannot be viewed simply as a process of maximizing protein percentage. It is better understood as an effort to optimize the balance among nutrition, productivity, and grain quality within a biologically constrained system (Li et al., 2023; Yang et al., 2022).

The first major trade-off involves the relationship between protein and yield. In cereals, grain filling represents a competition among major storage components, principally starch and protein, under the limits imposed by plant resources and sink capacity. When rice plants produce larger yields, much of the increase is driven by enhanced starch deposition, and this can dilute grain protein concentration even if the total

amount of nitrogen taken up by the plant does not decline. Conversely, management or genotypes that increase grain protein often do so by increasing nitrogen accumulation or reducing the dominance of starch deposition, and this may alter yield formation or biomass partitioning. As a result, high protein is sometimes associated with lower yield or with conditions that require greater nitrogen input to maintain acceptable productivity (Li et al., 2023; Wang et al., 2025).

This relationship, however, is not absolute. Some studies suggest that the negative association between yield and protein can be moderated by genotype and sink structure. Research on high-yielding rice with large sink capacity has shown that it is possible to achieve yields above 700 g m⁻² while maintaining comparatively moderate protein levels and acceptable grain quality, indicating that the yield–protein relationship can be managed through improved partitioning rather than being treated as fixed (Wada et al., 2018). This is an encouraging result for high-protein rice breeding because it implies that trade-offs may be softened through a better understanding of panicle architecture, nitrogen use, and grain filling dynamics. Even so, in practical breeding, large and stable gains in protein without any yield penalty remain difficult to achieve.

The second major trade-off concerns cooking and eating quality, which is often more limiting than yield from the standpoint of farmer adoption and consumer preference. Rice protein has a strong influence on cooked grain texture because proteins restrict water penetration into starch granules, reduce starch swelling, and interfere with gelatinization during cooking. As protein content increases, cooked rice often becomes harder, less sticky, and less glossy, all of which may reduce eating quality in markets where softness and palatability are highly valued (Li et al., 2023; Xiong et al., 2022). This is why many studies report a negative relationship between total grain protein content and taste quality. The effect is not only due to total protein percentage but also to the composition of protein fractions. Recent work indicates that glutelin and prolamin can both contribute to deterioration in cooking and eating quality, although their effects may differ in magnitude and mechanism. Increased protein accumulation can alter starch–protein interactions within the endosperm, reducing the mobility of water during cooking and changing the viscoelastic behaviour of the cooked grain. This helps explain why varieties with similar total protein can still differ noticeably in eating quality depending on the balance among storage-protein components and their interaction with starch properties (Association between rice protein components and eating quality, 2023; Li et al., 2023).

Rapid Visco Analyzer studies and taste-value analyses reinforce this conclusion. Xiong et al. (2022) showed that protein content explained a substantial share of variation in rice taste value, especially in high- and low-taste-value groups, and that higher protein was generally associated with poorer palatability. Their work suggested that protein content can serve as an important

indicator of eating quality because it affects pasting behaviour and, through that, the texture perceived by consumers. This is especially relevant for breeding because it means that even nutritionally desirable increases in protein must be interpreted through their consequences for cooking performance, not simply through compositional analysis.

A third layer of trade-off arises through agronomic management, especially nitrogen fertilization. Nitrogen application often increases grain protein content, but it can simultaneously reduce eating quality by making grains harder and less palatable. In some situations, nitrogen also increases chalkiness risk or changes grain appearance, especially when combined with heat stress during grain filling. Thus, management strategies that appear useful for improving protein concentration may not align with the quality characteristics demanded by consumers or processors (Li et al., 2023; Tang et al., 2021). This highlights an important principle for high-protein rice improvement: raising protein by agronomic manipulation alone is not always the same as breeding a nutritionally superior variety. The interaction between protein and grain appearance also deserves attention. Protein-related changes in endosperm structure can influence translucency, chalkiness, and milling behaviour. In some cases, loci that improve protein content are linked with undesirable effects on endosperm organization, which may reduce grain appearance and therefore market value. The example of *Flo5*, which affects both protein content and chalkiness, illustrates how intimately protein traits are connected with broader grain-quality genetics (Chen et al., 2022). This means that breeding high-protein rice requires more than identifying favourable protein alleles; it requires evaluating their pleiotropic effects on the physical and commercial quality of the grain. Despite these challenges, the goal of high-protein rice should not be abandoned. Instead, it should be reframed. The real objective is not to produce rice with the highest possible protein content regardless of consequence, but to develop varieties with improved protein levels and acceptable agronomic and sensory performance. In many contexts, a moderate increase in protein combined with stable yield and good eating quality may be far more valuable than a dramatic increase in protein accompanied by severe quality penalties. This is especially true in countries where consumer rejection of poor-tasting rice can quickly undermine the adoption of otherwise promising varieties.

Future progress will therefore depend on multi-trait optimization. Breeding programs need to combine protein content with amylose content, gel consistency, chalkiness, texture, and yield components in a coordinated selection framework. Functional genes and QTLs should be evaluated not only for their direct effect on protein accumulation but also for their influence on cooking behaviour and grain structure. Likewise, biotechnology strategies such as genome editing and metabolic engineering must be assessed in a whole-grain

context rather than judged only by target-trait improvement. Only by acknowledging and managing these trade-offs can rice biotechnology produce high-protein cultivars that are nutritionally meaningful, agronomically competitive, and acceptable to consumers.

CONCLUSION

Rice remains the principal staple food for a large proportion of the global population, enhancing grain protein content an important strategy for improving nutritional security. This review highlights that grain protein accumulation in rice is a complex quantitative trait governed by interconnected biochemical, physiological, and genetic mechanisms. Protein deposition is influenced by nitrogen uptake and remobilization, amino acid transport, storage-protein synthesis, endosperm development, and environmental conditions, particularly during grain filling. Among storage proteins, *glutelins* constitute the major fraction and play a crucial role in determining both nutritional quality and grain functionality.

Significant progress has been achieved in identifying key quantitative trait loci (QTLs) and functional genes associated with grain protein content, including *OsAAP6*, *OsGluA2*, and *OsNAC74*. Advances in genome-wide association studies, multi-omics approaches, genomic selection, and genome-editing technologies have greatly enhanced our understanding of the molecular networks regulating protein accumulation. These tools provide unprecedented opportunities for the development of high-protein rice cultivars with improved nutritional value.

However, increasing grain protein content remains challenging due to its frequent trade-offs with grain yield, cooking quality, texture, and consumer preference. Therefore, future breeding efforts should focus on multi-trait optimization rather than maximizing protein concentration alone. Integrating conventional breeding, marker-assisted selection, genomic prediction, allele mining, and precision genome editing will be essential for achieving sustainable genetic gains.

Overall, the development of high-protein rice represents a promising avenue for addressing protein malnutrition and enhancing food quality. Continued research integrating genomics, biotechnology, and breeding will facilitate the creation of nutritionally enriched rice varieties that combine superior protein content with high yield, desirable grain quality, and broad consumer acceptance.

REFERENCES

- Akhter, M., Sultana, N., Rahman, M. and colleagues, 2025. A systems review of grain proteins in rice and barley. *Frontiers in Plant Science*, **16**: 1658144. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12678259/>

- Association between rice protein components and eating quality under different nitrogen fertilizer levels, 2023. *Food Chemistry*. Available at: <https://www.sciencedirect.com/science/article/pii/S0733521023001339>
- Bharali, V., Yadla, S., Thati, S., Bitra, B., Karapati, D., Chirravuri, N.N., Badri, J., Sundaram, R.M. and Jukanti, A.K., 2025. Mapping genomic regions for grain protein content and quality traits in milled rice (*Oryza sativa* L.). *Plants*, **14**(6): 905. <https://doi.org/10.3390/plants14060905>
- Chen, P., Lou, G., Wang, Y., Chen, J., Chen, W., Fan, Z., Liu, Q., Sun, B., Mao, X., Yu, H., Jiang, L., Zhang, J., Lv, S., Xing, J., Pan, D., Li, C. and He, Y., 2022. The genetic basis of grain protein content in rice by genome-wide association analysis. *Molecular Breeding*, **43**(1): 1. <https://doi.org/10.1007/s11032-022-01347-z>
- Custodio, M.C., Cuevas, R.P.O., Ynion, J., Laborte, A.G., Velasco, M.L. and Demont, M., 2019. Rice grain quality: Current developments and future prospects. In: Waisundara, S.W. and Shiomi, N. (Eds.), *Rice Grain Quality*. IntechOpen, pp. 1–55.
- Han, C., Yang, P., Sakata, K. and Komatsu, S., 2010. Integrated proteomic and cytological study of rice endosperms at the storage phase. *Journal of Proteome Research*, **9**(10): 4906–4916. <https://doi.org/10.1021/pr100603z>
- Hasan, M.M., Raffi, M.Y., Ismail, M.R. and colleagues, 2023. Enhancing genetic gain through the application of genomic selection in developing irrigated rice for the favorable ecosystem in Bangladesh. *Frontiers in Genetics*, **14**: 1083221. <https://doi.org/10.3389/fgene.2023.1083221>
- Huang, M., Chen, J., Cao, F., Zou, Y. and Jiang, P., 2019. Quantifying accumulation characteristics of glutelin and prolamin in rice grains. *Rice*, **12**: 1–9. <https://doi.org/10.1186/s12284-019-0315-8>
- International Rice Research Institute (IRRI), n.d. Marker assisted breeding. *Rice Knowledge Bank*. Available at: http://www.knowledgebank.irri.org/ricebreeding/course/Marker_assisted_breeding.htm
- Kawakatsu, T., Hirose, S., Yasuda, H. and Takaiwa, F., 2010. Reducing rice seed storage protein accumulation leads to changes in nutrient quality and storage organelle formation. *Plant Physiology*, **154**(4): 1842–1854. <https://doi.org/10.1104/pp.110.163162>
- Li, R., Li, R., Luo, A., Wu, H. and Zhang, H., 2024. Breeding rice for yield improvement through CRISPR/Cas9 genome editing. *Rice*, **17**: 13. <https://doi.org/10.1186/s12284-024-00689-1>
- Li, X., Wang, M., Zhang, L. and Chen, Q., 2023. Research progress on the relationship between rice protein content and grain quality. *Seed Biology*, **2**: 1–17. <https://doi.org/10.48130/SeedBio-2023-0016>
- Mahender, A., Anandan, A., Pradhan, S.K. and Pandit, E., 2023. Rice biofortification: Breeding and genomic approaches for genetic enhancement of grain zinc and iron contents. *Frontiers in Plant Science*, **14**: 1138408. <https://doi.org/10.3389/fpls.2023.1138408>
- Max Planck Society, 2024. Rice with a high protein content developed. Available at: <https://www.mpg.de/23544458/high-protein-low-glycemic-index-rice>
- Rathour, R., Singh, A.K., Pathak, H. and colleagues, 2021. The need and current perspective for rice improvement in India. *Oryza*, **58**(2): 109–124.
- Ren, Y., Zhang, J., Wang, H. and colleagues, 2023. OsNAC74 affects grain protein content and various quality traits by regulating OsAAP6 in rice. *Rice*, **16**: 55. <https://doi.org/10.1186/s12284-023-00678-3>
- Shrestha, B., Man, K., Bhandari, B. and Prakash, S., 2022. A narrative review on rice proteins: Current scenario and food industrial application. *Foods*, **11**(15): 2283. <https://doi.org/10.3390/foods11152283>
- Singh, V.K., Ellur, R.K., Puspha, R. and colleagues, 2024. Marker-assisted breeding accelerates the development of multiple stress tolerant rice genotypes. *Frontiers in Plant Science*, **15**: 1402368. <https://doi.org/10.3389/fpls.2024.1402368>
- Sreenivasulu, N., Butardo, V.M., Bhatnagar-Mathur, P. and colleagues, 2024. Multiomics of a rice population identifies genes and genomic regions that bestow low glycemic index and high protein content. *Proceedings of the National Academy of Sciences*, **121**(35): e2410598121. <https://doi.org/10.1073/pnas.2410598121>
- Sukegawa, S. and Toki, S., 2022. Genome editing technology and its application to metabolic engineering in rice. *Plant Cell Reports*, **41**(6): 1263–1272. <https://doi.org/10.1007/s00299-022-02857-7>
- Tang, X., Li, Y., Sun, Y., Zhang, Y., Zhou, J. and Gu, J., 2021. Nitrogen fertilizer regulated grain storage protein accumulation and reduced chalkiness under high temperature in rice. *Plant Physiology and Biochemistry*, **166**: 921–930. <https://doi.org/10.1016/j.plaphy.2021.06.032>
- Tosi, P., Gritsch, C.S., He, J. and Shewry, P.R., 2014. The trafficking pathway of a wheat storage protein in transgenic rice endosperm. *Plant Journal*, **77**(2): 220–230. <https://doi.org/10.1111/tbj.12380>
- Wada, T., Nonami, H., Matsue, Y. and colleagues, 2018. Characteristics of growth and quality, and factors contributing to high yield in rice. *Plant Production Science*, **21**(4): 278–287. <https://doi.org/10.1080/1343943X.2018.1463165>

- Wang, X., Pang, Y., Wang, C., Chen, K., Zhu, Y., Shen, C., Ali, J., Xu, J., Li, Z. and Li, J., 2021. Genomic selection: A breakthrough technology in rice breeding. *The Crop Journal*, **9**(3): 669–677. <https://doi.org/10.1016/j.cj.2021.03.008>
- Wang, X., Zhang, Y., Liu, J. and colleagues, 2025. Mapping genomic regions for grain protein content and related traits in rice. *Plants*, **14**: 813.
- Wang, Y., Zhang, H., Liu, S. and colleagues, 2023. Stimulus-responsive proteins involved in multi-process regulation of storage substance accumulation during rice grain filling under elevated temperature. *Plants*, **12**(21): 3758. <https://doi.org/10.3390/plants12213758>
- Xiong, Y., Wang, S., Hao, W. and colleagues, 2022. Use of protein content, amylose content, and RVA parameters to evaluate the taste quality of rice. *Frontiers in Nutrition*, **8**: 758547. <https://doi.org/10.3389/fnut.2021.758547>
- Yang, W., Wang, W., Xin, Y. and colleagues, 2022. The genetic basis of grain protein content in rice by genome-wide association study. *Foods*, **12**(1): 23.
- Zhang, Y., Liu, J., Wang, X. and colleagues, 2026. Physiological and molecular mechanisms of nitrogen regulation on cereal grain quality. *International Journal of Molecular Sciences*, **27**: 1543.

Citation: Sheetanshu Gupta 2026. Toward high-protein rice: breeding, genomics, and metabolic strategies for improving grain protein accumulation. *International Journal of Agricultural and Applied Sciences*, 7(1): 77-90. <https://doi.org/10.52804/ijaas2026.719>

Copyright: © Gupta 2026. Creative Commons Attribution 4.0 International License. IJAAS allows unrestricted use, reproduction, and distribution of this article in any medium by providing adequate credit to the author(s) and the source of publication.

