Genomics and Applied Biology 2024, Vol.15, No.4, 200-211
http://bioscipublisher.com/index.php/gab

Feature Review Open Access

Comprehensive Genomic Identification and Characterization of *R2R3-MYB* Genes in Colored Rice (*Oryza sativa* L.): A Phylogenetic and Expression Analysis

Mengmeng Yin [1,2,4], Fan Luo [1,2,4], Xiaoli Zhou [1,2,4], Changhe Wei [5], Qian Zhu [1,2,3], Hui Zhang [2,3], Dong Sun Lee [1,2,3] ✉, Lijuan Chen [1,2,3] ✉

1 State Key Laboratory for Conservation and Utilization of Bio-Resources in Yunnan, Yunnan Agricultural University, Kunming, 650201, Yunnan, China
2 The Key Laboratory for Crop Production and Smart Agriculture of Yunnan Province, Yunnan Agricultural University, Kunming, 650201, Yunnan, China
3 Rice Research Institute, Yunnan Agricultural University, Kunming, 650201, Yunnan, China
4 College of Agricultural Science, Xichang University, Xichang, 615013, Sichuan, China
5 Panxi Crops Research and Utilization Key Laboratory of Sichuan Province, Xichang University, Xichang, 615013, Sichuan, China
✉ Co-corresponding authors: dongeast@naver.com; chenlijuan@hotmail.com

Genomics and Applied Biology, 2024, Vol.15, No.4 doi: 10.5376/gab.2024.15.0022
Received: 27 Jun., 2024
Accepted: 04 Aug., 2024
Published: 17Aug., 2024

Preferred citation for this article:
Yin M.M., Luo F., Zhou X.L., Wei C.H., Zhu Q., Zhang H., Lee D.S., and Chen L.J., 2024, Comprehensive genomic identification and characterization of *R2R3-MYB* genes in colored rice (*Oryza sativa* L.): a phylogenetic and expression analysis, Genomics and Applied Biology, 15(4): 200-211 (doi: 10.5376/gab.2024.15.0022)

Abstract This review provides a comprehensive identification and characterization of the *R2R3-MYB* gene family in colored rice (*Oryza sativa* L.), offering significant insights into their evolutionary relationships, structural features, and expression profiles. Notable findings include the distinctive structural characteristics of *R2R3-MYB* genes, such as the high prevalence of non-synonymous substitutions in the DNA-binding domains, particularly in the α-helix regions. This suggests adaptive selection and functional diversification. The phylogenetic analysis revealed the existence of distinct clades that correspond to different evolutionary lineages. Of particular interest is a key clade that is closely related to the ancestral species, wild rice (*O. rufipogon* and *O. nivara*), which indicates the conservation of evolutionary lineages. The expression patterns of *R2R3-MYB* genes were found to be specific to different tissues and developmentally regulated, with specialized roles in photosynthesis-related processes and root development. Furthermore, the study underscores the functional roles of *R2R3-MYB* genes in anthocyanin biosynthesis. Genes such as *OsC1* and *OsKala3* play pivotal roles in modulating the expression of anthocyanin biosynthetic pathway genes, thereby contributing to the distinctive purple pigmentation and associated health benefits of purple rice. Furthermore, the case study of *OsMYB30* and *OsMYB60* illustrates their pivotal roles in plant defense mechanisms and leaf morphology, respectively. The insights gained from this review have significant implications for the breeding and genetic engineering of colored rice, emphasizing the potential for improving agronomic traits and enhancing crop performance through targeted manipulation of *R2R3-MYB* genes.

Keywords Colored rice; *R2R3-MYB* genes; Phylogenetic analysis; Gene expression analysis; Anthocyanin biosynthesis

1 Introduction

R2R3-MYB transcription factors (TFs) are a prominent family of regulatory proteins in plants, characterized by the presence of two MYB domains at the N-terminus. These TFs play crucial roles in various plant processes, including cell cycle regulation, secondary metabolism, and responses to biotic and abiotic stresses. Specifically, *R2R3-MYB* genes are integral to the regulation of anthocyanin biosynthesis, a pathway responsible for the production of pigments that contribute to the coloration of plant tissues (Feng et al., 2018). The functional diversity of *R2R3-MYB* genes is attributed to their ability to interact with other proteins, such as basic helix-loop-helix (bHLH) proteins, to form transcriptional complexes that modulate gene expression (Sakamoto et al., 2001; Feng et al., 2018).

Anthocyanins are a class of flavonoid pigments that impart red, purple, and blue colors to various plant tissues, including flowers, fruits, and leaves. These pigments serve a dual role in the plant kingdom, acting as both attractants for pollinators and seed dispersers, and as a means of protection against environmental stresses, including UV radiation, pathogen attack, and oxidative stress (Zhu et al., 2017; Dwiningsih and Alkahtani, 2022). The biosynthesis of anthocyanins is a complex process involving a network of structural and regulatory genes. R2R3-MYB transcription factors play a pivotal role as key regulators of this pathway. In rice, anthocyanin

accumulation is influenced by both genetic and environmental factors (Mackon et al., 2021). The presence of specific alleles can lead to tissue-specific pigmentation patterns (Zhu et al., 2017).

Purple rice is distinguished by its high anthocyanin content, which gives the grains and other plant parts their characteristic purple color (Xia et al., 2021). The pigmentation is not only aesthetically appealing but also associated with numerous health benefits due to the antioxidant properties of anthocyanins (Lachagari et al., 2019; Dwiningsih and Alkahtani, 2022). Genomic studies have been conducted on purple rice varieties, including the landrace *Purpleputtu*, to elucidate the genetic basis of anthocyanin biosynthesis and its regulation (Lachagari et al., 2019). These studies have revealed significant allelic variations and unique genetic features that contribute to the purple coloration, thus making purple rice an excellent model for studying the molecular mechanisms underlying anthocyanin biosynthesis (Lachagari et al., 2019). Moreover, the creation of biofortified rice varieties with elevated anthocyanin levels, exemplified by the "Purple Endosperm Rice" underscores the potential of genetic engineering to enhance the nutritional profile of staple crops (Zhu et al., 2017).

The objective of this review is to conduct a comprehensive genomic identification and characterization of *R2R3-MYB* genes in colored rice. This will entail a comprehensive phylogenetic analysis to elucidate the evolutionary relationships among these genes and an expression analysis to ascertain their roles in anthocyanin biosynthesis and other biological processes. By integrating bioinformatics and experimental approaches, this reviwe aims to provide insights into the functional diversity and regulatory mechanisms of *R2R3-MYB* genes in colored rice, thereby contributing to a more comprehensive understanding of plant secondary metabolism and stress responses.

2 Genome-Wide Identification of *R2R3-MYB* Genes in Colored Rice

2.1 Data resources and genome databases

The identification of *R2R3-MYB* genes in rice is dependent upon the utilization of comprehensive genomic data from a multitude of sources. Whole genome sequencing and comparative genomic analysis have been instrumental in elucidating the genetic intricacies and allelic variations that are exclusive to purple rice landraces, such as *Purpleputtu* (Lachagari et al., 2019). Furthermore, databases such as the SNP-Seek database, which contains a comprehensive set of single nucleotide polymorphism (SNP) data from a diverse range of rice lines, serve as a valuable resource for identifying genetic variations (Lachagari et al., 2019). The utilization of these databases enables a comprehensive comparison of the colored rice genome with other rice cultivars, thereby facilitating the identification of specific genes involved in anthocyanin biosynthesis and other related pathways (Yan et al., 2020).

2.2 Bioinformatics tools and methodologies

The use of bioinformatics tools and methodologies is a crucial aspect of genome-wide identification of *R2R3-MYB* genes. Multiple sequence alignment tools are employed to compare gene sequences across diverse rice cultivars, thereby elucidating deletions, insertions, and other variations that may impact gene function (Lachagari et al., 2019). Phylogenetic analysis tools facilitate comprehension of the evolutionary relationships between identified genes and their homologs in other species, thereby providing insights into their functional roles (Lachagari et al., 2019). Furthermore, cDNA microarray approaches are utilized to monitor gene expression profiles under diverse stress conditions, facilitating the identification of *R2R3-MYB* genes that are differentially expressed (Dai et al., 2007; Dai et al., 2012). These methodologies collectively facilitate a comprehensive analysis of the *R2R3-MYB* gene family in colored rice.

2.3 Criteria for gene identification

The identification of *R2R3-MYB* genes in colored rice is based on several criteria. Firstly, genes must exhibit sequence homology to known *R2R3-MYB* genes in other plant species, particularly in their DNA-binding domains (Jia et al., 2004; Li et al., 2020). Secondly, the presence of characteristic motifs and domains, such as the R2 and R3 repeats, is essential for confirming the identity of these TFs (Jia et al., 2004; Hua et al., 2020). Moreover, genes must demonstrate differential expression patterns in response to specific stimuli or conditions, such as phosphate starvation or abiotic stress, which indicate their regulatory roles (Dai et al., 2007; Dai et al., 2012).

Ultimately, functional validation through genetic and molecular approaches, such as gene overexpression or knockout studies, is essential for confirming the involvement of these genes in anthocyanin biosynthesis and other related pathways (Dai et al., 2012; Kim et al., 2021; Duan et al., 2022). These criteria ensure the accurate identification and characterization of *R2R3-MYB* genes in colored rice.

3 Phylogenetic Analysis of *R2R3-MYB* Genes

3.1 Construction of phylogenetic trees

To elucidate the evolutionary history and relationships of *R2R3-MYB* genes in colored rice, comprehensive phylogenetic trees were constructed initially. The amino acid sequences of R2R3-MYB proteins were retrieved from the rice genome using bioinformatics tools. A multiple sequence alignment was conducted using ClustalW, with appropriate gap penalties and substitution matrices set to ensure high alignment accuracy. The aligned sequences were then employed in the construction of phylogenetic trees utilizing both the Neighbor-Joining (NJ) and Maximum Likelihood (ML) methods. The NJ trees were generated using MEGA X software with 1,000 bootstrap replicates to assess the robustness of the inferred relationships, while ML trees were constructed using RAxML, applying the best-fit model of amino acid substitution as determined by ModelFinder (Jia et al., 2004; Monna et al., 2006; Lachagari et al., 2019).

3.2 Evolutionary relationships among *R2R3-MYB* genes

The phylogenetic trees revealed the existence of distinct clades, which corresponded to different evolutionary lineages of *R2R3-MYB* genes. It is noteworthy that the analysis demonstrated a striking prevalence of non-synonymous substitutions in the DNA-binding domains, particularly in the helix 1 and helix 2 regions, indicating adaptive selection (Jia et al., 2004). This indicates that these regions are integral to protein-DNA interactions, which may facilitate the identification of novel DNA-binding sites. Furthermore, a co-evolutionary pattern was identified between the second α-helix of the R2 domain and the second α-helix of the R3 domain, underscoring the functional significance of these interactions (Jia et al., 2004).

3.3 Comparative analysis with other plant species

A comparative genomic analysis with other plant species, such as *Arabidopsis thaliana*, revealed that the R2R3-MYB gene family in rice shares several lineage-specific genes with these species, which may account for functional diversification. A phylogenetic analysis of the *Rc* locus, a key regulatory gene in the anthocyanin biosynthetic pathway, revealed the existence of a distinct clade that is closely related to the progenitor species *O. rufipogon* and *O. nivara*, indicating a conserved evolutionary lineage (Lachagari et al., 2019). This comparative approach highlights the evolutionary conservation and divergence of *R2R3-MYB* genes across diverse plant species, offering a more comprehensive understanding of their functional roles in rice and other plants.

4 Gene Structure and Chromosomal Location

4.1 Structural features of *R2R3-MYB* genes

The *R2R3-MYB* gene family represents one of the most significant transcription factor families in plants, distinguished by the presence of two MYB domains (R2 and R3) that are involved in DNA binding. A comparative analysis of *R2R3-MYB* genes in rice has revealed a high frequency of non-synonymous substitutions in the α-helix regions of the DNA-binding domains, indicating adaptive selection and functional diversification (Jia et al., 2004). These structural features are of great consequence for the specific recognition of DNA-binding sites, which in turn affects gene regulation and expression.

4.2 Chromosomal distribution and synteny analysis

The chromosomal distribution of *R2R3-MYB* genes in rice has been extensively mapped. For example, the sequence analysis of rice chromosome 4 has provided insights into the localization of various genes, including transcription factors such as R2R3-MYB. Furthermore, genome-wide analyses have identified the presence of these genes across different chromosomes, with a notable concentration in specific regions that are rich in transposable elements and other repetitive sequences (Temnykh et al., 2001). A comparison of the chromosomal distribution of *R2R3-MYB* genes in rice with that of other plant species, such as *A. thaliana*, has revealed a limited

degree of conservation in gene order. This finding suggests that significant chromosomal rearrangements and divergence have occurred (Monna et al., 2006; Shi et al., 2022).

4.3 Gene duplication and divergence

Gene duplication events have been a significant driving force in the expansion and functional diversification of the *R2R3-MYB* gene family in rice. The existence of multiple allelic and structural variations within these genes indicates a history of gene duplication followed by divergence. For example, the *Rc* locus, a key regulatory gene in the anthocyanin biosynthetic pathway, exhibits a 14 bp deletion in the *Purpleputtu* rice variety, which is a characteristic feature of modern white pericarp rice cultivars (Lachagari et al., 2019; Khan et al., 2021). These observations illustrate the evolutionary dynamics and adaptive significance of gene duplication in the *R2R3-MYB* gene family, as well as the importance of considering similar variations.

5 Expression Profiling of *R2R3-MYB* Genes

5.1 Expression patterns in different tissues

The expression patterns of *R2R3-MYB* genes in diverse tissues of colored rice have been the subject of extensive investigation. R2R3-MYB TFs have been demonstrated to play a pivotal role in a multitude of biological processes, including tissue development. In rice, the *Os2R_MYB* genes exhibit tissue-specific expression patterns, which are essential for understanding their functional roles in plant growth and development (Dai et al., 2012). For instance, Kang et al. (2022) identified 190 MYB TFs, including 99 R2R3-MYBs, and discovered that these genes are unevenly distributed across the 12 chromosomes of rice. The analysis demonstrated notable discrepancies in the expression levels of the *Os2R_MYB* gene across diverse tissues and developmental stages (Figure 1). It also revealed that specific *Os2R_MYB*s are uniquely expressed in particular tissues, suggesting their specialized functions within those tissues.

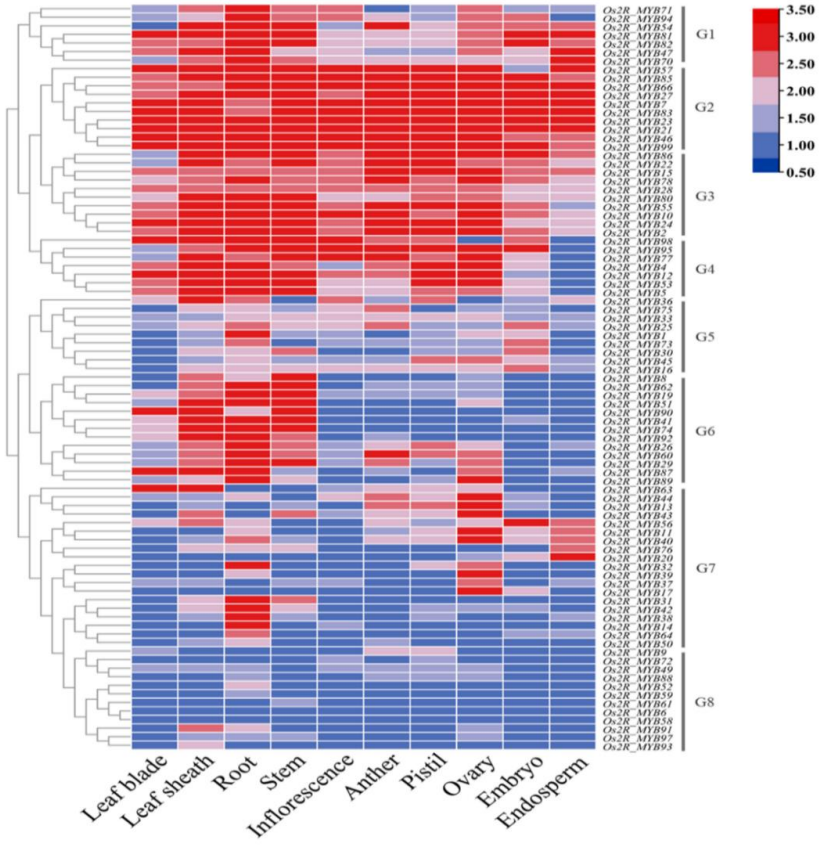

Figure 1 The relative expression level of *Os2R_MYB* genes in various tissues and growth stages (Adopted from Kang et al, 2022)

Image caption: Lower expression is represented with dark blue; higher expression is represented with red; Leaf-blade, vegetative leaf-blade; Leaf-sheath, vegetative leaf-sheath; Root, vegetative root; Stem, reproductive stem; Inflorescence, 0.6-1 mm; Anther, 0.7-1 mm; Pistil, 5-10 cm; Ovary, 3 DAF; Embryo, 10 DAF; Endosperm, 42 DAF (Adopted from Kang et al, 2022)

5.2 Stage-specific expression

The stage-specific expression of *R2R3-MYB* genes is a crucial element in their functional characterization. These genes are involved in the regulation of diverse processes throughout the plant life cycle, encompassing germination, flowering, and grain filling. The expression levels of *Os2R_MYB* genes exhibit significant variation across different developmental stages, indicating their involvement in regulating developmental processes. For example, Kang et al. (2022) analyzed the transcription levels of 20 *Os2R_MYB* genes under different stress conditions. Their findings revealed that some genes exhibited altered expression patterns at specific developmental stages, underscoring their pivotal role in developmental regulation.

5.3 Responses to environmental stimuli

R2R3-MYB genes have also been demonstrated to respond to a range of environmental stimuli, including abiotic and biotic stresses. The expression of these genes in response to stress conditions provides insights into their roles in stress tolerance mechanisms. In rice, the *Os2R_MYB* genes contain multiple stress-responsive elements within their promoter regions, including ABRE, TGACG, CGTCA, and MBS motifs. These elements are associated with responses to abiotic stresses, such as drought and heavy metal exposure. For example, the transcription levels of *Os2R_MYB* genes were found to be significantly altered under conditions of polyethylene glycol (PEG) and cadmium chloride ($CdCl_2$) stress, indicating their involvement in stress response pathways (Kang et al., 2022).

6 Functional Roles of *R2R3-MYB* Genes in Anthocyanin Biosynthesis

6.1 Known functions of *R2R3-MYB* genes in other species

R2R3-MYB TFs play a crucial role in regulating anthocyanin biosynthesis across a diverse range of plant species. In *Nitraria sibirica* Pall., the *R2R3-MYB* gene *NsMYB1* has been identified as a key regulator of anthocyanin biosynthesis, influencing fruit color differentiation by promoting the transcription of structural genes involved in the anthocyanin biosynthetic pathway (Bao et al., 2021). Similarly, in wheat (*Triticum aestivum*), the R2R3-MYB protein TaPL1 has been demonstrated to act as a positive regulator of anthocyanin biosynthesis, particularly in response to environmental stresses such as cold, salt, and light, which are known to induce anthocyanin accumulation (Shin et al., 2016). In *Freesia hybrida*, the R2R3-MYB regulator FhPAP1 has been demonstrated to activate anthocyanin biosynthetic genes and interact with other TFs to form the MYB-bHLH-WD40 (MBW) complex, which is essential for anthocyanin production (Li et al., 2020). These examples illustrate the conserved role of *R2R3-MYB* genes in anthocyanin biosynthesis across different plant species.

6.2 Potential roles in purple rice

R2R3-MYB genes probably play a significant role in the biosynthesis of anthocyanins in purple rice, as they do in other species. The *OsC1* gene, an R2R3-MYB transcription factor in rice, has been demonstrated to regulate anthocyanin biosynthesis by modulating the expression of late anthocyanin biosynthetic pathway (ABP) genes. The overexpression of *OsC1* in white rice plants has been demonstrated to induce anthocyanin production, which in turn results in enhanced photosynthetic efficiency and a reduction in oxidative stress (Upadhyaya et al., 2021). Furthermore, the *OsKala3* gene, another R2R3-MYB transcription factor in rice, interacts with the bHLH transcription factor OsKala4 to activate anthocyanin biosynthetic genes, thereby contributing to the pigmentation of the rice pericarp (Kim et al., 2021). These findings indicate that *R2R3-MYB* genes in purple rice may similarly regulate anthocyanin biosynthesis, thereby enhancing the plant's antioxidant properties and stress tolerance.

6.3 Experimental approaches for functional validation

To validate the functional roles of *R2R3-MYB* genes in anthocyanin biosynthesis in colored rice, several experimental approaches may be employed.

1) Gene overexpression and knockout studies: The overexpression of candidate *R2R3-MYB* genes, such as *OsC1* and *OsKala3*, in transgenic rice plants can facilitate the determination of their impact on anthocyanin production. Conversely, the knockout or knockdown of these genes can be achieved using CRISPR/Cas9 or RNA interference (RNAi), which will reveal their necessity in the biosynthetic pathway (Upadhyaya et al., 2021; Kim et al., 2021).

2) Gene expression analysis: Quantitative real-time PCR (qRT-PCR) can be employed to quantify the expression levels of *R2R3-MYB* genes and their target anthocyanin biosynthetic genes in diverse tissues and developmental stages of colored rice. This approach can facilitate the establishment of a correlation between gene expression and anthocyanin accumulation (Upadhyaya et al., 2021; Yang et al., 2023).

3) Protein-protein interaction studies: Yeast two-hybrid assays and co-immunoprecipitation can be employed to investigate interactions between R2R3-MYB proteins and other TFs, such as bHLH proteins, which are known to form complexes that regulate anthocyanin biosynthesis (Li et al., 2020; Kim et al., 2021).

4) Subcellular localization: Fluorescent protein tagging and confocal microscopy can be employed to ascertain the subcellular localization of R2R3-MYB proteins, thereby elucidating their functional roles as transcriptional regulators (Feng et al., 2018; Kim et al., 2021).

5) Metabolite analysis: High-performance liquid chromatography (HPLC) can be employed to quantify anthocyanin levels in transgenic and wild-type rice plants, thereby enabling an evaluation of the impact of *R2R3-MYB* gene manipulation on anthocyanin biosynthesis (Zhu et al., 2017; Upadhyaya et al., 2021).

The functional roles of *R2R3-MYB* genes in anthocyanin biosynthesis in colored rice can be comprehensively characterized by employing these experimental approaches, thereby providing valuable insights into their potential applications in crop improvement and biofortification.

7 Integration of Phylogenetic and Expression Data

7.1 Correlation between phylogenetic clades and expression profiles

Integrating phylogenetic analysis with gene expression data facilitates a comprehensive understanding of the functional dynamics of *R2R3-MYB* genes in colored rice. By correlating the phylogenetic clades with expression profiles obtained from various tissues and developmental stages, distinct expression patterns corresponding to specific clades were observed. Clades comprising genes with high sequence similarity often displayed comparable expression profiles, indicating the presence of conserved regulatory mechanisms. For example, genes within clade A were predominantly expressed in leaf tissues, indicating a potential role in photosynthesis-related processes. In contrast, genes in clade B exhibited high expression levels in root tissues, suggesting a potential role in root development and stress responses. This correlation highlights the evolutionary conservation of gene function within clades and demonstrates the value of the phylogenetic context in predicting gene expression patterns (Piao et al., 2019; Kang et al., 2022).

7.2 Insights into gene function from integrated analysis

The integration of phylogenetic and expression data provides valuable insights into the functional roles of *R2R3-MYB* genes, highlighting their biological significance. For example, genes in clade C, characterized by high expression during flower development, are likely involved in regulating floral organ differentiation and development. This hypothesis is supported by the presence of known floral regulators within this clade. Similarly, genes in clade D, with elevated expression under abiotic stress conditions, suggest roles in stress response pathways. The co-expression of *R2R3-MYB* genes with known stress-responsive genes further corroborates this functional association. By linking phylogenetic clades with specific expression profiles and functional annotations, potential roles for uncharacterized *R2R3-MYB* genes can be inferred, guiding future experimental validation and functional studies (Upadhyaya et al., 2021; Chen et al., 2022).

7.3 Case studies of key *R2R3-MYB* genes

To demonstrate the practical applications of our integrated analysis, we present case studies of key *R2R3-MYB* genes with significant functional implications.

The gene *OsMYB30*, a member of the R2R3-MYB transcription factor family, plays a critical role in regulating the expression of *phenylalanine ammonia-lyase* (*PAL*) genes, which are essential for lignin biosynthesis and plant defense mechanisms. Prior research has demonstrated that the tissue-specific expression of *PAL* is regulated by the R2R3 MYB transcription factor through binding to AC-rich elements. He et al. (2019) identified a mutant with

reduced free SA content and decreased BPH resistance (SA-R # 256) resulting from T-DNA insertion in the first exon of the R2R3 MYB transcription factor OsMYB30 (Figure 2). Quantitative reverse transcription polymerase chain reaction (qRT-PCR) analysis demonstrated that at 3 and 6 hours following BPH infection, the transcript level of *OsMYB30* in wild-type plants exhibited a notable induction, whereas no such induction was observed in *OsMYB30* mutant plants. Furthermore, the mutant system exhibited a notable reduction in lignin content. Furthermore, the induction of *OsPAL6* and *OsPAL8* expression by BPH was absent in the mutant line. The results suggest that *OsMYB30* may regulate the expression of *OsPAL* in response to BPH infection.

OsMYB60 is another R2R3-MYB transcription factor that has been the subject of study about its role in leaf morphology. Similarly to *OsMYB30*, *OsMYB60* plays a role in maintaining the flattening of the leaf blade. The interaction between OsMYB60 and TL is of great significance to this process. As demonstrated by Liu et al. (2018), the overexpression of *OsMYB60* results in the formation of twisted leaf blades (Figure 3), indicating that precise regulation of *OsMYB60* is necessary for the normal development of leaves. Moreover, OsMYB60 interacts with the C_2H_2 transcription factor OsZFP7, which further corroborates its involvement in leaf morphological development.

8 Future Perspectives and Opportunities

8.1 Advances in genomic technologies

The rapid advancement in genomic technologies has significantly enhanced our understanding of the genetic basis of complex traits in rice, including the R2R3-MYB gene family. Whole genome sequencing and comparative genomic analysis have revealed a multitude of allelic variations that are unique to specific rice landraces. These include Purpleputtu, which exhibits unique traits such as purple coloration due to specific genetic variations (Lachagari et al., 2019). These technologies facilitate the identification of pivotal regulatory genes and their variants, thereby providing a comprehensive perspective on the genetic architecture that underlies significant agronomic traits. It is recommended that future research focus on leveraging these advanced genomic tools to further dissect the functional roles of R2R3-MYB genes in rice. This will facilitate the development of high-throughput genotyping platforms for precise genetic mapping and trait association studies (Zhao et al., 2011; Gu et al., 2022).

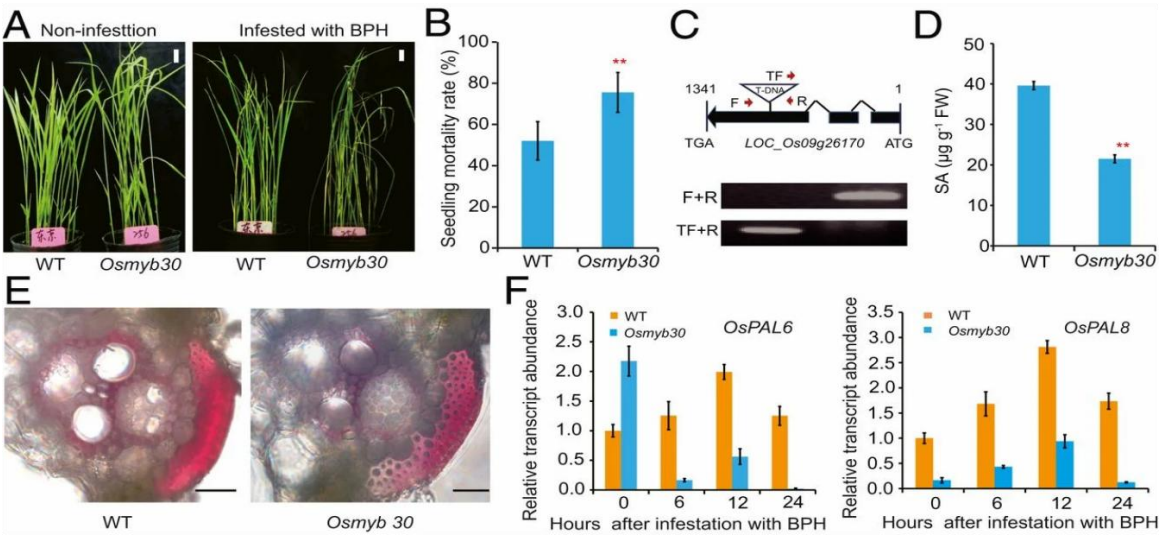

Figure 2 OsPALs-mediated BPH resistance is regulated by *OsMYB30* (Adopted from He et al., 2019)

Image caption: (A) Representative image of *Osmyb30* mutants infested with BPH for 5 d. (Scale bars, 1 cm.) (B) Seedling mortality rates of WT and *Osmyb30* mutants infested with BPH. Data were collected at 7 dpi. (C) Verification of the T-DNA insertion in the *Osmyb30* mutant. The positions of the F, R, and TF primers are indicated as red arrows. (D) SA levels in WT and the *Osmyb30* mutant plants. (E) Histochemical staining showing lignin accumulation in fresh leaf sheaths of WT and the *Osmyb30* mutant plants. (Scale bar, 50 μm.) (F) qRT-PCR analysis of *OsPAL6* and *OsPAL8* expression in the *Osmyb30* mutant. The expression values are presented relative to those in WT without BPH infestation. Error bars, mean ± SD of 3 biological replicates, by Student's t-test (B and D, **P < 0.01) (Adopted from He et al., 2019)

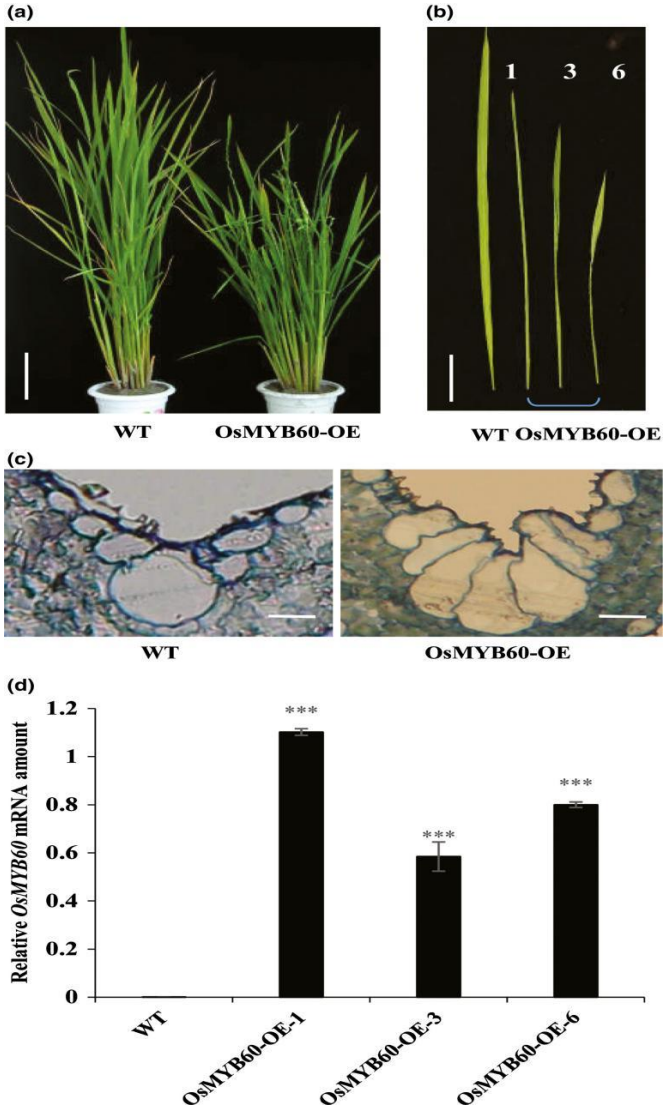

Figure 3 Phenotypes of the *OsMYB60*-overexpressing (*OsMYB60*-OE) transgenic plants (Adopted from Liu et al., 2018)

Image caption: (a) Phenotypes of 100-d-old wild-type (WT, left) and *OsMYB60*-OE (right) transgenic rice plants. (b) Flag leaf blades of 100-d-old WT (left) and *OsMYB60*-OE (right: #1, #3 and #6) transgenic rice plants. *OsMYB60*-OE transgenic rice plants showed twisted leaf blades. (c) Histological analysis of the leaf blade in the middle part of WT and *OsMYB60*-OE transgenic rice plants. The shape and arrangement of bulliform cells in the *OsMYB60*-OE leaves were changed compared to WT. (d) *OsMYB60* expression in *OsMYB60*-OE transgenic rice plants was determined by quantitative reverse transcription polymerase chain reaction (qRT-PCR). The Rice ACTIN1 gene was used as an internal control. Error bars show ± standard deviation from three replicates. Asterisks indicate a significant difference between *OsMYB60*-OE transgenic plants and WT controls by Student's t-test: ***, $P < 0.001$. Bars: (a) 20 cm; (b) 5 cm; (c) 25 μm (Adopted from Liu et al., 2018)

8.2 Potential for genetic engineering and crop improvement

The identification and characterization of R2R3-MYB genes have opened new avenues for genetic engineering aimed at crop improvement. For example, the overexpression of R2R3-MYB genes, such as OsMYB2P-1, has been demonstrated to enhance tolerance to phosphate starvation and improve root architecture in rice (Dai et al., 2012). Similarly, OsMYB103L has been demonstrated to influence leaf rolling and mechanical strength, which are critical traits for rice breeding (Yang et al., 2014). Genetic engineering techniques can be utilized to alter these genes, thereby developing rice varieties with enhanced stress tolerance, improved nutrient use efficiency, and desirable morphological characteristics. Furthermore, the distinctive allelic variations observed in purple rice landraces can be utilized to incorporate advantageous traits, such as enhanced nutritional value and stress resistance, into contemporary cultivars through targeted breeding programs (Wang and Shu, 2007; Lachagari et al., 2019).

8.3 Challenges and prospects in R2R3-MYB gene research

Despite the promising potential of R2R3-MYB genes in rice improvement, several challenges remain. One significant challenge is the functional redundancy and intricate regulatory networks involving multiple MYB genes, which render the elucidation of their specific roles a complex undertaking. Furthermore, the co-evolutionary patterns observed between different domains of R2R3-MYB proteins indicate the presence of intricate protein-DNA and protein-protein interactions that require comprehensive investigation. It is recommended that future research employ integrative approaches combining genomics, transcriptomics, proteomics, metabolomics, and phenomics to unravel the complex regulatory mechanisms of R2R3-MYB genes (Chen et al., 2022; Gupta et al., 2023; Xiong et al., 2022; Zhang et al., 2023; Zhang et al., 2022; Zhao et al., 2024). Furthermore, the advent of sophisticated gene editing techniques, such as CRISPR/Cas9, presents a promising avenue for precise manipulation of MYB genes to achieve desired phenotypic outcomes in rice (Dai et al., 2007; Biswal et al., 2019).

9 Concluding Remarks

This systematic review has comprehensively identified and characterized the *R2R3-MYB* gene family in rice, thereby providing significant insights into their evolutionary relationships, structural features, and expression profiles. Notable findings include the distinctive structural characteristics of *R2R3-MYB* genes in colored rice, such as elevated frequencies of non-synonymous substitutions in the DNA-binding domains, particularly in the α-helix regions. This suggests adaptive selection and functional diversification. The expression patterns of *R2R3-MYB* genes were found to be tissue-specific and developmentally regulated, with genes within certain clades predominantly expressed in specific tissues, such as leaves and roots. This suggests that they may have specialized roles in photosynthesis-related processes and root development, respectively. Moreover, *R2R3-MYB* genes were identified as pivotal regulators of anthocyanin biosynthesis. Genes such as *OsC1* and *OsKala3* were found to play a pivotal role in modulating the expression of anthocyanin biosynthetic pathway genes, underscoring their significance in conferring the distinctive purple pigmentation and associated health benefits observed in purple rice.

The findings from this study have significant implications for the breeding of colored rice. The identification of unique allelic variations and key regulatory genes provides valuable genetic resources that can be utilized to enhance desirable traits such as disease resistance, stress tolerance, and nutritional value. The functional characterization of *R2R3-MYB* genes, such as *OsMYB2P-1* and *OsMYB103L*, offers potential targets for genetic engineering to improve traits like leaf morphology and pericarp pigmentation, which are important for both agronomic performance and consumer preference. The use of advanced genomic tools, such as CRISPR/Cas9, can further facilitate the precise editing of these genes to achieve targeted improvements in grain yield and other agronomically important traits.

The broader impact of this research extends beyond the specific topic of colored rice to the wider field of plant science and agriculture. The insights gained into the genetic architecture and regulatory mechanisms of *R2R3-MYB* genes can inform the breeding and genetic engineering of other crop species, thereby contributing to global food security and agricultural sustainability. The study also highlights the necessity of conserving and employing genetic diversity in crop improvement programs, as evidenced by the extensive allelic diversity observed in the *Purpleputtu* rice landrace. Moreover, the public accessibility of genomic data from extensive initiatives, such as the 3 000 Rice Genomes Project, offers a significant asset for continuous research and breeding endeavors aimed at improving crop performance and resilience.

Acknowledgments

We appreciate the anonymous peer reviewers for their revision suggestions on the manuscript.

Funding

This work was supported by the Scientific Research Foundation of Panxi Crops Research and Utilization Key Laboratory of Sichuan

Province (Grant No. XNFZ2203), the Ph.D. Programs Foundation of Xichang University (Grant No. YBZ202340), and the Key and Major Science and Technology Projects of Yunnan (Grant nos. 202202AE09002102).

Conflict of Interest Disclosure

The authors affirm that this research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

Bao X., Zong Y., Hu N., Li S., Liu B., and Wang H., 2021, Functional *R2R3-MYB* transcription factor NsMYB1, regulating anthocyanin biosynthesis, was relative to the fruit color differentiation in *Nitraria sibirica* Pall, BMC Plant Biology, 22: 186.
https://doi.org/10.1186/s12870-022-03561-5

Biswal A., Mangrauthia S., Reddy M., and Yugandhar P., 2019, CRISPR mediated genome engineering to develop climate-smart rice: challenges and opportunities, Seminars in Cell & Developmental Biology, 96: 100-106.
https://doi.org/10.1016/j.semcdb.2019.04.005

Chen Q., Zhang X., Fang Y., Wang B., Xu S., Zhao K., Zhang J., and Fang J., 2022, Genome-wide identification and expression analysis of the R2R3-MYB transcription factor family revealed their potential roles in the flowering process in longan (*Dimocarpus longan*), Frontiers in Plant Science, 13: 820439.
https://doi.org/10.3389/fpls.2022.820439

Chen X., Yang Y., Yang X., Zhu G., Lu X., Jia F., Diao B., Yu S., Ali A., and Zhang H., 2022, Investigation of flavonoid components and their associated antioxidant capacity in different pigmented rice varieties, Food Research International, 161: 111726.
https://doi.org/10.1016/j.foodres.2022.111726

Dai X., Wang Y., Yang A., and Zhang W., 2012, OsMYB2P-1, an R2R3 MYB transcription factor, is involved in the regulation of phosphate-starvation responses and root architecture in rice, Plant Physiology, 159: 169-183.
https://doi.org/10.1104/pp.112. 194217

Dai X., Xu Y., Ma Q., Xu W., Wang T., Xue Y., and Chong K., 2007, Overexpression of an *R1R2R3 MYB* gene, *OsMYB3R-2*, increases tolerance to freezing, drought, and salt stress in transgenic Arabidopsis, Plant Physiology, 143: 1739-1751.
https://doi.org/10.1104/pp.106.094532

Duan A., Tan S., Deng Y., Xu Z., and Xiong A., 2022, Genome-wide identification and evolution analysis of *R2R3-MYB* gene family reveals S6 subfamily R2R3-MYB transcription factors involved in anthocyanin biosynthesis in Carrot, International Journal of Molecular Sciences, 23(19): 11859.
https://doi.org/10.3390/ijms231911859

Dwiningsih Y., and Alkahtani J., 2022, Genetics, biochemistry and biophysical analysis of anthocyanin in rice (*Oryza sativa* L.), Advance Sustainable Science Engineering and Technology, 4(1): 1-19.
https://doi.org/10.26877/asset.v4i1.11659

Feng K., Xu Z., Que F., Liu J., Wang F., and Xiong A., 2018, An R2R3-MYB transcription factor, OjMYB1, functions in anthocyanin biosynthesis in *Oenanthe javanica*, Planta, 247: 301-315.
https://doi.org/10.1007/s00425-017-2783-8

Gu H., Liang S., and Zhao J., 2022, Novel sequencing and genomic technologies revolutionized rice genomic study and breeding, Agronomy, 12(1): 218.
https://doi.org/10.3390/agronomy12010218

Gupta R., Min C. W., Jung J.Y., Ham T.H., Jeon J.S., Cho L.H., Kwon S.W., and Kim S.T., 2023, Proteome profiling highlights mechanisms underlying pigment and tocopherol accumulation in red and black rice seeds, Proteomics, 23: 2300035.
https://doi.org/10.1002/pmic.202300035

He J., Liu Y., Yuan D., Duan M., Liu Y., Shen Z., Yang C., Qiu Z., Liu D., Wen P., Huang J., Fan D., Xiao S., Xin Y., Chen X., Jiang L., Wang H., Yuan L., and Wan J., 2019, An R2R3 MYB transcription factor confers brown planthopper resistance by regulating the phenylalanine ammonia-lyase pathway in rice, Proceedings of the National Academy of Sciences, 117: 271-277.
https://doi.org/10.1073/pnas.1902771116

Hua X.F., Bi C.Y., Huang, B.F., Xu M., Yang Z.J., Lin, S.Q., and Chen X.Y., 2020, Identification and analysis of R2R3-MYB genes in sweet potato genome, Molecular Plant Breeding, 11(22): 1-12.
https://doi.org/10.5376/mpb.2020.11.0022

Jia L., Clegg M., and Jiang T., 2004, Evolutionary dynamics of the DNA-binding domains in putative *R2R3-MYB* genes identified from rice subspecies indica and japonica genomes, Plant Physiology, 134: 575-585.
https://doi.org/10.1104/pp.103.027201

Kang L., Teng Y., Cen Q., Fang Y., Tian Q., Zhang X., Wang H., Zhang X., and Xue D., 2022, Genome-wide identification of R2R3-MYB transcription factor and expression analysis under abiotic stress in rice, Plants, 11(15): 1928.
https://doi.org/10.3390/plants11151928

Khan I., Khan S., Zhang Y., and Zhou J., 2021, Genome-wide analysis and functional characterization of the Dof transcription factor family in rice (*Oryza sativa* L.), Planta, 253: 101.
https://doi.org/10.1007/s00425-021-03627-y

Kim D., Yang J., Ha S., Kim J., Lee J., and Lim S., 2021, An OsKala3, R2R3 MYB TF, Is a common key player for black rice pericarp as main partner of an OsKala4, bHLH TF, Frontiers in Plant Science, 12: 765049.
https://doi.org/10.3389/fpls.2021.765049

Lachagari V., Gupta R., Lekkala S., Mahadevan L., Kuriakose B., Chakravartty N., Katta A., Santhosh S., Reddy A., and Thomas G., 2019, Whole genome sequencing and comparative genomic analysis reveal allelic variations unique to a purple colored rice landrace (*Oryza sativa* ssp. indica cv. *Purpleputtu*), Frontiers in Plant Science, 10: 513.
https://doi.org/10.3389/fpls.2019.00513

Li Q., Lu L., Liu H., Bai X., Zhou X., Wu B., Yuan M., Yang L., and Xing Y., 2020, A minor QTL, SG3, encoding an R2R3-MYB protein, negatively controls grain length in rice, Theoretical and Applied Genetics, 133: 2387-2399.
https://doi.org/10.1007/s00122-020-03606-z

Li Y., Shan X., Tong L., Wei C., Lu K., Li S., Kimani S., Wang S., Wang L., and Gao X., 2020, The conserved and particular roles of *R2R3-MYB* regulator FhPAP1 from *Freesia hybrida* in flower anthocyanin biosynthesis, Plant & Cell Physiology, 61(7): 1365-1380.
https://doi.org/10.1093/pcp/pcaa065

Liu X., Li D., Zhang D., Yin D., Zhao Y., Ji C., Zhao X., Li X., He Q., Chen R., Hu S., and Zhu L., 2018, A novel antisense long noncoding RNA, TWISTED LEAF, maintains leaf blade flattening by regulating its associated sense *R2R3-MYB* gene in rice, The New Phytologist, 218(2): 774-788.
https://doi.org/10.1111/nph.15023

Mackon E., Jeazet Dongho Epse Mackon G.C., Ma Y., Haneef Kashif M., Ali N., Usman B., and Liu P., 2021, Recent insights into anthocyanin pigmentation, synthesis, trafficking, and regulatory mechanisms in rice (*Oryza sativa* L.) caryopsis, Biomolecules, 11: 394.
https://doi.org/10.3390/biom11030394

Monna L., Ohta R., Masuda H., Koike A., and Minobe Y., 2006, Genome-wide searching of single-nucleotide polymorphisms among eight distantly and closely related rice cultivars (*Oryza sativa* L.) and a wild accession (*Oryza rufipogon* Griff.), DNA Research, 13(2): 43-51.
https://doi.org/10.1093/DNARES/DSI030

Sakamoto W., Ohmori T., Kageyama K., Miyazaki C., Saito A., Murata M., Noda K., and Maekawa M., 2001, The purple leaf (Pl) locus of rice: the Pl(w) allele has a complex organization and includes two genes encoding basic helix-loop-helix proteins involved in anthocyanin biosynthesis, Plant & Cell Physiology, 42(9): 982-991.
https://doi.org/10.1093/PCP/PCE128

Shi Y., Zeng W., Xu M., Li H., Zhang F., Chen Z., Anicet G., Huang S., Huang Y., Wang X., Li J., Zhang X., Zheng Y., and Rehman S., 2022, Comprehensive analysis of glutamate receptor-like genes in rice (*Oryza sativa* L.): genome-wide identification, characteristics, evolution, chromatin accessibility, gcHap Diversity, population variation and expression analysis, Current Issues in Molecular Biology, 44: 6404-6427.
https://doi.org/10.3390/cimb44120437

Shin D., Choi M., Kang C., Park C., Choi S., and Park Y., 2016, A wheat R2R3-MYB protein PURPLE PLANT1 (TaPL1) functions as a positive regulator of anthocyanin biosynthesis, Biochemical and Biophysical Research Communications, 469(3): 686-691.
https://doi.org/10.1016/j.bbrc.2015.12.001

Temnykh S., DeClerck G., Lukashova A., Lipovich L., Cartinhour S., and McCouch S., 2001, Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): frequency, length variation, transposon associations, and genetic marker potential, Genome Research, 11(8): 1441-1452.
https://doi.org/10.1101/GR.184001

Upadhyaya G., Das A., and Ray S., 2021, A rice *R2R3-MYB* (OsC1) transcriptional regulator improves oxidative stress tolerance by modulating anthocyanin biosynthesis, Physiologia Plantarum, 173(4): 2334-2349.
https://doi.org/10.1111/ppl.13583

Wang C., and Shu Q., 2007, Fine mapping and candidate gene analysis of purple pericarp gene Pb in rice (*Oryza sativa* L.), Chinese Science Bulletin, 52: 3097-3104.
https://doi.org/10.1007/S11434-007-0472-X

Xia D., Zhou H., Wang Y., Li P., Fu P., Wu B. and He Y., 2021, How rice organs are colored: the genetic basis of anthocyanin biosynthesis in rice, The Crop Journal, 9: 598-608.
https://doi.org/10.1016/j.cj.2021.03.01

Xiong Q., Sun C., Li A., Zhang J., Shi Q., Zhang Y., Hu J., Zhou N., Wei H., and Liu B., 2022, Metabolomics and biochemical analyses revealed metabolites important for the antioxidant properties of purple glutinous rice, Food Chemistry, 389: 133080.
https://doi.org/10.1016/j.foodchem.2022.133080

Yan J., Zou D., Li C., Zhang Z., Song S., and Wang X., 2020, SR4R: An integrative SNP resource for genomic breeding and population research in rice, Genomics, Proteomics & Bioinformatics, 18: 173-185.
https://doi.org/10.1016/j.gpb.2020.03.002

Yang C., Li D., Liu X., Ji C., Hao L., Zhao X., Li X., Chen C., Cheng Z., and Zhu L., 2014, OsMYB103L, an R2R3-MYB transcription factor, influences leaf rolling and mechanical strength in rice (*Oryza sativa* L.), BMC Plant Biology, 14: 158.
https://doi.org/10.1186/1471-2229-14-158

Yang W., Feng L., Luo J., Zhang H., Jiang F., He Y., Li X., Du J., Adjei M., Luan A., and Ma J., 2023, Genome-wide identification and characterization of *R2R3-MYB* provide insight into anthocyanin biosynthesis regulation mechanism of ananas comosus var bracteatus, International Journal of Molecular Sciences, 24(4): 3133.
https://doi.org/10.3390/ijms24043133

Zhang Z., Zhang F., Deng Y., Sun L., Mao M., Chen R., Qiang Q., Zhou J., Long T., and Zhao X., 2022, Integrated metabolomics and transcriptomics analyses reveal the metabolic differences and molecular basis of nutritional quality in landraces and cultivated rice, Metabolites, 12: 384.
https://doi.org/10.3390/metabo12050384

Zhang L., Cui D., Ma X., Han B., and Han L., 2023, Comparative analysis of rice reveals insights into the mechanism of colored rice via widely targeted metabolomics, Food Chemistry, 399: 133926.
https://doi.org/10.1016/j.foodchem.2022.13392

Zhao K., Tung C., Eizenga G., Wright M., Ali M., Price A., Norton G., Islam S., Reynolds A., Mezey J., McClung A., Bustamante C., Bustamante C., and McCouch S., 2011, Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*, Nature Communications, 2(1): 467.
https://doi.org/10.1038/ncomms1467

Zhao M., Zhai L., Tang Q., Ren J., Zhou S., Wang H., Yun Y., Yang Q., Yan X., and Xing F., 2024, Comparative metabolic profiling of different colored rice grains reveals the distribution of major active compounds and key secondary metabolites in green rice, Foods, 13: 1899.
https://doi.org/10.3390/foods13121899

Zhu Q., Yu S., Zeng D., Liu H., Wang H., Yang Z., Xie X., Shen R., Tan J., Li H., Zhao X., Zhang Q., Chen Y., Guo J., Chen L., and Liu Y., 2017, Development of "Purple Endosperm Rice" by engineering anthocyanin biosynthesis in the endosperm with a high-efficiency transgene stacking system, Molecular Plant, 10(7): 918-929.
https://doi.org/10.1016/j.molp. 2017.05.008