

HAPLOTYPE-BASED IDENTIFICATION OF SUPERIOR *OsPYL9* ALLELES ASSOCIATED WITH GRAIN YIELD IN DIVERSE RICE GERMPLASM

Janani Santhi Kuppan¹, Williams Mohanvel¹, Veera Ranjani Rajagopalan¹, Sudha Manickam¹, Raveendran Muthurajan^{1*}

¹Department of Plant Biotechnology, Centre for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

* Correspondence Author: Raveendran Muthurajan, Email Id: raveendrantsnau@gmail.com

ABSTRACT

Rice (*Oryza sativa* L.) is one of the most important staple crops, and enhancing yield and ensuring its resilience to environmental stresses is vital to ensure food security in the future. Use of favourable natural alleles is a potential tool to improve complex traits like grain yield. The gene *OsPYL9* encodes ABA receptor that is associated with drought and grain yield responses in rice. But the role of naturally-occurring *OsPYL9* allelic variation in grain yield has not been adequately studied. In this study, 203 rice accessions from the 3K Rice Genome panel were tested for plant yield. Haplotype analysis was performed using single nucleotide polymorphisms (SNPs) in the coding region of *OsPYL9* (LOC_Os06g33670) in the Rice SNP-Seek database. Eight SNPs resulted in two major haplotypes, H1 (60.0%) and H2 (37.3%). The allelic combinations were GGGGAATG for H1 and AGGGTATG for H2. Haplo-pheno analysis showed significant difference between the two haplotypes. The mean single plant yield of the accessions carrying H1 allele was about 17.3 g while for H2 it was higher at about 19.5 g. The difference between the two haplotypes was 2.26 g and significant when tested by one-way analysis of variance ($P = 0.031$) and Hsu's simultaneous multiple-comparison test (adjusted $P = 0.015$). Thus, H2 was identified as a superior *OsPYL9* haplotype associated with higher plant yield in the evaluated germplasm. The identified haplotype H2 could be a potential genetic target for further validation and use in haplotype-assisted rice breeding.

KEYWORDS: *Oryza sativa*, *OsPYL9*, haplotype, Plant yield, ABA signalling, 3K Rice Genome

INTRODUCTION

Rice (*Oryza sativa* L.) is an important food crops for half of the world's population. The global population is expected to reach about 9.7 billion by 2050 (United Nations Department of Public Information) and global food demand is forecast to rise by nearly 60-70% relative to current levels, and thus increasing rice productivity is crucial to ensure food security for the future(1).

The development of high yielding rice cultivars can be achieved by improving key traits such as plant height, tiller number, panicle architecture, number of grains per panicle, grain size, grain filling and thousand grain weight(2,3). These traits were controlled by intricate genetic networks involving multiple phytohormones and signalling pathways(2). The discovery and use of superior alleles for crop improvement is an another approach that has been accomplished by recent advances in marker-assisted breeding, mutation breeding and genome-editing technologies(4,5).

Domestication of wild species and traditional breeding methods have identified natural genetic variations that were important sources of beneficial alleles for crop improvement. Haplotype analysis is used to characterize natural allelic variation, identifying superior haplotypes related to desirable agronomic traits and serve as important genetic resources for marker-assisted breeding and genome-editing programmes(6,7). In spite of the proven potential of *OsPYL9* to enhance grain yield, the natural allelic variation and haplotype diversity of *OsPYL9* in rice germplasm has not been studied. The introgression of superior *OsPYL9* haplotypes would contribute novel alleles for developing high yielding rice cultivars.

Abscisic acid (ABA) is one of the several phytohormones involved in regulating plant growth, development and grain yield through the PYRABACTIN RESISTANCE/PYRABACTIN RESISTANCE-LIKE (PYR/PYL) signalling pathway. *OsPYL9* is an important candidate gene for rice improvement among the ABA receptor genes. Creation of 1-4 bp deletions in *OsPYL9* gene through CRISPR-Cas9 approach improved grain yield under well-watered and drought conditions(8). This evidence highlights the importance of *OsPYL9* in regulating grain yield and provides a rationale for investigating naturally occurring allelic variation of this gene in rice germplasm. Identification of novel alleles in germplasm through haplotype analysis could provide useful genetic resources for breeding rice cultivars suitable for changing climatic conditions. Hence, the present study was designed to investigate the allelic variation and haplotype diversity of *OsPYL9* in diverse rice accessions of 3K Rice Genome panel.

MATERIALS AND METHODS

A diverse panel of 203 rice accessions from the 3K Rice Genome (3K RG) collection(9) was screened in an augmented design at the Paddy Breeding Station (PBS), Tamil Nadu Agricultural University (TNAU), Coimbatore, India (11°00' N, 76°54' E; 426 m above mean sea level) during kharif (wet season) 2021 and rabi (post-monsoon season) 2022. Twenty-five day old seedlings were transplanted at 20 × 20 cm spacing following augmented design and all the recommended agronomic practices were followed throughout the crop growth period with 150:50:50 kg ha⁻¹ (N, P₂O₅ and K₂O) respectively.

For grain yield per plant, five plants were randomly selected from each accession at physiological maturity. The plants were harvested individually, and the grains obtained from each plant were cleaned and dried to approximately 14% moisture content. Grain yield was determined by weighing the total grain obtained from each individual plant using a precision electronic balance and expressed in grams per plant (g plant⁻¹). The grain yield values of the five plants were averaged to obtain the mean grain yield for each accession, which was subsequently used for haplotype–phenotype association analysis.

Retrieval of SNP data and haplotype analysis

Allelic diversity for *OsPYL9* gene (LOC_Os06g36670) was obtained from the Rice SNP-Seek database that has genomic information from the 3K Rice Genome Project(10). Allele mining was employed only on single nucleotide polymorphisms (SNPs) in the coding region of the target gene. The SNP dataset was downloaded as PLINK format and then converted to Haploview-formatted files with the gPLINK program (version 1.07)(11).

The analysis of haplotypes was performed with Haploview version 4.1(12) and haplotype blocks were created based on linkage disequilibrium (LD) and SNPs with a minor allele frequency (MAF) ≥ 0.001 were included in the analysis. The genotypes with heterozygous alleles or missing alleles were not included in the analysis. Unique haplotypes of *OsPYL9* were identified using SNPs and the haplotypes were then correlated with grain yield. Only the haplotype groups with genotype frequency more than three were considered for haplotype-phenotype analysis.

Statistical analysis

The descriptive statistics such as mean, range and coefficient of variation (CV) were obtained by Minitab Statistical Software (Version 19.1). One-way analysis of variance (ANOVA) was used to test the differences in the identified haplotypes for grain yield per plant. A simultaneous multiple comparison test was used to find superior haplotypes with significantly better performance using Hsu's method(13). Box plots were created to show the distribution of the phenotypical traits in the different haplotype groups(14).

RESULTS

Phenotypic variation for grain yield per plant

A considerable variation in grain yield per plant was observed among the rice accessions evaluated. The descriptive statistics for grain yield per plant are presented in Table 1. The grain yield ranged from 3.61 to 42.39 g plant⁻¹, with a mean of 18.57 g plant⁻¹ and a coefficient of variation of 40.84%, indicating substantial phenotypic variation among the evaluated accessions. The frequency distribution further illustrated the broad variation in grain yield across the evaluated population, with most accessions concentrated around the intermediate yield classes (Fig. 1)(Table 1)

Identifying allelic variation present in the coding region of *OsPYL9*.

Allele mining of the *OsPYL9* coding region using the Rice SNP-Seek database identified eight SNPs distributed across the target region at positions 10621556600, 10621556606, 10621556612, 10621556937, 10621557034, 10621557172, 10621557307 and 10621557332. The haplotype analysis of these eight markers resulted in two haplotype groups in a single haplotype block, explaining strong linkage between the informative polymorphic sites (IPSS) within the population (Fig. 3). A very strong pairwise linkage disequilibrium (LD) was observed between the two SNPs located at positions 10621556600 and 10621557034 where D' value is 0.99 within the block. These polymorphic sites were important in the differentiation of the main haplotypic combinations found in the population.

The haplotype H1 with allelic combination “GGGGAATG” had 60.0% frequency and the haplotype H2 with allelic combination “AGGGTATG” had 37.3% frequency (Fig. 2). This is because, at the two differentiating SNP positions, H1 had the alleles G and A, while H2 had the alleles A and T respectively. All six other SNP positions had the same allele for both of the major haplotypes. The remaining 2.7% of accessions were minor haplotypic combinations and were not used for the analysis.(Fig. 2)(Fig. 3)

Association of *OsPYL9* haplotypes with yield

Significant variation in plant yield was observed between the two major *OsPYL9* haplotypes (Fig. 4). A mean plant yield of about 17.3 g plant⁻¹ was observed for the H1 haplotype (alleles GA) while the mean plant yield for the H2 haplotype (alleles AT) was about 19.5 g plant⁻¹. The yield difference between the two haplotypes was 2.26 g plant⁻¹ which was approximately 13.1% more plant yield by H2 than H1.

The plant yield showed a significant difference between the two haplotypes by one way analysis of variance (P = 0.031). The difference between H1 (GA) and H2 (AT) was further confirmed by Hsu's simultaneous multiple-comparison test, which showed a significant difference between the haplotypes with p<0.05 (P = 0.015). These results showed the AT haplotype combination (H2) had higher plant yield compared to the GA haplotype combination (H1), thus H2 is the better haplotype in the rice germplasm tested.(Fig. 4)(Table. 2)(Table. 3)

DISCUSSION

OsPYL9 encodes an ABA receptor involved in ABA-mediated regulation of plant growth and responses to abiotic stress. Previous functional characterization showed that alteration of *OsPYL9* through CRISPR/Cas9 affected rice grain yield under both well-watered and drought conditions, which highlights its importance for rice production(8). These findings establish *OsPYL9* as a functionally important candidate gene for rice improvement. Although earlier studies focused on functional characterization of *OsPYL9* using targeted mutagenesis, the current study aimed to explore naturally occurring allelic variation in this gene. Therefore, characterization of natural allelic variation at *OsPYL9* provides an approach for identifying favourable genetic variation that may be useful for rice breeding.

The previous studies reported a significant contribution of natural allelic and haplotypic variation to rice productivity. Superior haplotypes of the gene *CCD8* were associated with higher total and productive tiller numbers(15), whereas different alleles of the gene *SCM2* were associated with variation in culm diameter and lodging-related traits(16). Allele mining of the grain-number gene *An-1* also revealed favourable allelic variation associated with enhanced grain number and yield-related traits (17). Analysis of 120 functionally characterized genes across the 3K Rice Genome panel showed that more than 75% of genes had 2–15 haplotypes, further indicating the wide genetic diversity present in rice(18). Furthermore, identification of haplotypes of drought responsive genes showed that certain haplotypes were significantly associated with higher single plant yield under drought stress, suggesting their potential as donor haplotypes for breeding climate resilient rice cultivars(19). Similarly, the *EPN4-1* haplotype was associated with favourable variation in effective panicle number, grain number and grain yield(20), whereas an elite *STRONG1* haplotype was associated with improved plant architecture, lodging resistance and grain yield(21). Recently, haplotypes of several genes contributing to higher grain yield under reproductive-stage drought stress were identified, whereas *OsIRO2-H3* haplotype was validated through introgression, which resulted in 25.0–27.3% yield advantage under drought conditions(22). Favourable haplotypes of *SMG11/D2* and *GW7* were associated with variation in grain-size traits(23). Collectively, these studies demonstrate the importance of natural allelic and haplotypic variation represent a valuable genetic resources for rice breeding.

Haplotype-based breeding (HBB) provides an opportunity to utilize favourable combinations of linked variants rather than relying on individual polymorphisms alone(24,25). Haplotype analysis of the 3K Rice Genome panel identified superior haplotypes and haplotype combinations for 21 key genes associated with grain yield and quality traits, demonstrating the potential of natural haplotypic variation for rice improvement(18). Likewise, superior haplotypes for preferred grain and cooking quality and low glycemic index have been identified which shows the application of haplotype-based approaches beyond yield improvement(26). These studies support the potential use of favourable haplotypes as genetic resources in breeding programmes.

Based on the present results, the higher grain yield associated with H2 *OsPYL9* suggests that genotypes in this haplotype could serve as a potential donor for future rice breeding. The LD information for the two differentiating SNPs (Fig. 3) can be used to develop haplotype-specific markers for tracking H2 in breeding populations. Further evaluation of H2 across multiple populations, genetic backgrounds and environments are required to confirm its stability and breeding value. Functional characterization of the differentiating variants would further clarify their contribution to grain yield. These studies would help determine whether H2 and its associated SNPs can be reliably exploited through marker-assisted or haplotype-assisted breeding for rice yield improvement.

CONCLUSION

In 203 rice accessions in the 3K Rice Genome panel, two major naturally occurring *OsPYL9* haplotypes (H1: GGGGAATG and H2: AGGGTATG) were identified. The AT-type H2 haplotype showed higher grain yield per plant (19.5 g plant⁻¹) than the GA-type H1 haplotype (17.3 g plant⁻¹). The difference between the two haplotypes was statistically significant and was also verified by Hsu's simultaneous multiple-comparison test (adjusted P = 0.015).

The high linkage disequilibrium between the two differentiating SNPs will underpin for further study on their use as markers for discriminating the major *OsPYL9* haplotypes. The results of the present study suggest that naturally occurring allelic combination (H2) is a potential combination that may be correlated with higher grain yield in the rice germplasm tested.

The present study, however, provides a haplotype–phenotype association and does not provide evidence for the individual SNPs causing the phenotype. The identified variants need further validation in independent populations and multiple environments, and their functional characterization are required to conclude the stability and biological basis of the yield advantage associated with H2 and assess the utility of its application in marker-assisted and haplotype-assisted rice breeding.

ACKNOWLEDGEMENTS

Funding The authors acknowledge the Department of Biotechnology (DBT), Ministry of Science and Technology, India for providing financial assistance under the Grant No. BT/ INF/22/SP45584/2022.

REFERENCES

1. Shi J, An G, Weber APM, Zhang D. Prospects for rice in 2050. *Plant Cell Environ.* 2023 Apr;46(4):1037–45. doi:10.1111/pce.14565

2. Li N, Xu R, Li Y. Molecular Networks of Seed Size Control in Plants. *Annu Rev Plant Biol.* 2019 Apr 29;70(1):435–63. doi:10.1146/annurev-arplant-050718-095851
3. Xing Y, Zhang Q. Genetic and Molecular Bases of Rice Yield. *Annu Rev Plant Biol.* 2010 Jun 2;61(1):421–42. doi:10.1146/annurev-arplant-042809-112209
4. Mishra R, Joshi RK, Zhao K. Genome Editing in Rice: Recent Advances, Challenges, and Future Implications. *Front Plant Sci.* 2018 Sep 19;9:1361. doi:10.3389/fpls.2018.01361
5. Gao C. Genome engineering for crop improvement and future agriculture. *Cell.* 2021 Mar;184(6):1621–35. doi:10.1016/j.cell.2021.01.005
6. Bhat JA, Yu D, Bohra A, Ganie SA, Varshney RK. Features and applications of haplotypes in crop breeding. *Commun Biol.* 2021 Nov 4;4(1):1266. doi:10.1038/s42003-021-02782-y
7. Sivabharathi RC, Rajagopalan VR, Suresh R, Sudha M, Karthikeyan G, Jayakanthan M, et al. Haplotype-based breeding: A new insight in crop improvement. *Plant Sci.* 2024 Sep;346:112129. doi:10.1016/j.plantsci.2024.112129
8. Usman B, Nawaz G, Zhao N, Liao S, Liu Y, Li R. Precise Editing of the OsPYL9 Gene by RNA-Guided Cas9 Nuclease Confers Enhanced Drought Tolerance and Grain Yield in Rice (*Oryza sativa* L.) by Regulating Circadian Rhythm and Abiotic Stress Responsive Proteins. *Int J Mol Sci.* 2020 Oct 23;21(21):7854. doi:10.3390/ijms21217854
9. Li JY, Wang J, Zeigler RS. The 3,000 rice genomes project: new opportunities and challenges for future rice research. *GigaScience.* 2014 Dec;3(1):8. doi:10.1186/2047-217X-3-8
10. The 3,000 rice genomes project. The 3,000 rice genomes project. *Gigascience.* 2014 Dec 1;3(1):2047-217X-3–7. doi:10.1186/2047-217X-3-7
11. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, et al. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet.* 2007 Sep;81(3):559–75. doi:10.1086/519795
12. Barrett JC, Fry B, Maller J, Daly MJ. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics.* 2005 Jan 15;21(2):263–5. doi:10.1093/bioinformatics/bth457
13. Hsu JC. Multiple comparisons: theory and methods. 1. ed., 1. CRC Press repr. Boca Raton, Fla.: Chapman & Hall/CRC; 1999. 277 p.
14. Allen TT. Software Overview and Methods Review: Minitab. In: *Introduction to Engineering Statistics and Lean Six Sigma* [Internet]. London: Springer London; 2019 [cited 2026 Aug 7]. p. 575–600. Available from: http://link.springer.com/10.1007/978-1-4471-7420-2_24 doi:10.1007/978-1-4471-7420-2_24
15. Vrushali N Dhande, M. Williams, R. Veera Ranjani, A.R. Sakthi, M. Sudha, M. Raveendran. Identification of superior haplotypes for CCD8 regulating tiller number and grain yield in rice. *Electron J Plant Breed.* 2025 Nov 26;14(1):148–53. doi:10.37992/2023.1401.003
16. K. Madhuri, S. Manonmani, T. Kalaimagal, R. Veera Ranjani, M. Raveendran. Haplotype analysis of SCM2 loci for identification of donor for lodging resistance in rice. *Electron J Plant Breed.* 2025 Dec 19;12(3):794–9. doi:10.37992/2021.1203.110
17. R. Shobica Priya, T. Kalaimagal, S. Rajeswari, R. Ajay Prasanth, M. Raveendran. Allele mining for the grain number gene An-1 in rice (*Oryza sativa* L.). *Electron J Plant Breed.* 2025 Dec 19;12(3):772–9. doi:10.37992/2021.1203.107
18. Abbai R, Singh VK, Nachimuthu VV, Sinha P, Selvaraj R, Vipparla AK, et al. Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains. *Plant Biotechnol J.* 2019 Aug;17(8):1612–22. doi:10.1111/pbi.13087
19. Singh P, Sundaram KT, Vinukonda VP, Venkateshwarlu C, Paul PJ, Pahi B, et al. Superior haplotypes of key drought-responsive genes reveal opportunities for the development of climate-resilient rice varieties. *Commun Biol.* 2024 Jan 12;7(1):89. doi:10.1038/s42003-024-05769-7
20. Wang Y, Wang X, Zhai L, Zafar S, Shen C, Zhu S, et al. A novel Effective Panicle Number per Plant 4 haplotype enhances grain yield by coordinating panicle number and grain number in rice. *Crop J.* 2024 Feb;12(1):202–12. doi:10.1016/j.cj.2023.11.003
21. Zhao Y, Gao J, Wang X, Rashid MAR, Wu Z, Ma Z, et al. Elite haplotype of STRONG1 enhances rice yield by improving lodging resistance, panicle and plant architecture. *Nat Commun.* 2025 Jul 1;16(1):5894. doi:10.1038/s41467-025-60604-9
22. Singh P, Pahi B, Sundaram KT, Gnanapragasam N, Habde SV, Kumar NV, et al. Superior haplotypes of genes associated with higher grain yield under reproductive stage drought stress in rice. Gonzalez F, editor. *J Exp Bot.* 2025 Nov 4;76(17):5086–101. doi:10.1093/jxb/eraf150
23. Williams M, Veera RR, Bharathi A, Sudha M, Raveendran M. Haplotype characterisation of the SMG11/D2 and GW7 genes reveals novel allelic variants for grain size improvement in rice. *Plant Sci Today.* 2026 Jun 17. doi:10.14719/pst.13971
24. Zuo J, Li J. Molecular Genetic Dissection of Quantitative Trait Loci Regulating Rice Grain Size. *Annu Rev Genet.* 2014 Nov 23;48(1):99–118. doi:10.1146/annurev-genet-120213-092138
25. Bevan MW, Uauy C, Wulff BBH, Zhou J, Krasileva K, Clark MD. Genomic innovation for crop improvement. *Nature.* 2017 Mar;543(7645):346–54. doi:10.1038/nature22011

26.Selvaraj R, Singh AK, Singh VK, Abbai R, Habde SV, Singh UM, et al. Superior haplotypes towards development of low glycemic index rice with preferred grain and cooking quality. Sci Rep. 2021 May 12;11(1):10082. doi:10.1038/s41598-021-87964-8

Statistics	Whole population
Range	3.61-42.39
Mean	18.57
Median	17.78
SE mean	0.53
CV(%)	40.84

Table 1. Descriptive statistics for the studied population

Source of variation	Degrees of freedom	Adjusted sum of squares	Adjusted mean square	F- value	P- value
Haplotype	1	253.6	253.65	4.75	0.031
Error	197	10527.0	53.44		
Total	198	10780.6			

Table 2. Analysis of variance for grain yield per plant across *OsPYL9* haplotypes

Difference of Levels	Difference of Means	SE of Difference	95% CI	T-Value	Adjusted P-Value
H1 - H2	-2.26	1.04	(-3.97,0.00)	-2.18	0.015
H2 - H1	2.26	1.04	(0.00,3.97)	2.18	0.015

Table 3. Hsu's simultaneous test for level means indicating significant difference between *OsPYL9* haplotypes in grain yield per plant

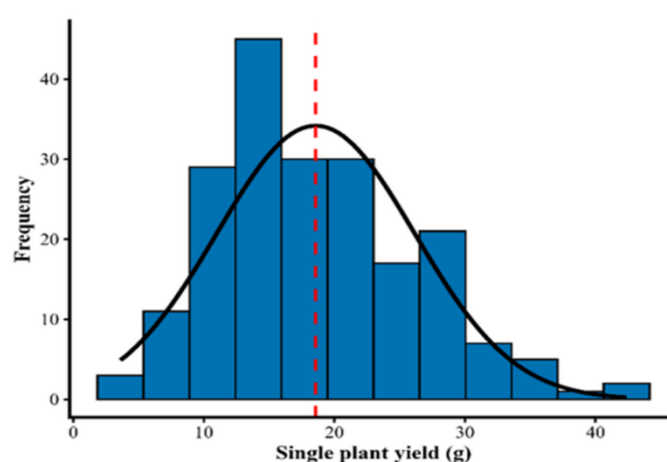


Fig. 1. Frequency distribution of grain yield per plant



Fig. 2. Formation of haplotype groups based on allelic variation

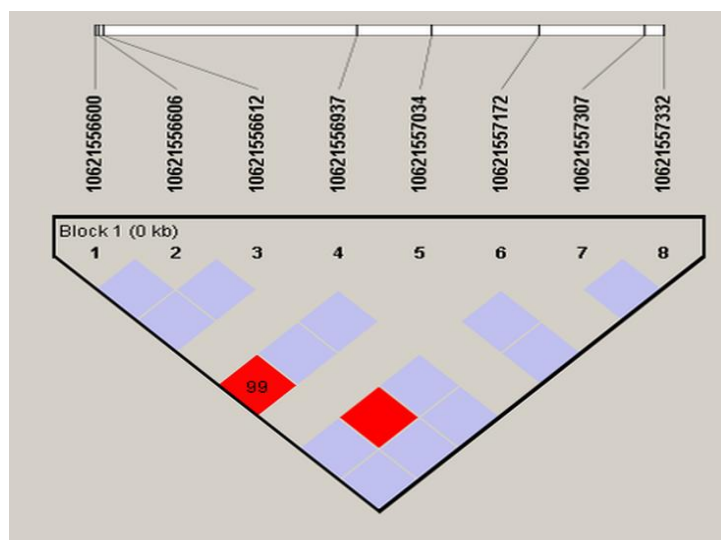


Fig. 3. Linkage Disequilibrium (LD) block of *OsPYL9* gene for 199 accessions. Values inside the block indicate LD D' .

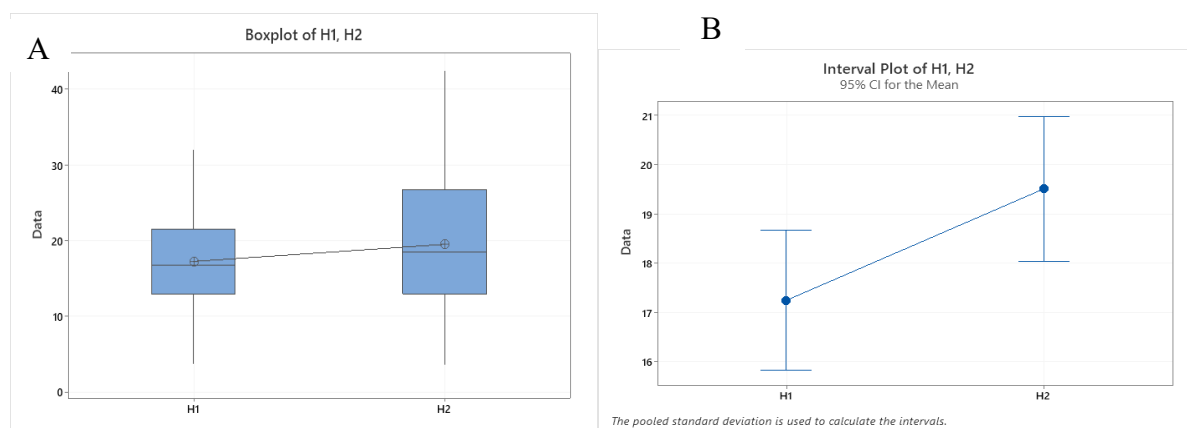


Fig. 4. Grain yield per plant (g) in the two *OsPYL9* haplotype groups. A) Box plot of grain yield per plant for haplotypes H1 and H2. B) Interval plot of means (95% CI) for grain yield per plant across haplotypes H1 and H2.