Haplotype Analysis of the Functionally characterized Genes Related to Yield in Rice

Vinay Kumar N.[1&2], Krishna Sundaram T.[2], Bandana P.[2] and Prasanna Rajesh A.[3*]
[1]Department of Genetics and Plant Breeding, Agricultural College, ANGRAU, Bapatla (Andhra Pradesh), India.
[2]International Rice Research Institute, Hyderabad (Telangana), India.
[3]Department of Genetics and Plant Breeding, RARS, ANGRAU, Kadiri (Andhra Pradesh), India.

(Corresponding author: Prasanna Rajesh A.*)

ABSTRACT: Rice is a staple food crop for billions of people and it meets around 20% calorie intake of the human population. Increasing global population, reducing arable land and climate change are limiting the rice production and productivity. In the era of narrowing genetic diversity and increased phenotypic similarity of cultivated germplasm, genomics data are indispensable to tap the genetic potential of the cache germplasm and utilize for crop improvement. In this context, 3K RGP data serves as a mine to dig out the genetic variation of the potential genes. Although several variants of biotic and abiotic stress tolerant genes are meticulously examined that are resistant to various races of pests, variants in yield related genes are disregarded. In this current experiment we have analyzed the haplotypic variants and identified the superior haplotypes for the functionally characterized genes which boost the yield. For the panicle architecture genes *OsFIB*, and *OsGA20ox,* H2 and H1 were identified as the superior haplotypes with the mean panicle length of 21.31 cm and 21.44 cm. Similarly for the single plant yield *DLT*- H8 and *DTH2*-H8 were found to be superior with a mean phenotypic value of 20.02 g and 17.51 g. These superior haplotypes can be introduced into the modern varieties to increase the genetic advance of the yield traits.

Keywords: Haplotype based breeding, Rice, Yield, Functionally characterized genes.

INTRODUCTION

The majority of people on the planet are fed by rice (*Oryza sativa* L.), one of the most significant cereal crops. Most Asian nations use rice as a staple diet, while African nations are observing a sharp increase in rice consumption. Crop yields are in jeopardy of becoming stable due to climate change, urbanization, and industrialization, as well as the ongoing growth in the human population. However, by 2050, the rates at which the major crops' worldwide yields are increasing will not be enough to feed the world's expected nine billion people (Ray *et al.*, 2013). One of the most crucial objectives of rice genetic research and breeding programs is the significant development of genetic yield potential per unit area.

Productivity in rice grains is a multifaceted attribute. Many additional characteristics, like heading date, plant architecture, and level of resilience to biotic and abiotic stresses, all have an impact on it, either directly or indirectly. Grain size, panicle architecture & length, grain numbers per panicle (GNPP), and tiller number per plant are the four main features that are directly linked to rice grain productivity.

Ever since genome-wide association studies were introduced, numerous genes and qtl regions controlling the yield-related traits were reported. However, the functional validation and real-world use of the discovered genes, are behind schedule. There are more than 400 genes reported to be functionally characterized for yield-related traits (Huang *et al.,* 2022).

DWARF AND LOW-TILLERING (*DLT*) is a positive regulator that mediates several BR responses in rice. *DLT* interacts with OSH15, which functions in the lower internodes, to modulate rice internode elongation *via* orchestrating BR signalling and metabolism (Li *et al.*, 2010). The gene, *DTH2* encodes a CONSTANS-like protein that promotes heading by inducing the florigen genes *Heading date 3a* and *RICE FLOWERING LOCUS T 1*, and it acts independently of the known floral integrators *Heading date 1* and *Early heading date 1*. In addition, *DTH2* served as a potential yield gene, and its deficiency reduced the grain number of the main panicle in plants (Liu *et al.*, 2016; Wu *et al.*, 2013). Similarly, *FIB* encodes an orthologue of *TAA* genes and loss of *FIB* function resulted in pleiotropic abnormal phenotypes, such as small leaves with large lamina joint angles, abnormal vascular development, small panicles, abnormal organ identity and defects in root development, together with a reduction in internal IAA levels. *FIB* plays a pivotal role in IAA biosynthesis in rice and that auxin biosynthesis, transport and sensitivity are closely interrelated (Guo *et al.*, 2020; Yoshikawa *et al.*, 2014). The gene *GNP1* had the pleiotropic effect of underlying grain number per panicle on sink, source and slow in Rice. *GNP1* encoding gibberellin biosynthesis gene *GA20ox1* affects grain number per panicle (GNP) in rice, thus increasing grain yield. To clarify the GNP1 effect on sink, source and flow in regulating rice grain yield, the researchers compared Lemont, a japonica (*geng*) cultivar, with its near-isogenic line (NIL-*GNP1* TQ) in

Lemont background with introgression of the allele at *GNP1* from Teqing, a high-yielding indica *(xian)* cultivar. The results indicated that significantly increased GNP from introgression of *GNP1* TQ into Lemont did not highly significantly improve grain yield of NIL-*GNP1* TQ as expected, due primarily to significant low sink activities and possible insufficient source supply which did not fully meet the increased sink capacity (Wu *et al.*, 2016; Zhai *et al.*, 2020).

Introgression of these genes into the modern cultivars would be remunerative for crop improvement. The marker-assisted breeding was a predominant tool used to accomplish the task. However, the MAB alone is not enough to improve the complex polygenic yield traits. Due to the low density, low coverage, and high cost of marker systems like simple sequence repeats (SSRs), gene mapping resolution is weak and plant selection and breeding efficiency are low. Similarly, The bi-allelic nature of single nucleotide polymorphisms (SNPs) and their potential for rare alleles and high levels of linkage drag make it difficult to pinpoint the precise loci responsible for important traits (Qian *et al.*, 2017). Haplotypes improve the accuracy of complex trait predictions in genomic selection models by better capturing local high-order allelic interactions, LD, and genomic similarity. Haplotype-based breeding (HBB) enhances prediction capabilities in populations with prominent linkage disequilibrium (LD) structures by more accurately identifying and using genomic areas associated with particular traits (Bhat *et al.*, 2021).

MATERIALS AND METHODS

The list of functionally characterized genes related to yield were downloaded from the Funrice genes (Huang *et al.*, 2022). Among the genes for yield related traits, the genes which were affecting more than one yield-related trait were considered for the haplotype analysis. Finally, four genes, *DLT*, *DHT2*, *FIB* & *GNP1* were selected for the haplotyping Table 1. Throughout the entire investigation, we used the '3k filtered' SNP set that is available in the SNP seek database. The following filtering criteria were used to extract the filtered from the Base SNP set: According to (Mansueto *et al.*, 2017) (http://snp-seek.irri.org/_download.zul) the alternative allele frequency must be at least 0.01 and the proportion of missed calls per SNP must not exceed 0.2. This SNP set was already accessible in the SNP seek database, which was used directly in this investigation. All of the genes' haplotype analyses have solely taken into account nonsynonymous SNPs *i.e.,* SNPs found in the exon region that alter the amino acid composition. The phenotypic data for the panicle length and single plant yield were extracted from the earlier data of 3K subset panel consisting 520 genotypes which were evaluated at IRRI, Hyderabad in 2019. Haplotype analysis was done using Duncan's multiple range test. The haplotype and haplo-pheno analysis were done using in house R script.

RESULTS AND DISCUSSION

Haplotype analysis of the candidate genes. *OsFIB* and *OsGA20ox1* were the two genes for the panicle architectural trait whose haplotypic variations were examined. In the same way, two genes *DLT; OsGRAS-32 and DTH2; OsCOL9* were taken into account for the trait single plant yield.

OsFIB. Four haplotypes were found throughout the 3K panel and the subset panel according to haplotype analysis of the *OsFIB* gene using the three SNPs (Table 2). Haplotype H3 was present in a large number of genotypes in both panels, occurring in 309 genotypes in the subset panel and in 1960 genotypes in the 3K panel. In the 3K panel, the frequency of the H8 was 0.70, while in the subset panel, it was 0.65. Haplotype H4 was also rare in the 3K and subset panels, occurring in just 41 and 4 genotypes, respectively, at frequencies of 0.01 and 0.008. There was a significant variation in the panicle length between the haplotype groups (Fig. 1). However, because of its peak mean panicle length of 21.31 cm, H2(GTA) is considered the superior haplotype. This haplotype, with frequencies between 0.18 and 0.27, was present in 505 genotypes of the 3K and 125 genotypes of the subset panels.

OsGA20ox1. The whole 3K and the subset panels were sorted into two groups i.e., haplotypes H1 and H2 based on the haplotype analysis with a single non-synonymous SNP (Table 3). Haplotype H2 is inordinately present in the 3K and the subset panels and represented by 1669 and 255 genotypes at a frequency of 0.58 and 0.51 respectively Haplotype H1 existed in 1229 and 241 genotypes with a frequency of 0.42 and 0.49 in the 3K and the subset panels respectively. Haplotype H1 existed in 1229 and 241 genotypes with a frequency of 0.42 and 0.49 in the 3K and the subset panels respectively. There were noticeable variations in the panicle length between the two haplotypic groups, H1 and H2 (Fig. 2). H1 has the highest mean value of 21.44 cm it was deemed superior to H2.

DLT. The haplotype analysis of this gene using six non-synonymous SNPs produced nine haplotypes in the 3K and the subset panels (Table 4). Haplotype H1 was substantially dispersed in the 3K and subset panels. It was discovered in 1359 and 233 genotypes in 3K and subset panels, with a frequency of 0.57 and 0.53 respectively. Haplotypes H2 and H9 in the 3K panel and H4 in the subset panel showed the lowest frequencies. With frequencies of 0.0004 and 0.01 respectively, these haplotypes were restricted to one and five genotypes in the relevant panels. Significant differences were observed between the haplotypic groups in the Duncan analysis (Fig. 3). Haplotype H8 (_GCCC) was found to be superior, yielding an average of 20.02 g per plant. The superior haplotype, which was represented by 34 and 19 genotypes, had frequencies of 0.01 and 0.04 in the 3K population and the subset populations, respectively.

DTH2. Haplotype analysis of the *DTH2* gene with the seven SNPs revealed the presence of 19 haplotypes across the 3K panel and 15 in the subset panel (Table 5). Haplotype H1 was widespread across the genotypes in both the panels with its occurrence in 878 and 137 genotypes in 3K and subset panels respectively. The frequency of the H1 was 0.31 in the 3K panel and 0.28 in the subset panel. Among the 3K panel, haplotype

H19 was scanty and limited to only one genotype with a frequency of 0.0003 and within the subset panel two haplotypes (H9 & H18) were recorded as rare with a frequency of 0.0002 and represented by single genotype each. Remarkable differences were observed among the haplotype groups for the single plant yield (Fig. 4). Nevertheless, H8 (GTCAG__) is regarded as the superior haplotype as it had the peak mean value of 17.57 for single plant yield. This haplotype was existed in 77 genotypes of the 3K and 18 genotypes of the subset panels with the frequencies of 0.02 and 0.03 respectively.

Haplotype diversity of the genes in study. Among the four genes analyzed for panicle architecture and plant yield, SNPs for the haplotype analysis ranged from one in the gene *OsGA20ox1* and seven in *DTH2* (Fig. 5). The haplotypes ranged from 2 to 19 in the 3K panel and 0 to 15 in the subset panel. Similar results were observed in the several haplotype analysis studies. For instance, 2 to 15 haplotypes were reported from the haplotype analysis of non- synonymous SNPs of the 120 genes related to yield and quality in rice (Abbai *et al.,* 2019), 2 to 7 haplotypes were reported by (Contreras-Soto *et al.,* 2017) & 3 to 9 haplotypes were observed by Bhat *et al.* (2022) for seed yield related genomic regions in soybean.

Among the four haplotypes of the *OsFIB* gene haplotype H3 was present in more than 1900 & 300 genotypes in 3K and subset panels respectively and recorded a frequency of more than 0.65 in both the panels. *OsGA20ox1*-H2 was only present in the 3K panel and absent from the subset panel and recorded a frequency of 0.58 and present across 1669 genotypes in the 3K panel. *DLT*-H1was widespread across the 3K and the subset panels, occurring in 1359 and 233 genotypes respectively. The frequency of this gene was more than 0.53 in both the panels. Interestingly for the gene *DTH2* haplotype H1 was present in higher frequencies in the 3K panel with a frequency of 0.31 and represented by around 878 genotypes and *DTH2*-H6 is represented by larger population of 137 genotypes and maintained a frequency of 0.28 (Table 6 and Fig. 6).

The lowest frequency haplotypes were represented by a range of 1 to 1229 genotypes in the 3K panel and 1 to 5 genotypes in the subset panel (Table 7). The frequencies of the lowest haplotype which is represented by one genotype were 0.0003 and 0.002. Gene *OsGA20ox1*- Haplotype H1 was present in 1299 and 241 genotypes of 3K and the subset panels respectively (Fig. 7). Similarly, H2 & H9 of *DLT*, and H19 of *DHT 2*genes were represented by single genotype each in the 3K panicle.

Superior haplotypes of the genes. The superior haplotypes were defined for all the four genes analyzed. For the panicle architecture traits *OsFIB*-H2 and *OsGA20ox1*-H1were recorded as the superior haplotypes with the mean panicle length of 21.31 cm and 21.44 cm respectively. The frequencies of the *OsFIB*-H2 gene were 0.18 and 0.26 in the 3K and the subset panels respectively. Superior haplotype H1 of *OsGA20ox1* was only present in the 3K panel. *DTL* H8 and *DHT2* emerged as the superior haplotype with the single plant yield of 20.02 g and 17.57 g. The superior haplotypes of the genes and their frequencies were given in the Table 8 and Fig. 8.

Even though the mean trait values of the haplotypes seem to be low for both the traits when compared to the current trend of phenotypic trait values of the existing elite lines in the breeding pipelines, it is possible to realize the expected outcome if the superior haplotypes from the best donors exhibiting the superior trait value can be introgressed into the suitable elite backgrounds. As the actual problem lies with the haplotype interactions with the host background, care should be taken while accessing the performance of superior haplotypes in different genetic backgrounds. For instance, though the mean single plant yield of the *DTL*-H8 was 20.02, the actual single plant yield of high yielding lines was about 30.9 g in the line NASOUAN (IRGC_11889-1) and 29.9 g in PERUNEL (IRGC_63113-1). Similarly, MUTA GANJE (IRGC 26744-1) evinced the higher panicle length of 27.66 cm and AUS 278 (IRGC 29068-1) exhibited 27.33 cm for the superior haplotype H2 of the *FIB* gene among the 3K panel.
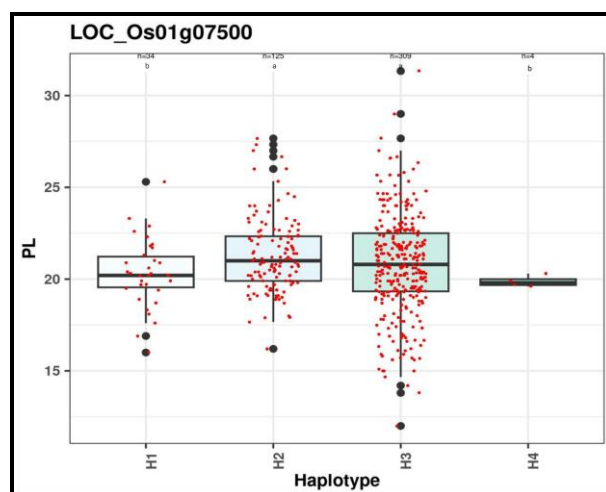

Fig. 1. Box plots representing the panicle length (PL in cm) distribution of the genotypes carrying haplotypic variants of the gene *OsFIB*.
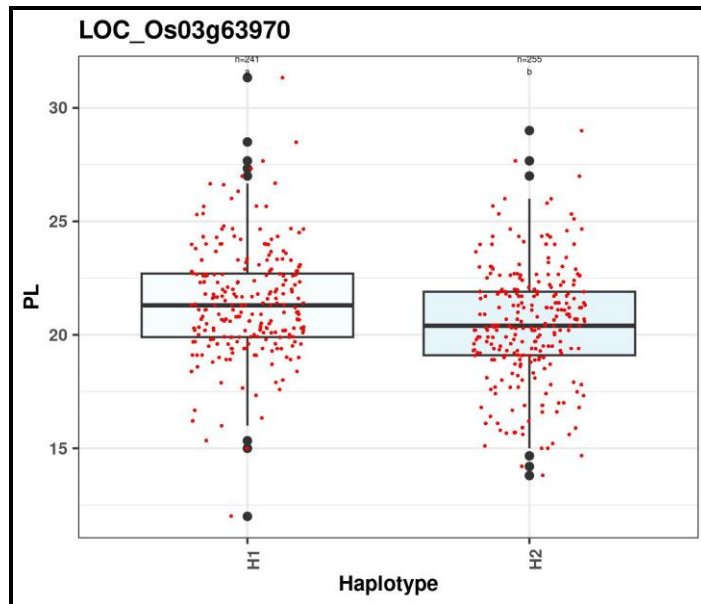
Fig. 2. Box plots representing the panicle length (PL in cm) distribution of the genotypes carrying haplotypic variants of the gene *OsGA20ox1*.
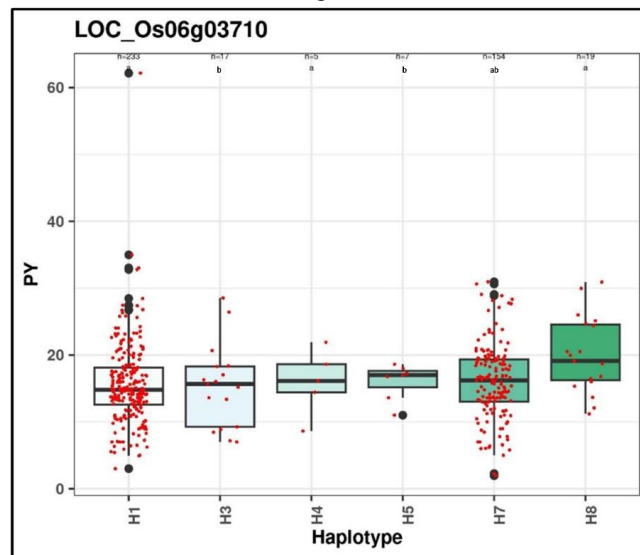

Fig. 3. Box plots representing the plant yield (PY in grams) distribution of the genotypes carrying haplotypic variants of the gene *DLT*.
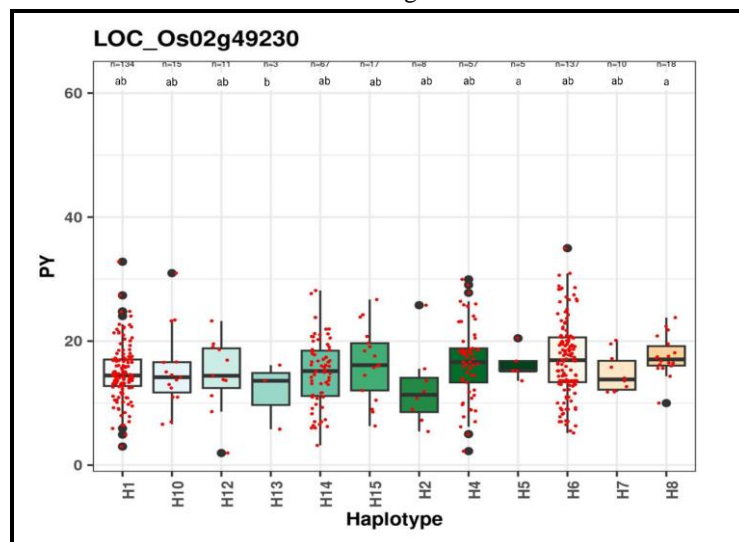

Fig. 4. Box plots representing the plant yield (PY in grams) distribution of the genotypes carrying haplotypic variants of the gene*DTH2*.
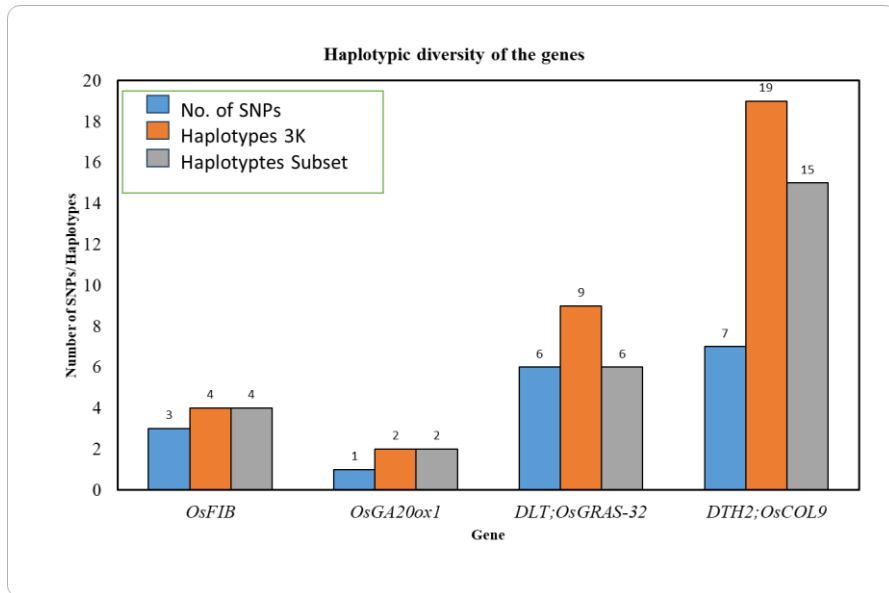
Fig. 5. Haplotype diversity of the 3K and subset panels for the four genes related to plant yield traits.
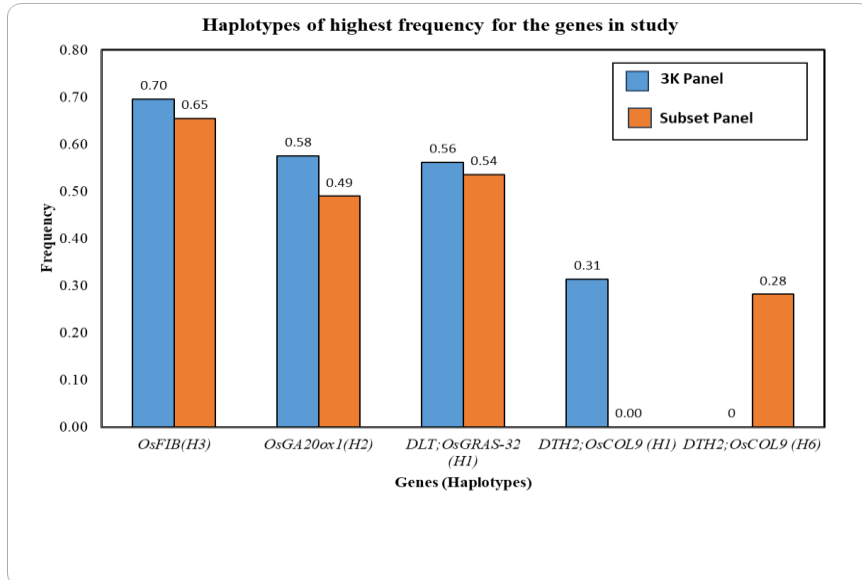

Fig. 6. Bar graph depicting the highest haplotype frequencies of the yield related genes.
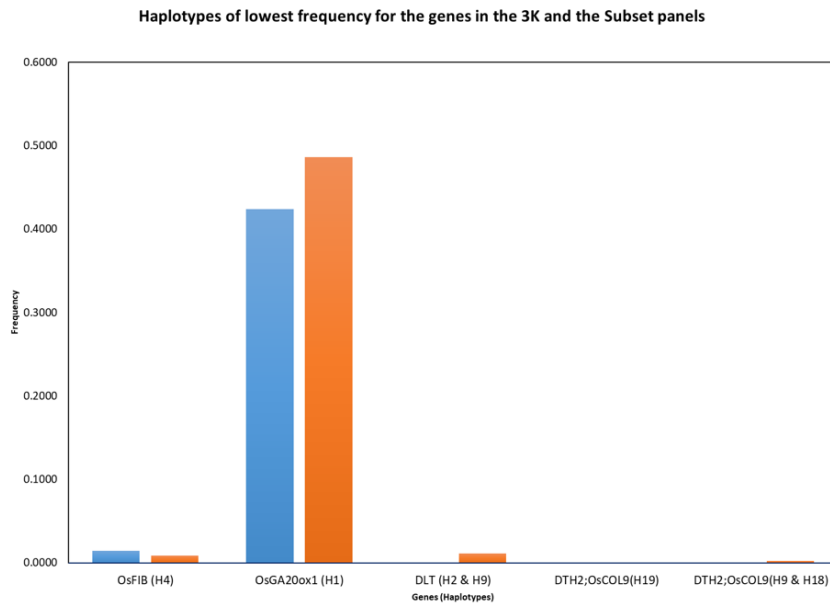

Fig. 7. Bar graph depicting the lowest haplotype frequencies of the yield-related genes in the subset panel.
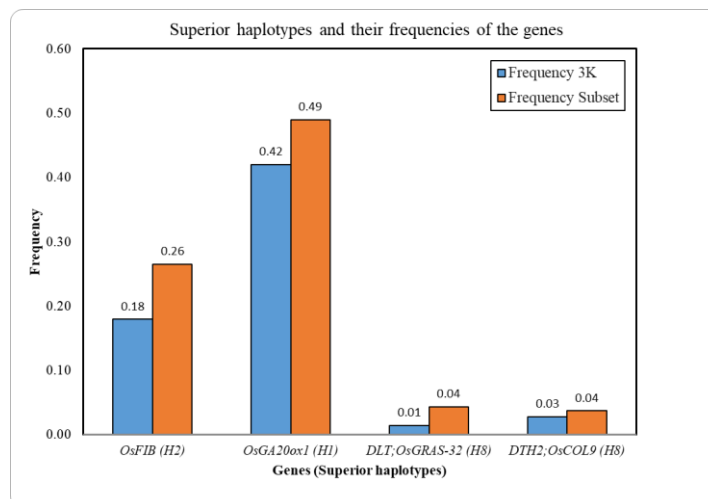
Fig. 8. Bar diagram representing superior haplotypes of the yield related genes and their frequency in the 3K and subset panels.

Table 1: Details of the genes selected for the haplotype analysis.

Sr. No.	Gene Symbol	Gene Name	Locus ID	Description
1.	DLT	DWARF AND LOW-TILLERING	LOC_Os06g03710	*DWARF AND LOW-TILLERING* (*DLT*) is a positive regulator that mediates several BR responses in rice. *DLT* interacts with OSH15, which functions in the lower internodes, to modulate rice internode elongation *via* orchestrating BR signalling and metabolism
2.	DTH2	CCT domain-containing gene 8, CCT (CO, CO-LIKE and TOC1) domain protein 8 CCT domain protein 8	LOC_Os02g49230	*DTH2* encodes a CONSTANS-like protein that promotes heading by inducing the florigen genes *Heading date 3a* and *RICE FLOWERING LOCUS T 1*, and it acts independently of the known floral integrators *Heading date 1* and *Early heading date 1*.
3.	FIB	FISH BONE	LOC_Os01g07500	*FIB* plays a pivotal role in IAA biosynthesis in rice and that auxin biosynthesis, transport and sensitivity are closely interrelated
4.	GNP1	GRAIN NUMBER PER PANICLE 1	LOC_Os03g63970	The gene *GNP1* had the pleiotropic effect of underlying grain number per panicle on sink, source and slow in Rice. *GNP1* encoding gibberellin biosynthesis gene *GA20ox1* affects grain number per panicle (GNP) in rice, thus increasing grain yield.

Table 2: Haplotypic diversity of the *OsFIB* gene in the 3K and the subset panels.

Sr. No.	Haplotype	SNP ID			3K Panel		Subset Panel	
		3578550	3578965	3581757	Genotypes	Frequency	Genotypes	Frequency
1.	H1	G	C	C	311	0.110	34	0.072
2.	H2	G	T	A	505	0.179	125	0.265
3.	H3	G	T	C	1960	0.696	309	0.655
4.	H4	T	T	C	41	0.015	4	0.008

Table 3: Haplotypic diversity of the *OsGA20ox1* gene in the 3K panel.

Sr. No.	Haplotype	SNP	3K Panel		Subset Panel	
			Genotypes	Frequency	Genotypes	Frequency
1	H1	C	1229	0.42	241	0.49
2	H2	T	1669	0.58	255	0.51

Table 4: Haplotypic diversity of the *DLT* gene in the3K and the subset panels.

Sr. No.	Haplotype	SNP ID					3K Panel		Subset Panel	
		1466088	1466091	1466206	1466715	1467723	Genotypes	Frequency	Genotypes	Frequency
1.	H1	-	C	-	C	C	1359	0.56	233	0.54
2.	H2	-	C	-	C	T	1	0.00	NA	NA
3.	H3	-	C	-	T	C	214	0.09	17	0.04
4.	H4	-	C	-	T	T	25	0.01	5	0.01
5.	H5	-	C	C	C	C	56	0.02	7	0.02
6.	H6	-	C	C	T	C	13	0.01	NA	NA
7.	H7	-	G	-	C	C	718	0.30	154	0.35
8.	H8	-	G	C	C	C	34	0.01	19	0.04
9.	H9	G	C	-	C	C	1	0.00	NA	NA

Table 5: Haplotypic diversity of the *DTH2* gene in the3K and the subset panels.

Sr. No.	Haplotype	SNP ID							3K Panel		Subset Panel	
		30096330	30096766	30097391	30099368	30099425	30099467	30099617	Genotypes	Frequency	Genotypes	Frequency
1.	H1	A	C	C	G	-	C	-	878	0.31	134	0.28
2.	H2	A	C	C	G	-	C	G	95	0.03	8	0.02
3.	H3	G	C	C	G	-	C	-	2	0.00	NA	NA
4.	H4	G	T	C	A	-	-	-	330	0.12	57	0.12
5.	H5	G	T	C	A	-	-	G	33	0.01	5	0.01
6.	H6	G	T	C	A	-	C	-	717	0.26	137	0.28
7.	H7	G	T	C	A	-	C	G	100	0.04	10	0.02
8.	H8	G	T	C	A	G	-	-	77	0.03	18	0.04
9.	H9	G	T	C	A	G	-	G	3	0.00	1	0.00
10.	H10	G	T	C	A	G	C	-	63	0.02	15	0.03
11.	H11	G	T	C	A	G	C	G	4	0.00	NA	NA
12.	H12	G	T	C	G	-	-	-	63	0.02	11	0.02
13.	H13	G	T	C	G	-	-	G	12	0.00	3	0.01
14.	H14	G	T	C	G	-	C	-	312	0.11	67	0.14
15.	H15	G	T	C	G	-	C	G	71	0.03	17	0.03
16.	H16	G	T	T	G	-	-	-	15	0.01	2	0.00
17.	H17	G	T	T	G	-	-	G	2	0.00	NA	NA
18.	H18	G	T	T	G	-	C	-	14	0.01	1	0.00
19.	H19	G	T	T	G	-	C	G	1	0.00	NA	NA

Table 6: Haplotypes with the highest frequencies in the 3K and subset panels for the yield related genes

Sr. No.	Gene	No. of genotypes		Frequency	
		3K	Subset	3K	Subset
1.	*OsFIB(H3)*	1960	309	0.70	0.65
2.	*OsGA20ox1(H2)*	1669	NA	0.58	NA
3.	*DLT (H1)*	1359	233	0.56	0.54
4.	*DTH2 (H1)*	878	-	0.31	-
5.	*DTH2 (H6)*	-	137	-	0.28

Table 7: Haplotypes with the lowest frequencies in the 3K and subset panels for the yield related genes.

Sr. No.	Gene (Haplotype)		Genotypes		Frequency	
	3K	Subset	3K	Subset	3K	Subset
1.	*OsFIB (H4)*	*OsFIB (H4)*	41	4	0.0146	0.0085
2.	*OsGA20ox1 (H1)*	*OsGA20ox1 (H1)*	1229	330	0.4241	0.7512
3.	*DLT (H2 & H9)*	*DLT (H4)*	1	5	0.0003	0.0115
4.	*DTH2 (H19)*	*DTH2 (H9 & H18)*	1	1	0.0003	0.0021

Table 8: Superior haplotypes of the genes and their distribution in the 3K and subset panels.

Gene	No. of SNPs	Haplotypes 3K	Haplotypes Subset	Superior haplotype	Mean Trait Value of S.H	S.H. Sequence	Genotypes		Frequency	
							3K	Subset	3K	Subset
OsFIB	3	4	4	H2	21.31 cm	GTA	505	125	0.18	0.26
OsGA20ox1	1	2	2	H1	21.44 cm	C	1229	330	0.42	0.75
DLT	6	9	6	H8	20.02 g	_GCCC	34	19	0.01	0.04
DTH2	7	19	15	H8	17.57 g	GTCAG__	77	18	0.03	0.04

CONCLUSIONS

In the current study the haplotypes were analyzed for the four yield enhancing functionally characterized genes. The integration of identified superior haplotypes into current breeding programs by introgression into the suitable elite genetic backgrounds through the superior haplotype based breeding strategy holds a considerable importance for improvement of the complex yield traits.

REFERENCES

Abbai, R., Singh, V. K., Nachimuthu, V. V., Sinha, P., Selvaraj, R., Vipparla, A. K., Singh, A. K., Singh, U. M., Varshney, R. K., and Kumar, A. (2019). 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 Biotechnology Journal, 17*(8), 1612–1622.

Bhat, J. A., Adeboye, K. A., Ganie, S. A., Barmukh, R., Hu, D., Varshney, R. K., and Yu, D. (2022). Genome-wide association study, haplotype analysis, and genomic prediction reveal the genetic basis of yield-related traits in soybean (*Glycine max* L.). *Frontiers in Genetics, 13,* 953833.

Bhat, J. A., Yu, D., Bohra, A., Ganie, S. A., and Varshney, R. K. (2021). Features and applications of haplotypes in crop breeding. *Communications Biology, 4*(1), 1266.

Contreras-Soto, R. I., Mora, F., de Oliveira, M. A. R., Higashi, W., Scapim, C. A., and Schuster, I. (2017). A Genome-Wide Association Study for Agronomic Traits in Soybean using SNP Markers and SNP-Based Haplotype Analysis. *PLOS ONE, 12*(2), e0171105.

Guo, T., Chen, K., Dong, N., Ye, W., Shan, J., and Lin, H. (2020). Tillering and small grain 1 dominates the tryptophan aminotransferase family required for local auxin biosynthesis in rice. *Journal of Integrative Plant Biology, 62*(5), 581–600.

Huang, F., Jiang, Y., Chen, T., Li, H., Fu, M., Wang, Y., Xu, Y., Li, Y., Zhou, Z., Jia, L., Ouyang, Y., and Yao, W. (2022). New Data and New Features of the Fun Rice Genes (Functionally Characterized Rice Genes) Database: 2021 Update. *Rice, 15*(1), 23.

Li, W., Wu, J., Weng, S., Zhang, Y., Zhang, D., and Shi, C. (2010). Identification and characterization of dwarf 62, a loss-of-function mutation in DLT/OsGRAS-32 affecting gibberellin metabolism in rice. *Planta, 232*(6), 1383–1396.

Liu, H., Gu, F., Dong, S., Liu, W., Wang, H., Chen, Z., and Wang, J. (2016). CONSTANS-like 9 (COL9) delays the flowering time in *Oryza sativa* by repressing the Ehd1 pathway. *Biochemical and Biophysical Research Communications, 479*(2), 173–178.

Mansueto, L., Fuentes, R. R., Borja, F. N., Detras, J., Abriol-Santos, J. M., Chebotarov, D., Sanciangco, M., Palis, K., Copetti, D., Poliakov, A., Dubchak, I., Solovyev, V., Wing, R. A., Hamilton, R. S., Mauleon, R., McNally, K. L., and Alexandrov, N. (2017). Rice SNP-seek database update: New SNPs, indels, and queries. *Nucleic Acids Research, 45*(D1), D1075–D1081.

Qian, L., Hickey, L. T., Stahl, A., Werner, C. R., Hayes, B., Snowdon, R. J., and Voss-Fels, K. P. (2017). Exploring and Harnessing Haplotype Diversity to Improve Yield Stability in Crops. *Frontiers in Plant Science, 8, 1534.*

Ray, D. K., Mueller, N. D., West, P. C., and Foley, J. A. (2013). Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLOS ONE, 8*(6), e66428.

Wu, W., Zheng, M., Lu, G., Zhong, Z., Gao, H., Chen, L., Wu, C., Wang, H., Wang, Q., Zhou, K., Wang, L., Wu, F., Zhang, X., Guo, X., Cheng, Z., Lei, C., Lin, Q., Jiang, L., Wang, H., Wan, J. (2013). Association of functional nucleotide polymorphisms at *DTH2* with the northward expansion of rice cultivation in Asia. *Proceedings of the National Academy of Sciences, 110*(8), 2775–2780.

Wu, Y., Wang, Y., Mi, F., Shan, X., Li, M., Xu, L., and Lin, X. (2016). The QTL GNP1 Encodes GA20ox1, Which Increases Grain Number and Yield by Increasing Cytokinin Activity in Rice Panicle Meristems. *PLOS Genetics, 12*(10), e1006386.

Yoshikawa, T., Ito, M., Sumikura, T., Nakayama, A., Nishimura, T., Kitano, H., Yamaguchi, I., Koshiba, T., Hibara, K., Nagato, Y., and Itoh, J. (2014). The rice FISH BONE gene encodes a tryptophan aminotransferase, which affects pleiotropic auxin-related processes. *The Plant Journal, 78*(6), 927–936.

Zhai, L., Wang, F., Yan, A., Liang, C., Wang, S., Wang, Y., and Xu, J. (2020). Pleiotropic Effect of GNP1 Underlying Grain Number per Panicle on Sink, Source and Flow in Rice. *Frontiers in Plant Science, 11,* 933.