



Wild species *Oryza rufipogon* as a potential donor for development of interspecific backcross population in the background of Samba Mahsuri (BPT5204) and detection of yield enhancing QTLs

M. Sridhar^{1,2} · C. Gireesh^{1,3} · Divya Balakrishnan¹ · M. Sujatha² · M. S. Anantha¹ · C. H. Raveendra¹ · P. Senguttuvel¹ · D. Srinivasa Chary² · V. P. Bhadana^{1,4} · R. M. Sundaram¹

Received: 3 July 2025 / Accepted: 26 September 2025
© Indian Society for Plant Physiology 2025

Abstract

The present rice cultivars have reached a yield plateau primarily due to the restricted use of germplasm within *Oryza sativa* cultivars as parental lines in breeding programmes. Wild ancestors and related species comprise a vital source of genetic material to break the yield barrier. The present investigation was carried out for identification and introgression of yield enhancing QTLs by developing 190 BC₂F₂ backcross population derived from a cross between Samba Mahsuri*2/*O. rufipogon* (IRGC106106). Polymorphic SSR markers spanning the genome were used to genotype the mapping population to construct a linkage map of 1893 cM (Kosambi). The CIM identified a total of 47 significant QTLs at $P < 0.001$ (LOD > 2.50) for ten yield traits distributed across 11 chromosomes. Among these, *O. rufipogon* alleles contributed a beneficial effect on 22 QTLs. The cumulative phenotypic variance explained by donor-derived QTLs for target traits was ranging from 2.12 to 32.65% with a single major QTL *qPL9.1* with 29.14% PVE. Yield increase in introgression lines over Samba Mahsuri (BPT5204) was attributed by beneficial allele from *O. rufipogon* for the traits such as productive tillers, panicle length, test weight and grain yield per plant. Three major QTLs with more than 10% phenotypic variance observed. Significant genomic region *qPL9.1* was detected from *O. rufipogon* associated with panicle length, with 29.14% of observed phenotypic variance. *qDF1.1* and *qPH2.1* are the other major QTLs from recurrent parent Samba Mahsuri with 29.33% and 29.20% phenotypic variance respectively. There were six QTL hotspots viz., *qY1a*, *qY1b*, *qY2*, *qY3*, *qY7* and *qY9* for yield and related traits mapped in our study on 5 chromosomes where QTLs for different traits co-localized. The present study demonstrated that *O. rufipogon* introgressions and QTLs as potential sources associated with positive transgressive variation for enhancement of yield traits.

Keywords Rice · Yield · Wild species · Backcrossing · QTL mapping

Introduction

The Food and Agriculture Organization predicts that the global population will reach 9 billion by 2050, necessitating a 60% increase in food production (Alexandratos & Bruin-sma, 2012). The yield of rice must be increased in order to cater to the food security needs of a growing population. Rice (*Oryza sativa* L.), an essential staple food sustaining half of the world's population, is therefore in much demand. Rice is much more than a grain; it has fuelled economies, been subject to science, and travelled the globe due to world trade. The slogan “*Rice is life*” is highly suitable for the Indian subcontinent, where it provides livelihood to millions of rural households and demonstrates its importance

✉ C. Gireesh
giri09@gmail.com

✉ Divya Balakrishnan
dbiirr23@gmail.com

¹ ICAR-Indian Institute of Rice Research, Hyderabad, Telangana 500030, India

² Department of Genetics and Plant Breeding, PJTSAU, Hyderabad 500030, India

³ ICAR-Indian Institute of Seed Science, Regional Station, Bengaluru, Karnataka 560065, India

⁴ ICAR-Indian Institute of Agricultural Biotechnology, Ranchi, Jharkhand 834010, India

in food and nutritional security. Globally, rice accounts for a large portion of Asian and Latin American diets, forms the basis of agriculture in these regions, and connects to the rest of the world through global markets. It is also a major contributor of foreign exchange and government revenue in many countries of Asia and Africa. The majority of rice production takes place in South and East Asia, contributing 90% of the world's total output. India is the 2nd largest producer after China, with approximately 42.95 million hectares under cultivation, producing 1.11 billion tonnes annually with a productivity of 2494 kg ha⁻¹ with Chhattisgarh, West Bengal, Uttar Pradesh, Punjab, Telangana, Tamil Nadu, and Andhra Pradesh as major rice growing states. In Telangana, rice occupies 1.72 million hectares with a productivity of 3075 kg ha⁻¹ (Indiastat, 2017–18).

Wild species of rice possess tremendous genetic variability for agronomically important traits and rice domestication resulted in the loss of many desirable traits which greatly reduced the genetic diversity (Balakrishnan et al., 2016). Several genes regulating key agronomic traits viz., days to heading, seed shattering, and seed dormancy are still retained in wild rice relatives. Advanced molecular biology studies have revealed that wild rice contains genes that have potential to improve yield contributing characters despite its inferior economic traits (Tian et al., 2006). Since the release of semi-dwarf varieties in the 1960s, yield potential of commercial *indica* cultivars has stagnated. Over the last few years, rice breeding has primarily focused on subduing the yield plateau, a repercussion of the narrow genetic base. Therefore, hybridizations with wild species and wild relatives have been performed to broaden the genetic base and increase yields.

Wild progenitor species, which are diverse and potential donor source in improving complex traits such as yield. The genus *Oryza* comprises 22 wild species and 2 cultivated species, with *O. rufipogon* Griff. the direct ancestor of *O. sativa* being genetically more diverse and a known source of alleles for yield and stress tolerance (Marri et al., 2005; Balakrishnan et al., 2016; Yadavalli et al., 2022; Chandu et al., 2024; Magudeeswari et al., 2024). Advanced backcross QTL (AB-QTL) analysis is a proven strategy to broaden the genetic base of elite cultivars while simultaneously identifying and transferring useful QTLs from native and untransformed genetic resources into modern cultivars (McCouch, 2008). Even though Indian rice germplasm is endowed with considerable genetic diversity, most of the agronomically important traits, including yield, have been genetically narrowed in the elite cultivars. Intensive modern breeding efforts, by selecting favourable alleles from early domesticates have further narrowed the rice gene pool by concentrating only the favourable alleles.

Rice yield and its component traits are governed by a complex network of integrated physiological processes, encompassing photosynthetic assimilation, partitioning, utilization, panicle exertion, spikelet fertility, grain filling and flowering synchrony. Quantitative trait loci (QTLs) associated with traits such as panicle length, tiller number, and spikelet fertility frequently employ their effects by modifying key physiological determinants demonstrated by enhancing light interception, improving source–sink balance, or increasing pollen viability and fertilization efficiency. The introgression of beneficial alleles from *Oryza rufipogon* into elite cultivars like Samba Mahsuri serves to broaden the genetic base and facilitates the dissection of physiological mechanisms underpinning yield enhancement. By establishing linkages between mapped QTLs and their physiological functions, this study provides insights pertinent to both plant breeding and physiological research.

In the rice crop, the grain yield potential is majorly affected by four factors viz., grain weight, grain number per panicle, filled grain percentage, and productive tiller number. These complex and quantitative characteristics have a pronounced influence on crop yield and are often reliant on the genetic potential of rice. About 20 genes attributed to yield-related traits were detected through QTL fine mapping or positional cloning using rice mutants and employed through marker assisted breeding (Wang & Li, 2005; Xing & Zhang, 2010; Miura et al., 2011; Huang et al., 2013; Liu et al., 2015; Li et al., 2022; Sen et al., 2024).

Samba Mahsuri, a prominent south Indian rice variety, is renowned for its superior grain quality, fine texture, and low glycaemic index. Despite its premium quality and commercial potential, its average yield under optimal conditions is typically around 4.5–5.0 t ha⁻¹, which is lower than many modern high-yielding cultivars (Marri et al., 2005; Sundaram et al., 2008). The lower yield levels and stress tolerance remained as major constrain during its production. Hence it is essential to break the yield barriers persisted due to genetic constraints to unlock higher productivity. In addition, it is highly susceptible to major biotic stresses such as bacterial blight and blast, as well as abiotic constraints like salinity, which collectively limit its productivity and stability across diverse environments.

Common wild rice (*Oryza rufipogon* Griff.), the progenitor of cultivated rice, exhibits greater genetic diversity than *O. sativa* cultivars and serves as a reservoir of beneficial alleles for yield and stress tolerance (Marri et al., 2005). Introgressed *O. rufipogon* QTLs enhance elite cultivars, improving panicle length/branching (Li et al., 2002), grain number per panicle (Xiao et al., 1998), test weight/yield components (Septiningsih et al., 2003; Thomson et al., 2003), spikelet fertility (Marri et al., 2005), and biomass (Xiao et al., 1998). Although undesirable traits like

shattering, long duration, low harvest index exist, advanced backcrossing minimizes these while retaining advantageous loci. Thus, *O. rufipogon* is a strategic donor for broadening the genetic base of cultivars like Samba Mahsuri and capturing alleles underpinning yield potential and physiological resilience.

Therefore, the present study was undertaken to develop an advanced backcross (BC_2F_2) mapping population derived from crossing the elite indica cultivar Samba Mahsuri and its wild progenitor *O. rufipogon* with the objective of broadening the genetic base of elite cultivar. This study also aimed in identifying yield-enhancing QTLs to overcome existing yield barriers imparted during domestication by integrating phenotypic evaluation with genome-wide SSR marker analysis, to dissect the genetic architecture of key yield-contributing traits through the AB-QTL approach and to generate introgression lines that can serve as valuable pre-breeding material for improving productivity of elite rice varieties.

Materials and methods

Plant material and population development

The experimental material consisted of 190 BC_2F_2 introgression lines derived from the cross Samba Mahsuri*2/*O. rufipogon* (IRGC106106). Samba Mahsuri (BPT-5204) is a medium slender high yielding rice variety developed from GEB24/TN-1/Mahsuri through a pedigree breeding method (Reddy et al., 1979) and notified in 1986. It is suitable for irrigated ecology and very popular in southern parts of India due to its medium slender grain type, premium cooking and eating qualities. Hence, Samba Mahsuri was used as a recurrent parent in the present study. Wild species accession, *O. rufipogon* (IRGC106106) was used as a donor parent to develop an advanced backcross population.

Crosses were made between Samba Mahsuri and *O. rufipogon* (IRGC106106) during *Kharif*-2016. The clipping method of emasculation followed by pollination was carried out for the crossing. The F_1 seeds were harvested from the

female parent and subjected to hybrid confirmation through molecular markers (Fig. 1). The interspecific F_1 developed by crossing Samba Mahsuri (♀) with *O. rufipogon* (♂) showed pollen sterility. Therefore, F_1 's used as female parent and Samba Mahsuri as pollen donor in developing the BC_1F_1 population. During *Rabi*-2016, three plants of confirmed F_1 's were backcrossed to recurrent parent and developed 78 BC_1F_1 . Five BC_1F_1 plants, selected for desirable plant type, maturity and fertility, were backcrossed a second time to recurrent parent to generate 620 lines of BC_2F_1 seeds during *Kharif*-2017. The resulting BC_2F_1 plants were self-pollinated in *Rabi*-2017 under irrigated conditions to ensure survival during population development. The detailed flow-chart on development of ILs is illustrated in Fig. 2. During phenotypic selection in segregating population undesirable traits like sterility, delayed flowering, spreading plant type, excessive shattering, long awns and tall plants were eliminated. In total, 190 BC_2F_2 linkage population developed from Samba Mahsuri*2/*O. rufipogon* were evaluated for yield enhancing traits along with two checks namely Samba Mahsuri and RNR 15048 during *Kharif* 2018 at ICAR-IIRR farm, ICRISAT campus, Patancheru, Hyderabad.

Phenotyping of the mapping population

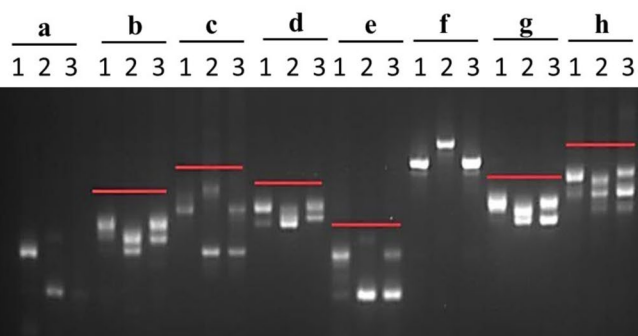
The experiment followed an Augmented Randomized Complete Block Design with five blocks, each containing 38 BC_2F_2 introgression lines (ILs) alongside two check varieties. Twenty one day old seedlings were transplanted in two rows, each consisting of 20 hills (3 m in length), at a spacing of 20 cm × 15 cm. Data were recorded for ten yield-related traits, including days to 50% flowering (DF), plant height (PH), panicle length (PL), total tillers per plant (TT), productive tillers per plant (PT), number of filled grains per panicle (FGP), number of chaffy grains per panicle (CGP), spikelet fertility percentage (SF), 1000-grain weight (TW), and single plant yield (GYP). The measurements for days to 50% flowering and 1000-grain weight were recorded on a plot basis, while the remaining traits were assessed from five plants chosen randomly within each plot. Spikelet fertility

Figure 1. Crossing and Hybrid confirmation of Samba Mahsuri/*Oryza rufipogon* (acc. IRGC 106106) through molecular markers. **A** Rice panicle showing the seed setting after pollination.

B Confirmation of F_1 hybrid using SSR markers. 1—Samba Mahsuri; 2—*O. rufipogon*; 3— F_1 hybrid. Marker order, a—RM 10033; b—RM 15630; c—RM 14250; d—RM 15251; e—RM 10053; f—RM 10126; g—RM 10009; h—RM 6759



(A)



(B)

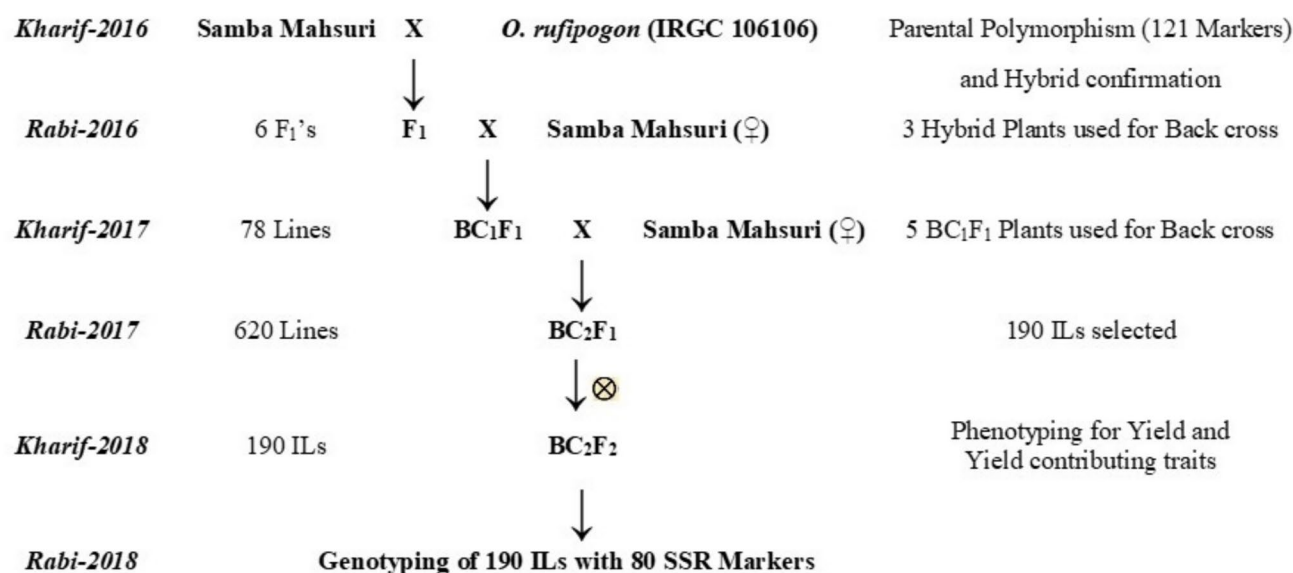


Figure 2. Schematic representation of development of advance backcross population derived from cross Samba Mahsuri/*O. rufipogon* (IRGC 106106). ILs—Introgression Lines

was determined as the percentage of filled grains relative to the total number of grains (filled and unfilled) per panicle.

Genotyping of mapping population

Total genomic DNA was isolated from 15 days old seedlings using 2% CTAB (Cetyl Try Methyl Ammonium Bromide) method (Doyle & Doyle, 1990). The genomic DNA quantity and quality were analyzed using 0.8% agarose gel using lambda (λ) Hind III DNA as standard. The DNA samples were normalized in T₁₀E₁ buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) to get a final concentration of 50 ng/ μ l for PCR amplification. In the present investigation, genome wide SSR (rice microsatellites) markers were employed for the parental polymorphism survey. The SSRs are co-dominant, multi-allelic, highly polymorphic, abundant with uniform distribution in the plant genome and availability of complete genome sequences for most of the crops. Hence they are considered ideal markers for genetic studies (McCouch et al., 1997). A total of 405 SSR markers distributed throughout 12 chromosomes of rice were used to detect polymorphism between Samba Mahsuri and *O. rufipogon* of which 121 (about 29.87%) were found polymorphic between parents. A marker was considered polymorphic if there were differences in the size of amplified products between the two parents. The list of parental polymorphic SSR markers along with their chromosome number, primer sequences, product size identified in the study, start point and end point in the genome is provided in Table S1. The polymorphic SSR primers were synthesized and supplied by IDT Avantor (100 pmol) and the working concentration of each primer was 5 μ M.

Polymerase Chain Reaction (PCR) amplification was conducted using Eppendorf Vapo Protect PCR cyclers at the Molecular Breeding Laboratory, IIRR, Hyderabad using 96-well plates (Watson Biolabs), with each 10 μ l reaction containing 3.5 μ l of 2X PCR Taq Mastermix (ABM with dye), 0.5 μ l each of forward and reverse SSR primers (5 pmol concentration), 2 μ l of genomic DNA (~50 ng/ μ l), and 3.5 μ l of nuclease-free water. Thermal cycling parameters carried with an initial denaturation at 94 °C for 3 min, followed by 35 cycles of denaturation (94 °C for 30 s), annealing (50–58 °C for 30 s), and extension (72 °C for 40 s), ending with a final extension step at 72 °C for 5 min. Amplified products were subsequently maintained at 4 °C for storage. PCR amplicons were separated using 3–4% agarose gels (prepared by dissolving 3–4 g of agarose in 1X TAE buffer) and subjected to electrophoresis at 100 V for 2 h in an iLIFE Biotech electrophoresis unit containing 1X TAE buffer. A 100 bp ladder (Genei) was used to characterize amplified fragments. Post-electrophoresis the gels were imaged under ultraviolet with trans-illuminator which stained with ethidium bromide (10 mg/ml). Documentation was performed with the GELSTAN gel documentation system (Medicare, India). Scoring of amplified fragments was done visually, with allele "A" assigned to the recurrent parent (Samba Mahsuri), allele "B" to the donor parent (*O. rufipogon*), and heterozygous alleles designated as "H." Missing alleles were marked as "NA" (Fig. S1).

Linkage map construction and QTL analysis

In this study, a linkage map was developed by employing 80 polymorphic SSR markers using the QTL IciMapping

software v4.2 (CIMMYT), following the Kosambi mapping function as outlined by Wang et al. (2019) with a minimum logarithm of odds (LOD) score threshold of 2.5 and a maximum recombination fraction (θ) of 0.5. The linkage map was constructed through steps involving 30 cM recombination frequency threshold for grouping loci, organizing loci within these groups via K-optimality and the 2-optMAP algorithm, and refining the final order using recombination frequency-based rippling with a window size parameter of 5. To analyze the relationship between marker genotypes and trait phenotypes, QTL mapping was conducted. Inclusive Composite Interval Mapping (ICIM) was performed with a walking step of 2 cM and a LOD threshold of 2.50. The significance threshold for QTL detection was determined empirically by conducting 1,000 permutation tests for each trait across all 12 chromosomes at a genome-wide significance level of $P=0.05$ (Churchill & Doerge, 1994). The permutation analysis yielded an average threshold LOD score of 2.50, which was adopted as the critical value for declaring significant QTLs in this study. Significant QTLs were positioned on the linkage map based on their LOD peaks, and gene action along with the percentage of phenotypic variance explained by each QTL were estimated at these peaks.

QTL nomenclature followed the standard guidelines outlined by "The Committee on Gene Symbolization, Nomenclature, and Linkage (CGSNL) of the Rice Genetic Cooperative" (McCouch, 2008). QTLs with LOD scores exceeding 2.5 and a substantial additive effect on the trait's total phenotypic variance were considered effective QTLs. The QTLs discovered in this research were cross-referenced with previously documented QTLs in the QTARO database (<http://qtaro.abr.affrc.go.jp/>) and Gramene QTL database (<https://gramene.org/>). Mendelian segregation ratio (3:1), segregation ratios of individual markers were analyzed using chi-square (χ^2) tests in the Graphical Genotypes (GGT v2.0) software (Van Berloo, 2008).

Statistical data analysis

The analysis of variance (ANOVA) was calculated using R studio v4.0.2 with the augmented RCBBD package (Aravind et al., 2019) (<https://cloud.r-project.org/web/packages/augmentedRCBD>) to estimate genotypic and phenotypic coefficients of variability, broad-sense heritability, genetic advance as a percentage of the mean, and the distribution of yield-related traits. Box plots and correlation matrices were generated using the 'ggplot2' and 'ggcorrplot' packages in R Studio to visualize data trends effectively. In addition, genetic variability parameters (PCV, GCV, heritability, genetic advance), frequency distribution, and correlation (Pearson, 1900) were calculated using R (R Core Team, 2013).

Results

The segregating population of 190 BC₂F₂ families were phenotyped together with two checks for ten yield and related traits and genotyped using genome-wide polymorphic markers. The mean values genotypic and phenotypic coefficient of variation, heritability and genetic advance were presented in Table 1 and mean performance for the 10 quantitative characters of BC₂F₂ introgression lines and standard checks are presented in Table S2. Analysis of variance in augmented design based on block adjusted and treatment adjusted ANOVA (Table S3) showed that the mean Sum of Squares for Blocks ignoring treatments were having profound effects for all the traits studied. Since evaluation of genotypes was in Augmented Design, significant differences are due to presence of different variable genotypes in each block. Whereas blocks eliminating treatments, as expected, were found to be non-significant for all the traits indicating the homogeneity of experimental site. The analysis of variance revealed substantial genotypic

Table 1 Descriptive Statistics for yield and its contributing characters in BC₂F₂ population derived from Samba Mahsuri x *O. rufipogon*

Characters	Variability			Coefficient of variation		Heritability (%) (Broad sense)	Genetic advance 5%
	Min	Max	Mean	GCV (%)	PCV (%)		
DF	110	135	119.60	5.42	6.42	71.43	9.46
PH	61	125	86.19	14.55	14.68	98.29	29.77
PL	16.67	27.5	21.80	9.13	11.04	68.35	15.57
TT	3	40	14.17	22.97	42.72	28.90	25.47
PT	2	34	12.71	39.93	41.55	92.36	79.18
FGP	40.50	443.5	209.40	37.31	38.00	96.36	75.56
CGP	0	150	28.17	71.18	72.30	96.93	144.59
SF	21.25	100	87.13	11.45	11.56	98.05	23.40
TW	10.29	38.73	15.48	17.10	21.99	60.46	27.43
GYP	3.1	97.78	27.96	55.58	56.19	97.84	113.42

DF days to 50% flowering, PH plant height, PL panicle length, TT Total number of tillers/plant, PT number of productive tillers/plant, GP filled grains per panicle, CGP chaffy grains per panicle, SF Spikelet fertility, TW test weight, GYP grain yield per plant. PCV, GCV & GA Scales: > 10 (Low); 10–20 (Moderate); > 20 (High). Heritability Scales (Broad sense): Up to 30 (Low); 30–60 (Moderate); > 60 (High)

effects among the introgressed lines and checks for all yield-contributing traits. The mean sum of squares due to treatments suggests that genotypes had high variability for the traits studied except for the DF, PL, TT and TW. However, when compared with checks *i.e.*, mean sum of squares due to Treatment vs Checks, PL and TW were also found to have significant variability among the introgression lines (Fig. 3). Since the two checks (Samba Mahsuri and RNR 15048) are prominent high yielding cultivated varieties, they were also showed significant differences for DF, PH, FGP, CGP and SF.

The traits such as DF, PH, PL, TT, PT, FGP and grain yield per plant showed nearly normal distribution whereas CGP, SF and TW were skewed towards Samba Mahsuri (Fig. 4). However, QTL analysis was performed after

eliminating data outliers for each trait to ensure a normal distribution of the phenotyping data which is important factor for QTL mapping. The transgressive segregants displaying phenotypic values beyond parental lines BPT5204 and RNR15048 were observed for all traits studied pointing beneficial genetic interactions where alleles from *O. rufipogon* enhance trait performance within the BPT5204 genetic background indicated a positive genotype-by-genotype ($G \times G$) interaction. Transgressive segregants of 3.15%, 23.15%, 2.10% and 27.89% were recorded over BPT5204 respectively for number of productive tillers, number of filled grains/panicle, test weight and grain yield per plant. However, for the same traits, the percentage of transgressive segregants observed over the check RNR15048 was 3.15%, 15.78%, 1.05% and 24.73% respectively.

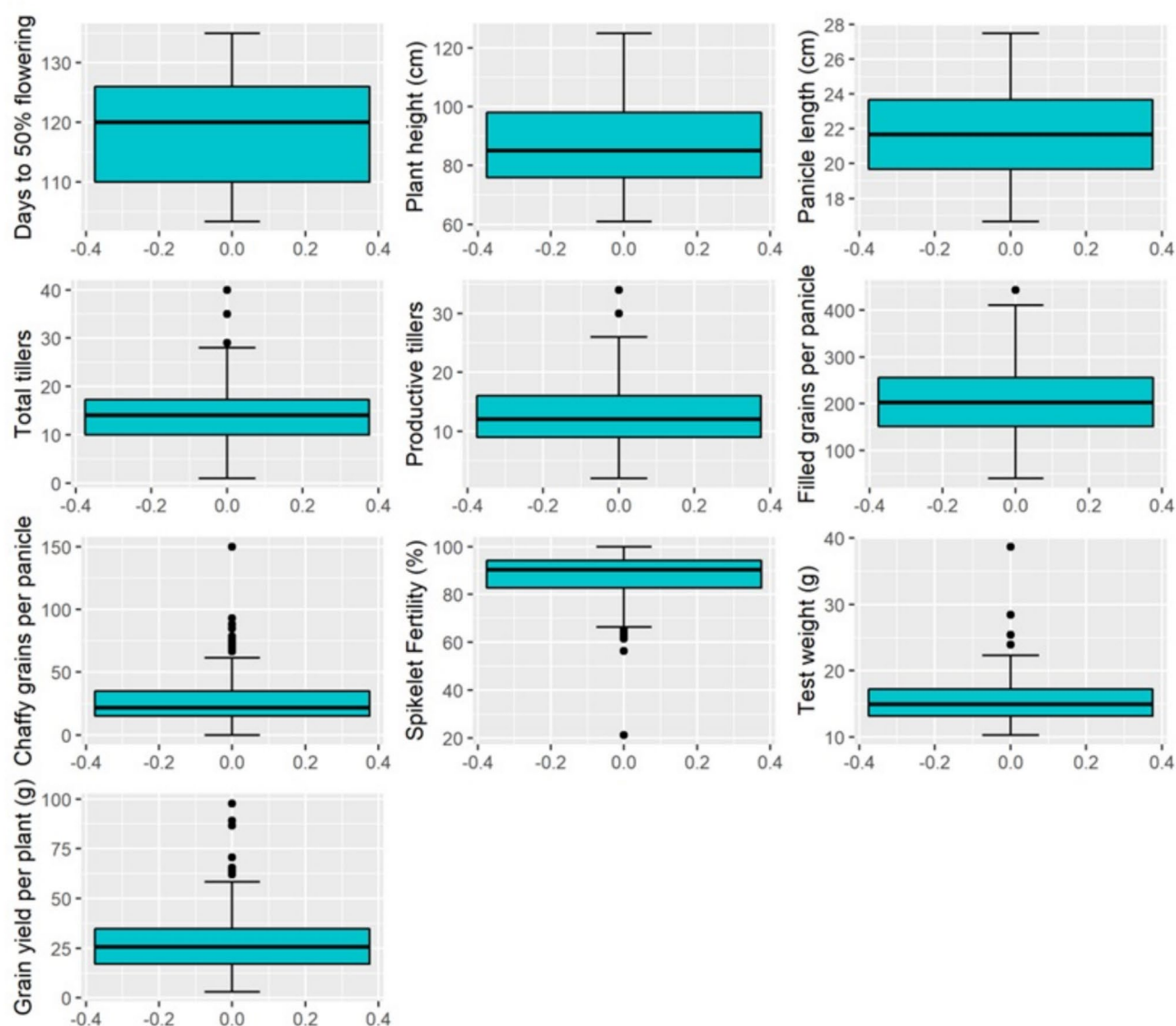


Figure 3. Box plots representing the phenotypic variation for yield and yield contributing traits in BC_2F_2 population of *O. sativa* (Samba Mahsuri)/*O. rufipogon*

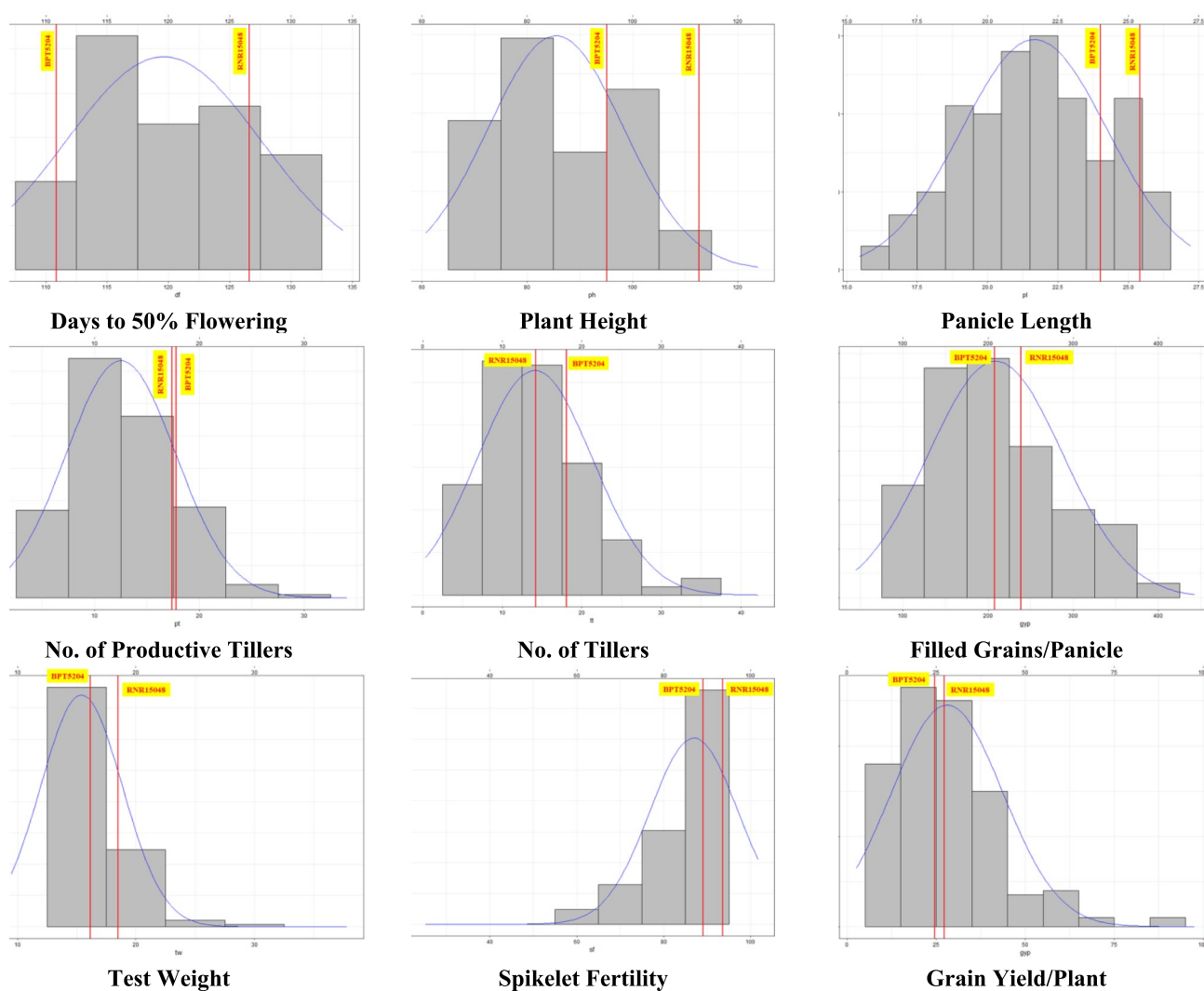


Figure 4. Frequency distribution of yield and yield contributing traits in BC_2F_2 population

The genotypic and phenotypic coefficients of variations for DF and CGP were low, while moderate values were observed for PH and SF. PL exhibited low (9.13) and moderate (11.04) GCV and PCV, respectively. Moderate genotypic (17.10) and high phenotypic (21.99) coefficients of variation recorded for TW. High genotypic and phenotypic coefficients of variability were noticed for the TT, PT, FGP and single plant yield. Rice grain yield is complex trait which in turn mainly determined by the traits viz., PL, PT, FGP, SF, TW and grain yield per plant. Among the 190 BC_2F_2 population, fourteen (7.3%) were considered superior viz., IL10, IL12, IL18, IL21, IL42, IL77, IL104, IL145, IL171, IL202, IL204, IL206, IL208 and IL216 with significant performance for at least any three of the above six characters (Fig. 5). Samba Mahsuri genome recovery in these 14 genotypes ranged from 65.50% to 73.1% whereas the *O. rufipogon* genome introgression ranged between 18% and 29.8% (Table S4; Fig. S2). Genetic advance expressed as a

percentage of the mean was categorized following Johnson and et al. (1955), as low (<10%) in DF (9.46%); moderate (10–20%) in PL (15.57%) and high (>20%) as observed in all other traits, including chaffy grains per panicle (CGP, 93.23%) and single-plant yield (91.42%). chaffy grains per panicle (CGP) indeed recorded the highest genetic advance (93.23%) with high heritability (~96%), suggesting strong additive variance. Although undesirable, this trait is under genetic control and therefore amenable to selection in the negative direction. Importantly, 28 introgression lines performed significantly better than Samba Mahsuri by producing fewer CGP, demonstrating that effective selection against this negative trait was possible. Hence, despite partial recovery of wild genome segments contributing to this trait, lines with reduced CGP and favourable yield attributes could be identified. The apparent skewness towards Samba Mahsuri in CGP and also in spikelet fertility and test weight reflects the directional selection imposed during

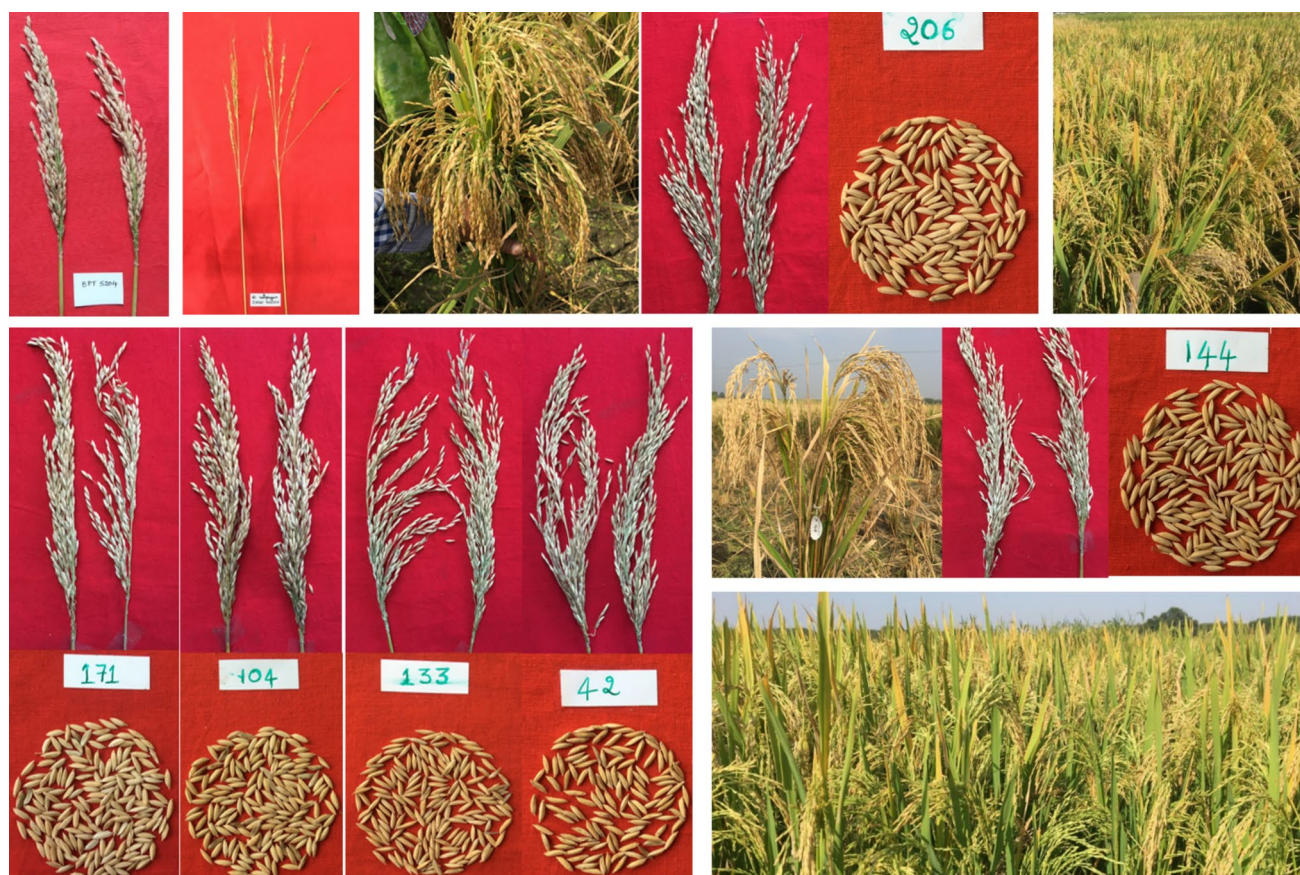


Figure 5. Plant, panicle, grain characteristics and field view of selected introgression lines

two backcross generations in favour of the recurrent parent genotype. Thus, the study addressed the negative trait by identifying lines with low CGP while retaining superior yield component traits. Together, these considerations demonstrated that, although the parental polymorphism level was lower compared to previously reported interspecific crosses, marker coverage and phenotypic variation for key yield traits were sufficient for meaningful QTL identification, and that negative traits such as CGP can be effectively reduced in subsequent selection cycles.

Heritability analysis provides precise genetic information that can be passed on to offspring. Broad-sense heritability values for most traits except TT (28.90%) in this study ranged from 60.46% to 98.29%, with plant height (PH, 98.29%) and spikelet fertility (SF, 98.05%) showing the highest values. Genetic advance (expressed as a percentage of the mean) was categorized as low (<10%) in DF (9.46%); moderate (10–20%) in PL (15.57%) and high (>20%) as observed in all other traits, including chaffy grains per panicle (CGP, 93.23%) and single-plant yield (91.42%). Notably, TT exhibited low heritability but high genetic advance, while DF showed high heritability paired with low genetic advance. For the remaining traits, high heritability coincided

with substantial genetic advance, indicating strong potential for trait improvement through selection.

Correlation analysis

Genotypic correlations between yield and its component traits were estimated and are presented in Fig. 6. Significant positive correlation observed for grain yield per plant with PH (0.34**), PL (0.43**), TT (0.51**), PT (0.63**), FGP (0.53**) and SF (0.19**). Hence selection for these yield component traits will be rewarding. However, inverse relationship among traits also observed which limits the potential gain attainable for yield improvements. DF negatively correlated to grain yield with significant (-0.14*) relationship inferring that selection of high yielding early maturing genotypes will be possible from the developed BC₂F₂ population.

Estimation of recurrent parent genome recovery

A total of 413 SSR markers spanning all 12 chromosomes of rice were initially employed to assess polymorphism. It was observed that 121 of these markers, accounting for

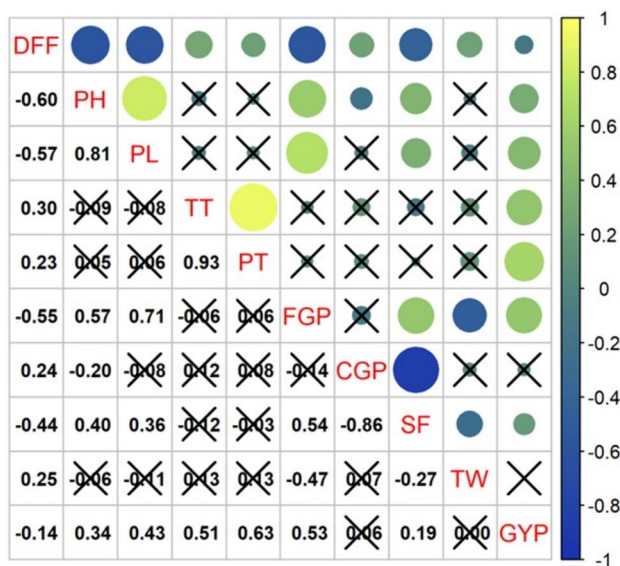


Figure 6. Correlation among yield and yield-related traits in BC₂F₂ mapping population DF: days to 50% flowering, PH: plant height, PL: panicle length, TT: total number of tillers, PT: number of productive tillers, FGP: filled grains per panicle, CGP: chaffy grains per panicle, SF: spikelet fertility, TW: test weight, GYP: grain yield per plant. Critical Values for correlation @ 190 df (n-2)=5%—0.13; 1%—0.18. Values with cross marks are non-significant

approximately 29.30%, exhibited polymorphism between parents. Subsequently, 92 markers were used for estimating the recurrent parent genome recovery. Analysis of the 92 markers showed that the proportion of *O. rufipogon* DNA ranged from 2.2 to 41.4% across the 190 individuals, with an overall average of 19%. In contrast, the average genome proportion of Samba Mahsuri was 71.7%, ranging from 55.5 to 92.1%. IL165 (41.40%) was observed to be having highest proportion of the introgression from *O. rufipogon* followed by IL166 (33%), IL119 (31%), IL28 (31%), IL24 (31%), IL103 (31%), IL175 (30%). Eight Lines (IL136, IL151, IL147, IL141, IL188, IL120, IL14 and IL22) showed lowest genome introgressions. Whereas the average genome proportion of Samba Mahsuri is 71.7% with a range of 55.5–92.1%. When the frequency distribution of the percentage of recovery of Samba Mahsuri was observed, few genotypes viz., IL14 (92.1%), IL22 (92.1%), IL193 (91.7%), IL4 (91.6%) showed highest recovery of Samba Mahsuri genome. Two genotypes viz., IL165 (55.5%), and IL75 (59.5%) showed the lowest recovery of recurrent parent genome (Table S2; Fig. S3).

From the chromosome wise percentage of polymorphism, it was observed that chromosome 8 showed 50% of polymorphism (7 out of 14 were polymorphic) followed by chromosome 3 (46.15%), chromosome 5 (41.94%), chromosome 7 (37.04%), chromosome 1 (33.33%), chromosome 4 (32.43%), chromosome 2 (25%), chromosome 10 (23.33%), chromosome 9 (21.43%), chromosome 6

(18.75%), chromosome 12 (17.39%) and chromosome 11 (14.71%). Chromosome 10, had 7 polymorphic markers in parental polymorphism survey but most of them were observed as monomorphic in BC₂F₂ population. Hence only two of seven were used for linkage map construction for chromosome 10. The average number of polymorphic markers per chromosome was 10.08, with the fewest (4) on chromosome 12 and the highest (18) on chromosomes 1 and 3. However, 29 out of 121 markers which were located at or near to telomere, centromere and which produced a non-scorable banding pattern and were excluded. Among these 92 markers interestingly 12 markers were monomorphic in mapping population even though being polymorphic among parents. Hence only 80 markers used for construction of linkage map.

Segregation distortion of markers

Segregation analysis of markers was carried out using chi-square goodness of fit test to decipher, whether markers following a typical Mendelian ratio or not. The markers which are not segregating according to the expected Mendelian ratio and skewed towards one of the two alleles are considered to be distorted and this phenomenon is called as segregation distortion (Xu & Hu, 2009). According to Wang and et al. (2015), in BC₂F₂ population, the expected segregation ratio (genotype frequency) would be 81.25% (AA), 12.5% (AB) and 6.25% (BB) resulting in an allele frequency of 87.5% (A) and 12.5% (B). A total of 80 SSR markers, which are segregating in BC₂F₂ population were subjected chi-square goodness of fit to decipher marker segregation pattern. The genotype frequency in this population was 72.3% Samba mahsuri, 9.1% Heterozygotes and 18.2% *O. rufipogon*, due to skewed allele frequencies at 56 out of 80 (70%) marker loci. 21 markers (26.25%) were skewed towards Samba Mahsuri while 35 markers (43.75%) showed skewness towards *O. rufipogon* (Table S5) and only 24 marker loci segregated as per the expected Mendelian ratio.

Identification of yield enhancing QTLs introgressed from *O. rufipogon*

A linkage map provides a visual depiction of genetic marker positions on chromosomes, inferred from recombination frequencies between marker pairs. In this study, a linkage map was developed using 80 SSR markers, grouped based on a recombination threshold of 30 cM. The final map spanned 1893 cM (Kosambi), with chromosome lengths varying from 6.38 cM (chromosome 10) to 306.35 cM (chromosome 7). The Marker intervals ranged between 3.19 cM (chromosome 10) and 38.29 cM (chromosome 7), averaging 22.32 cM between adjacent markers (Table S6).

QTL mapping

A total of 47 Significant QTLs were identified at $P < 0.001$ (CIM LOD > 2.50) for ten yield traits as summarized in Table 2 and Fig. 7. Among these, 22 QTLs derived from *Oryza rufipogon* alleles demonstrated beneficial effects on yield traits, indicating the similar contribution of *O. rufipogon* accession alleles in enhancing yield traits. Since, quantitative traits are polygenic in nature, most of the QTLs (44 out of 47 QTLs) were minor QTLs (phenotypic variance < 10) while three exhibited major effects ($> 10\%$ phenotypic variance explained PVE). Chromosomal distribution of QTLs varied markedly with chromosome 7 harbouring the highest number (12 QTLs), followed by chromosomes 1 (8 QTLs) and 2 (6 QTLs). QTLs were not detected on chromosome 10, likely due to limited marker polymorphism. Chromosome 3, 5 and 9 had 3 QTLs each. Chromosome 4 and 11 had four QTLs each. Chromosome 6 showed presence of two QTLs. Chromosome 8 and chromosome 12 had only 1 QTL each.

Among trait-specific QTLs, only one QTL *qPH2.1* from Samba Mahsuri, was detected on chromosome 2 with phenotypic variance of 29.20%. For DF, two QTLs (*qDF1.1* and *qDF3.1*) were identified. *qDF1.1*, a major QTL from recurrent parent Samba Mahsuri explained 29.33% PVE, while *qDF3.1* from *O. rufipogon* contributed 6.59%. Two significant genomic regions from *O. rufipogon* were associated with PL. *qPL7.1* and *qPL9.1* on chromosome 7 and chromosome 9 accounted for 3.50% and 29.14% of PVE respectively. Three genomic regions *qTT1.1* on chromosome 11 from Samba mahsuri (12.21% PVE), *qTT8.1* (2.11%) and *qTT11.1* (5.68%) were identified for total number of tillers which jointly explained 20% of the total variation. Three QTL significantly influenced the number of productive tillers. The contribution of each of QTL was rather small, around 2.81–4.40% with a total of 10.50% ariation. Four Samba Mahsuri-derived QTLs collectively explained 22.48% PVE for FGP on chromosomes 1, 2, 7 and 12 with individual range of 1.35–9.15%. Minor-effect of five QTLs observed for the trait CGP with 1.30% to 1.35% phenotypic variance per QTL. Spikelet fertility (SF) associated with five QTLs (15.08% cumulative PVE), including only one QTL *qSF2.1* with 4.45% phenotypic variance from *O. rufipogon*. Eight QTLs linked to test weight (TW) were identified on chromosomes 1, 2, 3, 4, 7, and 11. Each QTL individually contributed to phenotypic variance is ranging from 2.52% to 4.66%. Collectively, these loci explained 27.82% of the total observed variation, with *Oryza rufipogon* alleles responsible for 18.9% (5 QTLs) of this variance.

Notably, GYP showed the highest QTL number of 14 loci detected, explaining its polygenic inheritance. Fourteen QTL showed significant association with grain yield

which jointly explained 24.48% of the total variation. *O. rufipogon* alleles contributed to 13.77% (8 QTLs) of total variation. The variation explained by these individual QTLs ranged from 1.38% to 2.31% as determined by CIM. One QTL *qGYP4.1* on chromosome 4 exhibited 2.31% of phenotypic variance. Understanding the multifaceted nature of yield-related traits requires not only decoding the molecular mechanisms underlying yield-associated QTLs but also generating actionable insights for traditional breeding programs. Introgression lines (ILs) developed through advanced backcrossing which segregate for specific chromosomal regions serve as powerful tools for fine mapping of yield-associated QTLs. Coupled with genome-wide metabolic analyses, these ILs can illuminate the biological pathways and cellular interactions that collectively govern yield performance.

There were six co-localized QTLs for yield and related traits detected in our study on 5 chromosomes (Table 3). Two QTL hotspot regions each on chromosomes 1 followed by one each on chromosome 2, 3, 7 and 9. QTL hotspots were selected based on co-localization of minimum 3 QTLs and average phenotypic variance of more than 3 percent except *qY9* which had higher average phenotypic variance with 2 QTL regions from *O. rufipogon*. Two QTL hotspots on chromosome 1 between RM10009- RM10149 and RM10662-RM11125 harboured 3 QTLs (*qTW1.1*, *qDF1.1* and *qGYP1.1*) and 4 QTLs (*qGYP1.2*, *qTT1.1*, *qCGP1.1* and *qFGP1.1*) respectively. One QTL hotspot co-localized between RM13075-RM13263 harbouring 3 QTLs (*qSF2.1*, *qFGP2.1* and *qPH2.1*) on chromosome 2. In case of QTL hotspots on chromosome 3 (RM14735-RM15404), chromosome 7 (RM20818- RM21749) and chromosome 9 (RM5526- RM3808), the number of co-localized QTLs were 3 (*qDF3.1*, *qGYP3.1*), 3 (*qPL7.1*, *qGYP7.2* and *qTW7.2*) and 2 (*qPL9.1* and *qGYP9.2*) respectively.

Discussion

This study was aimed at identifying and incorporating QTLs associated with enhanced yield traits from *O. rufipogon*. This involved the generation of backcross populations derived from *Oryza sativa* (Samba Mahsuri) and *Oryza rufipogon*. Besides restoring pollen and seed fertility, backcross helps to surpass sterility in interspecific hybridizations as the fraction of the donor parent genome gets reduced after repeated backcrosses and thus reduces the undesirable effects of the wild genome on the elite background.

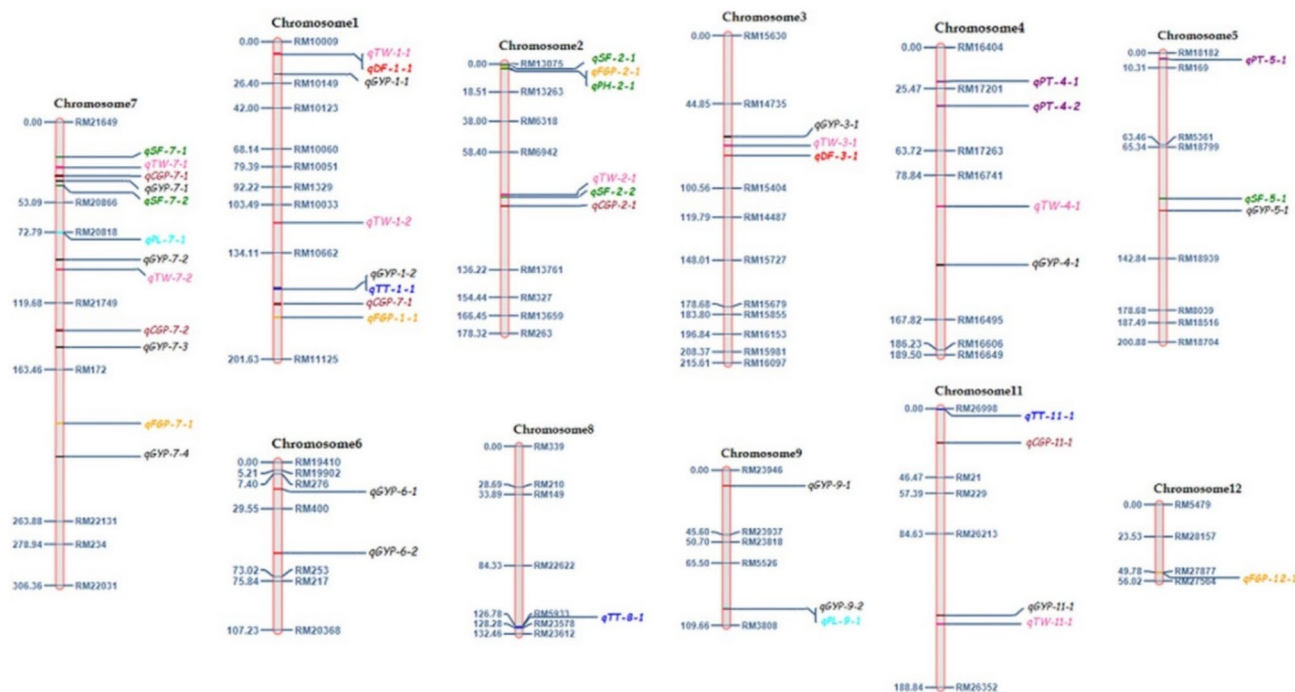
The analysis of variance using augmented RCBD revealed significant variability among the lines. Frequency distribution of yield contributing traits showed that chaffy grains/panicle, SF and test weight were skewed towards

Table 2 QTLs for yield and yield-related traits detected in BC₂F₂ generation of *O. sativa* (Samba Mahsuri)/*O. rufipogon*

Locus	Chromosome	QTL position (cM)	Left marker	Right marker	LOD	PVE (%)	Add
Days to 50% flowering							
<i>qDF1.1</i>	1	8	RM10009	RM10149	24.4978	29.3392	6.9973
<i>qDF3.1</i>	3	79	RM14735	RM15404	2.7664	6.5921	-2.402
Plant height							
<i>qPH2.1</i>	2	3	RM13075	RM13263	14.3266	29.2009	8.5045
Panicle length							
<i>qPL7.1</i>	7	73	RM20818	RM21749	3.999	3.5032	-1.5675
<i>qPL9.1</i>	9	98	RM5526	RM3808	12.4282	29.1464	-1.7929
Total no. of tillers							
<i>qTT1.1</i>	1	157	RM10662	RM11125	2.5281	12.214	1.9534
<i>qTT8.1</i>	8	128	RM5933	RM23578	2.96	2.1161	-0.9052
<i>qTT11.1</i>	11	0	RM26998	RM21	7.5168	5.6866	2.5393
Productive Tillers							
<i>qPT4.1</i>	4	21	RM16404	RM17201	2.6935	3.2865	-5.4209
<i>qPT4.2</i>	4	36	RM17201	RM17263	2.6886	4.4067	0.9884
<i>qPT5.1</i>	5	4	RM18182	RM169	4.2577	2.8145	1.8741
Filled grains/panicle							
<i>qFGP1.1</i>	1	175	RM10662	RM11125	4.7306	9.1587	36.9129
<i>qFGP2.1</i>	2	2	RM13075	RM13263	5.2452	3.0253	31.9865
<i>qFGP7.1</i>	7	199	RM172	RM22131	3.5692	8.9447	36.7312
<i>qFGP12.1</i>	12	50	RM27877	RM27564	3.2962	1.3558	6.4172
Chaffy grains/panicle							
<i>qCGP1.1</i>	1	165	RM10662	RM11125	6.2087	1.3089	-0.9121
<i>qCGP2.1</i>	2	94	RM6942	RM13761	6.9269	1.307	0.5722
<i>qCGP7.1</i>	7	35	RM21649	RM20866	5.9722	1.359	-16.7388
<i>qCGP7.2</i>	7	137	RM21749	RM172	5.296	1.308	-14.0595
<i>qCGP11.1</i>	11	23	RM26998	RM21	5.7184	1.3147	0.4993
Spikelet fertility							
<i>qSF2.1</i>	2	0	RM13075	RM13263	22.9603	4.4581	6.5916
<i>qSF2.2</i>	2	88	RM6942	RM13761	4.0324	2.6886	8.3553
<i>qSF5.1</i>	5	101	RM18799	RM18939	5.7052	2.7613	8.2164
<i>qSF7.1</i>	7	23	RM21649	RM20866	4.8832	2.6043	-0.4476
<i>qSF7.2</i>	7	42	RM21649	RM20866	7.4459	2.5748	7.8929
Test weight							
<i>qTW1.1</i>	1	7	RM10009	RM10149	3.7312	2.5298	0.432
<i>qTW1.2</i>	1	116	RM10033	RM10662	3.9104	2.5365	-2.3578
<i>qTW2.1</i>	2	87	RM6942	RM13761	4.3064	4.0061	-2.4095
<i>qTW3.1</i>	3	72	RM14735	RM15404	4.2216	3.6276	-0.1444
<i>qTW4.1</i>	4	98	RM16741	RM16495	3.346	3.25	0.5724
<i>qTW7.1</i>	7	30	RM21649	RM20866	4.2046	3.148	0.3905
<i>qTW7.2</i>	7	98	RM20818	RM21749	4.017	4.0577	-2.3315
<i>qTW11.1</i>	11	146	RM26213	RM26352	3.5635	4.6673	-1.4123
Grain yield/plant							
<i>qGYP1.1</i>	1	19	RM10009	RM10149	2.8474	1.4536	-2.0208
<i>qGYP1.2</i>	1	156	RM10662	RM11125	4.2995	1.9157	2.0706
<i>qGYP3.1</i>	3	66	RM14735	RM15404	2.5127	1.612	-13.3465
<i>qGYP4.1</i>	4	134	RM16741	RM16495	3.5052	2.3195	-15.9198
<i>qGYP5.1</i>	5	110	RM18799	RM18939	3.4821	1.6354	-13.9774
<i>qGYP6.1</i>	6	17	RM276	RM400	2.6364	1.3861	0.1332
<i>qGYP6.2</i>	6	58	RM400	RM253	2.7266	1.6928	1.3748
<i>qGYP7.1</i>	7	39	RM21649	RM20866	2.7114	1.6768	-2.2417
<i>qGYP7.2</i>	7	91	RM20818	RM21749	2.5656	1.4646	-14.0162
<i>qGYP7.3</i>	7	149	RM21749	RM172	2.6026	1.9362	0.1894
<i>qGYP7.4</i>	7	221	RM172	RM22131	4.5207	1.973	1.6855
<i>qGYP9.1</i>	9	11	RM23946	RM23937	5.0435	1.8101	1.3523

Table 2 (continued)

Locus	Chromosome	QTL position (cM)	Left marker	Right marker	LOD	PVE (%)	Add
<i>qGYP9.2</i>	9	98	RM5526	RM3808	3.4504	1.7579	-1.3485
<i>qGYP11.1</i>	11	140	RM26213	RM26352	4.0168	1.8522	-13.0503

**Figure 7.** QTLs identified for yield and yield related traits in BC₂F₂ Generation of *O. sativa* (Samba Mahsuri)/*O. rufipogon***Table 3** Co-Localization of QTLs for yield and related traits detected in BC₂F₂ generation of *O. sativa* (Samba Mahsuri)/*O. rufipogon*

Chr	QTL hotspot	Genomic region	Position (cM)	Physical size (Mb)	Locus	PVE %	Co-localizations with previous QTLs/genes
1	<i>qY1a</i>	RM10009- RM10149	0.00–26.40	2.85	<i>qTW1.1</i> , <i>qDF1.1</i> , <i>qGYP1.1</i>	11.11	<i>OsmiR393</i> (Xia et al., 2012); <i>gpl a</i> (Yu et al., 1997)
1	<i>qY1b</i>	RM10662-RM11125	134.11–201.63	10.06	<i>qGYP1.2</i> , <i>qTT1.1</i> , <i>qCGP1.1</i> , <i>qFGP1.1</i>	6.15	<i>phd1</i> (Li et al., 2011)
2	<i>qY2</i>	RM13075-RM13263	0.00–18.51	5.52	<i>qSF2.1</i> , <i>qFGP2.1</i> , <i>qPH2.1</i>	12.23	<i>qPH-2</i> (Mao et al., 2003); <i>np2.2</i> (Marri et al., 2005); <i>d50</i> (Sato et al., 2002); <i>qSS-2</i> (Wang et al., 2005); <i>MTR1</i> (Tan et al., 2012)
3	<i>qY3</i>	RM14735-RM15404	44.85–100.56	12.97	<i>qDF3.1</i> , <i>qGYP3.1</i> , <i>qTW3.1</i>	3.94	<i>Pdhk</i> (Mukherjee et al., 2012); <i>phyB</i> (Takano et al., 2005); <i>gw3.1</i> (Thomson et al., 2003); <i>qGY-3</i> (Mao et al., 2003); <i>GS3</i> (Fan et al., 2006); <i>qGL-3a</i> (Wan et al., 2005); <i>ki3.1</i> (Li et al., 2004a, 2004b); <i>qGW3</i> (Miyata et al., 2007)
7	<i>qY7</i>	RM20818- RM21749	72.79–119.68	20.04	<i>qPL7.1</i> , <i>qGYP7.2</i> , <i>qTW7.2</i>	3.01	<i>sp2(t)</i> (Yoshimura et al., 2004); <i>snb</i> (Lee et al., 2006); <i>chr729</i> (Hua et al., 2012); <i>OsMADS15</i> (Kobayashi et al., 2012); Lu et al., 2012); <i>qNFGP-7-1</i> (Zhuang et al., 2002); <i>qSSP7</i> (Xing et al., 2008); <i>gp7a</i> (Li et al., 2000); <i>yd7a</i> (Li et al., 2000); <i>gw7</i> (Li et al., 2000)
9	<i>qY9</i>	RM5526- RM3808	65.50–109.66	12.99	<i>qPL9.1</i> , <i>qGYP9.2</i>	15.45	<i>Dn-1</i> (Yu et al., 1995); <i>OsEMF2b</i> (Xie et al., 2014)

DF days to 50% flowering, PH plant height, PL panicle length, TT total number of tillers/plant, PT number of productive tillers/plant, FGP filled grains per panicle, CGP chaffy grains per panicle, SF spikelet fertility, TW test weight, GYP grain yield per plant

Samba Mahsuri which is due to selections in two back crosses for recurrent parent genotype. *O. rufipogon* was late flowering and maturing (Xiao et al., 1998); however, 32% of the introgressed lines were on par with BPT5204 in terms of their flowering and maturity. Despite the fact that *O. rufipogon* panicles were consistently shorter than those of cultivated (Moncada et al., 2001), the *O. rufipogon* accession (IRGC106106) used here reported to have 31–40 cm PL at reproductive stage which could be the reason for 16.32% and 4.21% transgressive segregants for PL over BPT5204 and RNR15048 respectively. Of the 190 BC₂F₂ lines, fourteen introgression lines were found to be superior. 190 BC₂F₂ population used in the study considering statistical power, genetic resolution, and resource availability which can be sufficient to capture recombination events and donor-derived alleles of interest while maintaining manageable resource requirements for precise trait measurement and genotyping. Population sizes in the range of 150–250 individuals have been widely employed in several QTL studies (Marri et al., 2005; McCouch et al., 1997; Septiningsih et al., 2003), providing sufficient statistical power for detecting QTLs of moderate-to-large effect while allowing detailed phenotyping. These results indicate that genes from *O. rufipogon*, despite a low-yielding agronomically inferior species, into an elite genetic background can enhance key agronomic characters.

The average performance of introgression lines confirmed the well-known fact that wild species (*O. rufipogon*) have alleles to ameliorate traits that can boost the yield of present day cultivars. Similar to previous reports (Marri et al., 2005; Moncada et al., 2001; Septiningsih et al., 2003; Thomson et al., 2003; Xiao et al., 1998) high-yielding positive transgressive segregants were obtained in this study from interspecific hybridization with *O. rufipogon*. These ILs may be useful for breeding, providing diverse germplasm to boost elite cultivar yields. Genetic variability is basis for selection of diverse parents for hybridization which was emphasized by Falconer (1981), to exert selection pressure in breeding material. Critical analysis of adequate genetic variability is necessary to start crop improvement programs and use appropriate selection techniques. The extent of genetic variability and the heritability of traits are key factors for the success of crop improvement. The genetic coefficient of variability can be accurately measured by partitioning the heritable and non-heritable components of the observed variability. The phenotypic coefficient of variation (PCV) and genotypic coefficient of variation (GCV) provide valuable insights into the magnitudes of variability, highlighting both heritable variation and the environmental influence on the genotypes. PCV values in this study are slightly higher than the GCV indicating the insignificant environmental

impact on the characters. Thus, employing direct selection could prove effective for enhancing these traits.

The quantitative traits in the study had high heritability, implying that they are less affected by the environment and that there may be a high concurrence between phenotypic value and breeding value. It should be noted that genetic variation in traits may not solely stem from additive gene action, and relying exclusively on heritability estimates cannot predict the genetic improvement achievable through selecting top-performing plants. Therefore, distinguishing between the relative contributions of additive and non-additive genetic components becomes crucial when analysing target traits. While high heritability values not necessarily translate to substantial genetic gains, combining it with genetic advance assessments provides a more accurate final effect for developing improved cultivars. The trait, TT showed low heritability with high genetic advance, implying that it is primarily determined by additive gene effects, although the low heritability is due to significant environmental influences. Conversely, DF exhibited little genetic advance with high heritability, suggesting a non-additive gene action and a propitious environmental influence rather than the genotype. Selection for this trait is perhaps ineffective as reported by Mohan et al. (2016) and Singh and Verma (2018). PL recorded a high heritability with moderate genetic advance, which is consistent with Singh and Verma (2018). All other traits showed high heritability with greater genetic advance, implying preeminence of additive gene action in controlling the traits. As a result, direct selection for these traits is expected to be effective. These results followed the findings of Singh and Verma (2018) and Abebe et al. (2017).

The main aim of character association studies is to get an insight how different traits are suitable for selection. Selecting a particular trait could lead to desirable or undesirable changes in the correlated characters. Since its complexity due to multifactor interactions and high environmental influence, yield trait was not aimed directly for selection. In addition to their association with yield, component characters exhibit various associations among themselves. In order to plan a sound breeding program, it is imperative to know the strength of the correlation between yields and attribute characters since unfavourable associations may limit genetic progress. Positive significant association of PL with single plant yield and FGP indicating that there is an increase in single plant yield of introgressed lines of BC₂F₂ population by accommodating more number of grains per panicle. For most correlations, the degree and direction (positive or negative) of the correlation was consistent with that observed in populations derived by Xiao and et al. (1998), Moncada and et al. (2001), Thomson and et al. (2003) and Fu et al. (2010) from an interspecific cross with *O. rufipogon*. Previous

studies reported negative correlation of test weight with total number of tillers (Marri et al., 2005), Productive Tillers (Marri et al., 2005; Fu et al., 2010), Filled grains per panicle (Septiningsih et al., 2003; Thomson et al., 2003).

The parental polymorphism, an estimate of genetic diversity was lower in the present study compared to prior studies with different accessions of *O. rufipogon*, where polymorphism stretched from 30 to 90% with *indica* and *japonica* parents (Lee et al., 2005; Li et al., 2002; Marri et al., 2005; Moncada et al., 2001; Qi et al., 2017; Tan et al., 2008; Thomson et al., 2003; Tian et al., 2006; Xiao et al., 1998). However, it varies with the parental sources used, gene pool they belong, position and distribution of used markers on chromosomes. In some regions of the genome, the frequency of polymorphism is found low, such as at the short arms of chromosome 10, 11 and 12. The wild accession IRGC106106, used in this study is an Indian origin *O. rufipogon* which was collected from the areas of West Bengal. Hence, the low polymorphism may be attributed to genetic similarity or reduced recombination (Causse et al., 1994; Grandillo & Tanksley, 2005) between *O. rufipogon* (IRGC 106106) and Samba mahsuri (Indian cultivar). This pattern of low polymorphism on short arm of 10th chromosome was observed by Xiao and et al. (1998) and Septiningsih and et al. (2003) between *O. rufipogon* (IRGC 105491) and *indica* species but high polymorphism was observed in crosses with *japonica* cultivars implying that the *O. rufipogon* genome is less diverse to the *indica* than the *japonica* subgroup. Comparing linkage maps across different studies having *O. rufipogon* as the donor parent is less critical because variations in the number of markers, recombination rates between markers, and the type of mapping population used can significantly alter the length of linkage map. Despite this, a genome-wide coverage of 80 well-distributed polymorphic SSR markers with scorable segregation patterns were used for linkage map construction, ensuring genome-wide coverage for QTL mapping. QTL analysis was conducted after removing phenotypic outliers and normalizing trait distributions to reduce the risk of false positives and support a reliable QTL detection. It is also noteworthy that low GCV was trait-specific (e.g., DF and CGP), while major yield components such as productive tillers, filled grains per panicle, and single-plant yield exhibited high GCV, providing adequate phenotypic variation for QTL mapping. Therefore, although overall polymorphism was modest, the marker density and phenotypic variability were adequate, making the QTL map conservative but reliable for identifying useful alleles.

The phenomenon of segregation distortion, a potent evolutionary drive, is often encountered whenever wild species are used as donor parents (Reffinur et al., 2014; Xian-Liang et al., 2006). Segregation distortion has power of reducing

accuracy of QTL detection by altering the order and genetic distance of markers on the linkage group. The segregation distortion observed for 56 out of 80 (70%) markers in the mapping population used in the study was attributed to selection pressure during backcrosses and the presence of sterility loci (gamete eliminators and pollen killers) which affects pollen fertility and gene segregation (Bimpong et al., 2011). In interspecific crosses with *O. rufipogon*, which utilized the SSR and RFLP markers, reported 19.2%–68.5% segregation distortion on all chromosomes (Lee et al., 2005; Marri et al., 2005; Moncada et al., 2001; Septiningsih et al., 2003; Tan et al., 2008; Thomson et al., 2003). The literatures also cited, variable range of marker distortion with *O. sativa*/*japonica* × *O. glaberrima* populations i.e., Aluko et al., 2004 (39%) and Bimpong et al., 2011 (93.1%). Muralidhara et al., 2020 reported 100% segregation distortion in interspecific hybridization with *O. glaberrima*. Studies reported different sterility loci viz., *S1* (Yang et al., 2016), *S21* (Miyazaki et al., 2007) in the hybrid combinations obtained from *O. sativa* / *O. rufipogon* hybridizations. On the other hand, sterility loci act as gamete eliminators or pollen eradicators and result in segregation distortion as found in *O. sativa* × *O. glaberrima* (Li et al., 2018).

Due to the selection pressures applied for recurrent parent genotypes during population development in BC₁ and BC₂, the observed skewing toward the adapted, elite parent is plausible. However, the skewing towards *Oryza rufipogon* may be attributed to factors such as genetic distance or linkage drag in certain genomic regions, which can hinder recombination. In the BC₁ generation, selection might cause segregation distortion, leading to the retention of some *O. rufipogon* regions instead of the expected segregation of QTLs in the BC₂F₂ generation. However, Zhang et al. (2010) suggested that while segregation distortion occurs, it does not significantly influence the identification of QTL positions or effects. Its impact remains minimal, particularly when using co-dominant markers such as SSRs in backcross populations, as noted by Xian-Liang et al., 2006. Various statistical tools, including Proc QTL (SAS) (Xu & Hu, 2009) and MapDisto 2.0 (Aluko et al., 2004), have been effectively utilized to manage distorted markers in F₂, BC₁F₁, RIL, and DH populations without adverse effects. However, there remains a need to develop an optimized package specifically tailored for BC₂F₂ populations. The QTLs identified in this study are likely minimally affected by segregation distortion, as the majority aligns with previously reported QTLs. The cumulative phenotypic variance explained by donor-derived QTLs was estimated at 6.59%, 32.65%, 2.12%, 3.29%, 3.98%, 2.60%, 18.90%, and 13.77% for days to 50% flowering, panicle length, total number of tillers, number of productive tillers, chaffy grains per panicle, spikelet fertility, test weight, and grain yield per

plant, respectively including the major QTL *qPL9.1* with 29.14% PVE.

In this study, alleles from *Oryza rufipogon* positively influenced 46.80% (22 QTLs) of the identified QTLs associated with yield traits. These findings align with previous research that reported beneficial effects of wild species alleles on 15.5–74% of QTLs (Xiao et al., 1998; Moncada et al., 2001; Thomson et al., 2003; Septiningsih et al., 2003; Marri et al., 2005; Tian et al., 2006; Tan et al., 2008; Qi et al., 2017). The higher percentage inferred in this study suggests that the *O. rufipogon* accession used may possess more advantageous alleles compared to those examined in prior studies.

Given that *O. rufipogon* is known for its late flowering and maturity characteristics (Xiao et al., 1998), selecting for *qDF1.1* could aid breeders in developing early-maturing introgression lines through interspecific hybridization. The *qDF1.1* locus, positioned between markers RM10009 and RM10149, coincides with the *OsmiR393* gene (Xia et al., 2012). Overexpression of *OsmiR393*, a microRNA in *Oryza sativa*, reduces the expression of auxin receptor genes (*OsTIR1* and *OsAFB2*), leading to an increase in tiller number and early flowering, though at the cost of auxin sensitivity. Additionally, the *qDF3.1* region introgressed from *O. rufipogon* overlaps with previously identified genes such as pyruvate dehydrogenase kinase (*pdhk*) (Mukherjee et al., 2012) and phytochrome B (*phyB*) (Takano et al., 2005), both of which have been linked with early flowering and improvements in grain number and size.

Plant height (PH) is closely linked to the inherent gigantism observed in wild rice (Xiong et al., 1999). The QTL *qPH2.1*, which accounts for 29.2% of the phenotypic variation from the recurrent parent, may facilitate the development of introgression lines with PH similar to Samba Mahsuri in wide hybridizations. Multiple studies reported PH-related QTLs in this genomic region, including *qPH-2* (Mao et al., 2003) and *np2.2* (Marri et al., 2005). The observed effects of these QTLs may be associated with the finely mapped dwarf locus *d50* (Sato et al., 2002), which is located within the same genomic region.

Samba mahsuri though it is a high yielding *indica* cultivar, exhibits incomplete panicle emergence with susceptibility to multiple biotic stresses (Chandu et al., 2024; Potupureddi et al., 2017). *O. rufipogon* accession IRGC106106 used in this study had long panicles with 31–40 cm length at reproductive stage. However, none of the lines showed superiority in PL over the wild parent but showed full panicle emergence for most of the lines, unlike Samba mahsuri. Among the 190 introgressed lines, 32 exhibited panicle exertion. Panicle emergence was measured based on the length of exertion of the panicle base above the sheath of the flag leaf or $\geq 90\%$ of the panicle emerged from the flag leaf sheath comparable

to the recurrent parent, Samba Mahsuri is considered as full panicle emergence. The general mean of the introgressed lines for PL was 21.8 cm, with a range from 16.67 cm (IL84) to 27 cm (IL206 and IL216). Among the 190 introgressed lines, 32 exhibited panicle exertion comparable to the recurrent parent, Samba Mahsuri. Consequently, the major QTL *qPL9.1* and minor QTL *qPL7.1* derived from *O. rufipogon* could be beneficial for selecting lines with complete panicle emergence. The QTL *qPL7.1* is located in a genomic region bound with several known loci, including *sp2(t)* (Yoshimura et al., 2004), *snb* (SUPERNUMERARY BRACT) (Lee et al., 2006), *chr729* (CHD-related gene 729) (Hua et al., 2012) and *OsMADS15* (Kobayashi et al., 2012; Lu et al., 2012), all of which play roles in transitioning from spikelet meristem to floral meristem, floral organ development, secondary panicle branching, and overall PL regulation. Meanwhile, *qPL9.1* overlaps with genes such as *Dn-1* (Yu et al., 1997) and *OsEMF2b* (Xie et al., 2014), which regulate traits including panicle density, floral organ differentiation, number of stamens, and pistil morphology.

Five QTLs for negative trait CGP were detected in this study, of which three QTLs were introgressed from *O. rufipogon* since the chaffy grains are a common phenomenon in wild species. In contrary to this, all the four QTLs for FGP were introgressed from recurrent parent Samba Mahsuri. The QTLs *qFGP1.1* and *qFGP7.1* explaining the phenotypic variance of 9.15% and 8.94% respectively will be useful for good filled grain number in distant hybridization programmes. A previous study (Zhuang et al., 2001) also reported QTL for FGP in *qFGP1.1* location. The QTL *qSF2.1* along with the other 3 QTL regions (*qSF2.2*, *qSF5.1*, *qSF7.2*) for SF from Samba mahsuri helps breeders in developing fertile offspring in interspecific crosses because sterility is a common phenomenon in hybridization programmes with wild species. For TW, 8 QTLs were detected in this study, among them, *qTW2.1* is in the same region as *yld2.1* (plot yield) and *np2.2* (number of panicles per plant) QTLs reported by Marri et al. (2005) in an advanced backcross IR58025A $\times$ *O. rufipogon* (IC 22015) population. Eight yield enhancing minor QTLs were introgressed from *O. rufipogon* contributing additively 13.77% phenotypic variance to the trait. The QTL *qGYP4.1* alone got 2.31% of phenotypic variance. Previous studies reported the QTLs *gwt4a* (Lin et al., 1996) for grain weight, *spp4-2* (Xiao et al., 1998) and *qSNP-4a* (Mei et al., 2006) for spikelets per panicle in the same region of *qGYP4.1*.

Yield in rice is ascertained by a combination of characteristics such as tillers, panicles, spikelets and grains number, SF, grain weight, etc. The QTLs for these yield-related traits are often clustered in a few chromosome segments as QTL hotspots, i.e., genomic regions where groups of traits co-localize. This allude the presence of a single gene with a

pleiotropic effect or closely linked loci affecting two or more traits. Moncada et al. (2001), Septiningsih et al. (2003), Thomson et al. (2003), Li et al., (2004a, 2004b) and Marri et al. (2005) reported the co-localization of QTLs for yield and related traits on different chromosomes for different co-localized traits. The presence of QTLs for yield and yield-related traits in the same marker interval of the chromosome was found in several other studies indicating a non-random distribution of yield-related traits in the rice genome.

Linkage or pleiotropy cause co-localization resulting in significant correlations between co-localized traits that provide possible explanation for QTL “hotspots”. Desirable significant trait correlations were observed in all co-localized regions. Co-localized regions on chromosomes 1 (RM10662-RM11125), 2 (RM13075-RM13263) and 9 (RM5526-RM3808) were found to be useful and enhancing the yield with all positive significant correlations among traits. QTLs for PL (*qPL7.1* and *qPL9.1*) have co-localized with grain yield per plant at both the regions. Fine mapping and cloning of QTL hot spots provide an opportunity for manipulating or improving multiple traits simultaneously in addition to understanding the biology of the genes controlling underlying QTLs (Ashikari et al., 2005;). Further, QTL hotspots identified in multiple environments could be ideal for studying their regulatory functions. In the present study, QTL hotspots *qY1a* (2.85 Mb) and *qY1b* (10.06 Mb) with average phenotypic variance of 11.11% and 6.15% were reported on chromosome 1. As reported in previous studies, *qY1a* sharing common region with the gene *OsmiR393* (Xia et al., 2012) and the QTL *gp1a* (Yu et al., 1997) governing the traits early flowering and grain number per panicle respectively whereas *qY1b* sharing with FGP QTL (Zhuang et al., 2001) and *phd1* (photo assimilate defective1) (Li et al., 2011) governing grain production. The QTL hotspots on chromosome 2, *qY2* (5.52 Mb) found to harbour three QTLs with average phenotypic variance of 12.23%. The co-localization of *qTT1.1* (tiller number) and *qCGP1.1* (chaffy grains per panicle) within the same marker interval may reflect pleiotropy or tight linkage between loci controlling these traits. Similar clustering of QTLs for favorable and unfavorable traits has been reported in rice (Moncada et al., 2001; Thomson et al., 2003; Xiao et al., 1998). Such QTL clusters are typical in interspecific populations where beneficial alleles from wild relatives may be linked with deleterious ones. Marker-assisted selection with fine mapping can help break this linkage, enabling retention of the favorable tiller trait while selecting against excess chaffy grains.

Co-localization of identified QTLs with functionally characterized genes which are directly associated with key physiological processes influencing yield, indicating a linkage of several mapped QTL hotspots to known genes and pathways regulating physiological traits. Panicle exertion

and branching (*qPL7.1*, *qPL9.1*; *sp2(t)*, *OsMADS15*, *chr729*) influence spikelet exposure, pollination efficiency, and light interception during grain filling. Spikelet fertility and reproductive organ development (*qSF2.1*, *qSS-2*, *MTR1*) affect pollen viability, fertilization success, and seed set. Early flowering, grain production and number per panicle (*OsmiR393*, *gp1a*, *phd1*) is controlling developmental timing and synchrony with optimal environmental conditions. Grain production, grain filling and grain size determination (*gw3.1*, *GS3*, *qGL-3a*, *qGW3*) impact the assimilate partitioning and sink strength in rice plants. Source-sink balance through pleiotropic or linked loci directly affect traits like tiller number, panicle number, and grain weight.

Positive allele contribution from female parent Samba Mahsuri, is desirable for the traits like SF, FGP and PH in wide crosses. The genomic region (12666016-18187278) on chromosome 2 having *qSS-2* and *MTR1* (MICROSPORE AND TAPETUM REGULATOR1) might be providing SF to fertile progeny in crosses with wild species (Tan et al., 2012; Wang et al., 2005). However, this region also associated with PH QTL *qPH-2* (Mao et al., 2003), panicle number QTL *np2.2* (Marri et al., 2005) and a dwarf gene *d50* (Sato et al., 2002) reported in earlier studies. Similarly, QTL hotspots *qY3* (12.97 Mb) and *qY7* (20.04 Mb) were co-localized with three QTLs each with an average phenotypic variance of 3.94% and 3.01%, respectively, with favourable allelic contributions from *O. rufipogon*. The loci associated with grain number and size (*pdhk*-pyruvate dehydrogenase kinase; Mukherjee et al., 2012), flowering time (*phyB*-phytochromeB; Takano et al., 2005), grain weight (*gw3.1*; Thomson et al., 2003), grain size (*GS3*; Fan et al., 2006), grain length (*qGL-3a*; Wan et al., 2005 and *ki3.1*; Li et al., 2004a, 2004b) grain width (*qGW3*; Miyata et al., 2007) and grain yield (*qGY-3*; Mao et al., 2003) were identified in *qY3* region. The QTL hotspot *qY7* of our study on chromosome 7 was co-localized with QTLs/gene for floral meristem formation gene *snb* (Lee et al., 2006), panicle branching locus *chr729* (Hua et al., 2012), PL locus *sp2(t)* (Yoshimura et al., 2004), floral organ formation locus *OsMADS15* (Kobayashi et al., 2012; Lu et al., 2012) in the marker interval RM20818-RM21749. The QTLs *qNFGP-7-1* (Zhuang et al., 2001) and *qSSP7* (Xing et al., 2008) for filled grains per panicle also mapped to the same region of chromosome 7. Li et al., 2000 reported co-localisation of *gp7a* for FGP, *yd7a* for yield per plant and *gw7* for TW in the *qY7* region. QTL hotspot *qY9* (12.99 Mb) harboured two QTLs with average phenotypic variance of 15.45% with positive allele contribution from *O. rufipogon*. The QTL hotspots *qY1a* and *qY1b* have positive allele contributions from both the parents.

Conclusion

In conclusion, this study aimed to enhance yield-related traits in rice through the introgression of desirable alleles from the wild rice species, *O. rufipogon* into the *indica* cultivar Samba Mahsuri. The development of advanced backcross populations, spanning multiple generations, allowed for the creation of introgression lines with improved traits while retaining the desired characteristics of the recurrent parent. The study detected a total of 47 significant QTLs associated with ten yield-related traits, with 46.80% of these QTLs demonstrating beneficial effects from *O. rufipogon* alleles. The segregation distortion observed in the marker analysis, although a common phenomenon in interspecific crosses, did not significantly impact the identification of QTLs. The co-localization of QTLs revealed hotspot regions on chromosomes 1, 2, 3, 7, and 9 indicating clusters of genes influencing multiple yield-related traits. These regions could be targeted for further exploration in breeding programs aimed at enhancing rice yield and productivity. The genetic variability analysis highlighted the presence of superior genotypes among the ILs, with 14 lines showing significant performance for multiple yield-related traits with higher recovery of the recurrent parent genome Samba Mahsuri and the introgression of beneficial alleles from *O. rufipogon*. The study contributes valuable insights into the genetic basis of yield-related traits in rice and provides a foundation for future breeding efforts. The identified QTLs and ILs with superior performance offer promising resources for marker-assisted breeding programs, allowing for the development of rice varieties with improved yield and agronomic traits.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s40502-025-00904-w>.

Acknowledgements Authors acknowledge the research support received from ICAR-Indian Institute of Rice Research for carrying out the research.

Author contributions GC has designed the experiment. MS carried out development of population and genotyping with the guidance and support of GC and DB. MS, GC and DB were involved in the data analysis. MS, GC, DB, SM, AMS, RCH, SP, VB, SCD and RMS were involved in the various stages of research and preparation and peer review of the manuscript. All authors have reviewed and approved the final manuscript.

Funding The authors declare that no funds, grants, or other support were received during the preparation of this manuscript.

Data availability All relevant data generated during this study are provided within the manuscript and its supplementary materials.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethical approval and consent to participate Not applicable.

References

- Abebe, T., Alamerew, S., & Tulu, L. (2017). Genetic variability, heritability and genetic advance for yield and its related traits in rainfed lowland rice (*Oryza sativa* L.) genotypes at Fogera and Pawe, Ethiopia. *Advances in Crop Science and Technology*, 5(2), 2–8.
- Alexandratos, N., & Bruinsma, J. (2012). *World agriculture towards 2030/2050: The 2012 revision*. FAO, Rome. ESA Working Paper No. 12–03.
- Aluko, G., Martinez, C., Tohme, J., Castano, C., Bergman, C., & Oard, J. H. (2004). QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* x *O. glaberrima*. *Theoretical and Applied Genetics*, 109, 630–639.
- Aravind, J., Shankar, M. S., Wankhede, D. P., & Kaur, V. (2019). *Analysis of Augmented Randomised Complete Block Designs* (pp. 1–17).
- Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., Angeles, E. R., Qian, Q., Kitano, H., & Matsuoka, M. (2005). Cytokinin oxidase regulates rice grain production. *Science*, 309(5735), 741–745.
- Balakrishnan, D., Subrahmanyam, D., Badri, J., Raju, A. K., Rao, Y. V., Beerelli, K., Mesapogu, S., Surapaneni, M., Ponnuswamy, R., Padmavathi, G., & Babu, V. R. (2016). Genotype × environment interactions of yield traits in backcross introgression lines derived from *Oryza sativa* cv. Swarna/*Oryza nivara*. *Frontiers in Plant Science*, 7, Article 1530.
- Bimpong, I. K., Serraj, R., Chin, J. H., Joie, R., Mendoza, E. M. T., Hernandez, J. E., Mendioro, M. S., & Brar, D. S. (2011). Identification of QTLs for drought-related traits in alien introgression lines derived from crosses of rice (*Oryza sativa* cv. IR64) × *O. glaberrima* under lowland moisture stress. *Journal of Plant Biology*, 54, 237–250.
- Causse, M. A., Fulton, T. M., Cho, Y. G., Ahn, S. N., Chunwongse, J., Wu, K., Xiao, J., Yu, Z., Ronald, P. C., & Harrington, S. E. (1994). Saturated molecular map of the rice genome based on an interspecific backcross population. *Genetics*, 138, 1251–1274.
- Chandu, G., Balakrishnan, D., Munnam, S. B., Mangrauthia, S. K., Sanjeeva Rao, D., Neeraja, C. N., Sundaram, R. M., & Neelamraju, S. (2024). Mapping QTLs for grain iron, zinc, and yield traits in advanced backcross inbred lines of Samba mahsuri (BPT5204)/*Oryza rufipogon*. *Journal of Plant Biochemistry and Biotechnology*, 33(1), 68–84.
- Churchill, G. A., & Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics*, 138(3), 963–971.
- Doyle, J. J., & Doyle, J. L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12, 13–15.
- Falconer, D. S. (1981). *Introduction to Quantitative Genetics* (2nd ed.). London: Oliver and Boyd.
- Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., Li, X., & Zhang, Q. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative trans membrane protein. *Theoretical and Applied Genetics*, 112, 1164–1171.
- Fu, Q., Zhang, P., Tan, L., Zhu, Z., Ma, D., Fu, Y., Zhan, X., Cai, H., & Sun, C. (2010). Analysis of QTLs for yield-related traits in

- Yuanjiang common wild rice (*Oryza rufipogon* Griff.). *Journal of Genetics and Genomics*, 37(2), 147–157.
- Grandillo, S., & Tanksley, S. D. (2005). Advanced backcross QTL analysis results and perspectives. In R. Tuberosa, R. L. Phillips, & M. Gale (Eds.), *In the wake of the double helix: From the green revolution to the gene revolution. Proceedings of the International Congress, Bologna, Italy, 27–31 May 2003* (pp. 115–132). Dordrecht: Kluwer Academic Publishers.
- Hua, Y., Liua, D., Zhonga, X., Zhanga, C., Zhanga, Q., & Zhou, D. (2012). CHD3 protein recognizes and regulates methylated histone H3 lysines 4 and 27 over a subset of targets in the rice genome. *Proceedings of the National Academy of Sciences of the United States of America*, 109(15), 5773–5778.
- Johnson, H. W., Robinson, H. F., & Comstock, R. E. (1955). Estimates of genetic and environmental variability in soybeans. *Agronomy Journal*, 47(7), 314–318.
- Kobayashi, K., Yasuno, N., Sato, Y., Yoda, M., Yamazaki, R., Kimizu, M., Yoshida, H., Nagamura, Y., & Kyozuka, J. (2012). Inflorescence meristem identity in rice is specified by overlapping functions of three AP1/FUL-like MADS Box Genes and PAP2, a SEPALLATA MADS Box Gene. *The Plant Cell*. <https://doi.org/10.1105/tpc.112.097105>
- Lee, D., Lee, J., Moon, S., Park, S. Y., & Gynheung. (2006). The rice heterochronic gene SUPERNUMERARY BRACKET regulates the transition from spikelet meristem to floral meristem. *The Plant Journal*, 49, 64–78.
- Lee, S. J., Oh, C. S., Su, J. P., McCouch, H. S. R., & Ahn, S. N. (2005). Identification of QTLs for domestication-related and agronomic traits in an *Oryza sativa* x *O. rufipogon* BC1F7 population. *Plant Breeding*, 124, 209–219.
- Li, D., Sun, C., Fu, Y., Li, C., Zhu, Z., Chen, L., Cai, H., & Wang, X. (2002). Identification and mapping of genes for improving yield from Chinese common wild rice (*O. rufipogon* Griff.) using advanced backcross QTL analysis. *Chinese Science Bulletin*, 47(18), 1533–1537.
- Li, C., Wang, Y., Liu, L., Hu, Y., Zhang, F., Mergen, S., Wang, G., Schlappi, M. R., & Chu, C. A. (2011). Rice plastidial nucleotide sugar epimerase is involved in galactolipid biosynthesis and improves photosynthetic efficiency. *PLoS Genetics*, 7(7), Article e1002196.
- Li, J., Thomson, M., & McCouch, S. R. (2004a). Fine mapping of a grain-weight quantitative trait locus in the pericentromeric region of rice chromosome 3. *Genetics*, 168(4), 2187–2195.
- Li, J., Xiao, J., Grandillo, S., Jiang, L., Wan, Y., Deng, Q., Yuan, L., & McCouch, S. R. (2004b). QTL detection for rice grain quality traits using an interspecific backcross population derived from cultivated Asian (*O. sativa* L.) and African (*O. glaberrima* S.) rice. *Genome*, 47, 697–704.
- Li, J. X., Yu, S. B., Xu, C. G., Tan, Y. F., Gao, Y. J., Li, X. H., & Zhang, Q. (2000). Analyzing quantitative trait loci for yield using a vegetatively replicated F2 population from a cross between the parents of an elite rice hybrid. *Theoretical and Applied Genetics*, 101, 248–254.
- Li, J., Zhou, J., Xu, P., Deng, X., Deng, W., He, M., Yang, Y., Zhang, Y., & Tao, D. (2018). Neutral alleles at hybrid sterility loci of *Oryza glaberrima* from AA genome relatives in genus *Oryza*. *Breeding Science*, 68, 343–351.
- Li, M., Pan, X., & Li, H. (2022). Pyramiding of *gn1a*, *gs3*, and *ipa1* exhibits complementary and additive effects on rice yield. *International Journal of Molecular Sciences*, 23(20), 12478.
- Lin, H.-X., Qian, H.-R., Zhuang, J.-Y., Lu, J., Min, S.-K., Xiong, Z.-M., Huang, N., & Zheng, K. L. (1996). RFLP mapping of QTLs for yield and related characters in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, 92, 920–927.
- Liu, L., Tong, H., Xiao, Y., Che, R., Xu, F., Hu, B., Liang, C., Chu, J., Li, J., & Chu, C. (2015). Activation of big grain1 significantly improves grain size by regulating auxin transport in rice. *Proceedings of the National Academy of Sciences of the United States of America*, 112(35), 11102–11107.
- Lu, S. J., Wei, H., Wang, Y., Wang, H. M., Yang, R. F., Zhang, X. B., & Tu, J. M. (2012). Overexpression of a transcription factor *OsMADS15* modifies plant architecture and flowering time in rice (*Oryza sativa* L.). *Plant Molecular Biology Reporter*, 30, 1461–1469.
- Magudeeswari, P., Balakrishnan, D., Surapaneni, M., Krishnam Raju, A., Rao, Y. V., Pranay, G., Valarmathi, P., Bhadana, V. P., Neelamraju, S., & Sundaram, R. M. (2024). Exploring stable low soil phosphorous stress tolerance in rice using novel allele recombination from *Oryza rufipogon*. *Plant Breeding*. <https://doi.org/10.1111/pbr.13238>
- Mao, B. B., Cai, W. J., Zhang, Z. H., Hu, Z. L., Li, P., Zhu, L. H., & Zhu, Y. G. (2003). Characterization of QTLs for harvest index and source-sink characters in a DH population of rice (*Oryza sativa* L.). *Yi Chuan Xue Bao*, 30, 1118–1126.
- Marri, P. R., Sarla, N., Reddy, L. V., & Siddiq, E. A. (2005). Identification and mapping of yield and yield related QTLs from an Indian accession of *Oryza rufipogon*. *Bmc Genetics*, 6, 33.
- McCouch, S. R., et al. (1997). Microsatellite marker development, mapping and applications in rice genetics and breeding. *Plant Molecular Biology*, 35(1–2), 89–99.
- McCouch, S. R. (2008). Gene nomenclature system for rice. *Rice*, 1, 72–84.
- Mei, H. W., Xu, J. L., Li, Z. K., Yu, X. Q., Guo, L. B., Wang, Y. P., Ying, C. S., & Luo, L. J. (2006). QTLs influencing panicle size detected in two reciprocal introgressive line (IL) populations in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, 112, 648–656.
- Miura, K., Ashikari, M., & Matsuoka, M. (2011). The role of QTLs in the breeding of high-yielding rice. *Trends in Plant Science*, 16(6), 319–326.
- Miyata, M., Yamamoto, T., Komori, T., & Nitta, N. (2007). Marker-assisted selection and evaluation of the QTL for stigma exertion under japonica rice genetic background. *Theory of Applied Genetics*, 114, 539–548.
- Miyazaki, Y., Doi, K., & Yasui, H. (2007). Identification of a new allele of F1 pollen sterility gene, *S21*, detected from the hybrid between *Oryza sativa* and *O. rufipogon*. *Rice Genetics Newsletter*, 23, 36–38.
- Mohan, C. Y., Srinivas, B., Thippeswamy, S., & Padmaja, D. (2016). Diversity and variability analysis for yield parameters in rice (*Oryza sativa* L.) genotypes. *Indian Journal of Agricultural Research*, 50(6), 609–613.
- Moncada, P., Martinez, C. P., Borrero, J., Chatel, M., Gauch, H. J., Guimaraes, E., Tohme, J., & McCouch, S. R. (2001). Quantitative trait loci for yield and yield components in an *Oryza sativa* X *Oryza rufipogon* BC2F2 population evaluated in an upland environment. *Theoretical and Applied Genetics*, 102, 41–52.
- Mukherjee, R., Gayen, S., Chakraborty, A., Bhattacharyya, J., Maiti, M. K., Basu, A., & Sen, S. K. (2012). Double-stranded RNA-mediated downregulation of *pdhk* gene expression to shorten maturation time of a late maturing native indica rice cultivar, Badshahbhog. *Crop Science*, 52, 1743–1753.
- Muralidhara, B., Anantha, M. S., Senguttuvel, P., Ramappa, L., Patil, B., Manoj, A., Madhav, M. S., Sundaram, R. M., Shankarappa, S. K., Mangal Deep, T., Sreedevi, B., Parmar, B., Rathod, S., Barbadikar, K. M., Suneetha, K., Subbarao, L. V., Mondal, T. K., & Gireesh, C. (2020). Mapping QTL hotspots associated with weed competitive traits in backcross population derived from *Oryza sativa* L. and *O. glaberrima* Steud. *Scientific Reports*, 10, 22103.
- Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from

- random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 50(302), 157–175.
- Potupureddi, G., Suneel, B., Subba Rao, L. V., Sonti, R. V., Sundaram, R. M., Laha, G. S., Sudhakar, P., Satya, A. K., & Madhav, M. S. (2017). Identification of agro-morphological characters in sheath blight tolerant lines of Samba Mahsuri (BPT-5204) rice variety. *Bulletin of Environment, Pharmacology and Life Sciences*, 6(10), 41–45.
- Qi, L., Sun, Y., Li, J., Su, L., Zheng, X., Wang, X., Li, K., Yang, Q., & Qiao, W. (2017). Identify QTLs for grain size and weight in common wild rice using chromosome segment substitution lines across six environment. *Breeding Science*, 67, 472–482.
- R Core Team. (2013). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from: <http://www.R-project.org/> (Version 3.0.0).
- Reddy, M. V., Prasad, S. S. N. D. B., Reddy, B. M., & Rao, L. V. S. (1979). BPT 5204— a new rice variety for kharif season for coastal districts of Andhra Pradesh. *Andhra Agricultural Journal*, 26, 66–67.
- Reflinur, K., Kim, B., Jang, S. M., Chu, S. H., Bordiya, Y., Akter, Md. B., Lee, J., Chin, J. H., & Koh, H. J. (2014). Analysis of segregation distortion and its relationship to hybrid barriers in rice. *Rice*, 7(3), 1–12.
- Sato, K., Ashikari, M., Tamura, K., Mase, K., Kitano, H., Matsuoka, M., & Katayama, Y. (2002). High resolution map of DWARF 50 (D50) in rice. *Rice Genetics Newsletter*, 19, 31–33.
- Sen, P., Chatterjee, A., Kumar, D., Bhattacharyya, S., Bandyopadhyay, S., & Sarkar, A. (2024). Genetic and functional mechanisms of yield-related genes in rice. *Acta Physiologiae Plantarum*, 46(3), 29.
- Septiningsih, E. M., Prasetyono, J., Lubis, E., Tai, T. H., Tjubaryat, T., Moeljopawiro, S., & McCouch, S. R. (2003). Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryza sativa* variety IR64 and the wild relative *O. rufipogon*. *Theoretical and Applied Genetics*, 107, 1419–1432.
- Singh, N., & Verma, O. P. (2018). Genetic variability, heritability and genetic advance in rice (*Oryza sativa* L.) under salt stressed soil. *Journal of Pharmacognosy and Phytochemistry*, 7(3), 3114–3117.
- Sundaram, R. M., Naveenkumar, B., Biradar, S. K., Balachandran, S. M., Mishra, B., IlyasAhmed, M., Viraktamath, B. C., Ramesha, M. S., & Sarma, N. P. (2008). Identification of informative SSR markers capable of distinguishing hybrid rice parental lines and their utilization in seed purity assessment. *Euphytica*, 163(2), 215–224.
- Takano, M., Inagaki, N., Xie, X., Yuzurihara, N., Hihara, F., Ishizuka, T., Yano, M., Nishimura, M., Miyao, A., Hirochika, H., & Shinomura, T. (2005). Distinct and cooperative functions of phytochromes A, B and C in the control of deetiolation and flowering in rice. *The Plant Cell*, 17, 3311–3325.
- Tan, H., Liang, W., Hu, J., & Zhang, D. (2012). MTR1 encodes a secretory fasciclin glycoprotein required for male reproductive development in rice. *Developmental Cell*, 22, 1127–1137.
- Tan, L., Zhang, P., Liu, F., Wang, G., Ye, S., Fu, Z. Y., Cai, H., & Sun, C. (2008). Quantitative trait loci underlying domestication-and yield related traits in an *Oryza sativa* x *Oryza rufipogon* advanced backcross population. *Genome*, 51, 692–704.
- Thomson, M. J., Tai, T. H., McClung, A. M., Lai, X. H., Hinga, M. E., Lobos, K. B., Xu, Y., Martinez, C. P., & McCouch, S. R. (2003). Mapping quantitative trait loci for yield, yield components and morphological traits in an advanced backcross population between *Oryza rufipogon* and the *Oryza sativa* cultivar Jefferson. *Theoretical Applied Genetics*, 107, 479–493.
- Tian, F., Li, D. J., Fu, Q., Fu, Z. F. Z. Y. C., Wang, X. K., & Sun, C. Q. (2006). Construction of introgression lines carrying wild rice (*Oryza rufipogon* Griff.) segments in cultivated rice (*Oryza sativa* L.) background and characterization of introgressed segments associated with yield-related traits. *Theoretical Applied Genetics*, 112, 570–580.
- Van Berloo, R. (2008). Computer note: GGT 2.0: Versatile software for visualization and analysis of genetic data. *Journal of Heredity*, 99, 232–236.
- Wan, X. Y., Wan, J. M., Weng, J. F., Jiang, L., Bi, J. C., Wang, C. M., & Zhai, H. Q. (2005). Stability of QTLs for rice grain dimension and endosperm chalkiness characteristics across eight environments. *Theoretical and Applied Genetics*, 110, 1334–1346.
- Wang, Q., Li, K., Hu, X., Shi, H., Liu, Z., Wu, Y., Wang, H., & Huang, C. (2019). Genetic analysis and QTL mapping of stalk cell wall components and digestibility in maize recombinant inbred lines from B73 × By804. *The Crop Journal*, 8(1), 132–139.
- Wang, C., Zhu, C., zhai, H., & Wan, J. (2005). Mapping segregation distortion loci and quantitative trait loci for spikelet sterility in rice (*Oryza sativa* L.). *Genetical Research*, 86, 97–106.
- Wang, J., Xu, H., Li, N., Fan, F., Wang, L., Zhu, Y., & Li, S. (2015). Artificial selection of *Gn1a* plays an important role in improving rice yields across different ecological regions. *Rice*, 8, 37.
- Wang, Y., & Li, J. (2005). The plant architecture of rice (*Oryza sativa*). *Plant Molecular Biology*, 59(1), 75–84.
- Xia, K., Wang, R., Ou, X., Fang, Z., Tian, C., Duan, J., Wang, Y., & Zhang, M. (2012). *OsTIR1* and *OsAFB2* downregulation via *OsmiR393* overexpression leads to more tillers, early flowering and less tolerance to salt and drought in rice. *PLoS ONE*, 7(1), Article e30039.
- Xian-Liang, S., Xue-Zhen, S., & Tian-Zhen, Z. (2006). Segregation distortion and its effect on genetic mapping in plants. *Chinese Journal of Agricultural Biotechnology*, 3(03), 163–169.
- Xiao, J., Li, J., Grandillo, S., Ahn, S. N., Yuan, L., Tanksley, S. D., & McCouch, S. R. (1998). Identification of trait improving quantitative trait loci alleles from a wild rice relative, *Oryza rufipogon*. *Genetics*, 150, 899–909.
- Xie, S., Min Chen, M., Pei, R., Ouyang, Y., & Yao, J. (2014). *OsEMF2b* acts as a regulator of flowering transition and floral organ identity by mediating H3K27me3 deposition at *OsLFL1* and *OsMADS4* in rice. *Plant Molecular Biology Reporter*. <https://doi.org/10.1007/s11105-014-0733-1>
- Xing, Y. Z., Tang, W. J., Xue, W. Y., Xu, C. G., & Zhang, Q. (2008). Fine mapping of a major quantitative trait loci, *qSSP7*, controlling the number of spikelets per panicle as a single Mendelian factor in rice. *Theoretical and Applied Genetics*, 116, 789–796.
- Xing, Y., & Zhang, Q. (2010). Genetic and molecular bases of rice yield. *Annual Review of Plant Biology*, 61, 421–442.
- Xiong, L. Z., Liu, K. D., Dai, X. K., Xu, C. G., & Zhang, Q. (1999). Identification of genetic factors controlling domestication related traits of rice using an F2 population of a cross between *Oryza sativa* and *O. rufipogon*. *Theoretical and Applied Genetics*, 98, 243–251.
- Xu, S., & Hu, Z. (2009). Mapping quantitative trait loci using distorted markers. *International Journal of Plant Genome*, 2009, 1–11.
- Yadavalli, V. R., Balakrishnan, D., Surapaneni, M., Addanki, K., Mesapogu, S., Beerelli, K., Desiraju, S., Voleti, S. R., & Neelamraju, S. (2022). Mapping QTLs for yield and photosynthesis-related traits in three consecutive backcross populations of *Oryza sativa* cultivar Cottondora Sannalu (MTU1010) and *Oryza rufipogon*. *Planta*, 256(4), 71.
- Yang, Y., Zhou, J., Li, J., Xu, P., Zhang, Y., & Tao, D. (2016). Mapping QTLs for hybrid sterility in three AA genome wild species of *Oryza*. *Breeding Science*, 66, 367–371.
- Yoshimura, A., Takano-Kai, N., & Anno, C. (2004). Linkage mapping of genes for short panicle and awn in rice. *Rice Genetics Newsletter*, 21, 17–19.
- Yu, S. B., Li, J. X., Xu, C. G., Tan, Y. F., Gao, Y. J., Li, X. H., Zhang, Q., & Maroof, M. A. S. (1997). Importance of epistasis as the genetic

- basis of heterosis in an elite rice hybrid (hybrid vigor molecular markers quantitative trait loci interaction between loci). *Proceedings of National Academy of Sciences*, 94, 9226–9231.
- Yu, Z. H., McCouch, S. R., Tanksley, S. D., Kinoshita, T., & Sato, S. (1995). Association of morphological and RFLP markers in rice (*Oryza sativa* L.). *Genome*, 38(3), 566–574.
- Zhang, L., Wang, S., Li, H., Deng, Q., Zheng, A., Li, S., Li, P., Li, Z., & Jiankang Wang, J. (2010). Effects of missing marker and segregation distortion on QTL mapping in F2 populations. *Theory of Applied Genetics*, 121, 1071–1082.
- Zhuang, J. Y., Fan, Y. Y., Wu, J. L., Xia, Y. W., & Zheng, K. L. (2001). Comparison of the detection of QTL for yield traits in different generations of a rice cross using two mapping approaches. *Yi Chuan Xue Bao*, 28, 458–464.
- Zhuang, J.-Y., Fan, Y.-Y., Rao, Z.-M., Wu, J.-L., Xia, Y.-W., & Zheng, K.-L. (2002). Analysis on additive effects and additive-by-additive epistatic effects of QTLs for yield traits in a recombinant inbred line population of rice. *Theor Appl Genet*, 105, 1137–1145.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.