Genomics 114 (2022) 110269

Contents lists available at ScienceDirect

Genomics

journal homepage: www.elsevier.com/locate/ygeno

Development and validation of a novel core set of KASP markers for the traits improving grain yield and adaptability of rice under direct-seeded cultivation conditions

Nitika Sandhu [a,*,1], Jasneet Singh [a,1], Gursewak Singh [a], Mehak Sethi [a], Mohini Prabha Singh [a], Gomsie Pruthi [a], Om Prakash Raigar [a], Rupinder Kaur [a], Rupinder Kaur [a], Preetinder Singh Sarao [a], Jagjeet Singh Lore [a], Uma Maheshwar Singh [b], Shilpi Dixit [b], Deepti B. Sagare [c], S.P. Singh [d], Vanisri Satturu [e], Vikas Kumar Singh [c], Arvind Kumar [b,f]

[a] Punjab Agricultural University, Ludhiana 141004, Punjab, India
[b] International Rice Research Institute (IRRI) South Asia Regional Centre (ISARC), Varanasi, Uttar Pradesh 221106, India
[c] International Rice Research Institute (IRRI), South-Asia Hub (SAH), Patancheru, Telangana 502324, India
[d] Bihar Agricultural University, Bhagalpur Rd, Sabour, Bihar 813210, India
[e] Professor Jayashankar Telangana State Agricultural University (PJTSAU), Rajendranagar, Hyderabad 500030, Telangana, India
[f] International Crops Research Institute For The Semi-Arid Tropics (ICRISAT), Patancheru, Telangana 502324, India

ARTICLE INFO

Keywords:
Abiotic
Biotic
Direct seeded rice
Genomics-assisted breeding
Grain yield
KASP marker
Whole genome resequencing

ABSTRACT

The development and utilization of molecular-markers play an important role in genomics-assisted breeding during pyramiding of valuable genes. The aim of present study was to develop and validate a novel core-set of KASP (Kompetitive Allele-Specific PCR) markers associated with traits improving rice grain yield and adaptability under direct-seeded cultivation conditions. The 110 phenotypically validated KASP assays out of 171 designed KASP, include assays for biotic-resistance genes, anaerobic germination, root-traits, grain yield, lodging resistance and early-uniform emergence. The KASP assays were validated for their robustness and reliability at five different levels using diverse germplasm, segregating and advanced population, comparison with SSR markers and on F_{1s}. The present research work will provide (i) breeding material in form of anticipated pre-direct-seeded adapted rice varieties (ii) single improved breeding line with many useful genes and (iii) KASP assay information for the useful QTL/genes providing grain yield and adaptability to rice under direct-seeded cultivation conditions.

1. Introduction

Rice (*Oryza sativa*) is the staple food for more than 3.5 billion people which comprises around 50% of the world's population [1]. According to an estimate, a yield increases of around 1.5 to 2.4% per year is required to sustain the growing demand [2] On the contrary, the studies indicated that the rice yield increases seem to have plateaued around various parts of the world [2]. Limited water supplies, reduced cultivation area, fluctuating climatic conditions, and labor shortage are the major challenges faced by conventional puddled transplanted rice.

Direct-seeded rice (DSR) is a feasible alternative to conventional puddled transplanted rice (PTR) [3] having a potential for sustaining the future rice demand due to low water requirement, reduced labor requirements, mitigation of greenhouse gas emissions, adaptation to climatic risks, and the yield comparable with that of transplanted rice [4]. The DSR cultivation method has not gained the required popularity because of certain issues such as the poor crop stand, low yield [5], weeds [4], poor adaptability, reduced nutrient uptake (especially of phosphorus, nitrogen, and iron) [6], and lodging [7]. The inefficient uptake of water and nutrients under aerobic cultivation conditions resulted from the poor root architecture [8] leading to yield reduction [4] DSR cultivation system is generally more favorable for the growth of weeds that compete with rice for nutrients, moisture, and sunlight, and can cause large yield losses compared to the transplanted system of rice

* Corresponding author.
E-mail address: nitikasandhu@pau.edu (N. Sandhu).
[1] Both authors contributed equally.

https://doi.org/10.1016/j.ygeno.2022.110269
Received 25 June 2021; Received in revised form 12 November 2021; Accepted 16 January 2022
Available online 19 January 2022

cultivation [9].

In order to improve the crop establishment during the initial stages, DSR adapted rice varieties with higher germination and faster seedling emergence with more vigorous growth are needed to minimize the risks encountered in direct seeding [10–13]. This could help in reducing soil evaporation and accelerate root access to soil water and nutrients [14]. The development of DSR adapted rice varieties depends on the selection of suitable traits, identification, and introgression of genomic regions associated with those particular traits of interest in different genetic backgrounds. The traits reported to play an important role in providing yield stability and adaptability under DSR include anaerobic germination (ability to germinate under water) [15], early uniform seedling emergence [6,16], vegetative vigor, root phenotypic plasticity, proper nutrient uptake [6], and lodging resistance [17–19]. The biotic stress (disease and insect) resistance includes blast [20], brown spot [21], bacterial blight [22], sheath blight, brown planthopper [23], gall midge [24], and nematode resistance [25].

To meet the gaps between rice crop yields and global rice consumption, rice breeders need to continuously release new rice varieties with better yield potential, adaptability under DSR, acceptable grain quality traits, high nutrient-use efficiency, and resistance to various biotic/abiotic stresses. Breeding methods such as QTL (quantitative trait loci) /gene pyramiding and multiparent application have been reported to be effective in development of rice varieties [18,26,27]. The conventional breeding approach suffers from the problem of linkage drag, leading to the transfer of undesired traits closely linked with traits of interest. The use of molecular markers in breeding programs improves the efficiency of traditional breeding by enabling breeders to select trait-linked molecular markers [28]. Molecular marker-assisted selection (MAS) combined with conventional breeding approaches, is a good strategy to identify individual genotypes associated with different economically important traits, which can dramatically improve rice breeding efficiency [29]. The development and utilization of genetic markers play a pivotal role in marker-assisted breeding of rice cultivars during pyramiding of valuable genes. High polymorphism, co-dominant inheritance, high density, high throughput, and easy automation and data exchange are the characteristics of ideal DNA markers [30].

Recent advancements in next-generation sequencing (NGS) and single nucleotide polymorphism (SNP) genotyping can help to accelerate crop improvement provided their proper integration and deployment into breeding programs [31]. To date, the marker-assisted breeding programs attempting to use these QTL/genes have almost universally relied on old SSR (simple-sequence repeats) marker systems. SSRs are not so useful in marker-assisted introgression program involving multiple donors as there is possibility of getting same allelic pattern for multiple parents. The explosion in genomics resources and the low-cost genome resequencing approaches now enables the development of highly accurate SNP (single nucleotide polymorphism) marker systems. SNPs are the markers of choice for most high-throughput genotyping applications as they are abundant, co-dominant and evenly distributed along the genome. Also, high-throughput SNP genotyping platforms based on Kompetitive Allele-Specific PCR (KASP) and other methods have provided routine, rapid and cost-effective genotyping solutions for the targeted MAS (marker-assisted selection) [32–35].

The next generation sequencing methods ranged from whole genome re-sequencing/skim sequencing [36] to the reduced representation sequencing [37]. A number of SNP arrays such as GoldenGate 1536 SNPs [38], 384-plex BeadXpress [39], Illumina Infinium-based 6 K arrays, RiceSNP6K [40], C6AIR (Cornell_6K_Array_Infinium_Rice) [41], HDRA700K (High Density Rice Array) [42], RiceSNP50K array and Affymatrix 50 K array [43,44], 44 K array [45] have been developed for rice and their utility has also been demonstrated across a range of applications. Instead, all these tremendous efforts, a large gap between the development of SNP markers and their application in marker-assisted breeding still exists. In 2005, total of 408,898 SNP based DNA polymorphisms were identified using the draft genome sequence of rice cultivars Nipponbare (*japonica*) [46]. Non redundant SNPs were identified using the sequencing data of 20 different cultivars in 100 Mb genomic region [47]. Various databases have been constructed for SNP markers in rice, e.g., the Gramene database (http://ensembl.gramene.org/genome_browser/index.html), the Rice Diversity Project database (https://ricediversity.org/), the Rice Genome Annotation Project database (http://rice.plantbiology.msu.edu/), and the Rice SNP-Seek Database (http://snp-seek.irri.org). Considering this, the present study was designed keeping the following objectives in mind (i) to develop SNP/allele specific trait-linked markers for the traits improving grain yield and adaptability under DSR (ii) to validate the targeted SNPs in parental lines and different mapping populations (iii) to develop a trait linked SNP set targeting specific QTL/genes of interest.

2. Results

2.1. Phenotyping

A total of 108 F_3 introgression lines having genes/QTL for early uniform emergence (*qEUE*$_{3.1}$, *qEUE*$_{11.1}$), anaerobic germination (*qAG*$_{9.1}$, *qAG*$_{9.2}$), seedling vigor (*qEVV*$_{9.1}$), high nutrient uptake (*qNR*$_{5.1}$), root density (*qRHD*$_{1.1}$), culm strength (*qCS*$_{1.1}$), lodging resistance (*qLDG*$_{3.1}$, *qLDG*$_{4.1}$), grain yield under DSR (*qGY*$_{1.1}$, *qGY*$_{6.1}$, *qGY*$_{9.1}$, *qGY*$_{10.1}$), and biotic stress (blast-*Pi9*, gall midge-*Gm4*, *Gm8*, BPH-*Bph3*, *Bph17*) tolerance were tested for agronomical performance under DSR in 2019WS (wet season) (Fig. 1A). The grain yield of selected single plant ranged from 14.9 to 48.5 g/plant whereas for PR126 single plant yield was 29.3 g/plant (Table S1), the days to 50% flowering ranged from 68 to 91 days and plant height ranged from 95 to 125 cm (data not shown). Based on plant and grain type, a total of 599 single plants were selected and advanced under controlled conditions in 2020DS (dry season) (Fig. 1B). The 599 F_5 introgression lines were tested for agronomical performance under DSR (Fig. 1C). A total of 54 breeding lines performed better in terms of grain yield and yield related traits over the best yielded local check (PR126). In addition, the seeds of 42 advanced breeding lines developed using 12 donors, including donors for biotic and abiotic stress tolerance and DSR-adapted traits were procured from IRRI, Philippines and multiplied/evaluated at ISARC-Varanasi in 2019WS and evaluated at PAU in 2020WS (Fig. 1D). Both the sets of introgression lines were evaluated for capability of anaerobic germination (Fig. 1E, 1F), resistance to bacterial blight (Fig. 2G), blast (Fig. 2H) and for root traits improving nutrient uptake (Fig. 1I) (Table S2). The advanced introgression lines were evaluated for agronomic traits at ISRAC, Varanasi in 2019WS (Table S2). Analysis of variance to study genotype × location interactions (Table S3) indicated significant differences among genotypes suggesting the presence of variability among genotypes and among locations. The breeding lines IR 129477–1510–100-7-5-4, IR 129477–1510–100-7-5-6, IR 129477–1629-14-1-4-2, IR 129477–3343–109-13-1-1, IR 129477–4026–249-15-1-7, IR 129477–4139–439-1-1-2, IR 129477–4139–439-2-4-2, IR 129477–991–430-1-9-4 and MTU1010 were identified as stable yielder across three tested locations i.e. PAU (Ludhiana), ISARC (Varanasi) and BAU (Sabour) (Fig. S4).

2.2. Genome wide discovery of polymorphism among different donors and recipients

The whole genome resequencing of twenty diverse genotypes (14 donors, 6 recipient background) resulted in a total of 840,724,425 paired end reads of 150 bp (Table S4). The read based GC content estimate ranged from 42 to 45%. A total of 98% of the filtered reads were mapped on the Nipponbare reference genome. The average genome coverage was 98% with the highest in Tadukan (98.55%) and lowest in IR 94225-B-82-B (96.74%). From the high-quality sequences, a total of 96,56,366 SNPs and 13,06,524 InDels were detected. The largest

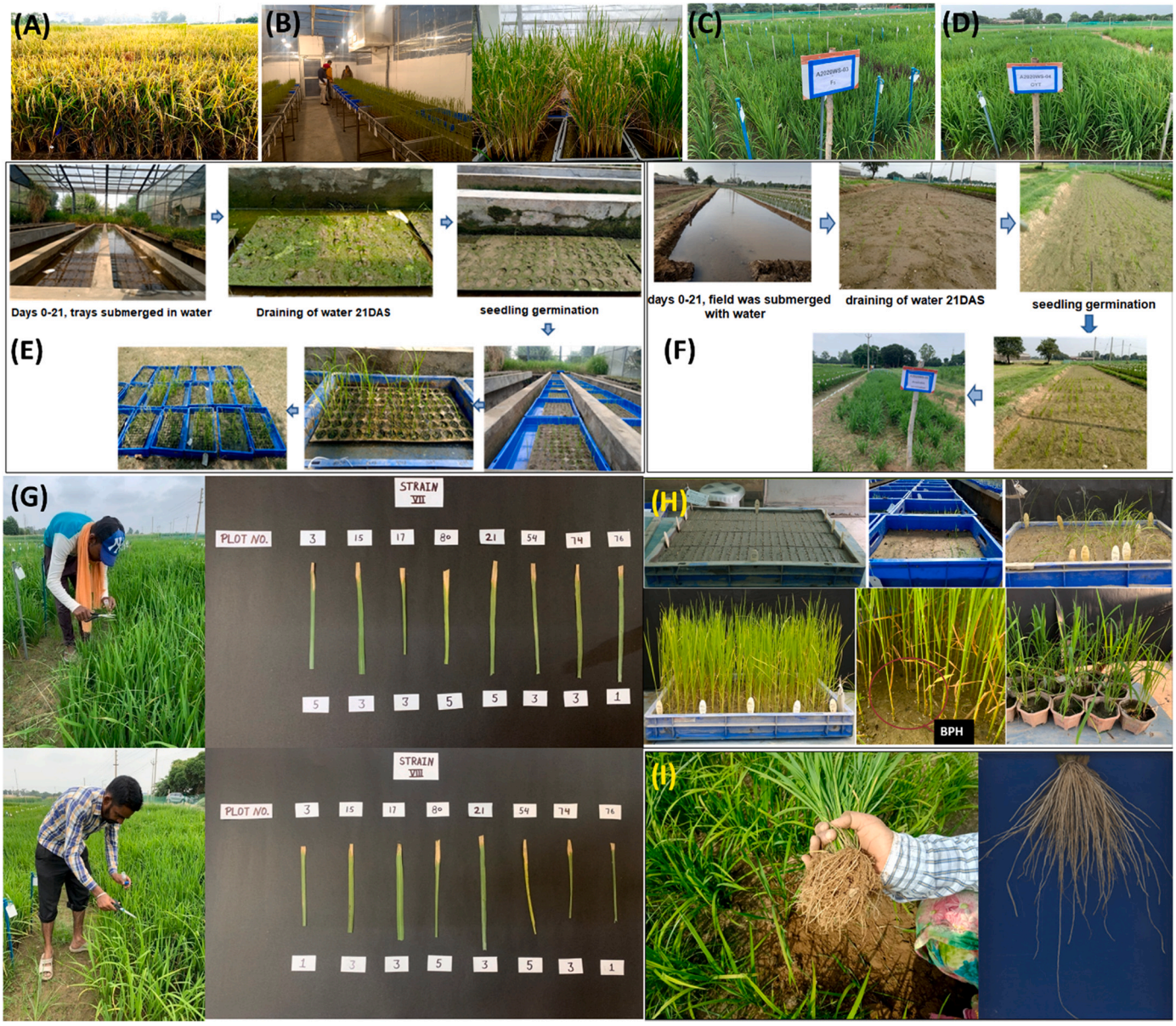

Fig. 1. Phenotypic evaluation of segregating breeding lines and advanced breeding lines for different traits improving biotic stress resistance/tolerance, establishment, grain yield and adaptability of rice under direct seeded cultivation conditions (A) Field view of phenotypic evaluation of F_3 breeding lines (B) Screenhouse view of advancing F_4 generation under controlled conditions (C) Field view of phenotypic evaluation of F_5 breeding lines (D) Field view of phenotypic evaluation of advanced breeding lines procured from IRRI, Varanasi (E) Screening of segregating and advanced breeding panel for anaerobic germination under screen house conditions (F) Screening of segregating and advanced breeding panel for anaerobic germination under field conditions (G) Screening of segregating and advanced breeding panel for bacterial blight under field conditions (H) Screening of segregating and advanced breeding panel for blast in blast nursery (I) Screening of segregating and advanced breeding panel for root traits under field conditions and scanning of roots using WinRhizo STD4800.

number of variants at 10× was detected in *oryza nivara* accession (46,92,725) and minimum in Tadukan (23,30,227). The largest number of SNPs and InDels were identified on chromosome 7 (1,03,511), whereas chromosome 8 harbored the least number of variants (34,819). The chromosome wise distribution of mapped SNPs across whole genome of rice is presented in Fig. S5. The genome sequence of each of the donor for particular traits was compared with each of six recipient backgrounds for the discovery of the SNPs. The designed KASP markers were explicitly informative for the *O. sativa* L. ssp. *indica* rice germplasm constituting 14 donors and 6 recipient backgrounds.

A total of already identified 81 polymorphic SSR markers, one indel, eight gene-specific markers (Sandhu et al. 2021) and 110 polymorphic KASP markers were used for the genotyping and validation purpose (Table S5). Further, the detailed information on the KASP markers associated with particular trait of interest showing polymorphism to each of the six recipient backgrounds is presented in the Table S6. The average physical distance between the two validated KASP markers across the whole genome was 260 kb or ~ 1.066 cM considering 1 cM equal to ~244 kb (Chen et al. 2002). Of the total 110 KASP markers identified and validated from a set of 171 KASP markers designed initially, 107 KASP markers were localized within the MSUv7 gene models (http://rice.plantbiology.msu.edu), and 7 KASP markers were located within the intergenic regions (Table S7). The highest quality SNPs were detected for the KASPs, K_10,607,281 (T - > C 10607281), K_8,791,450 (A - > G 8791450), K_18,666,765 (A - > G 18666765), K_1,029,593 (T - > C 1029593), K_628,981 (C - > T 628981), K_788,995 (T - > A 788995) and K_4,974,371 (G - > C 4974372) in IRGSP1.0 (International Rice Genome Sequencing Project. The clusters for the 61

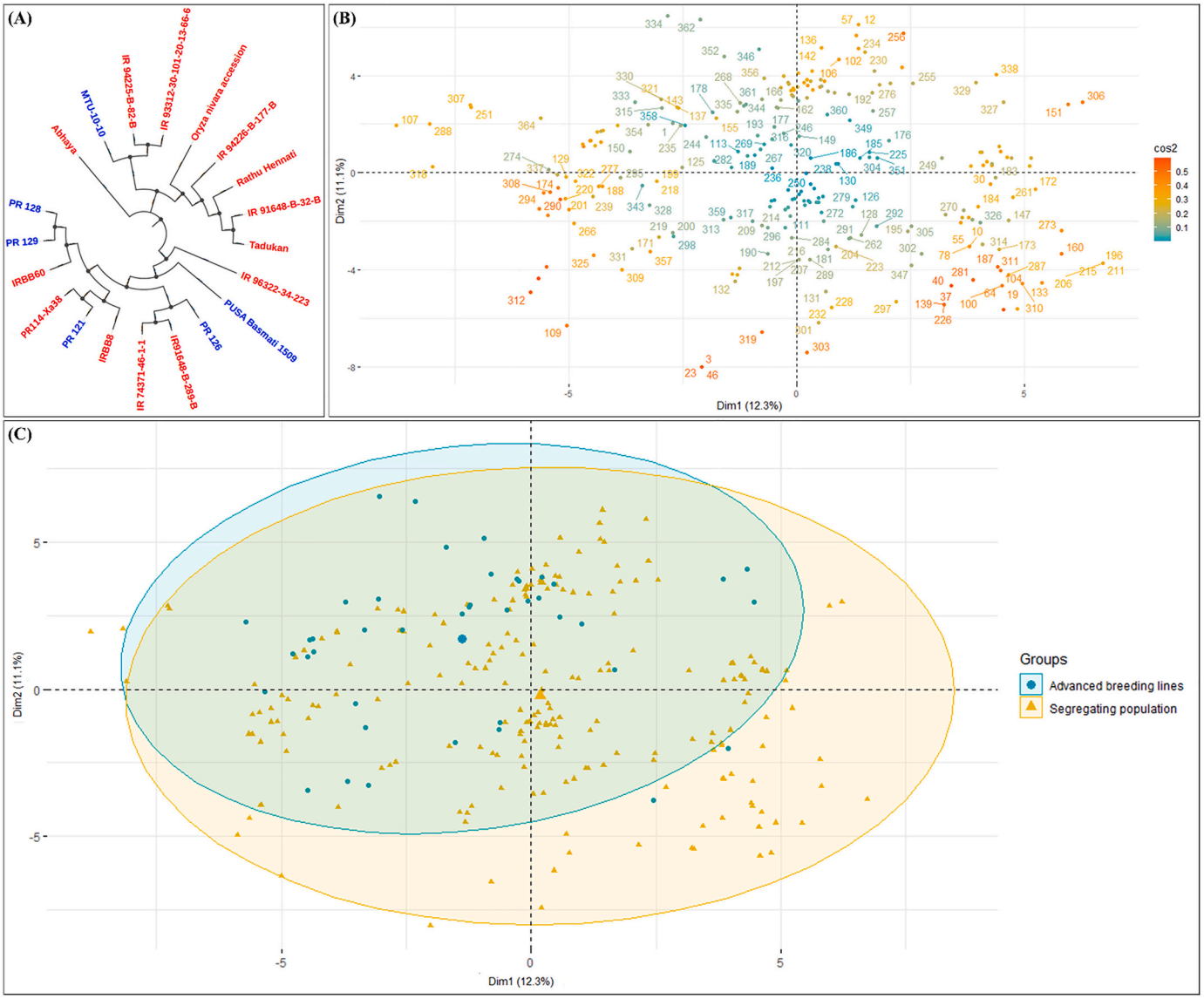

Fig. 2. (A) Genetic diversity analysis of the 20 diverse accessions using the whole genome resequencing data (B, C) Principal component analysis of the 364 breeding lines (322 segregating lines +42 advanced breeding lines) used for the validation of the 110 KASP markers.

KASPs were not well separated properly in the 20 diverse genotypes in parental polymorphism survey and not validated on the F_1 genotypes, therefore not considered in the further analysis.

2.3. Quality control assessment of the KASP markers

The quality control parameters such as the KASP utility, KASP false positive rate (FPR) and the KASP false negative rate' (FNR) were estimated for each of the 110 trait-specific KASP markers (Table 1) on a population size of 384 samples. The utility of the KASP markers ranged from 16.7 to 100%, FPR from 0.00 to 19.20% and FNR from 0.00 to 12.20%. The detailed description of the validated KASP marker analysis results and their allelic interpretation for the selected KASP markers is presented in Table 1. The allelic effects on the phenotypes of the segregating and advanced breeding panel are described for each of the biotic/abiotic stress resistance/tolerance traits, root traits improving nutrient uptake, agronomic, grain yield and yield related traits in Table 1.

2.4. Genetic diversity analysis and Principal component analysis (PCA)

The genetic relationships among donor and recipient parents as determined by UPGMA (unweighted pair group method with arithmetic mean) cluster analysis showed that the 20 rice genotypes were divided into two major groups (Fig. 2A). All the recipients except MTU1010 along with IRBB60, IRBB8, PR114-*Xa38*, IR 74371–46–1-1, IR 96322–34-223 and IR 91468-B-289-B were present in Group I. The remaining donors along with MTU1010 constituted the Group II, which is further divided into two subgroups. The subgroup I had MTU1010, IR 94225-B-82-B and IR 93312–30–101-20-13-66-6 where, the subgroup II had *oryza nivara* accession, IR 94226-B-177-B, Rathu Hennati, IR 91648-B-32-B and Tadukan.

To determine the ability of the KASP markers to access the genetic diversity among the 322 segregating breeding lines and 42 advanced breeding lines derived from the crosses involving multi-parent and to validate the KASP markers, a PCA was performed using the 110 KASP markers. The first principal component (PC1) explained 12.3% and the second principal component (PC2) explained 11.1% of the total genetic variations (Fig. 2B). The segregating and the advanced breeding lines were well distributed across both the principal components (Fig. 2C).

Table 1
The quality control assessment results and allelic effects of the 110 trait-specific KASP markers validated on the phenotypes of the segregating and advanced breeding panel.

SNP ID	Chr	MSU7 pos. (bp)	*QTL/ gene*	Ref allele	Positive allele	Negative trait	Positive trait	KASP utility	KASP FPR	KASP FNR	Segregating breeding lines panel				Advanced breeding lines panel			
											Frequency (%)		Phenotypic mean		Frequency (%)		Phenotypic mean	
											Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait
K_1,594,302	6	1,594,302	*Bph3*	A	G	Susceptible	Resistant	66.7	0.55%	0.00%	80 (24.85%)	220 (68.32%)	9	3	8 (19.05%)	26 (61.90%)	9	3
K_1,606,487	6	1,606,487	*Bph3*	C	T			50.0	19.20%	9.40%	78 (24.22%)	225 (69.88%)	9	3	8 (19.05%)	26 (61.90%)	7	1
K_1,631,244	6	1,631,244	*Bph3*	A	G			66.7	1.80%	0.55%	80 (24.85%)	220 (68.32%)	9	1	8 (19.05%)	27 (64.29%)	9	3
K_1,637,921	6	1,637,921	*Bph3*	A	G			66.7	3.55%	0.00%	80 (24.85%)	220 (68.32%)	7	1	8 (19.05%)	26 (61.90%)	7	1
K_1,693,594	6	1,693,594	*Bph3*	G	A			66.7	2.80%	0.76%	199 (61.80%)	119 (39.96%)	9	1	8 (19.05%)	26 (61.90%)	9	3
K_1,738,798	6	1,738,798	*Bph3*	C	A			66.7	19.10%	0.00%	88 (27.32%)	228 (70.81%)	7	1	8 (19.05%)	27 (64.29%)	9	3
K_5,382,269	4	5,382,269	*Bph17*	G	C			66.7	3.11%	0.00%	67 (20.81%)	221 (68.63%)	9	1	12 (28.57%)	27 (64.29%)	7	1
K_5,650,499	4	5,650,499	*Bph17*	T	C			66.7	0.00%	0.00%	74 (22.98%)	221 (68.63%)	9	3	15 (35.71%)	23 (54.76%)	9	1
K_5,847,738	4	5,847,738	*Bph17*	C	G			66.7	2.11%	0.00%	74 (22.98%)	222 (68.94%)	7	3	15 (35.71%)	25 (59.52%)	9	1
K_6,336,060	4	6,336,060	*Bph17*	G	A			66.7	0.00%	0.00%	80 (24.84%)	220 (68.32%)	9	1	12 (28.57%)	25 (59.52%)	9	3
K_5,585,652	8	5,585,652	*Gm4*	C	T			83.3	–	–	128 (39.75%)	194 (60.25%)	–	–	16 (38.10%)	24 (57.14%)	–	–
K_5,585,772	8	5,585,772	*Gm4*	G	A			83.3	–	–	132 (40.99)	188 (58.39%)	–	–	16 (38.10%)	24 (57.14%)	–	–
K_5,585,944	8	5,585,944	*Gm4*	G	A			83.3	–	–	135 (41.93%)	184 (57.14%)	–	–	16 (38.10%)	24 (57.14%)	–	–
K_5,585,994	8	5,585,994	*Gm4*	G	T			83.3	–	–	128 (39.75%)	194 (60.25%)	–	–	16 (38.10%)	24 (57.14%)	–	–
K_5,586,508	8	5,586,508	*Gm4*	C	T			83.3	–	–	120 (37.27%)	187 (58.07%)	–	–	16 (38.10%)	24 (57.14%)	–	–
K_10,607,281	12	10,607,281	*Pita2*	T	C			100	–	–	309 (95.96%)	13 (4.04%)	–	–	42 (100%)	0 (0.00%)	–	–
K_27,030,975	11	27,030,975	*Xa4*	T	A			66.7	0.00%	2.38%	96 (29.81%)	215 (66.77%)	7	3	5 (11.90%)	37 (88.10%)	7	1
K_27,183,490	11	27,183,490	*Xa4*	G	A			66.7	0.00%	0.00%	102 (31.67%)	216 (67.08%)	7	3	6 (14.29%)	36 (85.71%)	5	3
K_27,357,030	11	27,357,030	*Xa4*	C	T			66.7	0.00%	0.00%	95 (29.50%)	220 (68.32%)	9	3	5 (11.90%)	37 (88.10%)	7	3
K_27,413,024	11	27,413,024	*Xa4*	T	C			66.7	0.00%	0.00%	99 (30.75%)	219 (68.01%)	9	1	6 (14.29%)	36 (85.71%)	7	3
K_27,468,611	11	27,468,611	*Xa4*	C	T			66.7	0.00%	0.00%	100 (31.06%)	215 (66.77%)	7	1	3 (7.14%)	39 (92.86%)	9	3
K_27,489,986	11	27,489,986	*Xa4*	G	T			66.7	0.00%	2.38%	97 (30.12%)	219 (68.01%)	7	1	6 (14.29%)	36 (85.71%)	9	1
K_27,553,861	11	27,553,861	*Xa4*	C	T			66.7	0.00%	0.00%	104 (32.30%)	218 (67.70%	7	3	6 (14.29%)	36 (85.71%)	9	1
K_27,584,500	11	27,584,500	*Xa4*	G	A			66.7	0.00%	2.38%	98 (30.43%)	219 (68.01%)	7	1	4 (9.52%)	38 (9.05%)	7	1
K_27,606,117	11	27,606,117	*Xa4*	C	A			66.7	0.00%	0.00%	97 (30.12%)	219 (68.01%)	9	1	7 (16.67%)	35 (83.33%)	7	3

(*continued on next page*)

Table 1 (*continued*)

SNP ID	Chr	MSU7 pos. (bp)	*QTL/ gene*	Ref allele	Positive allele	Negative trait	Positive trait	KASP utility	KASP FPR	KASP FNR	Segregating breeding lines panel				Advanced breeding lines panel			
											Frequency (%)		Phenotypic mean		Frequency (%)		Phenotypic mean	
											Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait
K_437,375	5	437,375	*xa5*	G	T			100	0.00%	0.00%	141 (43.79%)	175 (54.35%)	9	3	12 (28.57%)	27 (64.29%)	9	3
K_438,206	5	438,206	*xa5*	C	T			100	0.00%	0.00%	150 (46.58%)	168 (52.17%)	9	3	12 (28.57%)	27 (64.29%)	7	1
K_438,578	5	438,578	*xa5*	C	T			100	0.00%	0.00%	141 (43.79%)	175 (54.35%)	9	1	13 (30.95%)	27 (64.29%)	7	1
K_439,616	5	439,616	*xa5*	G	A			100	0.00%	0.00%	141 (43.79%)	175 (54.35%)	7	1	12 (28.57%)	27 (64.29%)	7	1
K_441,684	5	441,684	*xa5*	T	C			10	4.76%	9.50%	130 (40.37%)	188 (55.90%)	7	1	13 (30.95%)	27 (64.29%)	9	1
K_21,274,518	11	21,274,518	*Xa21*	A	G			66.7	0.00%	7.14%	318 (98.76%)	4 (1.24%)	9	1	34 (80.95%)	8 (19.05%)	9	3
K_21,276,435	11	21,276,435	*Xa21*	T	C			66.7	2.40%	7.14%	318 (98.76%)	4 (1.24%)	7	1	34 (80.95%)	8 (19.05%)	9	1
K_26,727,222	8	26,727,222	*xa13*	G	C			66.7	0.00%	4.80%	322 (100%)	0 (0.00%)	9	–	34 (80.95%)	6 (14.29%)	7	1
K_31,579,287	4	31,579,287	*Xa38*	T	C			100	0.00%	0.00%	322 (100%)	0 (0.00%)	9	–	34 (80.95%)	8 (19.05%)	9	3
K_31,579,822	4	31,579,822	*Xa38*	T	C			100	0.00%	0.00%	322 (100%)	0 (0.00%)	9	–	34 (80.95%)	8 (19.05%)	9	1
K_31,580,379	4	31,580,379	*Xa38*	T	A			100	0.00%	0.00%	322 (100%)	0 (0.00%)	7	–	34 (80.95%)	8 (19.05%)	9	1
K_12,253,431	9	12,253,431	$qAG_{9.1}$	T	C	no/very low germination under anaerobic conditions	good germination under anaerobic conditions	66.7	0.00%	0.54%	151 (46.89%)	116 (36.02%)	34.29%	82.86%	17 (40.48%)	17 (40.48%)	36.97%	95.49%
K_12,253,887	9	12,253,887	$qAG_{9.1}$	A	G			66.7	0.00%	0.54%	151 (46.89%)	116 (36.02%)	34.29%	82.86%	17 (40.48%)	18 (42.86%)	36.97%	96.33%
K_1,198,917	5	1,198,917	$qNR_{5.1}$	T	C	less number of nodal roots	more number of nodal roots	66.7	7.15%	2.40%	240 (74.53%)	47 (14.6%)	68	95	30 (71.43%)	11 (26.19%)	165	244
K_1,296,530	5	1,296,530	$qNR_{5.1}$	A	G			83.3	4.76%	2.40%	240 (74.53%)	49 (15.22%)	72	94	31 (73.81%)	11 (26.19%)	174	257
K_1,688,209	5	1,688,209	$qNR_{5.1}$	G	A			16.7	2.40%	0.00%	242 (75.16%)	48 (14.91%)	70	99	31 (73.81%)	11 (26.19%)	170	248
K_1,888,209	5	1,888,209	$qNR_{5.1}$	C	A			66.7	7.15%	0.00%	245 (76.09%)	48 (14.91%)	77	102	31 (73.81%)	11 (26.19%)	177	252
K_5,786,391	4	5,786,391	$qNR_{4.1}$	T	G			100	0.00%	2.40%	277 (86.02%)	45 (13.98%)	65	92	36 (85.71%)	6 (14.29%)	142	222
K_5,787,060	4	5,787,060	$qNR_{4.1}$	A	T			100	0.00%	2.40%	270 (83.85%)	52 (16.15%)	72	95	37 (88.10%)	5 (11.90%)	133	210
K_5,787,840	4	5,787,840	$qNR_{4.1}$	C	G			100	0.00%	0.00%	277 (86.02%)	45 (13.98%)	77	90	37 (88.10%)	5 (11.90%)	133	210
K_5,819,155	4	5,819,155	$qNR_{4.1}$	C	T			100	0.00%	0.00%	277 (86.02%)	45 (13.98%)	77	90	36 (85.71%)	6 (14.29%)	144	225
K_6,494,690	4	6,494,690	$qNR_{4.1}$	T	C			100	0.00%	0.00%	277 (86.02%)	45 (13.98%)	77	90	37 (88.10%)	5 (11.90%)	140	220
K_6,941,654	4	6,941,654	$qNR_{4.1}$	C	T			100	0.00%	0.00%	270 (83.85%)	52 (16.15%)	70	95	36 (85.71%)	6 (14.29%)	139	206
K_7,154,131	4	7,154,131	$qNR_{4.1}$	G	A			100	0.00%	0.00%	270 (83.85%)	52 (16.15%)	77	95	39 (92.86%)	3 (7.14%)	148	228
K_7,073,586	1	7,073,586	$qRHD_{1.1}$	G	T	Sparse root hairs	Dense root hairs	66.7	7.40%	0.00%	69 (21.29%)	238 (73.91%)	12,205	15,652	2 (4.76%)	40 (95.24%)	23,346	33,073
K_7,081,184	1	7,081,184	$qRHD_{1.1}$	C	T			66.7	2.40%	0.00%			12,205	15,652	2 (4.76%)		23,346	33,073

(*continued on next page*)

Table 1 (*continued*)

SNP ID	Chr	MSU7 pos. (bp)	*QTL/ gene*	Ref allele	Positive allele	Negative trait	Positive trait	KASP utility	KASP FPR	KASP FNR	Segregating breeding lines panel				Advanced breeding lines panel			
											Frequency (%)		Phenotypic mean		Frequency (%)		Phenotypic mean	
											Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait
											69 (21.29%)	238 (73.91%)				40 (95.24%)		
K_8,791,450	1	8,791,450	$qRHD_{1.1}$	A	G			33.3	2.40%	0.00%	84 (26.09%)	238 (73.91%)	13,245	15,545	2 (4.76%)	40 (95.24%)	23,346	33,073
K_8,855,279	1	8,855,279	$qRHD_{1.1}$	T	C			33.3	2.40%	0.00%	84 (26.09%)	238 (73.91%)	13,245	15,545	2 (4.76%)	40 (95.24%)	23,346	33,073
K_8,887,013	1	8,887,013	$qRHD_{1.1}$	C	T			33.3	2.40%	0.00%	84 (26.09%)	238 (73.91%)	13,245	15,545	2 (4.76%)	40 (95.24%)	23,346	33,073
K_8,935,224	1	8,935,224	$qRHD_{1.1}$	G	T			33.3	2.40%	0.00%	69 (21.29%)	238 (73.91%)	12,205	15,652	2 (4.76%)	40 (95.24%)	23,346	33,073
K_15,997,367	5	15,997,367	$qRHD_{5.1}$	A	G			100.0	0.00%	2.40%	208 (26.52%)	114 (35.40%)	16,554	18,463	20 (47.62%)	14 (33.33%)	30,641	32,525
K_2,086,213	8	2,086,213	$qRHD_{8.1}$	G	T			100	2.40%	2.40%	318 (98.76%)	4 (1.24%)	15,563	17,455	20 (47.62%)	14 (33.33%)	29,782	31,416
K_9,653,331	2	9,653,331	$qDTY_{2.1}$	A	T	yield penalty under reproductive stage drought stress	Improved yield under reproductive stage drought stress	83.3	0.00%	2.38%	–	–	–	–	38 (9.05%)	3 (7.14%)	4276	4480
K_9,780,606	2	9,780,606	$qDTY_{2.1}$	C	A			83.3	0.00%	2.38%	–	–	–	–	38 (9.05%)	3 (7.14%)	4276	4480
K_30,279,242	3	30,279,242	$qDTY_{3.1}$	G	T			83.3	2.38%	2.38%	–	–	–	–	27 (64.29%)	11 (26.19%)	4151	4688
K_30,270,860	3	30,270,860	$qDTY_{3.1}$	G	A			83.3	2.38%	2.38%	–	–	–	–	27 (64.29%)	11 (26.19%)	4151	4688
K_30,738,342	3	30,738,342	$qDTY_{3.1}$	G	C			83.3	2.38%	0.00%	–	–	–	–	27 (64.29%)	14 (33.33%)	4151	4562
K_30,796,109	3	30,796,109	$qDTY_{3.1}$	T	A			83.3	0.00%	0.00%	–	–	–	–	27 (64.29%)	14 (33.33%)	4151	4562
K_30,821,108	3	30,821,108	$qDTY_{3.1}$	T	C			83.3	0.00%	0.00%	–	–	–	–	27 (64.29%)	14 (33.33%)	4151	4562
K_17,464,660	12	17,464,660	$qDTY_{12.1}$	G	A			83.3	2.37%	4.72%	168 (52.17%)	154 (47.83%)	3001	3356	22 (52.38%)	18 (42.86%)	4468	4915
K_17,486,676	12	17,486,676	$qDTY_{12.1}$	C	T			83.3	3.16%	4.72%	220 (68.32%)	102 (31.68%)	3156	3446	27 (64.29%)	11 (26.19%)	4578	4865
K_17,489,337	12	17,489,337	$qDTY_{12.1}$	T	C			83.3	2.38%	2.38%	220 (68.32%)	102 (31.68%)	2979	3463	27 (64.29%)	11 (26.19%)	4578	4865
K_17,537,562	12	17,537,562	$qDTY_{12.1}$	G	T			83.3	2.38%	2.38%	168 (52.17%)	154 (47.83%)	3001	3356	27 (64.29%)	11 (26.19%)	4578	4865
K_39,538,807	1	39,538,807	$qGY_{1.1}$	G	T	low yield under DSR	Improved yield under DSR	83.3	2.36%	8.67%	273 (84.78%)	35 (10.87%)	3025	3400	27 (64.29%)	3 (7.14%)	4664	5079
K_39,538,868	1	39,538,868	$qGY_{1.1}$	C	T			83.3	3.94%	5.26%	273 (84.78%)	35 (10.87%)	3025	3400	27 (64.29%)	3 (7.14%)	4664	5079
K_39,610,271	1	39,610,271	$qGY_{1.1}$	T	G			83.3	2.40%	6.56%	273 (84.78%)	35 (10.87%)	3025	3400	27 (64.29%)	3 (7.14%)	4664	5079
K_39,612,639	1	39,612,639	$qGY_{1.1}$	G	T			83.3	5.51%	8.67%	273 (84.78%)	35 (10.87%)	3025	3400	27 (64.29%)	3 (7.14%)	4664	5079
K_16,734,396	10	16,734,396	$qGY_{10.1}$	T	C			16.7	3.94%	8.67%	151 (46.89%)	116 (36.03%)	3095	3320	4 (9.52%)	38 (9.05%)	4651	4995
K_18,666,765	10	18,666,765	$qGY_{10.1}$	A	G			16.7	6.56%	7.34%	46 (14.29%)	276 (85.71%)	3012	3465	4 (9.52%)	38 (9.05%)	4651	4995
K_18,703,848	10	18,703,848	$qGY_{10.1}$	C	T			16.7	7.09%	8.14%	46 (14.29%)	276 (85.71%)	3012	3465	4 (9.52%)	38 (9.05%)	4651	4995
K_18,734,396	10	18,734,396	$qGY_{10.1}$	T	C			16.7	4.20%	9.18%			3033	3401	4 (9.52%)		4651	4995

(*continued on next page*)

Table 1 (*continued*)

SNP ID	Chr	MSU7 pos. (bp)	*QTL/ gene*	Ref allele	Positive allele	Negative trait	Positive trait	KASP utility	KASP FPR	KASP FNR	Segregating breeding lines panel				Advanced breeding lines panel			
											Frequency (%)		Phenotypic mean		Frequency (%)		Phenotypic mean	
											Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait
											47 (14.70%)	275 (85.40%)				38 (9.05%)		
K_18,749,935	10	18,749,935	$qGY_{10.1}$	T	C			16.7	8.67%	5.77%	46 (14.29%)	276 (85.71%)	3012	3465	4 (9.52%)	38 (9.05%)	4651	4995
K_18,797,856	10	18,797,856	$qGY_{10.1}$	A	T			16.7	0.52%	1.05%	46 (14.29%)	276 (85.71%)	3012	3465	4 (9.52%)	38 (9.05%)	4651	4995
K_18,901,381	10	18,901,381	$qGY_{10.1}$	G	A			16.7	3.94%	5.26%	47 (14.70%)	273 (84.78%)	3033	3401	4 (9.52%)	38 (9.05%)	4651	4995
K_19,779,370	10	19,779,370	$qGY_{10.1}$	C	G			50.0	4.20%	2.40%	47 (14.70%)	273 (84.78%)	3033	3401	23 (54.76%)	19 (45.24%)	4476	4893
K_19,909,715	10	19,909,715	$qGY_{10.1}$	G	A			83.3	5.51%	5.77%	47 (14.70%)	275 (85.40%)	3033	3401	23 (54.76%)	19 (45.24%)	4476	4893
K_20,036,149	10	20,036,149	$qGY_{10.1}$	C	T			83.3	11.81%	9.97%	89 (27.64%)	225 (69.88%)	3109	3328	23 (54.76%)	19 (45.24%)	4476	4893
K_456,591	3	456,591	$qLDG_{3.1}$	G	A	susceptible to lodging	resistant to lodging	83.3	0.00%	0.00%	104 (32.30%)	203 (63.04%)	105	87	10 (23.81%)	29 (69.04%)	108	90
K_499,461	3	499,461	$qLDG_{3.1}$	T	C			66.7	0.00%	0.00%	104 (32.30%)	203 (63.04%)	105	87	11 (26.19%)	31 (73.81%)	108	92
K_610,549	3	610,549	$qLDG_{3.1}$	C	G			83.3	1.83%	2.36%	104 (32.30%)	203 (63.04%)	105	87	11 (26.19%)	31 (73.81%)	108	92
K_648,480	3	648,480	$qLDG_{3.1}$	T	C			66.7	0.08%	4.70%	104 (32.30%)	203 (63.04%)	105	87	11 (26.19%)	31 (73.81%)	108	92
K_673,659	3	673,659	$qLDG_{3.1}$	T	C			33.3	0.00%	0.00%	189 (58.70%)	133 (41.30%)	101	85	11 (26.19%)	31 (73.81%)	108	92
K_767,400	3	767,400	$qLDG_{3.1}$	G	A			83.3	0.00%	0.00%	104 (32.30%)	203 (63.04%)	105	87	11 (26.19%)	31 (73.81%)	108	92
K_1,029,593	3	1,029,593	$qLDG_{3.1}$	T	C			83.3	0.00%	0.00%	104 (32.30%)	203 (63.04%)	105	87	11 (26.19%)	31 (73.81%)	108	92
K_1,083,767	3	1,083,767	$qLDG_{3.1}$	C	T			83.3	0.00%	0.00%	189 (58.70%)	133 (41.30%)	101	85	11 (26.19%)	31 (73.81%)	108	92
K_1,111,246	3	1,111,246	$qLDG_{3.1}$	G	C			83.3	0.00%	0.00%	189 (58.70%)	133 (41.30%)	101	85	11 (26.19%)	31 (73.81%)	108	92
K_1,179,119	3	1,179,119	$qLDG_{3.1}$	G	C			83.3	0.00%	0.00%	189 (58.70%)	133 (41.30%)	101	85	11 (26.19%)	31 (73.81%)	108	92
K_1,298,111	3	1,298,111	$qLDG_{3.1}$	C	T			83.3	0.00%	0.00%	188 (58.39%)	133 (41.30%)	101	85	11 (26.19%)	31 (73.81%)	108	92
K_16,279,477	4	16,279,477	$qLDG_{4.1}$	T	A			100	2.40%	0.00%	289 (89.76%)	33 (10.25%)	105	82	38 (9.05%)	4 (9.52%)	96	89
K_16,402,594	4	16,402,594	$qLDG_{4.1}$	G	A			100	2.40%	0.00%	289 (89.76%)	33 (10.25%)	105	82	38 (9.05%)	3 (7.14%)	98	92
K_17,734,727	4	17,734,727	$qLDG_{4.1}$	G	A			100	0.00%	0.00%	289 (89.76%)	33 (10.25%)	105	82	40 (95.24%)	2 (4.76%)	97	92
K_18,522,689	4	18,522,689	$qLDG_{4.1}$	G	A			100	0.00%	3.12%	289 (89.76%)	33 (10.25%)	101	88	38 (9.05%)	3 (7.14%)	98	92
K_19,027,599	4	19,027,599	$qLDG_{4.1}$	T	G			100	0.00%	2.49%	268 (83.23%)	40 (12.42%)	101	83	40 (95.24%)	2 (4.76%)	97	92
K_19,221,785	4	19,221,785	$qLDG_{4.1}$	C	T			100	0.00%	3.12%	268 (83.23%)	40 (12.42%)	101	83	42 (100%)	0 (0.00%)	98	–
K_19,760,969	4	19,760,969	$qLDG_{4.1}$	G	A			100	0.00%	2.18%	268 (83.23%)	40 (12.42%)	101	83	42 (100%)	0 (0.00%)	98	–
K_264,232	1	264,232	$qEUE_{1.1}$	C	T	poor emergence		100	2.62%	8.85%		12 (3.73%)	55%	91%			58%	95%

(*continued on next page*)

Table 1 (*continued*)

SNP ID	Chr	MSU7 pos. (bp)	*QTL/ gene*	Ref allele	Positive allele	Negative trait	Positive trait	KASP utility	KASP FPR	KASP FNR	Segregating breeding lines panel				Advanced breeding lines panel			
											Frequency (%)		Phenotypic mean		Frequency (%)		Phenotypic mean	
											Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait	Negative trait	Positive trait
							early and uniform emergence				309 (95.96%)				31 (73.81%)	11 (26.19%)		
K_628,981	1	628,981	$qEUE_{1.1}$	C	T			100	3.93%	3.12%	309 (95.96%)	9 (2.80%)	55%	91%	31 (73.81%)	11 (26.19%)	58%	95%
K_728,679	1	728,679	$qEUE_{1.1}$	G	T			100	4.72%	3.93%	310 (96.27%)	10 (3.11%)	48%	87%	31 (73.81%)	11 (26.19%)	58%	95%
K_742,759	1	742,759	$qEUE_{1.1}$	T	C			100	0.00%	12.20%	307 (95.34%)	11 (3.42%)	52%	85%	29 (69.04%)	10 (23.81%)	55%	96%
K_780,655	1	780,655	$qEUE_{1.1}$	G	T			100	4.99%	5.51%	311 (96.58%)	10 (3.11%)	50%	88%	31 (73.81%)	10 (23.81%)	58%	95%
K_788,995	1	788,995	$qEUE_{1.1}$	T	A			100	3.67%	3.93%	313 (97.20%)	9 (2.80%)	53%	92%	31 (73.81%)	10 (23.81%)	58%	95%
K_4,033,148	11	4,033,148	$qEUE_{11.1}$	C	T			100	2.62%	4.72%	116 (36.02%)	206 (63.98%)	66%	82%	31 (73.81%)	10 (23.81%)	58%	95%
K_4,974,371	11	4,974,371	$qEUE_{11.1}$	T	C			100	4.12%	5.78%	102 (31.68%)	202 (62.73%)	60%	80%	31 (73.81%)	10 (23.81%)	58%	95%
K_5,786,143	11	5,786,143	$qEUE_{11.1}$	C	T			100	3.41%	4.47%	116 (36.02%)	206 (63.98%)	66%	82%	31 (73.81%)	10 (23.81%)	58%	95%
K_6,581,858	11	6,581,858	$qEUE_{11.1}$	A	G			100	3.93%	7.89%	116 (36.02%)	206 (63.98%)	66%	82%	31 (73.81%)	10 (23.81%)	58%	95%

Chr: chromosome, bp: base pair, Ref allele: allele present in the reference genome, positive allele: allele present in the donor parent, negative trait: trait present in the recipient parent, positive trait: targeted trait present in the donor parent, FPR: false positive rates, FNR: false negative rates, frequency (%) negative trait: number (percent to the total) of the breeding lines possessing recipient parent allele, frequency (%) positive trait: number (percent to the total) of the breeding lines possessing donor parent allele, phenotypic mean negative trait: mean value of the breeding lines possessing recipient parent allele, phenotypic mean positive trait: mean value of the breeding lines possessing donor parent allele,

SNP utility: the percentage of a prospective breeding pool across which the SNP marker could be used to introgress a QTL/gene or the proportion of the breeding pool which does NOT carry the donor allele of the SNP marker. It was calculated as: number of breeding lines without favorable allele/Total number of breeding lines assessed, False Positive Rate' (FPR): the proportion of breeding lines with recipient allele but identified as not having an unfavorable/recipient allele of the SNP marker. It was calculated as the number of breeding lines withOUT recipient allele/Total number of breeding lines with recipient allele, False Negative Rate (FNR): the proportion of breeding lines with donor allele but identified as not having the desired QTL/donor allele. It was calculated as: # number of breeding lines with-OUT favorable allele/Total number of breeding lines with donor allele.

The most of the advanced breeding lines grouped exclusively in the upper left side of the PC1 vs. PC2 scatter plot while the segregating breeding lines were scattered across the scatter plot (Fig. 2C).

2.5. Comparison of the KASP markers with conventional SSR markers

The detailed information on the conventional SSR markers and KASP markers used in the present study is presented in Table S5. The 110 KASP assays produced consistent results when compared to the conventional PCR based SSR markers in 20 diverse rice genotypes (Fig. S6). The similar allelic pattern was observed in the parental polymorphism assay and during the genotyping of the segregating and advanced breeding lines when tested with both SSR and KASP markers.

2.6. Phenotypic validation of the KASP markers

All the KASP markers which produced satisfactory results in the 20 diverse genotypes in parental polymorphism survey were validated against the phenotypes except assays for the *Gm4* as phenotypic screening data was not available. The allelic patterns of the 20 diverse accessions for the identified KASP associated with biotic stress resistance/tolerance traits; anaerobic germination, early-uniform seedling emergence, root traits associated with nutrient uptake, lodging resistance; and grain yield under reproductive stage drought stress, under DSR are presented in Figs. 3A, 3B and 3C respectively. The allelic effects on the mean phenotypic values of the segregating and advanced breeding lines panel are described in Table 1. The effects of all the alleles were significant at $P \leq 0.05$ in segregating and the advanced breeding panel. The 110 phenotypically validated KASP assays include 36 assays for biotic resistance genes (6 for *Bph3,* 4 for *Bph17,* 5 for *Gm4,* 1 for *Pita2,* 9 for *Xa4,* 5 for *xa5,* 2 for *Xa21,* 1 for *xa13,* 3 for *Xa38*), 2 assays for anaerobic germination ($qAG_{9.1}$), 19 assays for the root traits (7 for $qNR_{4.1}$, 4 for $qNR_{5.1}$, 6 for $qRHD_{1.1}$, 1 for $qRHD_{5.1}$ and 1 for $qRHD_{8.1}$) (Table 1, Fig. 4). The remaining assays include 11 assays for grain yield under reproductive stage drought stress (2 for $qDTY_{2.1}$, 5 for $qDTY_{3.1}$, 5 for $qDTY_{12.1}$) and 14 assays for grain yield under DSR (4 for $qGY_{1.1}$ and 10 for $qGY_{10.1}$), 18 assays for lodging resistance trait (11 for $qLDG_{4.1}$ and 7 for $qLDG_{3.1}$) and 10 assays for the early uniform emergence trait (6 for $qEUE_{1.1}$ and 4 for $qEUE_{11.1}$) (Table 1, Fig. 4). The few examples of KASP assays on the 20 diverse accessions including parents that were used to develop the breeding panel and KASP assays on the breeding panel (segregating and advanced breeding lines) are presented in Fig. 5.

On average, the assays for biotic resistance genes clearly distinguished the resistant (1: highly resistant, 3: resistant) and susceptible (7: susceptible, 9: highly susceptible) breeding lines (Table 1). The breeding lines carrying the alleles for anaerobic germination tolerance showed improved germination (83 to 96%) compared to breeding lines carrying reference alleles (34 to 37%) (Table 1). The phenotypic effect of the alleles associated with root traits showed significant variations in number of nodal roots and root hair density (Table 1). The alleles

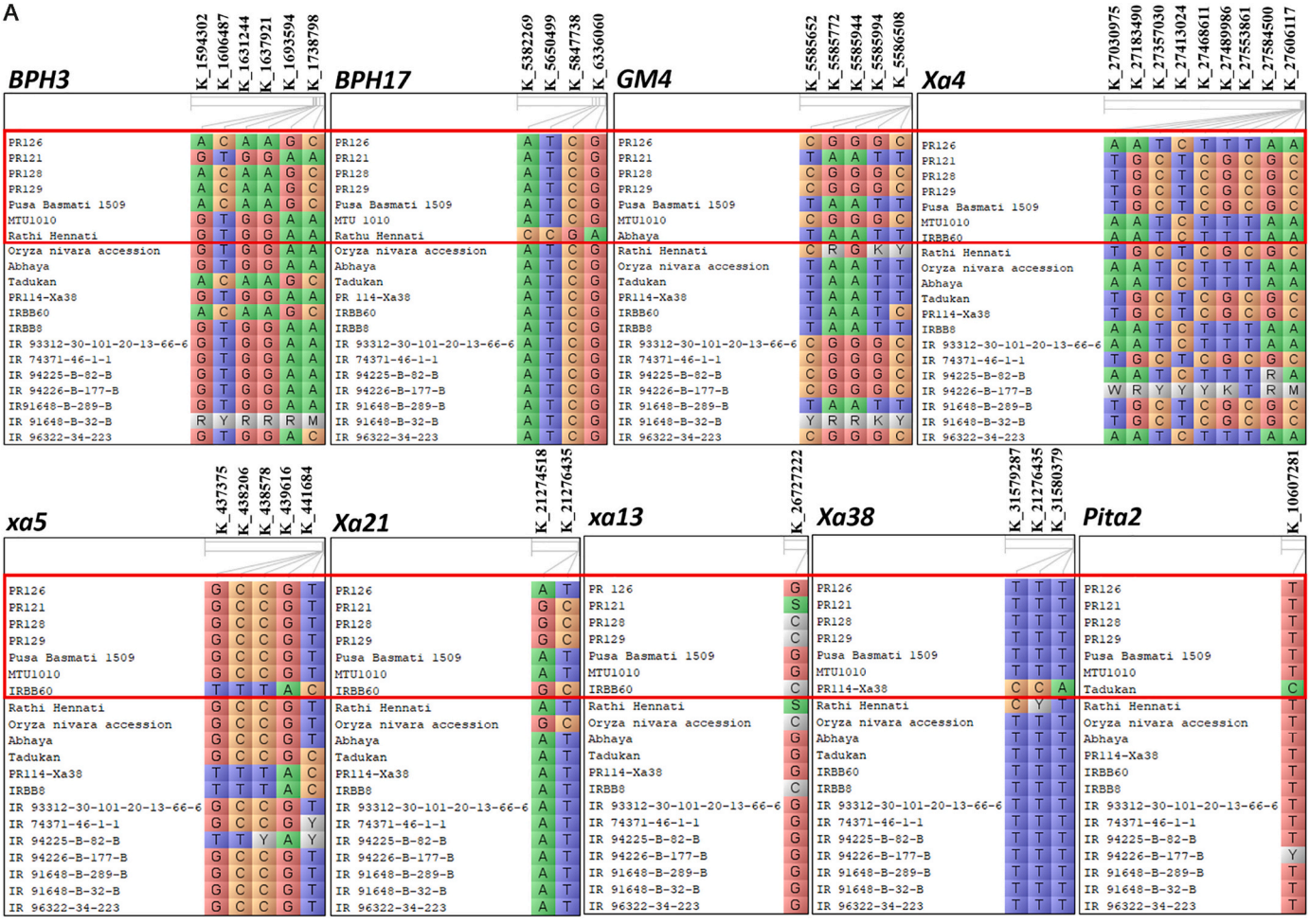

Fig. 3A. The allelic constitution of the 20 diverse accessions for the identified KASP associated with the biotic stress resistance/tolerance traits such as brown plant hopper (*Bph3, Bph17*), gall midge (*GM4*), bacterial blight (*Xa4, xa5, Xa21, xa13, Xa38*) and blast resistance (*Pita2*) traits. The accessions in the red box represents the six recipient backgrounds (PR126, PR121, PR128, PR129, Pusa Basmati 1509, MTU1010) and every seventh accession in each Fig. is the donor for the respective trait. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

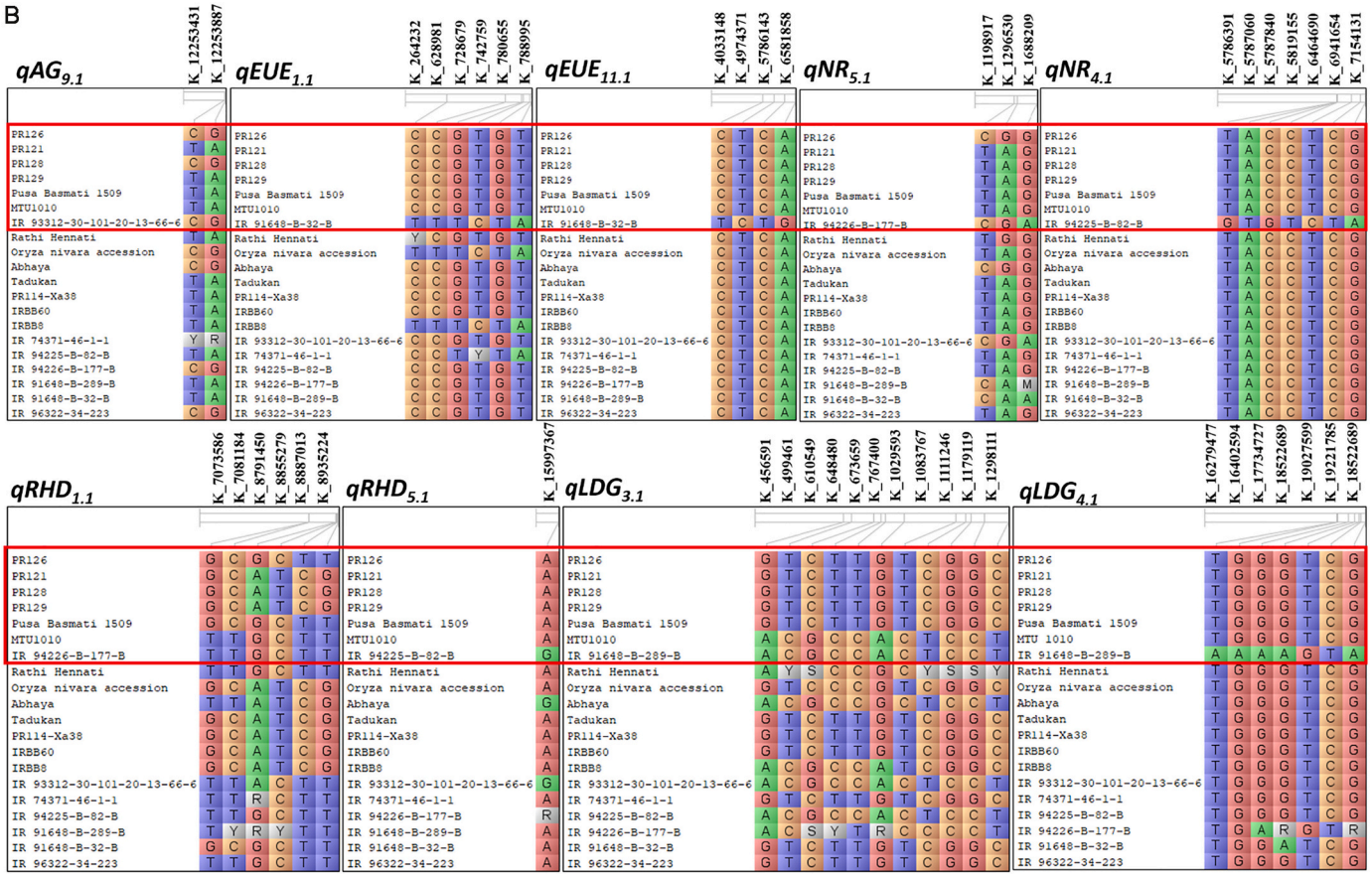

Fig. 3B. The allelic constitution of the 20 diverse accessions for the identified KASP associated with the anaerobic germination ($qAG_{9.1}$), early-uniform seedling emergence ($qEUE_{1.1}$, $qEUE_{11.1}$), root traits associated with nutrient uptake ($qNR_{5.1}$, $qNR_{4.1}$, $qRHD_{1.1}$, $qRHD_{5.1}$) and lodging resistance ($qLDG_{3.1}$, $qLDG_{4.1}$) traits under direct seeded cultivation conditions. The accessions in the red box represents the six recipient backgrounds (PR126, PR121, PR128, PR129, Pusa Basmati 1509, MTU1010) and every seventh accession in each Fig. is the donor for the respective trait. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

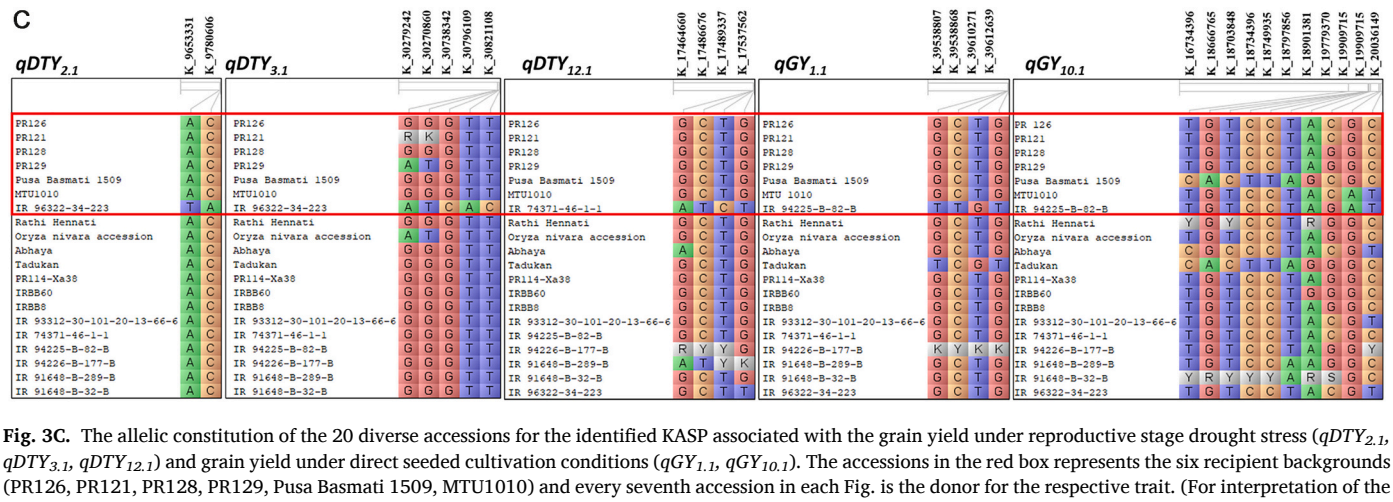

Fig. 3C. The allelic constitution of the 20 diverse accessions for the identified KASP associated with the grain yield under reproductive stage drought stress ($qDTY_{2.1}$, $qDTY_{3.1}$, $qDTY_{12.1}$) and grain yield under direct seeded cultivation conditions ($qGY_{1.1}$, $qGY_{10.1}$). The accessions in the red box represents the six recipient backgrounds (PR126, PR121, PR128, PR129, Pusa Basmati 1509, MTU1010) and every seventh accession in each Fig. is the donor for the respective trait. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

associated with the $qLDG_{3.1}$, $qLDG_{4.1}$ showed significant reduction in the plant height and the alleles associated with $qGY_{1.1}$, $qGY_{10.1}$ improved the grain yield under DSR. Favorable alleles for *Bph3*, *Bph17*, *Gm4*, *Xa4*, *xa5*, $qRHD_{1.1}$, $qGY_{10.1}$ and $qLDG_{3.1}$ were present at high frequency in both the segregating and the advanced breeding lines panel (Table 1, Fig. 6).

2.7. Reliability of the developed KASP markers

The identification and validation of KASP markers is necessary for the high-throughput and cost-effective screening of large population developed involving multiple parents for different biotic and abiotic stress resistance. Therefore, it is very much important to validate the KASP markers in different ways. In the present study, the developed

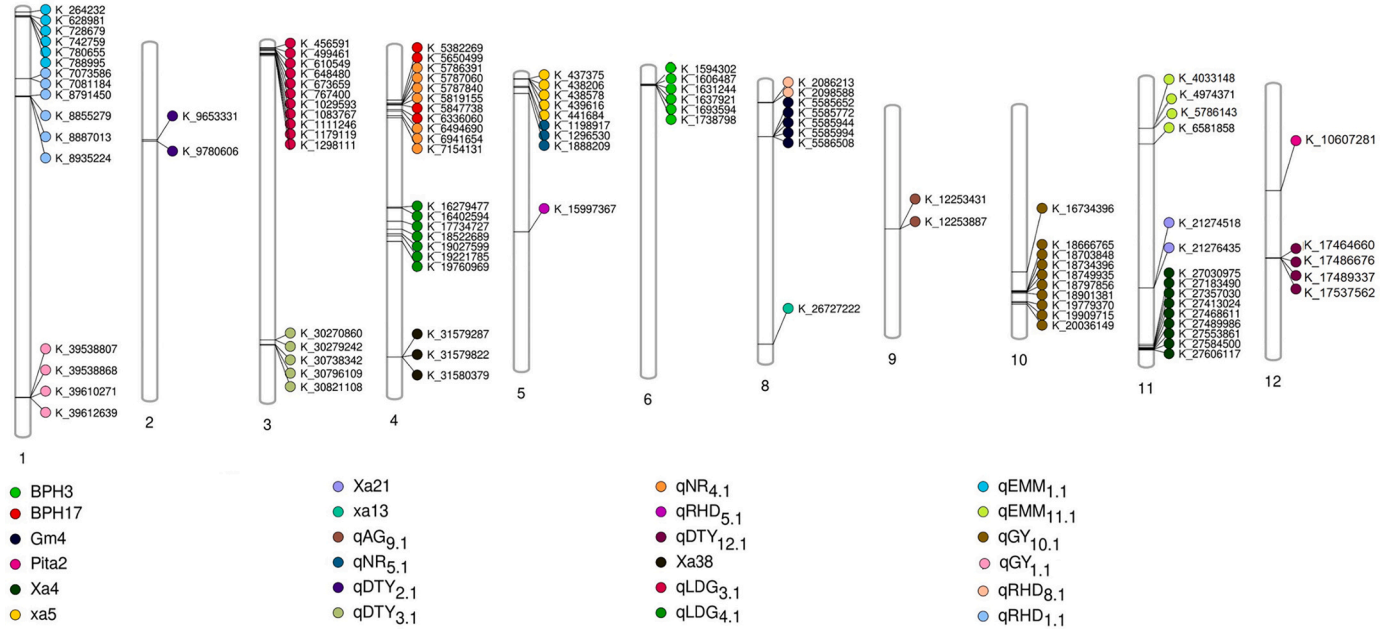

Fig. 4. Schematic representation of the distribution of validated KASP assays associated with different biotic/abiotic resistance/tolerance trait, root traits improving nutrient uptake, seedling establishment, yield and yield related traits along the 12 chromosomes of rice. The alternate SNP ID (K_followed by numeric value) showing genomic position in base pairs representing the physical position of the SNPs on the chromosome. The numbers below each chromosome indicate chromosome numbers. The. The 110 phenotypically validated KASP assays include 36 assays for biotic resistance genes, 2 assays for anaerobic germination, 10 assays for the early uniform emergence trait, 19 assays for the root traits improving nutrient uptake, 11 assays for grain yield under reproductive stage drought stress, 14 assays for grain yield under DSR, and 18 assays for lodging resistance trait.

KASP markers were validated first on a set of 20 diverse parents followed by the second level validation on a set of segregating breeding material and third level validation on advanced breeding material. Further, at the fourth level, the KASP markers were checked on the 10 predicted F_{1s} plants developed for each trait considered in the present study. At the fifth level, the KASP markers were validated by comparing the KASP assays with the already available gel-based PCR markers associated with the DSR traits. In addition, the repeatability of the KASP assays was accessed on a set of random 70 samples using 10 random markers.

3. Discussion

An increase in rice productivity through the introgression of multiple traits for abiotic/biotic stresses as well as traits improving adaptability under DSR cultivation conditions is a feasible breeding strategy to adapt with changing climate, limited resources and to develop high-yielding DSR varieties. The development of marker-assisted derived DSR breeding lines with pyramided QTL/genes with different specificities or broad-spectrum QTL may be expected to increase grain yield, durability and adaptability under DSR. The newly identified QTLs/genes for traits related to DSR adaptation [6,18,19,48,49] shall provide opportunities to develop new rice varieties with higher yield potential of 6.0–6.5 tha^{-1} in the background of popular recipient varieties and broader adaptability to diverse DSR situations. This means that substantial opportunities exist for the exploitation of these QTL/genes in the modern breeding programs, and efforts are underway to achieve this. To date, the genomics-assisted breeding programs attempting to use the identified QTL/genes have almost universally relied on the old SSR marker systems [50]. SSRs are not so useful in marker-assisted introgression programs involving multiple donors as there is possibility of getting the same allelic pattern for multiple parents. The identification of a cost-effective, high-throughput and straightforward genotyping approach that does not jeopardize the prediction accuracies is must.

The introduction of new genomics tools and markers offer great solution in meeting the challenge of genetic gain improvement [39,43] ensuring productivity under DSR. The explosion in genomics resources and the low-cost genome resequencing approaches now enables the development of highly accurate SNP marker systems. To the best of our knowledge this is the first study targeting development of trait-based SNP panel for the traits improving yield and adaptability of rice under DSR. The concepts of marker-assisted selection are now transformed into the high-throughput and cost-effective whole genome sequencing platform to maximize the genetic gains especially for complex traits [51,52]. While a multiple of genotyping platforms have been developed, very less efforts have been made in developing informative, cost-effective and high-throughput genotyping solution specifically designed for applied DSR breeding programs. The existing SNP databases such as the Rice 3 K project provides about 18 million different types of SNP information which creates difficulty for the breeders to get quick analysis and useful information about the SNPs [35]. Consequently, it is very much important in the genomics-assisted breeding to identify the core trait-linked SNPs representing the key information. Not much research work has been done in developing tightly linked, functional and diagnostic SNP markers to be used directly in molecular breeding except the development of diagnostic markers such as *ALK* [53,54], *Wx* [55,56], *GS3* [57], *Pikh*, *GW5*, and *CHALK5* [34], *rtsv1* [58], *Xa4* [59], *xa5* [60,61], *xa13* [62], *Xa23* [63], *Xa21* [64], *Xa7* [61,65], and Sub1A [66].

The KASP-SNP genotyping panel underpinning polymorphism among different rice genera [67], linked with genes associated with the economic traits in crop breeding [68], and disease resistance have been developed. The approach of the present study was to bring in a large number of molecularly characterized strains amenable to rapid transfer and consolidation along with high-throughput genotyping solution. We demonstrated here the effectiveness of newly developed KASP assays for QTL/genes conferring adaptability, biotic and abiotic stress resistances and grain yield stability of rice under DSR. Unlike the previously developed SNP arrays [39–41,43], the set of KASP markers developed in the present study target specific QTL/genes of interest, and first provide a platform for the foreground MAS for multiple DSR related QTL/genes.

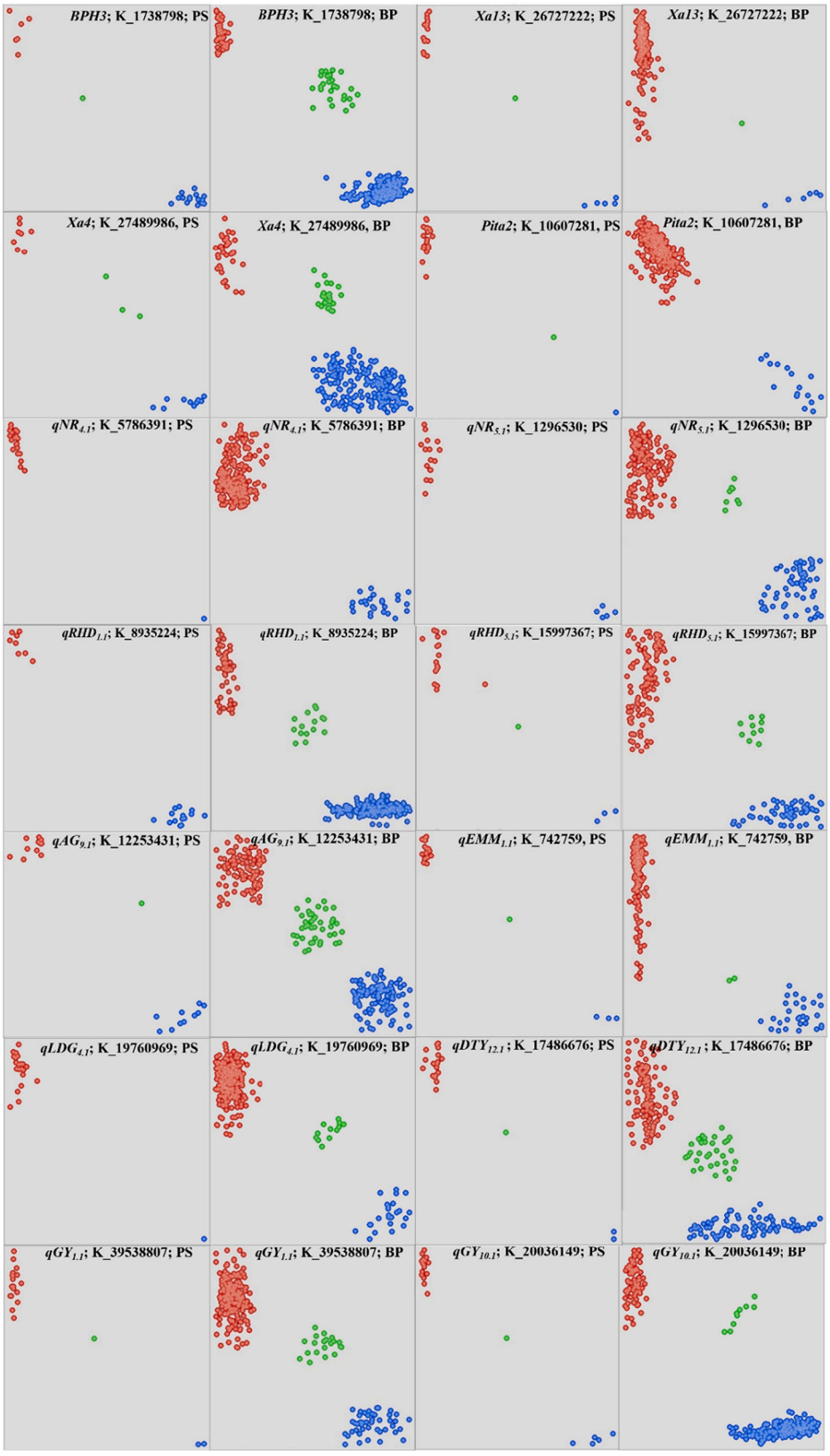

Fig. 5. The pictorial representation of the KASP assays conducted on the 20 diverse accessions including parents used to develop the breeding panel and KASP assays on the breeding panel constituting segregating and advanced breeding lines. PS: polymorphism survey on the 20 diverse accessions; BP: breeding panel. Blue colour indicates the donor allele, red colour indicates the recipient allele and green colour indicates the heterozygotes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

One hundred and seventy-one KASP markers classified into five subsets including biotic stress resistance, early seedling establishment (anaerobic germination, early uniform emergence), root traits improving nutrient uptake (number of nodal roots, root hair density), plant type (lodging resistance) and grain yield (reproductive stage drought stress and DSR conditions) were designed. The designed 171 markers fulfilled the criterion of quality control, allelic variation of targeted donor to other 13 donors and to 6 recipient backgrounds, key/functional genes targeting sites and strong association with the important DSR trait. The major challenge in designing KASP markers was to identify the SNPs

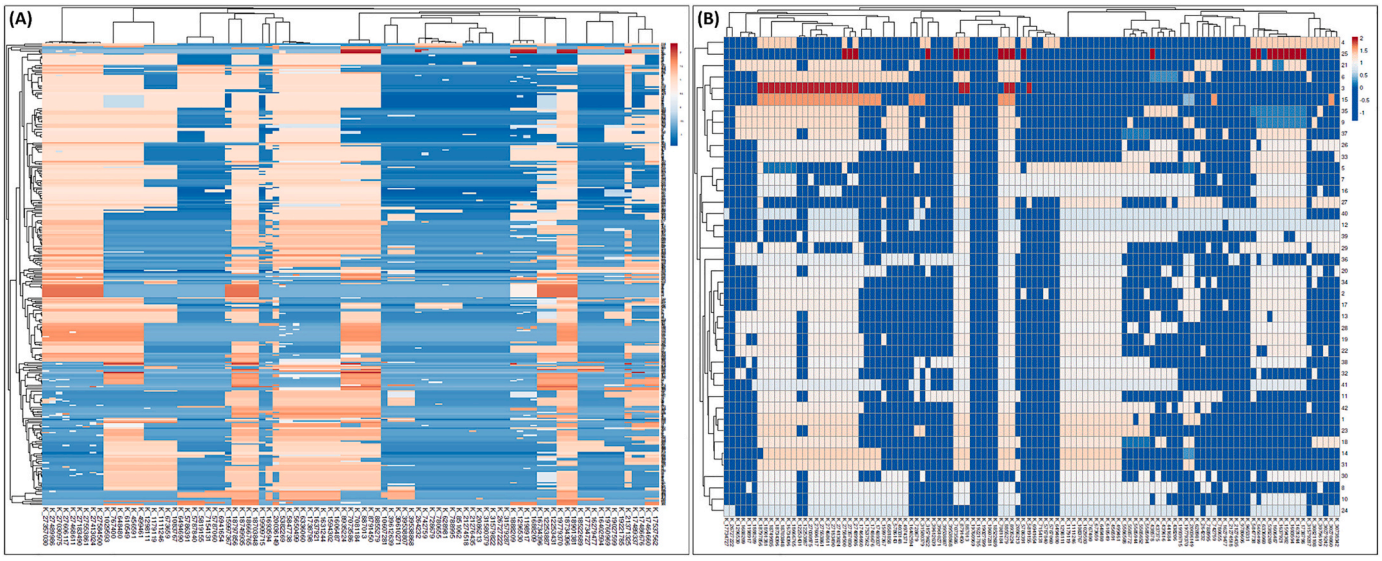

Fig. 6. The heat map indicating the frequency of favorable alleles associated with different biotic/abiotic resistance/tolerance trait, root traits improving nutrient uptake, seedling establishment, yield and yield related traits in (A) the segregating and (B) the advanced breeding lines panel.

specifically linked with the particular donor and trait of interest as the marker-assisted introgression program involve multiple parents. Finally, 110 out of the 171 successfully-designed SNPs were able to display diversity at the particular loci.

The KASP assays analyzed in the present study were validated for their robustness and reliability at five different levels using diverse germplasm, segregating population, advanced breeding lines, comparison with PCR based markers and on the developed F_{1s}. Overall, the newly developed KASP assays worked for most of the targeted genomic regions with few exceptions like *qEVV*$_{9.1}$ (early uniform emergence) and the genomic regions associated with nematode tolerance. Several KASP assays may need to be developed to reveal the good clusters and to get reliable results. All the 110 KASP assays being reported here showed significant association with the relevant phenotypes in the diverse germplasm, segregating and advanced breeding populations panel, thus revealing their potential application in DSR breeding programs.

The KASP array developed in the present study may be useful in constructing a set of nearly isogenic lines (NILs) suitable for the DSR trait evaluation because the identified SNPs can be used to select favorable alleles in a wide range of genetic backgrounds. The tightly linked set of SNPs such as the SNPs detected for *Gm4*, *Xa4*, *xa5*, *xa13*, *Xa21*, *Xa38*, *qAG*$_{9.1}$, *qDTY*$_{3.1}$, *qDTY*$_{12.1}$ and *qGY*$_{1.1}$ can also be used for dissecting the "linkage drag". The detection of haplotypes around the target QTL/genes can further be utilized for the fine genetic dissection of the genomic regions near the targeted QTL/genes. As we have large segregating population available with us, the validated KASP array can be readily used to construct a NIL with a very small introgressed chromosome segment from the respective donor parents and a MAGIC (multi-parent advanced generation inter-cross) population with multiple useful genes. The simultaneous detection of multiple useful alleles using KASP arrays potentially allows a dramatic decrease in the time required and labor cost to develop pyramid lines. In addition, the direct pyramiding of multiple abiotic and biotic stress tolerance/resistance related genes will contribute in increasing the genetic diversity in the breeding pool. However, the cost for KASP assay is still expensive but the multiplexing of both samples and markers may reduce the cost. The versatile KASP array developed in the present study will largely contribute in facilitating the DSR breeding activities. Both the plant materials and the KASP arrays are available for the rice breeders to improve the already existing rice varieties.

KASP technology has recently emerged as a single-plex high-throughput genotyping technology and this is the first report on high throughput screening of KASP markers associated with traits improving adaptability, grain yield and biotic/abiotic stress resistance/tolerance under DSR in a major crop. A set of core SNPs was built via targeting variations in the already identified genomic region associated with DSR traits. Development and validation of such 110 KASP assays may provide useful ways to deploy these important QTLs/genes in DSR breeding programs in a timely and cost-effective manner. The development of trait-based SNP panel may strengthen DSR breeding programs and enable the broader distribution and the active use of SNP based markers for both the public and private sectors marker-assisted DSR breeding programs. It may allow the access of trait-based markers to the DSR breeders to screen their germplasm for the DSR adapted traits, including traits providing grain yield improvement and tolerance to biotic and abiotic stresses. Building on these achievements, development of DSR rice varieties and the development of trait-based KASP-SNP markers can provide opportunities for significant increase in DSR productivity. Hence, we believe that this toolkit can significantly accelerate the efforts to select and pyramid favorable alleles/genes in DSR breeding programs.

4. Materials and methods

The study on the development and validation of KASP markers was carried out at School of Agricultural Biotechnology, Punjab Agricultural University, Ludhiana, Punjab (India). In the past decades, efforts have been made in the identification of donors/genomic regions associated with the traits improving grain yield and adaptability of rice under DSR cultivation conditions. A panel of diverse accessions consist of 14 donors (listed in Table S8) and six recipient backgrounds (PR126, PR121, PR128, PR129, PB1509, and MTU1010 were chosen to test the diagnostic ability of SSR and KASP markers. Two marker-assisted derived breeding populations (segregating and advanced breeding populations) were also used for the validation of markers. A total 108 marker-assisted derived F_3 breeding lines developed through forward breeding approach were procured from IRRI, South Asia Hub, Hyderabad in 2019WS. These breeding lines were evaluated and advanced to F_4 generation under DSR at PAU and a total of 599 single plant selection were selected in 2019WS. The 599 single plant selection were sown in controlled screenhouse conditions during 2020DS and advanced to F_5 generation. The 599 breeding lines and the local and upland adapted checks were evaluated under DSR cultivation conditions and advanced to F_6 generation in 2020WS. The 599 breeding lines were genotyped using the identified polymorphic SSR markers. Considering the genotyping cost of KASP

markers in mind, a total of 322 breeding lines segregating for different combinations of QTLs/genes were screened for KASP assay along with the 14 donors and 6 recipient backgrounds. The detailed information on the donors and the breeding strategy used to develop the breeding lines at IRRI-SAH is presented in Fig. S1. In addition, seeds of 42 advanced breeding lines developed at International Rice Research Institute (IRRI), Philippines (Sandhu et al. 2021) were procured and evaluated at PAU, ISARC (IRRI-South Asia regional center), Varanasi and BAU, Sabour under DSR. The F_1s were developed utilizing the donors possessing the particular trait of interest in six recipient backgrounds to further check the efficacy of designed markers in determining the heterozygosity (Fig. S2).

4.1. Phenotypic characterization of the breeding populations

To combine traits providing grain yield and adaptability under DSR, true F_1s were selected using trait-linked markers. A plot size of 3.2, 1.6, 1.6, 2.0 and 3.6 m^2 was maintained at F_3, F_4, F_5, F_6 and in advanced generations, respectively; maintaining 20 cm (hill to hill) × 20 cm (row to row) distance across the two replications. The complete package of field management used was described in detail in Sandhu et al. [18]. Depending on the generation, six plants were randomly chosen to record data on days to 50% flowering (days), plant height (cm, at physiological maturity), panicle length (cm), number of grains/panicles, number of tillers/0.5 m^2 and grain yield of single plant selection (g). The plot yield was measured in kg ha^{-1}. The method for the measurement of all the observed agronomic traits was followed from Sandhu et al. [18]. The root traits measurement on six random plants were attempted following the procedure described in detail in Sandhu et al. [6] at 60 days after sowing was done. At F_5 and advanced generation stages, screening for bacterial blight (6 plants per breeding lines), brown planthopper (BPH) (7 to 10 seedlings per breeding lines) and anaerobic germination (7 to 10 seeds per breeding lines) were attempted. Mixed inoculum for the races presents in Punjab was used for the blast inoculation. For the bacterial blight screening, the most prevalent pathotypes *PbXo-7* and *PbXo-8* were used. The pure virulent culture of each pathotype was inoculated separately at the maximum tillering stage using clip inoculation technique [69]. The scoring for bacterial blast was carried out following Narayanan et al. [70] and for BPH following Heinrichs [71]. The screening for anaerobic germination was carried out in screenhouse as well as in the field following Angaji et al. [72]. Already identified donors and local check varieties were used for positive and negative control in each of the screening experiments.

4.2. Genotyping

4.2.1. Whole genome resequencing

Genomic DNA of the 14 donors and 6 recipient backgrounds were prepared using modified CTAB method [73], integrity was analyzed on gel electrophoresis and then subjected to high throughput whole genome ReSequencing using Illumina HiSEQ 4000. In the sequencing analysis, gDNA library was constructed using Illumina Truseq protocol v3 and 150 bp paired-end short reads were generated in fastq format. A total of 4 Gb raw sequence data was generated. The obtained raw data were processed using the following procedure. The schematic representation of the complete workflow used for the development of core trait-linked KASP marker panel to be further used in genomics-assisted breeding program is presented in Fig. S3.

4.2.2. Sequencing, read processing and read alignment

Paired-end sequencing, using the Illumina HiSeq 4000 platform, and read processing were carried out at NGB Diagnostics Private Limited, New Delhi (India). For bioinformatics analysis, Illumina adaptor sequences were removed and quality trimming of adaptor-clipped reads was performed, removing reads containing Ns and 3′-end trimming reads to get a minimum average Phred quality score of 20 over a window of ten bases. Reads with a final length of less than 20 bases were discarded.

The reference genome sequence used was of *O. sativa* (version 7.0), retrieved from Rice genome annotation project (http://rice.plantbiology.msu.edu/ pub/ data/ Eukaryotic Projects/osativa/ annotationdbs/ pseudomolecules/version_7.0/all.dir/). Sequencing reads were mapped against this reference using bwa (version 0.7.17-r1188). Discordant or mixed paired-read alignments were not permitted, with all other alignment parameters kept as default. Only read pairs with both reads aligning in the expected orientation were used in subsequent analyses.

4.2.3. Variant calling

SAMtools (version 0.1.19) [74] were used for conversion of mapping files from Sam alignment format into bam binary format and the duplicates were marked for the sorted bam files using Picard software (version 1.48). Bam file generated was then used for variant calling using software Unified Genotyper of GATK pipeline (Genome Analysis Toolkit, version 3.6). For comparative analysis and identification of unique SNPs in the donor parent variant files for all samples were merged using Bcftools (version 1.9) and samples with MAF 2% and 80% were kept and finally, filtering for variant calling was done using Vcftools (version 0.1.17).

4.2.4. Designing of KASP markers

KASP markers were designed using offline Polymarker software [75]. MAFFT, Primer3, Exonerate and, Blast software's, Samtools, Bamtools, Bio-samtools and, Glib 2.0 were used in the system's path. Database for the reference genome was generated using BLAST tool and indexing of reference genome was done using samtools to generate an index file for the genome. Variant calls for the specified gene/QTL regions were retrieved from the VCF files generated for SNP calling. Flanking region for the SNPs extracted from the reference genome using bedtools and the final files for marker designing were created in the format desired for the Polymarker softer, containing ID, chromosome number, and the variant call having flanking region of 100 bp at each side in the CSV (Comma-separated values) format. These files were used as input files for the Polymarker software.

4.2.5. Filtering and selection of KASP markers

The markers located in the earlier identified genomic region associated with the traits considered in the present study were screened. All gene files for the reference genome were retrieved from Rice Genome Annotation Project. Selected markers then screened for high donor specificity using the merged variant file created using BCF tools. All the markers shortlisted were aligned with the reference genome using BLAST (*Basic Local Alignment Search Tool)* and markers showing alignment at multiple loci were rejected. High specificity markers aligning at the desired locus with low e-value were selected.

4.2.6. SSR and KASP assay

To test the utility and accuracy of the SSR and KASP markers, the genomic DNA of both sets of the breeding lines was extracted. Parental polymorphism survey using the SSR and KASP markers was performed. A total of already identified 81 polymorphic SSR markers, one indel, eight gene-specific markers [48] and 171 newly designed KASP markers were used for the polymorphic survey of the donor and the recipient backgrounds used in the present study. SSR marker assay was carried out following the procedure as described by Sandhu et al. [48]. The KASP genotyping assays were tested and further validated in a 384-well format having a total reaction volume of 4 μl (2 μl of template DNA (25 ng), 0.056 μl of primer mix and 1.944 μl of Kasp mix). Touchdown PCR was performed with the following configuration: Initial denaturation at 95 °C for 15 min, 10 touchdown cycles (95 °C for 20s, touchdown at 65 °C, − 1 °C per cycle, 25 s) and then 20 cycles of amplification (95 °C for 10 s; 57 °C for 60 s). Fluorescence data was collected using the infinite F200 pro micro-plate reader and analyzed using the Tecan i-

control 1.11 software and clusters were marked (XX, XY, YY) based on their graphical location using the KlusterCaller.

4.2.7. Hierarchical clustering and principal component analysis

Hierarchical clustering of the diverse accessions based on whole genome resequencing data was generated. The genetic distance matrix was calculated using TASSEL v.5.2.37 (Bradbury et al., 2007) and the tree was visualized using iTOL [76]. The principal component analysis (*PCA) was performed* using prcomp command of the R statistical software (http://www.R-project.org) [77].

4.2.8. Quality assurance of the KASP markers

The SNP quality control methods and variables [(i) utility (ii) false positive rate; FPR (iii) false negative rate; FNR)] as described by Platten et al. [78] were used to check the ability of the newly designed trait linked KASP markers. It leads to the accurate identification of the plant samples with desirable alleles. The parameters (utility, FPR and FNR) were measured and analyzed separately for each individual trait associated marker and/or for trait associated haplotypes in case the trait with more than one KASP marker associated with it.

4.3. Statistical analysis

Analysis of variance (ANOVA), experiment and experiment-wise mean was calculated using mixed model analysis in PBTools V 1.4.0. To evaluate the phenotypic stability and grain yield adaptability of the breeding lines across seasons and locations, the yield stability analysis was performed. The location effect was considered as an "environment (E)." The best linear unbiased prediction (BLUP) values of the G (genotype) and G x E (genotype x environment) effects were calculated. The multiplicative model in PB tool version 1.3 (bbi.irri.orgbbi.irri.org) was used to explain the relationship between genotype and seasons. The stability models were fitted into the genotype × environment means within a mixed-model framework where the effect of the genotypes was considered as fixed and the trials were random [79–81]. The allelic effect of all the significant markers associated with all the traits measured in the present study was determined comparing the mean phenotypic values and the significant allelic variations for the particular traits were determined performing the Kruskal–Wallis test in "R".

Author contributions

N.S. conceptualized the study; N.S., J.S., G.S., M.S., M.P., G.P., O.P.R. and R.K. performed the experiments; N.S. and J.S. participated in bioinformatic analyses; N.S. wrote the paper; P.S.S. performed phenotyping of brown plant hopper; J.S.L. performed phenotyping of bacterial blight; U.M.S. and S.D. performed experiments at IRRI, Varanasi; D.B.S. and V. K.S. were involved in development of segregating breeding lines at IRRI, Hyderabad; S.P.S. performed experiments at BAU, Sabour; R.K. and V.S. provides critical inputs in phenotyping; N.S. and A.K. were involved in development of advanced breeding lines at IRRI Philippines. All authors contributed in finalizing the original draft of the manuscript.

Declaration of Competing Interest

No conflict of interest is declared.

Acknowledgments

We acknowledge the Department of Biotechnology and BBSRC, GCRF Global Research Translation Awards, UKRI, UK for providing financial support to the present study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ygeno.2022.110269.

References

[1] R.A. Wing, M.D. Purugganan, Q. Zhang, The rice genome revolution: from an ancient grain to green super Rice, Nat. Rev. Genet. 19 (8) (2018) 505–517.
[2] D.K. Ray, N.D. Mueller, P.C. West, J.A. Foley, Yield trends are insufficient to double global crop production by 2050, PLoS One 8 (6) (2013), e66428.
[3] A. Anandan, J. Meher, R.P. Sah, S. Samantaray, C. Parameswaran, P. Panneerselvam, S.K. Dash, P. Swain, M. Annamalai, S.R. Prabhukarthikeyan, G. Kumar, Enhancing input use efficiency in direct-seeded rice with classical and molecular breeding, in: Rice Research for Enhancing Productivity, Profitability and Climate Resilience, 2018, pp. 73–89.
[4] V. Kumar, J.K. Ladha, Direct seeding of rice: recent developments and future research needs, Adv. Agron. 111 (2011) 297–413.
[5] B.A.M. Bouman, Y. Xiaoguang, W. Huaqi, W. Zhiming, Z. Junfang, W. Changgui, C. Bin, May. Aerobic rice (Han Dao): a new way of growing rice in water-short areas, in: Proceedings of the 12th International Soil Conservation Organization Conference 26, Tsinghua University Press, Beijing, China, 2002, p. 31, 175–181.
[6] N. Sandhu, R.O. Torres, M.T. Sta Cruz, P.C. Maturan, R. Jain, A. Kumar, A. Henry, Traits and QTLs for development of dry direct-seeded rainfed rice varieties, J. Exp. Bot. 66 (1) (2015) 225–244.
[7] K. Yano, T. Ookawa, K. Aya, Y. Ochiai, T. Hirasawa, T. Ebitani, Isolation of a novel lodging resistance QTL gene involved in strigolactone signaling and its pyramiding with a QTL gene involved in another mechanism, Mol. Plant 8 (2015) 303–314.
[8] L. Comas, S. Becker, V.M.V. Cruz, P.F. Byrne, D.A. Dierig, Root traits contributing to plant productivity under drought, Front. Plant Sci. 4 (2013) 442.
[9] H. Pathak, A.N. Tewari, S. Sankhyan, D.S. Dubey, U. Mina, V.K. Singh, N. Jain, A. Bhatia, Direct-seeded rice: potential, performance and problems–a review, Curr. Adv. Agric. Sci. 3 (2) (2011) 77–88.
[10] T. Azhiri-Sigari, H. Gines, L.S. Sebastian, L. Wade, Seedling vigor of rice cultivars in response to seeding depth and soil moisture, Phil. J. Crop. Sci. 30 (1) (2005) 53–58.
[11] H. Poorter, R.O.B. De Jong, A comparison of specific leaf area, chemical composition and leaf construction costs of field plants from 15 habitats differing in productivity, New Phytol. 143 (1) (1999) 163–176.
[12] B. Shipley, Net assimilation rate, specific leaf area and leaf mass ratio: which is most closely correlated with relative growth rate? A meta-analysis, Funct. Ecol. 20 (2006) 565–574.
[13] M. Dingkuhn, D.E. Johnson, A. Sow, A.Y. Audebert, Relationships between upland rice canopy characteristics and weed competitiveness, Field Crop Res. 61 (1) (1999) 79–95.
[14] D.L. Zhao, G.N. Atlin, L. Bastiaans, J.H.J. Spiertz, Comparing rice germplasm groups for growth, grain yield and weed-suppressive ability under aerobic soil conditions, Weed Res. 46 (2006) 444–445.
[15] S. Ghosal, C. Casal, F.A. Quilloy, E.M. Septiningsih, M.S. Mendioro, S. Dixit, Deciphering genetics underlying stable anaerobic germination in rice, phenotyping, QTL identification, and interaction analysis, Rice 12 (1) (2019) 50.
[16] S. Dixit, A. Grondin, C.R. Lee, A. Henry, T.M. Olds, A. Kumar, Understanding rice adaptation to varying agro-ecosystems, trait interactions and quantitative trait loci, BMC Genet. 16 (1) (2015) 86.
[17] S. Yadav, U.M. Singh, S.M. Naik, C. Venkateshwarlu, P.J. Ramayya, K.A. Raman, N. Sandhu, A. Kumar, Molecular mapping of QTLs associated with lodging resistance in dry direct-seeded rice (*Oryza sativa* L.), Front. Plant Sci. 8 (2017) 1431.
[18] N. Sandhu, S.R. Subedi, V.K. Singh, P. Sinha, S. Kumar, S.P. Singh, S.K. Ghimire, M. Pandey, R.B. Yadaw, R.K. Varshney, A. Kumar, Deciphering the genetic basis of root morphology, nutrient uptake, yield, and yield-related traits in rice under dry direct-seeded cultivation systems, Sci. Rep. 9 (1) (2019) 1–16.
[19] S.R. Subedi, N. Sandhu, V.K. Singh, P. Sinha, S. Kumar, S.P. Singh, S.K. Ghimire, M. Pandey, R.B. Yadaw, R.K. Varshney, A. Kumar, Genome-wide association study reveals significant genomic regions for improving yield, adaptability of rice under dry direct seeded cultivation condition, BMC Genomics 20 (1) (2019) 471.
[20] S. Qu, G. Liu, B. Zhou, M. Bellizzi, L. Zeng, L. Dai, B. Han, G.L. Wang, The broad-spectrum blast resistance gene Pi9 encodes a nucleotide-binding site–leucine-rich repeat protein and is a member of a multigene family in rice, Genetics 172 (3) (2006) 1901–1914.
[21] H. Sato, H. Hirabayashi, Y. Takeuchi, S. Arase, J. Kihara, H. Kato, T. Imbe, H. Nemoto, QTL analysis of brown spot resistance in rice (*Oryza sativa* L.), Breed. Sci. 58 (1) (2008) 93–96.
[22] I. Ullah, S. Jamil, M.Z. Iqbal, H.L. Shaheen, S.M. Hasni, S. Jabeen, A. Mehmood, M. Akhter, Detection of bacterial blight resistance genes in basmati rice landraces, Genet. Mol. Res. 11 (3) (2012) 1960–1966.
[23] J. Jairin, K. Phengrat, S. Teangdeerith, A. Vanavichit, T. Toojinda, Mapping of a broad-spectrum brown planthopper resistance gene, Bph3, on rice chromosome 6, Mol. Breed. 19 (1) (2007) 35–44.

[24] S. Nair, A. Kumar, M.N. Srivastava, M. Mohan, PCR-based DNA markers linked to a gall midge resistance gene, Gm4t, has potential for marker-aided selection in rice, Theor. App. Genet. 92 (6) (1996) 660–665.
[25] K. Das, D. Zhao, D. De Waele, R.K.S. Tiwari, D.K. Shrivastava, A. Kumar, Reactions of traditional upland and aerobic rice genotypes to rice root knot nematode (Meloidogyne graminicola), J. Plant Breed. Crop Sci. 3 (7) (2011) 131–137.
[26] Y. Koide, N. Kobayashi, D. Xu, Y. Fukuta, Resistance genes and selection DNA markers for blast disease in rice (*Oryza sativa* L.), Japan Agric. Res. 43 (4) (2009) 255–280.
[27] N. Bandillo, C. Raghavan, P.A. Muyco, M.A.L. Sevilla, I.T. Lobina, C.J. Dilla-Ermita, C.W. Tung, S. McCouch, M. Thomson, R. Mauleon, R.K. Singh, Multi-parent advanced generation inter-cross (MAGIC) populations in rice: progress and potential for genetics research and breeding, Rice 6 (1) (2013) 11.
[28] B.C. Collard, D.J. Mackill, Marker-assisted selection, an approach for precision plant breeding in the twenty-first century, Phil, Trans. Royal Society B Bio. Sci. 363 (1491) (2008) 557–572.
[29] Z. Tian, C. Yan, Q. Qian, S. Yan, H. Xie, F. Wang, J. Xu, G. Liu, Y. Wang, Q. Liu, S. Tang, Development of gene-tagged molecular markers for starch synthesis-related genes in rice, Chin. Sci. Bull. 55 (33) (2010) 3768–3777.
[30] R. Tripathi, K. Kumar, R. Balachandar, P. Marimuthu, M. Varghese, S. Bharath, Cognitive markers of mild cognitive impairment: an Indian experience, Ann. Indian Acad. Neurol. 19 (1) (2016) 164.
[31] M.J. Thomson, High-throughput SNP genotyping to accelerate crop improvement, Plant Breed. Biotechnol. 2 (3) (2014) 195–212.
[32] G. Ramkumar, G.D. Prahalada, S.L. Hechanova, R. Vinarao, K.K. Jena, Development and validation of SNP-based functional codominant markers for two major disease resistance genes in rice (*O. sativa* L.), Mol. Breed. 35 (6) (2015) 1–11.
[33] Y. Kurokawa, T. Noda, Y. Yamagata, R. Angeles-Shim, H. Sunohara, K. Uehara, T. Furuta, K. Nagai, K.K. Jena, H. Yasui, A. Yoshimura, Construction of a versatile SNP array for pyramiding useful genes of rice, Plant Sci. 242 (2016) 131–139.
[34] G. Yang, S. Chen, L. Chen, K. Sun, C. Huang, D. Zhou, Y. Huang, J. Wang, Y. Liu, H. Wang, Z. Chen, Development of a core SNP arrays based on the KASP method for molecular breeding of rice, Rice 12 (1) (2019) 1–8.
[35] J.D. Arbelaez, M.S. Dwiyanti, E. Tandayu, K. Llantada, A. Jarana, J.C. Ignacio, J. D. Platten, J. Cobb, J.E. Rutkoski, M.J. Thomson, T. Kretzschmar, 1k-RiCA (1K-Rice custom amplicon) a novel genotyping amplicon-based SNP assay for genetics and breeding applications in rice, Rice 12 (1) (2019) 1–15.
[36] A. Scheben, J. Batley, D. Edwards, Revolution in genotyping platforms for crop improvement, in: Plant Genetics and Molecular Biology, Springer, Cham, 2018, pp. 37–52.
[37] R.J. Elshire, J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, S. E. Mitchell, A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species, PLoS One 6 (5) (2011), e19379.
[38] K. Zhao, M. Wright, J. Kimball, G. Eizenga, A. McClung, M. Kovach, W. Tyagi, M. L. Ali, C.W. Tung, A. Reynolds, C.D. Bustamante, S.R. McCouch, Genomic diversity and introgression in *O. sativa* reveal the impact of domestication and breeding on the rice genome, PLoS One 5 (2010), e10780.
[39] H. Chen, H. He, Y. Zou, W. Chen, R. Yu, X. Liu, Y. Yang, Y.M. Gao, J.L. Xu, L. M. Fan, Y. Li, Development and application of a set of breeder-friendly SNP markers for genetic analyses and molecular breeding of rice (*Oryza sativa* L.), Theor. App. Genet. 123 (6) (2011) 869–879.
[40] H. Yu, W. Xie, J. Li, F. Zhou, Q. Zhang, A whole-genome SNP array (RICE6K) for genomic breeding in RICE, Plant Biotechnol. J. 12 (2014) 28–37.
[41] M.J. Thomson, N. Singh, M.S. Dwiyanti, D.R. Wang, M.H. Wright, F.A. Perez, G. DeClerck, J.H. Chin, G.A. Malitic-Layaoen, C.J. Juanillas, R. Dilla-Ermita, T. Mauleon, S.R. McCouch Kretzschmar, Large-scale deployment of a rice 6 K SNP array for genetics and breeding applications, Rice 10 (2017) 40.
[42] S.R. McCouch, M.H. Wright, C.W. Tung, L.G. Maron, K.L. McNally, M. Fitzgerald, N. Singh, G. DeClerck, F. Agosto-Perez, P. Korniliev, A.J. Greenberg, Open access resources for genome-wide association mapping in rice, Nat. Commun. 7 (1) (2016) 1–14.
[43] H. Chen, W. Xie, H. He, H. Yu, W. Chen, J. Li, R. Yu, Y. Yao, W. Zhang, Y. He, X. Tang, A high-density SNP genotyping array for rice biology and molecular breeding, Mol. Plant 7 (3) (2014) 541–553.
[44] N. Singh, P.K. Jayaswal, K. Panda, P. Mandal, V. Kumar, B. Singh, S. Mishra, Y. Singh, R. Singh, V. Rai, A. Gupta, Single-copy gene based 50 K SNP chip for genetic studies and molecular breeding in rice, Sci. Rep. 5 (2015) 11600.
[45] K. Zhao, C.W. Tung, G.C. Eizenga, M.H. Wright, M.L. Ali, A.H. Price, G.J. Norton, M.R. Islam, A. Reynolds, J. Mezey, A.M. McClung, C.D. Bustamante, S.R. McCouch, Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*, Nat. Commun. 2 (2011) 467.
[46] F.A. Feltus, J. Wan, S.R. Schulze, J.C. Estill, N. Jiang, A.H. Paterson, An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments, Genome Res. 14 (9) (2004) 1812–1819.
[47] K.L. McNally, K.L. Childs, R. Bohnert, R.M. Davidson, K. Zhao, V.J. Ulat, G. Zeller, R.M. Clark, D.R. Hoen, T.E. Bureau, R. Stokowski, Genome wide SNP variation reveals relationships among landraces and modern varieties of rice, Proceed. Nat. Acad. Sci. 106 (30) (2009) 12273–12278.
[48] N. Sandhu, S. Yadav, M. Catolos, M.T.S. Cruz, A. Kumar, Developing climate-resilient, direct-seeded, adapted multiple-stress-yolerant rice applying genomics-assisted breeding, Front. Plant Sci. 12 (2021), 637488.
[49] N. Sandhu, K.A. Raman, R.O. Torres, A. Audebert, A. Dardou, A. Kumar, A. Henry, Rice root architectural plasticity traits and genetic regions for adaptability to variable cultivation and stress conditions, Plant Physiol. 171 (4) (2016) 2562–2576.
[50] J.N. Cobb, R.U. Juma, P.S. Biswas, J.D. Arbelaez, J. Rutkoski, G. Atlin, T. Hagen, M. Quinn, E.H. Ng, Enhancing the rate of genetic gain in public-sector plant breeding programs: lessons from the breeder's equation, Theor. App. Genet. 132 (3) (2019) 627–645.
[51] K. Semagn, R. Babu, S. Hearne, M. Olsen, Single nucleotide polymorphism genotyping using Kompetitive allele specific PCR (KASP): overview of the technology and its application in crop improvement, Mol. Breed. 33 (1) (2014) 1–14.
[52] Y. Zhao, M.F. Mette, M. Gowda, C.F. Longin, J.C. Reif, Bridging the gap between marker-assisted and genomic selection of heading time and plant height in hybrid wheat, Heredity 112 (2014) 638–645.
[53] Z. Gao, Map-based cloning of the ALK gene, which controls the gelatinization temperature of rice, Sci. China Ser. Comm. 46 (2003) 661.
[54] J.S. Bao, H. Corke, M. Sun, Nucleotide diversity in starch synthase IIa and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.), Theor. App. Genet. 113 (7) (2006) 1171–1183.
[55] M. Dobo, N. Ayres, G. Walker, W.D. Park, Polymorphism in the GBSS gene affects amylose content in US and European rice germplasm, J. Cereal Sci. 52 (3) (2010) 450–456.
[56] B. Teng, Y. Zhang, S. Du, J. Wu, Z. Li, Z. Luo, J. Yang, Crystalline, thermal and swelling properties of starches from single-segment substitution lines with different Wx alleles in rice (Oryza sativa L.), J. Sci. Food Agric. 97 (2017) 108–114.
[57] N. Takano-Kai, H. Jiang, T. Kubo, M. Sweeney, T. Matsumoto, H. Kanamori, B. Padhukasahasram, C. Bustamante, A. Yoshimura, K. Doi, S. McCouch, Evolutionary history of GS3, a gene conferring grain length in rice, Genetics 182 (4) (2009) 1323–1334.
[58] J.H. Lee, M. Muhsin, G.A. Atienza, D.Y. Kwak, S.M. Kim, T.B. De Leon, E. R. Angeles, E. Coloquio, H. Kondoh, K. Satoh, R.C. Cabunagan, Single nucleotide polymorphisms in a gene for translation initiation factor (eIF4G) of rice (*Oryza sativa*) associated with resistance to Rice tungro spherical virus, Mole. Plant-Microbe Inter. 23 (1) (2010) 29–38.
[59] Z.K. Li, A. Sanchez, E. Angeles, S. Singh, J. Domingo, N. Huang, G.S. Khush, Are the dominant and recessive plant disease resistance genes similar? A case study of rice R genes and *Xanthomonas oryzae* pv. oryzae races, Genetics 159 (2) (2001) 757–765.
[60] A.S. Iyer, S.R. McCouch, The rice bacterial blight resistance gene xa5 encodes a novel form of disease resistance, Mole. Plant-Microbe Inter. 17 (12) (2004) 1348–1354.
[61] C.J. Dilla-Ermita, E. Tandayu, V.M. Juanillas, J. Detras, D.N. Lozada, M. S. Dwiyanti, C.V. Cruz, E.G.N. Mbanjo, E. Ardales, M.G. Diaz, M. Mendioro, Genome-wide association analysis tracks bacterial leaf blight resistance loci in rice diverse germplasm, Rice 10 (1) (2017) 1–17.
[62] Z. Chu, M. Yuan, J. Yao, X. Ge, B. Yuan, C. Xu, X. Li, B. Fu, Z. Li, J.L. Bennetzen, Q. Zhang, Promoter mutations of an essential gene for pollen development result in disease resistance in rice, Genes Dev. 20 (10) (2006) 1250–1255.
[63] C. Wang, X. Zhang, Y. Fan, Y. Gao, Q. Zhu, C. Zheng, T. Qin, Y. Li, J. Che, M. Zhang, B. Yang, Y. Liu, K. Zhao, Xa23 is an executor R protein and confers broad-spectrum disease resistance in rice, Mol. Plant 8 (2015) 290–302.
[64] H. Peng, Z. Chen, Z. Fang, J. Zhou, Z. Xia, L. Gao, L. Chen, L. Li, T. Li, W. Zhai, W. Zhang, Rice Xa21 primed genes and pathways that are critical for combating bacterial blight infection, Sci. Rep. 5 (1) (2015) 1–12.
[65] P. Römer, S. Recht, T. Lahaye, A single plant resistance gene promoter engineered to recognize multiple TAL effectors from disparate pathogens, Proc. Natl. Acad. Sci. 106 (48) (2009) 20526–20531.
[66] E.M. Septiningsih, A.M. Pamplona, D.L. Sanchez, C.N. Neeraja, G.V. Vergara, S. Heuer, A.M. Ismail, D.J. Mackill, Development of submergence-tolerant rice cultivars: the Sub1 locus and beyond, Ann. Bot. 103 (2) (2009) 151–160.
[67] J. Pariasca-Tanaka, M. Lorieux, C. He, S. McCouch, M.J. Thomson, M. Wissuwa, Development of a SNP genotyping panel for detecting polymorphisms in *Oryza glaberrima/O. sativa* interspecific crosses, Euphytica 201 (1) (2015) 67–78.
[68] A. Rasheed, W. Wen, F. Gao, S. Zhai, H. Jin, J. Liu, Q. Guo, Y. Zhang, S. Dreisigacker, X. Xia, Z. He, Development and validation of KASP assays for genes underpinning key economic traits in bread wheat, Theor. App. Genet. 129 (10) (2016) 1843–1860.
[69] H.E. Kauffman, An improved technique for evaluating resistance of rice varieties to *Xanthomonas oryzae*, Plant Dis. Rep. 57 (1973) 537–541.
[70] N.N. Narayanan, N. Baisakh, C.M. Vera Cruz, S.S. Gnanamanickam, K.B. Datta, S. K. Datta, Molecular breeding for the development of blast and bacterial blight resistance in rice cv. IR50, Crop Sci. 42 (6) (2002) 2072–2079.
[71] E.A. Heinrichs, Genetic evaluation for insect resistance in rice, Int. Rice Res. Inst. (1985) 22–29.
[72] S.A. Angaji, E.M. Septiningsih, D.J. Mackill, A.M. Ismail, QTLs associated with tolerance of flooding during germination in rice (*Oryza sativa* L.), Euphytica 172 (2) (2010) 159–168.
[73] M.G. Murray, W.F. Thompson, Rapid isolation of high molecular weight DNA, Nucleic Acids Res. 8 (1980) 4221–4235.
[74] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, The sequence alignment/map format and SAMtools, Bioinformatics 25 (2009) 2078–2079.
[75] R.H. Ramirez-Gonzalez, C. Uauy, M. Caccamo, PolyMarker: a fast polyploid primer design pipeline, Bioinformatics 31 (12) (2015) 2038–2039.
[76] I. Letunic, P. Bork, Interactive tree of life (iTOL) v4: recent updates and new developments, Nucleic Acids Res. 47 (W1) (2019) W256–W259.
[77] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2015.

[78] J.D. Platten, J.N. Cobb, R.E. Zantua, Criteria for evaluating molecular markers: comprehensive quality metrics to improve marker-assisted selection, PLoS One 14 (1) (2019), e0210529.
[79] H.P. Piepho, Methods for comparing the yield stability of cropping systems - a review, J. Agron. Crop Sci. 180 (1998) 193–213.
[80] H.P. Piepho, Stability analysis using the SAS system, Agron. J. 9 (1999) 154–160.
[81] A. Raman, J.K. Ladha, V. Kumar, S. Sharma, H.P. Piepho, Stability analysis of farmer participatory trials for conservation agriculture using mixed models, Field Crop Res. 121 (2011) 450–459.