Current Plant Biology 34 (2023) 100285

Contents lists available at ScienceDirect

Current Plant Biology

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

Development and validation of a genome-wide InDel marker set discriminating the alleles between the BB-genome *Oryza* species and rice (*O. sativa*)☆

Katrina B. Malabanan-Bauan [a,b], Sherry Lou Hechanova [a], Eok-Keun Ahn [c], Charng-Pei Li [d], Il-Ryong Choi [a,c], Jose E. Hernandez [b], Kshirod K. Jena [e], Sung-Ryul Kim [a,*]

[a] *Rice Breeding Innovation Platform, International Rice Research Institute (IRRI), Philippines*
[b] *Institute of Crop Science (ICropS), University of the Philippines Los Baños (UPLB), Philippines*
[c] *National Institute of Crop Science, Rural Development Administration (RDA), Republic of Korea*
[d] *Taiwan Agricultural Research Institute (TARI), Council of Agriculture, Taiwan*
[e] *School of Biotechnology, KIIT Deemed University, Bhubaneswar 751024, Odisha, India*

ARTICLE INFO

Keywords:
Oryza genus
Wild rice species
BB genome
InDel marker
Oryza punctata
Oryza minuta

ABSTRACT

Wild relatives of cultivated rice (*Oryza sativa*) belonging to the genus *Oryza* are regarded as 'genetic reservoirs' for rice improvement. Their genes are transferable to cultivated rice through crossing and they have been contributing significantly to rice variety improvement by introgression of their diverse valuable traits, especially biotic stress resistance. Moreover, active use of wild rice resources might be one of the ideal solutions to seek novel or superior alleles/genes to cope with climate change and stable high-yield rice production. DNA markers are essential tools for genetic analysis and breeding. However, to date, there are no suitable DNA marker sets for BB-genome species (*O. punctata*). In this study, we developed a genome genome-wide InDel marker set evenly distributed (~2 Mb intervals) across the 12 chromosomes for BB-genome species. The markers were validated by PCR-agarose gel analysis with BB-genome containing species: four accessions each of *O. punctata* (BB) and *O. minuta* (BBCC). Out of the 191 InDel markers designed, 184 (96.3%) and 138 (72.2%) were able to differentiate the alleles between *O. sativa* and *O. punctata* and between *O. sativa* and *O. minuta* respectively. The same marker sets were also tested in other genome types species (CC, CCDD, EE, FF, GG, HHJJ, and KKLL) and one accession of *Leersia perrieri* (a sister genus to *Oryza*). The number of polymorphic markers *(O. sativa* vs other genome types) was drastically reduced in other genome types. In contrast, the number of markers showing no PCR amplification increased, especially in FF, GG, HHJJ, and KKLL species. This suggests that the development of genome type-specific marker sets would be more efficient rather than testing random InDel markers. The newly developed InDel markers maybe be useful for the identification of valuable genetic factors from the BB or BBCC-genome species and also for transferring the identified genes/QTLs into elite variety backgrounds.

1. Introduction

Rice is one of the most economically important crops in the world and serves as the staple food for around four billion people [1]. Over the years, many breeding programs have produced improved varieties with better yield potential. However, stable high-yield production of rice has been affected by diverse biotic/abiotic factors and climate changes [2]. Cultivated rice has limited variability for resistance to different stresses and other agronomic traits such as heading date, biomass, grain quality and nutrition due to limited genetic variations. In contrast, wild rice species in the genus *Oryza* have adapted to harsh environmental conditions without any protection for few thousand to 15 million years and thus they might have genetically developed resistance or tolerance to many biotic/abiotic stresses as well as valuable agronomic traits. For examples, many biotic stress-resistant genes including *Xa21* (*O. longistaminata*), *Xa23* (*O. rufipogon*), and *Xa27* (*O. minuta*) for bacterial blight resistance, *Pi9* (*O. minuta*), *Pi40* (*O. australiensis*), and *Pi54* (*O. rhizomatis*) for blast resistance, and *BPH14* (*O. officinalis*) and *BPH18*

☆ This article is part of a special issue entitled "VSI:crop improvement".

* Corresponding author.
E-mail address: s.r.kim@irri.org (S.-R. Kim).

https://doi.org/10.1016/j.cpb.2023.100285
Received 20 March 2023; Received in revised form 3 May 2023; Accepted 3 May 2023
Available online 4 May 2023

(*O. australiensis*) for brown planthopper resistance were identified and used in many rice breeding programs [3,4]. In addition, lack of some traits in rice germplasm such as early morning flowering from *O. officinalis* for heat stress escape during pollination/fertilization [5], long-exserted stigma from an out-crossing preferred species (*O. longistaminata*) for high-outcrossing rate in hybrid rice production [6], and rhizome traits from *O. longistaminata* for perennial rice variety development [7], were identified and successfully introgressed into the rice. In order to broaden the genetic variability of cultivated rice, valuable genes from the wild rice species need to be screened, identified, and transferred into modern rice varieties [8].

Genetic diversity is fundamental in crop breeding for variety improvement. Wild relatives of cultivated rice (*Oryza sativa*) belonging to the genus *Oryza* are regarded as 'genetic reservoirs' for rice improvement. The genus *Oryza* consists of 27 species (diploid and allotetraploid species) including two cultivated species (*O. sativa* – Asian rice and *O. glaberrima* – African rice) with 11 different genome types (AA, BB, CC, BBCC, CCDD, EE, FF, GG, HHJJ, HHKK and KKLL) [9]. All *Oryza* species can be hybridized with *O. sativa* through interspecific cross, and even *Leersia perrieri,* a sister genus of *Oryza,* was successfully crossed with a cultivar and produced true F_1 [10]. The BB genome is present in the diploid BB-genome species, *O. punctata* that originated from Africa and three allotetraploid BBCC-genome species: *O. minuta* originated from the Philippines and Papua New Guinea, *O. malampuzhaensis* originated from India, and the allotetraploid form of *O. punctata* which some classify as *O. schweinfurthiana* [11]. Several valuable traits such as insect resistance (brown planthopper and green leaf-hopper), disease resistance (bacterial blight and blast), and abiotic stress resistance (heat and drought) were reported in *O. punctata* and *O. minuta* [3,4,8,12]. Currently, 71 accessions of *O. punctata* and 63 of *O. minuta* are available from the International Rice Genebank [4]. Harnessing of the BB and BBCC genome germplasm will contribute to isolation of valuable genetic factors for rice varietal improvement.

Molecular markers are the essential tools for an efficient use of wild germplasm for rice improvement in the genetics aspect including genetic study, gene mapping, and marker-assisted breeding using wild germplasm. To date SSR (simple sequence repeat) and STS (sequence-tagged site) markers which were designed based on the *O. sativa* sequences are commonly used for wild rice species. However, these SSR and STS markers have shown less polymorphism or no amplification when applied to other rice genomes. Thus, there are limitations in obtaining enough polymorphic markers as well as even marker distribution after the screening with those markers in wild rice species-derived materials. Due to this reason, we recently developed and published a genome-wide InDel marker set (338–416 polymorphic markers) which can discriminate the alleles between *O. sativa* and the other AA-genome species through multiple sequence alignments [13]. In order to develop the BB-genome introgression lines, Jena et al. [12] screened 700 SSR and STS markers and obtained only 60 polymorphic markers (8.57%) between rice and *O. punctata*. Thus, ~100 BB-genome polymorphic markers were developed and used in their study. However, this marker set was tested in only one *indica* rice (IR31917–45–3–2) and one *O. punctata* accession (IRGC105690). Development of additional polymorphic markers for the BB-genome and validation of the markers with several accessions of the BB-genome containing species, *O. punctata* (BB) and *O. minuta* (BBCC) will be helpful to harness useful traits from BB and BBCC genome wild species.

Here we developed a genome-wide InDel marker set consisting of 191 markers across 12 rice chromosomes by multiple sequence alignment among five rice cultivars and *O. punctata*. Furthermore, we validated the markers with four accessions each of *O. punctata* and *O. minuta* and also with seven other genome types of wild species.

2. Materials and methods

2.1. Sequence preparation and comparison for DNA marker development

Chromosome level sequence assembly of the BB-genome species (*O. punctata*, IRGC:105690) was performed by The *Oryza* Map Alignment Project (OMAP, http://www.omap.org/) and was deposited to the National Center for Biotechnology Information (NCBI, https://www.ncbi.nlm.nih.gov/) database with the BioProject number (PRJNA13770). Whole genome sequences of both BB-genome *O. punctata* and five AA-genome Asian cultivars (*O. sativa*) including IR8, IR64, Minghui 63, Zhenshan 97, and Nipponbare were obtained from the NCBI nucleotide database.

Around 20-kb bait sequences were prepared using the rice reference genome sequence, *O. sativa* subsp. *japonica* cv. Nipponbare from RAP-DB (https://rapdb.dna.affrc.go.jp/). Bait sequences were manually selected from top to bottom of each chromosome with 1.5–2.5 Mb intervals while avoiding long repeat sequences by using 'Repeat regions' marking option in RAP-DB. Chromosome sequences of the five varieties and *O. punctata* were loaded into the NCBI Genome Workbench software to isolate the orthologous region to the bait sequence by using BLAST tool. After extraction of the orthologous sequences, multiple sequence alignment was conducted using the web-based tool mVISTA (http://genome.lbl.gov/vista/). BioEdit software [14] was used to highlight polymorphic regions among the aligned sequences and for export of the aligned sequence into text file format.

2.2. Primer design

To design primers, multiple sequence alignments processed by mVISTA and BioEdit were opened using Microsoft Word. The InDels with > 20-bp gap between the cultivars and *O. punctata* were selected for marker development. Forward and reverse primers were manually designed on the conserved regions among the sequences flanking the polymorphic InDel sequences to ensure proper annealing of primers in all species used. GC content of about 50%, annealing temperature of 55 °C, and 100–500 bp PCR product sizes were maintained for all primer pairs so that all the markers can be PCRed at the same PCR conditions. Primer sequences were checked using BLAST tool in the RAP-DB for redundancy, and the unique hits were selected for primer sequences.

2.3. Plant materials

For validation of the newly developed marker sets, PCR-agarose gel analyses were conducted with four accessions of BB-genome species (*O. punctata*) and four accessions of BBCC-genome species (*O. minuta*) together with two Asian cultivars (*O. sativa* subsp. *indica* cv. IR24 and subsp. *japonica* cv. Nipponbare) and one African rice (*O. glaberrima*). To check the functionality of the BB-genome markers in the other genome types, wild rice accessions from the seven different genome types including CC, CCDD, EE, FF, GG, HHJJ, and KKLL for the genus *Oryza* and one accession of *Leersia perrieri* were tested together. All plant materials are listed in Table 1. Seeds of the wild species were obtained from the IRRI Genebank (https://gringlobal.irri.org/gringlobal/search) and all the plant materials were grown in the glasshouse at IRRI.

2.4. DNA preparations

Leaf samples were collected from each accession at the seedling stage. DNA was prepared using a modified simple DNA preparation method [15]. Briefly, 2–4 cm length of leaf samples were placed in 2 mL tubes with two steel balls. In each tube, 400 μL of TP (100 mM tris-HCl pH9.5, 1 M KCl) buffer was added and the leaf samples were ground by using a 2010 Geno Grinder (SPEX SamplePrep, http://www.spexsampleprep.com) for 7 m at 1000RPM. Then, the tubes were incubated in a 65 °C water bath for 30–60 m. Tubes were centrifuged for 15 m at 13,

Table 1
Plant materials used in this study.

Code	Species	Accession no.	Genome	Origin
NB	*O. sativa* subsp. *japonica*	cv. Nipponbare	AA	Japan
IR24	*O. sativa* subsp. *indica*	cv. IR24	AA	Philippines
Pun_B01	*O. punctata*	IRGC 105690	BB	Kenya
Pun_B02	*O. punctata*	IRGC 101417	BB	Kenya
Pun_B03	*O. punctata*	IRGC 103896	BB	Tanzania
Pun_B04	*O. punctata*	IRGC 104067	BB	Chad
Min_B05	*O. minuta*	IRGC 101141	BBCC	Philippines
Min_B06	*O. minuta*	IRGC 93257	BBCC	Philippines
Min_B07	*O. minuta*	IRGC 101080	BBCC	Philippines
Min_B08	*O. minuta*	IRGC 103877	BBCC	Philippines
Eich_B09	*O. eichingeri*	IRGC 101424	CC	Uganda
Eich_B10	*O. eichingeri*	IRGC 99567	CC	Tanzania
Offi_B11	*O. officinalis*	IRGC 100896	CC	Thailand
Offi_B12	*O. officinalis*	IRGC 80777	CC	Philippines
Rhiz_B13	*O. rhizomatis*	IRGC 105432	CC	Sri Lanka
Rhiz_B14	*O. rhizomatis*	IRGC 103421	CC	Sri Lanka
Alta_B15	*O. alta*	IRGC 105143	CCDD	Guyana
Grand_B16	*O. grandiglumis*	IRGC 101405	CCDD	Brazil
Lati_B17	*O. latifolia*	IRGC 100914	CCDD	Mexico
Austral_B18	*O. australiensis*	IRGC 100882	EE	Australia
Austral_B19	*O. australiensis*	IRGC 86536	EE	Australia
Brachy_B20	*O. brachyantha*	IRGC 101232	FF	Sierra Leone
Brachy_B21	*O. brachyantha*	IRGC 81950	FF	Zambia
Granu_B22	*O. granulata*	IRGC 102118	GG	Thailand
Meyeri_B23	*O. meyeriana*	IRGC 89244	GG	Philippines
Longiglu_B24	*O. longiglumis*	IRGC 105148	HHJJ	Indonesia
Ridle_B25	*O. ridleyi*	IRGC 100821	HHJJ	Thailand
Coarc_B26	*O. coarctata*	IRGC 104502	KKLL	Bangladesh
L. perrieri_B27	*L. perrieri*	IRGC 105164	-	Madagascar
Glabe_B28	*O. glaberrima*	IRGC 102486	AA	Liberia

000 x *g* and supernatant was transferred into 1.5 mL tubes and stored at 4 °C. Each sample was diluted by placing 30 μL of DNA extracts into a 96-well plate containing 120 μL of double-distilled water and used as template DNA for PCRs.

2.5. PCR and agarose gel analysis

For PCR analysis, 1.5 μL of 10x PCR buffer, 1.5 μL dNTPs (2 mM of each dNTP), 0.15 μL each of forward and reverse primer (5 μM each), 0.3 μL *Taq* polymerase (1 unit of *Taq* pol) (BioFact Co., https://www.bio-ft.com/eng/index.html), and 9.4 μL of double distilled water were mixed with 2 μL diluted-crude DNA extract. Thermal cycle program used was as follows: 94 °C for 2.5 m, 35 cycles of 95 °C for 25 s, 55 °C for 25 s, 72 °C for 50 s, and a final extension at 72 °C for 5 m. PCR products were resolved in 1.5–3% agarose gel based on the PCR product sizes. Gel images were taken and presented in the manuscript. For the markers with poor results (weak or no amplifications, unexpected sizes of bands, and multiple bands), PCRs were repeated. Only clearly amplified bands with the expected band sizes in *O. sativa* and BB-genome species were regarded as true PCR bands. Faint or poorly amplified bands were ignored in all the samples.

2.6. Graphical mapping of the markers in the rice genome

The web-based tool, PhenoGram (http://visualization.ritchielab.psu.edu/), was used for mapping the physical location of each marker on the chromosomes.

2.7. Development of interspecific hybrids through embryo rescue

IR24 variety (female donor) was crossed with *O. punctata* (IRGC 105609, male donor). After 7–10 days from pollination, immature embryos were collected in 50-mL conical tube, sterilized using 50% Clorox with gentle shaking for 40 min, washed five times with autoclaved water, and placed on a half-strength of Murashige & Skoog (MS) media.

3. Results

3.1. Development of a genome-wide InDel marker set for BB genome Oryza species

To develop evenly distributed genome-wide InDel markers discriminating the alleles between *O. sativa* and BB-genome species (*O. punctata*), we used a multiple sequence alignment in every ~2 Mb intervals along the chromosome and primer designing strategy which was used for the development of a genome-wide InDel marker set distinguishing the alleles between rice cultivars and AA-genome wild species [13]. Briefly, the bait sequences were obtained from the rice reference genomes at RAP-DB and used to extract orthologous regions from the *O. punctata* (BB-genome) and cultivars (AA-genome). In some cases, orthologous regions were not detected, suggesting that the selected region might not be present or highly variable in some cultivars or in wild species. Orthologous sequences found in wild species and at least three cultivars were used for multiple sequence alignment. The mVISTA webtool was used to generate multiple sequence alignments while BioEdit highlighted the SNPs and InDels among the aligned sequences. The InDels between the cultivars and *O. punctata* with at least 20-bp gap were selected for marker development. A total of 191 InDel markers were designed across the 12 chromosomes. Based on the multiple sequence alignment, the expected PCR band sizes for five varieties and *O. punctata* were calculated for each marker (Supplemental Table 1). Intervals of 1.5–2.5 Mb interval were intentionally maintained between neighboring markers to ensure even marker distribution in the rice genome. As expected, 191 markers were evenly distributed across the 12 chromosomes (Fig. 1A). Markers were named following the format BxxPxxxxx where Bxx stands for the genome type (B) and chromosome number while Pxxxxx indicates the physical location (kb) of the marker along the chromosome. This nomenclature will help the users to easily recognize the chromosome number and physical location of the markers.

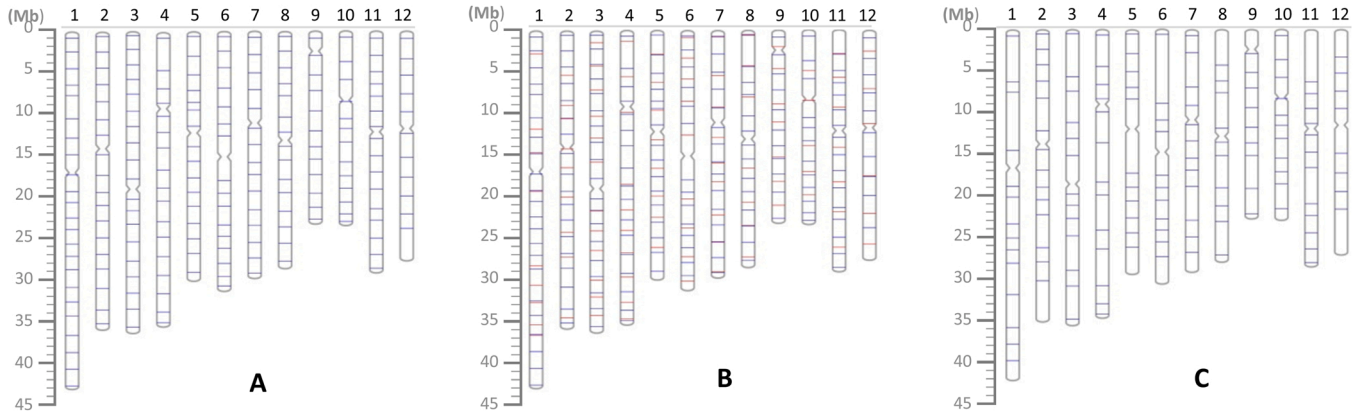

Fig. 1. Graphical mapping of physical location of the BB-genome InDel markers on 12 rice chromosomes. (A) 191 InDel markers developed in this study. (B) 184 polymorphic markers between *O. sativa* (AA genome) and *O. punctata* (BB genome). The previously reported 99 BB-genome markers by Jena et al. [12] are also mapped together (red bar) and their marker information are available at the Supplemental Table 2. (C) 138 polymorphic markers between *O. sativa* and *O. minuta* (BBCC genome).

3.2. Marker validation with BB and BBCC genome Oryza species

For experimental validation, all 191 markers were tested by PCR-agarose gel analysis with an *indica* cultivar (IR24), *japonica* cultivar (Nipponbare), four accessions of *O. punctata* (BB genome), and four accessions of *O. minuta* (BBCC genome). Of the 191 markers, 186 markers (97.3%) were successfully amplified while five markers showed no/weak amplification or unexpected multiple bands (Supplemental Fig. 1 and Supplemental Table 1). Among the four *O. punctata* accessions used (Pun_B01 to Pun_B04), Pun_B02 and Pun_B03 samples had the highest number of polymorphic markers at 181 against IR24 and Nipponbare while only 136 and 139 markers exhibited polymorphism in Pun_B01 versus IR24 and Nipponbare, respectively (Table 2). In total, 184 markers (96.3%) showed polymorphism at least in one *O. punctata* accession and these markers can be used in the genetic study of other BB genome accessions. The physical locations of 184 potential BB-genome polymorphic markers together with the previously reported BB-genome markers [12] were mapped on to the 12 chromosomes (Fig. 1B). Moreover, some of the markers were also tested to confirm true hybridity in F_1s between cultivar and *O. punctata*. Nine potential interspecific F_1 hybrids (IR24 x *O. punctata* _IRGC 105690) obtained by embryo rescue were tested by four markers (B01P39081, B02P02074, B04P20407, and B06P00234). The true F_1 plants (#6, #7, #8, and #9) clearly showed both expected parental band sizes (Fig. 2).

In another BB-genome containing species *O. minuta* (BBCC), a higher number of markers (135–138 markers) still showed polymorphism between cultivated rice (IR24 and Nipponbare) and *O. minuta* except for the Min_B08 accession (Table 2, Supplemental Fig. 1). The PCR band patterns of the Min_B08 were highly similar to those of cultivars (IR24 and Nipponbare) in most of the markers (> 160), suggesting that the Min_B08 accession might be *O. sativa* and probably something went wrong during seed amplification or conservation processes. Physical

Table 2
Summary of the BB-genome marker validation results.

Code	Species	Genome	No. of polymorphic marker (vs IR24)	No. of polymorphic marker (vs NB)	No. of no amplification markers
Pun_B01	*O. punctata*	BB	136	139	7
Pun_B02	*O. punctata*	BB	181	181	8
Pun_B03	*O. punctata*	BB	181	181	6
Pun_B04	*O. punctata*	BB	179	179	8
Min_B05	*O. minuta*	BBCC	135	138	6
Min_B06	*O. minuta*	BBCC	135	136	8
Min_B07	*O. minuta*	BBCC	136	139	7
Min_B08	*O. minuta*	BBCC	6	11	6
Eich_B09	*O. eichingeri*	CC	137	138	9
Eich_B10	*O. eichingeri*	CC	118	121	19
Offi_B11	*O. officinalis*	CC	123	124	21
Offi_B12	*O. officinalis*	CC	119	126	22
Rhiz_B13	*O. rhizomatis*	CC	117	118	22
Rhiz_B14	*O. rhizomatis*	CC	112	115	29
Alta_B15	*O. alta*	CCDD	116	120	13
Grand_B16	*O. grandiglumis*	CCDD	113	113	21
Lati_B17	*O. latifolia*	CCDD	121	121	13
Austral_B18	*O. australiensis*	EE	95	96	51
Austral_B19	*O. australiensis*	EE	95	97	52
Brachy_B20	*O. brachyantha*	FF	59	59	108
Brachy_B21	*O. brachyantha*	FF	49	49	123
Granu_B22	*O. granulata*	GG	71	72	98
Meyeri_B23	*O. meyeriana*	GG	63	63	105
Longiglu_B24	*O. longiglumis*	HHJJ	19	19	163
Ridle_B25	*O. ridleyi*	HHJJ	20	2	164
Coarc_B26	*O. coarctata*	KKLL	25	26	157
L. perrieri_B27	*L. perrieri*	-	48	48	119
Glabe_B28	*O. glaberrima*	AA	24	21	18

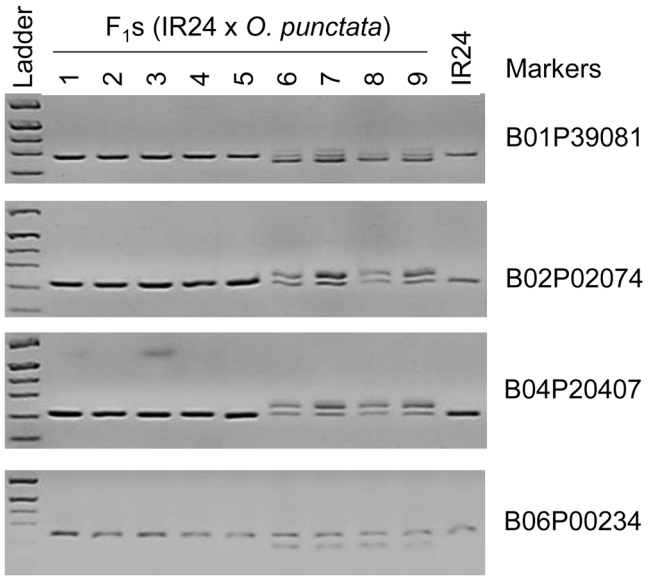

Fig. 2. Selection of true F_1s (IR24 x *O. punctata*) by using the new BB-genome markers. Nine potential interspecific F_1 hybrids derived by the cross between IR24 and *O. punctata* (IRGC 105690) were tested by four markers (B01P39081, B02P02074, B04P20407, and B06P00234). The results showed that four plants (#6, #7, #8, and #9) were true hybrids.

locations of the 138 polymorphic markers to the BBCC genome were mapped on to the 12 chromosomes (Fig. 1C).

3.3. Application of the markers on other Oryza genome types and L. perrieri

In addition to BB-genome containing species, these markers were also tested in other seven genome types of the genus *Oryza* (CC, CCDD, EE, FF, GG, HHJJ, and KKLL) and one accession of *L. perrieri* (a sister genus to *Oryza*) (Table 1) to determine their applicability to other *Oryza* genomes. Two accessions from each CC genome species (*O. eichingeri*, *O. officinalis*, and *O. rhizomatis*) were tested with the 191 markers. Although the number of polymorphic markers between *O. sativa* and the CC-genome species were reduced compared to that of the BB genome species, 112–138 markers (58.6–72.2%) exhibited polymorphism. In comparison with the CCDD-genome germplasm (one accession each from *O. alta*, *O. grandiglumis*, and *O. latifolia*), the number of polymorphic markers ranged from 113 to 121 (59.1–63.3%). For the EE (*O. australiensis*), FF (*O. brachyantha*), GG (*O. granulata*) genome-types accessions, the number of polymorphic markers ranged from 49 to 97 (25.6–50.7%). The genomes of HHJJ (*O. longiglumis* and *O. ridleyi*) and KKLL (*O. coarctata*) showed only 2–26 polymorphic markers (1.0–13.6%) against the cultivars. Between rice cultivars and *L. perrieri*, 48 polymorphic markers (25.1%) were found. Interestingly, the number of markers without amplification increased in the other genome types and more than half of the markers tested (98–164 markers) showed no amplification in the FF, GG, HHJJ, KKLL, and *L. perrieri* materials (Table 2, Supplemental Fig. 1). This phenomenon might be caused by sequence variation in the primer annealing sites or complete deletion in some other genome-species.

4. Discussion

Wild relatives of rice exhibit wide genetic diversity as they have been exposed and adapted to various adverse environments. These wild species have been known as reservoir of useful genes conferring resistance to multiple biotic and abiotic stresses as well as improved yield [2,9]. For genetic analysis, gene mapping, gene identification and transferring the identified genetic factors (QTLs/genes) from wild relatives of rice, molecular markers are the essential tools. One of the most common markers used in rice is the STS markers detecting InDels and the SSRs amplifying potential polymorphism based on the number of simple repeat sequences (mono, di, tri, tetra nucleotides) among rice germplasm. While they may be powerful tools, they were designed for the cultivated species and need to be screened to identify polymorphism for the specific genotypes/materials which may be more time-consuming and costly. Generally, the number of polymorphic markers and marker distributions on the genome are poor in wild rice species, and many markers tend to show no PCR bands, especially in more distant species. Thus, the need for markers designed for other genomes to allow genetic analysis and identify valuable genes. In this study, we developed a genome-wide InDel marker set for the BB genome species. In total, 191 InDel markers were designed by using multiple sequence alignment among the sequences of five rice cultivars and one *O. punctata* with an intended ~2.0 Mb marker intervals. Of the 191 InDel markers, 186 successfully amplified with the majority (184 markers) showing polymorphism between *O. sativa* and *O. punctata* accessions. These results support that the InDel markers developed in this study are highly effective in distinguishing the alleles with even intervals between cultivated and BB-genome wild rice species compared to the SSR and STS markers. Moreover, we designed the markers showing the same/similar band sizes among the five cultivars including *japonica* and *indica* (Supplemental Table 1) so that the markers can be applied to many diverse rice backgrounds. Thus, these markers have strong potentials to show polymorphism between diverse rice cultivars and other accessions of *O. punctata* collections.

For BBCC genome species, Wang et al. [16] developed a set of 495 SSR markers by using the next-generation sequencing of BBCC species *O. minuta* (acc. W303) and following isolation of the SSR sequences. These markers will be useful to study BBCC genome species but the polymorphism need to be resolved by a DNA fragments analyzer or polyacrylamide gel electrophoresis (PAGE) which are costly and cumbersome methods. The markers we developed can be clearly separated in agarose gel which is a simple, inexpensive, and user-friendly method compared to the methods above. Marker application in three accessions (Min_B05, B06, and B07) of *O. minuta* (BBCC) showed the high numbers of polymorphic markers (135–138 markers: 70.7–72.2%) between rice cultivars (IR24 and Nipponbare) and *O. minuta*. This might be caused by mainly possessing the BB genome but some cases caused by the CC genomes. Around 48–55 markers showed two or more PCR bands in the three accessions of *O. minuta* (Supplemental Fig. 1), suggesting that B and C genomes possess different sequences at the marker locus and some potential heterozygosity in B or C genomes. Although the markers are not able to discriminate the B or C genomes, the markers showing polymorphism with *O. sativa* will be useful to harness *O. minuta* germplasms as well.

When the markers were applied to other genome types, the number of polymorphic markers was quite reduced especially in the germplasm of EE, FF, GG, HHJJ, and KKLL-genome species. In contrast, the number of markers without amplification increased. Similar results which showed high tendency of no PCR amplification in the distant species (FF, GG, HHJJ, and KKLL) with subspecies-specific InDel markers were observed by Chin et al. [17]. These results suggested that the sequences at the marker loci are much different among the genome types. The *Oryza* phylogenetic tree showing that the AA, BB, and CC genomes are relatively close to each other with other genomes present in different clusters [4] supports our results. In our study, one accession of the African cultivated species *O. glaberrima* (Glabe_B28) was included but only 21–24 markers showed polymorphism between *O. sativa* and *O. glaberrima* (Table 2). All these results suggested that the development of genome type-specific markers would be efficient to utilize diverse genome types of wild species in terms of saving time, effort, and cost.

5. Conclusions

In this study, a genome-wide InDel marker set consisting of 186 potential polymorphic markers between *O. sativa* and the BB-genome-containing species was successfully developed and validated. Seeds of more than 100 accessions of BB- or BBCC-genome species are available from the International Rice Genebank. The newly developed markers together with the previously reported BB-genome markers by Jena et al. [12] will be useful to identify valuable genetic factors from the BB-genome species for rice improvement. In addition to marker information, all the gel images of the 186 markers are available in the Supplemental Fig. 1. Thus, any users can select potential polymorphic markers based on the gel images for their own wild rice materials.

Funding

This research was funded by the Temperate Rice Research Consortium (TRRC) project and partially by the cooperative research program between IRRI and the Council of Agriculture (COA), Taiwan (Grant No. DRPC_A-2020-22).

CRediT authorship contribution statement

Katrina B. Malabanan-Bauan: experiment, data collection and analysis, and writing draft. **Sherry Lou Hechanova:** plant material preparation and experiments. **Charng-Pei Li:** testing markers, investment, and editing. **Eok-Keun Ahn:** conceiving idea, curation, and editing. **Il-Ryong Choi:** conceiving idea, editing and investment. **Jose E. Hernandez:** conceiving idea and editing. **Kshirod K. Jena:** conceiving idea and editing. **Sung-Ryul Kim:** conceiving idea, set up the marker designing strategy, writing and editing. All authors read and approved the manuscript.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All the data are presented in the main manuscript, Figures, and Supplementary data.

Acknowledgments

We thank Ms. Marianne Ilagan for embryo rescue and Mr. Allan Trinidad and Mr. Aldrich Llanera for growing and maintaining plant materials in glasshouses.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.cpb.2023.100285.

References

[1] J. Kim, X. An, K. Yang, S. Miao, Y. Qin, Y. Hu, B. Du, L. Zhu, G. He, R. Chen, Molecular mapping of a new brown planthopper resistance gene Bph43 in rice (*Oryza sativa L.*), Agronomy 12 (2022) 808, https://doi.org/10.3390/agronomy12040808.

[2] R.A. Wing, M.D. Purugganan, Q. Zhang, The rice genome revolution: from an ancient grain to Green Super Rice, Nat. Rev. Genet. 19 (2018) 505–517, https://doi.org/10.1038/s41576-018-0024-z.

[3] Z.K. Li, T.Q. Zheng, Utilization of exotic germplasm, in: Q.F. Zhang, R.A. Wing (Eds.), Genetics and Genomics of Rice: Crops and Models, Springer, 2013, pp. 349–361.

[4] P.L. Sanchez, R.A. Wing, D.S. Brar, The wild relative of rice: genomes and genomics, in: Q. Zhang, R. Wing (Eds.), Genetics and Genomics of Rice, Springer, 2013, pp. 9–25.

[5] H. Hirabayashi, K. Sasaki, T. Kambe, R.B. Gannaban, M.A. Miras, M.S. Mendioro, E. V. Simon, P.D. Lumanglas, D. Fujita, Y. Takemoto-Kuno, Y. Takeuchi, R. Kaji, M. Kondo, N. Kobayashi, T. Ogawa, I. Ando, K.S. Jagadish, T. Ishimaru, qEMF3, a novel QTL for the early-morning flowering trait from wild rice, *Oryza officinalis*, to mitigate heat stress damage at flowering in rice, *O. sativa*, J. Exp. Bot. 66 (5) (2015) 1227–1236, https://doi.org/10.1093/jxb/eru474.

[6] G.D. Prahalada, B. Marathi, R. Vinarao, S.R. Kim, R. Diocton IV, J. Ramos, K. K. Jena, Q.T.L. Mapping, of a novel genomic region associated with high outcrossing rate derived from *Oryza longistaminata* and development of new CMS lines in rice O. sativa L. Rice 14 (2021) 80, https://doi.org/10.1186/s12284-021-00521-9.

[7] S. Zhang, G. Huang, Y. Zhang, X. Lv, K. Wan, J. Liang, Y. Feng, J. Dao, S. Wu, L. Zhang, X. Yang, X. Lian, L. Huang, L. Shao, J. Zhang, S. Qin, D. Tao, T.E. Crews, E.J. Sacks, J. Lyu, L.J. Wade, F. Hu, Sustained productivity and agronomic potential of perennial rice, Nat. Sustain. 6 (2023) 28–38, https://doi.org/10.1038/s41893-022-00997-3.

[8] K.K. Jena, The species of the genus *Oryza* and transfer of useful genes from wild species into cultivated rice O. sativa, Breed. Sci. 60 (2010) 518–523.

[9] J.C. Stein, Y. Yu, D. Copetti, D.J. Zwickl, L. Zhang, C. Zhang, K. Chougule, D. Gao, A. Iwata, J.L. Goiocoechea, S. Wei, J. Wang, Y. Liao, M. Wang, J. Jacquemin, C. Becker, D. Kudrna, J. Zhang, C.E.M. Londono, X. Song, S. Lee, P. Sanchez, A. Zuccolo, J.S.S. Ammiraju, J. Talag, A. Danowitz, L.F. Rivera, A.R. Gschwend, C. Noutsos, C. Wu, S. Kao, J. Zeng, F. Wei, Q. Zhao, Q. Feng, M. El Baidouri, M. Carpentier, E. Lasserre, R. Corke, L.C. d. da Rosa Farias, R.S. da Maia, K.G. dos Santos, K.L. Nyberg, R. McNally, N. Mauleon, J. Alexandrov, D. Schmutz, C. Flowers, D. Fan, K.K. Weigel, T. Jena, M. Wicker, B. Chen, R. Han, Y.C. Henry, N. Hsing, A.C. Kurata, O. de Oliveira, S.A. Panaud, C.A. Jackson, M.J. Machado, M. Sanderson, D.Ware Long, R.A. Wing, Genomes of 13 domesticated and wild rice relatives highlight genetic conservation, turnover and innovation across the genus Oryza, Nat. Genet. (2018) 285–296, https://doi.org/10.1038/s41588-018-0040-0.

[10] M.L.E. Ballesfin, R.B. Vinarao, J. Sapin, S.R. Kim, K.K. Jena, Development of an intergeneric hybrid between *Oryza sativa* L. and *Leersia perrieri* (A. Camus) Launert, Breed. Sci. 68 (2018) 474–480, https://doi.org/10.1270/jsbbs.18045.

[11] M.C.N. Banaticla-Hilario, A.G. Sajise, Recent developments in wild rice conservation, research, and use, in: S. Ramamoorthy, I.J. Buot, R. Chandrasekaran (Eds.), Plant Genetic Resources, Inventory, Collection and Conservation, Springer, Singapore, 2022, https://doi.org/10.1007/978-981-16-7699-4_3.

[12] K.K. Jena, M.L. Ballesfin, R.B. Vinarao, Development of *Oryza sativa* L. by *Oryza punctata* Kotschy ex Steud. monosomic addition lines with high value traits by interspecific hybridization, Theor. Appl. Genet. 129 (10) (2016) 1873–1886, https://doi.org/10.1007/s00122-016-2745-8. Epub 2016 Jun 18.

[13] S.L. Hechanova, K. Bhattarai, E.V. Simon, G. Clave, P. Karunarathne, E.K. Ahn, C. P. Li, J.S. Lee, A. Kohli, N.R.S. Hamilton, J.E. Hernandez, G.B. Gregorio, K.K. Jena, G. An, S.R. Kim, Development of a genome-wide InDel marker set for allele discrimination between rice (*Oryza sativa*) and the other seven AA-genome *Oryza* species, Sci. Rep. 11 (2021) 8962, https://doi.org/10.1038/s41598-021-88533-9.

[14] T.A. Hall, BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT, Nucl. Acids Symp. Ser. 41 (1999) 95–98.

[15] S.R. Kim, J. Yang, G. An, K.K. Jena, A simple DNA preparation method for high quality polymerase chain reaction in rice, Plant Breed. Biotechnol. 4 (2016) 99–106, https://doi.org/10.9787/PBB.2016.4.1.99.

[16] C. Wang, X. Liu, S. Peng, Q. Xu, X. Yuan, Y. Feng, H. Yu, Y. Wang, X. Wei, Development of novel microsatellite markers for the BBCC *Oryza* genome (Poaceae) using high-throughput sequencing technology, e91826, PLoS One 9 (3) (2014), https://doi.org/10.1371/journal.pone.0091826. PMID: 24632997; PMCID: PMC3954799..

[17] J.H. Chin, Y.J. Lee, W. Jiang, H.J. Koh, M.J. Thomson, Characterization of *indica–japonica* subspecies-specific InDel loci in wild relatives of rice (*Oryza sativa* L. subsp. *indica* Kato and subsp. *japonica* Kato), Genet. Resour. Crop Evol. 64 (2017) 405–418, https://doi.org/10.1007/s10722-016-0368-1.