Genomic analysis of polycarpellary rice (Oryza sativa L.) through whole genome resequencing

Article *in* Journal of Plant Biochemistry and Biotechnology · November 2020

DOI: 10.1007/s13562-020-00602-8

6 authors, including:

Soumya Prakash Das
Visva Bharati University
5 PUBLICATIONS **31** CITATIONS
SEE PROFILE

Rahul Singh Jasrotia
Florida State University
49 PUBLICATIONS **100** CITATIONS
SEE PROFILE

Debal Deb
Centre for Interdisciplinary Studies
70 PUBLICATIONS **1,920** CITATIONS
SEE PROFILE

Mir Asif Iquebal
Indian Agricultural Statistics Research Institute
171 PUBLICATIONS **565** CITATIONS
SEE PROFILE

J. Plant Biochem. Biotechnol.
https://doi.org/10.1007/s13562-020-00602-8

ORIGINAL ARTICLE

Genomic analysis of polycarpellary rice (*Oryza sativa* L.) through whole genome resequencing

Soumya Prakash Das[1] · Rahul Singh Jasrotia[2] · Debal Deb[3] · Mir Asif Iquebal[2] · Sarika Jaiswal[2] · Narottam Dey[1]

Received: 20 March 2020 / Accepted: 3 September 2020

Abstract

There is a natural floral organ mutant of rice (var. Jugal) where the florets, popularly known as spikelet bear multiple carpels and produce multiple kernels in most of its grain. In our earlier work a detailed study has been done on its morpho-anatomical structure with allelic diversity and expression study of the major genetic loci associated with floral organ development. In present study high throughput whole genome sequencing was done which generated about of 3.7 million base pair genomic data for downstream analysis. The reads were about 101 bases long and mapped to the *Oryza sativa* var. Nipponbare as reference genome. Genome wide variant analysis detected 1,096,419 variants which included 943,033 SNPs and 153,386 InDels. A total of 24,920 non-synonymous SNPs were identified for 11,529 identified genes. Chromosome-wise distribution of uniquely mapped reads onto reference genome showed that maximum reads were mapped to 1st chromosome and least to 9th chromosome. 10th chromosome showed highest density of variations (about 325.6 per 100 kb genome sequence). Detailed sequence analysis of 23 floral organ developmental genes detected 419 potent variants where *DL* (*Drooping Leaf*) and *OSH1* (*Oryza sativa Homeobox*1) genes showed highest number (32) of variants; whereas, *MADS21* (*Minichromosome Agamous Deficient Serum Factor* 21) gene have lowest number (5) of variants. The information generated in this study will enrich the genomics of floral organ development in indica rice and cereal crops in general.

Keywords Rice · Floral organ mutants · Multiple seeded rice · Whole genome sequencing · SNP · InDels

Abbreviations

BAM	Binary alignment map
CTX	Inter chromosomal translocation
DEL	Deletion

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13562-020-00602-8) contains supplementary material, which is available to authorized users.

✉ Narottam Dey
narottam.dey@visva-bharati.ac.in

Soumya Prakash Das
soumyaprakash.das@visva-bharati.ac.in

Rahul Singh Jasrotia
rahuljasrotia86@gmail.com

Debal Deb
debaldeb01@yahoo.com

Mir Asif Iquebal
jiqubal@gmail.com

Sarika Jaiswal
sarika@icar.gov.in

[1] Rice Biotechnology Laboratory, Department of Biotechnology, Visva-Bharati, Santiniketan, West Bengal 731235, India

[2] Centre for Agricultural Bioinformatics, ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, PUSA, New Delhi 110012, India

[3] Centre for Interdisciplinary Studies, Basudha Biotechnology Laboratory for Conservation (Basudha Trust), 9 Old Calcutta Road, Barrackpore, Kolkata, West Bengal 700123, India

Published online: 05 November 2020

DL	Drooping leaf
FON	Floral organ mutant
GB	Gigabyte
INS	Insertion
INV	Inversion
IRGSP	International Rice Genome Project
ITX	Intra chromosomal translocation
MADS	M for minichromosome maintenance factor, a for agamous, d for deficient s for serum response factor
NCBI	National Center for Biotechnology Information
NGS	Next generation sequencing
OSH1	*Oryza sativa* Homeobox1
SAM	Sequence alignment map
SAM tool	Sequence alignment map tool
SNP	Single nucleotide polymorphism
SPAdes	Sequential pattern discovery using equivalence classes
SRA	Sequence alignment map
Ts	Transition
Tv	Transversion
UTR	Untranslated region
Vcf	Variant call format

Introduction

Floral structure and successful development of healthy mature grain is the most crucial phenomenon in all cereal crops including rice. Study related to mechanism of floral organ development is an interesting subject which has been studied in details for both the two model plants (*Arabidopsis* and rice), for which a number of mutants have been reported. The regulatory mechanism for floral organ number and identity are closely associated with floral meristem size which is controlled by a group of developmental genes known as FON genes. Mutation in these genes results in altered position and number of different floral parts like stamen, carpel etc. Though these natural mutants are very unique for their morpho-anatomical structure, most of these genotypes suffer from lower yield and poor quality which discourage farmers from their cultivation. One such natural mutant line from South Bengal and Odisha of Eastern India is var. Jugal which naturally produce multiple kernels (Fig. 1. a, b) in most of its grains and for the first time reported by Prain (1903) from British India. A possible genetic basis of this trait was studied by another group (Pandian and Thiyagarajan 2004), who reported this trait as Mendelian recessive trait. The next detailed study on morpho-anatomy and genetic diversity was performed by Das et al. (2018) and Priya et al. (2015) followed by expression analysis (Das et al. 2020) of selected genetic loci in respect to normal rice was done by our laboratory. This present study was aimed to understand the basis of genomic variation of this line through whole genome re-sequencing followed by comparative genomics in respect to the reference genome sequence of var. Nipponbare.

Materials and methods

Genomic DNA extraction

Seeds of rice var. Jugal having multiple kernels were germinated and genomic DNA was isolated from harvested leaves of a 10 days young seedling according to the earlier standard protocol of our laboratory (Karmakar et al. 2012). Purity of the isolated DNA was checked by resolving on 0.8% agarose gel and quality was measured by spectroscopic method (Nano-drop—JENWAY Genova Nano Ver 1.51.4).

Library construction and genome sequencing

A paired-end library was prepared from the isolated genomic DNA as per the manufacturer's protocol (TruSeq DNA PCR Free) and sequenced with Illumina's Hiseq 2500 systems (Illumina Inc. USA). The generated paired-end reads were in FASTQ format and studied through downstream analysis by considering *Oryza sativa*

Fig. 1 Matured grains (**a**) and Kernels (**b**) of Jugal

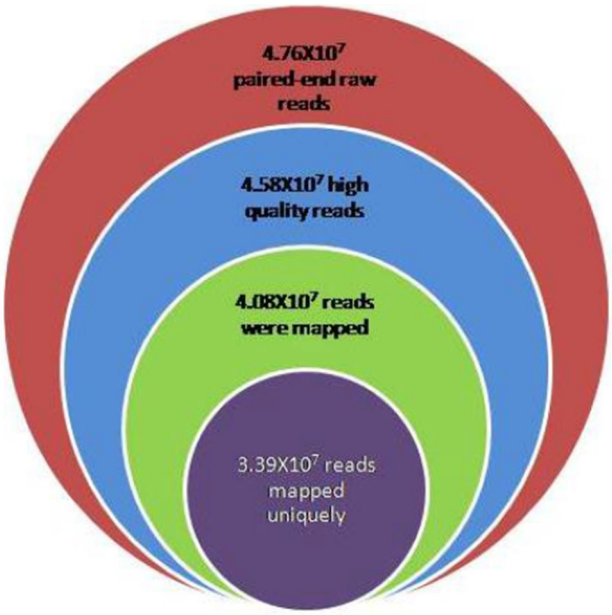

Fig. 2 Summary of the reads

L. cv. Nipponbare genome (International Rice Genome Sequencing Project 2005) as reference genome.

Pre-processing and mapping of reads

Visualization and assessment of the detected reads were performed using FastQC tool (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adaptor sequences, ambiguous read (3 bps from 5′ end) and low quality of reads (phred score < 30) were removed using Trimmomatic tool version 0.33 (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) (Bolger et al. 2014). Reference genome of *Oryza sativa* japonica cv. Nipponbare (IRGSP-1.0; GCA_001433935.1) was downloaded

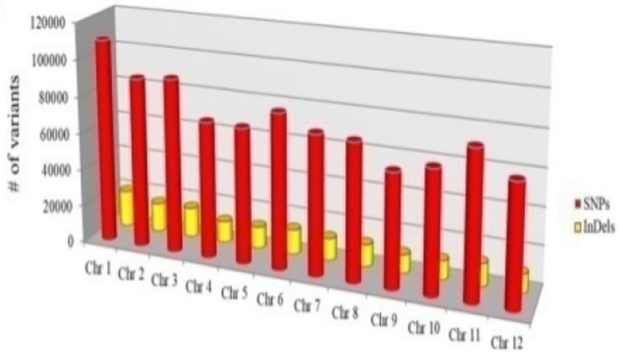

Fig. 3 Chromosome wise distribution of SNPs and InDels

from NCBI (https://www.ncbi.nlm.nih.gov/genome/genomes/10). Cleaned reads of *Oryza sativa* var. Jugal were mapped onto the reference genomes with default parameters using Burrows–Wheeler Aligner tool (BWA-MEM algorithm) (version 0.7.17) (Li and Durbin 2009). Conversion of aligned files from SAM to BAM and variants (SNPs and InDels) discovery were performed using SAMtools-v1.10 (Li et al. 2009) (The stepwise description is presented in supplementary file, SI). We have used the command line flags for generating a circular map using Circos available at http://circos.ca/intro/tabular_visualization/. Since the progam has been run on command line and not the GUI based server, it surpasses the limit of SNPs for generating circos figure. For this, all the variants were divided into various bin sizes (based on base pairs per chromosome wise) and the path was added in the "circos.conf". The commands are much available and reported in literature as well as its tutorial available at http://circos.ca/documentation/tutorials/reference/command_line/. Vcftools (version 0.1.16) (Danecek et al. 2011) were used for filtering of variants and analysis of transition to transversion ratio, read depth and read quality. Frequency

Table 1 List of chromosome number with uniquely mapped reads on *Oryza sativa japonica* cv. Nipponbare reference genome

Sl. no.	Chromosome (Accession)	*Oryza sativa japonica* cv. Nipponbare bp	Uniquely paired mapped reads
1	Chromosome 1 (NC_029256.1)	43270923	3966575
2	Chromosome 2 (NC_029257.1)	35937250	3358620
3	Chromosome 3 (NC_029258.1)	36413819	3550428
4	Chromosome 4 (NC_029259.1)	35502694	3007257
5	Chromosome 5 (NC_029260.1)	29958434	2816056
6	Chromosome 6 (NC_029261.1)	31248787	2829818
7	Chromosome 7 (NC_029262.1)	29697621	2637964
8	Chromosome 8 (NC_029263.1)	28443022	2593109
9	Chromosome 9 (NC_029264.1)	23012720	2002568
10	Chromosome 10 (NC_029265.1)	23207287	2208380
11	Chromosome 11 (NC_029266.1)	29021106	2433673
12	Chromosome 12 (NC_029267.1)	27531856	2290523

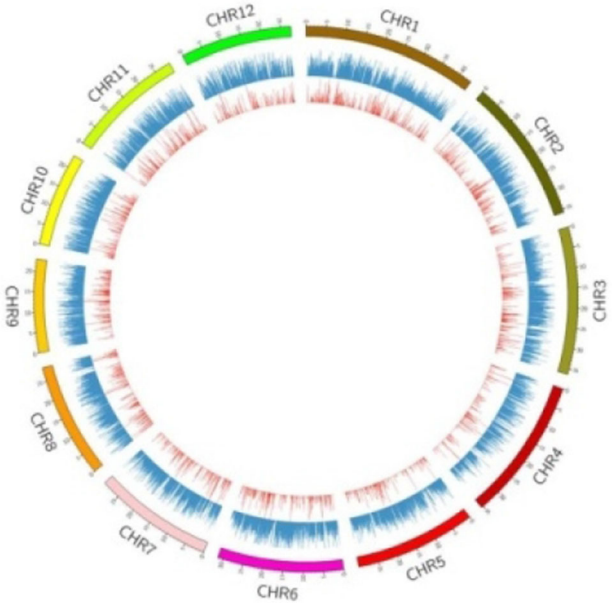

Fig. 4 Circular plot presenting variations across the rice chromosome

distribution of variants per 100 kb were analysed for each rice chromosome. Chromosome wise distribution of variants (SNPs and InDels) was plotted using Circos tool (0.69–6 s) (Krzywinski et al. 2009). Structural variation analysis was done with BreakDancer version 1.1 (http://breakdancer.sourceforge.net/) (Chen et al. 2009). Further, annotation of variants and their function was performed using SnpEff (version 4.3T) toolbox (Cingolani et al. 2012) which provided a number of selected information (splice sites, intergenic, intragenic, 5′UTRs, 3′UTRs, exons, introns, downstream region and upstream regions).

De novo assembly, annotation and variants analysis of unmapped reads

De novo assembly of unmapped reads was performed using SPAdes assembler (version 3.13.0) (Bankevich et al. 2012) with default parameters. Homology search was performed against NCBI NR (non-redundant) database using blastx program (Camacho et al. 2009). Gene ontology and annotation were carried out using Blast2GO software (version 5.2.5) (Conesa et al. 2005), which categorized unigenes into three sub divisions such as biological process, cellular component and molecular functions. Variants were identified from de novo assembly of unmapped reads using SAMtools and Vcftools.

Results

Pre-processing and mapping of reads

Whole genome sequencing using NGS technology generated a total 4.09 GB high quality sequence data. Total 4.76×10^7 paired-end raw reads with 101 bp length were obtained with an average of 10 × depth coverage. After quality check 4.58×10^7 (96.2%) cleaned and high-quality reads were identified and studied for downstream analysis including variant analysis. Around 4.08×10^7 (89.02%) reads were mapped properly onto the reference genome and out of which 3.39×10^7 reads were mapped uniquely which were further used for variant identification analysis. A summary of the read statistics is graphically presented Fig. 2. The aligned reads were converted to BAM file and submitted to the Sequence Read Archive (SRA) at NCBI database with the accession number SRP131720. Analysis of mapped reads showed that maximum reads were mapped to chromosome 1, followed by chromosome 3 and 2, and minimum for chromosome 9. Chromosome wise comparison of aligned reads in respect to Japonica (var. Nipponbare) is presented in Table 1. The unmapped reads were subjected to de novo assembly using SPAdes assembler.

Variant (SNP/InDels) discovery

After removing the false positive variants, a total of 1,096,419 variants were obtained among which 962,268 and 134,151 were homozygous and heterozygous respectively and categorically 943,033 were SNPs and 153,386 were InDels. Out of the 153,386 InDels, 75,051 were insertion and 78,335 were deletions. Chromosome 1 has the maximum number of variants i.e. 129,870 followed by Chromosome 3 (109,464 variants) and Chromosome 2 (107,230 variants); least number of variants were found in chromosome 9 (only 70,611). Chromosome wise distribution of the detected variants is presented in Figs. 3 and 4 respectively. Density distribution of the detected variants across different chromosomes were calculated per 100 kb, in which chromosome 10 has the maximum density i.e. 325.6, followed by 311.6 and 308.4 in chromosome 11 and 6 respectively (Table 2). Chromosome 4 has the lowest density i.e. 238.6. Structural variant analysis detected 12,143 structural variants (SV) which included 896 intra-chromosomal translocation (ITX), 3895 inter-chromosomal translocation (CTX), 6754 deletion (DEL), 459 inversion (INV), 87 unknown (UN) and 52 insertion (INS) (Fig. 5).

Table 2 Chromosome wise variant distribution and density per 100 kb

Sl. no.	Chromosome (Accession)	*Oryza sativa japonica* cv. Nipponbare bp	Total variants (density per 100 kb)	SNPs (density per 100 kb)	InDels (density per 100 kb)
1	Chromosome 1 (NC_029256.1)	43270923	129,870 (300.1)	110,196 (254.7)	19,674 (45.5)
2	Chromosome 2 (NC_029257.1)	35937250	107,230 (298.4)	91,546 (254.7)	15,684 (43.6)
3	Chromosome 3 (NC_029258.1)	36413819	109,464 (300.6)	93,481 (256.7)	15,983 (43.9)
4	Chromosome 4 (NC_029259.1)	35502694	84,725 (238.6)	73,253 (206.3)	11,472 (32.3)
5	Chromosome 5 (NC_029260.1)	29958434	84,282 (281.3)	72,643 (242.5)	11,639 (38.9)
6	Chromosome 6 (NC_029261.1)	31248787	96,356 (308.4)	83,055 (265.8)	13,301 (42.6)
7	Chromosome 7 (NC_029262.1)	29697621	86,599 (291.6)	74,765 (251.8)	11,834 (39.8)
8	Chromosome 8 (NC_029263.1)	28443022	85,179 (299.5)	73,370 (258)	11,809 (41.5)
9	Chromosome 9 (NC_029264.1)	23012720	70,611 (306.8)	60,850 (264.4)	9,761 (42.4)
10	Chromosome 10 (NC_029265.1)	23207287	75,563 (325.6)	65,582 (282.6)	9,981 (43)
11	Chromosome 11 (NC_029266.1)	29021106	90,428 (311.6)	78,460 (270.4)	11,968 (41.2)
12	Chromosome 12 (NC_029267.1)	27531856	75,075 (272.7)	64,922 (235.8)	10,153 (36.9)
13	Mitochondrion genome (NC_011033.1)	490520	66 (13.5)	50 (10.2)	16 (3.3)
14	Other genomic regions	–	971	860	111
15	Total		1,096,419	943,033	153,386

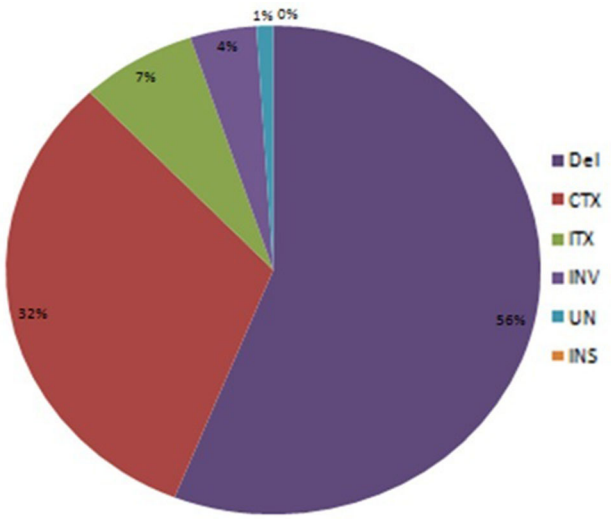

Fig. 5 Distribution of structural variation across genome

Analysis of variants (SNPs and InDels)

Among the 943,033 SNPs detected, 668,660 were transitions (T/C and G/A) and 274,373 were transversion (T/G, C/G, T/A, and C/A). The ratio of transition to transversion was 2.44. The frequency of transversions was comparatively lower than transitions. In transition, maximum SNPs were C to T which is 181,812 in number and A to G was minimum (152,353) whereas in case of transversion, T to A is maximum in count i.e. 39,127 and C to G has 25,035 which is minimum in count. The Category wise distribution of the detected variants is presented in Table 3. In InDels, the insertions variants are ranged from 1 to 29 bp and deletions are ranged in between − 1 and − 44 bp (Fig. 6).

Variants annotation

Annotation of the variants against the reference genome detected 821,472 (about 16%) variants are in intergenic

Table 3 Category wise distribution of the detected variants

Total number of variants (1096419) (Through Reference based assembly)					
Homozygous (962268)					Heterozygous (134151)
SNP (943033)					INDELS (153386)
TS (668660)			TV (274373)		
C/T (181812)	A/G (152353)	T/C and G/A (334495)	T/A (39127)	C/G (25035)	A/T and G/C (210211)

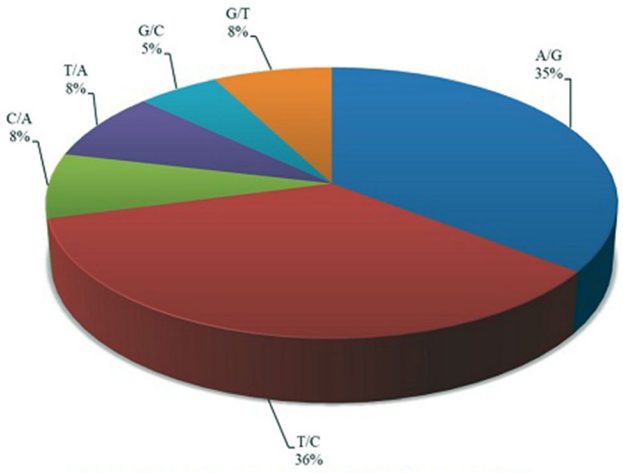

Fig. 6 Transitions/transversion (Ts/Tv) distribution in SNPs

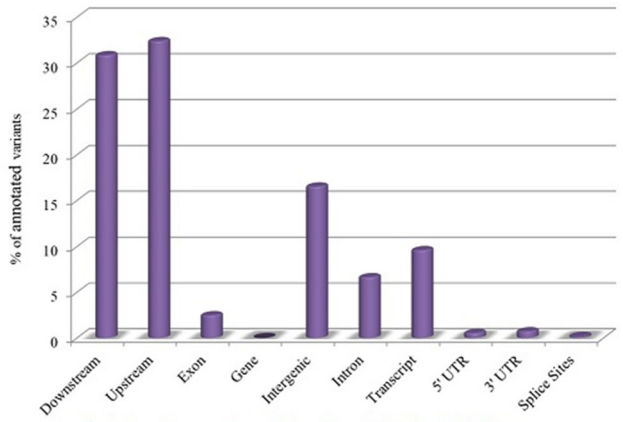

Fig. 7 Annotation and distribution of Variants (SNPs and InDels)

region and 123,739 (3%) in exon region. In addition to this, 40,408 missense variants, 508 nonsense, and 34,924 silent variants were detected. Ratio of Missense/Silent ratio was 1.157 (Fig. 7). Out of 1,096,419 variants, 24,920 were non-synonymous SNPs which were found to be distributed in 11,529 genes. SNPs per gene in non-synonymous SNPs were ranged from 1 to 49.

De novo assembly of unmapped reads

Total 1,601,994 paired end unmapped reads were subjected to de novo assembly which generated 16,585 contigs with minimum read length 56 bp with kmer length of 55 bp. Contigs with sequence length less than 200 bp were removed and 16,159 contigs were retained with N50 of 977 bp. Maximum contigs have length in between 200 and 299 i.e. 6653, followed by 2430 and 1343 in base pair ranged 300 to 399 and 400 to 499, respectively. Out of 16,159 contigs, 9075 contigs were not showed any similarity against NR database, while 7084 contigs matched with known genes. Maximum hits were found to *Oryza sativa Japonica* and *Oryza sativa Indica* with 1478 and 1298 hits respectively. Variant study was mined against de novo assembly of unmapped reads. A total of 2772 variants were detected with 2636 SNPs and 136 InDels. Out of 2772 variants, 2542 and 230 were heterozygous and homozygous, respectively. Out of 2636 SNPs, 1670 and 966 were transitions (Ts) and transversion (Tv), respectively with Ts/Tv ratio of 1.73. In case of InDels, 87 were insertions and 49 were deletions. In insertion and deletion both were ranged from 1 to 5 and − 1 to − 5.

Analysis of the variants in the floral organ developmental genes

Analysis of the variations of the 23 flowering genes shows that the 23 flowering genes in Jugal contained 419 SNPs. The detail information of the genes studied was provided in Table 4 and the of variants in the studied genes presented in Table 5. Among the 23 floral organ developmental genes studied, *DL* and *OSH1* genes found to have highest number of variants (32 variants) whereas *MADS21* gene had lowest number of variants (5 variants).

Discussion

Chromosome-wise distribution of uniquely mapped reads onto reference genome showed that maximum reads were mapped to chromosome 1 while Chromosome 9 has the

Table 4 Details of the flowering related genes studied

Sl. no.	Gene names	RAP-DB Accession no.	Function
1	*FON1*	Os06g0717200	Enlargement of floral meristem
2	*FON2*	Os11g0595400	Enlargement of floral meristem
3	*DL*	Os03g0215200	Carpel development
4	*LOG*	Os01g0588900	Maintenance of floral meristem
5	*OSH1*	Os03g0727000	Maintenance of reproductive meristem
6	*SNB*	Os07g0235800	Maintenance of floral meristem
7	*OsMADS1*	Os03g0215400	Class E gene
8	*OsMADS2*	Os01g0883100	Class B gene
9	*OsMADS3*	Os01g0201700	Class C gene
10	*OsMADS4*	Os05g0423400	Class B gene
11	*OsMADS5*	Os06g0162800	Class E gene
12	*OsMADS6*	Os02g0682200	Floral organ development
13	*OsMADS7*	Os08g0531700	Class E gene
14	*OsMADS8*	Os09g0507200	Class E gene
15	*OsMADS13*	Os12g0207000	Class D gene
16	*OsMADS14*	Os03g0752800	Class A gene
17	*OsMADS15*	Os07g0108900	Class A gene
18	*OsMADS16*	Os06g0712700	Class B gene
19	*OsMADS17*	Os04g0580700	Floral organ development
20	*OsMADS18*	Os07g0605200	Class A gene
21	*OsMADS21*	Os01g0886200	Class D gene
22	*OsMADS34*	Os03g0753100	Class E gene
23	*OsMADS58*	Os05g0203800	Class C gene

minimum number of mapped reads. Analysis of mapped reads also showed that the detected variants were diversely distributed in the Jugal genome. Although, Chromosome 1 has the maximum number of variants (129,870 variants) and chromosome 9 (70,611) possesses least number of variants, density of variants per 100 kb varies within chromosomes. Chromosome 10 has highest density of variants per 100 kb (325.6), i.e. the 23,207,287 bp long Chromosome 10 possesses 75,563 variants, Chromosome 4 has the lowest density per 100 kb (238.6). These uneven distributions of SNPs and InDels indicate presence of regions of low SNPs, which was reported previously (Wang et al. 2009; Hu et al. 2014; Singhabahu et al. 2017) and described as "SNP deserts". It is hypothesised that SNP deserts possess highly conserved functions and also are key sites in the process of domestication. The number of transition SNPs were identified higher than transversion SNPs in Jugal genome which ranges between 2.0 and 2.53. The higher frequency of transitional SNPs over transversion SNPs was also reported previously (Subbaiyan et al. 2012; Hu et al. 2014; Hwang et al. 2014) and designated as "transition bias". The higher ratio of transitional SNPs over transversion SNPs indicates conservation of the protein structures to facilitate conformational advantages in natural selection (Wakeley 1996; Rathinasabapathi et al. 2015; Zhang et al. 2016). Earlier works on the mutant rice (var. Jugal) showed that it possesses increased number of carpels and enlarged floral meristems in comparison to normal rice (Das et al. 2018). Thus, from the whole genome sequencing data, the Jugal flowering genes were further investigated. A total of twenty-three genes were investigated in this present study; all the genes were involved in either floral organ development or floral meristem development and maintenance. Among the twenty genes investigated, fifteen genes namely *OsMADS1*, *OsMADS2*, *OsMADS3*, *OsMADS4*, *OsMADS5*, *OsMADS7*, *OsMADS8*, *OsMADS13*, *OsMADS14*, *OsMADS15*, *OsMADS16*, *OsMADS18*, *OsMADS21*, *OsMADS34,* and *OsMADS58* belong to ABCDE model. Three genes (LOG, *OSH1*, and *SNB*) studied were associated with development and maintenance of floral meristem. Two rice *AGL6* genes namely *OsMADS6* and *OsMADS1*, two genes responsible for increased floral meristem—*FON1* and *FON2*, and a carpel marker gene—*DL* were also investigated. *DROOPING LEAF (DL)* is a carpel marker gene (Nagasawa et al. 1996; Yamaguchi et al. 2004) and the Jugal flowers bear increased number of carpels. Among the genes studied, the presence of highest number of SNPs in Jugal *DL* gene indicates that the Jugal *DL* is rich in natural variations and will be the potential candidate gene responsible for

Table 5 Summary of variants in the selected 23 flowering genes

Sl. no.	Gene name	Gene_ID	Total no. of variants
1	*FON1*	Os06g0717200	9
2	*FON2*	Os11g0595400	13
3	*DL*	Os03g0215200	32
4	*LOG*	Os01g0588900	11
5	*SNB*	Os07g0235800	25
6	*OSH1*	Os03g0727000	32
7	*OsMADS1*	Os03g0215400	23
8	*OsMADS2*	Os01g0883100	7
9	*OsMADS3*	Os01g0201700	20
10	*OsMADS4*	Os05g0423400	7
11	*OsMADS5*	Os06g0162800	17
12	*OsMADS6*	Os02g0682200	23
13	*OsMADS7*	Os08g0531700	16
14	*OsMADS8*	Os09g0507200	21
15	*OsMADS13*	Os12g0207000	15
16	*OsMADS14*	Os03g0752800	26
17	*OsMADS15*	Os07g0108900	24
18	*OsMADS16*	Os06g0712700	13
19	*OsMADS17*	Os04g0580700	22
20	*OsMADS18*	Os07g0605200	18
21	*OsMADS21*	Os01g0886200	5
22	*OsMADS34*	Os03g0753100	20
23	*OsMADS58*	Os05g0203800	20

increased number of carpels. Whereas *Oryza sativa* Homeobox1 (*OSH1*) is a marker of shoot apical and reproductive meristem cells (Sato et al. 1996; Hu et al. 2015). Thus, similar like the *DL* gene, the presence of highest number of SNPs in Jugal *OSH1* gene indicates its richness in natural variations and will be responsible for enlargement of Jugal floral meristems. The BAM (Binary Alignment Map) files generated from the aligned reads were submitted to the Sequence Read Archive (SRA) data base of NCBI as a Bioproject with the accession number SRP131720.

Conclusion

Whole genome sequencing and genome wide variant analysis of the Jugal rice has provided insights to the possible genetic basis of its multiple carpelled and increased floral meristems trait which lead to development of multiple seeded rice. The present investigation detected several SNPs in Jugal floral organ developmental genes. To correlate these SNPs with their function, in silico functional predictions of the selected proteins is presently underway. Sequence variation study of the selected 23 flowering genes will further confirm and validate the findings resulted from the Bioinformatics study performed. Furthermore, performing functional analysis and investigating crystal structures of the proteins would be the future focuses of this research.

Acknowledgements The work was financially supported by SERB, Department of Science and Technology, Government of India in form of a research project (Ref. No. SB/YS/LS-187/2013).

References

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA (2012) SPAdes a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19(5):455–477

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, BealerK MT (2009) BLAST + : architecture and applications. BMC Bioinform 10:421

Chen K, Wallis JW, McLellan MD, Larson DE, Kalicki JM, Pohl CS, McGrath SD, Wendl MC, Zhang Q, Locke DP, Shi X, Fulton RS, Ley TJ, Wilson RK, Ding L, Mardis ER (2009) BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods 6(9):677–681

Cingolani P, Platts A, le Wang L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. Fly (Austin) 6(2):80–92

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M (2005) Blast2GO: a universal tool for annotation, visualizationand analysis in functional genomics research. Bioinformatics 21:3674–3676

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R, 1000 Genomes Project Analysis Group (2011) The variant call format and VCFtools. Bioinformatics 27(15):2156–2158

Das SP, Deb D, Dey N (2018) Micromorphic and molecular studies of floral organs of a multiple seeded rice (*Oryza sativa* L.). Plant Mol Biol Report. https://doi.org/10.1007/s11105-018-1116-9

Das SP, Deb D, Dey N (2020) Expression study of five genes involved in floral organ development in multiple seeded rice. J Plant Biochem Biotechnol 29(2):348–351. https://doi.org/10.1007/s13562-019-00526-y

Hu Y, Mao B, Peng Y, Sun Y, Pan Y, Xia Y, Sheng X, Li Y, Tang L, Yuan L, Zhao B (2014) Deep re-sequencing of a widely used maintainer line of hybrid rice for discovery of DNA polymorphisms and evaluation of genetic diversity. Mol Genet Genom 289:303–315

Hu Y, Liang W, Yin C, Yang X, Ping B, Li A, Jia R, Chen M, Luo Z, Cai Q, Zhao X, Dabing Zhang D, Yuan Z (2015) Interactions of OsMADS1 with floral homeotic genes in rice flower development. Mol Plant 8(9):1366–1384

Hwang SG, Hwang JG, Kim DS, Jang CS (2014) Genome-wide DNA polymorphism and transcriptome analysis of an early-maturing rice mutant. Genetica 142:73–85

International Rice Genome Sequencing Project (2005) The map-based sequence of the rice genome. Nature 436:793–800. https://doi.org/10.1038/nature03895

Karmakar J, Roychowdhury R, Kar RK, Deb D, Dey N (2012) Profiling of selected indigenous rice (*Oryza sativa* L.) landraces of Rarh Bengal in relation to osmotic stress tolerance. Physiol Mol Biol Plants 18:125–132

Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA (2009) Circos: an information aesthetic for comparative genomics. Genome Res 19(9):1639–1645

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 25:1754–1760

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079

Nagasawa N, Miyoshi M, Nagato Y (1996) DL regulates both leaf and pistil development in rice. Rice Genet Newsl 13(12):102–105

Pandian RT, Thiyagarajan K (2004) Inheritance of floral traits in spontaneous mutant in rice (*Oryza sativa* L.). Curr Sci 87:1051–1052

Prain D (1903) Bengal plants, vol 2. Botanical Survey of India, Calcutta, p 1184

Priya A, Das SP, Goswami S, Adak MK, Deb D, Dey N (2015) An exploratory study on allelic diversity for five genetic loci associated with floral organ development in rice. Am J Plant Sci 6:1973–1980

Rathinasabapathi P, Purushothaman N, Ramprasad VL, Parani M (2015) Whole genome sequencing and analysis of Swarna, a widely cultivated indica rice variety with low glycemic index. Sci Rep 5:11303. https://doi.org/10.1038/srep11303

Sato Y, Hong SK, Tagiri A, Kitano H, Yamamoto N, Nagato Y, Matsuoka M (1996) A rice homeobox gene, OSH1, is expressed before organ differentiation in a specific region during early embryogenesis. Proc Natl Acad Sci U S A 93:8117–8122

Singhabahu S, Wijesinghe C, Gunawardana D, Senarath-Yapa MD, Kannangara M, Edirisinghe R, Dissanayake VHW (2017) Whole genome sequencing and analysis of Godawee, a salt tolerant indica rice variety. J Rice Res 5:1. https://doi.org/10.4172/2375-4338.1000177

Subbaiyan GK, Waters DL, Katiyar SK, Sadananda AR, Vaddadi S, Henry RJ (2012) Genome-wide DNA polymorphisms in elite indica rice inbreds discovered by whole-genome sequencing. Plant Biotechnol J 10:623–634. https://doi.org/10.1111/j.1467-7652.2011.00676.x

Wakeley J (1996) The excess of transitions among nucleotide substitutions: new methods of estimating transition bias underscore its significance. Trends Ecol Evol 11:158–162

Wang L, Hao L, Li X, Hu S, Ge S (2009) SNP desert of Asian cultivated rice: genomic regions under domestication. J Evol Biol 22:751–761

Yamaguchi T, Nagasawa KS, Matsuoka M, Nagato Y, Hirano HY (2004) The YABBY gene DROOPING LEAF regulates carpel specification and midrib development in *Oryza sativa*. Plant Cell 16:500–509

Zhang F, Xu T, Mao L, Yan S, Chen X, Wu Z, Chen R, Luo X, Xie J, Gao S (2016) Genome-wide analysis of Dongxiang wild rice (*Oryza rufipogon* Griff.) to investigate lost/acquired genes during rice domestication. BMC Plant Biol 16:103