SSR Marker Based Genetic Diversity Analysis and SNP Haplotyping of Genes Associating Abiotic and Biotic Stress Tolerance, Rice Growth and Development and Yield Across 93 Rice Landraces

Smitha Kunhiraman Vasumathy

Central University of Kerala

Manickavelu Alagu (✉ amanicks@cukerala.ac.in)

Central University of Kerala <https://orcid.org/0000-0003-2875-2290>

Research Article

Keywords: Rice landraces, SSR, Genetic diversity, SNP, Haplotype

Posted Date: April 29th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-466659/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Molecular Biology Reports on July 28th, 2021. See the published version at <https://doi.org/10.1007/s11033-021-06595-4>.

Abstract

I. Background: As rice is the staple food for more than half of the world population, enhancing grain yield irrespective of the variable climatic conditions is indispensable. Many of the traditionally cultivated rice landraces are well adapted to severe environmental conditions and have high genetic diversity that could play an important role in crop improvement.

II. Methods and Results: The present study disclosed high level of genetic diversity among the unexploited rice landraces cultivated by farmers of Kerala. Twelve polymorphic markers detected a total of seventy- seven alleles with an average of 6.416 alleles per locus. PIC value ranged from 0.459 to 0.809 and to differentiate the rice genotypes, RM 242 was found to be the most appropriate marker with the highest value of 0.809. The current study indicated that the rice landraces were highly diverse with higher values of the effective number of alleles, PIC, and Shannon information index and utilizing these informative SSR markers for future molecular characterization and population genetic studies in rice landraces are advisable. Haplotypes are sets of genomic regions within a chromosome that are inherited together and haplotype-based breeding is a promising strategy for designing next-generation rice varieties. Here, haplotype analysis explored 270 haplotype blocks and 775 haplotypes from all the chromosomes of landraces under study. The number of SNPs in each haplotype block ranged from two to 28. Haplotypes of genes related to biotic and abiotic stress tolerance, yield-enhancing, and growth and development in rice landraces were also elucidated in the current study.

III. Conclusions: The present investigation revealed genetic diversity of rice landraces and the haplotype analysis will open the way for genome wide association studies, QTL identification, and marker assisted selection in the unexplored rice landraces collected from Kerala.

Introduction

Growing human population, unhinged environmental conditions and decreasing agricultural resources have severe impact on global crop production. The current global yield increase rate (0.9-1.3% per year) of major crops like rice, wheat, maize, etc., is inadequate to meet the food demand for the estimated nine billion people in 2050 [1]. Rice yield need to be increased by almost 40% by 2030 to satisfy the growing demand without affecting the resource adversely [2]. As rice is one of the staple food crops, the ultimate way to increase yield within the obtainable agricultural lands is genetic improvement in rice.

Landraces are natural cultivars of farmers with adaptability to varying environments and evolve to generate new genetic traits that address the challenge against the biotic and abiotic stresses with copious diversity [3]. Rice landraces are crucial requirement for food security, but the germplasm of the rice gene pool is poorly understood [4]. Molecular markers are a powerful means of determining genotypes and investigating genetic diversity. Among the available molecular marker systems, microsatellite loci, or simple sequence repeats (SSR) and SNP's became important DNA markers as they display a high allelic variance between organisms. In the present scenario, genetic information of rice

landraces is essential for improving the existing crop varieties with recent molecular techniques to defend ourselves from food scarcity through the concept of sustainable agricultural development.

Haplotypes are sets of genomic regions within a chromosome which tend to be inherited together due to the low probability of recombination among them in the given population [5]. Haplotype blocks consist of a combination of adjacent allelic markers. The number of polymorphisms in the haplotype blocks developed by the combination of SNP alleles represents the genetic diversity in rice [6]. Determination of SNP in the regions linked with a particular trait increases the probability of detecting variants located on the same haplotype and develops the accuracy of identifying candidate genes in association mapping studies [7]. The principles of genomic selection (GS) has been used considerably in genetic improvement programs of complex traits [8]. Recent advances in DNA sequencing and genotyping techniques help the plant breeding community to achieve high yield even in the incompatible environments. GS can be executed for plant improvement programs using individual SNPs and haplotypes for complex traits.

Plants are sessile and struggle to cope up environmental stresses, which recurrently limit the overall plant growth. Haplotype analysis has been conducted on several important genes from diverse rice genomic sequences like quality and yield traits [9], disease resistance [10], flowering time [9], etc. As haplotypes are genomic sequence variants co-occur in a chromosome, inherited genetic variations have created diversity among these haplotypes [11]. So haplotype identification and analysis will explain the genetic patterns with higher accuracy, and produce knowledge for effective utilization of the existing rice germplasm.

Generally, excessive inbreeding has resulted in decreasing genetic diversity and creating a more uniform genome pattern in rice. To overcome the scenario, genome or haplotype shuffling should be carefully considered in future rice breeding. Our objectives are to explore the genetic diversity and characterize the novel haplotypes exhibiting association with different stress tolerance, growth and development, and yield-enhancing properties in the rice landraces collected from Kerala.

Materials And Methods

1. Plant material and DNA extraction

For the current study ninety-three rice landraces were collected from the northern region of Kerala (Supplementary Table S1). Genomic DNA was isolated from two weeks old plant leaf sample using CTAB plant DNA extraction protocol [12] and DNA quality check was carried out in Nanodrop by determining A260/280 ratio followed by visualization in agarose gel electrophoresis.

2. SSR markers and PCR amplification

Twelve different SSR markers having high polymorphism information content distributed across the chromosomes were selected for the genetic diversity analysis (Supplementary Table S2). The PCR amplification was carried out in 10 µl reaction mixture containing Origin Taq PCR mixture (2x) – 2.5 µl, Forward primer (10pmole/ µl) - 0.5 µl, Reverse primer (10pmole/ µl) - 0.5 µl, Template DNA (<200ng) - 1 µl

and 5.5 µl distilled water, using a thermal cycler. PCR amplification was done using the following temperature profile, 94°C for 5 minutes; 2*(94°C for 1 minute, 65°C for 1 minute and 72°C for 2 minutes); 2*(94 °C for 1 minute, 62°C for 1 minute and 72°C for 2 minutes); 4 *(94°C for 1 minute. 59°C for 1 minute and 72°C for 2 minutes); 25 *(94°C for 1 minute. 57°C for 1 minute and 72°C for 2 minute; 72°C for 2 minutes); 72°C for 5 minutes. Amplification products along with 100bp DNA ladder were resolved on 1% agarose gel using horizontal gel electrophoresis unit and the DNA fragments were visualized and documented using Gel documentation system.

3. SSR data analysis

The size of amplified fragments was estimated using 100bp ladder and only unambiguous bands of SSR markers were scored. Polymorphic information content (PIC) values were calculated for each SSR primer pairs based on the presence or absence of bands, producing a binary data matrix of 1 or 0 for each marker. Genetic Diversity analysis using SSR markers, analysis of allelic variability and major allelic frequency and PIC is computed using PowerMarker V3.0 [13]. Genetic diversity parameters like effective allele number (Ne), Shannon's information index (I), Expected heterozygosity (He) and Fixation index (F) were computed. Cluster analysis based on the genetic distance matrix using Neighbor- Joining (NJ) algorithm implemented in Darwin v6.0.018 [14] was performed. Principal Coordinate Analysis (PCoA) was done using GenAlex 6.5 [15].

4. Genotyping and haplotype block construction

In our previous study, Genotyping by sequencing was performed on Illumina Nextseq500 platform with 2 x 150 bp v2 chemistry [16]. Monomorphic SNPs and SNPs with a call rate <90% and SNPs with a minor allele frequency (MAF) <0.05 were also excluded from the analysis.

The software haploview v.4.2 [17] was used to estimate the haplotype block structure in the 93 rice genotypes with the command: java -jar Haploview.jar -memory 4096

Input data file (.ped file) and locus information file (.info file) was created by using PLINK v1.90b6.21. Based on the confidence interval algorithm, the haplotype blocks were constructed [18] with the default parameters : 'ignore pair-wise comparisons of markers >500kb apart, exclude individuals with >50% of missing genotypes'.

5. Decoding of haplotype linked rice functional genes

To analyze the variants and predict the functional consequence of the variants, the chromosome number and marker position in the tagger output data of each chromosome were uploaded on Variant Effect Predictor (VEP) tool in the Ensembl Plants dataset (<http://plants.ensembl.org>). Ensembl results were compared with Gramene database (<http://www.gramene.org>) for getting curated and integrative information about the genes which are depicted in Supplementary Table S3. To characterize the selected genes, functions of each haplotype linked genes were decoded by using Rice Annotation Project

Database (RAP-DB) (<http://rapdb.dna.affrc.go.jp/>) and funRiceGenes database (<https://funricegenes.github.io/>) which will promote the functional genomics studies in rice landraces.

Results

1. Polymorphism and marker efficiency

The current study evaluated the genetic diversity of 93 rice landraces. Here, the 12 SSR markers screened to assess the genetic variability and relatedness were produced reproducible and polymorphic bands. A total of 77 alleles were detected among rice genotypes with an average of 6.416 alleles per locus. The number of alleles per locus ranged from 4 (RM232 & RM248) to 9 (RM 206 & RM242), which indicates the richness of the population (Table 1A). The PIC value for SSR markers was ranged from 0.459 (RM232) – 0.809 (RM 242) with an average of 0.658. Markers RM242, RM26, RM228, and RM206 were the most informative primers on the basis of highest PIC of 0.809, 0.764, 0.763, and 0.688 respectively. Among the 12 SSR markers used in this study, eleven markers had PIC value exceeding 0.5 (Table 1A).

From the entire microsatellite loci analyzed in the study, the mean value of Fixation index (F), Shannon's information index (I), number of effective alleles (Ne), number of different alleles (Na) were 0.98, 1.23, 3.02, 5.5 respectively. The heterozygosity value ranged from 0.41 (RM 349) to 0.8 (RM 242) (Table 1B). The genotypes showed Shannon information index ranged from 0.8 to 1.7 and the highest 'I' value was exhibited by RM242 which was having the maximum PIC value also. From the SSR markers evaluated in this study, RM242, RM206, RM48 produced maximum Na and Ne.

2. Genetic relatedness using cluster analysis and Principal co ordinate analysis

Cluster analysis based on microsatellite allelic diversity clearly demarcated rice landraces into four groups based on genetic distance and dissimilarity matrix and from which 'Jaiva' and 'Mavilan' showed the highest genetic differences by indicating genetic dissimilarity coefficient greater than 0.5. Cluster I, which is the largest cluster, comprised of Black jasmine, Veluthan, Choverian, Anamodan, Chenkayama, Karakayama etc, Cluster II encompassed Rajakayama, Allikkannan, Assam black, Palliyaral, Okkapuncha etc. The third cluster included Rakthasali, Kunjinellu, Chettuchomala, Rasagatham, Chennellu etc. whereas cluster IV consist of Urunikayama, Karuthanjavara, Manja njavara (Fig.1A).

PCoA with SSR markers showed that large diversity existed in the rice genotypes. The first three principal coordinates explained 26.56%, 14.79%, 12.17% individually and overall, rice genotypes possess 54% of cumulative variation in their genotypes (Fig. 1B). The genotypes viz. Chennellu, Chenkayama, Mavilan, Palliyaral, Kalajeera, Kooran, White Basmati, etc. were found placing far away from the centroid of the cluster. Whereas the genotypes viz. Poonaran, Kavunni, Rasagatham, Kannichennellu, Karutha allikkannan, etc. were placed more or less around the centroid.

3. Haplotype block construction

In our recent study, Genotyping by sequencing of germplasm used for the current study yielded an average of 20 million reads per sample with an average depth on the reference genome (without Ns) in 5.37X to 13.01X range [16]. Haploblocks were constructed by using SNPs obtained from GBS of 93 rice landraces. A total of 270 haplotype blocks and 775 haplotypes were identified from all the chromosomes (Table 2). Out of 79,953 filtered SNPs, 1402 tag SNPs were grouped into haplotype blocks. The maximum number of haplotype blocks was determined by combinations of SNPs located on chromosome 8 (n=31). A lesser number of blocks were constructed by SNPs located on chromosome 10 (n=16). From the total blocks formed by the chromosomes, 775 haplotypes were obtained, while each block had three haplotype variants on average. The number of SNPs in each haplotype block ranged from two to 28. The largest haplotypes had a length of 499 kb obtained from chromosome 12.

4. Analysis of haplotype linked genes

Several genes determining various traits have been reported in rice. But the haplotypes for developing supreme variety from the landraces collected from Kerala remain intangible. To elucidate the relationship between haplotypes and traits, selected SNPs related to different genes and their corresponding haplotypes were described in (Fig.2). The functions of each gene were explored from the Rice Annotation Project Database and FunRiceGenes database.

4.1 Haplotypes of abiotic stress tolerance genes

Plant, being sessile organisms cannot escape from the extremes in environmental factors which can results in stressful conditions. These inauspicious effects can inhibit plant growth and development, and limit crop productivity up to 70%. Abiotic stress tolerance is a multigenic trait and we identified the genes, *OsRZFP34*, *OsNADP-ME2*, *OsDHSRP1*, *OsSIRH2-14*, *OsGRAS23*, *CK2alpha2*, *OsSWI3C/OsCHB705*, *OsAIR2* and corresponding haplotypes responsible for tolerance against different abiotic stresses (Table 3A, Fig.3A).

4.2 Haplotypes of biotic stress tolerance genes

Biotic stresses include pathogens, insects, pests and weeds, leads severe yield losses or crop failure during infestation. During the past few decades, an ample variety of genes concerned in rice defense response have been recognized. But most of these cloned genes effective against one or a few strains of the pathogen and is effective only for few years. Therefore, diverse rice landrace is an ideal option to broaden the rice gene pool. A few genes, *OsPSKR1*, *OsLOX10*, *Pita/Pi-4a* and their haplotypes involved in biotic stress tolerance in rice landraces are depicted here in Table 3B Fig.3B.

4.3 Haplotypes of genes related to rice growth and development

Plant growth and development comprise a continuous process which is highly compatible to the change in environmental conditions. By regulating cell proliferation and expansion, several phytohormones control plant growth and these phytohormones regulate meristem function and organogenesis which entails controlling cell division. Some of the genes involved in plant growth and development, *RAD51C*,

HAZ1/HOX1a, *TIG1*, *WEG1*, *OspTAC2*, *OsABA2*, *OsCCC2* and their haplotypes are listed in Table 3C, Fig.3C.

4.4 Haplotypes of yield enhancing genes

The ideal way to increase yield within the existing agricultural land is genetic improvement of yield related traits in rice. The major traits like grain length, grain width, number of well filled grains per panicle, panicle number per plant, weight of 1000 grains etc. are directly associated with rice grain productivity. Some of the genes, *OsSTAP1/OsLG3*, *OsMFT1*, *OsABCG3* and their haplotypes that strongly depend on the yield potential of rice landraces are illustrated in Table 3D, Fig.3D.

Discussion

Polymorphism of microsatellite markers

Allelic variation among SSR markers was high enough to categorize rice landraces [23]. The number of alleles per locus (6.416), observed in the current study, was high compared with some earlier reports, which reported 2.274 [19], 5.79 [20]. And this result was consistent with other studies which showed 6.13 alleles per locus [21]. The current result showed greater diversity than the results scrutinized in previous studies showing 3-9 alleles with an average of 4.53 alleles per locus [22]. The variability in the number of alleles per locus might be allied to the diversity in germplasm and specificity of the markers.

PIC is the value of a marker for distinguishing polymorphism within a population based on the number of detectable alleles and major allelic frequency. The higher the PIC value of a locus, the higher the number of alleles detected which indicates higher diversity of germplasm and thus PIC values unveil allelic frequency and the diversity among rice germplasm could be assorted from one to another SSR locus [23]. The mean PIC value (0.658) explored from the current study is higher than the reported PIC value in a previous work [24]. Markers with PIC value greater than 0.5 are considered highly polymorphic [24]. Thus the highly polymorphic markers explored in the current study are significantly informative for genetic studies and revealing of more alleles at a specific locus and the association with the alleles could be utilized for marker assisted selection programs.

Gene diversity often designated to as expected heterozygosity and higher the heterozygosity values, the broader the genetic diversity. The average 'I' value (1.23) was higher than the value described earlier [25] which was found to be 0.58. Higher value (>0.5) of genetic diversity indices like Shannon's information index and expected heterozygosity among the markers were indicates the heterozygous nature of the population [25]. The highest levels of actual and effective number of alleles showed in such locus also contribute to their great levels of expected heterozygosity (here, 0.64).

Genetic relatedness using cluster analysis and Principal co ordinate analysis

Cluster analysis based on microsatellite allelic diversity clearly demarcated rice landraces into different groups based on genetic distance and dissimilarity matrix. The clustering patterns in the current study

illustrated that the selected markers were consistent as they can detect advanced levels of variations among closely related germplasm. These highlight the precision of SSR markers in tracking the phylogeny of rice landraces. The cumulative variation found out from PCoA analysis was higher than the earlier report which stated 13.33% of variance using 36 SSR markers [26]. The results designated that the genotypes placed far away from the centroid were more genetically diverse while the genotypes placed near around the centroid hold more or less similar genetic backgrounds [22]. Thus the results indicated that the rice genotypes used for the current study were highly diverse with relatively high values of the effective number of alleles, PIC, and Shannon information index. Moreover, utilizing these informative SSR markers for future molecular characterization and population genetic studies in rice landraces are advisable.

Haplotype block construction

Several studies had explored the use of haplotypes in different GS models, and the major advantage of using haplotypes in GS is the capability to detect and include mutations as genomic information. In a rice study, among 177 rice accessions, 3211 haplotype blocks were found out from 3259 SNP combinations in the 12 rice chromosomes [6] whereas 270 haplotype blocks and 775 haplotypes were identified from the chromosomes of 93 landraces used in the current study. Eighteen consensus haplotype blocks were identified in Koshihikari, Hitomebore, Akitakomachi, and Hinohikari, which are the major cultivars of Japan [27]. Besides, these haplotype blocks identified could be allocated to genomic regions of the ancestral landraces and is crucial for identifying characteristic of recent Japanese rice cultivars. Still, they did not develop functional annotations for these haplotype blocks. The association between particular haplotypes and phenotypic differences is necessary. So in the current study, we had revealed functional genes associated with the detected haplotypes.

Haplotypes of abiotic stress tolerance genes

Rice *RZFP* gene, ***OsRZFP34*** (OS01G0719100), had three haplotypes grouped in the 20th hap block present in chromosome 1. Heat stress due to elevated temperature is one of the major environmental limitations of plant growth which can denature cellular proteins and reduce photosynthetic electron transport. Plants have evolved molecular responses like synthesis of the heat shock proteins associated with thermo tolerance to acclimatize to stressful high temperatures. After exposure to a higher temperature, plant leaves can physiologically respond by regulating stomata opening, which results in heat dissipation from the leaves. Higher transpiration cooling lowers leaf temperature. *OsRZFP34* has a physiological function in rice leaves with exposure to elevated temperature and drought for mediating stomatal aperture in leaf cooling [28].

NADP Malic-enzyme (***NADP-ME2***) [29], major role in adapting to carbonate and salt stress, had three haplotypes clustered in the 20th hap-block of chromosome 1. ***OsDHSRP1***, involved in regulating abiotic stress response, had three haplotypes clustered in the first hap-block in chromosome 2. Ubiquitination is one of the mechanisms that plants have developed to survive abiotic stresses. RING finger E3 ligase,

Oryza sativa Drought, Heat and Salt-induced RING finger protein 1 (OsDHSRP1), gene transcripts were extremely expressed under various abiotic stresses such as NaCl, drought, and heat and the phytohormones Absciscic acid (ABA) and regulate stress response through the Ub/26S proteasome system [30].

OsSIRH2-14, plays an important role in salinity tolerance through ubiquitin/26S proteasome-mediated degradation of salt related proteins [31], had three haplotypes clustered in the 22nd hap-block in chromosome 4. **OsGRAS23**, entailed in drought stress response [32], had three haplotypes grouped in the 23rd hap-block in chromosome 4. **CK2alpha2**, which regulates response to phosphate level [33], had 5 haplotypes in the 1st hap-block in chromosome 7. **OsSWI3C**, plays a major in drought tolerance [34], had three haplotypes grouped in 3rd hap-block in chromosome 11. **OsAIR2**, associated with arsenic stress tolerance [35], had four haplotypes grouped in 17th hap-block in chromosome 11.

Haplotypes of biotic stress tolerance genes

OsPSKR1, acts as an immunity regulating signal against the pathogen *Xanthomonas oryzae oryzicola* (*Xoc*), had three haplotypes grouped in hap-block 21 of chromosome 2. Bacterial leaf streak caused by the gram negative bacteria *Xanthomonas oryzae* pv. *oryzicola* is an important disease affecting rice production in lowlands of Asia and *OsPSKR1* function as a phytosulfokine receptor and regulates resistance to *Xoc* [36]. **OsLOX10**, entailed in blast resistance [37], had four haplotypes clustered in 17th hap-block in chromosome 11. **Pi-ta**, involved in blast resistance [38], had four haplotypes grouped in 6th hap-block in chromosome 12.

Haplotypes of genes related to rice growth and development

RAD51C, required for double strand break repair in rice [39], had five haplotypes clustered in 15th hap block present in chromosome 1. **OsPTAC2**, regulates rice chloroplast development [40], had two haplotypes grouped in 20th hap-block in chromosome 3. **HAZ1**, entailed in gibberelin response [41], had four haplotypes clustered in the 4th hap-block in chromosome 6. **TIG1**, promotes cell elongation [42], had four haplotypes grouped in 24th hap-block in chromosome 8. **WEG1**, entailed in root elongation [43], had four haplotypes clustered in 16th hap-block in chromosome 9.

Haplotypes of yield enhancing genes

OsLG3, plays an important role in grain yield and length [44], had three haplotypes clustered in the 3rd hap-block in chromosome 3. **OsTIF1**, involved in caryopses development [45], had three haplotypes grouped in the 2nd hap-block in chromosome 4. **OsMFT1**, regulates rice heading and panicle architecture [46], had three haplotypes grouped in the 12th hap-block in chromosome 6. **OsABCG3**, which is an ATP binding cassette transporter, vital for pollen development [47], had three haplotypes grouped in 25th hap-block in the chromosome 1. **OsPHS8**, [48] entailed in seed development and germination in rice, had three haplotypes grouped in the 29th hap-block in chromosome 8.

Many regions with low haplotype diversity were not allied with QTLs or genes. This finding may be due to undiscovered genes affecting important traits in the selected landrace germplasm. As haplotype is a set of nearby SNPs with strong linkage disequilibrium between them, the development of GWAS and competence of QTL identification based on haplotype blocks are more efficient than individual markers. Moreover, the use of haplotypes reduces the degree of freedom in the models of genomic association which increase the accuracy in QTL identification.

Global climate change and increasing population draw attention to the demand for novel allelic variants of rice stress tolerance genes. For increasing rice productivity, new yield-enhancing genes or superior alleles should be identified, and gene specific markers were developed quickly. Overall, this study revealed genetic diversity of rice landraces using different criterion such as allele number (N_a), polymorphism information content (PIC), and genetic diversity indices like Shannon information index (I) and expected heterozygosity (H_e), which endows with adequate facts in the identification of rice genotypes. As far as we know, this is the first study in exploring the haplotypes in rice landraces from Kerala and these results will therefore pave the way for evaluating genetic diversity and genomic composition in the unexplored rice landraces collected from Kerala and marker development and genome wide association studies for next-generation breeding in rice.

Declarations

Funding

Supported by Department of Science and Technology – Science and Engineering Research Board, Government of India (Grant No. ECR/2016/001934).

Conflict of interest/Competing interests

The authors declare that they have no conflict of interest.

Ethical Approval

The authors declare that the study compliance with Ethical standards.

Consent to participate

Not applicable

Consent for publication

All authors agreed to submit the manuscript

Availability of data and material

Plant material can be shared for scientific purposes.

Code availability

Not applicable

Authors' contribution

Conceptualization: Manickavelu Alagu, Smitha Kunhiraman Vasumathy; Methodology: Smitha Kunhiraman Vasumathy; Formal analysis and investigation: Smitha Kunhiraman Vasumathy; Writing - original draft preparation: Smitha Kunhiraman Vasumathy; Writing - review and editing: Manickavelu Alagu; Funding acquisition: Manickavelu Alagu; Supervision: Manickavelu Alagu

References

1. Ray DK, Mueller ND, West PC, Foley JA (2013) Yield Trends Are Insufficient to Double Global Crop Production by 2050. *PLoS One* 8:e66428. <https://doi.org/10.1371/journal.pone.0066428>
2. Tester M, Langridge P (2010) Breeding technologies to increase crop production in a changing world. *Science* 327:818–822
3. Casañas F, Simó J, Casals J, Prohens J (2017) Toward an Evolved Concept of Landrace. *Front Plant Sci* 08:145. <https://doi.org/10.3389/fpls.2017.00145>
4. Cui D, Lu H, Tang C et al (2020) Genomic analyses reveal selection footprints in rice landraces grown under on-farm conservation conditions during a short-term period of domestication. *Evol Appl* 13:290–302. <https://doi.org/10.1111/eva.12866>
5. Andersen JR, Lübberstedt T (2003) Functional markers in plants. *Trends Plant Sci* 8:554–560
6. Yonemaru J, Yamamoto T, Ebana K et al (2012) Genome-Wide Haplotype Changes Produced by Artificial Selection during Modern Rice Breeding in Japan. *PLoS One* 7:e32982. <https://doi.org/10.1371/journal.pone.0032982>
7. Nyine M, Wang S, Kiani K et al (2019) Genotype imputation in winter wheat using first-generation haplotype map SNPs improves genome-wide association mapping and genomic prediction of traits. *G3 Genes, Genomes. Genet* 9:125–133. <https://doi.org/10.1534/g3.118.200664>
8. Resende MDV, Resende MFR, Sansaloni CP et al (2012) Genomic selection for growth and wood quality in Eucalyptus: Capturing the missing heritability and accelerating breeding for complex traits in forest trees. *New Phytol* 194:116–128. <https://doi.org/10.1111/j.1469-8137.2011.04038.x>
9. Abbai R, Singh VK, Nachimuthu VV et al (2019) Haplotype analysis of key genes governing grain yield and quality traits across 3K RG panel reveals scope for the development of tailor-made rice with enhanced genetic gains. *Plant Biotechnol J* 17:1612–1622. <https://doi.org/10.1111/pbi.13087>
10. <https://doi.org/10.4038/tar.v30i1.8278>
11. Snyder MW, Adey A, Kitzman JO, Shendure J (2015) Haplotype-resolved genome sequencing: Experimental methods and applications. *Nat Rev Genet* 16:344–358

12. Doyle J, Focus JD- (1990) U (1990) Isolation of plant DNA from fresh tissue. *researchgate.net* 2:13–15
13. Liu K, Muse SV (2005) PowerMaker: An integrated analysis environment for genetic marker analysis. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/bti282>
14. Perrier X, Jaccquemoud, Collet J (2006) DARwin - Dissimilarity Analysis and Representation for Windows
15. Peakall R, Smouse PE (2012) GenALEX 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28:2537–2539. <https://doi.org/10.1093/bioinformatics/bts460>
16. Vasumathy SK, Peringottillam M, Sundaram KT et al (2020) Genome- wide structural and functional variant discovery of rice landraces using genotyping by sequencing. *Mol Biol Rep* 47:7391–7402. <https://doi.org/10.1007/s11033-020-05794-9>
17. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265. <https://doi.org/10.1093/bioinformatics/bth457>
18. Gabriel SB, Schaffner SF, Nguyen H et al (2002) The structure of haplotype blocks in the human genome. *Science* 296:2225–2229. <https://doi.org/10.1126/science.1069424>
19. Ahmad F, Hanafi MM, Hakim MA et al (2015) Genetic Divergence and Heritability of 42 Coloured Upland Rice Genotypes (*Oryza sativa*) as Revealed by Microsatellites Marker and Agro-Morphological Traits. *PLoS One* 10:e0138246. <https://doi.org/10.1371/journal.pone.0138246>
20. Shakil S, Shahanaz S, Hasan MM, Prodhan SH (2015) SSR marker based genetic diversity analysis of modern rice varieties and coastal landraces in Bangladesh. *Indian J Biotechnol* 14:33–41
21. Kumar S, Bisht IS, K.V.Bhat (2010) Assessment of genetic diversity among rice (*Oryza sativa* L.) landrace populations under traditional production using microsatellite (SSR) markers-Indian Journals. *Electron J Plant Breed* 1
22. Hossain MZ, Rasul MG, Ali MS et al (2007) Molecular characterization and genetic diversity in fine grain and aromatic landraces of rice (*Oryza sativa* L.) using microsatellite markers. *Bangladesh J Plant Breed Genet* 20:01–10. <https://doi.org/10.3329/bjpb.v20i2.17028>
23. Rashmi D, Bisen P, Saha S et al (2017) Genetic Diversity Analysis in Rice (*Oryza sativa* L.) Accessions using SSR Markers. *Int J Agric Environ Biotechnol* 10:457. <https://doi.org/10.5958/2230-732x.2017.00057.2>
24. Getachew M, Shilai Z, Teklehaymanot H (2018) Comparative Evaluation of Rice SSR Markers on Different *Oryza* Species. *J Rice Res Dev* 1:. <https://doi.org/10.36959/973/418>
25. Aljumaili SJ, Rafii MY, Latif MA et al (2018) Genetic Diversity of Aromatic Rice Germplasm Revealed by SSR Markers. *Biomed Res Int* 2018:. <https://doi.org/10.1155/2018/7658032>
26. Singh N, Choudhury DR, Singh AK et al (2013) Comparison of SSR and SNP Markers in Estimation of Genetic Diversity and Population Structure of Indian Rice Varieties. *PLoS One* 8:e84136. <https://doi.org/10.1371/journal.pone.0084136>

27. Yamamoto T, Nagasaki H, Yonemaru J-I et al (2010) Fine definition of the pedigree haplotypes of closely related rice cultivars by means of genome-wide discovery of single-nucleotide polymorphisms
28. Hsu KH, Liu CC, Wu SJ et al (2014) Expression of a gene encoding a rice RING zinc-finger protein, *OsRZFP34*, enhances stomata opening. *Plant Mol Biol* 86:125–137. <https://doi.org/10.1007/s11103-014-0217-6>
29. Liu S, Cheng Y, Zhang X et al (2007) Expression of an NADP-malic enzyme gene in rice (*Oryza sativa* L.) is induced by environmental stresses; over-expression of the gene in Arabidopsis confers salt and osmotic stress tolerance. *Plant Mol Biol* 64:49–58. <https://doi.org/10.1007/s11103-007-9133-3>
30. Kim JH, Lim SD, Jang CS (2020) *Oryza sativa* drought-, heat-, and salt-induced RING finger protein 1 (OsDHSRP1) negatively regulates abiotic stress-responsive gene expression. *Plant Mol Biol* 103:235–252. <https://doi.org/10.1007/s11103-020-00989-x>
31. Park YC, Lim SD, Moon JC, Jang CS (2019) A rice really interesting new gene H2-type E3 ligase, OsSIRH2-14, enhances salinity tolerance via ubiquitin/26S proteasome-mediated degradation of salt-related proteins. *Plant Cell Environ* 42:3061–3076. <https://doi.org/10.1111/pce.13619>
32. Xu K, Chen S, Li T et al (2015) *OsGRAS23*, a rice GRAS transcription factor gene, is involved in drought stress response through regulating expression of stress-responsive genes. *BMC Plant Biol* 15:141. <https://doi.org/10.1186/s12870-015-0532-3>
33. Wang F, Deng M, Chen J et al (2020) CASEIN KINASE2-Dependent Phosphorylation of PHOSPHATE2 Fine-tunes Phosphate Homeostasis in Rice1. *Plant Physiol* 183:250–262. <https://doi.org/10.1104/pp.20.00078>
34. Yang J, Chang Y, Qin Y et al (2020) A lamin-like protein OsNMCP1 regulates drought resistance and root growth through chromatin accessibility modulation by interacting with a chromatin remodeller OsSWI3C in rice. *New Phytol* 227:65–83. <https://doi.org/10.1111/nph.16518>
35. Hwang SG, Chapagain S, Han AR et al (2017) Molecular characterization of rice arsenic-induced RING finger E3 ligase 2 (OsAIR2) and its heterogeneous overexpression in Arabidopsis thaliana. *Physiol Plant* 161:372–384. <https://doi.org/10.1111/ppl.12607>
36. Yang W, Zhang B, Qi G et al (2019) Identification of the phytosulfokine receptor 1 (OsPSKR1) confers resistance to bacterial leaf streak in rice. *Planta* 250:1603–1612. <https://doi.org/10.1007/s00425-019-03238-8>
37. Marla SS, Singh VK (2012) LOX genes in blast fungus (*Magnaporthe grisea*) resistance in rice. *Funct Integr Genomics* 12:265–275. <https://doi.org/10.1007/s10142-012-0268-1>
38. Lee S, Jia Y, Jia M et al (2011) Molecular Evolution of the Rice Blast Resistance Gene Pi-ta in Invasive Weedy Rice in the USA. *PLoS One* 6:e26260. <https://doi.org/10.1371/journal.pone.0026260>
39. Fu R, Wang C, Shen H et al (2020) Rice OsBRCA2 Is Required for DNA Double-Strand Break Repair in Meiotic Cells. *Front Plant Sci* 11:. <https://doi.org/10.3389/fpls.2020.600820>
40. Wang D, Liu H, Zhai G et al (2016) OspTAC2 encodes a pentatricopeptide repeat protein and regulates rice chloroplast development. *J Genet Genomics* 43:601–608.

<https://doi.org/10.1016/j.jgg.2016.09.002>

41. Wen BQ, Xing MQ, Zhang H et al (2011) Rice homeobox transcription factor HOX1a positively regulates gibberellin responses by directly suppressing EL1. *J Integr Plant Biol* 53:869–878. <https://doi.org/10.1111/j.1744-7909.2011.01075.x>
42. Zhang W, Tan L, Sun H et al (2019) Natural Variations at TIG1 Encoding a TCP Transcription Factor Contribute to Plant Architecture Domestication in Rice. *Mol Plant* 12:1075–1089. <https://doi.org/10.1016/j.molp.2019.04.005>
43. Lucob-Agustin N, Kawai T, Takahashi-Nosaka M et al (2020) WEG1, which encodes a cell wall hydroxyproline-rich glycoprotein, is essential for parental root elongation controlling lateral root formation in rice. *Physiol Plant* 169:214–227. <https://doi.org/10.1111/ppl.13063>
44. Yu J, Xiong H, Zhu X et al (2017) OsLG3 contributing to rice grain length and yield was mined by Ho-LAMap. *BMC Biol* 15:28. <https://doi.org/10.1186/s12915-017-0365-7>
45. Chang S, Chen Y, Jia S et al (2020) Auxin apical dominance governed by the OsAsp1-OsTIF1 complex determines distinctive rice caryopses development on different branches. *PLOS Genet* 16:e1009157. <https://doi.org/10.1371/journal.pgen.1009157>
46. Song S, Wang G, Hu Y et al (2018) OsMFT1 increases spikelets per panicle and delays heading date in rice by suppressing Ehd1, FZP and SEPALLATA-like genes. *J Exp Bot* 69:4283–4293. <https://doi.org/10.1093/jxb/ery232>
47. Chang Z, Jin M, Yan W et al (2018) The ATP-binding cassette (ABC) transporter OsABCG3 is essential for pollen development in rice. *Rice* 11:58. <https://doi.org/10.1186/s12284-018-0248-8>
48. Li QF, Zhang GY, Dong ZW et al (2009) Characterization of expression of the OsPUL gene encoding a pullulanase-type debranching enzyme during seed development and germination in rice. *Plant Physiol Biochem* 47:351–358. <https://doi.org/10.1016/j.plaphy.2009.02.001>

Tables

Due to technical limitations, table 1-3 is only available as a download in the Supplemental Files section.

Figures

1A

1B

Percentage of variation explained by the first 3 Co-ordinates			
Co-ordinate	1	2	3
% of variation	26.56	14.79	12.17

Figure 1

A) Neighbor joining (NJ) tree: NJ tree clustering constructed based on SSR markers in Darwin v6.0.018.

B) Principal coordinate analysis: PCoA of rice germplasm using GenAlex 6.5.Coord.1 and Coord.2 represent first and second coordinates, respectively

Figure 2

Genes and their haplotypes related to biotic and abiotic stress tolerance, grain yield and growth and development in rice landraces. Numbers in blue within the bracket specifies the number of haplotypes for that particular gene

3A

OsHTAS

OsGMST1

NADP-ME2 (OsNADP-ME2)

OsPR1a

3B

OsSL

OsLG3

OsPRR1

OsABCG3

3C

OsLOX10

OsGER1

Pita

3D

WEG1

OsTOP1

RAD51C

Figure 3

Haplotypes of genes. A) Genes related to abiotic stress tolerance B) Genes related to yield C) Genes related to biotic stress tolerance D) Genes related to rice growth and development

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementarytableS1.xlsx](#)

- [SupplementaryTableS2.xlsx](#)
- [SupplementaryTableS3.xlsx](#)
- [Table1SSRmarkeranalysis.pdf](#)
- [Table2Summaryofhaplotypeblocks.pdf](#)
- [Table3Detailsofhaplotypes.pdf](#)