International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 11 | Nov - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Genetic Analysis of Regions Linked to Root System Structure in Rice Utilizing MUTMAP QTL-SEQ

Pritam Kanti Guha[1], Gopalakrishna Kadambari[2], Merugu Krishna Vamshi[1], Degati Vijayalakshmi[3]

[1]PhD Scholar, Yogi Vemana University, Kadapa- 516005

[2]Post Doctoral Fellow, IISER Tirupati- 517619

[3]Professor, Yogi Vemana University, Kadapa- 516005

ABSTRACT

The architecture of the root system is a crucial complex trait in rice. Due to the changing climate and deficiencies in soil nutrients, there is an urgent need to develop rice varieties that are efficient in nutrient use and possess strong root system architecture (RSA) traits. To identify the genomic regions linked to essential RSA component traits such as root length and root volume, a biparental F2 mapping population was created using TI-128, an Ethyl Methane Sulphonate (EMS) mutant of a major variety BPT-5204 known for its high root length (RL) and root volume (RV), along with the wild type BPT-5204. Extreme bulks exhibiting high RL and RV as well as low RL and RV were whole genome re-sequenced alongside the parent plants. Genetic mapping through the MutMap QTL-Seq method identified two genomic regions on Chromosome 12 (3.14–3.74 Mb, 18.11–20.85 Mb) and one on Chromosome 2 (23.18–23.68 Mb) as possible positions linked to both root length (RL) and root volume (RV). In a panel of sixty-two genotypes with various root lengths and volumes, Kompetitive Allele Specific PCR (KASP) testing for SNPs with delta SNP indexes approaching 1 were related with larger RL and RV. KASP_SNPs include Chr12_S4 (C→T; Chr12:3243938) in the 3' UTR region of LOC_Os12g06670, which encodes a protein kinase domain-containing protein, and Chr2_S6 (C→T; Chr2:23181622) upstream in LOC_Os2g38350, which regulates chromosomal condensation protein.

I. INTRODUCTION

Rice (Oryza Sativa) is the major staple meal for over half of the global population. It demands a thorough understanding of its production, particularly in varied ecosystems such as those found in India and categorised according to water availability (Rasheed et al., 2020). With the ever-changing climate and freshwater shortage, enhancing fertiliser usage efficiency has been a primary emphasis of rice research. Roots are extremely plastic organs that are moulded for growth initiation, elongation, and development based on the nutritional circumstances and element levels in the soil. The root detects the below-ground soil environment and may rearrange its architecture to improve water and nutrient intake. A number of genes related to phytohormones, signalling, transporters, transcription factors, etc., as well as regulators for proper gene expression of physiological responses are involved in root architecture, which includes crown roots, lateral roots, root length, volume, and root hairs. The early establishment and growth of plants are influenced by improved root system architecture (RSA), indicating that these root-related characteristics are essential for comprehending adaption through the linked gene interaction.

Causative mutations, including SNPs, candidate genes, and QTLs linked to agronomically significant rice properties, have been effectively identified using the MutMap and its modifications. To generate resource-use-efficient rice varieties, focused breeding efforts can benefit from the discovery and mapping of QTLs linked to root characteristics (Sagare et al., 2020, Sandhu and Kumar, 2017). Furthermore, NGS-based mapping techniques have been effectively used to identify genomic positions and genes linked to physiological features, disease resistance, salt tolerance, and other attributes in a variety of crops, including wheat, maize, pigeon pea, chickpea, mustard, and others. NGS-based QTL-seq techniques have been used to rice in order to quickly identify candidate genes and genomic positionss for various characteristics, such as dwarfism (Oh et al., 2020) and light green leaves (Abe et al., 2012).

I. DNA extraction and root length and volume phenotyping

Root length (RL) and root volume (RV) were measured by phenotyping the F2 population in the field. A measuring scale was used to carefully dig the soil around the roots down to the specified depth of 75 cm using a spade, an iron instrument. For the purpose of measuring RL and RV, the roots were gently cleaned under low water pressure. During the dry season of 2020–2021, 134 F2 individuals were phenotyped for RV using the water displacement method (ml) and RL using a metric scale (cm) during the panicle initiation stage. The DNA bulks were made for extreme phenotypes, such as low RL and RV (PWT) that exhibited phenotypes similar to BPT-5204 (n = 15) and high RL and RV (PMT) that showed phenotypes similar to mutant TI-128 (n = 15). Each sample's DNA was separated as previously mentioned and put onto an agarose gel for quality control (Additional File 3). Nanodrop (Thermo Fisher Scientific, USA) was used to quantify and quality-check the DNA. A sample known as PMT was created by pooling the DNA of 15 severe bulk PMT people at an equimolar quantity. Likewise, a sample known as PWT was created by pooling the DNA of 15 extreme bulk PWT people at an equimolar quantity. Figure 1 shows the MutMap-QTL-Seq flowchart for RV and RL.

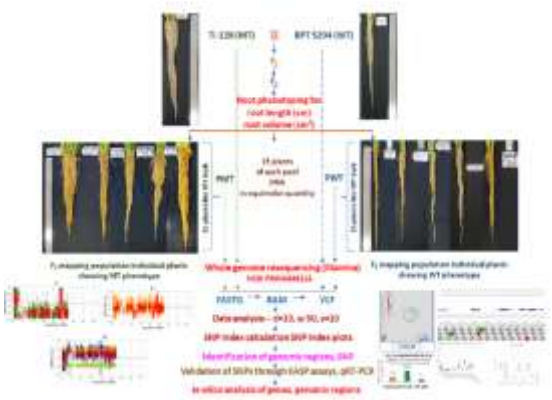

Figure 1 Flowchart for the MutMap QTL-Seq data analysis for elucidating the genomic regions associated with high RL and RV (P1- parent 1 i.e. TI-128, P2- parent 2 i. e. BPT-5204, Bulk 1 (pooled wild type bulk_PWT), Bulk 2 (pooled mutant type bulk_PMT).

II. Whole genome resequencing coupled to Data Analysis

Four samples—PMT, PWT, TI-128, and BPT-5204—were subjected to whole genome resequencing using the Illumina HiSeq 2500 platform (Illumina, United States) with 20X genome coverage, producing 10 GB of data per sample. The MutMap QTL seq pipeline was used to the data (Sugihara et al., 2022). The Oryza sativa Cv. japonica Nipponbare sequence (MSU release IRGSP 7.0) used as the reference genome for sequence alignment in this investigation. BPT-5204 produced 78.15 million paired-end short reads, TI-128 produced 97.01 million, PWT produced 92.31, and PMT produced 72.70 million.

The readings were subjected to quality tests using Trimmomatic v0.39 and FastP v0.20.1. Potential adaptor contamination was eliminated, and low-quality readings were filtered out. To guarantee a mean base Phred quality score over 30 in 10-bp sliding windows, iterative pruning was carried out. Orphan single-end reads and reads less than 40 bp after trimming were eliminated. The NCBI Sequence Read Archive (SRA) now contains the data's cleaned reads. Samtools v1.9 was then used to align the high-quality short reads to the Oryza sativa Cv. japonica Nipponbare reference sequence. Using bcftools v1.9, alignment and variant files were acquired. To find trustworthy causal SNPs, the MutMap QTL-seq pipeline employed variant call format (vcf) files.

The bulked DNA from the mutant F2 progeny's paired-end sequencing readings were compared to those of Oryza sativa Cv. japonica. SNPs and the Nipponbare reference sequence were classified as either heterozygous (SNP index ≥0.3 and <0.6) or homozygous (SNP index ≥0.9). After identifying genomic positionss with SNP clusters and an SNP index close to 1, all of the SNPs in the region were regarded as potential candidates for the causative mutation for additional study. By averaging SNP indices from a moveable window of 10 consecutive SNPs and adjusting the window size to 50 SNPs at a time and depth of 10X, SNP index plot regression lines were produced. The midway point between the first and fifth SNPs was chosen as the x-axis value for each averaged SNP index. Each averaged SNP index's x-axis value was set at the halfway point between the window's first and fifth SNPs. Plotting of the genome's delta SNP-index was done using the confidence intervals for each of the 12 rice chromosomes.

III. Kompetitive Allele Specific PCR (KASP) assay designing and panel validation

The delta SNP index was extensively analyzed at the p95 and p99 levels across the chromosomes. SNPs chosen for validation via KASP, based on their allelic variations and functional significance, were classified as candidate

SNPs, with a delta SNP index close to 1 found within genes associated with abiotic stress, including traits related to nutrient use efficiency or root development. The SNP indices, their positions, and functional annotations were meticulously reviewed.

The relevant genes with SNP indices close to ~1 were chosen. Kompetitive Allele Specific PCR (KASP) assays were created for particular alleles by extracting sequences that are 100 bp upstream and downstream, utilizing rice databases like RAP-DB (https://rapdb.dna.affrc.go.jp/) and RGAP (http://rice.uga.edu/). LGC (Teddington, UK; www.lgcgenomics.com) designed six KASP assays. These assays were confirmed using qRT-PCR (Applied Biosystems Veriti Thermal Cycler in 384-well format, USA) across a panel of 62 genotypes, which included 15 individuals from each extreme bulk (15 PWT and 15 PMT), as well as the parent rice mega-varieties, landraces, and BPT-5204 mutant lines exhibiting different root lengths and volumes. The PCR reaction comprised five µL, which included 2.5 µL of template DNA (10 ng), 2.3 µL of 2 × KASP master mix, and 0.14 µL of the primer mix. This reaction was carried out on a Step One Plus machine following a defined protocol: a pre-read phase at 30 ◦C for 1 min, a hot start at 94 ◦C for 15 min, followed by 10 touchdown cycles (94 ◦C for 20 s; touchdown starting at 61 ◦C, decrementing by 1 ◦C each cycle, for 60 s), and then 26 amplification cycles (94 ◦C for 20 s; 55 ◦C for 60 s). Fluorescence data were detected during both the pre-read and post-read phases at 30 ◦C for 1 min, utilizing dyes such as FAMAbs (485 nm), HEXAbs (535 nm), and ROXAbs (575 nm).

IV. Phenotyping for F2 mapping population and preparation of bulk for WGRS Results

A genetically advantageous line for root length (RL) and root volume (RV), known as TI-128, was crossed with BPT-5204 to create the mapping population. The F1 seeds underwent heterozygosity checks utilizing highly variable SSR markers, and verified true F1 lines were progressed to the F2 mapping population for genetic analysis. In the F2 cohort, the RL varied between 20 and 43 cm, while the RV ranged from 20 to 100 ml of water displaced. Individuals with root characteristics akin to BPT-5204 (RL - 25.90 cm and RV - 43.33 ml) were classified as wild type, whereas those resembling TI-128 (RL - 39.03 cm and RV - 96.67 ml) were categorized as mutant type. In the F2 generation, there were 20 wild type plants and 114 mutant type plants for RL, along with 29 wild type and 105 mutant type for RV. Dominant inhibitory epistasis had a notable impact, showing a complete dominance effect at both interacting gene pairs; when one gene is dominant, it obscures the effects of the other gene. A dominant allele at one locus inhibits the expression of both the dominant and recessive alleles at the second locus. Based on their phenotypes, 15 individuals (n = 15) were chosen from each bulk to form two groups: the high RL and RV bulk designated PMT (RL 32–43 cm and RV 70–100 ml) and the low RL and RV bulk labeled PWT (RL 20–30 cm and RV 20-< 60 ml).

V. Whole genome Re-sequencing and alignment of reads Results

The statistics for whole genome resequencing at 20X depth for PMT, PWT, TI-128, and BPT-5204 have been recorded, including total read count, clean reads, aligned read count, and alignment percentage. We generated 13.847 GB of cleaned bases for PWT, 10.906 GB for PMT, 11.724 GB for BPT-5204, and 14.552 GB for TI-128. In total, 59.25 million reads were aligned, resulting in the identification of 5339 SNPs compared to the reference genome.

VI. Genomic regions, candidate genes and SNPs associated with root length (RL) and root volume (RV)

The comparison of SNPs from the mutant bulk (PMT) with those from the wild-type (WT) led to the identification of causal SNPs related to RL and RV. The distribution of these distinct SNPs across each chromosome was analyzed to gain insights into their genomic locations, functional annotations, and potential implications. In the genomic regions on Chr.12 (3.14 Mb–3.74 Mb, 18.11 Mb–20.85 Mb) and Chr.2 (23.18 Mb–23.68 Mb), regions exhibiting a peak in SNP index

close to 1 were identified (Figure 2).

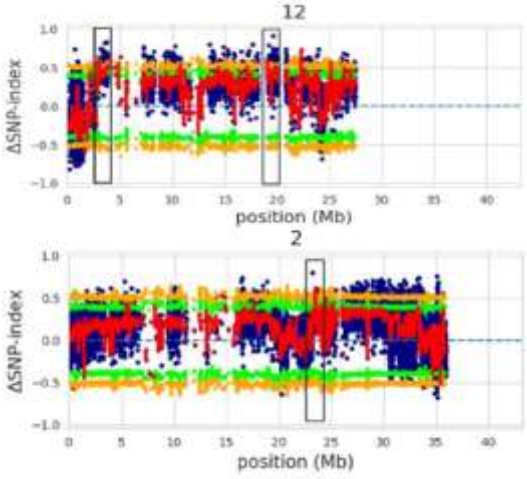

Figure 2: A genomic region on Chr.12 and Chr.2 x-axis indicates the physical position on the chromosome, and the y-axis indicates the average SNP index with a 50 kb sliding window *p99 and *p95 confidence interval of the delta SNP-index. Delta SNP index (blue dots) between the high RL & RV (PMT) and low RL & RV (PWT) for causal SNPs with a window size of 50 kb, minimum depth of variants-10, site size- 10 Kb. The X and Y axes depict the chromosomal position (Mb) and delta SNP index respectively. The red line depicts mean SNP-index, orange line depicts 99% confidence interval of simulated delta SNP-index, green line depicts 95% confidence interval of simulated delta SNP-index.

The SNPs point to genomic regions that show significa nt differences between the two opposing phenotypic gro ups, indicating potential connections to the characteristi c being studied (RL and RV).
29 functionally significant SNPs were found, all of whi ch had indices near 1.

This study shed light on the genetic differences found in the mutant lines and how they relate to particular chromosomal genomic positionss.17 of the 29 functionally significant SNPs were discovered in intergenic positions, which denotes that they are situated outside of regions known to code for proteins. Within intron regions, five SNPs were found. Seven SNPs were also synonymous, which means that the encoded protein does not vary in amino acids as a result of them. Most significantly, two SNPs were labelled missense, meaning that the final protein had different amino acids. Six of the candidate SNPs were chosen for KASP validation because their delta SNP index in the PMT was higher than 0.7. The detected genomic regions were underpinned by five of these six SNPs (Chr12_S1, Chr12_S2, Chr12_S3, Chr12_S4, and Chr2_S6), while Chr.12_S5 (SNP index ~1) was located extremely near to the Chr.12 (18.11–20.85 Mb) genomic region. The gene LOC_Os12g32420, which codes for an undetermined expressed protein, contains the SNP (Chr12_S5). In the 3′ UTR region of LOC_Os12g06670, another SNP (Chr12_S4) was found. This region encodes ACT domain-like protein kinases 6 (ACTPK6), which are known to aid in root growth in rice seedlings when ammonium levels are sufficient. The gene LOC_Os2g38350, which codes for a regulator of the chromosomal condensation protein RCC1, had the SNP (Chr2_S6) upstream. This protein is essential for stress signalling and has been strongly linked to strong root system architecture, highlighting its vital function in stress adaptation. All things considered, the SNPs found in these candidate genes offer important new information on possible genetic variants linked to root development, growth, and stress tolerance. Clarifying the precise functions and contributions of these genes and their corresponding SNPs to the observed phenotypic differences would be made easier with additional functional characterization.

VII. Validation of causal SNP for root length and root volume by Kompetitive allele specific PCR (KASP)

Six SNPs were chosen, Chr12_S1, Chr12_S2, Chr12_S3, Chr12_S4, Chr.12_S5, and Chr2_S6, and their associations with RL and RV were validated using the KASP genotyping assay. Aerobic adapted lines, landraces with high root length and root volume like the mutant (TI-128), flooding adapted lines with low root length and root volume like the wild type (BPT-5204), and a total of 62 genotypes with 15 individual plants each from PWT and PMT bulks. Three of the six KASPs—Chr12_S4, Chr12_S5, and Chr2_S6—showed polymorphism and were discovered to be closely linked to the root length and root volume characteristics (Figure 3). A protein kinase domain-containing protein is encoded by the KASP Chr12_S4 (Chr12:3243938), which is found in the 3′ UTR region of the gene LOC_Os12g06670. It is well known that this gene is essential for rice seedling root development, particularly when there is enough ammonium present. In the positions of the gene LOC_Os12g32420, which codes for a putatively uncharacterized expressed protein, another KASP, Chr12_S5, was discovered. Additionally, the regulator of chromosomal condensation protein (RCC1) is encoded by the KASP Chr2_S6

(Chr2:23181622), which is located upstream of the gene LOC_Os2g38350. Based on genotyping, the validation of these KASPs, in particular Chr12_S4, Chr12_S5, and Chr2_S6, demonstrated a substantial correlation with root length and root volume features. This helps to clarify the fundamental mechanics of root formation and offers important insights into the genetic foundation of high RL and RV traits, which may improve crop performance.

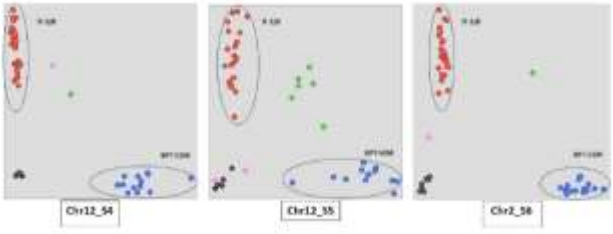

Figure 3 Validation through KASP assays developed for RL and RV on a panel of 62 rice genotypes varying in RL and RV (A) KASP Chr12_S4 (Chr12:3243938), (B) KASP Chr12_S5(Chr12: 19554787) and (C) KASP Chr2_S6(Chr2:23181622). Red: Sample is homozygous for the HEX allele; Green: Sample is heterozygous: one FAM allele and one HEX allele; Blue: Sample is homozygous for the FAM allele; Black: NTC.

VIII. DISCUSSION

Plants' root system design allows them to respond to a variety of abiotic challenges, including as drought, salt, and metal toxicity. As climate change brings new difficulties to global agriculture, knowing root features and their genetic foundation becomes critical in generating adaptable rice varieties (Hossain et al., 2021). The root system architecture is a complicated feature that has a big influence on rice crop yield. Furthermore, effective nutrition utilisation is critical for maximum development, stress tolerance, and higher output. The roots are primary indicators of abiotic stressors caused by inadequate water, fertiliser, and element delivery in plants. Robust root length and volume let plants absorb enough water and nutrients to grow and develop properly. As a result, rice lines with large root lengths and volumes may have the intrinsic potential to tolerate abiotic stress and nutritional deficiency.

We used a combination of MutMap and QTL-seq to sequence the entire genome of BPT-5204, its EMS mutant TI-128, and the two extreme phenotype bulks PMT and PWT, with the goal of mapping genomic regions, genes, and SNPs associated with major root architecture determinants RL and RV under field conditions. We used MutMap QTL-Seq to identify putative genetic positionss related with RL and RV. A genomic position on Chromosome 2 (23.18 Mb-23.68 Mb) was linked to elevated RL and RV levels. This position is 0.50 Mb in size and 0.38 Mb upstream of a known QTL, QUICK ROOTING 1 (Kitomi et al., 2018). The QUICK ROOTING 1 gene (23.56 Mb-25.27 Mb), which is related with maximal root length, was previously identified on chromosome 2 between markers RM5651 and RM6107 over a 1.7-Mb gap (Kitomi et al., 2018).

This genomic position has been connected to QTLs associated with maximal root development under control circumstances (Yue et al., 2006; Courtois et al., 2009), and it is located near MQTL-RDR2, which is associated with root dry weight, with flanking markers RM5320 and RM2737 (Khahani et al., 2021). The findings indicate that this positions of Chr.2 contains genetic factors that regulate several aspects of root formation and growth in rice. Furthermore, the mapped positions are fine (23.18-23.68 Mb) and represents a shorter segment of the genomic region associated with high RL and RV, making it appropriate for use in breeding programs (Jaganathan et al., 2020). The genomic region between 3.14 Mb and 3.74 Mb on Chr.12 is 0.44 Mb upstream of the QTL rv12-2 (3.58 Mb-8.82 Mb). A 0.16 Mb positions overlaps rv12-2, which has been linked to drought resistance and root volume (Qu et al., 2008).

This co-mapped region is much shorter than rv12-2 and is presented as a finely mapped region of 0.60 Mb with strong RL and RV. Rice's root-related properties are genetically complicated, with many QTLs in distinct genomic positions contributing to root formation, growth, and responses to environmental circumstances. The highly mapped genomic position can be pyramided in a single genetic background to improve root system architecture and nutrient usage efficiency. Furthermore, the high resolution, efficiency, increased marker density, and cost-effectiveness of the SNPs pave the path for genomics-assisted breeding to improve RL and RV.

Another position on Chr.12, spanning 18.11 Mb to 20.85 Mb, has been mapped and reported for the first time for its relationship with RL and RV in rice. The mapped positions on Chr. 12 contains several genes involved in root formation, water and nutrient intake. These genes contain transcription factors, transporters, signalling molecules, and phytohormones including auxin and cytokinin.

The region includes genes such as OsbZIP84, GRAS transcription factor, CYCLIC NUCLEOTIDE-GATED ION CHANNELs, Ethylene responsive transcription factor (ERF-110), sodium symporter family protein, C2H2 zinc factor, Phosphorus utilization AtPR5, receptor-like cytoplasmic kinase, phosphoserine phosphatase, OsFbox665, which play roles in auxin regulation and lateral root initiation, as well as OsCLT2, auxin-regulating genes, and qCTR12 that are associated with root development and stress tolerance in rice. Notably, there is a greater abundance of genes related to auxin regulation within the identified genomic regions. Previous studies have indicated that GRAS genes operate specifically under drought conditions to assist in nutrient uptake and increasing yields (Dutta et al., 2021). This genomic region represents a unique position that could be leveraged for transferring root traits in molecular breeding efforts focused on enhancing root characteristics and nutrient use efficiency, which was further validated by KASP assay.

CONCLUSION

Understanding the genetic foundation and regulatory mechanisms for root characteristics is important for breeders because it provides precise genetic markers to speed up the production of rice varieties that are suited to water-limited aerobic conditions and have enhanced nutrient utilisation efficiency. The deliberate integration of highly mapped genomic positionss into a single genetic background, with flanking SNPs serving as powerful markers, provides a technically sophisticated method to improving root system design for greater nutrient usage efficiency. SNPs' high resolution, efficiency, and cost-effectiveness make them a potential road forward for genomics-assisted breeding approaches.

REFERENCES

1. Rasheed R, Ashraf MA, Iqbal M, Hussain I, Akbar A, Farooq U, Shad MI. Major constraints for global rice production: changing climate, abiotic and biotic stresses. Rice research for quality improvement: genomics and genetic engineering: Volume 1: breeding techniques and abiotic stress tolerance. 2020:15-45.
2. Sagare, D.B., Abbai, R., Jain, A., Jayadevappa, P.K., Dixit, S., Singh, A.K., Challa, V., Alam, S., Singh, U.M., Yadav, S., 2020. More and more of less and less: is genomics-based breeding of dry direct- seeded rice (DDSR) varieties the need of hour? Plant Biotechnol. J. 18, 2173–2186.
3. Sandhu, N., Kumar, A., 2017. Bridging the rice yield gaps under drought: QTLs, genes, and their use in breeding programs. Agronomy 7, 27.
4. Abe, A., Kosugi, S., Yoshida, Kentaro, Natsume, S., Takagi, H., Kanzaki, H., Matsumura, H., Yoshida, Kakoto, Mitsuoka, C., Tamiru, M., 2012. Genome sequencing reveals agronomically important loci in rice using MutMap. Nat. Biotechnol. 30, 174–178.
5. Oh, J., Cheon, K.-S., Kang, D.-Y., Kim, S.L., Lee, E., Kim, N., Oh, H., Choi, I., Baek, J., Yoon, I.S., 2020. MutMap analysis of a rice dwarf mutant line. Korean Society of Breeding Science 52, 9–19.
6. Sugihara, Y., Young, L., Yaegashi, H., Natsume, S., Shea, D.J., Takagi, H., Booker, H., Innan, H., Terauchi, R., Abe, A., 2022. High-performance pipeline for MutMap and QTL-seq. PeerJ 10, e13170.
7. Hossain, A., Islam, M.T., Maitra, S., Majumder, D., Garai, S., Mondal, M., Ahmed, A., Roy, A., Skalicky, M., Brestic, M., 2021. Neglected and Underutilized crop Species: are they Future Smart crops in Fighting Poverty, Hunger and Malnutrition under changing climate?. In: Neglected and Underutilized Crops-Towards Nutritional Security and Sustainability, pp. 1–50.
8. Kitomi, Y., Nakao, E., Kawai, S., Kanno, N., Ando, T., Fukuoka, S., Irie, K., Uga, Y., 2018. Fine mapping of QUICK ROOTING 1 and 2, quantitative trait loci increasing root length in rice. G3: genes, Genomes. Genetics 8, 727–735.
9. Yue, B., Xue, W., Xiong, L., Yu, X., Luo, L., Cui, K., Jin, D., Xing, Y., Zhang, Q., 2006. Genetic basis of drought resistance at reproductive stage in rice: separation of drought tolerance from drought avoidance. Genetics 172, 1213–1228.
10. Khahani, B., Tavakol, E., Shariati, V., Rossini, L., 2021. Meta-QTL and ortho-MQTL analyses identified genomic regions controlling rice yield, yield-related traits and root architecture under water deficit conditions. Sci. Rep. 11, 6942.

11. Jaganathan, D., Bohra, A., Thudi, M., Varshney, R.K., 2020. Fine mapping and gene cloning in the post-NGS era: advances and prospects. Theoretical and Applied Genetics 133 (5), 1791–1810.
12. Qu, Y., Mu, P., Zhang, H., Chen, C.Y., Gao, Y., Tian, Y., Wen, F., Li, Z., 2008. Mapping QTLs of root morphological traits at different growth stages in rice. Genetica 133, 187–200.
13. Dutta, M., Saha, A., Moin, M., Kirti, P.B., 2021. Genome-wide identification, transcript profiling and bioinformatic analyses of GRAS transcription factor genes in rice. Front. Plant Sci. 12, 777285.