

Genome Wide Association Analysis for Uniform Coleoptiles Emergence and Early Seedling Growth in Rice

Bhagyarabi Pani^{1*}; G.M. Lal²; Parameswaran C³; Selvaraj Sabarinathan⁴; Tejasmita Prusty⁵; Asit Prasad Dash⁶

^{1,3}ICAR-National Rice Research Institute, Cuttack, 753006, India

²SHUATS-Naini Agricultural Institute, Prayagraj, 211007, India

⁴OUAT-Orissa University of Agriculture & Technology, BBSR, 751003, India

^{5,6}SOA-Siksha O Anusandan University, BBSR, 751003, India

Corresponding Author: Bhagyarabi Pani^{1*}

Abstract:- Early seedling vigour (ESV) is a complex trait in rice. Detection of QTLs/genes controlling these traits can help us in enhancing the yield potential in rice varieties. Association mapping is a technique based on the principle of linkage disequilibrium that is used to find genes or quantitative trait loci (QTL) underlying the complex traits. In this study of haplotype breeding 281 rice genotypes were taken. ANOVA analysis showed P-value for traits and genotypes was found significant. Similarly, P-value for interaction between the traits and genotypes was also found to be highly significant (1.8663*10⁻²⁰⁸). Further, mean germination data positively correlated with mean shoot length, mean leaf number, mean culm diameter, mean shoot dry weight, and mean shoot area of 21st day among the 281 genotypes. Among the 281 number of genotypes, 111 genotypes are found to be in PCA1 and 170 genotypes are found to be in PCA2 based on the phenotypic analysis. PCA1 component constituted 29.93% and PCA2 constituted the 13.68% of total variation in the analysis. Besides, whole genome phylogenetic analysis showed three major groups of which Group 1 consists of 215 genotypes, group 2 consists of 38 genotypes and group 3 consists of 28 genotypes respectively. Especially, both the subgroups II and III comprised of the unique genotypes from the *indica* and *aus* subpopulations of rice. In this analysis, 16 significant associations (LOD Score >7) for different traits were identified using the three different models (MLM, farmCPU, and blink) for GWAS studies. Especially, one major QTL was identified for the mean coleoptiles' emergence for 10 DAS on 11th chromosome (18983591) which explained 49% of the phenotypic variance. Additionally, another major QTL contributing to the shoot length variation of 29.75% was identified in the Chr02 (32954393) for shoot length trait on 28 DAS. A candidate gene namely Os02g0778400 UMP/CMP kinase A/adenylate kinase (LOC_Os02g53790) was located in the significant SNP region of the GWAS analysis. Further characterization of this gene would assist in elucidation of the mechanism regulating the early seedling length in rice under direct seeded rice.

Keywords:- ESV, GWAS, PCA, ANOVA, QTLs.

I. INTRODUCTION

Rice (*Oryza sativa* L.) is one of the most important staple food crops widely cultivated in the Asian countries and the world in an area of about ~ 162 million hectares **Gadal et al. (2019)**. As a major cereal crop, rice production is highly crucial to food security (Fairhurst and Doberman, 2002) and it provides 20 percent of dietary energy to the feeding population **Alexandratos and Jelle, (2012)**. Green revolution in Asian countries and hybrid rice adoption especially in China greatly influenced the rice productivity and sustainability in the second half of the last century **Xu et al. (2015)**. Recently rice production in the world has reached ~750 million tons (FAOSTAT, 2019). However, risk of climate change, degradation of soils, and less availability of the irrigation water pose a serious challenge to sustain the world rice production **Saud et al. (2022)**.

Direct seeded rice (DSR) rather than transplanted rice is relatively easy method of rice cultivation. In DSR, field requires proper leveling and seeds are sown directly in the soil with one pre-irrigation facilitating the germination of rice seeds **Pathak et al. (2011)**. The major advantage of DSR relative to transplanted rice is the saving of the water, time, cost, and energy in the field preparation and no requirement for the nursery **Rao et al. (2007)**. However, the basic advantage of seedling establishment and initial weed control in case of the transplanted rice is compromised in DSR due to improper field leveling and high rainfall during the germination stage of the crop. Both of these conditions severely affect the seedling establishment in DSR **Farooq et al. (2011)**. Besides, weed infestation also hinders the seedling establishment and yield **Rao et al. (2007)**.

A yield loss of 5-15% was observed in the multi-local meta-analysis between the cultivation practices of the DSR and transplanted rice **Xu et al. (2019)**. Further, seed priming, seed pelleting, and optimization of the sowing depth in the anaerobic tolerant rice varieties are found to positively affect the yield of rice in DSR **Adhikary et al. (2022)**, **Mei et al. (2017)** and **Chamara et al. (2018)**. Thus, rapid and uniform germination, early seedling vigor, weed competitiveness, and anaerobic germination tolerance are

favorable traits for enhancing the yield in the DSR **Mahender et al. (2015)**. Besides, mean germination rate, germination index, and early seedling vigour were found to be positively correlated with the rice yield in DSR **Mahajan et al. (2011)**.

The favorable traits in rice for DSR namely early germination, uniform emergence, and early seedling vigor are controlled by multiple genes in rice. The QTL hotspots for early seed germination has been identified in 11 and 12th chromosome of rice **Dimaano et al. (2020)**. Similarly, genomic region regulating the seed vigor indexes were also identified in the 11 and 12th chromosome of rice **Barik et al. (2022)**. Besides, QTL regions for the seedling vigour index have been reported in the 2nd chromosome of rice **Anandan et al. (2016)**. Additionally, hotspots for the seedling vigour for DSR conditions were identified in the 3 and 5th chromosomes of rice **Singh et al. (2017)**. Besides, GA20ox gene involved in gibberellic acid synthesis has also been identified for regulating the seedling vigor in rice **Abe et al. (2012)**. In addition, several candidate genes have been identified for regulating the speed of germination and seedling vigor in rice **Yang et al. (2021)** and **Kalluru et al. (2023)**.

➤ Justification

The performance of different Paddy varieties varies under different Argo-climatic conditions due to their specific Argo-climatic requirement. Therefore, an appraisal of varieties for their uniform germination and coleoptile emergence with respect to growth, necessary to improve the production in Direct Seeded Rice. In this context it is very much necessary to evaluate the gene regulating early uniform coleoptiles emergence and early seedling growth for identification of candidate genes

➤ Objectives

The present study has been formulated to identify the candidate genes for the speed of germination and seedling vigour traits in rice with the following objectives.

- Phenotyping diverse rice germplasm for uniform coleoptiles emergence and early seedling growth in rice.
- Association mapping of the candidate genes associated with uniform coleoptiles emergence and early seedling growth in rice.

II. MATERIALS AND METHODS

The germination and early seedling vigor of plants are fundamental processes that determine the establishment and growth of plants. The present investigation on the seedling vigor and germination rate differences in the rice genotypes was carried out during the *Rabi* with the hypothesis to dissect the genomic regions associated with the germination and early seedling vigour in direct seeded rice conditions. In this section, we will outline the materials and methods commonly used in this research to study germination and early seedling vigor in plants.

A. Description of the Experimental Site

The experiment was conducted at a shade house facility at ICAR-NRRI, Cuttack, Odisha. Cuttack is located in the eastern coastal plain region of India, with geographical coordinates of 20.4625° N latitude and 85.8830° E longitude, and 36 MSL. The average temperature during the experiment ranged from 18-32°C and the relative humidity was 53%.

B. Experimental Materials

A collection of 281 rice genotypes in the rice 3K panel were collected with a combination of twelve sub-populations were used in the study. Supplementary Table 1 presents the details of the genotypes utilized in this study, including the sub-populations and countries from which the data were obtained.

C. Phenotypic Screening and Data Observation

Soil was collected from Aerobic field. By the help of hand hammer soil was given to a uniform shape of fine powder. Uniform filling of measured soil in the seedling tray was done. Seeds were placed uniformly 1cm below the soil surface and little amount of soil was added to it. A total of 281 rice accessions were sown in black colour germination trays with a length of 49.5×31 cm and a height of 4 cm. In each pit, 10 seeds were sown and maintained in a natural environment with two replication of the randomized block design. For each replication, six data points were collected and used in the analysis. The emergence of the plumule from the soil was considered as the germination. The various stages of germination and seedling growth differences were observed and the traits viz., mean germination, shoot length on the 7th, 11th, 13th, 14th day, 21st day and 28th day, the number of leaves on the 14th day, 21st day and 28th day, culm diameter, shoot area and shoot dry weight on 14th day, and 21st day, respectively were collected for the analysis. The germination data was counted regularly from 4 days after sowing up to 9 days after sowing. Further, the shoot area was observed through the image analysis method. For each seedling, ruler was placed near and recorded photos using a Nikon D700 camera. All the images were preprocessed and analyzed manually using ImageJ software (<https://imagej.nih.gov/ij/index.html>).

D. Phenotype and GWAS Analysis

Phenotypic data were evaluated through descriptive statistics in the data analysis option available in MS EXCEL. Similarly, two replications of the genotypes were used for understanding the interaction between the genotypes and traits using the two-way ANOVA with significance at 5% and 1% level of significance. Further, distribution of the traits was understood using the histogram plot analyzed using 'hist' function in the R software (Version 3.6.0). Besides, correlation between the traits was understood using the 'corrplot' package in R following the 'hclust' method of correlation. The total variation in the traits between the genotypes was understood using the 'ggfortify' package in R. Further, for the association analysis, 3K RG 29 million SNPs were downloaded from the SNPs datasets and the genotype data for the 281 genotypes were retrieved for further analysis (https://snp-seek.irri.org/_snp.zul). The heterozygous data were considered and missing value and minimum allele

frequency was set to 10% and LD KNNI imputation was followed (Money *et al.*, 2015) for data imputation using Tassel software version. Additionally, the mixed linear model was performed through GAPIT version 3 (Wang *et al.*, 2021) for the identification of the genomic regions associated with the different traits. The significant SNPs with LOD value of more than seven threshold value ($-\log_{10}P$) was considered for candidate gene identification. The annotated genomic regions of Nipponbare in RAP-DB database was used for the identification of the candidate genes.

III. RESULTS

➤ Descriptive Statistics for 10th Day of Germination

The mean value of 10th day germination was found to be 97.82%. The lowest and highest number of germinations on 10th day was found to be 50% and 100%, respectively. The kurtosis value for mean germination on 10th day was found to be 21.45 and the skewness value was found to be -4.38. In the histogram analysis, it was found that 235 number of genotypes are found to be germinated 100% on 10th day, 19 genotypes having 90%, 13 genotypes having 85%, five genotypes having 80% germination, two genotype having 70-80% germination, 3 genotypes having 60-70% germination, and three genotypes having 50-60% germination rate (Table 1).

➤ Descriptive Statistics for 7th Day Shoot Length

The mean value of 7th day shoot length was found to be 3.54 cm. The lowest and highest number of shoot length on 7th day was found to be 0 and 7.06 cm, respectively. The kurtosis value for mean shoot length on 7th day was found to be 0.14 cm and the skewness value was found to be -0.028 cm, respectively. In the histogram analysis, it was found that maximum number of genotypes, i.e., 87 having the shoot length 3-4 cm, 77 number of genotypes having shoot length in between 4-5 cm and 71 genotypes had shoot length of 2-3 and 22 genotypes having shoot length of 0-2 cm (Fig. 5).

➤ Descriptive Statistics for 11th Day Shoot Length

The mean value of 11th day shoot length was found to be 6.05 cm. The lowest and highest number of shoot length on 11th day was found to be 2.98 cm and 8.54 cm, respectively. The kurtosis value for mean shoot length on 11th day was found to be 0.27 cm and the skewness value was found to be -0.044 cm. In the histogram analysis, it was found that 65 number of genotypes are found to be in between 5.5-6 cm of shoot length, 47 and 47 number of genotypes are found to be in between 5-5.5cm and 6-6.5cm, 39 number of genotypes having 6.5-7 cm, 40 number of genotypes having 6-8 cm, 9 number of genotypes having 8-9 cm and then 34 number of genotypes having 0-5 cm shoot length, respectively (Fig. 6).

➤ Descriptive Statistics for 13th Day Shoot Length

The mean value of 13th day shoot length was found to be 7.17 cm. The lowest and highest number of shoot length on 13th day was found to be 3.53 cm and 11.31 cm respectively. The kurtosis value for mean shoot length on 13th day was found to be 1.05 and the skewness value was found to be 0.26. In the histogram analysis, it was found that 105

number of genotypes having 6-7 cm, 83 number of genotypes having 7-8 cm, 47 number of genotypes having 8-9 cm, 11 number of genotypes having 9-10 cm, 3 number of genotypes having 10-12 cm, 30 genotypes having 4-6 cm and 2 genotypes having 0-2 cm of shoot length, respectively (Fig. 7).

➤ Descriptive Statistics for 14th Day Shoot Length

The mean value of 14th day shoot length was found to be 9.29 cm. The lowest and highest number of shoot length on 14th day was found to be 5.26 cm and 15.03 cm respectively. The kurtosis value for mean shoot length on 14th day was found to be 0.82 and the skewness value was found to be 0.64. In the histogram analysis, it was found that 72 genotypes having 8-9cm of shoot length, 68 genotypes having 9-10 cm, 46 genotypes having 7-8 cm, 45 genotypes having 10-11 cm, 10 genotypes having 6-7 cm, 3 genotypes having 0-3cm, 19 genotypes having 11-12 cm, 15 number of genotypes having 12-14 cm and 3 genotypes having 14-16 cm of shoot length, respectively (Fig. 8).

➤ Descriptive Statistics for 14th Day Leaf Number

The mean value of 14th day leaf number was found to be 2.00. The lowest and highest number of leaf on 14th day was found to be 1 and 2.83, respectively. The kurtosis value for mean leaf number on 14th day was found to be 21.01 and the skewness value was found to be -0.72. In the histogram analysis, it was found that 254 number of genotype having leaf number 2, 20 genotypes having leaves number in between 2-2.5, one genotype having leaves number 3, 5 genotypes having leaves number in between 1-1.5, one genotype having leaves number 1.8, respectively (Fig. 11).

➤ Descriptive Statistics for 14th Day Culm Diameter

The mean value of 14th day culm diameter was found to be 0.68 cm. The lowest and highest culm diameter on 14th day was found to be 0.34 cm and 0.89 cm, respectively. The kurtosis value for mean culm diameter on 14th day was found to be 0.66 and the skewness value was found to be -0.39. In the histogram analysis, it was found that 71 genotypes having culm diameter of 0.65-0.70 cm, 94 genotypes having culm diameter of 0.7-0.8, 28 genotypes having culm diameter of 0.8-0.9 cm, 43 genotypes having culm diameter of 0.6-0.65 cm, 38 genotypes having culm diameter of 0.5-0.6 cm, and 7 genotypes having culm diameter of 0.3-0.5 cm, respectively (Fig. 14).

➤ Descriptive Statistics for 14th Day Shoot Dry Weight

The mean value of 14th day shoot dry weight was found to be 0.008 gm. The lowest and highest shoot dry weight on 14th day was found to be 0.003 gm and 0.016 gm, respectively. The kurtosis value for mean shoot dry weight on 14th day was found to be 0.54 and the skewness value was found to be 0.31. In the histogram analysis, it was found that 118 genotypes having shoot dry weight of 0.008-0.010 gm, 129 genotypes having shoot dry weight of 0.006-0.008 gm, 24 genotypes having shoot dry weight of 0.004-0.006 gm, 3 genotypes having shoot dry weight of 0-0.004 gm, 42 genotypes having shoot dry weight of 0.010-0.012 gm, 13 genotypes having shoot dry weight of 0.012-0.014 gm, 2 genotypes having shoot dry weight of 0.014-0.016gm,

respectively (Fig. 16).

➤ *Descriptive Statistics for 14th Day Shoot Area*

The mean value of 14th day shoot area was found to be 2.20 cm². The lowest and highest shoot on 14th day was found to be 1.11cm² and 4.02cm², respectively. The kurtosis value for mean shoot area on 14th day was found to be 0.02 and the skewness value was found to be 0.47, respectively. In the histogram analysis, it was found that 22 genotypes having shoot area of 1-1.5 cm², 83 genotypes having shoot area of 1.5-2.0 cm², 97 genotypes having shoot area of 2.0-2.5 cm², 56 genotypes having shoot area of 2.5-3.0 cm², 17 genotypes having shoot area of 3.0-3.5cm², 5 genotypes having shoot area of 3.5-4.0 cm², and one genotype having shoot area of 4.0-4.5cm² respectively (Fig. 18).

➤ *Descriptive Statistics for 21st Day Shoot Length*

The mean value of 21st day shoot length was found to be 12.22 cm. The lowest and highest number of shoot length on 21st day was found to be 2 and 18.75 cm, respectively. The kurtosis value for mean shoot length on 21st day was found to be 2.53 and the skewness value was found to be -0.39, respectively. In the histogram analysis, it was found that 227 genotypes have shoot length of 10-15cm, 12 genotypes having shoot length of 15-20 cm, 41 genotypes having shoot length of 5-10 cm, and one genotype having shoot length of 0-5cm, respectively (Fig. 9).

➤ *Descriptive Statistics for 21st Day Leaf Number*

The mean value of 21st day leaf number was found to be 2.82. The lowest and highest number of leafs on 21st day was found to be 0 and 4, respectively. The kurtosis value for mean leaf number on 21st day was found to be 3.59 and the skewness value was found to be -1.04. In the histogram analysis, it was found that 136 genotypes having leaves number in between 2.5-3, 48 genotypes having leaves number in between 2-2.5, 72 genotypes having leaves number in between 3-4, 22 genotypes having leaves number in between 1-2, and three genotypes having leaves was found to be one, respectively (Fig. 12).

➤ *Descriptive Statistics for 21st Day Culm Diameter*

The mean value of 21st day culm diameter was found to be 0.88cm. The lowest and highest culm diameter on 21st day was found to be 0 and 1.17 cm, respectively. The kurtosis value for mean culm diameter on 21st day was found to be 8.38 and the skewness value was found to be -1.86. In the histogram analysis, it was found that 174 genotypes having culm diameter of 0.8-1.0 cm, 54 genotypes having culm diameter of 0.6-0.8 cm, 46 genotypes having culm diameter of 1.0-1.2 cm, 3 genotypes having culm diameter of 0.4-0.6 cm, 3 genotypes having culm diameter of 0.2-2.4 cm and one genotype having culm diameter of 0.2 cm, respectively (Fig.15).

➤ *Descriptive Statistics for 21st Day Shoot Dry Weight*

The mean value of 21st day shoot dry weight was found to be 0.014 gm. The lowest and highest shoot dry weight on 21st day was found to be 0 and 0.024 gm respectively. The kurtosis value for mean shoot dry weight on 21st day was found to be 0.99 gm and the skewness value was found to be

-0.02. In the histogram analysis, it was found that 185 genotypes have shoot dry weight of 0.010-0.015 gm, 57 genotypes having shoot dry weight of 0.015-0.020 gm, 14 genotypes having shoot dry weight of 0.020-0.025 gm, 24 genotypes having shoot dry weight of 0.005-0.010 gm and two genotypes having shoot dry weight of range 0-0.005, respectively (Fig. 17).

➤ *Descriptive Statistics for 21st Day Shoot Area*

The mean value of 21st day shoot area was found to be 4.38 cm². The lowest and highest shoot area on 21st day was found to be 0 and 7.35 cm², respectively. The kurtosis value for mean shoot on 21st day was found to be -0.16 and the skewness value was found to be -0.31. In the histogram analysis, it was found that 79 genotypes having shoot area of 4-5 cm², 75 genotypes having shoot area of 5-6 cm², 23 genotypes having shoot area of 6-7 cm², one genotype having shoot area of 7-8 cm², 64 genotypes having shoot area of 3-4 cm², 33 genotypes having shoot area of 2-3cm², 5 genotypes having shoot area of 1-2cm², and one genotype having shoot area of 1 cm², respectively (Fig.19).

➤ *Descriptive Statistics for 28th Day Shoot Length*

The mean value of 28th day shoot length was found to be 12.95 cm. The lowest and highest number of shoot length on 28th day was found to be 0 and 19.81 cm, respectively. The kurtosis value for mean shoot length on 28th day was found to be 1.80 and the skewness value was found to be -0.45. In the histogram analysis it was found that 212 genotypes were having shoot length of 10-15 cm, 37 genotypes having shoot length of 15-20 cm, 30 genotypes were having shoot length of 5-10 cm, and two genotypes were having shoot length of 0-5 cm, respectively (Fig.10).

➤ *Descriptive Statistics for 28th Day Leaf Number*

The mean value of 28th day leaf number was found to be 3.05. The lowest and highest number of leaf on 28th day was found to be 0 and 3.5, respectively. The kurtosis value for mean leaf number on 28th day was found to be 89.6 and the skewness value was found to be -7.68. In the histogram analysis, it was found that 174 genotypes are found to be having leaves number in between 2.5-3.0, 102 genotypes were having leaves number in between 3.0-3.5, 3 genotypes having leaves number in between 2.0-2.5, one genotype were having leaves number in the range between 1.0-1.5 and one genotype were having in the range between 0-0.5, respectively (Fig.13).

➤ *Correlation Analysis for Different Traits*

The mean germination data positively correlated with mean shoot length, mean leaf number, mean culm diameter, mean shoot dry weight, and mean shoot area of 21st day among the 281 genotypes. Similarly, mean shoot length data on 7th day was positively correlated with mean shoot length of 11th day, mean shoot length of 13th day, mean shoot length of 14th day, mean shoot dry weight of 14th day, mean shoot area of 14th day, and mean shoot length of 28th days. The mean shoot length data on 11th day positively correlated with mean shoot length of 13th day, mean shoot length of 14th day, mean shoot dry weight of 14th day, mean shoot length of 21st day, and mean shoot dry weight of 21st day. Besides, mean shoot

length data on 13th day positively correlated with mean shoot length 14th day, mean shoot dry weight of 14th day, mean shoot length of 21st day, and mean shoot length of 28th day.

The mean shoot length data on 14th day was positively correlated with mean shoot dry weight of 14th day, mean shoot area of 14th day, mean shoot length of 21st day, and mean shoot length of 28th days. Further, mean leaf number data on 14th day correlated with mean shoot dry weight of 14th day. Also, mean culm diameter on 14th day positively correlated with mean shoot dry weight of 14th day and mean culm diameter of 21st day. The mean shoot dry weights on 14th day positively correlated with mean shoot area of 14th day, mean shoot length of 21st day, mean shoot dry weight of 21st day and mean shoot length of 28th day. In addition, mean shoot area on 14th day correlated with mean shoot length of 21st day, mean shoot length of 28th days and mean shoot dry weight of 21st days.

The mean shoot length of 21st day correlated positively with mean shoot dry weight of 21st day, mean shoot area of 21st day, and mean shoot length of 28th day. The mean leaf number of 21st day correlated positively with mean culm diameter of 21st day and mean shoot dry weight of 21st day. The mean culm diameter of 21st day correlated positively with mean shoot dry weight of 21st day and mean shoot area of 21st day. The mean shoot dry weight of 21st day correlated with mean shoot area of 21st day and mean shoot length of 28th day. The mean shoot area of 21st day correlated positively with mean shoot length of 28th days. The mean shoot length of 28th day positively correlated with mean leaf number of 28th days.

➤ HC Cluster Correlation Analysis

In the cluster diagram of correlation analysis, it was found that mean germination of 9th day was highly correlated with mean germination of 10th day. Similarly, mean germination of 6th day was grouped and correlated with mean germination of 5th day, and mean germination of 8th day was correlated with mean germination of 7th day. Further, mean shoot length of 13th day was clustered and positively correlated with mean shoot length of 11th day. The mean shoot length of 28th day was highly correlated with mean shoot length of 21st day, mean shoot length of 14th day was correlated with mean shoot length of 28th day respectively and all three traits were found to be grouped together in the analysis. The mean germination data of 5th, 6th, 7th, 8th day and mean shoot length of 7th, 11th and 21st day are highly correlated and grouped with each other, respectively. The mean shoot length of 21st, 28th and mean shoot area of 14th and mean shoot length of 14th and mean shoot dry weight are highly correlated with each other. The mean leaves number of 21st, mean culm diameter of 21st, mean shoot dry weight of 21st, mean shoot area of 21st, mean germination data of 9th and mean germination data of 10th day are highly correlated and clustered together in the analysis (Fig. 1).

➤ Principal Component Analysis and Phylogenetic Tree

• Principal Component Analysis for Phenotypic Traits

Among the 281 number of genotypes taken for experiment purpose, 111 genotypes are found to be in PCA1 and 170 genotypes are found to be in PCA2. PCA1 component constituted 29.93% and PCA2 constituted the 13.68% of total variation in the analysis. Mean culm diameter of 21st (MCD21) and mean leaves number of 21st (MNL21) are highly correlated and grouped in the same principal component. The mean germination of 10th day and mean germination of 9th day are correlated and grouped together in the same principal component. The mean shoot length of 11th and mean shoot length of 13th day are highly correlated with each other. The mean shoot length of 21st day and mean germination of 7th day are highly correlated with each other. The mean culm diameter of 21st day and mean leaves number of 21st days are highly correlated with each other. The mean germination of 5th, 6th and mean shoot area of 14th, mean shoot dry weight of 14th day and mean shoot length of 28th, mean shoot length of 14th day and mean shoot length of 7th day are correlated with each other. The mean shoot area of 21st day, mean germination of 8th day and mean shoot dry weight of 21st day are correlated with each other and grouped in the same principal component. (Fig. 2).

• Principal Component Analysis for Genotypic Traits

Among the 281 number of genotypes taken for experiment purpose, for PCA it was found that 1st principal component having 28%, 2nd principal component having 8%, 3rd principal component having 4%, 4th principal component having 3%, 5th principal component having 2.8%, 6th principal component having 2.5%, 7th principal component having 2.4%, 8th principal component having 2.3%, 9th principal component having 2.1% and 10th principal component having 2.0% respectively (Fig. 3)

➤ Phylogenetic Tree

The total 281 number of rice accession taken for our experiment purpose it was found that Phylogenetic tree classified into three groups. Group 1 consists of 215 genotypes; group 2 consists of 38 genotypes and group 3 consists of 28 genotypes respectively. The group 2 classified into 3 sub population among them ind1A subpopulation consists of 3 genotypes, ind1B subpopulation consists of 7 genotypes, ind2 subpopulation consists of 3 genotypes, ind3 subpopulation consists of 8 genotypes, indx subpopulation consists of 13 genotypes and aus subpopulation consists of 4 genotypes respectively. The group 3 classified into 7 subpopulation ind1A subpopulation consists of 3 genotypes, ind1B subpopulation consists of 6 genotypes, ind3 subpopulation consists of 4 genotypes, ind2 subpopulation consists of 3 genotypes, indx subpopulation consists of 4 genotypes, aus subpopulation consists of 7 genotypes and admix subpopulation consists of 1 genotypes respectively (Fig. 24).

➤ ANOVA for Single Factor Analysis

Among the 281 genotypes taken for experiment purpose, sum of squares between traits was found to be 2432108 and sum of square within traits was found to be 22371.91, mean sum square between traits was found to be 110550.4 and mean sum of square within traits was found to be 3.473899. P-value for the ANOVA was found to be 0 which is found to be highly significant (Table 2).

➤ ANOVA for Two Factors with Replication Analysis

Among the 281 genotypes taken for experiment purpose sum of square between traits was found to be 206290.9 and sum of square between genotypes was found to be 4244.678, sum of square due to interaction was found to be 14647.54 and sum of square within the trait and genotype was found to be 6912.969. Mean sum of square between the traits was found to be 9823.375 and mean sum of square between genotype was found to be 15.1597, mean sum of square due to interaction was found to be 2.491078 and mean sum of square within the trait and genotype was found to be 1.118242. Further, P-value for traits was found to be 0 which is significant, and P-value for genotypes was found to be 0 which is also significant. Similarly, P-value for interaction between the traits and genotype was also found to be highly significant (1.8663×10^{-208}) (Table 3)

➤ ANOVA Analysis for Different Traits Related to Germination and Seedling Vigour in Rice

The completely randomized design was used to analysis the significance of the trait and genotypes. The germination of 4th day was having sum of square 12.1388, mean sum of square 0.0434, P-value of 0.7066 which was not significant and test statistics of 2.1047. Germination of 5th day was having sum of square 398.9929, mean sum of square 1.425, and significant P-value of 0.0007. Germination of 6th day was having sum of square 1803.4484, mean sum of square 6.4409, and highly significant P-value of 0. Germination of 7th day was having sum of square 3349.9502, mean sum of square 11.9641, and highly significant P-value of 0. Germination of 8th day was having sum of square 1401.448, mean sum of square 5.0052, and highly significant P-value of 0. Germination of 9th day was having sum of square 424.8719, mean sum of square 1.5174, and highly significant P-value of 0.0968. Germination of 10th day was having sum of square 258.516, mean sum of square 0.9233, non-significant P-value of 0.2919.

The shoot length of 7th day was having sum of square of 682.7678, mean sum of square 2.4385, P-value of 0 which was highly significant and test statistics of 4.5994. The shoot length of 11th day was having sum of square of 552.5948, mean sum of square of 1.9736, P-value of 0 which was highly significant. The shoot length of 13th day was having sum of square of 700.1553, mean sum of square of 2.5006, P-value of 0 which was highly significant. The shoot length of 14th day was having sum of square of 1429.7618, mean sum of square of 5.1063, P-value of 0 which was highly significant. The shoot length of 21st day was having sum of square of 3050.3935, mean sum of square of 10.8493, P-value of 0 which was highly significant. The shoot length of 28th day was having sum of square of 3635.2954, mean sum of square

of 12.9832, P-value of 0 which was highly significant.

The Number of leaves 14th day was having sum of square of 10.2681, mean sum of square of 0.0367, P-value of 0 which was significant. The number of leaves of 21st day was having sum of square of 150.762, mean sum of square of 0.5384, P-value of 0 which was highly significant. The number of leaves of 28th day was having sum of square of 34.0668, mean sum of square of 0.1217, P-value of 0 which was highly significant.

The culm diameter of 14th day was having sum of square of 4.9591, mean sum of square of 0.0177, P-value of 0 which was highly significant. The culm diameter of 21st day was having sum of square of 10.7527, mean sum of square of 0.0384, P-value of 0 which was highly significant.

The Shoot dry weight of 14th day was having sum of square of 0.0023, mean sum of square of 0, P-value of 0 which was highly significant. The shoot dry weight of 21st day was having sum of square of 0.0066, mean sum of square of 0, P-value of 0 which was highly significant.

The shoot area of 14th day was having sum of square of 166.5707, mean sum of square of 0.5949, P-value of 0 which was highly significant. The shoot area of 21st day was having sum of square of 814.4944, mean sum of square of 2.9089, P-value of 0 which was highly significant respectively.

In maximum number of traits, it was found that P-value was found to be 0 which is highly significant. (Table.4)

➤ GWAS Analysis

The mean germination trait was found to associated with SNP value of 335676148 and this SNP value was linked with chromosome number 11 having position of 18983591 and P-value of 6.99E-12 having minimum allelic frequency of 0.11744 and the SNP value was highly matched with i.e-49.98% with the phenotypic variance. The mean Leaves number of 28th day trait was found to be associated with SNP value of 171071000 and this SNP value linked with chromosome number 5 having position of 19946314 and P-value of 4.99E-43 having minimum allelic frequency of 0.14057 and the SNP value was matched with i.e-41.01% with the phenotypic variance. The mean Shoot Length of 28th day trait was found to associated with SNP value of 181062773 and this SNP value was linked with chromosome number 5 having position of 76225316 and P-value of 3.42E-13 having minimum allelic frequency of 0.13167 and the SNP value was highly matched with i.e-29.75% with the phenotypic variance. (Table.5)

The mean germination trait having 1SNP of LOD value of 11 and found to be under chromosome number 11. The Mean leaves number of 28th day having 37 SNP of LOD value 7-43 and are found to be under chromosome number 3 having 11 SNP ,4 having 2SNP ,5 having 8SNP ,7 having 1SNP ,8 having 8SNP,9 having 4SNP ,10 having 3SNP respectively .The mean Shoot length of 28th day having 2SNP of LOD value more than 8 and are found to be under chromosome number 5.The mean Leaves number of 28th day having 2 SNP

of LOD value 8-12 and are found to be under chromosome number 2 having 1SNP, 6 having 1SNP respectively.

➤ *BLINK Model for GWAS Analysis*

In the BLINK model of GWAS analysis for mean germination trait was found to be associated with SNP value of 335676148 having chromosome number 11, with P-value of 6.99E-12, with minimum allelic frequency of 0.117438 and phenotypic variance of 49.98054587 respectively. In the BLINK model of GWAS analysis for mean germination of 5th day trait was found to be associated with SNP value of 11164482, 201271236, 322371914 having chromosome number 1,6,11, with P-value of 4.54E-11, 1.05E-12, 3.49E-09, with minimum allelic frequency of 0.101423, 0.117438, 0.1121 and phenotypic variance of 8.887864772, 36.0951012, 9.850987732 respectively. In the BLINK model of GWAS analysis for mean germination of 6th day trait was found to be associated with SNP value of 19345482, having chromosome number 1, with P-value of 1.08E-08, with minimum allelic frequency 0.217082 and phenotypic variance of 19.87062946 respectively. In the BLINK model of GWAS analysis for mean germination of 10th day trait was found to be associated with SNP value of 335676148, having chromosome number 11, with P-value of 6.99E-12, with minimum allelic frequency 0.117438 and phenotypic variance of 49.98054572 respectively. In the BLINK model of GWAS analysis for mean leaves number of 14th day trait was found to be associated with SNP value of 64224893, 265534668, 265599883, having chromosome number 2,8,8, with P-value of 9.52E-20, 2.18E-08, 4.68E-15, with minimum allelic frequency 0.163701, 0.129893, 0.129893 and phenotypic variance of 12.93783064, 13.13537133, 21.35824367 respectively. In the BLINK model of GWAS analysis for mean leaves number of 28th day trait was found to be associated with SNP value of 171071000, 270183386, having chromosome number 5,8, with P-value of 4.99E-43, 2.10E-23 with minimum allelic frequency 0.140569, 0.144128 and phenotypic variance of 41.01505233, 12.50369798 respectively.

In the BLINK model of GWAS analysis for mean shoot length of 14th day trait was found to be associated with SNP value of 138950562, 277190265, having chromosome number 4,9, with P-value of 3.98E-13, 6.79E-0923 with minimum allelic frequency 0.418149, 0.140569 and phenotypic variance of 11.41384459, 9.739466622 respectively. In the BLINK model of GWAS analysis for mean shoot length of 21st day trait was found to be associated with SNP value of 193032228, having chromosome number 6, with P-value of 9.21E-11 with minimum allelic frequency 0.469751 and phenotypic variance of 15.63346482 respectively. In the BLINK model of GWAS analysis for mean shoot length of 28th day trait was found to be associated with SNP value of 129578003, 277189360, having chromosome number 4,9, with P-value of 7.93E-10, 5.73E-11 with minimum allelic frequency 0.266904, 0.172598 and phenotypic variance of 9.190663758, 10.8705620 respectively.

➤ *Farm CPU MODEL*

In the Farm CPU model of GWAS analysis for mean germination trait was found to be associated with SNP value of 335676148, having chromosome number 11, with P-value of 9.99E-10 with minimum allelic frequency 0.117438 and phenotypic variance of 49.98054587 respectively. In the Farm CPU model of GWAS analysis for mean germination of 6th day trait was found to be associated with SNP value of 19345482, 189253131, having chromosome number 1,6 with P-value of 5.32E-09, 9.29E-11 with minimum allelic frequency 0.217082, 0.379004 and phenotypic variance of 9.021996887, 7.040325044 respectively. In the Farm CPU model of GWAS analysis for mean germination of 10th day trait was found to be associated with SNP value 335676148 having chromosome number 11 with P-value of 9.99E-10 with minimum allelic frequency 0.117438 and phenotypic variance of 49.98054572 respectively. In the Farm CPU model of GWAS analysis for mean leaves number of 14th day trait was found to be associated with SNP value of 64224893, 199177888 having chromosome number 2,6 with P-value of 1.04E-12, 2.65E-11 with minimum allelic frequency 0.163701, 0.115658 and phenotypic variance of 19.56053538, 11.47634817 respectively.

In the Farm CPU model of GWAS analysis for mean leaves number of 28th day trait was found to be associated with SNP value of 76225316, 198678646 having chromosome number 2,6 with P-value of 3.42E-13, 4.39E-12 with minimum allelic frequency 0.131673, 0.128114 and phenotypic variance of 29.7529929, 12.10665396 respectively.

➤ *MLM Model*

In the MLM model of GWAS analysis for mean leaves number of 14th day trait was found to be associated with SNP value 63812633, 64224893 having chromosome number 2,2 with P-value of 1.41E-08, 4.36E-09 with minimum allelic frequency 0.16726, 0.163701 and phenotypic variance of 15.46362467, 22.27478475 respectively. In the MLM model of GWAS analysis for mean shoot length of 28th day trait was found to be associated with SNP value 76225316 having chromosome number 2 with P-value of 1.82E-08 12 with minimum allelic frequency 0.131673 and phenotypic variance of 55.75397769 respectively.

IV. IDENTIFICATION OF CANDIDATE GENES

In the GWAS analysis we have got the position of SNP of various parameters taken for our experiment purpose. For the parameters like mean germination the SNP position is 18983591 on chromosome number 11, from this position we add 150kb and subtract 150kb to find the exact position of candidate genes and we found the candidate gene that is Os11g052500, similarly for the parameters like mean shoot length of 28th day the SNP region that is 32954393 on chromosome number 2 and the candidate gene found to be Os02g0776900.

V. DISCUSSION

This work attempted to identify the genomic regions associated with the early germination and seedling growth in rice through genome-wide association mapping. Both these traits are highly important for the seedling establishment in the direct seeded rice cultivation for reducing the yield penalty due to the weed proliferation under DSR conditions Xu *et al.* (2019). The major findings of this study include the identification of the genomic hotspots regulating the speed of germination and seedling establishment in rice. Further, several genotypes were identified for having the relatively higher germination rate and seedling growth in early stage of the rice growth. The findings of the study are discussed below,

➤ Early Emergence and Uniform Coleoptile Emergence

The genotypes emerged early (4 or 5 DAS) was poorly correlated with the coleoptile's emergence percentage on 10th DAS, i.e., late emergence genotypes (6 DAS) also showed 100% emergence similar to the early emerged genotypes on 10 DAS. This indicates the early emergence and uniform coleoptiles emergence are distinct process and the combination of the germination, mesocotyl elongation, and coleoptiles growth could determine the uniform coleoptile emergence trait in rice. In this study, 27 genotypes were identified which showed late coleoptiles emergence on 5 or 6 DAS, but uniform cent percent emergence was observed on the 7 DAS. In direct seeded rice cultivation, uniform coleoptiles emergence is an important trait which is related to the seedling establishment Ogiwara and Terashima. (2007) and Alibu *et al.*, (2012). Further, elongation ability of the mesocotyl and coleoptiles were found to be relatively higher in the weedy rice rather than bred varieties Chung *et al.* (2010). Previously, *OsSUT* (Sucrose Transporters) has been identified for regulating the germination and coleoptiles length in rice Scofield *et al.* (2007) and Julius *et al.* (2017). Thus, understanding the mechanism regulating the uniform coleoptiles emergence would be providing valuable information for better seedling establishment in direct seeded rice

➤ Coleoptile Emergence and its Relationship with Early Shoot Length (14 DAS)

Interestingly, mean coleoptiles emergence on 7th DAS was found to be highly correlated with the early mean shoot length (7, 9, and 13 DAS) but not for 28 DAS. This indicates the early coleoptiles emergence also promotes the early shoot vigour in rice. Previously, coleoptiles growth variation in the varieties was associated with the seedling vigour in the direct seeded rice under submerged conditions Ogiwara and Terashima. (2001). Further, seedling vigour was considered as an important trait in the crop establishment in the direct seeded rice Quilloy *et al.* (2021). Also, coleoptiles length was associated with the seedling vigour in rice Xie *et al.* (2014). Further, 27 genotypes which showed high emergence on 7 DAS also had high values for the shoot length (~11.06% high than other genotypes) up to 13 DAS. Therefore, selection of the genotypes having both the high coleoptiles emergence and high early shoot growth could be useful in DSR conditions in seedling establishment.

➤ Relationship between Culm Diameter and Number of Leaves (21 DAS)

Strong culm is a major trait in the direct seeded rice especially during the heading stage of the crop Yadav *et al.* (2017). In this work, a significant correlation was found between the culm diameter and number of leaves only in the 21 DAS but not in 14 DAS. This indicates the new leaf development coupled with the increase in the culm diameter would be a desirable trait in the direct seeded rice cultivation. Further, lodging resistant and grain yield was highly correlated in the direct seeded rice cultivation Yadav *et al.* (2017). Further, plant architecture with strong culm was suggested as an important trait in DSR. In this study, ~46 genotypes have been identified having culm diameter of > 1 mm at 21 DAS. The relationship between the culm diameter in the early seedling establishment and yield in DSR needs to be properly understood. Apart from the number of leaves on 21 DAS, mean emergence on 10 DAS was also found to be highly correlated with the culm diameter. The culm diameter of the genotypes having high emergence was found to be 9% higher than other genotypes having lesser emergence percentage. This indicates the selection of the genotypes with high emergence in direct seeded rice would also have added benefit of the strong culm indicating the association of these two traits in different stages of the crop growth. The genomic regions associated with the culm diameter in 21 DAS needs to be properly understood.

➤ Relationship between the Mean Coleoptiles' Emergence on 10th DAS with Shoot Dry Weight and Shoot Area (28 DAS)

The mean coleoptiles emergence (10 DAS) was also found to be highly correlated with the shoot dry weight and shoot area of 28 DAS. In support of the findings, germination rate and shoot dry weight for seedling vigour was found to be highly correlated in a recombinant inbred line mapping population developed between the Minghui 63 and Zhenshan 97 parents Cui *et al.* (2002). This indicates the selection of genotypes for the higher coleoptiles emergence during the germination stage is associated with the better seedling growth and vigour in rice. Thus, the characterization of the genomic regions would greatly assist in the identification of the genes regulating the seedling vigour and coleoptiles emergence in rice.

➤ Principal Component and Phylogenetic Analysis using the Genome-Wide Polymorphisms in Different Sub-Populations of Rice

The principal component and phylogenetic tree grouped the 281 genotypes into three subgroups. Especially, both the subgroups II and III comprised of the unique genotypes from the *indica* and *aus* subpopulations of rice. This indicates few of the genotypes (64 nos.) of *indica* and *aus* subpopulations are genetically different from the remaining *indica* subpopulations. Additionally, the remaining genotypes of *indica* were found to be related to the genotypes of japonica rice genotypes. As reported by the Wang *et al.* (2022), the regional adaptation could be a major factor in the differentiation of the specific indica lines in the studied genotypes of rice.

➤ *Genome-Wide Association Analysis for Early Coleoptiles Emergence and Seedling Vigour in Rice*

In this analysis, 16 significant associations (LOD Score >7) for different traits were identified using the three different models (MLM, farmCPU, and blink) for GWAS studies. The phenotypic variation ranged between 7 and 49% for different traits. Especially, one major QTL was identified for the mean coleoptiles emergence for 10 DAS on 11th chromosome (18983591) which explained 49% of the phenotypic variance. Previously, Li *et al.* (2021) identified a significant QTL for germination and seed vigour in the Chr11 (13.93 Mbp) and a causal gene was identified and characterized (*OsOMT*; LOC_Os11g24550). Similarly, Su *et al.* (2021) also identified a major QTL in the 11th chromosome (S11_19165397) associated with the coleoptile diameter Su *et al.* (2021) in anaerobic germination condition and Ma *et al.* (2022) have identified a QTL within the Chr11 (22Mbp) associated with the early shoot length in rice. Similarly, Yang *et al.* (2022) also identified a major QTL in the 11th chromosome regulating the improved seed germination Yang *et al.* (2022). Further, yield under direct seeded rice was found to be regulated by the QTL in the S11_17412133 genomic region indicating the co-localization of genomic regions regulating the yield and coleoptiles emergence in rice Subedi *et al.* (2019). All these reports indicate many QTLs has been identified in chr11 regulating the seed germination in rice. Therefore, identification of the QTL having 46% phenotypic variance and further characterization of the candidate genes would greatly assist in understanding the early coleoptiles emergence trait in direct seeded rice.

➤ *Shoot Length QTL and Candidate Gene Analysis*

In this study, another major QTL contributing to the shoot length variation of 29.75% was identified in the Chr02 (32954393) for shoot length trait on 28 DAS. Previous study by Ma *et al.* (2022) also identified a major QTL in the nearby genomic region (10 Mbp) regulating the shoot length in rice. A candidate gene namely Os02g0778400 UMP/CMP kinase A/adenylate kinase (LOC_Os02g53790) was located in the significant SNP region of the GWAS analysis. Further characterization of this gene would assist in elucidation of the mechanism regulating the early seedling length in rice under direct seeded rice.

VI. SUMMARY AND CONCLUSION

A. Summary

- ANOVA analysis showed P-value for traits was found to be 0 which is significant, and P-value for genotypes was found to be 0 which is also significant. Similarly, P-value for interaction between the traits and genotype was also found to be highly significant (1.8663*10⁻²⁰⁸).
- The mean germination data positively correlated with mean shoot length, mean leaf number, mean culm diameter, mean shoot dry weight, and mean shoot area of 21st day among the 281 genotypes.
- Similarly, mean germination of 6th day was grouped and correlated with mean germination of 5th day, and mean germination of 8th day was correlated with mean germination of 7th day.

- Among the 281 number of genotypes taken for experiment purpose, 111 genotypes are found to be in PCA1 and 170 genotypes are found to be in PCA2.
- PCA1 component constituted 29.93% and PCA2 constituted the 13.68% of total variation in the analysis.
- Among the 281 number of genotypes taken for experiment purpose, it was found that PCA of 1st principal component having 28%, 2nd principal component having 8%, 3rd principal component having 4% of variation.
- Whole genome phylogenetic analysis showed three major groups of which Group 1 consists of 215 genotypes, group 2 consists of 38 genotypes and group 3 consists of 28 genotypes respectively.
- The genotypes emerged early (4 or 5 DAS) was poorly correlated with the coleoptile's emergence percentage on 10th DAS, i.e. late emergence genotypes (6 DAS) also showed 100% emergence similar to the early emerged genotypes on 10 DAS.
- Interestingly, mean coleoptiles emergence on 7th DAS was found to be highly correlated with the early mean shoot length (7, 9, and 13 DAS) but not for 28 DAS.
- Strong culm is a major trait in the direct seeded rice especially during the heading stage of the crop.
- The culm diameter of the genotypes having high emergence was found to be 9% higher than other genotypes having lesser emergence percentage.
- The genomic regions associated with the culm diameter in 21 DAS needs to be properly understood.
- The mean coleoptiles emergence (10 DAS) was also found to be highly correlated with the shoot dry weight and shoot area of 28 DAS.
- The principal component and phylogenetic tree grouped the 281 genotypes into three subgroups. Especially, both the subgroups II and III comprised of the unique genotypes from the *indica* and *aus* subpopulations of rice.
- In this analysis, 16 significant associations (LOD Score >7) for different traits were identified using the three different models (MLM, farmCPU, and blink) for GWAS studies.
- Especially, one major QTL was identified for the mean coleoptiles emergence for 10 DAS on 11th chromosome (18983591) which explained 49% of the phenotypic variance.
- In this study, another major QTL contributing to the shoot length variation of 29.75% was identified in the Chr02 (32954393) for shoot length trait on 28 DAS.
- A candidate gene namely Os02g0778400 UMP/CMP kinase A/adenylate kinase (LOC_Os02g53790) was located in the significant SNP region of the GWAS analysis.
- Further characterization of this gene would assist in elucidation of the mechanism regulating the early seedling length in rice under direct seeded rice.

B. Conclusion

Significant results were found in this association analysis of this 281 rice germplasm accession and are found to be considering its emerging importance of DSR in present agriculture scenario and scope for its cultivation in different Agro-climatic region, the present study was carried out with

phenotype diverse rice germplasm for uniform coleoptiles emergence and early seedling growth in rice and analyse the association mapping of the candidate genes associated with uniform coleoptiles emergence and early seedling growth in rice and the outcomes are found to be one major QTL was identified on **11th chromosome with 49%** of phenotypic variance, and another QTL associated with **chromosome number 02 with 29.75%** of phenotypic variance and a candidate gene namely **Os02g0778400 UMP/CMP kinase A**.

REFERENCES

- [1]. **Abbai, R. Singh, V. K. Nachimuthu, V. V. Sinha, P. Selvaraj, R. Vipparla, A. K. and Kumar, A. (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 biotechnology journal*. **17**(8), 1612-1622.
- [2]. **Abe, A. Takagi, H. Fujibe, T. Aya, K. Kojima, M. Sakakibara, H. and Terauchi, R. (2012)** OsGA20ox1, a candidate gene for a major QTL controlling seedling vigor in rice. *Theoretical and Applied Genetics*. **125**, 647-657.
- [3]. **Adhikary, S. Biswas, B. Chakraborty, D. Timsina, J. Pal, S. Chandra Tarafdar, J. and Roy, S. (2022)** Seed priming with selenium and zinc nanoparticles modifies germination, growth, and yield of direct-seeded rice (*Oryza sativa* L.). *Scientific Reports*. **12**(1), 7103.
- [4]. **Anandan, A. Anumalla, M., Pradhan, S. K., and Ali, J. (2016)** Population structure, diversity and trait association analysis in rice (*Oryza sativa* L.) germplasm for early seedling vigor (ESV) using trait linked SSR markers. *PloS one*. **11**(3), e0152406.
- [5]. **Anandan, A. Panda, S. Sabarinathan, S. Travis, A. J. Norton, G. J. and Price, A. H. (2022)** Superior Haplotypes for Early Root Vigor Traits in Rice Under Dry Direct Seeded Low Nitrogen Condition Through Genome Wide Association Mapping. *Frontiers in plant science*. **13**.
- [6]. **Barik, S. R. Pandit, E. Sanghamitra, P. Mohanty, S. P. Behera, Mishra, J. and Pradhan, S. K. (2022)** Unraveling the genomic regions controlling the seed vigour index, root growth parameters and germination per cent in rice. *Plos one*. **17**(7), e0267303.
- [7]. **Campbell, M. T. Du, Q. Liu, K. Brien, C. J. Berger, B. Zhang, C. and Walia, H. (2017)** A comprehensive image-based phenomic analysis reveals the complex genetic architecture of shoot growth dynamics in rice (*Oryza sativa* L.). *The Plant Genome*. **10**(2), plantgenome 2016-07.
- [8]. **Chamara, B. S. Marambe, B. Kumar, V. Ismail, A. M. Septiningsih, E. M. and Chauhan, B. S. (2018)** Optimizing sowing and flooding depth for anaerobic germination-tolerant genotypes to enhance crop establishment, early growth, and weed management in dry-seeded rice (*Oryza sativa* L.). *Frontiers in Plant Science*. **9**, 1654.
- [9]. **Chen, K. Zhang, Q. Wang, C. C. Liu, Z. X. Jiang, Y. J. Zhai, L. Y. and Li, Z. K. (2019)** Genetic dissection of seedling vigour in a diverse panel from the 3,000 Rice (*Oryza sativa* L.) Genome Project. *Scientific Reports*. **9**(1), 4804.
- [10]. **Chung, N. J. (2010)** Elongation habit of mesocotyls and coleoptiles in weedy rice with high emergence ability in direct-seeding on dry paddy fields. *Crop and Pasture Science*. **61**(11), 911-917.
- [11]. **Counce, P. A. Keisling, T. C. and Mitchell, A. J. (2000)** A uniform, objective, and adaptive system for expressing rice development. *Crop Science*. **40**(2), 436-443.
- [12]. **Cui, K. Peng, S. Xing, Y. Xu, C. Yu, S. & Zhang, Q. (2002)** Molecular dissection of seedling-vigor and associated physiological traits in rice. *Theoretical and Applied genetics*. **105**, 745-753.
- [13]. **Dimaano, N. G. B. Ali, J. Mahender, A. Sta. Cruz, P. C. Baltazar, A. M. Diaz, M. G. Q. and Li, Z. (2020)** Identification of quantitative trait loci governing early germination and seedling vigor traits related to weed competitive ability in rice. *Euphytica*. **216**, 1-20.
- [14]. **Farooq, M. K. H. M. Siddique, K. H. Rehman, H., Aziz, T. Lee, D. J. and Wahid, A. (2011)** Rice direct seeding: experiences, challenges and opportunities. *Soil and Tillage Research*. **111**(2), 87-98.
- [15]. **Guo, T. Yang, J. Li, D. Sun, K. Luo, L. Xiao, W and Chen, Z. (2019)** Integrating GWAS, QTL, mapping and RNA-seq to identify candidate genes for seed vigor in rice (*Oryza sativa* L.). *Molecular Breeding*. **39**, 1-16.
- [16]. **Jang, S. G. Park, S. Y., Lar, S. M., Zhang, H., Lee, A. R. Cao, F. Y. and Kwon, S. W. (2021)** Genome-Wide Association Study (GWAS) of Mesocotyl Length for Direct Seeding in Rice. *Agronomy*. **11**(12), 2527.
- [17]. **Julius, B. T. Leach, K. A. Tran, T. M., Mertz, R. A. and Braun, D. M. (2017)** Sugar transporters in plants: new insights and discoveries. *Plant and Cell Physiology*. **58**(9), 1442-1460.
- [18]. **Kalluru, S. Vemireddy, L. R. Ramireddy, E. Mohan Reddy, D. and Umamahesh, V. (2023)** Identification of molecular markers and putative candidate genes associated with early seedling vigour traits in rice (*Oryza sativa* L.). *Brazilian Journal of Botany*. **46**(1), 35-49.
- [19]. **Kumari, S. Sharma, N., and Raghuram, N. (2021)** Meta-analysis of yield-related and N-responsive genes reveals chromosomal hotspots, key processes and candidate genes for nitrogen-use efficiency in rice. *Frontiers in Plant Science*. **12**, 627955.
- [20]. **Li, M. Sun, P. Zhou, H. Chen, S. and Yu, S. (2011)** Identification of quantitative trait loci associated with germination using chromosome segment substitution lines of rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*. **123**, 411-420.

- [21]. **Li, W. Yang, B. Xu, J. Peng, L. Sun, S. Huang, Z. and Wang, Z. (2021).** A genome-wide association study reveals that the 2-oxoglutarate/malate translocator mediates seed vigor in rice. *The Plant Journal*. **108**(2), 478-491.
- [22]. **Li, X. Zheng, H. Wu, W. Liu, H. Wang, J. Jia, Y. and Zhao, H. (2020).** QTL mapping and candidate gene analysis for alkali tolerance in Japonica rice at the bud stage based on linkage mapping and genome-wide association study. *Rice*. **13**, 1-11.
- [23]. **Ma, Y. Wang, J. Yang, T. Dong, J. Yang, W. Chen, L. and Zhao, J. (2022).** Genome-wide association mapping and gene expression analysis identify OsCPS1 as a new candidate gene controlling early seedling length in rice. *Frontiers in Plant Science*. **13**, 3014.
- [24]. **Mahajan, G. Sarlach, R. S. Japinder, S. & Gill, M. S. (2011).** Seed priming effects on germination, growth and yield of dry direct-seeded rice. *Journal of Crop Improvement*. **25**(4), 409-417.
- [25]. **Mahender, A. Anandan, A. and Pradhan, S. K. (2015).** Early seedling vigour, an imperative trait for direct-seeded rice: an overview on physio-morphological parameters and molecular markers. *Planta*. **241**, 1027-1050.
- [26]. **Mei, J. Wang, W. Peng, S. and Nie, L. (2017).** Seed pelleting with calcium peroxide improves crop establishment of direct-seeded rice under waterlogging conditions. *Scientific Reports*. **7**(1), 1-12.
- [27]. **Menard, G. Sandhu, N. Anderson, D. Catolos, M. Hassall, K. L. Eastmond, P. J. and Kurup, S. (2021).** Laboratory phenomics predicts field performance and identifies superior indica haplotypes for early seedling vigour in dry direct-seeded rice. *Genomics*. **113**(6), 4227-4236.
- [28]. **Money, D. Gardner, K. Migicovsky, Z. Schwaninger, H., Zhong, G. Y. and Myles, S. (2015).** Link Impute: fast and accurate genotype imputation for non-model organisms. *G3: Genes, Genomes, Genetics*. **5**(11), 2383-2390.
- [29]. **Ogiwara, H. and Terashima, K. (2001).** A varietal difference in coleoptile growth is correlated with seedling establishment of direct seeded rice in submerged field under low-temperature conditions. *Plant production science*. **4**(3), 166-172.
- [30]. **Pathak, H. Tewari, A. N. Sankhyani, S. Dubey, D. S. Mina, U. Singh, V. K. and Jain, N. (2011).** Direct-seeded rice: potential, performance and problems-Areview. *Current Advances in Agricultural Sciences (An International Journal)*. **3**(2), 77-88.
- [31]. **Peng, L. Sun, S. Yang, B. Zhao, J. Li, W. Huang, Z. and Wang, Z. (2022).** Genome-wide association study reveals that the cupin domain protein OsCDP3.10 regulates seed vigour in rice. *Plant Biotechnology Journal*. **20**(3), 485-498.
- [32]. **Quilloy, F. A. Labaco, B. Casal, C. and Dixit, S. (2021).** Crop establishment in direct-seeded rice: Traits, physiology, and genetics. *Rice Improvement*. **27**(8), 171-202.
- [33]. **Rao, A. N. Johnson, D. E. Sivaprasad, B. Ladha, J. K. and Mortimer, A. M. (2007).** Weed management in direct-seeded rice. *Advances in agronomy*. **93**, 153-255.
- [34]. **Roy, S. K. S. Hamid, A. Miah, M. G. and Hashem, A. (1996).** Seed size variation and its effects on germination and seedling vigour in rice. *Journal of Agronomy and Crop Science*. **176**(2), 79-82.
- [35]. **Sandhu, N. Torres, R. O. Sta Cruz, M. T. Maturan, P. C. Jain, R., Kumar, A. and Henry, A. (2015).** Traits and QTLs for development of dry direct-seeded rainfed rice varieties. *Journal of Experimental Botany*. **66**(1), 225-244.
- [36]. **Saud, S. Wang, D., Fahad, S. Alharby, H. F. Bamagoos, A. A. Mjrashi, A. and Hassan, S. (2022).** Comprehensive Impacts of Climate Change on Rice Production and Adaptive Strategies in China. *Frontiers in Microbiology*. **13**.
- [37]. **Scofield, G. N. Aoki, N. Hirose, T. Takano, M. Jenkins, C. L. and Furbank, R. T. (2007).** The role of the sucrose transporter, OsSUT1, in germination and early seedling growth and development of rice plants. *Journal of Experimental Botany*. **58**(3), 483-495.
- [38]. **Singh, U. M. Yadav, S. Dixit, S. Ramayya, P. J. Devi, M. N. Raman, K. A. and Kumar, A. (2017).** QTL hotspots for early vigor and related traits under dry direct-seeded system in rice (*Oryza sativa* L.). *Frontiers in plant science*. **8**, 286.
- [39]. **Singh, U. M., Yadav, S. Dixit, S. Ramayya, P. J. Devi, M. N. Raman, K. A. and Kumar, A. (2017).** QTL hotspots for early vigor and related traits under dry direct-seeded system in rice (*Oryza sativa* L.). *Frontiers in plant science*. **8**, 286.
- [40]. **Su, L. Yang, J. Li, D. Peng, Z. Xia, A. Yang, M. and Guo, T. (2021).** Dynamic genome-wide association analysis and identification of candidate genes involved in anaerobic germination tolerance in rice. *Rice*. **14**, 1-22.
- [41]. **Subedi, S. R. Sandhu, N. Singh, V. K. Sinha, P. Kumar, S. Singh, S. P. and Kumar, A. (2019).** Genome-wide association study reveals significant genomic regions for improving yield, adaptability of rice under dry direct seeded cultivation condition. *BMC genomics*. **20**(1), 1-20.
- [42]. **Thapa, R. and Septiningsih, E. M. (2021).** Genome-wide association study for traits related to seedling vigor in rice. *Crop Science*. **61**(6), 3931-3946.
- [43]. **Tyagi, W. Rai, M. and Dohling, A. (2012).** Haplotype analysis for Pup1 locus in rice genotypes of North-Eastern and Eastern India to identify suitable donors tolerant to low phosphorus. *SABRAO Journal of Breeding and Genetics*. **44**(2), 398-405.
- [44]. **Wang, F. Longkumer, T. Catausan, S. C. Calumpang, C. L. F. Tarun, J. A. Cattin-Ortola, J. Kretschmar, T. (2018).** Genome-wide association and gene validation studies for early root vigour to improve direct seeding of rice. *Plant, Cell & Environment*. **41**(12), 2731-2743.

- [45]. **Wang, J. and Zhang, Z. (2021).** GAPIT version 3: boosting power and accuracy for genomic association and prediction. *Genomics, proteomics & bioinformatics*. **19**(4), 6.
- [46]. **Wang, M. Chen, J. Zhou, F. Yuan, J. Chen, L. Wu, R. and Zhang, Q. (2021).** The ties of brotherhood between japonica and indica rice for regional adaptation. *Science China Life Sciences*. 1-11.
- [47]. **Wu, J. Feng, F. Lian, X. Teng, X. Wei, H. Yu, H. and Mei, H. (2015).** Genome-wide Association Study (GWAS) of mesocotyl elongation based on re-sequencing approach in rice. *BMC plant biology*. **15**(1), 1-10.
- [48]. **Xie, L. Tan, Z. Zhou, Y. Xu, R. Feng, L. Xing, Y. and Qi, X. (2014).** Identification and fine mapping of quantitative trait loci for seed vigor in germination and seedling establishment in rice. *Journal of Integrative Plant Biology*. **56**(8), 749-759.
- [49]. **Xu, L. Li, X. Wang, X. Xiong, D. and Wang, F. (2019).** Comparing the grain yields of direct-seeded and transplanted rice: A meta-analysis. *Agronomy*. **9**(11), 767.
- [50]. **Yadav, S. Singh, U. M. Naik, S. M. Venkateshwarlu, C., Ramayya, P. J. Raman, K. A. and Kumar, A. (2017).** Molecular mapping of QTLs associated with lodging resistance in dry direct-seeded rice (*Oryza sativa* L.). *Frontiers in plant science*. **8**, 1431.
- [51]. **Yang, B. Chen, M., Zhan, C., Liu, K. Cheng, Y. Xie, T. and Cheng, J. (2022).** Identification of OsPK5 involved in rice glycolytic metabolism and GA/ABA balance for improving seed germination via genome-wide association study. *Journal of Experimental Botany*. **73**(11), 3446-3461.
- [52]. **Yang, J. Guo, Z. Luo, L. Gao, Q. Xiao, W. Wang, J. and Guo, T. (2021).** Identification of QTL and candidate genes involved in early seedling growth in rice via high-density genetic mapping and RNA-seq. *The Crop Journal*. **9**(2), 360-371.
- [53]. **Yonemaru, J. I. Yamamoto, T. Ebana, K. Yamamoto, E. Nagasaki, H. Shibaya, T. and Yano, M. (2012).** Genome-wide haplotype changes produced by artificial selection during modern rice breeding in Japan. *Plosone*. **7**(3), e32982.
- [54]. **Zeng, M. Yang, J. Wu, K. Wang, H. Sun, K. Chen, Z. and Chen, C. (2021).** Genome-wide association study reveals early seedling vigour-associated quantitative trait loci in indica rice. *Euphytica*. **217**, 1-16.
- [55]. **Zhang, J. Guo, T. Yang, J. Hu, M. Wang, H. Sun, K. and Wang, H. (2020).** QTL mapping and haplotype analysis revealed candidate genes for grain thickness in rice (*Oryza sativa* L.). *Molecular Breeding*. **40**, 1-12.

Table 1: Descriptive Statistics for all the Trait

	MG	MG ₄	MG ₅	MG ₆	MG ₇	MG ₈	MG ₉	MG ₁₀	MSL ₇	MSL ₁₁	MSL ₁₃	MSL ₁₄	MNL ₄	MCD ₁	MSDW ₁₄	MSA ₁	MSL ₁	MNL ₁	MCD ₁	MSDW ₂₁	MSA ₂	MSL ₂	MNL ₂
Mean	97.83	0.04	0.49	1.78	6.63	8.71	9.59	9.78	3.549	6.057	7.18	9.291	2.004	0.689	0.0086	2.207	12.22	2.822	0.881	0.014	4.3858	12.96	3.06
Standard Error	0.405	0.01	0.05	0.10	0.14	0.09	0.05	0.04	0.066	0.059	0.07	0.095	0.008	0.006	0.0001	0.633	0.139	0.031	0.008	0.0002	0.0719	0.152	0.01
Median	100	0	0	1	7	9.5	10	10	3.542	5.967	7.06	9.1	2	0.69	0.0086	2.141	12.13	2.833	0.892	0.0138	4.3937	13.03	3
Mode	100	0	0	0	10	10	10	10	4.533	6.642	6.82	9.25	2	0.78	0.0076	#N/A	13	3	0.943	0.011	#N/A	12.87	3
Standard Deviation	6.794	0.15	0.84	1.79	2.44	1.58	0.87	0.68	1.104	0.993	1.12	1.598	0.135	0.094	0.002	0.545	2.334	0.519	0.139	0.0034	1.206	2.548	0.25
Sample Variance	46.16	0.02	0.71	3.22	5.98	2.50	0.76	0.46	1.219	0.987	1.25	2.553	0.018	0.009	4E-06	0.297	5.447	0.269	0.019	1E-05	1.4545	6.492	0.06
Kurtosis	21.45	16.5	7.46	0.35	-0.73	3.37	9.47	21.5	0.142	0.272	1.06	0.822	21.01	0.669	0.5422	0.02	2.537	3.595	8.381	0.9965	-0.162	1.809	89.7
Skewness	-4.39	3.97	2.62	1.07	-0.46	-1.75	-2.94	-4.39	-0.03	-0.04	0.26	0.648	-0.73	-0.4	0.3147	0.476	-0.4	-1.04	-1.86	-0.0265	-0.317	-0.452	-7.69
Range	50	1	4.5	7.5	9	8	5	5	7.067	5.558	7.96	9.767	1.833	0.55	0.0126	2.91	18.58	4	1.17	0.0245	7.3548	19.82	3.5
Minimum	50	0	0	0	1	2	5	5	0	2.983	3.36	5.267	1	0.348	0.0036	1.114	0	0	0	0	0	0	0
Maximum	100	1	4.5	7.5	10	10	10	10	7.067	8.542	11.3	15.03	2.833	0.896	0.0162	4.024	18.58	4	1.17	0.0245	7.3548	19.82	3.5
Sum	27490	11	139.5	501.5	1865	2449	2695	2749	997.2	1702	2017	2611	563.2	193.5	2.4284	620.2	3434	793	247.7	3.9364	1232.4	3641	859
Count	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281	281
Largest(1)	100	1	4.5	7.5	10	10	10	10	7.067	8.542	11.3	15.03	2.833	0.896	0.0162	4.024	18.58	4	1.17	0.0245	7.3548	19.82	3.5
Smallest(1)	50	0	0	0	1	2	5	5	0	2.983	3.36	5.267	1	0.348	0.0036	1.114	0	0	0	0	0	0	0
Confidence Level(95.0%)	0.798	0.02	0.09	0.21	0.28	0.18	0.1	0.08	0.13	0.117	0.13	0.188	0.016	0.011	0.0002	0.064	0.274	0.061	0.016	0.0004	0.1416	0.299	0.03

Table 2: ANOVA Analysis for Different Traits Related to Germination and Seedling Vigour in Rice

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Traits	2432108	22	110550.4	31823.14	0	1.543687
Within Traits	22371.91	6440	3.473899			
Total	2454480	6462				

Table 3: Two-Way ANOVA Analysis for Different Traits Related to Germination and Seedling Vigour in Rice

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Traits	206290.9	21	9823.375	8784.663	0	1.557462
Genotypes	4244.678	281	15.15957	13.55661	0	1.146809
Trait x Genotype	14647.54	5880	2.491078	2.227674	1.8663E-208	1.043274
Within	6912.969	6182	1.118242			
Total	232096.1	12363				

Table 4: ANOVA Analysis for Different Traits Related to Germination and Seedling Vigour in Rice

	Source	DF	SS	MSS	F-Value	P-Value	CV(%)	SE	EDF	EMS	CV	Test statistics
G4	Genotype	280	12.1388	0.0434	0.94	0.7066 ^{NS}	549.46	0.2151	281	0.5499	6.795	2.1047
	Error	281	13	0.0463								
	Total	561	25.1388									
G5	Genotype	280	398.9929	1.425	1.46	0.0007	198.73	0.9866	281	0.9733	6.795	4.7402
	Error	281	273.5	0.9733								
	Total	561	672.4929									
G6	Genotype	280	1803.448	6.4409	1.92	0	102.56	1.83	281	3.3505	6.795	8.7949
	Error	281	941.5	3.3505								
	Total	561	2744.94									
G7	Genotype	280	3349.95	11.9641	1.61	0	41.03	2.72	281	7.4164	6.795	13.0849
	Error	281	2084	7.4164								
	Total	561	5433.95									
G8	Genotype	280	1401.448	5.0052	1.82	0	19.01	1.66	281	2.7483	6.795	7.9588
	Error	281	771	2.7438								
	Total	561	2172.448									
G9	Genotype	280	424.8719	1.5174	1.17	0.0968 ^{NS}	11.88	1.14	281	0.3794	6.795	7.364
	Error	281	365	1.2989								
	Total	561	789.8719									
G10	Genotype	280	258.516	0.9233	1.07	0.2919 ^{NS}	9.51	0.9299	281	1.6876	6.795	5.6791

	Error	281	243	0.8648								
	Total	561	501.516									
SL7	Genotype	280	682.7678	2.4385	2.66	0	26.97	0.9572	281	0.9163	6.795	4.5994
	Error	281	257.4877	0.9163								
	Total	561	940.2555									
SL11	Genotype	280	552.5948	1.9736	2.45	0	14.83	0.8983	281	0.8069	6.795	4.3159
	Error	281	226.7291	0.8069								
	Total	561	779.3239									
SL13	Genotype	280	700.1553	2.5006	2.4	0	14.23	1.02	281	1.0426	6.795	4.9062
	Error	281	292.9829	1.0426								
	Total	561	993.1382									
SL14	Genotype	280	1429.762	5.1063	5.94	0	9.98	0.9274	281	0.86	6.795	4.4558
	Error	281	241.6578	0.86								
	Total	561	1671.42									
	Source	DF	SS	MSS	F-Value	P-Value	CV(%)	SE	EDF	EMS	CV	Test statistics
NL14	Genotype	280	10.2681	0.0367	3.04	0	5.48	0.1098	281	0.0121	6.795	0.5227
	Error	281	3.3889	0.0121								
	Total	561	13.657									
CD14	Genotype	280	4.9591	0.0177	4.34	0	9.27	0.0638	281	0.0041	6.795	0.3068
	Error	281	1.1455	0.0041								
	Total	561	6.1046									
SDW14	Genotype	280	0.0023	0	5.77	0	13.78	0.0012	281	0	6.795	0.0057
	Error	281	0.0004	0								
	Total	561	0.0027									
SA14	Genotype	280	166.5707	0.5949	2.57	0	21.79	0.481	281	0.2313	6.795	2.3109
	Error	281	64.999	0.2313								
	Total	561	231.5706									
SL21	Genotype	280	3050.394	10.8943	6.27	0	10.79	1.32	281	1.7388	6.795	6.3358
	Error	281	488.5987	1.7388								
	Total	561	3538.992									
NL21	Genotype	280	150.762	0.5384	5.1	0	11.51	0.3249	281	0.1056	6.795	1.5612
	Error	281	29.6667	0.1056								
	Total	561	180.4286									
CD21	Genotype	280	10.7527	0.0384	5.36	0	9.6	0.0847	281	0.0072	6.795	0.4067
	Error	281	2.0136	0.0072								
	Total	561	12.7663									
SDW21	Genotype	280	0.0066	0	6.15	0	13.95	0.002	281	0	6.795	0.0094
	Error	281	0.0011	0								
	Total	561	0.0077									
SA21	Genotype	280	814.4944	2.9089	9.47	0	12.64	0.5544	281	0.3073	6.795	2.6636
	Error	281	86.3574	0.3073								
	Total	561	900.8518									
SL28	Genotype	280	3635.295	12.9832	7.12	0	10.42	1.35	281	1.8242	6.795	6.4895
	Error	281	512.6061	1.8242								
	Total	561	4147.902									
NL28	Genotype	280	34.0668	0.1217	2.39	0	7.39	0.2259	281	0.051	6.795	1.0852
	Error	281	14.3333	0.051								
	Total	561	48.4002									

Table 5: Association Analysis for Different Traits Along with the Effect of the Phenotypic Variance

Sl.No.	SNP	Chromosome	Position	P.value	maf	nobs	effect	Phenotype_Variance_Explained (%)
1	MG							
	335676148	11	18983591	6.99E-12	0.1174	281	-6.3116	49.98054587
2	MG5							
	11164482	1	11164482	4.54E-11	0.1014	281	-0.427	8.887864772
	201271236	6	20188116	1.05E-12	0.1174	281	-0.9552	36.0951012
	322371914	11	5679357	3.49E-09	0.1121	281	-0.5171	9.850987732
3	MG6							
	19345482	1	19345482	1.08E-08	0.2171	281	0.6478	19.87062946
4	MG10							
	335676148	11	18983591	6.99E-12	0.1174	281	-0.6312	49.98054572
5	MNL14							
	64224893	2	20953970	9.52E-20	0.1637	281	-0.146	12.93783064
	265534668	8	23505140	2.18E-08	0.1299	281	0.1512	13.13537133
	265599883	8	23570355	4.68E-15	0.1299	281	0.1019	21.35824367
6	MNL28							
	171071000	5	19946314	4.99E-43	0.1406	281	-0.6423	41.01505233
	270183386	8	28153858	2.10E-23	0.1441	281	-1.2804	12.50369798
7	MSL14							
	138950562	4	23328570	3.98E-13	0.4181	281	-0.6565	11.41384459
	277190265	9	6717715	6.79E-09	0.1406	281	0.6782	9.739466622

8	MSL21							
	193032228	6	11949108	9.21E-11	0.4698	281	0.8674	15.63346482
9	MSL28							
	129578003	4	13956011	7.93E-10	0.2669	281	1.0369	9.190663758
	277189360	9	6716810	5.73E-11	0.1726	281	-1.1778	10.87056209
10	MG							
	335676148	11	18983591	9.99E-10	0.1174	281	-6.3116	49.98054587
11	MG6							
	19345482	1	19345482	5.32E-09	0.2171	281	0.4952	9.021996887
	189253131	6	8170011	9.29E-11	0.379	281	-0.5219	7.040325044
12	MG10							
	335676148	11	18983591	9.99E-10	0.1174	281	-0.6312	49.98054572
13	MNL14							
	64224893	2	20953970	1.04E-12	0.1637	281	-0.1002	19.56053538
	199177888	6	18094768	2.65E-11	0.1157	281	-0.071	11.47634817
14	MSL28							
	76225316	2	32954393	3.42E-13	0.1317	281	2.4106	29.7529929
	198678646	6	17595526	4.39E-12	0.1281	281	-1.3659	12.10665396
15	MNL14							
	63812633	2	20541710	1.41E-08	0.1673	281	0.0508	15.46362467
	64224893	2	20953970	4.36E-09	0.1637	281	0.0508	22.27478475
16	MSL28							
	76225316	2	32954393	1.82E-08	0.1317	281	0.0068	55.75397769

Table 6: Haplotype Analysis of Locus Os11g0525500

	19056 216	19056 227	19056 251	19056 277	19056 287	19056 298	19056 315	19056 325	19056 326	19056 339	19056 353	19056 367					
	His18 0Gln	Lys17 7Glu	Gln16 9Glu	Lys16 0Arg	Arg15 7Trp	Cys15 3Tyr	Glu14 7Asp	Phe14 4Cys	Phe14 4Val	Asp13 9Glu	Arg13 5Trp	Thr13 0Met	Gro ups	Co unt	Su m	Aver age	Varia nce
Hap 1	G	T	C	T	G	C	T	A	A	A	G	G	Hap 1	36	3	0.08 3333	0.03 5714
Hap 2	G	C	G	C	A	C	T	A	C	A	G	G	Hap 2	48	0. 5	0.01 0417	0.00 5208

Table 6 (a): Anova For Single Factor

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	0.109375	1	0.109375	6	0.016436	3.957388
Within Groups	1.494792	82	0.018229			
Total	1.604167	83				

Table 7: Haplotype Analysis of Locus Os02g0776900

	3282305 2	3282319 7	3282344 2	3283597 9	3284110 0	32841526	3284152 9	3284169 1	SUMMARY				
					His357A sp	Asp215A sn	Leu214I le	Thr160Pr o	Group s	Coun t	Sum	Averag e	Varianc e
Hap 1	A	A	A	A	G	C	T	G	Hap1	32	465.383 3	14.5432 3	3.04345 6
Hap 2	G	A	A	A	G	C	T	T	Hap2	161	2049.08 3	12.7272 3	6.73540 9
Hap 3	G	G	G	T	G	C	G	T	Hap3	27	349.6	12.9481 5	4.13077 7

Table 7 (a): Anova for Single Factor

ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	88.18656	2	44.09328	7.478621	0.000723	3.037472
Within Groups	1279.413	217	5.895911			
Total	1367.599	219				

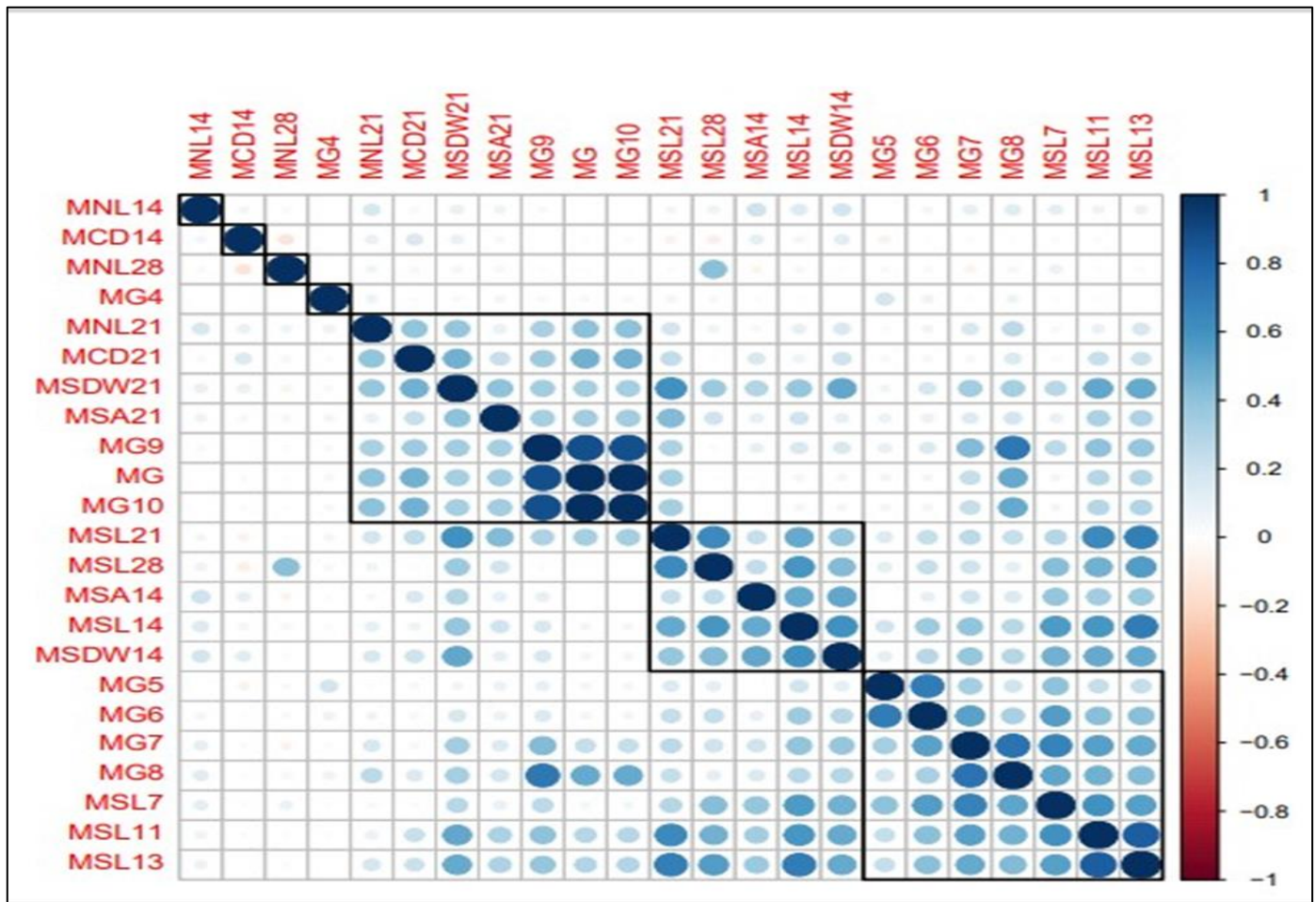


Fig 1: Correlation Analysis of all the Traits in 281 Rice Genotypes

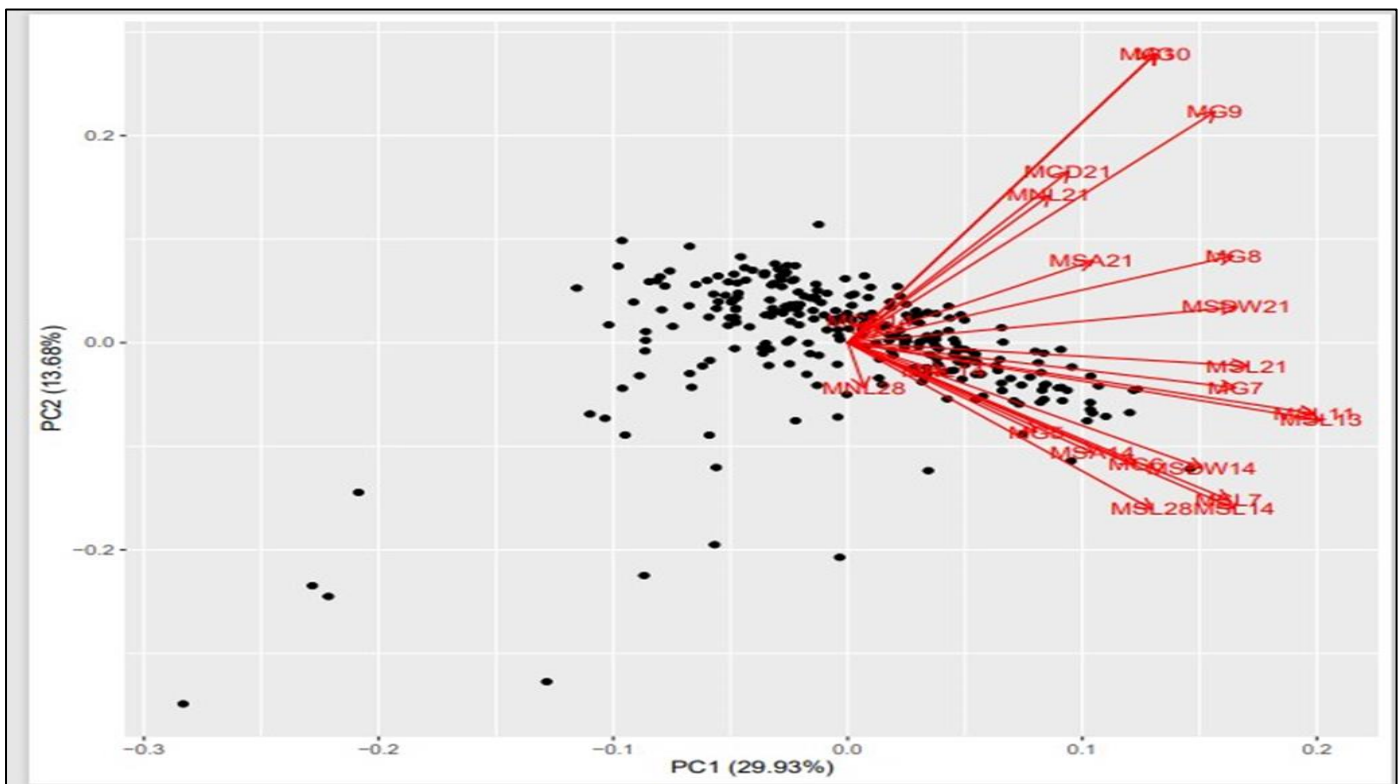


Fig 2: Principal Component Analysis of all the Genotypic Traits in 281 rice Genotypes

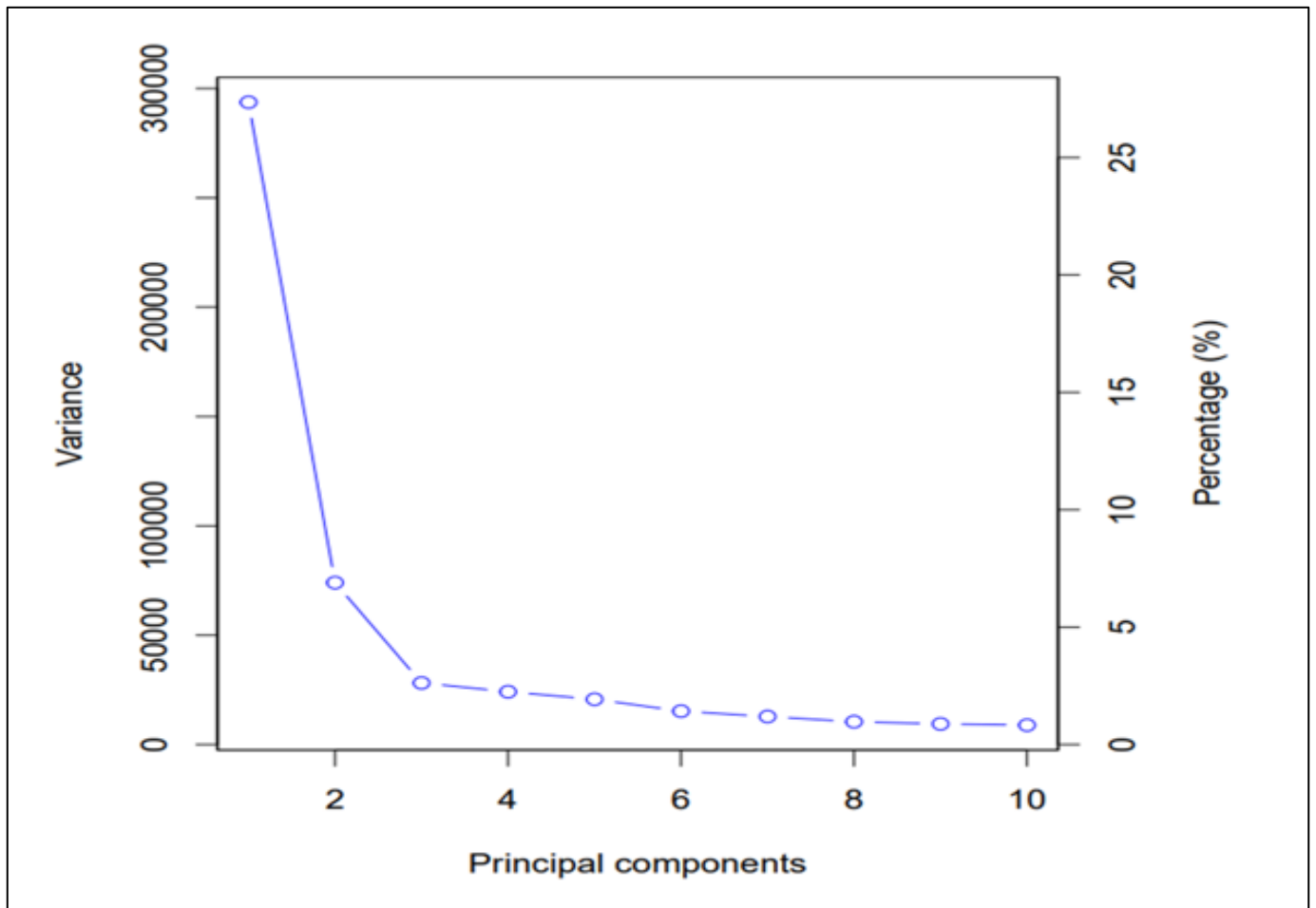


Fig 3: Principal Component Analysis of all the Genotypic Traits in 281 Rice Genotypes

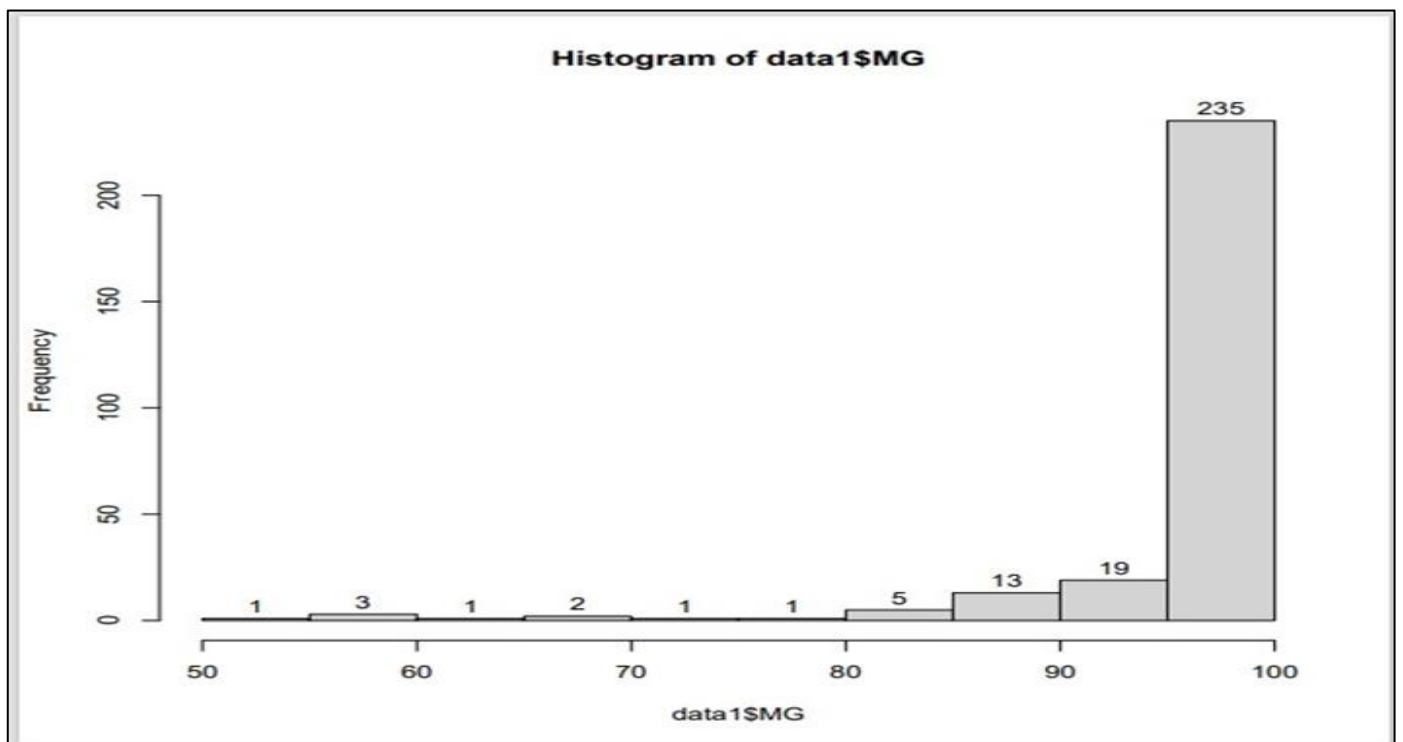


Fig 4: Histogram Analysis of Mean Germination Rate in the 281 Rice Genotypes

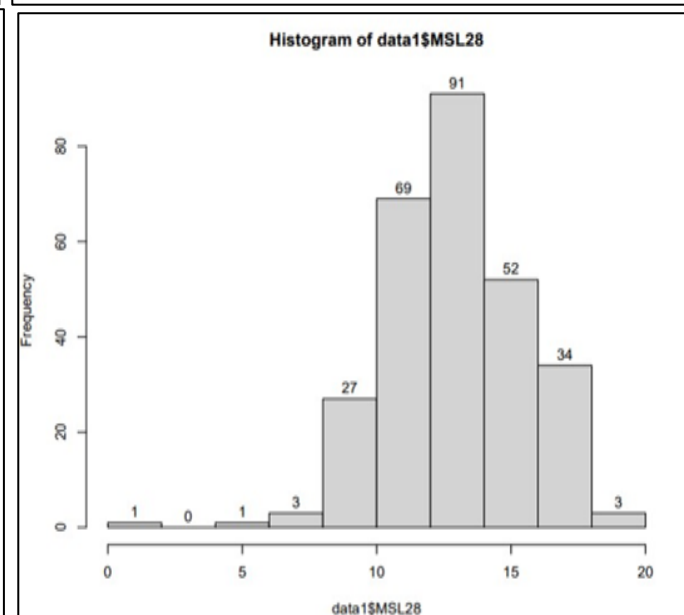
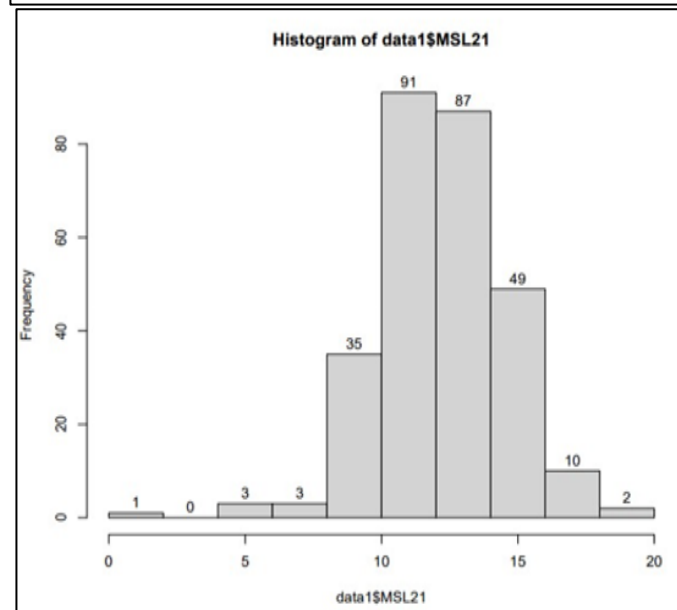
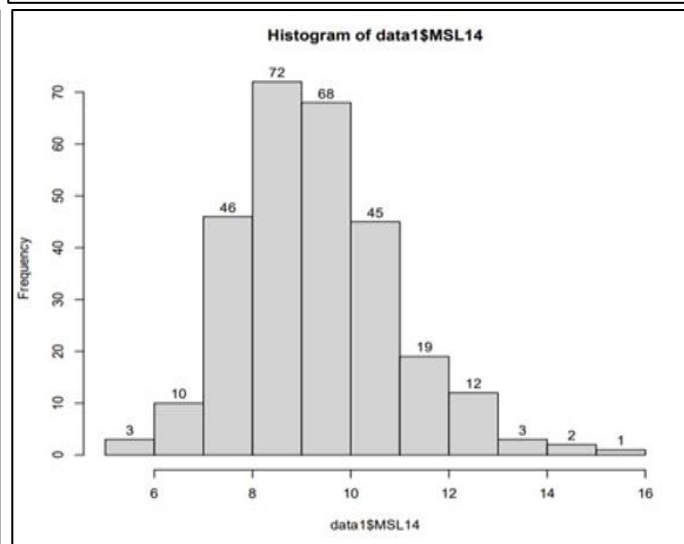
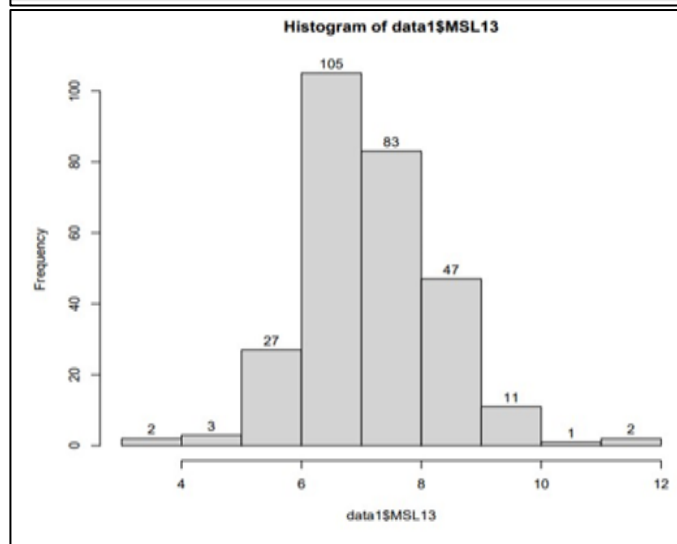
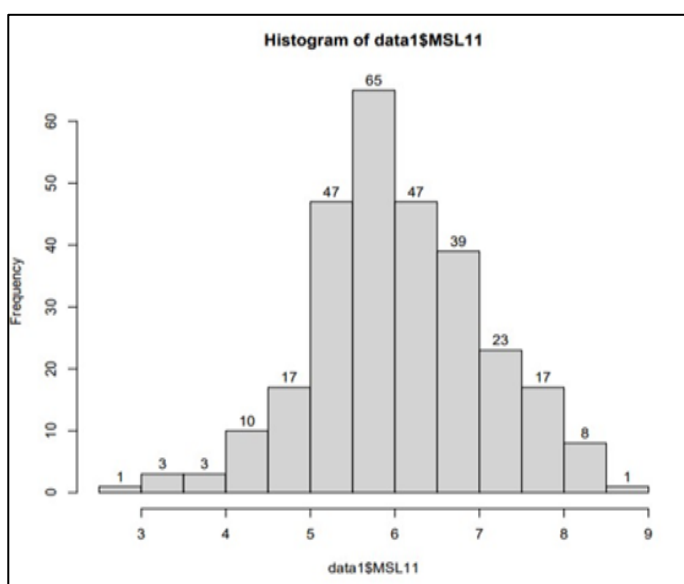
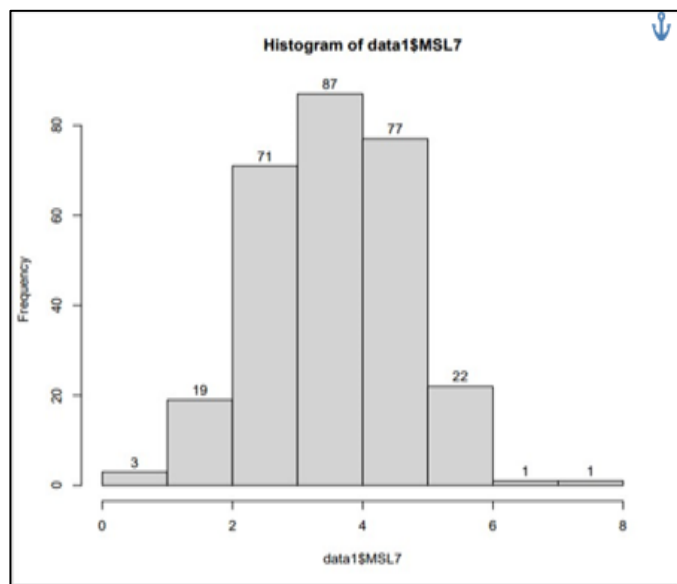
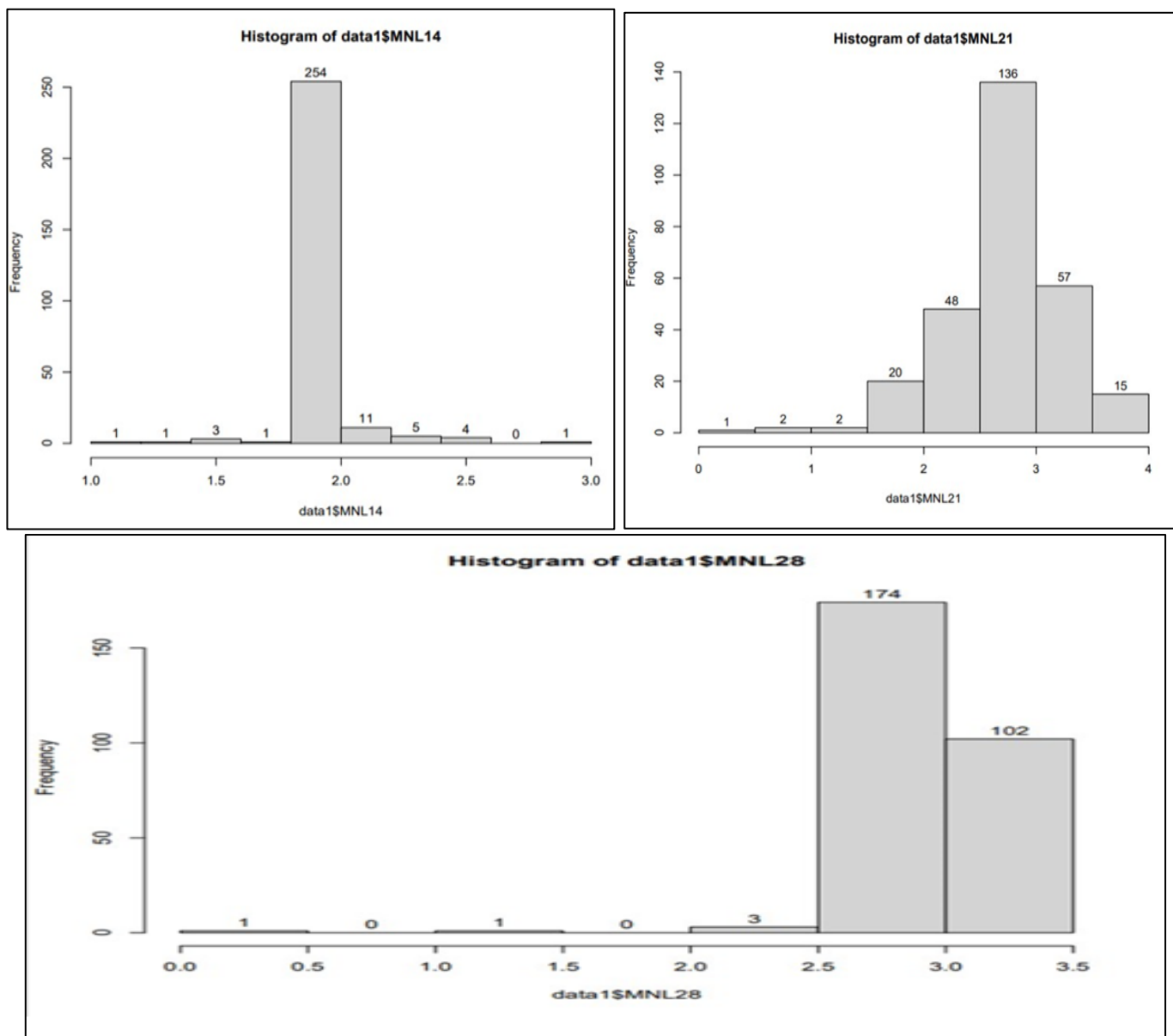
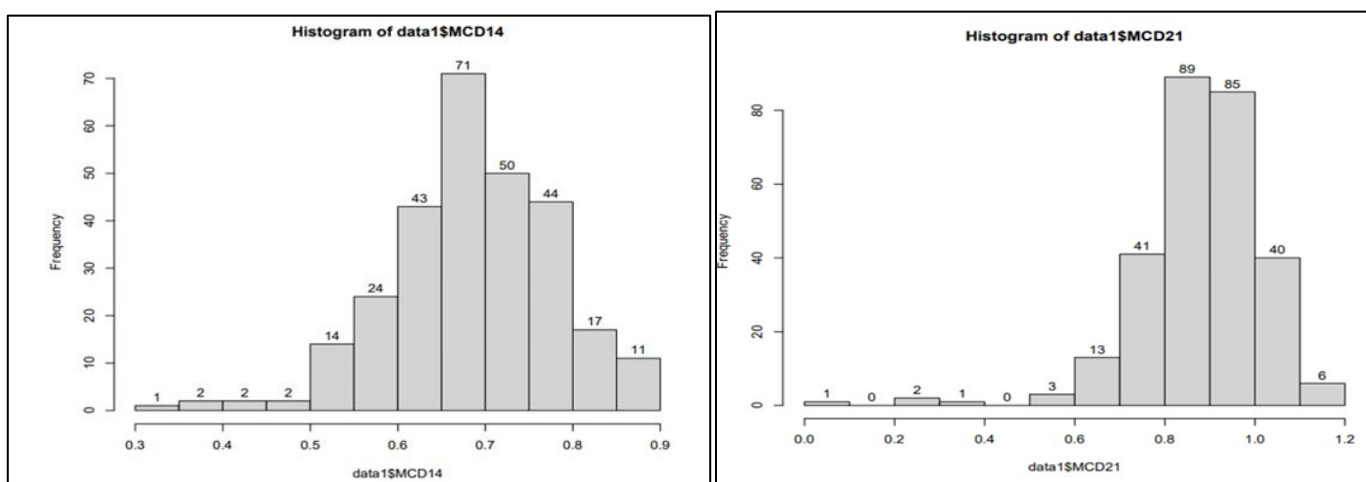
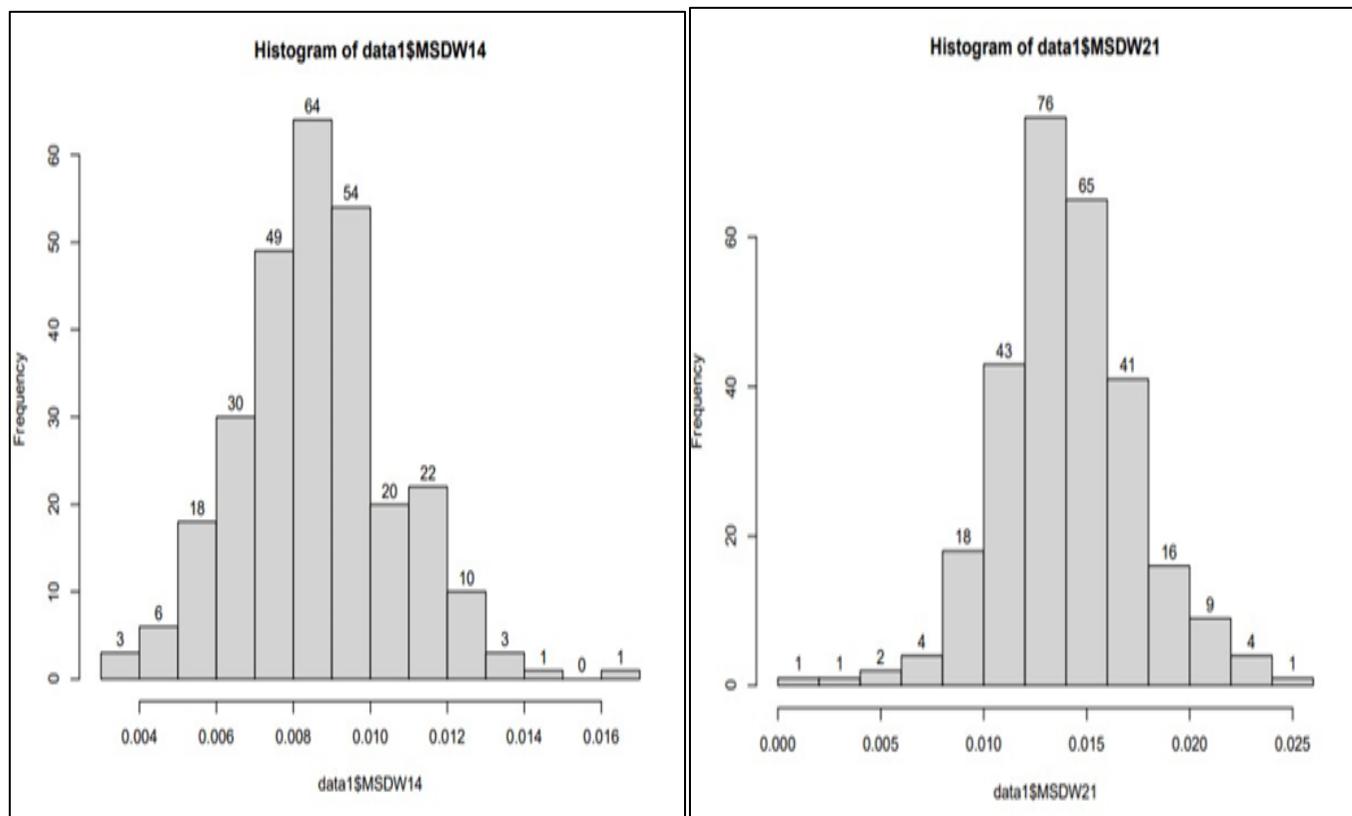
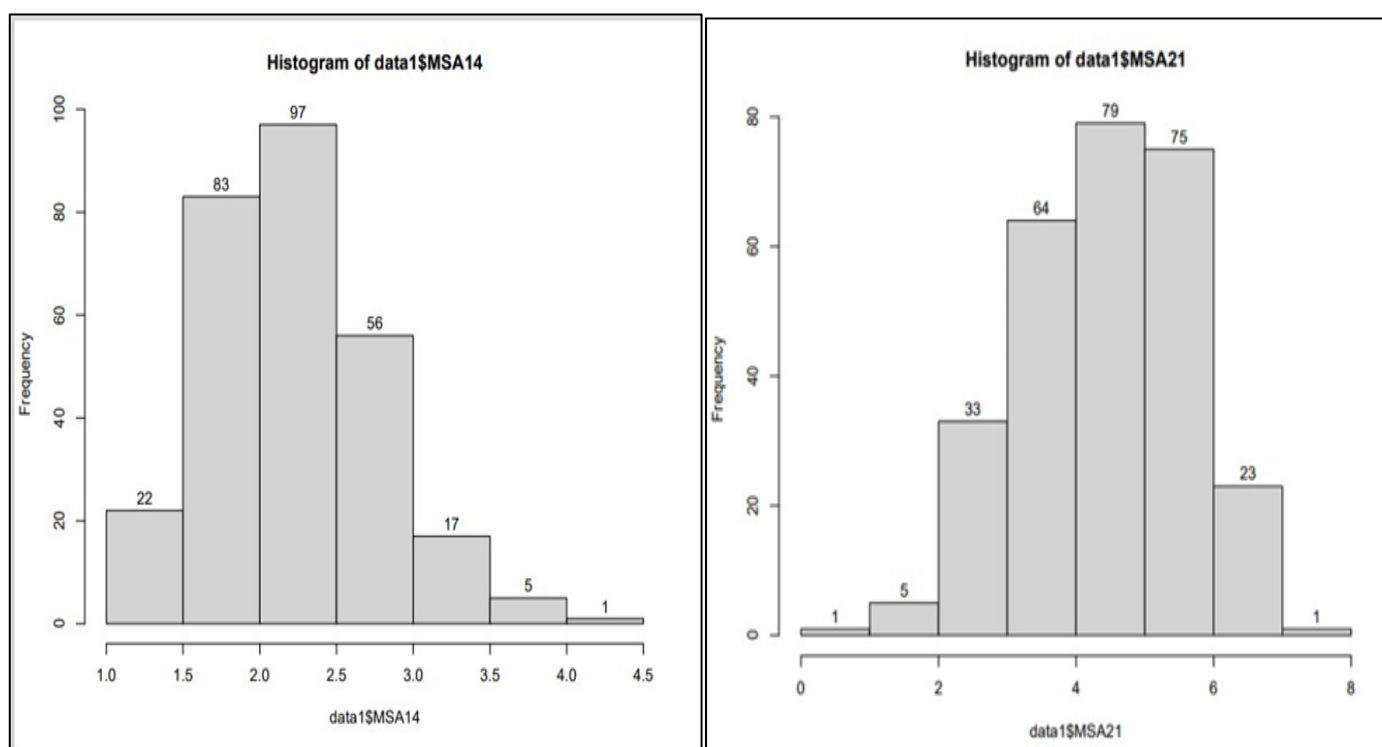


Fig 5,6,7,8,9,10: Histogram Analysis of Mean Shoot Length of 7th, 11th, 13th, 14th, 21st, 28th Day in the 281 Rice Genotypes

Fig 11,12,13: Histogram Analysis of Mean Leaves Number of 14th, 21st, 28th, Day in the 281 Rice GenotypesFig 14,15: Histogram Analysis of Mean Culm Diameter of 14th and 21st Day in the 281 Rice Genotypes

Fig 16,17: Histogram Analysis of Mean Shoot Dry Weight of 14th and 21st day in the 281 Rice GenotypesFig 18,19: Histogram Analysis of Mean Shoot Area of 14th and 21st day in the 281 Rice Genotypes

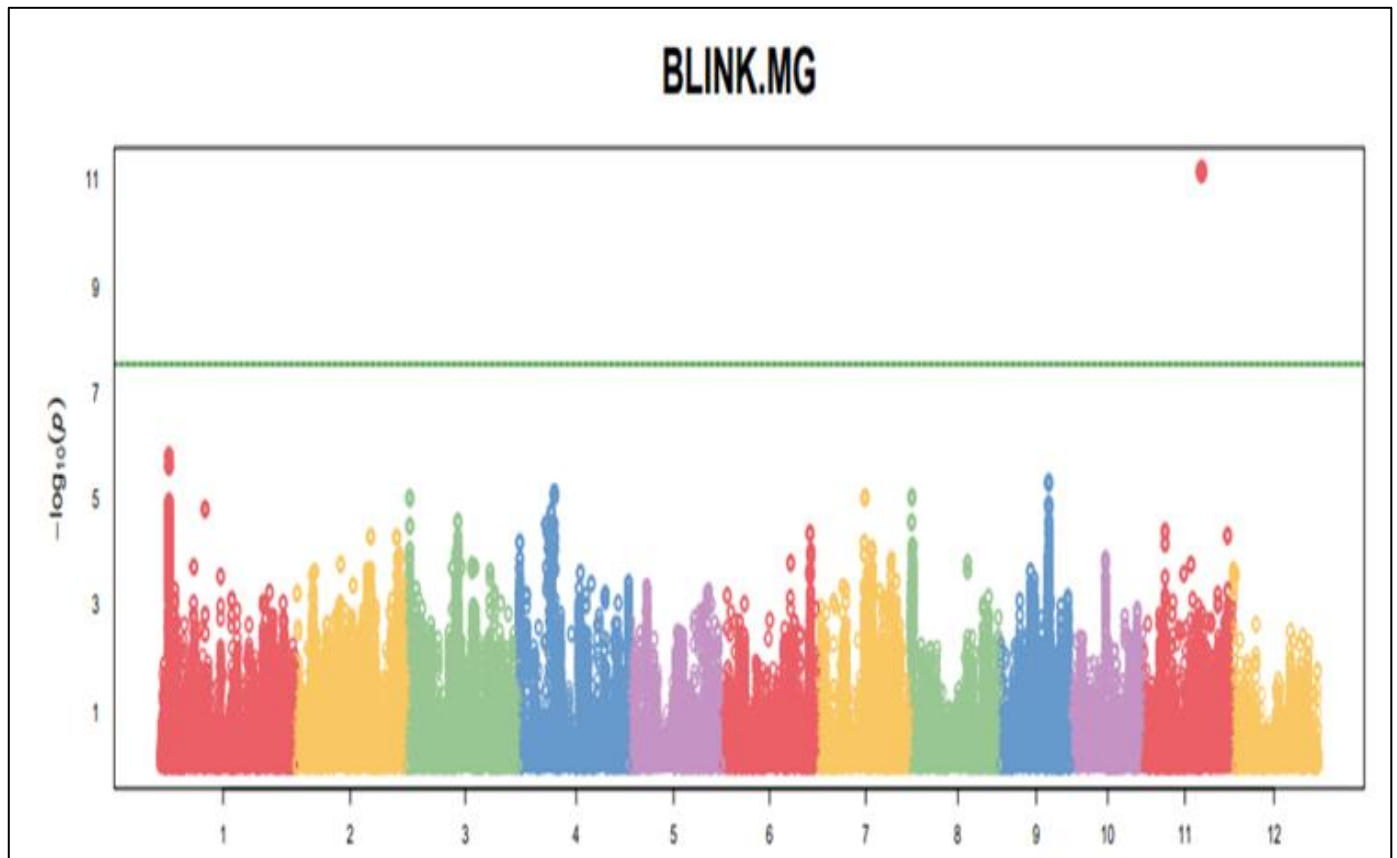


Fig 21: GWAS Analysis for Phenotypic Mean Shoot Length on 28th Day

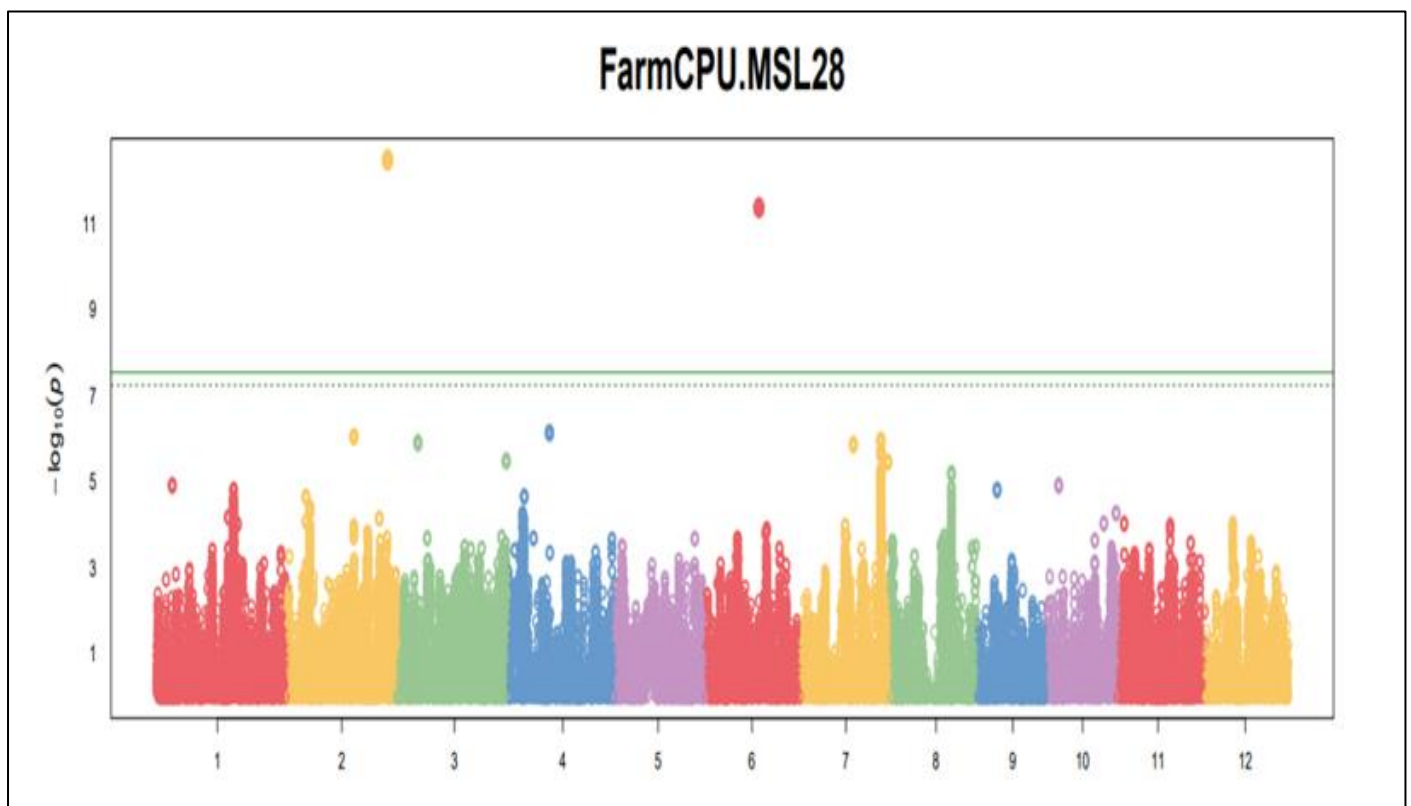


Fig 20: GWAS Analysis for Phenotypic Mean Germination

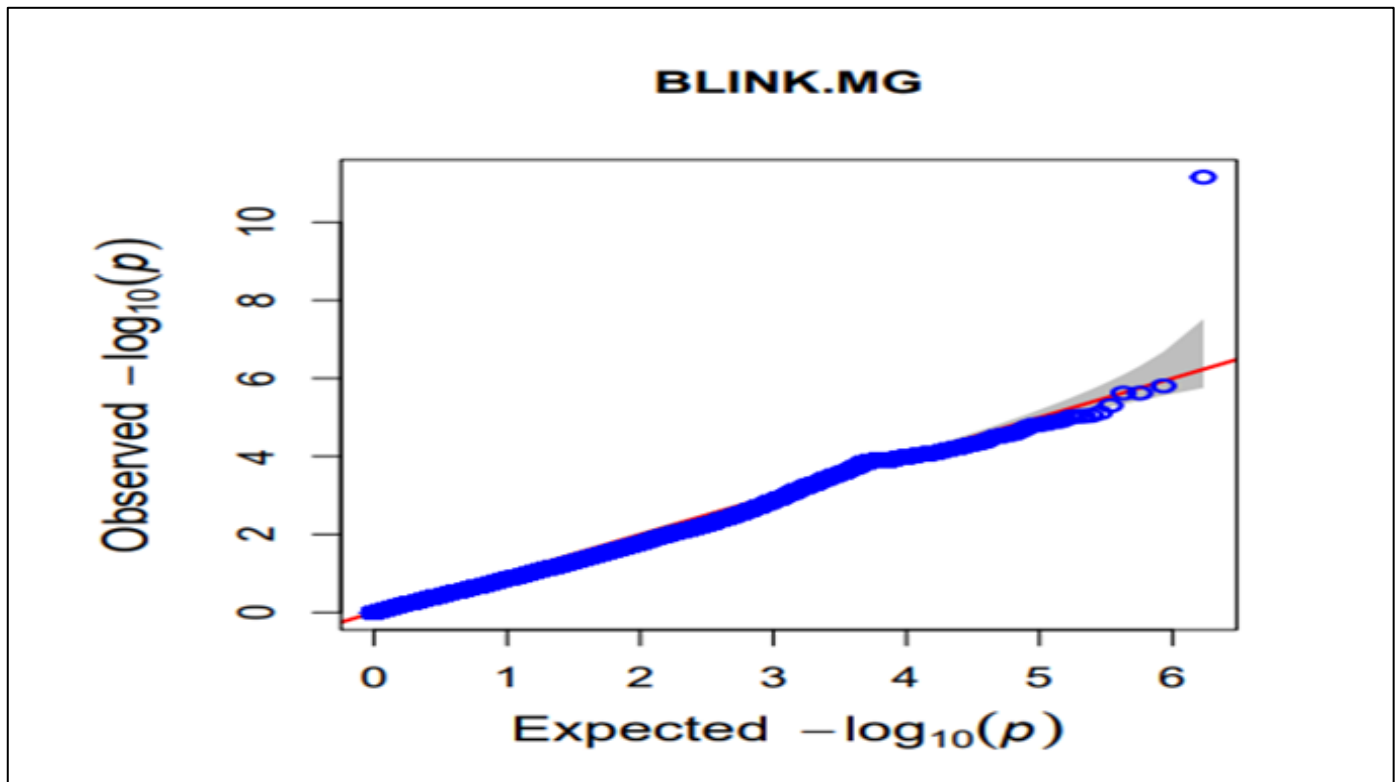


Fig 22: GWAS Analysis for Genotypic Mean Germination

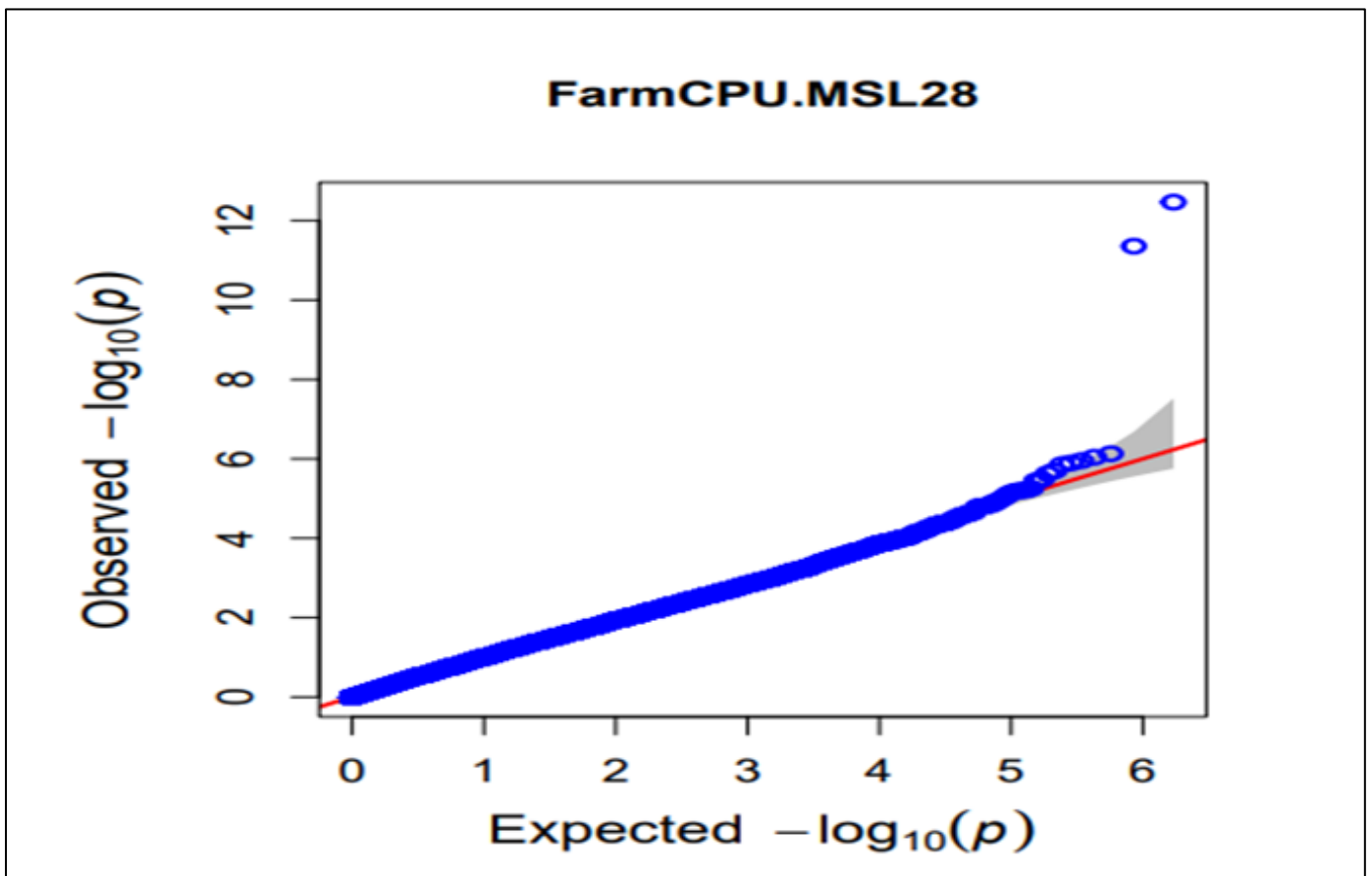


Fig 23: GWAS Analysis for Genotypic Mean Shoot Length of 28th Day

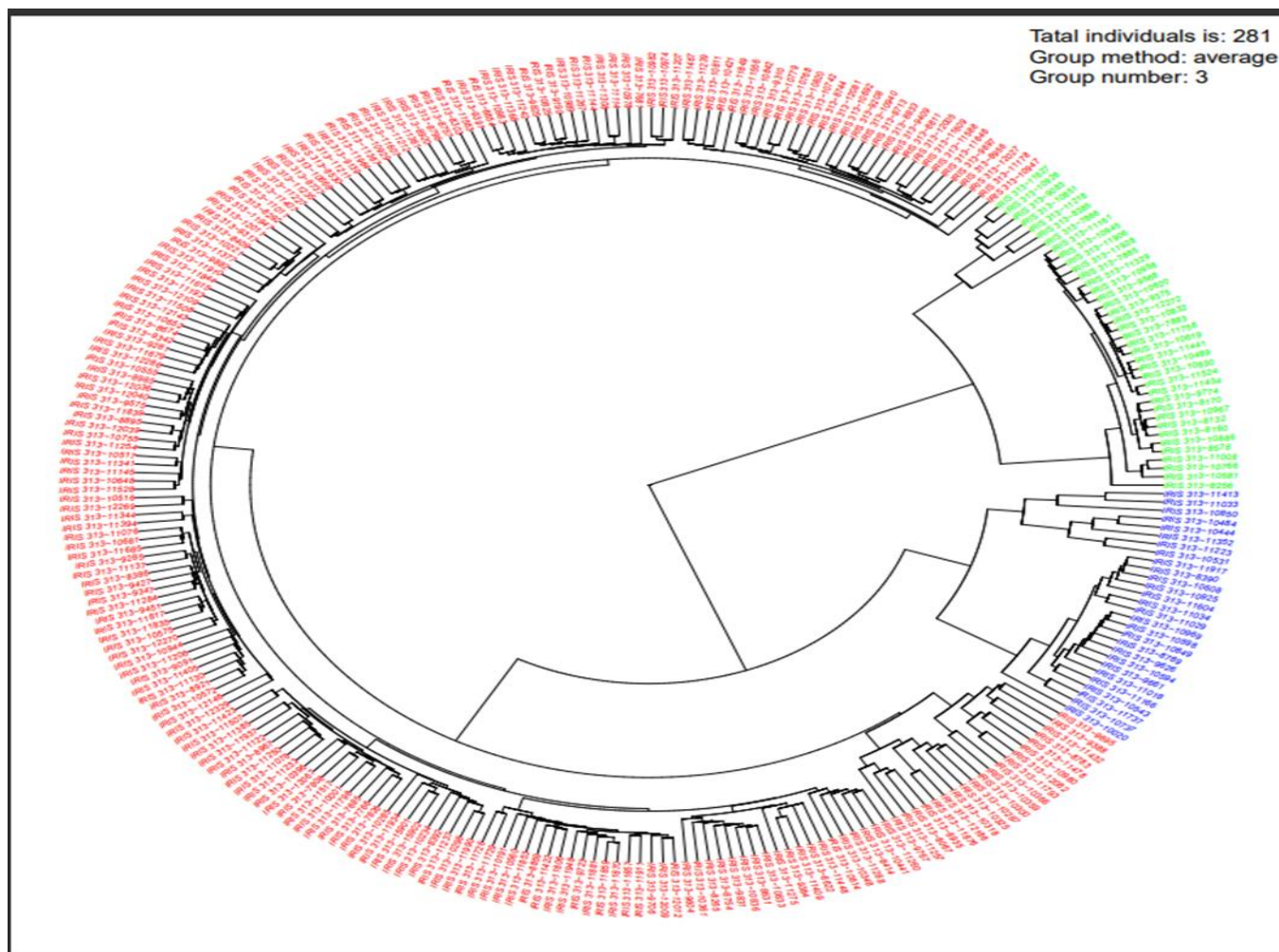
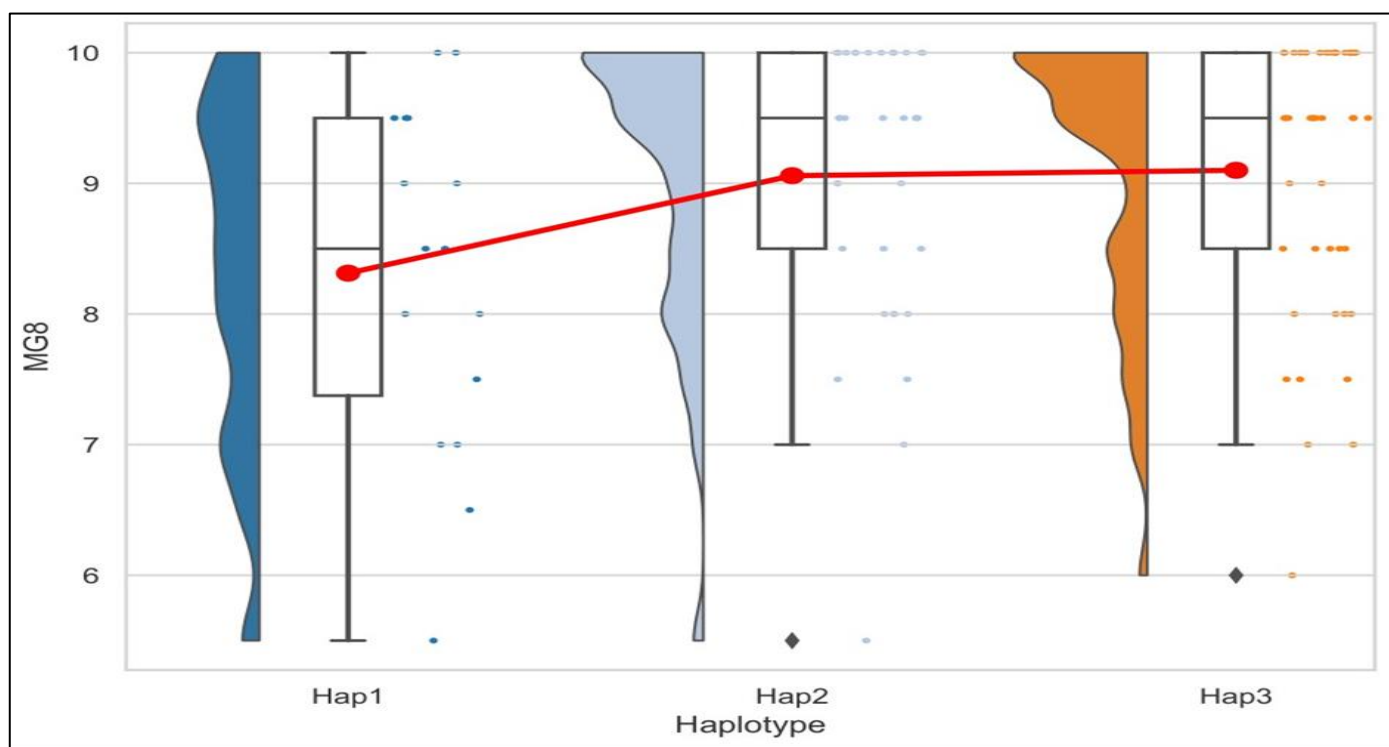


Fig 24: Phylogenetic Tree Analysis of all the Traits in 281 Rice Genotypes

Fig 25: Haplotype Analysis of Mean Germination of 8th Day of 281 Rice Genotypes

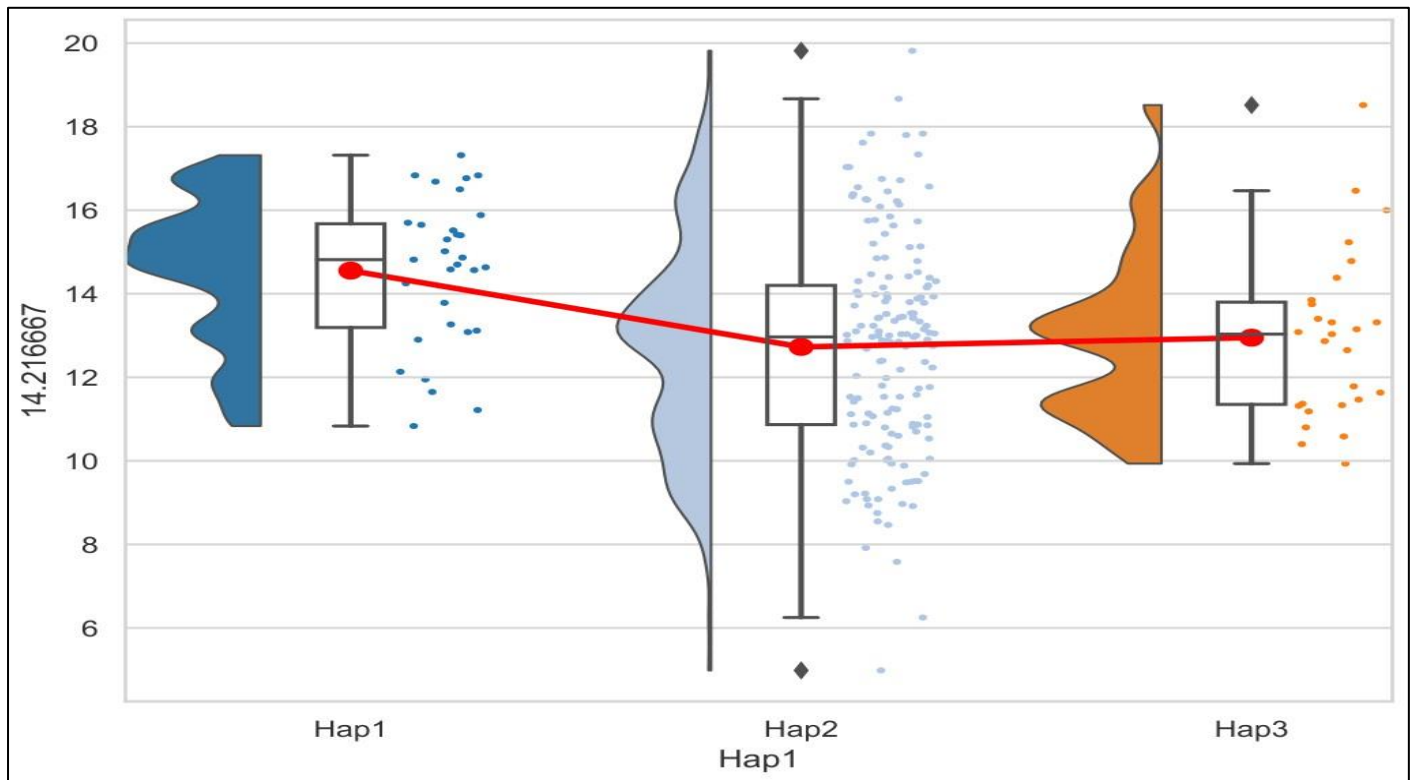
Fig 26: Haplotype analysis of Mean Germination of 8th day of 281 Rice Genotypes

Table 1: Supplementary

Sl. No.	Accession Id	Subpopulation	Country	Sl. No.	Accession Id	Subpopulation	Country
1	IRGC 126156	admix	Bangladesh	142	IRGC 128448	ind1A	China
2	IRGC 125607	aro	Bangladesh	143	IRGC 126008	ind1A	China
3	IRGC 125879	aro	Bangladesh	144	IRGC 127586	ind1A	China
4	IRGC 122024	aus	Bangladesh	145	IRGC 128338	ind1A	China
5	IRGC 128510	aus	Bangladesh	146	IRGC 128273	ind1A	China
6	IRGC 125866	aus	Bangladesh	147	IRGC 121038	ind1A	China
7	IRGC 127355	aus	Bangladesh	148	IRGC 126007	ind1B	China
8	IRGC 132327	aus	Bangladesh	149	IRGC 126173	indx	China
9	IRGC 122166	aus	Bangladesh	150	IRGC 125674	indx	China
10	IRGC 128085	ind1B	Bangladesh	151	IRGC 125933	indx	China
11	IRGC 121005	ind2	Bangladesh	152	IRGC 125608	indx	China
12	IRGC 127910	ind2	Bangladesh	153	IRGC 127101	indx	China
13	IRGC 127242	ind2	Bangladesh	154	IRGC 127226	indx	China
14	IRGC 125738	ind2	Bangladesh	155	IRGC 127761	indx	Colombia
15	IRGC 127887	ind2	Bangladesh	156	IRGC 132420	trop	Colombia
16	IRGC 127430	ind2	Bangladesh	157	IRGC 125841	ind3	Cote d'Ivoire
17	IRGC 128239	ind2	Bangladesh	158	IRGC 127156	trop	Cote d'Ivoire
18	IRGC 127786	ind2	Bangladesh	159	IRGC 117442	indx	Ecuador
19	IRGC 121036	ind2	Bangladesh	160	IRGC 127637	ind3	Guinea
20	IRGC 117474	ind2	Bangladesh	161	IRGC 126000	indx	Guyana
21	IRGC 125663	ind2	Bangladesh	162	IRGC 127888	indx	Hong Kong
22	IRGC 127241	ind2	Bangladesh	163	IRGC 127417	admix	India
23	IRGC 127030	ind2	Bangladesh	164	IRGC 125863	admix	India
24	IRGC 127572	indx	Bangladesh	165	IRGC 127225	admix	India
25	IRGC 127729	indx	Bangladesh	166	IRGC 127411	admix	India
26	IRGC 127113	indx	Bangladesh	167	IRGC 124442	admix	India
27	IRGC 125629	aus	Brazil	168	IRGC 128103	aro	India
28	IRGC 125654	temp	Brazil	169	IRGC 127227	aro	India
29	IRGC 127166	trop	Brazil	170	IRGC 132309	aro	India
30	IRGC 128297	ind3	Cambodia	171	IRGC 127830	aro	India
31	IRGC 126199	ind3	Cambodia	172	IRGC 126011	aus	India
32	IRGC 128315	ind3	Cambodia	173	IRGC 122165	aus	India
33	IRGC 117497	ind3	Cambodia	174	IRGC 126236	aus	India
34	IRGC 128258	ind3	Cambodia	175	IRGC 127835	aus	India
35	IRGC 127224	ind3	Cambodia	176	IRGC 125869	aus	India
36	IRGC 127339	ind3	Cambodia	177	IRGC 128118	aus	India
37	IRGC 127344	indx	Cambodia	178	IRGC 127832	aus	India
38	IRGC 127622	trop	Central America	179	IRGC 127966	ind1B	India

39	IRGC 127535	ind1A	China	180	IRGC 127907	ind2	India
40	IRGC 131919	ind1A	China	181	IRGC 126964	ind2	India
41	IRGC 127570	ind1A	China	182	IRGC 127100	ind2	India
42	IRGC 127485	ind1A	China	183	IRGC 126140	ind2	India
43	IRGC 125645	ind1A	China	184	IRGC 125931	ind2	India
44	IRGC 128287	ind1A	China	185	IRGC 125877	ind2	India
45	IRGC 127722	ind1A	China	186	IRGC 128059	ind2	India
46	IRGC 117454	ind1A	China	187	IRGC 127713	ind2	India
47	IRGC 127294	ind1A	China	188	IRGC 127970	ind2	India
48	IRGC 117551	ind1A	China	189	IRGC 125987	ind2	India
49	IRGC 127641	ind1A	China	190	IRGC 125872	ind2	India
50	IRGC 128094	ind1A	China	191	IRGC 127735	ind2	India
51	IRGC 127283	ind2	India	192	IRGC 127736	ind3	Indonesia
52	IRGC 127978	ind2	India	193	IRGC 128126	ind3	Indonesia
53	IRGC 128137	ind2	India	194	IRGC 125754	ind3	Indonesia
54	IRGC 121128	ind2	India	195	IRGC 127982	ind3	Indonesia
55	IRGC 132387	ind2	India	196	IRGC 127322	ind3	Indonesia
56	IRGC 122229	ind2	India	197	IRGC 128038	ind3	Indonesia
57	IRGC 125789	ind2	India	198	IRGC 128046	ind3	Indonesia
58	IRGC 127429	ind2	India	199	IRGC 128310	ind3	Indonesia
59	IRGC 126294	ind2	India	200	IRGC 125696	ind3	Indonesia
60	IRGC 127095	ind2	India	201	IRGC 125647	indx	Indonesia
61	IRGC 122285	ind2	India	202	IRGC 120968	indx	Indonesia
62	IRGC 127442	ind2	India	203	IRGC 121154	indx	Indonesia
63	IRGC 127393	ind2	India	204	IRGC 126129	indx	Indonesia
64	IRGC 127944	indx	India	205	IRGC 125801	trop	Indonesia
65	IRGC 125873	indx	India	206	IRGC 135538	trop	Indonesia
66	IRGC 122076	indx	India	207	IRGC 126966	trop	Indonesia
67	IRGC 128098	indx	India	208	IRGC 132408	trop	Indonesia
68	IRGC 127974	indx	India	209	IRGC 127630	trop	Indonesia
69	IRGC 127956	indx	India	210	IRGC 128160	trop	Indonesia
70	IRGC 132240	indx	India	211	IRGC 126153	trop	Indonesia
71	IRGC 128181	indx	India	212	IRGC 127848	japx	Iran
72	IRGC 127448	indx	India	213	IRGC 128000	ind3	Lao People's Democratic Republic
73	IRGC 128090	indx	India	214	IRGC 125970	ind3	Lao People's Democratic Republic
74	IRGC 127533	indx	India	215	IRGC 132037	ind3	Lao People's Democratic Republic
75	IRGC 126148	indx	India	216	IRGC 126085	indx	Lao People's Democratic Republic
76	IRGC 127989	indx	India	217	IRGC 127802	indx	Lao People's Democratic Republic
77	IRGC 132429	indx	India	218	IRGC 127034	indx	Lao People's Democratic Republic
78	IRGC 127365	indx	India	219	IRGC 132359	trop	Liberia
79	IRGC 126159	indx	India	220	IRGC 128055	ind2	Madagascar
80	IRGC 132305	indx	India	221	IRGC 135572	ind2	Madagascar
81	IRGC 122232	indx	India	222	IRGC 128004	ind3	Indonesia
82	IRGC 127614	indx	India	223	IRGC 122292	ind3	Indonesia
83	IRGC 128049	indx	India	224	IRGC 127975	ind3	Indonesia
84	IRGC 125913	indx	India	225	IRGC 127689	ind3	Indonesia
85	IRGC 127829	subtrop	India	226	IRGC 127268	ind3	Indonesia
86	IRGC 127745	temp	India	227	IRGC 127950	ind3	Indonesia
87	IRGC 127908	ind3	Indonesia	228	IRGC 128110	ind3	Indonesia
88	IRGC 125965	ind3	Indonesia	229	IRGC 125976	ind3	Indonesia
89	IRGC 127864	ind3	Indonesia	230	IRGC 126175	ind3	Indonesia
90	IRGC 127121	ind3	Indonesia	231	IRGC 128063	ind3	Indonesia
91	IRGC 128350	ind2	Madagascar	232	IRGC 132051	trop	Philippines
92	IRGC 128154	indx	Madagascar	233	IRGC 128121	trop	Philippines
93	IRGC 122018	trop	Madagascar	234	IRGC 122052	trop	Philippines
94	IRGC 127636	ind3	Malaysia	235	IRGC 128071	trop	Philippines
95	IRGC 127443	ind3	Malaysia	236	IRGC 127289	trop	Philippines
96	IRGC 126165	ind3	Malaysia	237	IRGC 127307	trop	Philippines
97	IRGC 125960	ind3	Malaysia	238	IRGC 127814	ind3	Sierra Leone
98	IRGC 127380	ind3	Malaysia	239	IRGC 125956	indx	Sierra Leone
99	IRGC 128128	ind3	Malaysia	240	IRGC 127438	admix	Sri Lanka
100	IRGC 127823	trop	Malaysia	241	IRGC 128186	aus	Sri Lanka
101	IRGC 128439	trop	Malaysia	242	IRGC 132310	ind1B	Sri Lanka
102	IRGC 135582	trop	Malaysia	243	IRGC 128100	ind1B	Sri Lanka
103	IRGC 125973	ind3	Myanmar	244	IRGC 125992	indx	Sri Lanka
104	IRGC 127600	ind3	Myanmar	245	IRGC 132367	indx	Sri Lanka
105	IRGC 126270	ind3	Myanmar	246	IRGC 127643	indx	Surinam

106	IRGC 125882	ind3	Myanmar	247	IRGC 127870	ind1A	Taiwan
107	IRGC 128099	indx	Myanmar	248	IRGC 131940	ind1B	Taiwan
108	IRGC 125959	indx	Myanmar	249	IRGC 132405	ind1B	Taiwan
109	IRGC 127053	indx	Myanmar	250	IRGC 131946	ind1B	Taiwan
110	IRGC 125690	indx	Myanmar	251	IRGC 128189	indx	Taiwan
111	IRGC 128057	indx	Myanmar	252	IRGC 127839	ind1B	Thailand
112	IRGC 126142	indx	Myanmar	253	IRGC 128122	ind3	Thailand
113	IRGC 127243	indx	Myanmar	254	IRGC 127370	ind3	Thailand
114	IRGC 128377	indx	Myanmar	255	IRGC 127068	ind3	Thailand
115	IRGC 128092	indx	Myanmar	256	IRGC 132364	ind3	Thailand
116	IRGC 125692	aro	Nepal	257	IRGC 128281	ind3	Thailand
117	IRGC 127119	aro	Nepal	258	IRGC 132244	ind3	Thailand
118	IRGC 122089	aus	Nepal	259	IRGC 128152	ind3	Thailand
119	IRGC 126346	aus	Nepal	260	IRGC 132238	ind3	Thailand
120	IRGC 131970	ind2	Nepal	261	IRGC 128208	ind3	Thailand
121	IRGC 125890	indx	Nigeria	262	IRGC 131962	ind3	Thailand
122	IRGC 128070	admix	Pakistan	263	IRGC 128116	indx	Philippines
123	IRGC 132300	aus	Pakistan	264	IRGC 121119	indx	Philippines
124	IRGC 126266	aus	Pakistan	265	IRGC 128521	indx	Philippines
125	IRGC 128040	aus	Pakistan	266	IRGC 128142	indx	Philippines
126	IRGC 128247	aus	Pakistan	267	IRGC 127701	indx	Philippines
127	IRGC 135638	ind1B	Peru	268	IRGC 132423	indx	Philippines
128	IRGC 127441	ind2	Peru	269	IRGC 127917	indx	Philippines
129	IRGC 132277	indx	Peru	270	IRGC 128067	indx	Philippines
130	IRGC 128399	ind1A	Philippines	271	IRGC 125711	indx	Philippines
131	IRGC 131916	ind1B	Philippines	272	IRGC 132266	trop	Philippines
132	IRGC 127594	ind1B	Philippines	273	IRGC 127371	trop	Philippines
133	IRGC 128074	ind1B	Philippines	274	IRGC 127209	trop	Philippines
134	IRGC 128157	ind1B	Philippines	275	IRGC 122063	ind3	Philippines
135	IRGC 126170	ind1B	Philippines	276	IRGC 128372	ind3	Philippines
136	IRGC 132395	ind1B	Philippines	277	IRGC 127741	ind3	Philippines
137	IRGC 127325	ind1B	Philippines	278	IRGC 132269	indx	Philippines
138	IRGC 127587	ind1B	Philippines	279	IRGC 127812	indx	Philippines
139	IRGC 128183	ind3	Philippines	280	IRGC 128450	indx	Philippines
140	IRGC 128443	ind3	Philippines	281	IRGC 132285	indx	Philippines
141	IRGC 131937	ind3	Philippines				