Indian Journal of Experimental Biology
Vol. 59, May 2021, pp. 328-338

# Population structure and association studies for reproductive stage salinity tolerance in rice (*Oryza sativa* L.)

Warraich AS[1,2], Krishnamurthy SL[1]*, Sooch BS[2], Lokeshkumar BM[1], Vinaykumar NM[3], Dushyanthkumar BM[4], Rai V[5], Singh NK[5] & Sharma PC[1]*

[1]Indian Council of Agriculture Research- Central Soil Salinity Research Institute, Karnal-132 001, Haryana, India
[2]Department of Biotechnology, Faculty of Life Science, Punjabi University, Patiala-147 001, Punjab, India
[3]Department of Biotechnology, Kuvempu University, Shankaraghatta, Shimvaogga–577 451, Karnataka, India
[4]Department of Genetics and Plant Breeding, University of Agricultural and Horticultural Science, Shimvaogga-577 204, Karnataka, India
[5]ICAR-National Institute for Plant Biotechnology, New Delhi-110 012, India

*Received 07 October 2019*; *revised 24 July 2020*

Salinity is a major abiotic stress responsible for yield loss in rice as it severely affects various yield contributing traits. Rice is categorised as salt sensitive crop and it is important to identify genomic regions associated to salinity tolerance. In the present study, association mapping was performed to investigate the functional relationship between microsatellite markers and salinity related traits in a set of 180 diverse rice accessions. Association analysis was carried out by employing mixed linear model (MLM) approach. Population structure analysis revealed four subgroups in entire study panel while the admixture level ranged from 0.7-57.2%. A total of 22 marker-trait associations were discovered and four marker-trait associations explained phenotypic variation ($R^2$) greater than 10%. Furthermore, 7 markers were found close to the candidate genes loci. Several markers were significantly associated with more than one trait, suggesting pleiotropic effects. The phenotypic variation explained by associated markers ranged from 2.92 to 18.50%. Comparative genomic search revealed that associated markers were close to candidate genes which play significant role in signal transduction, metabolic pathways and transcription factor activity. The significant associations identified in the present study could be used to improve salt tolerance in rice with introgression of favourable alleles through marker assisted breeding.

**Keywords:** Abiotic stress, Association mapping, GWAS, Quantitative trait loci

Rice is a cereal grain widely consumed as a staple food. It is a main component of food security programs in many developing nations which includes rice producing countries of Asia, Africa and South America. More than 90% rice production comes from Asia because of its adaptablity across different climatic conditons[1]. However, production of rice is restricted by many abiotic stresses and soil salinity is one major abiotic stress that affects its yield significantly. Soil salinity is the second most wide spread problem after drought and more than 6% land area of the world is salt affected[2]. Rice is sensitive to salinity having threshold E.C. (electrical conductivity) of 3 dS $m^{-1}$ (decisiemens per meter) and slop mean of 12[3]. Varied impact on rice crop can be obseved by different concentration of salt and duration of exposure[4]. Soil salinity hindres plant growth at both vegetative and reproductive stage[5]. Reproductive stage is the important stage in rice life cycle since it determines the crop yield. Effects of salinity on rice yield and yield contributing traits are manifested in the form of increased spikelet sterility, less florets/panicle, low seed weight and stunted growth of plant[6,7]. Like most complex traits in higher organisms, salinity tolerance in rice is also a multifactorial trait governed by multiple genes. Many quantitative trait loci (QTL) studies have been conducted in the past to discover these regions in rice genome and few identified QTLs were successfully introgressed into sensitive genotypes through marker assisted breeding[8,9]. Extensive mapping studies has been conducted at vegetative stage in the past however, studies at reproductive stage are scarce. Therefore, dissection of salt tolerance governing genes at reproductive becomes imperative to completely understand the mechanism of salt tolerance at reproductive phase of rice.

*Correspondence:
E-Mail: krishnagene@gmail.com (SLK);
pcsharma.knl@gmail.com (PCS)

Mapping studies by linkage analysis using $F_2$, recombinant inbreed lines (RILs) and backcross populations are used to identify genomic regions associated with salt tolerance. However, linkage analysis has its own shortfalls. Mapping population used in these studies are allele specific due to limited number of parents used for mapping. Moreover, linkage analysis can only localizes QTLs to 10-20cM region due to limited crossing over event during the development of these mapping populations which is not sufficient to completely cover entire genome[10]. On the other hand, association mapping is a suitable replacement of linkage analysis as it offers high mapping resolution, high allelic variation and it is more time efficient. Several association studies were conducted for salinity tolerance in rice at seedling stage in past few years and these studies has lead to identification of QTLs which were controlling different morphological and physiological traits[11-13]. Nonetheless, only one genome wide association study (GWAS) was conducted to identify QTLs for reproductive stage salinity tolerance using 6000 single nucleotide polymorphism (SNPs)[14]. Apart from this, no GWAS reported marker-trait association for reproductive stage salinity tolerance. Considering the importance of reproductive stage salinity tolerance, here, we have evaluated rice lines in both control and saline environment for identifying potential marker-trait associations under saline environment.

## Materials and Methods

### Plant material

The present study panel consisted of 180 diverse rice accessions (Suppl. Table 1. *All supplementary data are available only online along with the respective paper at NOPR repository at http://nopr.res.in*) which included landraces, agronomically important varieties and salt tolerant varieties with differential response to salinity. To summarize the entire panel, there were 48 landraces collected from all over India, 15 elite lines from International Rice Research Institute (IRRI), Philippines, 117 elite lines/varieties from ICAR Institutes and State Agriculture Universities. The set had a mixture of highly sensitive to highly tolerant rice lines which included salt sensitive basmati lines, high yielding lines (non-basmati type) like Jaya, Moti, Pusa 44 and varieties like CSR10, CSR13, CSR 27, CSR 30, CSR 36, CSR 43, CST 7-1, Amalmana, which are released for salt affected area. Pokalli, Trichy, Nona Bokra, Kalanamak which are traditional landraces and highly adapted to saline conditions were also included in the study.

### Phenotypic screening

Rice accessions were evaluated in microplots (rainout shelter) at ICAR- Central Soil Salinity Research Institute (CSSRI) Karnal, India during wet seasons of 2015 and 2016. Rice panel of 180 rice accessions were evaluated to determine the number of days to reach the 50 % flowering stage during 2014. Finally, for subsequent wet season of 2015 and 2016, the rice accessions were transplanted as per their days 50% flowering in microplots and were evaluated for different phenotypic characters. Thirty day old rice seedlings, uprooted from the nursery beds were transplanted in microplots in two environments i.e. saline ($EC_{iw}$ ~ 10.0 dS $m^{-1}$) and normal ($EC_{iw}$ ~ 1.0 dS $m^{-1}$). The experiment was designed in randomized complete block design with two replications. Two seedlings per hill were transplanted and raised as per recommended agronomic package of practices. Salinity was imposed at maximum tillering stage with three salt combinations (7NaCl: $1Na_2SO_4$: $2CaCl_2$). The irrigation water was regularly monitored for its conductivity before discharging it into the microplots. The saline irrigation was continued till the plants completed their physiological maturity. Normal water ($EC_{iw}$~1.0 dS $m^{-1}$) was used for irrigation of control plots. The traits were recorded and sampled at appropriate time and stage of the crop.

### Evaluation of morphological traits

Ten plants from the middle of the row for each genotype were randomly tagged and used for recording phenotypic data. A total of ten traits were measured, namely, days to 50% flowering, plant height (cm), panicle length (cm), total tillers/plant, productive tillers/plant, 1000 grain weight (g), spikelet fertility (%), biomass (g), grain yield/plant (g) and vigor (salt injury) score. The vigor score was assigned to rice genotypes based on visual symptoms of salinity. The recorded traits were averaged from 10 plants during two seasons for further computational analysis.

### Genotyping

Total genomic DNA (Deoxyribonucleic acid) of 180 rice accessions was isolated from the leaf tissue of rice seedlings. Leaf tissues (200-250 mg) were grinded in mortar pestle using liquid nitrogen to form dry powder which was subsequently used in DNA

extraction by modified CTAB (cetyltrimethylammonium bromide) method[15]. For PCR amplification, extracted DNA was diluted using molecular grade water to 30-50 ng/μL. A total of 94 simple sequence repeats (SSR) positioned on all 12 chromosomes in rice (Suppl. Table 2) were used for genotyping. The marker's primers were downloaded from Gramene database (http://www.gramene.org/) and few of the markers were designed using Batch Primer 3[16]. Primers were synthesized by Sigma-Aldrich Corporation, Bengaluru, India. PCR (polymerase chain reaction) reactions were carried out on a Biometra T Gradient Thermocycler (Imperial Life Science (P) Limited, Gurgaon, India) with a final PCR reaction mix containing 50 ng template DNA, 1.8 μL 10 × Taq polymerase assay buffer (with 16 mM $MgCl_2$), 1.0 μL of dNTPs, 0.5 μL of forward and reverse primer and 1 U TaqDNA polymerase (HiMedia, India). The profile for PCR running program was set as follows: Initial denaturation for 5 min at 94°C before actual cycles begins. Each cycle begins with denaturation for 30 s at 94°C, annealing for 30 s at 55°C, extension for one minute at 72°C and after completion of 35$^{th}$ cycles final extension for 7 min at 72°C. The final PCR amplified product was run on 3% agarose gel stained with DNA intercalating dye ethidium bromide along with DNA ladder in 1 × TBE buffer. Gels were scanned with a gel documentation system (Alpha-Imager private limited, Bangalore, India) and polymorphic bands were scored.

**Population structure and association analysis**

To understand the genetic composition and estimation for total number of subgroups in the study panel Structure 2.3.4 program was used. The markers were filtered using Tassel 5.0 using filter function with site minimum count of 150, minimum allelic frequency 0.05 and site maximum allelic frequency of 1.0. Structure was run using admixture and allele frequencies correlated model with typical burn-in period of 10,000 and 100,000 MCMC (Monte Carlo Markov Chain) replications after burn-in period[17]. To estimate the correct value of K from log probability of data [LnP(D)], STRUCTURE output was joined with ad hoc statistic ΔK. ΔK was calculated using second order rate of change in log probability [LnP (D)] between consecutive K values. STRUCTURE batch run was performed for K (2-10) for 20 iterations and the log probability data (LnP (D)) generated in structure output of each iteration is averaged for their respective K value and ad hoc statistics (ΔK) was computed. Finally, Delta ΔK was plotted against K[18,19].

The data of each trait was averaged for two years and descriptive statistics, correlations among various traits were summarized using SPSS statistical software package (version16.0) and MS Windows Excel. Association analysis was performed using agronomic traits data and genotypic data for 94 SSR markers. TASSEL 5.2.30 (Trait Analysis by Association, Evolution and Linkage) with MLM (Mixed Linear Model) (Q+K) approach was employed during the analysis[20]. Kinship file (K) was derived from TASSEL to assess genetic resemblances between individuals to avoid false associations. Robust markers trait associations were recognized based on significant P value ($P$ <0.05) and phenotypic variance ($R^2$) value.

## Results

**Genetic variation under saline environment**

Significant genotypic variations in agronomic traits of 180 rice accessions were found under saline conditions (Table 1). Descriptive statistics over pooled mean values of both years showed consistency between the experiments and mixed response was found among genotypes under saline conditions on the basis of salt injury score (Fig. 1). Forty seven genotypes that constitute 26.1% of our rice accessions were high to moderately salt tolerant. Furthermore,

Table 1 — Analysis of variance for yield and its contributing traits under saline stresses for two seasons in 180 rice genotypes

| Source of variation | df | Sum of Squares | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Biomass (g) | Days to flowering | Panicle length (cm) | Plant height (cm) | Productive tillers | 1000 Seed wt. (g) | Spikelet fertility (%) | Total tillers | Yield (g) | Vigor Score |
| Genotypes | 179 | 98754$^{**}$ | 844824$^{**}$ | 7231$^{**}$ | 546788$^{**}$ | 2911$^{**}$ | 15024$^{**}$ | 164625$^{**}$ | 3280$^{**}$ | 597$^{**}$ | 1054$^{**}$ |
| Replications | 1 | 90$^{**}$ | 18$^{ns}$ | 4$^{ns}$ | 16$^{ns}$ | 0.01$^{ns}$ | 148$^{**}$ | 87$^{**}$ | 4$^{ns}$ | 0.004$^{ns}$ | 0.02$^{ns}$ |
| Stresses | 1 | 16824$^{**}$ | 5832$^{**}$ | 19411$^{**}$ | 484585$^{**}$ | 1421$^{**}$ | 9188$^{**}$ | 378066$^{**}$ | 883$^{**}$ | 13722$^{**}$ | 11346$^{**}$ |
| Genotypes*Stresses | 179 | 17324$^{**}$ | 8063$^{**}$ | 6413$^{**}$ | 235334$^{**}$ | 681$^{**}$ | 6473$^{**}$ | 111968$^{**}$ | 477$^{**}$ | 2902$^{**}$ | 1054$^{**}$ |
| Genotypes*Replications | 179 | 487$^{ns}$ | 981$^{ns}$ | 243$^{ns}$ | 822$^{ns}$ | 172$^{ns}$ | 5489$^{**}$ | 814$^{ns}$ | 168$^{ns}$ | 319$^{ns}$ | 61$^{ns}$ |
| Seasons | 1 | 6$^{ns}$ | 146$^{**}$ | 140$^{**}$ | 349$^{**}$ | 10$^{**}$ | 2$^{ns}$ | 196$^{**}$ | 1$^{ns}$ | 67$^{**}$ | 4$^{**}$ |
| Genotypes*Seasons | 179 | 539$^{ns}$ | 1278$^{ns}$ | 465$^{**}$ | 1766$^{**}$ | 149$^{ns}$ | 1367$^{ns}$ | 2344$^{**}$ | 171$^{ns}$ | 147$^{ns}$ | 58$^{ns}$ |
| Stresses*Seasons | 1 | 210$^{**}$ | 0.37$^{ns}$ | 12$^{**}$ | 374$^{**}$ | 31$^{**}$ | 283$^{**}$ | 1$^{ns}$ | 0.42$^{ns}$ | 36$^{**}$ | 4$^{**}$ |

[**&* = Significant at 0.01 & 0.05 level of probability, respectively and ns, non significant]

mean value of each trait under saline condition has reduced significantly in comparison to normal condition as expected (Table 2). Under saline conditions yield per plant (g) ranged from 0.10 (Kempu Jaddu) to 10.07 (CSR 50), with mean yield of 3.42. The mean reduction of yield under saline condition over control was 62.90% which was highest among other recorded traits. The upper limit of spikelet fertility (%) was 91.81 (CSR 50) and the lower limit was 9.78 (PB 3) with mean of 52.45. The mean reduction of spikelet fertility under saline condition over control was 38.18%. Total tillers ranged from 9.67 (CSR 16) to 2.30 (Kempu doddagidda) with mean of 5.18 and mean reduction of 23.21%. Productive tiller was recorded maximum in CSR 37 (8.6) and lowest in Puttabatta (0.10). The mean productive tiller of genotypes was 3.15 and it was reduced by 38.35% when compared to the mean value under controlled conditions. Plant height (cm) was recorded highest in Pokalli (166.30), while least in Balaji (45.50), with overall mean of 85.12, a mean reduction of 30.11% was recorded under saline condition. The upper limit of biomass (g) was 58.13 (Iputtu) and lower limit was 3.58 (IR 64) with mean of 22.21 g. The mean reduction of biomass under saline conditions was 22.34% over control. CSR 56 had longest panicle of 24 cm, while, Holesaalechipiga (6.37cm) recorded shortest panicle length under salinity stress. The mean panicle length was 17.96cm which was reduced by 29% when compared to control. The 1000 seed weight (g) ranged from 6.30 (Jerrige Samba) to 27.7 (Maranellu), with mean of 18.1 g and mean reduction of 23.30% was recorded in the analysis. Correlation analysis was conducted to understand the effects of traits on each other under saline stress (Table 3). The data clearly represented the interdependence of traits on one another. Yield contributing traits such as total tillers (0.311), spikelet fertility (0.550), 1000 seed weight (0.408), productive tillers (0.662) and panicle length (0.604) showed high positive ($P \leq 0.001$) correlation with grain yield. Vigor (salt injury) score on the other hand was negatively correlated (-0.799) with yield.

**Genetic structure of rice accessions**

Three hundred fifty seven alleles were found during marker analysis of 180 rice accessions (Suppl. Table 2). The distribution of $\Delta K$ clearly indicated that there are 4 subgroups in our study panel (Fig. 2). These four subgroups are subgroup 1, subgroup 2, subgroup 3 and subgroup 4. These subgroups were comprised of 14, 105, 41 and 20 genotypes,

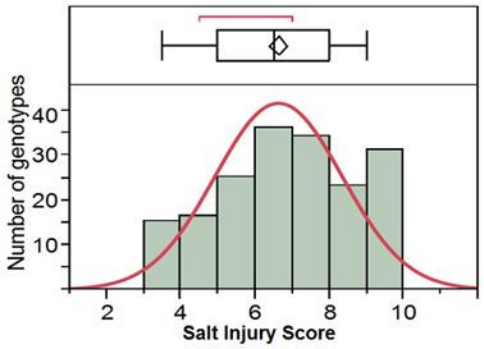


Fig. 1 — Frequency distribution of rice accessions based on vigor (salt injury) score under saline conditions

Table 2 — Descriptive statistics for comparison of mean and percent change in morphological characters under saline condition

| Traits | Mean (control) | Mean (saline) | Percent reduction over control |
|---|---|---|---|
| Biomass (g) | 28.60 | 22.21 | 22.34 |
| Days to flowering (days) | 110.46 | 114.19 | -3.36 |
| Panicle length (cm) | 25.30 | 17.96 | 29.01 |
| Plant height (cm) | 121.80 | 85.12 | 30.11 |
| Productive tillers | 5.11 | 3.15 | 38.35 |
| 1000-Seed weight (g) | 23.6 | 18.1 | 23.30 |
| Spikelet fertility (%) | 84.85 | 52.45 | 38.18 |
| Total tillers | 6.74 | 5.18 | 23.21 |
| Yield (g) | 9.22 | 3.42 | 62.90 |

Table 3 — Correlation co-efficient among different morphological traits under salinity stress (EC ~ 10 dS $m^{-1}$)

| Trait | 50% flowering (days) | Panicle length | Plant height | Productive tillers | 1000 Seed wt. | Spikelet fertility | Total tillers | Vigor score | Yield |
|---|---|---|---|---|---|---|---|---|---|
| Biomass | 0.400** | 0.171* | 0.549** | 0.109 | –0.072 | 0.054 | 0.088 | –0.202** | 0.248** |
| Days-50% flowering | 1 | –0.424** | 0.290 | –0.477** | –0.520** | –0.236** | –0.266** | 0.380** | –0.423** |
| Panicle length | | 1 | 0.366** | 0.582** | 0.297** | 0.368** | 0.096 | –0.715** | 0.604** |
| Plant height | | | 1 | 0.137 | –0.081 | 0.193** | –0.074 | –0.388** | 0.266** |
| Productive Tillers | | | | 1 | 0.303** | 0.295** | 0.657** | –0.704** | 0.662** |
| 1000 Seed weight | | | | | 1 | 0.189* | 0.097 | –0.322** | 0.408** |
| Spikelet Fertility | | | | | | 1 | 0.048 | –0.550** | 0.550** |
| Total tillers | | | | | | | 1 | –0.234** | 0.311** |
| Vigor score | | | | | | | | 1 | –0.799** |

[**&* = Significant at 0.01 & 0.05 level of probability, respectively]

respectively (Fig. 3). The mean performance of each subgroup was diverse with significant phenotypic variation (Table 4). The mean performance of genotypes under subgroup 2 was the best among other subgroups in terms of yield related traits. It had the highest mean productive tillers (3.74), yield (4.29 g), panicle length (19.34 cm) and spikelet fertility (56.22%), furthermore it comprised genotypes of high salt tolerance capacity with the lowest mean salt injury score (6). The subgroup 3 had mostly landraces with maximum plant height (91.60 cm), days to 50% flowering (143.22 days) and biomass per plant (25.71 g). Subgroup 4 comprised basmati type genotypes which were highly sensitive to salinity stress with highest mean vigor score of 7.90, lowest yield per plant (1.92 gm) and spikelet fertility (42.43%). Meanwhile, subgroup 1 consisted of mixed origin genotypes. Among 180 rice accessions, 90 (50%) genotypes has admixture level of less than 10%, 19 (10.5%) genotypes with admixture level of 10-20% and 71(39.4%) genotypes with admixture level of more than 20% (Fig. 3). Minimum admixture level of 0.4% and maximum admixture level of 57.2% was detected during the analysis.

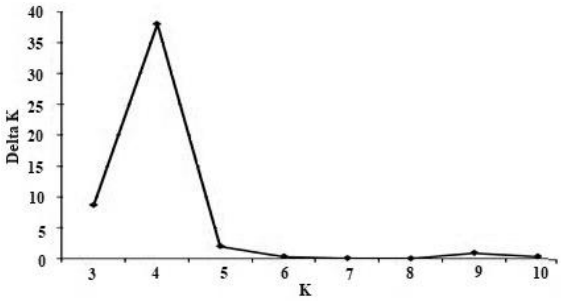


Fig. 2 — Estimation for optimal value of K using adhoc statistcs

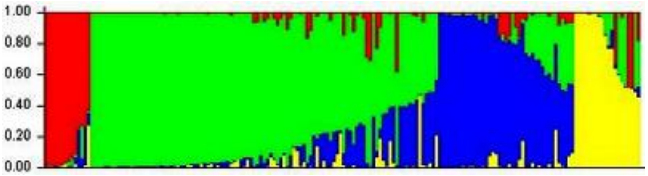


Fig. 3 — Bayesian clustering analysis of rice genotypes using STRUCTURE 2.3.4

**Marker trait associations**

The 180 rice accessions were evaluated for 10 agro-morphological traits and allelic information of 94 SSR markers was used to identify robust marker-trait associations. A total of 22 marker-trait associations were found to be significant at $P$ <0.05 using MLM approach under saline environment (Table 5). In our

Table 4 — Mean performance of 4 subgroups under saline stress for different morphological traits

| Subgroups | Plant height (cm) | Panicle length (cm) | 50% flowering (days) | Total tillers | Productive tillers | Biomass (g) | Vigor score | Yield (g) | 1000 Seed wt. (g) | Spikelet fertility (%) |
|---|---|---|---|---|---|---|---|---|---|---|
| Subgroup 1 | 89.01 | 15.27 | 126.32 | 5.74 | 2.79 | 21.35 | 7.32 | 2.31 | 14.6 | 51.17 |
| Subgroup 2 | 83.93 | 19.24 | 103.48 | 5.39 | 3.74 | 21.53 | 6 | 4.29 | 18.7 | 56.22 |
| Subgroup 3 | 91.6 | 16.32 | 143.22 | 4.35 | 2.1 | 25.71 | 7.3 | 2.43 | 16.7 | 48.45 |
| Subgroup 4 | 76.12 | 16.23 | 103.53 | 5.37 | 2.39 | 19.53 | 7.9 | 1.92 | 20 | 42.43 |

Table 5 — List of markers trait associations under saline condition using MLM model (Q+K)

| Traits (saline) | Marker | Marker position | Chromo-some | P- value | $R^2$- value (%) | Adjacent/ Overlapping genes |
|---|---|---|---|---|---|---|
| Biomass | OsCML8 | 11198205 | 104 | 9.05E-06 | 18.50 | - |
| | RM518 | 20130135 | | 0.0017 | 10.41 | *Os04g0132400*, receptor-like cytoplasmic kinase 127 |
| Days to 50% | RM432 | 18958597 | 73 | 0.00127 | 8.96 | *Os07g0502200*, *Os07g0502550* - |
| flowering | RM7389 | 36155173 | | 0.01175 | 5.60 | |
| Panicle | RM1261Os | 17531111 | 1210 | 0.00563 | 9.38 | - |
| length | CML8 | 11198205 | | 0.1284 | 6.57 | - |
| Plant height | RM6283 | 16968876 | 310 | 0.02837 | 2.92 | HAP2 subunit of CCAAT box binding complex |
| | OsCML8 | 11198205 | | 0.03893 | 5.36 | - |
| Productive | RM340 | 28599181 | 665 | 0.00104 | 10.73 | - |
| tillers | RM3827 | 22297146 | | 0.00148 | 11.40 | - |
| | RM6320 | 471605 | | 0.00506 | 8.62 | With no lysine kinase 2 (WNK2), *Os05g0108350* |
| 1000-Seed | RM201 | 20174289 | 99 | 0.01127 | 6.35 | - |
| weight | RM7175 | 16871728 | | 0.02232 | 6.41 | - |
| Spikelet | RM434 | 15662573 | 94 | 0.00234 | 9.86 | -*Os04g0102500* |
| fertility | RM551 | 177080 | | 0.00551 | 9.77 | |
| Total tillers | RM11 | 19256914 | 75 | 0.00731 | 7.05 | - |
| | RM421 | 23976333 | | 0.01024 | 5.34 | *Os05g0489000*, (putative inositol-1, 4, 5-trisphosphate 5-phosphatase) |
| Grain yield | RM80 | 24478642 | 85 | 0.00681 | 7.83 | - |
| | RM413 | 2212736 | | 0.02443 | 7.34 | - |
| Vigor (salt | RM504 | 28129335 | 345 | 0.0092 | 5.31 | *Os03g0701000* |
| injury) score | RM3310 | 34545624 | | 0.02574 | 5.41 | - |
| | RM6320 | 471605 | | 0.03148 | 6.09 | With no lysine kinase 2 (WNK2), *Os05g0108350* |

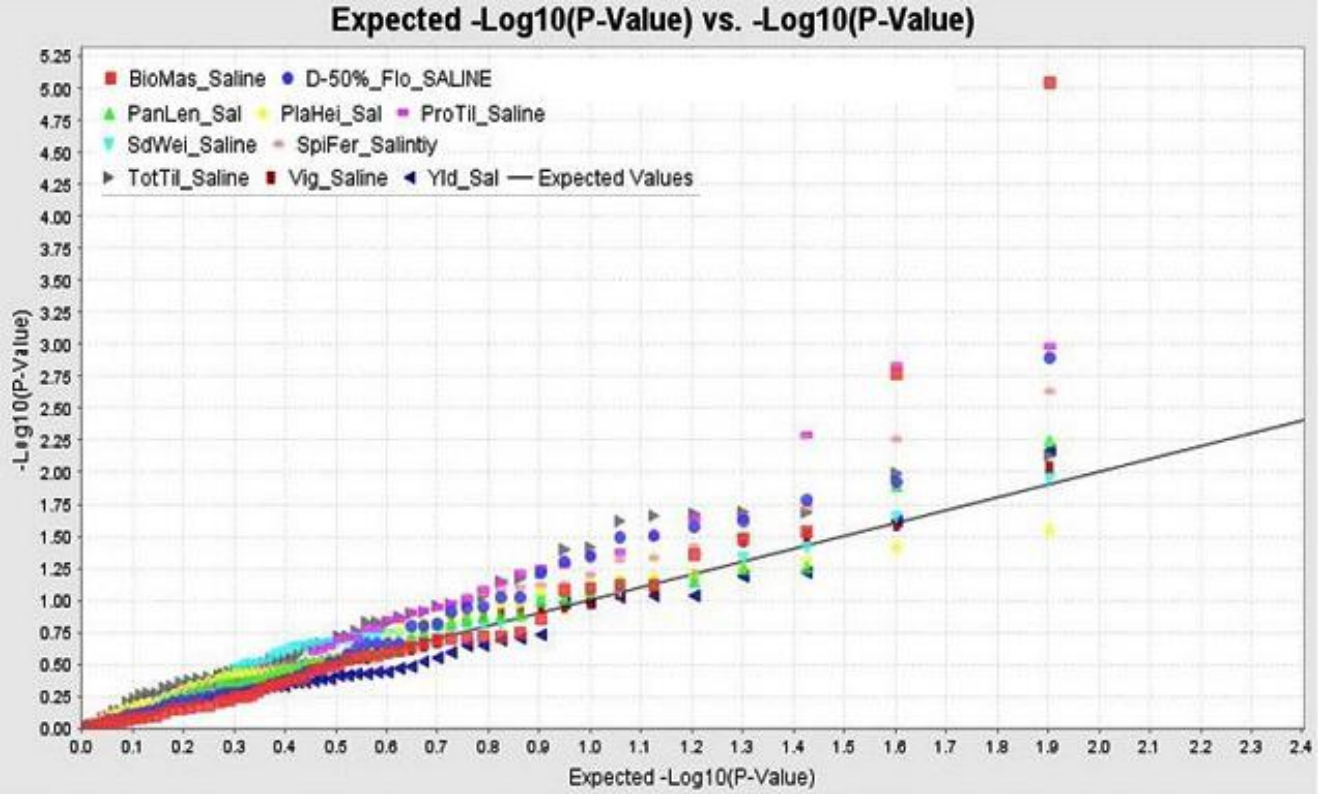


Fig. 4 — Quantile-Quantile (QQ) plot for selected traits under saline conditions. [The bold black line represents the expected P-values and coloured dots along the bold line represents observed P-values for the respective traits under saline conditions]

study, uniform distributions between expected and observed P -values for all traits were observed during the analysis (Fig. 4) and the marker-trait associations were present on different regions of rice chromosomes (Fig. 5). Under saline environment, marker OsCML8 showed significant association with three traits i.e. biomass, plant height and panicle length. Other marker RM6320 showed significant association with both productive tillers and vigor (salt injury) score (Table 5). Like most of the traits, biomass was associated with two significant markers OsCML8 and RM518, explaining notable phenotypic variability of 18.5 and 10.41%, respectively. Panicle length was associated with marker RM1261 and OsCML8 which accounted for phenotypic variance of 9.38% and 6.57%. Three markers were associated with productive tillers and out of which marker RM340 and RM3827 explained phenotypic variance of more than 10% each. RM431 and RM551 on chromosome 9 and 4 were associated with spikelet fertility. Yield under saline stress showed associations with markers RM80 and RM413, both collectively explained phenotypic variation of ~14%. Three markers (RM540, RM3310 and RM6320) were significantly linked to vigor score under salinity stress on chromosome 3, 4 and 5, respectively.

## Discussion

Salinity is a major constraint to productivity in rice and it has kept million of hectres of salt affected land uncultivated. The problem could be overcome by improving our conventional varieties so that rice can withstand the harsh saline environment. Finding some novel genes and their suitable allelic variants is an important first step toward development of these salt reselient varieties. Over the years many QTLs studies were carried out using biparental QTL mapping approach to identify these genes and they have been successful in tagging several QTLs time to time. But, the biparental approach of tagging genes has certain shortfalls in some aspects; as these studies are not comprehensive due to limited parental usage and non uniform recombination pattern[21].

In an effort to identify QTLs more comprehensively, we have adopted GWAS approach. Association mapping in the present study was successful in tagging the regions of genome to complex traits of salinity tolerance. Like most complex traits, the salt

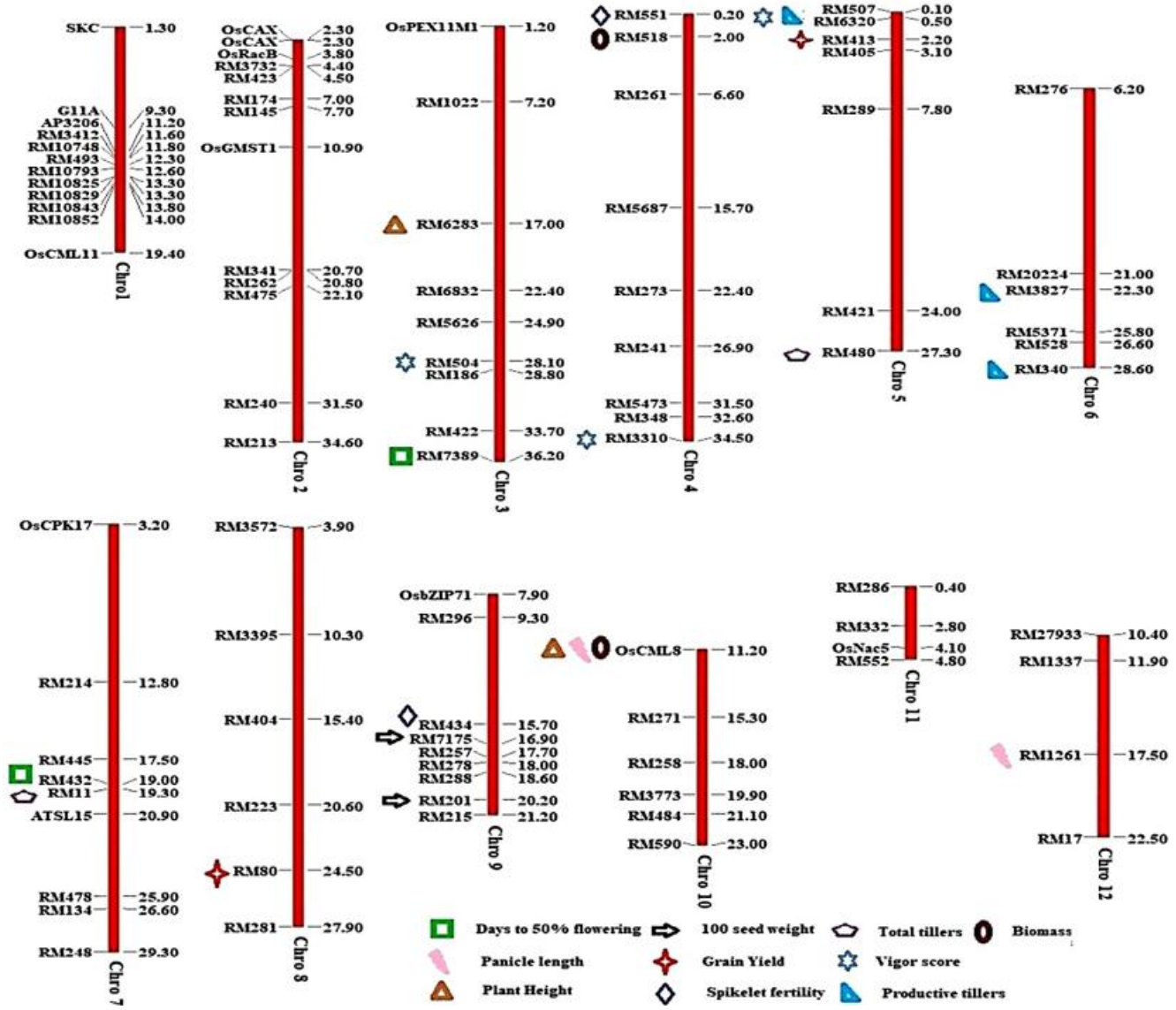


Fig. 5 — Framework map of 12 chromosomes of rice and position of different marker-trait associations present of different chromosomes

tolerance in rice is governed by many genes and genetic diversity plays an important role to maintain productivity as it offers wide platform to identify new genes for biotic and abiotic stresses. Various studies thatexplored genetic diversity in rice[22-24]have shown that diversity in the agro-morphological characters as suitable criteria in screening and selection of genotypes. In this experiment, we discovered significant genetic variation for different traits which was evident from the range values of each trait. Beside correlation, coefficient of variation (CV) also emerged as good selection criteria for breeding purpose. In our experiment, yield, productive tillers and spikelet fertility exhibited average CV of 65.88, 53.23 and 34.40%, respectively. These figures suggested the importance of these traits as selection criteria for breeding program. Furthermore, by analyzing the population structure, the entire rice panel was divided in to 4 subgroups which explained the genetic composition of each subgroup by assembling similar genotypes in same subgroup. Mean value of every traits for each subgroup explained the tolerance level and the composition of each subgroup. During the comparative analysis based on performance of each subgroup, subgroup 2 has performed best under stress conditions and it was mostly comprised of genotypes

which were released for salt affected areas. This subgroup also comprised of some new salt tolerant lines which could easily become ideal donors for development of salt tolerant varieties in future.

Genotypes of diverse geographical origin contain either population structure credited to local adaptations or familial relatedness due to co-ancestry and both these factors were prominent in our study panel too. Population structure is universal among organisms, particularly in the self fertilizing species and it can easily lead to false positive results if it is not controlled during association studies. Rice has a reputation of showing sluggish LD decay rate, stronger population structure, making GWAS study complex[25,26]. Successful utilization of the population structure in the present association analysis helped in identification of 22 robust marker trait associations by eliminating false positive. Some of these associations were common with previous studies. Like marker RM201 in our study was linked to 1000 seed weight and it was in concordance with previous findings[27]. Similarly, RM80 was associated with grain yield under saline stress; the similar results have also been reported in the past[28]. Moreover, RM80 was also reported to be associated with osmotic adjustment and spikelet sterility under draught stress[29]. Osmotic adjustments and spikelet sterility are important parameters of salinity tolerance too. Both salinity and drought stresses share similar symptoms that undermine the importance of this marker to be considered for molecular screening of genotypes. An important parameter of salinity tolerance is vigor (salinity injury) score which in our case was linked to three markers viz., RM504, RM3310 and RM6320, located on different chromosomes. Earlier study[30] had also reported few of these markers for vigor score in biparental RIL population. In addition, RM6320 was reported to be significantly linked with shoot $Na^+/K^+$ ratio[41]. Marker RM518 in our analysis was linked to biomass however, in the past this marker was reported for days to maturity and plant height[31,32]. Marker RM518 was also reported for total spikelet number in salt stress condition[33]. Throughout the association analysis, we have found that same marker showed association to different traits i.e. marker OsCML8 was linked to biomass, panicle length and plant height. Similarly, marker RM6320 was linked to productive tillers and vigor (salt injury) score which could be attributed to pleiotropic effect and interdependence of traits.

Comparative genomic search of reported markersusing Gramene BLAST program (http://ensembl.gramene.org/common/Tools/Blast?db=core) revealed multiple genes in close proximity of these markers and few of these markers shared overlapping regions with these genes Marker RM518 was overlapping with gene *Os04g0132400* and receptor-like cytoplasmic kinase 127 (*OsRLK127*). *OsRLK* gene family codes for transmembrane proteins which helps in intra cellular signal transduction through ATP binding kinase activity[34]. Furthermore, these proteins are functionally annotated to play crucial role in development of salt response in rice plant[35]. Similarly, RM432 was found next to gene *Os07g0502200* and *Os07g0502550*, the former enables the active transport of a solute across a membrane while the function of later is not characterized. Another marker RM6283 was found overlapping with gene HAP2 subunit of CCAAT-box binding complex which codes for CCAAT-binding transcription factor which binds toregulatory region of gene to modulate transcription. The heteromeric transcription factor complex binds to the CCAAT-box upstream of promoters and it could either function as activator and repressor depending on the interacting cofactors[36]. Over expression of this subunit in rice seedling leads to enhanced root and shoot growth under saline conditions[37,38].

Similarly, RM6320 was found in close proximity of gene no lysine kinase 2 whose alternative is WNK2. This gene was involved in intracellular transduction and high transcript accumulation of this gene was also observed under draught stress. RM551 overlapped with gene *Os04g0102500* which is cytosolic in origin however, the function of this gene was not characterized. Comparative genomic analysis revealed marker RM421 next to gene *Os05g0489000*, the given gene codes for a protein that has putative inositol-1,4,5-trisphosphate 5-phosphatase activity. The given gene plays crucial role in metabolic pathway of inositol. Inositol metabolic intermediates like pinitol and inositol polyphosphates play significant role in signal transduction and stress tolerance. Pinitol accumulation in halophytic wild rice *Porteresia coarctata* and its absence in domesticated rice showed its importance in salt tolerance[39]. Finally, RM504 was overlapping with gene *Os03g0701000* which has protein transporter activity, it direct the movement of a protein bearing a nuclear localization signal (NLS) from the cytoplasm into the nucleus.

The fact that all these markers were overlapping and lying next to genes is a useful finding; it could be presumed that they remained conserved and inherited together over the course of evolution, making these markers worthy to be considered for future indepth studies and role of these genes in salt tolerance should be futher examined through transcriptomic studies.

Over the few years, sequencing of genome has gain momentum as it become more economical and fast. One such implications of sequencing had its direct impact on the discovery of SNPs across different genotypes throughout the genome. High quality reference genome sequence of rice are also available. Therefore, a worldwide effort to collect rice accessions for whole-genome sequencing and comprehensive phenotyping is under way and results of association mapping and data for these accessions are being made public, so that researchers engaged in rice GWAS study could use this information for genotyping and reference for different environments and different set of parameters. Earlier, some studies have made their GWAS results public and these results culd be visualized under GWAS viewer for reference (http://rs-bt-mccouch4.biotech.cornell.edu/GWAS_Viewer/plot/)[40] Similarly, SNPs based hapmap for 1,529 rice accessions of *Oryza rufipogon* and *Oryza sativa* is available at http://server.ncgr.ac.cn/RiceHap3 and some other independend GWAS results are also made avialable at http://iric.irri.org/news/gwasviewer. Application of association mapping to plant breeding seems to be a promising means of overcoming the limitations of conventional linkage mapping. Our results have shown that association studies are efficient means of identifying novel genes for important agronomic characters under saline stress. The newly identified salt tolerant lines will help in broadening the gentic base for development of varities with enhanced level of salt tolerance. A number of different marker-trait associations detected during this study, and their resemblance with previous studies, established their importance in introgression and screening purpose. This study, therefore, gives an overview of long term collective effort of discovering valuable genes that play an important role in governing salt tolerance in rice.

## Conclusion

The results have demonstrated the ability of GWAS to successfully identify potential marker trait associations for complex traits of salinity tolerance in rice. The present marker trait associations cumulatively accounted for significant phenotypic variation for each trait with few associations explained phenotypic variation greater than 10%. Furthermore, the associations with small to moderate effects indicated that salinity is a complex trait controlled at multiple levels. Besides, its comparative genomic search of associated markers revealed numbers of genes either overlapping or were present in near proximity to these associated markers which make these markers crucial for further genomic studies. These genes are involved in signaling under saline stress, transcription factor activity, enzymatic activity and transporter function. The novel associations identified in the present study could become suitable candidates for further functional characterization through transcriptomic and proteomic analysis under saline stress. Furthermore, the genetic importance of these associations should be further confirmed by using dense marker system on different genetic background to establish their role in salinity tolerance.

## Conflict of interest

Authors declare no conflict of interests.

## References

1 Hossain M, Sustaining food security for fragile environments in Asia: Achievements, challenges, and implications for rice research, Fragile Lives in Fragile Ecosystems. *Int. Rice Res Inst*, 286 (1995) 3.

2 Wallender WW & Tanji KK, Agricultural Salinity Assessment and Management, 2nd ed. (American Society of Civil Engineers, Reston, VA, USA), 2012.

3 Mass EV & Hoffman GJ, Crop salt tolerance current assessment. *J Irrig Drain Eng*, 103 (1997) 115.

4 Krishnamurthy SL, Gautam RK, Sharma PC & Sharma DK, Effect of different salt stresses on agro-morphological traits and utilization of salt stress indices for reproductive stage salt tolerance in rice. *Field Crop Res*, 190 (2016) 26.

**5** Rahman MA, Thomson MJ, Ocampo MD, Egdane JA, Salam MA, Alam MS & Ismail AM, Assessing trait contribution and mapping novel QTL for salinity tolerance using the Bangladeshi rice landrace Capsule. *Rice*, 12 (2019) 63.

6 Krishnamurthy SL, Sharma PC, Sharma DK, Ravikiran KT, Singh YP, Mishra VK,Burman D, Maji B, Mandal S, Sarangi SK, Gautam RK, Singh RK, Manohara BC, Marandi KK, Padmavathi G, Vanve PB, Patil KD, Thirumeni S, Verma OP, Khan AH, Tiwari S, Geetha S, Shakila M, Gill R, Yadav VK, Roy SKB, Prakash M, Bonifacio J, Ismail A, Gregorio GB & Singh RK, Identification of mega-environments and rice genotypes for general and specific adaptation to saline and alkaline stresses in India. *Sci Rep*, 7 (2017) 7968.

7 Tiwari S, Krishnamurthy SL, Kumar V, Singh B, Rao AR, Amitha SV, Rai V, Singh AK & Singh NK, Mapping QTLs for salt tolerance in ice (*Oryza sativa* L.) by bulked segregant analysis of recombinant inbred lines using 50K SNP chip. *PLoS ONE*, 11 (2016) e0153610. Doi-101371/journalpone0153610.

8 Babu NN, Gopala Krishnan S, Vinod KK, Krishnamurthy SL, Vivek SK, Singh MP, Singh R, Ellur RK, Rai V, Bollinedi H, Bhowmick PK, Yadav A, Nagarajan M, Singh NK, Prabhu KV & Singh AK, Marker aided incorporation of Saltol, a major QTL associated with seedling stage salt tolerance, into *Oryzasativa* 'Pusa Basmati 1121'. *Front Plant Sci*, 8 (2017) 41 doi: 103389/fpls201700041.

9 Krishnamurthy SL, Pundir P, Warraich AS, Rathor S, Lokeshkumar BM, Singh NK & Sharma PC, Introgressed Saltol QTL Lines Improves the Salinity Tolerance in Rice at Seedling Stage. *Front. Plant Sci*, 11 (2020) 833. doi: 10.3389/fpls.2020.00833.

10 Holland JB, Genetic architecture of complex traits in plants. *Curr Opin Plant Biol*, 10 (2007) 156.

11 Naveed SA, Zhang F, Zhang J, Zheng TQ, Meng LJ, Pang YL, Xu JL & Li ZK, Identification of QTN and candidate genes for Salinity Tolerance at the Germination and Seedling Stages in Rice by Genome-Wide Association Analyses. *Sci Rep*. 8 (2018) 6505 doi.org/10.1038/s41598-018-24946-3.

12 Li N, Zheng H, Cui J, Wang J, Liu H, Sun J, Liu T, Zhao H, Lai Y & Zou D, Genome-wide association study and candidate gene analysis of alkalinity tolerance in *japonica* rice germplasm at the seedling stage. *Rice*, 12 (2019).

13 Yuan J, Wang X, Zhao Y, Khan NU, Zhao, Z, Zhang Y, Wen X, Tang F, Wang F & Li Z, Genetic basis and identification of candidate genes for salt tolerance in rice by GWAS. *Sci Rep*, 10 (2020) 9958.

14 Kumar V, Singh A, Mithra ASV, Krishnamurthy SL, Parida SK, Jain S, Kapil K, Tiwari KK, Kumar P, Rao AR, Sharma SK, Khurana JP, Singh NK & Mohapatra T, Genome-wide association mapping of salinity tolerance in rice (*Oryzasativa*). *DNA Res*, 22 (2015) 133.

15 Xalxo SM., Saxena, RR, Pali V, Kurrey LN & Verulkar S, A Rapid and Low Cost Method of DNA Isolation for Marker Assisted Selection in Rice. *Agric Res*, 3 (2014) 83.

16 You FM, Huo N, Gu YQ, Luo MC, Ma Y, Hane D, Lazo GR, Dovark J & Anderson OD, BatchPrimer3: A high throughput web application for PCR and sequencing primer design. *BMC Bioinformatics*, 9 (2008) 253.

17 Pritchard JK, Stephens JK & Donnelly P, Inference of Population Structure Using Multilocus Genotype Data. *Genetics*, 155 (2000) 945.

18 Evanno G, Regnault S & Goudet J, Detecting the number of clusters of individuals using the software STRUCTURE- a simulation study. *Mol Ecol*, 14 (2005) 2611.

19 Warraich AS, Krishnamurthy SL, Sooch BS, Vinaykumar NM, Dushyanthkumar BM, Bose J & Sharma PC, Rice GWAS reveals key genomic regions essential for salinity tolerance at reproductive stage. *Acta Physiol Plant*, 42 (2020) 134.

20 Swamy BPM, Shamsudin NA, Rahman SN, Mauleon R, Ratnam W, Cruz TS & Kumar A, Association Mapping of Yield and Yield-related Traits Under Reproductive Stage Drought Stress in Rice (*Oryza sativa* L.). *Rice*, 10 (2017) 21.

21 Bartoli C & Roux F, Genome-Wide Association Studies In Plant Pathosystems: Toward an Ecological Genomics Approach. *Front Plant Sci*, 8 (2017) 763 DOI=10.3389/ fpls.2017.00763.

22 Ravikiran KT, Krishnamurthy SL, Warraich AS & Sharma PC, Diversity and haplotypes of rice genotypes for seedling stage salinity tolerance analyzed through morpho-physiological and SSR markers. *Field Crop Res*, 220 (2018) 10.

23 Ganie SA, Karmakar J, Roychowdhury R, Mondal TK & Dey N, An exploratory study on allelic diversity among rice an its wild species as well as relatives with simple sequence repeat and inter simple sequence repeat markers. *Indian J Exp Biol*, 15 (2016) 357.

24 Salgotra RK, Bhat JA, Gupta BB & Sharma S, Determination of genetic relationship among basmati and non basmati genotypes for North-West Himalayas using microsatellite markers. *Indian J Exp Biol*, 16 (2017) 68.

25 Zhu C, Gore M, Buckler ES & Yu J, Status and prospects of association mapping in plants. *Plant Gen*, 1 (2008) 5.

26 Lekklar C, Pongpanich M, Duangjai SA, Chinpongpanich A, Tsai H, Comai L, Chadchawan S & Buabooch T, Genome-wide association study for salinity tolerance at the flowering stage in a panel of rice accessions from Thailand. *BMC genomics*, 20 (2019) 76.

27 Mohammadi NG, Arzani A, Rezai AM, Singh RK & Gregorio GB, Assessment of rice genotypes for salt tolerance using microsatellite markers associated with the saltol QTL. *Af J Biotechnol*, 7 (2008) 730.

28 Robin S, Pathan MS, Courtois B, Lafitte R, Carandang S, Lanceras S, Amante M, Nguyen HT & Li Z, Mapping osmotic adjustment in an advanced back-cross inbred population of rice. *Theor Appl Genet*, 107 (2003) 1288.

29 Lanceras JC, Pantuwan G, Jongdee B & Toonjinda T, Quantitative trait loci associated with drought tolerance at reproductive stage in rice. *Plant Physiol*, 135 (2004) 384.

30 Cheng L, Wang Y, Meng L, Hu X, Cui Y, Sun Y, Zhu L, Ali J, Xu J & Li Z, Identification of salt-tolerant QTLs with strong genetic background effect using two sets of reciprocal introgression lines in rice. *Genome*, 55(2011) 45.

31 Ghomi K, Rabiei B, Sabouri H & Sabouri A, Mapping QTLs for traits related to salinity tolerance at seedling stage of rice (*Oryzasativa* L). *Omics*, 17 (2013) 242.

32 Septiningsih EM, Prastiyono J, Lubis E, Tai TH, Tjubaryat T, Moeljopawiro S & McCouch SR, Identification of quantitative trait loci for yield and yield components in an advanced backcross population derived from the *Oryzasativa* variety IR64 and the wild relative O rufipogon. *Theor Appl Genet*, 107 (2003) 1419.

33 Mohammadi R, Mendioro MS, Diaz GQ, Gregorio GB & Singh RK, Mapping quantitative trait loci associated to yield and yield components under reproductive stage salinity stress in rice (*Oryzasativa L*). *J Genet*, 92 (2013) 433.

34 Shiu SH & Bleecker AB, Plant receptor-like kinase gene family: diversity, function, and signalling. *Science*'s *SKTE*, 113 (2001) 22.

35 Zhou YB, Liu C, Tang DY, Yan L, Wang D., Yang YZ, Gui JS, Zhao XY, Li LG, Tang XD, Yu F, Li JL, Liu LL, Zhu YH, Lin JZ & Liu XM, The Receptor-Like Cytoplasmic Kinase STRK1 Phosphorylates and Activates Cat C, Thereby

Regulating $H_2O_2$ Homeostasis and Improving Salt Tolerance in Rice. *Plant Cell*, 30 (2018) 1100.

36 Ceribelli M, Dolfini D, Merico D, Gatta R, Vigano AM, Pavesi G & Mantovani R, The Histone-Like NF-Y Is a Bifunctional Transcription Factor. *Mol Cell Biol*, 28 (2008) 2047.

37 Alam M, Tanaka T, Nakamura H, Ichikawa H, Kobayashi K, Yaeno T, Yamaoka K, Takayama K, Nishina H & Nishiguchi M, Overexpression of a rice heme activator protein gene (OsHAP2E) confers resistance to pathogens, salinity and drought, and increases photosynthesis and tiller number. *Plant Biot J*, 13 (2015) 85.

38 Krishnamurthy SL, Sharma PC, Sharma SK, Batra V & Kumar K, Effect of salinity and use of stress indices of morphological and physiological traits at the seedling stage in rice. *Indian J Exp Biol*, 54 (2016) 843.

39 Sengupta S, Patra B, Ray S & Majumdar AL, Inositol methyl tranferase from a halophytic wild rice, Porteresia coarctata Roxb. (Tateoka): regulation of pinitol synthesis under abiotic stress. *Plant Cell Environ*, 31 (2008) 1442.

40 McCouch SR, Wright MH, Tung CW, Maron LG, McNally KL, Fitzgerald M, Singh N, DeClerck G, Agosto-Perez F, Korniliev P, Greenberg AJ, Naredo Ma EB, Mercado SMQ, Harrington SE, Shi Y, Branchini DA, Kuser-Falcão PR, Leung H, Ebana K, Yano M, Eizenga G, McClung A & Mezey J, Open access resources for genome-wide association mapping in rice. *Nat Commun*, 7 (2016) Article no. 10532. https://doiorg/101038/ncomms10532.