Diversity in Rice Genotypes under Salt Affected Soil Based on Multivariate Analysis

A. Anandan[*], R. Eswaran and M. Prakash

Department of Genetics and Plant Breeding, Faculty of Agriculture,
Annamalai University, Chidambaram – 608 002, Tamil Nadu, India
[*]E-mail: anandanau@yahoo.com

ABSTRACT

Diversity of forty four salt tolerant rice genotypes from different geographic regions were assessed using Mahalanobis D^2 and principal component analysis (PCA). The D^2 statistics grouped the genotypes into 12 distinct clusters. Consisting of 19 genotypes, Cluster I was the largest cluster, followed by Cluster II with nine genotypes. Meanwhile, Clusters III, IV, VI, VII, IX, and XII were mono-genotypic clusters. The maximum intercluster distance was exhibited between Cluster IX and X (144.91), followed by Clusters II and X (131.87), as well as clusters VII and X (126.27). The number of grains per panicle (42.71%), followed by the grain yield per plot (29.81%), was the major contributor to the total divergence. The PCA revealed that axes 1 and 2 accounted for 82.88% and 11.14% of the variance, respectively. The highest contributing variable was the number of grains per panicle in PC1 and the plant height in PC2. The genotypes from more than one place of origin were grouped in one cluster, whereas the genotypes from one state were grouped in more than one cluster. Both D^2 and PCA revealed that the morphometric diversity was based on the pedigree and independent of geographical origin. Hybridization among the genotypes which had the maximum inter-cluster distances could produce heterotic combinations and wide variability in segregating generations for many beneficial traits.

Keywords: Morphometric diversity, Mahalanobis D^2, *Oryza sativa*, principal component analysis, rice

INTRODUCTION

The presence of excess salt is one of the most widespread soil toxicity problems in many rice growing areas. In particular, it accounts for 8.5 million hectares of land in India and the yield reduction is estimated at 30-50% (Babu *et al*., 2005). The success of any breeding programme is dependent on the available genetic divergence in the crop. Rice germplasm is known to be a rich source of salt tolerant genes (Yeo and Flowers, 1982) for improving salinity tolerance in high yielding varieties.

A narrow genetic base among released cultivars and the practice of using elite line x elite line crosses have been implicated in slowing the rate of genetic advance for yield (Lal and Rana, 2000). Meanwhile, divergent parents in a cross have a greater scope of obtaining heterotic F_1s and a broad spectrum of variability in the segregating generations in any varietal improvement programme (Arunachalam, 1981). Hence, analysis of genetic relationships among extant genotypes is an important component of crop improvement programme as it provides information about genetic diversity and helps to stratify breeding populations.

---

[*]Corresponding Author

Clustering germplasm into various groups, using hierarchical or non-hierarchical algorithms based on the multivariate statistical techniques and sampling from within discrete groups, is a common method used for maximizing diversity. Thus, the present investigation was designed and carried out to: (i) investigate the extent of genetic diversity in saline tolerant rice genotypes, (ii) identify promising genotypes for future utilization in hybridization for developing saline tolerant genotype with high yield, and (iii) estimate the nature and magnitude of the relationship between quantitative traits associated with yield under salinity environment.

MATERIALS AND METHODS

In the present study, forty four genetically diverse salt tolerant genotypes of rice taken from the different regions of India were investigated (Table 1). It is important to note that these genotypes were grown in saline soil with electrical conductivity (EC) of 2.83 ds m^{-1} during the dry seasons of 2006 and 2007. The experiments were conducted at the experimental farm of Plant Breeding (11° 24' N latitude and 79° 44' E longitude, + 5.79 m ASL), located in Annamalai University, Tamil Nadu, India. The seeds were sown in a raised nursery bed with good irrigation water, whereas the 25-day old seedlings were transplanted at the field. The experiment was laid out in a randomized block design with three replications, using 20 x 20 cm spacing in 10 m^2 plots. Crop management corresponded to intensive farming methods with full insecticide and fungicide cover. The recommended fertilizer doses of 150 kg nitrogen ha^{-1}, 50 kg phosphorus ha^{-1}, and 50 kg potash ha^{-1} were applied. Manual weeding was carried out twice. Observations were recorded on ten randomly selected plants in each experimental unit for seven characters, namely, days to 50% flowering, plant height, number of productive tillers, panicle length, number of grains per panicle, yield per plot, and percentage of survival. Meanwhile, the data gathered from the two years were pooled in the analysis and subjected to estimate genetic divergence. The

D^2 analysis was done as per Mahalanobis's D^2–statistic and extended by Rao (1970). The estimation of the D^2 values was according to the following formula:

$$D^2 = \sum \sum w^{ij} \left(\overline{x}_i^1 - \overline{x}_i^2 \right)\left(\overline{x}_j^1 - \overline{x}_j^2 \right)$$

Where, w^{ij} is the inverse of variance and co-variance matrix.

Furthermore, $_{the}$ computation of the D^2 values was reduced to simple summation of the differences in the mean values of various characters of two populations, i.e. Σd^2_i. Therefore, the correlated variables were first transformed to the uncorrelated ones so as to work out the D^2 values.

$$D^2 = \sum d_i^2 = \sum \left(Y_i^j - Y_i^k \right)^2$$

Where, Y is the uncorrelated variable which varied from i = 1 to x, i.e. the number of characters. Y_i^j and Y_i^k represented the transformed uncorrelated mean of the ith character for the genotypes *j* and *k*, respectively. The genotypes were then grouped on the basis of minimum generalised distance using the Tocher's method as described by Rao (1970), and this was followed by the principal component analysis (PCA). These analyses were performed using the software Windowstat 7.5 (Indostat Services, Hyderabad, India).

RESULTS AND DISCUSSION

Based on the D^2 analysis, the genotypes under saline soil condition could be grouped into 12 clusters (Table 2). Cluster I comprised of 19 genotypes, whereas Cluster II had nine genotypes. The genotypes from a few states (Andhra Pradesh, Haryana, Maharastra, Orissa, and Uttar Pradesh) or with similar pedigree were grouped in the same cluster, as revealed by Clusters I and II. This is in agreement with the findings reported by Arunachalam and Ram (1967). In contrast, some genotypes taken from Andhra Pradesh, Maharastra, and West Bengal were grouped in Cluster V,

TABLE 1
List of rice genotypes selected for diversity analysis, with their parentage and place of origin

S. No.	Genotype code	Parentage	Origin
1	G1	KDML 105 x IR 4630-22-2-5-1-3 x IR 20925-33-3-1-28	Haryana, India
2	G2	Savithri x Lunishree	Orissa, India
3	G3	Jaya x Lunishree	Orissa, India
4	G4	Savithri x Lunishree	Orissa, India
5	G5	Mahsuri x Ormundakan	Orissa, India
6	G6	TCCP 266-249-B-B-3 x IR 262-43-8-1	Uttar Pradesh, India
7	G7	IR 55182-3B-14-3-2-2- x IR 4499-29-2-2-2	Uttar Pradesh, India
8	G8	IR 70804-9-NDR-3-7-91	Uttar Pradesh, India
9	G9	IET 8320 x PNL 1	Maharastra, India
10	G10	PNL 2 x IET 8320	Maharastra, India
11	G11	PNL -2 x IET 8320	Maharastra, India
12	G12	Jaya x CSR 23	Haryana, India
13	G13	IR 72 x CSR 23	Haryana, India
14	G14	Mahsuri x Madhukar 105	Uttar Pradesh, India
15	G15	IR 68661-16-8 x NDR-2-B-2-1	Uttar Pradesh, India
16	G16	Usar 1 x Mahsuri	Uttar Pradesh, India
17	G17	Selection from Kalanamak	Uttar Pradesh, India
18	G18	CSR 21 x CSR 10	Haryana, India
19	G19	IR 64 x IR 4630-22-2-5-1-3 x IR 9764-45-2-2	Haryana, India
20	G20	IR 64 x PNL 2	Maharastra, India
21	G21	IET 13845 x GJ 11	Maharastra, India
22	G22	IR 28 x Chakrakonda	Orissa, India
23	G23	Mahsuri x Chakrakonda	Orissa, India
24	G24	Jaya x Lunishree	Orissa, India
25	G25	Jaya x Lunishree	Orissa, India
26	G26	CSR 23 x CSR 10	Haryana, India
27	G27	Bipasa x Kalojira	West Bengal, India
28	G28	Pankaj x SR 26B	West Bengal, India
29	G29	Nonabokra x IR 36	West Bengal, India
30	G30	Mutant of IR 4630-22-2-5-1-3 x Pokkali	Andhra Pradesh, India
31	G31	IET 9993 (MS) x N 52	Andhra Pradesh, India
32	G32	CSR 3 x Kasturi	Andhra Pradesh, India
33	G33	CSR 3 x Kasturi	Andhra Pradesh, India
34	G34	CSR 3 x Kasturi	Andhra Pradesh, India
35	G35	CSRL-04-2366	Uttar Pradesh, India
36	G36	CSRL-01-IR 75	Uttar Pradesh, India
37	G37	CSRL-01-IR 97	Uttar Pradesh, India
38	G38	IET 9993 (MS) x BM 47	Andhra Pradesh, India
39	G39	Pankaj x SR 26B	West Bengal, India
40	G40	CSR 1 x Basmati 370 x CSR 5	Haryana, India
41	G41	Nona Bokra x IR 5657-33-2	Haryana, India
42	G42	CST 7-1	Haryana, India
43	G43	TN1 x T141	Andhra Pradesh, India
44	G44	IR 578-172-2-2 x BR-1-2-B-19	Tamil Nadu, India

TABLE 2
Clustering pattern of rice genotypes based on the D2 analysis

Cluster No.	No. of genotypes	Genotypes
I	19	G18, G19, G42, G36, G37, G16, G22, G31, G27, G38, G6, G32, G25, G7, G20, G29, G33, G12 and G15
II	09	G13, G30, G23, G8, G2, G1, G10, G9 and G26
III	01	G34
IV	01	G4
V	04	G21, G39, G28 and G43
VI	01	G35
VII	01	G44
VIII	02	G5 and G17
IX	01	G11
X	02	G3 and G41
XI	02	G24 and G40
XII	01	G14

i.e. containing four genotypes (Sinha *et al.*, 2001). The remaining genotypes were scattered into monogenic clusters, namely III, IV, VI, VII, IX, and XII. This seems to suggest that geographical distribution is not necessarily related to genetic divergence (Sarawgi and Shrivastava, 1996). For instance, the scattering of the genotypes from the same geographic region into different clusters may be due to genetic heterogeneity (different genes producing identical phenotypes), genetic architecture and history of selection (Murty and Arunachalam, 1966). In particular, Clusters VIII, X and XI comprised of two genotypes each. The genotypes belonging to these three clusters originated from the states of Haryana, Orissa and Uttar Pradesh. These findings are in an agreement with that of other similar rice studies conducted by Ratho (1984), Rathore *et al.* (2001) and Sharma *et al.* (2008). The grouping of the genotypes from the different states into a single cluster might be due to unidirectional selection practiced by the breeders or free exchange of germplasm among them.

The intracluster distance varied from 0.00 to 40.16, as shown in Table 3. The maximum distances for Clusters XI, V, X, VIII, I, and II were 40.16, 37.12, 36.16, 35.9, 34.54, and 33.73, respectively. Intracluster distance was zero in Clusters III, IV, VI, VII, IX, and XII since each contained only a single genotype. Meanwhile, the minimum intercluster distance was obtained between Clusters III and IV (24.95), and this was followed by Clusters III and V (38.99). The maximum intercluster distance was observed between Clusters IX and X (144.91). The figures stated for II vs. X and VII vs. X were 131.87 and 126, respectively. The presence of high intercluster distanced genotypes permits the selection of divergent parents. This avoids the selection of parents from genetically homogeneous clusters, and thus reducing the breeding of a population with the likelihood of a narrow genetic base. Higher genetic distance between the clusters suggested a wide diversity among the genotypes. The crosses made between the genotypes from the above clusters might give useful transgressive segregants (Sharma and Bhuyan, 2004). According to Rieseberg *et al.* (1999), transgressive segregation is the production of F_2 or later-generation hybrid progeny with phenotypes that could fall outside the phenotypic range of the parental populations from which they were derived. Meanwhile, the improvement of self-pollinated crops is connected with the production of homozygous

TABLE 3
Intra- (bold) and inter-cluster distances for 12 clusters in rice in relation to Table 2

Cluster No.	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
I	**34.54**	50.78	45.41	45.29	72.73	42.05	52.85	83.55	53.75	107.78	88.69	56.33
II		**33.73**	76.01	67.73	106.60	51.08	79.09	102.46	60.36	131.87	114.45	69.51
III			**0.00**	24.95	38.99	45.35	55.74	55.38	78.51	80.15	75.68	63.18
IV				**0.00**	53.40	39.18	59.44	46.03	76.88	89.43	92.01	73.74
V					**37.12**	74.93	68.17	66.15	99.13	79.23	78.03	85.53
VI						**0.00**	76.54	66.54	79.01	97.62	92.84	49.80
VII							**0.00**	99.00	41.83	126.27	101.35	79.91
VIII								**35.90**	119.79	75.67	104.30	102.87
IX									**0.00**	144.91	112.82	82.60
X										**36.16**	66.05	116.27
XI											**40.16**	93.27
XII												**0.00**

selections superior to the parental genotypes. It is widely known that the frequency of the transgression effects in homozygous populations is dependent on cross combinations, i.e. on the genotypes of crossing parents. The characters contributing more to the divergence are given greater emphasis for deciding on the cluster for the purposes of further selection and the choice of patterns for hybridization (Jagadev *et al.*, 1991). The differences between the genotypes for the number of grains per panicle, grain yield per plot and percentage of survival were more pronounced as compared to the other traits (Table 4). In particular, the number of grains per panicle seems to be the most important as the percentage of its contribution was the maximum

TABLE 4
Mean performance of different clusters with respect to different traits in rice

Cluster / traits	Days to 50% flowering	Plant height	No. of productive tillers/plant	Panicle length	No. of grain/ panicle	Grain yield/ plot	Survival percent
I	86.86	104.02	19.33	29.20	109.91	5.28	61.37
II	85.63	92.47	17.46	25.52	78.41	1.91	61.34
III	86.33	87.22	14.83	28.10	166.33	6.15	62.00
IV	92.00	110.17	17.33	28.13	165.00	4.06	66.34
V	87.79	113.56	21.96	29.25	190.58	8.70	62.77
VI	77.17	121.53	27.50	28.45	132.67	2.39	61.17
VII	96.67	109.16	16.50	25.03	110.33	9.51	72.28
VIII	90.75	134.84	17.33	26.14	220.75	2.12	64.30
IX	97.83	94.30	22.83	26.48	59.00	7.71	66.78
X	85.58	131.38	17.67	30.21	222.33	5.13	34.85
XI	82.67	113.90	24.83	26.13	144.33	9.12	31.33
XII	60.33	93.85	20.50	26.50	102.33	5.78	60.67
Contribution to diversity (%)	7.08	1.16	1.59	0.42	42.07	29.81	17.86

(42.07%) in its genetic divergence. This was followed by the grain yielded per plant and the percentage of survival (De and Rao, 1987; Chaturvedi and Maurya, 2005).

The cluster mean values showed a wide range of variations for all the traits undertaken in this study (Table 4). In more specific, Cluster VIII was characterized with the high mean values for plant height and the number of grains per panicle. Meanwhile, Cluster XI exhibited a high mean for the number of productive tillers per plant and grain yield per plot. Clusters X, I, and V had the highest mean of panicle length. The single genotypic clusters were quite different from the other clusters by either the highest or the lowest value for a particular character. The monogenotypic Cluster XII had a low mean for days to 50% flowering and plant height. Cluster VI had a high mean for productive tillers per plant and percentage of survival. The putative parents for a systematic crossing programme should belong to diverse clusters which are characterized by a large intercluster distance. Such genotypes have genes with different magnitude of effects and higher probability in term of the chances to obtain recombinants outside the range of parents. Besides the high genetic divergence, considering the contribution of different characters towards the total divergence and the magnitude of cluster means for different characters should

also be given due importance in the selection of genotypes for hybridization programme. Based on the above data, the genotypes belonging to Clusters X (G3 and G41), VI (G35), VII (G44), and XII (G14) could therefore be used in the hybridization programme so as to obtain useful recombinants for saline environmental condition.

The graphical representation of the PCA analysis was applied to identify the genetic diversity among the genotypes and the traits responsible for the main source of variability. The first and second principal components (PC) accounted for 82.88% and 11.14% of the variance, respectively (Table 5). Since the first two PC accounted for about 94.02% of the total variability, a two dimensional representation of the relative position of the varieties in the biplot graph was found adequate.

The 82.88% variation in the first PC was mainly due to the variation in the number of grains per panicle. Loadings (latent vector) with large absolute values corresponding to the variables should have a greater discriminating ability. The first PC was positively correlated with all the characters, except for the percentage of plant survival. The largest absolute value for plant height in the second PC indicated that this trait was mainly responsible for explaining 11.14% of the total variance. Meanwhile, the second PC was positively correlated with the

TABLE 5
Principal components for rice genotypes based on seven characters

| Parameters	PC1	PC2	PC3	PC4
Eigen values	102366	13755	3934	1717
Percentage variance (%)	82.88	11.14	3.19	1.39
Cumulative variance (%)	82.88	94.02	97.21	98.60
Character	Latent vectors (loadings)			
Days to 50% flowering	0.009	-0.010	0.150	0.979
Plant height	0.218	-0.972	0.080	-0.023
No. of productive tillers/plant	0.005	-0.009	-0.013	-0.127
Panicle length	0.016	-0.003	0.029	0.000
No. of grain/panicle	0.974	0.222	0.034	-0.013
Grain yield/plot	0.014	0.004	-0.041	0.065
Survival percent	-0.053	0.073	0.984	-0.145
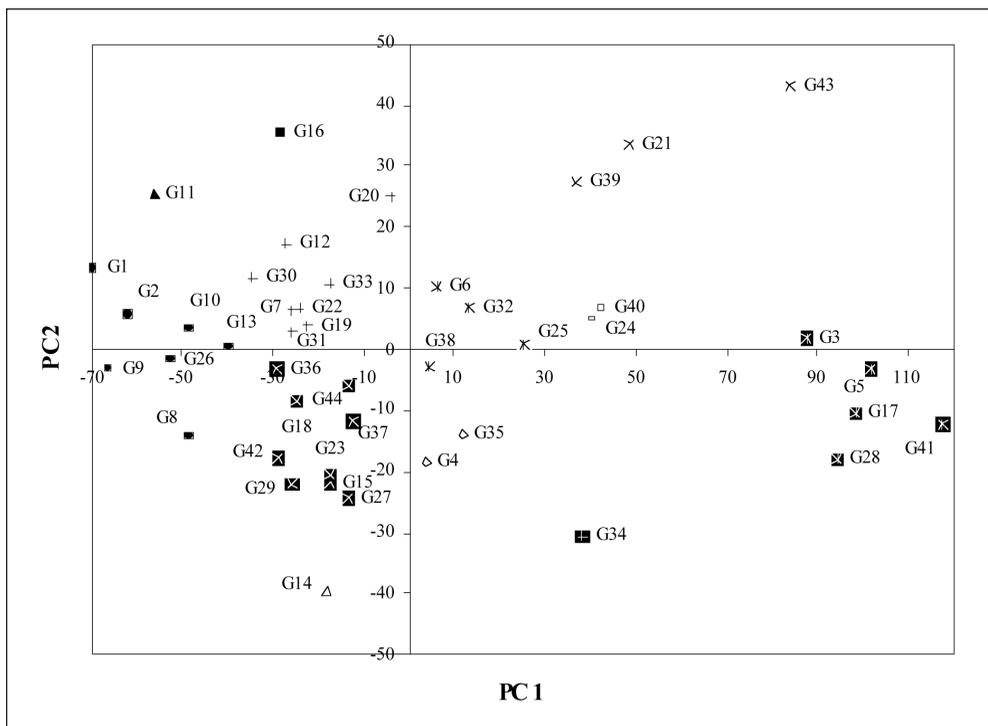
Fig. 1: Spatial distribution of 44 rice genotypes for the first two principal components.
Legend – ⊠ cluster I, ● cluster II, ➕ cluster III, ◊ cluster IV, × cluster V, ⊠ cluster VI,
▲ cluster VII, □ cluster VIII, Δ cluster IX, ■ cluster X, + cluster XI and ✳ cluster XII

number of grains per panicle, percentage of plant survival, and grain yield per plot.

Fig. 1 shows that the clusters were distinctly delineated to their respective positions that are similar to their position in the D^2 analysis (Tyagi *et al.*, 1999). In particular, genotypes G3 and G41 of Cluster X and G11 of Cluster IX were situated opposite to each other, indicating a considerable divergence between them. Just like the D^2 analysis, the PCA also clustered the genotypes based on their pedigree and not by their geographic origin. Thus, it was suggested that the selection of lines for hybridization programme should be based on genetic diversity rather than geographic distance.

CONCLUSIONS

Hierarchical and non-hierarchical algorithms, based on the multivariate statistical techniques, are common methods used by breeders to identify diverse genotypes for developing varieties that suit the target environment. Thus, it provides a chance to obtain recombinants resulting from recombination of favourable genes. In the present study, the results indicated that the genotypes selected from Clusters X, VI, VII, and XII could be used in hybridization programme. The genotypes from these clusters exhibited the maximum diversity with respect to the aggregate effects of the characters, such as earliness, high productive tillers per plant, longest panicle, maximum number of grains per panicle, high survival percentage, and high grain yield per plot. Therefore, the segregants from the above clusters would yield promising genotypes for salt affected soil with high survival percentage and yield.

The PCA and D^2 statistic exhibited a high level of variability among the genotypes and

it also allowed the selection of highly diverse genotypes which differ in their phenotype performance.

REFERENCES

Arunachalam, V. (1981). Genetic distances in plant breeding. *Indian Journal of Genetics and Plant Breeding*, *41*, 226-236.

Arunachalam, V. and Ram, J. (1967). Geographical diversity in relation to genetic divergence in cultivated sorghum. *Indian Journal of Genetics and Plant Breeding*, *27*, 369-380.

Babu, S., Anbumalarmathi, J., Sheeba, A., Yogameenakshi, P. and Rangasamy, P. (2005). Stability in performance of salt tolerant rice hybrids. *Oryza*, *42*, 222-224.

Chaturvedi, H.P. and Maurya, D.M. (2005). Genetic divergence analysis in rice (*Oryza sativa* L.). *Advances in Plant Sciences*, *18*, 349-353.

De, R.N. and Rao, A.V.S. (1987). Genetic divergence in rice under low land situation. *Crop Improvement*, *14*, 16-21.

Jagadev, P.N., Samal, K.M. and Lenka, L. (1991). Genetic divergence in rape mustard. *Indian Journal of Genetics and Plant Breeding*, *51*, 465-466.

Lal, S.K. and Rana, V.K.S. (2000). Genetic enhancement of soybean. *Indian Journal of Genetics and Plant Breeding*, *60*, 483-494.

Murty, B.R. and Arunachalam, V. (1966). The nature of genetic diversity in relation to breeding system in crop plant. *Indian Journal of Genetics and Plant Breeding*, *26*, 188-198.

Rao, C.R. (1970). *Advanced Statistical Method in Biometrical Research*. New York: John Wiley and Sons Inc.

Ratho, S.N. (1984). Genetic divergence in scented varieties of rice. *Indian Journal of Agricultural Sciences*, *54*, 699-701.

Rathor, A.G., Zargar, M.A. and Sheikh, F.A. (2001). Genetic divergence in rice (*Oryza sativa* L.) under temperate conditions. *Indian Journal of Agricultural Sciences*, *71*, 344-345.

Rieseberg, L.H., Archer, M.A. and Wayne, R.K. (1999). Transgressive segregation, adaptation and speciation. *Heredity*, *83*, 363-372.

Sarawgi, A.K. and Shrivastava, M.N. (1996). Genetic divergence in rice under irrigated and rainfed conditions. *Advances in Plant Sciences*, *9*, 93-100.

Sharma, A., Gupta, K.R. and Kumar, R. (2008). Genetic divergence in Basmati Rice (*Oryza sativa*) under irrigated ecosystem. *Crop Improvement*, *35*, 8-10.

Sharma, M.K. and Bhuyan, J. (2004). Genetic divergence in rice. *Advances in Plant Sciences*, *17*, 323-328.

Sinha, P.K., Prasad, C., Prasad, K. and Mandal, N.P. (2001). Genetic divergence in upland rice varieties. *Oryza*, *38*, 28-34.

Tyagi, P.C., Kumar, N. and Agarwal, M.C. (1999). Genetic divergence in early maturing cowpea. *Agricultural Science Digest*, *19*, 162-166.

Yeo, A.R. and Flowers, T.J. (1982). Accumulation and localization of sodium ions within the shoots of rice (*Oryza sativa* L.) varieties differing in salinity resistance. *Physiologia Plantarum*, *56*, 343-348.