

Original Research Article

<https://doi.org/10.20546/ijcmas.2020.911.043>

## K- Mean and Euclidian Cluster Analysis for Salt Tolerance Rice Genotypes under Alkaline Soil Condition

Ashutosh Kashyap\*, Vijay Kumar Yadav, Poonam Singh, P. K. Singh and Shweta

Department of Genetics and Plant Breeding, Chandra Shekhar Azad University of Agriculture & Technology, Kanpur- (U.P.) India

\*Corresponding author

### ABSTRACT

#### Keywords

Rice, genotypes, K-clustering, Euclidian clustering, Salt tolerance, Rice, Sodidity

#### Article Info

Accepted:  
04 October 2020  
Available Online:  
10 November 2020

An experiment was conducted to examine K- Mean Cluster and Euclidian Cluster analysis on 78 genotypes including seven standards (checks) varieties viz., CSR36, CSR10, CST7-1, CSR27, Usar Dhan 3 for salinity and alkalinity tolerant, while Sambha Sub1 as for general stress, and PUSA 44 as salt stress sensitive were grown in Augmented Randomized Block Design to selecting salt tolerance and breaking the yield barrier under alkaline soil condition. All genotypes were grouped into nine clusters by both k-Means Clustering, and Euclidian revealed the genotypes of heterogeneous origin were frequently present in same cluster. Low conformity was observed in placing of genotypes in both clustering techniques but it was provided important information on some genotypes which have common placing in both clustering pattern. In merit of mean yield performance, CSR -2016-IR-18-10 placed as highest second yielder followed by CSA -2016, CARI dhan 10, Usar Dhan 3 possessed 4<sup>th</sup>, 16<sup>th</sup> and 25<sup>th</sup> rank. These genotypes were considered with high yielder and more stable across the environments.

### Introduction

Rice is the most important staple food crop of the world. It is the principal food of half of the world's human population inhabiting the humid tropics and subtropics. World population is increasing rapidly by every passing year and there will be a need to produce 87% more of what we are producing today especially food crops such as rice, wheat, soy and maize by 2050 (Kromdijk and Long, 2016). Sodidity is one of the major soil constraints to crop production and is expected to increase due to global climate changes and as a consequence of many irrigation practices.

Clustering analysis is an important branch of data mining, and it is an active field. It is commonly used in data mining, clustering algorithm with hierarchical clustering method. The partitioning clustering based on the density clustering and grid clustering method analysis is based on specific requirements and rules to distinguish things and classification process. It belongs to the category of unsupervised classification by generic classification on the basis of the similarity between things. K-means algorithm is one of the most important algorithms in the field of clustering techniques. The subtlety of the

algorithm is simple, efficient, high and easy to handle data has been applied to many areas. However, K-means algorithm is very sensitive to initialize, the better center. This study was undertaken to determine the genetic diversity in salt tolerant rice genotypes for the maximum utilization of the genetic resources and proper selection of donor parents with using both K Cluster Mean and Euclidian cluster analysis.

## Materials and Methods

The experiment was conducted during year 2017 and 2018, at Crop Research Farm, Nawabganj and Seed Multiplication Farm Bojha, Chandra Sheker Azad University of Agriculture and Technology, Kanpur (U.P.) India on 71 rice genotypes and seven checks varieties *viz.*, CSR36, CSR10, CST7-1, CSR27, Sambha Sub1, Usar Dhan 3 for sodicity resistant and, and PUSA44 as salt stress sensitive in Augmented Randomized Block Design with replications of check under three environments taking into consideration of soil types and days of sowing. The details of the environments are given below: Environments: E-1: Environment I, Year 2017, high stress, pH 9.8, Ec 1.43  $\text{dsm}^{-1}$ , Seed Multiplication Farm, Bojha; E-2: Environment II, Year 2018, high stress, pH 9.8, Ec 1.41  $\text{dsm}^{-1}$ , Seed Multiplication Farm, Bojha; E-3: Environment III, Year 2018, Normal stress, pH 8.8, Ec 0.96  $\text{dsm}^{-1}$  CRF, Nawabganj.

Five plants in all genotype and checks were selected at random from each replication for recording of observations on characters of these genotype were used for recording all the below mentioned characters. The average of observations recorded on these five plants was considered for statistical analysis. Plant morphological characters of each genotype were recorded by selecting single or group of plants depending on all characters at different

stages of crop growth. Days to 50% flowering Plant height (cm), Total no of tillers  $\text{plant}^{-1}$ , Number of panicle bearing tillers  $\text{plant}^{-1}$ , Panicle Length (cm) Filled grain panicle  $^{-1}$ , Spikelet fertility percentage, 1000- grain weight (g), Stress score at reproductive stage and Grain yield  $\text{plant}^{-1}$ .

The genotypes were grouped into clusters based on Mahalanobis's D2 statistics and canonical variate analysis and K cluster mean analysis by K-means method (Hartigan and Wang, 1979; Lloyd, 1957; Mac Queen, 1967 on the basis of average distance of k-means and the accessions in each cluster were then analyzed for basic statistics.

## Results and Discussion

The aim of clustering is to provide measures and criteria that are used for determining whether two objects are similar or dissimilar. In present study, two types of clustering techniques k-Means Clustering and Hierarchical Euclidian clustering were used to characterization of genotypes based on genetic divergence for selection of suitable and diverse genotypes (Manju *et al.*, 2014). These procedures characterize genetic divergence using the criterion of similarity or dissimilarity based on the aggregate effects of a number of yield contributing important characters.

The k-means clustering algorithm is a centroid based approach using cluster distortion to decide when sufficient progress has been made but also can be restricted to a certain number of iterations (Hartigan and Wong 1979). Convergence of the algorithm is based on the change in distance of the mean cluster distance metric. This distance metric is often the squared Euclidean distance or squared normal distance between an observation and the centroid (Fig. 1–3).

Table.1 Mean performance of 78 genotypes for 10 characters in *Oryza sativa*

S. No.	Character	Days to 50% flowering	Plant height (cm)	Tillers/Plant	Productive Tillers/plant	Panicle Length (cm)	Filled grains/panicle	Spikelet Fertility (%)	Test Weight	Stress score at reproductive stage	Grain Yield g/plant
1	CARI Dhan 10	85.33	82.90	8.53	5.67	19.06	149.33	76.28	19.00	2.00	31.46
2	CARI Dhan 11	85.33	153.83	15.83	12.43	23.39	134.67	82.40	20.73	1.00	36.08
3	CARI Dhan 6	104.00	105.63	11.10	9.17	20.92	130.67	83.92	23.40	2.00	27.45
4	CR 2851-S-1-6-B-B-4	121.33	76.67	10.93	8.93	21.83	105.67	74.83	19.67	7.00	12.35
5	CR 2851-S-B-1-2B—1	96.67	92.07	15.03	11.27	19.55	122.67	75.10	23.47	3.00	27.00
6	CR 2851-S-B-1-B-B-1	127.33	104.53	14.33	10.67	21.07	96.00	71.54	22.68	7.00	8.58
7	CR 3437-1*-S-200-83-1	105.00	90.17	11.50	8.43	23.33	138.67	77.91	17.20	5.00	22.28
8	CR 3878-245-2-4-1	95.33	102.00	10.70	9.27	22.09	123.67	75.44	20.82	3.00	28.48
9	CR 3880-10-1-9-2-2-1	101.00	79.37	11.57	6.37	18.81	114.00	73.88	19.80	3.00	24.16
10	CR 3881-4-1-3-7-2-3	114.00	112.10	13.50	10.03	21.14	121.00	73.37	19.97	7.00	12.88
11	CR 3883-3-1-5-2-1-2	94.00	116.20	7.70	5.60	20.22	97.67	72.50	23.47	3.00	23.92
12	CR 3884-244-8-5-11-1-1	99.33	67.40	11.33	7.67	22.11	104.00	78.89	22.57	7.00	9.47
13	CR 3887-15-1-2-1	91.67	116.20	11.07	9.53	19.25	113.67	71.53	17.90	2.00	31.41
14	CR 3890-35-1-3-4	104.33	113.07	14.67	10.20	24.99	132.67	80.89	22.70	2.00	35.92
15	CR 3904-162-1-5-1	125.33	97.07	7.33	5.67	22.75	94.67	69.56	22.33	7.00	10.87
16	CR3881-M-3-1-5-1-1-1	94.33	98.10	13.33	10.20	22.52	115.00	71.22	20.28	3.00	30.96
17	CR3882-7-1-6-2-2-1	92.67	77.47	10.27	8.07	24.57	126.67	78.47	19.01	3.00	24.57
18	CR3884-244-8-5-6-1-1	100.67	96.68	15.10	10.27	22.13	111.67	77.62	18.82	7.00	14.58
19	CR3903-161-1-3-2	119.33	104.67	10.33	8.07	18.93	114.67	76.93	22.67	6.33	19.95
20	CSA 2016-IR18-6	91.00	113.87	13.17	11.93	20.87	149.00	87.20	22.83	1.00	36.72
21	CSAR 1604	92.33	78.23	10.33	7.60	19.82	122.67	73.28	24.60	3.00	27.81
22	CSAR 1610	91.33	90.20	11.77	9.00	20.43	112.67	68.92	23.73	2.00	29.88
23	CSAR 1628	105.33	104.51	9.93	7.90	19.67	78.67	67.50	24.03	5.00	18.94
24	CSAR1620	96.67	83.23	13.50	11.87	20.99	151.67	84.42	23.30	2.00	36.77
25	CSR 2016-IR18-1	96.67	75.53	12.00	9.33	21.61	119.67	76.69	21.54	7.00	16.41
26	CSR 2016-IR18-10	94.33	121.23	13.80	11.00	19.48	123.67	79.85	22.62	1.00	37.10
27	CSR 2016-IR18-11	108.33	103.07	13.37	10.17	19.75	120.67	76.64	20.83	5.00	18.28
28	CSR 2016-IR18-12	95.33	103.20	12.03	9.03	26.03	103.67	72.35	24.37	3.00	31.87
29	CSR 2016-IR18-14	114.00	77.28	12.87	10.93	19.90	108.33	77.49	18.97	6.67	12.34
30	CSR 2016-IR18-15	105.00	88.70	13.23	8.80	24.37	105.67	71.78	21.30	3.00	25.74
31	CSR 2016-IR18-16	105.67	84.03	12.43	7.67	24.97	64.67	65.84	22.48	3.00	19.41
32	CSR 2016-IR18-17	100.00	102.80	14.60	12.67	20.60	120.33	79.20	24.61	3.00	27.18
33	CSR 2016-IR18-18	99.67	103.83	10.77	8.60	24.58	115.00	77.35	24.37	3.00	30.28
34	CSR 2016-IR18-2	116.33	94.00	11.17	8.17	24.70	121.67	78.02	21.57	4.33	23.75
35	CSR 2016-IR18-3	107.33	85.03	11.93	9.93	20.91	128.00	85.47	24.65	3.00	26.07
36	CSR 2016-IR18-5	110.33	90.48	15.00	12.67	25.60	79.33	70.20	20.83	3.00	23.17
37	CSR 2016-IR18-7	95.33	101.23	13.07	9.37	21.04	99.00	83.14	22.43	5.00	20.38
38	CSR 2016-IR18-8	96.00	102.60	11.50	9.33	24.13	124.00	75.81	22.50	2.00	30.78
39	CSR 2016-IR18-9	81.33	75.57	13.90	11.00	18.70	98.33	74.85	22.13	5.00	12.61
40	CSR RIL-01-IR165	101.67	106.40	10.77	7.93	22.23	121.00	79.11	21.88	2.00	32.93
41	CSR-2748-197	84.00	102.07	17.87	13.93	22.75	120.33	77.19	20.77	2.00	30.88
42	CSR-2748-4441-193	77.00	125.23	9.43	7.10	23.41	118.67	82.54	23.50	1.00	35.22
43	CSR-2748-4441-195	95.00	82.93	14.60	11.97	22.46	119.67	75.47	18.73	2.00	27.26
44	CSRC(S)47-7-B-B-1-1	87.00	137.87	11.17	7.57	20.56	102.67	77.49	19.03	3.00	30.90
45	CSR-C27SM-117	92.00	96.27	12.93	10.73	18.25	105.00	72.18	24.90	2.00	33.95
46	IR 83421-6-B-3-1-1 CR 3364-S-2	96.33	91.20	14.77	11.20	20.69	119.00	76.51	23.53	3.00	21.30
47	IR10206-29-2-1-1	105.67	105.80	10.77	7.50	19.88	120.00	79.06	18.07	7.00	16.90
48	IR52280-117-1-1-3	93.33	128.30	12.87	9.40	25.31	116.00	72.27	22.55	2.00	34.62
49	IR84649-81-4-1-3B-CR3397-S-B-4	96.00	95.27	7.23	5.50	19.80	108.33	81.69	22.69	3.00	28.51
50	KR 15010	78.00	74.97	10.50	8.77	24.94	129.33	79.04	17.73	7.00	17.92
51	KR 15016	82.00	87.17	11.80	9.23	17.48	136.00	79.81	18.77	5.00	20.18
52	KR15006	78.67	90.83	15.77	12.77	25.92	134.00	83.16	25.23	3.00	29.33
53	KS -12	84.00	70.03	14.53	11.50	21.79	113.00	84.86	25.24	1.00	38.12
54	NDRK 11-20	86.00	94.50	16.13	12.03	21.04	123.67	76.29	23.37	2.00	29.08
55	NDRK 11-21	86.00	92.77	12.17	9.60	21.44	99.00	75.46	27.23	3.00	29.36

56	NDRK 11-22	94.33	100.23	12.93	10.27	22.07	103.00	67.10	24.67	3.00	20.91
57	NDRK 11-24	84.67	91.17	10.97	8.47	23.65	124.67	76.73	25.43	2.00	31.66
58	PAU 3835-12-1-1-1	103.33	86.94	16.00	11.67	20.74	121.33	82.45	22.00	5.00	22.03
59	PAU 3835-36-6-3-3-4	84.67	77.53	10.43	8.43	23.80	109.67	69.57	21.70	5.00	18.40
60	PAU 4254-14-1-2-2-2-4-1	115.33	101.40	12.43	9.83	21.26	114.33	72.53	21.07	5.00	16.38
61	PAU 5563-23-1-1	108.00	75.53	10.27	8.33	20.46	104.33	71.37	18.50	5.00	26.08
62	PAU 7114-3480-1-1-1-0	85.67	103.70	12.17	10.00	22.07	140.00	80.51	20.90	2.00	28.05
63	RAU 1397-14	104.67	85.23	15.33	12.60	19.60	111.00	79.52	20.40	2.00	32.86
64	RP 5440-302-100-7-6-3-2	110.00	66.40	13.53	9.27	19.98	113.00	79.09	18.07	5.00	16.25
65	RP 5687-420-111-5-4-2-1	115.00	100.87	14.13	8.40	22.63	140.33	81.22	19.70	7.00	18.23
66	RP 5694-36-9-5-1-1	109.00	107.07	16.80	13.53	23.13	113.00	76.76	20.85	5.00	17.80
67	RP-320-4-3-2-1	115.33	67.27	11.53	7.67	24.29	156.33	75.90	18.80	7.00	10.85
68	RP-5683-101-85-30-2-3-1	96.00	71.53	10.27	8.70	18.21	113.00	78.15	23.00	3.00	19.86
69	Sambha Sub1	120.00	87.97	15.43	12.83	18.71	128.00	77.64	22.58	7.00	19.59
70	TR 09027	87.00	63.87	13.33	11.10	19.60	87.33	72.42	18.63	5.33	13.05
71	TR 09030	95.67	87.53	11.60	9.20	21.21	122.33	83.21	21.65	2.00	35.03
72	CSR10 ©	88.67	77.03	16.93	14.73	19.29	117.33	81.76	22.03	1.78	30.43
73	CST7-1 ©	106.33	94.30	15.40	12.17	22.71	131.00	79.60	21.30	4.78	21.05
74	CSR36 ©	106.67	101.07	14.30	12.00	23.14	136.00	83.61	24.50	1.56	31.59
75	CSR27 ©	95.67	95.60	10.23	8.17	21.88	125.67	83.51	24.33	2.00	30.85
76	Sambha Sub1 ©	120.67	77.73	14.13	11.70	20.03	122.67	79.94	21.60	4.78	25.83
77	UsarDhan 3 ©	115.00	95.43	13.53	11.50	24.13	131.67	79.86	23.50	2.00	29.89
78	PUSA44 ©	93.67	79.23	10.27	8.17	20.10	115.00	76.51	19.60	7.00	18.21
	<b>Mean</b>	<b>99.20</b>	<b>94.12</b>	<b>12.56</b>	<b>9.74</b>	<b>21.63</b>	<b>117.44</b>	<b>77.04</b>	<b>21.77</b>	<b>3.75</b>	<b>24.73</b>
	<b>C.V.</b>	2.74	2.63	7.08	9.87	3.10	18.48	5.90	2.58	6.26	20.06
	<b>C.D. 5%</b>	4.38	3.99	1.43	1.55	1.08	35.00	7.32	0.90	0.38	8.38
	<b>C.D. 1%</b>	5.79	5.27	1.89	2.04	1.43	46.22	9.68	1.19	0.50	10.59

**Table.2 K - Clustering pattern of 78 salt tolerant rice genotype**

Group	n	Within SS	Cluster Members
1	10	20.117	CARI Dhan 10, CR 2851-S-B-1-2B-1, CR 3878-245-2-4-1, CR3881-M-3-1-5-1-1-1, CSAR 1628, CSR 2016-IR18-7, IR 83421-6-B-3-1-1 CR 3364-S-2B-14-2B-1, IR84649-81-4-1-3B-CR3397-S-B-4B-1, RP 5694-36-9-5-1-1, CST7-1 ©
2	14	52.643	CR 2851-S-B-1-B-B-1, CR 3437-1*-S-200-83-1, CR 3880-10-1-9-2-2-1, CR 3881-4-1-3-7-2-3, CR3884-244-8-5-6-1-1, CR3903-161-1-3-2, CSR 2016-IR18-11, CSR 2016-IR18-9, IR10206-29-2-1-1, KR 15010, KR 15016, PAU 3835-12-1-1-1, PAU 4254-14-1-2-2-2-4-1, RP 5687-420-111-5-4-2-1
3	4	6.400	CSAR 1610, CSAR1620, KS -12, Usar Dhan 3 ©
4	9	7.588	CARI Dhan 6, CSR 2016-IR18-17, CSR 2016-IR18-18, CSR 2016-IR18-8, CSR-C27SM-117, NDRK 11-20, NDRK 11-22, TR 09030, CSR27 ©
5	8	24.328	CR 3883-3-1-5-2-1-2, CR 3887-15-1-2-1, CR 3890-35-1-3-4, CSA 2016-IR18-6, CSR 2016-IR18-10, CSR RIL-01-IR165, CSR-2748-197, PAU 7114-3480-1-1-1-0
6	4	40.597	CARI Dhan 11, CSR-2748-4441-193, CSRC(S)47-7-B-B-1-1, IR52280-117-1-1-3
7	5	0.807	CSR 2016-IR18-12, KR15006, NDRK 11-21, NDRK 11-24, CSR36 ©
8	12	30.198	CR 2851-S-1-6-B-B-4, CR 3884-244-8-5-11-1-1, CR 3904-162-1-5-1, CSR 2016-IR18-1, CSR 2016-IR18-14, PAU 5563-23-1-1, RP 5440-302-100-7-6-3-2, RP-320-4-3-2-1, Sambha Sub1, TR 09027, Sambha Sub1 ©, PUSA44 ©
9	12	19.108	CR3882-7-1-6-2-2-1, CSAR 1604, SR 2016-IR18-15, CSR 2016-IR18-16, CSR 2016-IR18-2, CSR 2016-IR18-3, CSR 2016-IR18-5, CSR-2748-4441-195, PAU 3835-36-6-3-3-4, RAU 1397-14, RP-5683-101-85-30-2-3-1, CSR10 ©

**Table.3 K- Cluster mean for 9 clusters in salt tolerant rice genotypes**

Cluster	Days to 50% Flowering	Plant Height (cm)	Tillers Plant <sup>-1</sup>	Productive Tillers Plant <sup>-1</sup>	Panicle Length (cm)	Filled Grains Panicle <sup>-1</sup>	Spikelet Fertility (%)	1000 Seed Weight (g)	Stress at reproductive stage	Grain Yield (gm/plant)
<b>1 Cluster</b>	98.000	96.864	12.480	9.607	21.027	115.967	76.323	21.841	3.678	24.59
<b>2 Cluster</b>	104.024	94.520	12.802	9.314	20.771	119.738	76.918	20.102	5.810	17.50
<b>3 Cluster</b>	96.750	84.725	13.333	10.967	21.833	127.250	79.514	23.944	1.750	33.66
<b>4 Cluster</b>	95.926	98.778	12.422	10.019	21.630	118.852	77.619	23.755	2.333	29.50
<b>5 Cluster</b>	93.333	111.592	12.650	10.017	21.483	124.750	78.597	21.633	1.875	32.11
<b>6 Cluster</b>	85.667	136.308	12.325	9.125	23.169	118.000	78.675	21.455	1.750	34.21
<b>7 Cluster</b>	90.267	95.807	13.047	10.373	24.035	119.467	78.260	25.353	2.511	30.76
<b>8 Cluster</b>	109.278	75.996	11.914	9.300	20.949	113.250	75.860	20.238	6.315	15.94
<b>9 Cluster</b>	99.889	82.684	12.661	9.944	22.358	109.944	75.628	21.693	3.009	24.94

**Table.4 Cluster Member: Ward of salt tolerant genotypes**

Cluster	Number	Cluster Members
<b>1</b>	14	CARI Dhan 10,CR 3880-10-1-9-2-2-1,CR3882-7-1-6-2-2-1,CSR-2748-4441-195,RAU 1397-14,CSR10 ©,CSAR 1604,RP-5683-101-85-30-2-3-1,CSR 2016-IR18-3,CSAR 1610,CSR-C27SM-117,CSAR1620,TR 09030,KS -12
<b>2</b>	4	CSR 2016-IR18-15,CSR 2016-IR18-16,CSR 2016-IR18-5,CSR 2016-IR18-2
<b>3</b>	22	CARI Dhan 6,CSR RIL-01-IR165,CSR 2016-IR18-8,CSR27 ©,CSR36 ©,Usar Dhan 3 ©,CR 3883-3-1-5-2-1-2,IR84649-81-4-1-3B-CR3397-S-B-4B-1,CR 2851-S-B-1-2B—1,IR 83421-6-B-3-1-1 CR 3364-S-2B-14-2B-1,CSR 2016-IR18-17,NDRK 11-22,CR 3878-245-2-4-1,CR3881-M-3-1-5-1-1-1,PAU 7114-3480-1-1-1-0,CSR-2748-197 NDRK 11-20,CSR 2016-IR18-12,CSR 2016-IR18-18,NDRK 11-21,NDRK 11-24,KR15006
<b>4</b>	2	CARI Dhan 11,CSRC(S)47-7-B-B-1-1
<b>5</b>	6	CR 3890-35-1-3-4,IR52280-117-1-1-3,CSR-2748-4441-193,CSA 2016-IR18-6,CSR 2016-IR18-10,CR 3887-15-1-2-1
<b>6</b>	6	CR 2851-S-1-6-B-B-4,CSR 2016-IR18-14,RP-320-4-3-2-1,PAU 5563-23-1-1,RP 5440-302-100-7-6-3-2,Sambha Sub1 ©
<b>7</b>	9	CR 3437-1*-S-200-83-1,PAU 3835-36-6-3-3-4,CSR 2016-IR18-9,KR 15016,TR 09027,CSR 2016-IR18-1,PUSA44 © CR 3884-244-8-5-11-1-1,KR 15010
<b>8</b>	8	CR 2851-S-B-1-B-B-1,CR3903-161-1-3-2,CR 3904-162-1-5-1,Sambha Sub1,CR 3881-4-1-3-7-2-3,RP 5687-420-111-5-4-2-1,IR10206-29-2-1-1,CR3884-244-8-5-6-1-1
<b>9</b>	7	CSAR 1628, CSR 2016-IR18-7,CSR 2016-IR18-11,PAU 4254-14-1-2-2-2-4-1,RP 5694-36-9-5-1-1,PAU 3835-12-1-1-1,CST7-1 ©



**Table.5** Euclidean<sup>2</sup>: Cluster Distances: Ward of salt tolerant genotypes

	1 Cluster	2 Cluster	3 Cluster	4 Cluster	5 Cluster	6 Cluster	7 Cluster	8 Cluster	9 Cluster
1 Cluster	116.673	172.029	183.177	968.354	409.288	422.851	400.443	579.947	279.600
2 Cluster		51.753	168.866	945.709	419.343	295.620	337.771	371.517	164.089
3 Cluster			107.323	608.137	216.157	572.199	477.321	512.713	215.185
4 Cluster				178.996	279.854	1649.706	1332.879	1126.068	773.294
5 Cluster					133.85	1025.527	851.416	798.784	425.782
6 Cluster						101.840	216.753	278.966	275.870
7 Cluster							152.255	300.357	245.122
8 Cluster								125.463	160.785
9 Cluster									64.182

**Table.6** Cluster Mean of 10 traits for salt tolerant genotypes

	Days to 50% flowering	Plant height (cm)	Tillers/Plant	Productive Tillers/plant	Panicle Length (cm)	Filled grains/panicle	Spikelet Fertility (%)	Test Weight	Stress score at reproductive stage	Grain Yield g/plant
1 Cluster	94.476	81.929	12.436	9.852	20.385	121.881	78.276	22.147	2.270	2987.818
2 Cluster	109.333	89.302	12.958	9.325	24.908	92.833	71.462	21.546	3.333	2302.250
3 Cluster	94.818	99.394	12.527	9.932	22.339	119.955	77.437	23.424	2.525	2874.303
4 Cluster	86.167	145.850	13.500	10.000	21.975	118.667	79.945	19.883	2.000	3349.333
5 Cluster	91.944	119.650	12.500	9.861	22.219	125.611	79.046	22.017	1.500	3516.889
6 Cluster	114.889	73.481	12.211	9.472	21.083	118.389	76.436	19.267	5.907	1728.593
7 Cluster	89.741	76.826	11.674	9.126	21.298	115.333	76.187	19.986	5.926	1650.852
8 Cluster	115.917	101.210	12.617	9.179	20.905	115.792	75.867	20.852	6.917	1520.167
9 Cluster	106.143	99.789	13.857	10.662	21.186	111.143	76.945	21.789	4.968	1926.952

**Fig.1**

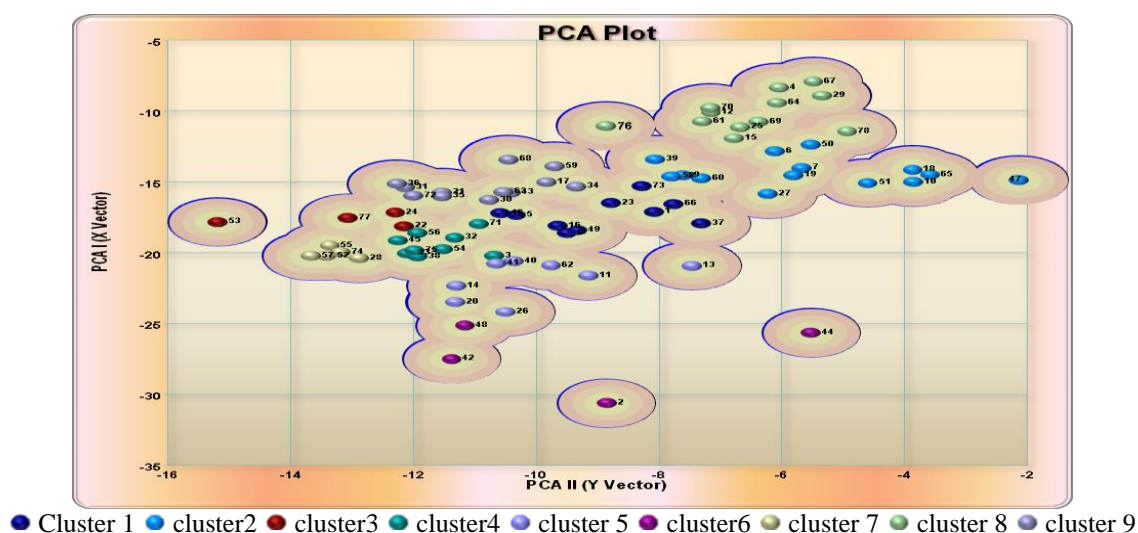


Fig.2

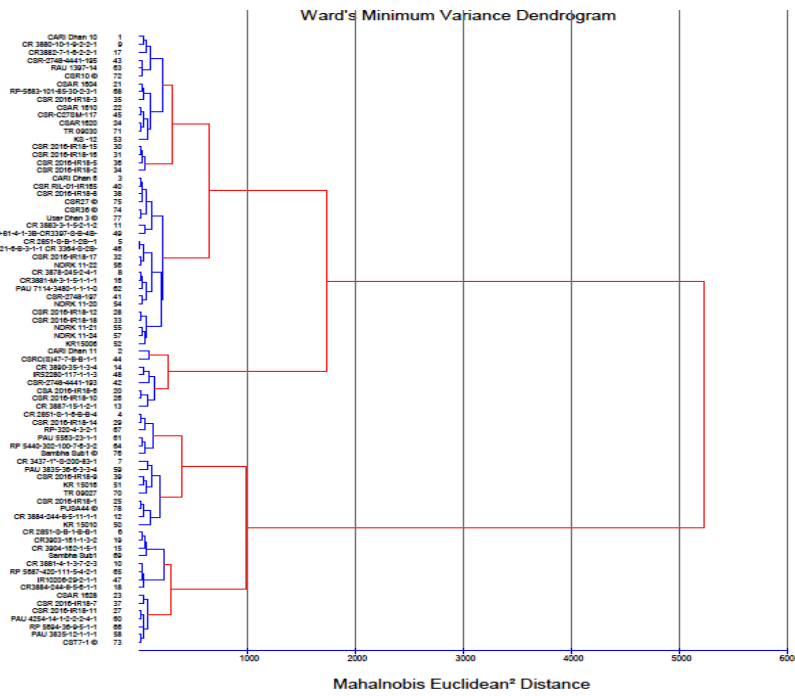
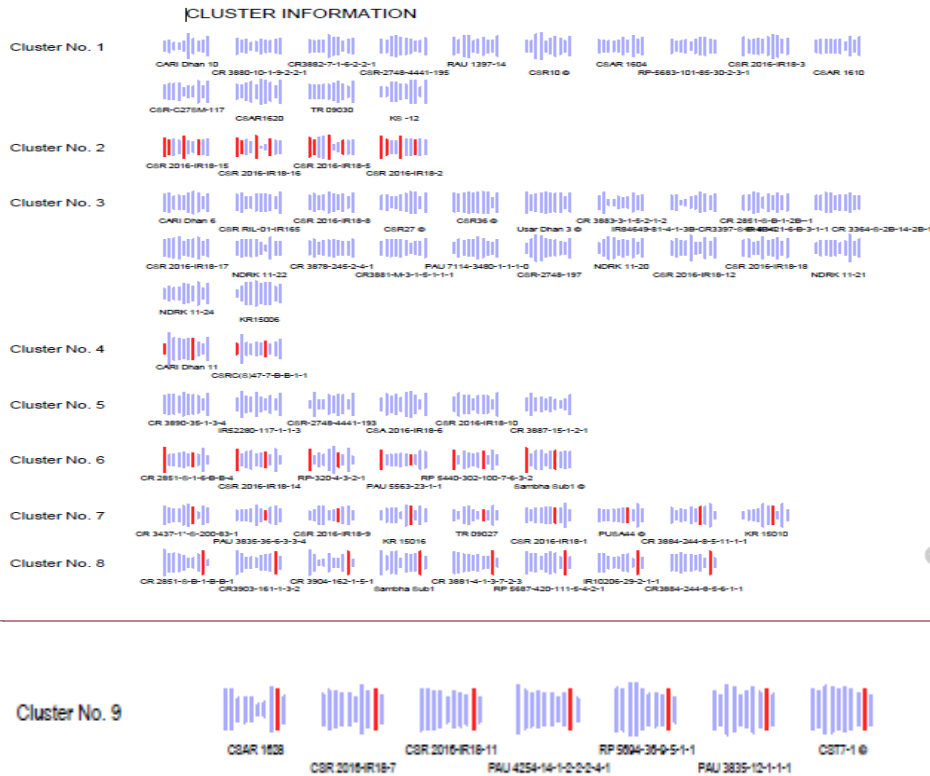


Fig.3



On the basis of difference within SS, seventy eight genotypes were grouped into nine clusters in the present study by both k-Means Clustering, and Euclidian revealed the genotypes of heterogeneous origin were frequently present in same cluster (Groenendyk *et al.*, 2014).

Although the genotypes originated in same place or geographic region were also found to be grouped together in same cluster, the instances of grouping of genotypes of different origin or geographical regions in same cluster were observed in case of all the clusters. k-Means Clustering showed that Cluster II, VIII, IX, I, IV, V consisted of 14, 12, 12, 10, 9 and 8 entries and Cluster III, IV and VII contains 4,4 and 5 genotypes, respectively, while in Euclidian, cluster V, III,I,IV,II comprised 23,19,15,9 and 7 entries, respectively. Although, cluster IV have equal numbers of entries but all the genotypes were different.

The average maximum inter cluster difference within SS values was observed between cluster II&VII followed by cluster II&III, cluster II &IV, cluster VI &VII, and cluster III & VI indicated great extent of diversity between these groups (Table 2 and 3). Cluster differences observed highest between cluster IV and six followed by cluster V and six. Therefore, it is suggested that any superior genotypes of cluster II and VI may be crossed with any superior genotype of cluster VII and III to produce desirable recombinants in hybridization programme and also revealed that the genotypes present in a cluster have little genetic divergence from each other with respect to aggregate effect of ten characters under study, while much more genetic diversity was observed between the genotypes belonging to different clusters Ranjbar *et al.*, (2007); Sapra and Lal (2003); Maqbool *et al.*, (2010) and Ahmadizadeh *et al.*, (2011).

A comparison of cluster mean for the studied characters indicated significant divergence between the groups. Some groups showed highest and other showed lowest value for the different characters in respect of the traits as fall in to different clusters in both types of cluster analysis. Low conformity was observed in placing of genotypes in both clustering techniques but it was provided important information on some genotypes which have common placing in both clustering pattern.

In cluster I genotype CARI dhan 10, cluster third Usar dhan 3 and cluster five CR 3890-35-1-3-4, CSA -2016 and CSR -2016-IR-18-10 are placed as common genotypes by both clustering pattern.

In merit of mean yield performance, CSR -2016-IR-18-10 placed as highest second yielder followed by CSA -2016, CARI dhan 10, Usar Dhan 3 possessed 4<sup>th</sup>, 16<sup>th</sup> and 25<sup>th</sup> rank. These genotypes were considered more stable across the environment (Table 1, 2 and 4).

In conclusion, it is clearly reflected wide variation from one cluster to another in respect of cluster means for ten characters, which indicated that genotypes having distinctly different mean performance for various characters were separated into different clusters (Table 5 and 6). Both clustering techniques have different results in placing of genotypes in respective cluster but it was provided important information on some genotypes which have common placing in both clustering pattern. The crossing between the entries belongings to cluster pairs having large difference within sum of square and possessing high cluster means for one or other characters to be improved may be recommended for isolating desirable salt tolerant rice lines.



## References

- Ahmadzadeh, M., Valizadeh, M., Shahbazi, H., Zaefizadeh, M. and Habibpor, M. 2011. Morphological diversity and interrelationships traits in durum wheat landraces under normal irrigation and drought stress conditions. *Adv. Environ. Biol.*, 5(7): 1934-1940.
- Derek Groenendyk Kelly Thorp Ty Ferre Wade Crow Doug Hunsaker 2014. A K-Means Clustering approaches To assess Wheat Yield Prediction Uncertainty with a HYDRUS -1D coupled crop model. *international Environmental Modeling and Software Society*.
- Manju Kaushik and Bhawana mathur 2014; comparative of K-Means and Hierarchical Clustering Techniques. *International journal of software & hardware research in Engineering*. Vol.2 Issue 6.
- Escobar-Hernandez, A., 2005; Troyo-dieguez, E., Garcia-hernandezcontreras, J.L., Murillo-amador, B. and Lopez-aguilar, R. Principal component analysis to determine forage potential of salt grass *Distichlis spicata* L. (Grrene) in coastal ecosystems of Baja Californiasur, Mexico. *Tech. Pec. Mex.*, 43: 13-25.
- Escobar-Hernandez, A., Troyo-dieguez, E., Garcia-hernandezcontreras, J.L., Murillo-amador, B. and Lopez-aguilar, R. 2005. Principal component analysis to determine forage potential of salt grass *Distichlis spicata* L. (Grrene) in coastal ecosystems of Baja Californiasur, Mexico. *Tech. Pec. Mex.*, 43: 13-25.
- Hartigan, J., and Wang, M. 1979. A K-means clustering algorithm. *Applied Statistics*, 28, 100–108.
- Kromdijk J, Long S P. 2016. One crop breeding cycle from starvation? How engineering crop photosynthesis for rising CO2 and temperature could be one important route to alleviation. *Proc Royal Soc B: Biol Sci*, 283: 20152578
- Lloyd, S. 1957. Least squares quantization in pcm. Bell Telephone Laboratories Paper, Marray Hill.
- MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. *Proc. 5th Berkeley Symposium*, 281–297.
- Mahalanobis, P C. 1930; On Test and Measures of groups divergence Part I. Theoretical Formulae *J.Asianic Sco. Bengal* 26, 541-586.
- Maqbool, R., Sajjad, M. and Khaliq, I. 2010. Morphological diversity and traits association in bread wheat (*Triticum aestivum* L.). *American-Eur. J. Agric. Environ. Sci.*, 8(2): 216- 224.
- R. Shivramakrishnan, R. Vinoth, Ajay Arora, G.P. Singh, B. Kumar and V.P. Singh, 2016 Characterization of wheat genotypes for stay green and physiological traits by principal component analysis under drought condition; *International Journal of Agricultural Sciences*, 12 (2) :245-251
- Sapra, R.L. and Lal, S.K. 2003. A strategy for selecting diverse accessions using principal component analysis from a large germplasm collection of soybean. *Pl. Genetic. Resour.*, 1: 151-156.

### How to cite this article:

Ashutosh Kashyap, Vijay Kumar Yadav, Poonam Singh, P. K. Singh and Shweta. 2020. K-Mean and Euclidian Cluster Analysis for Salt Tolerance Rice Genotypes under Alkaline Soil Condition. *Int.J.Curr.Microbiol.App.Sci*. 9(11): 359-367.  
doi: <https://doi.org/10.20546/ijcmas.2020.911.043>