Indian J. Genet., 55 (2): 138–147 (1995)

CLUSTERING OF RICE MUTANTS BY DIFFERENT METHODS OF ANALYSIS

K. C. MAHAPATRA,* C. H. P. MISHRA AND B. ACHARYA

Department of Plant Breeding and Genetics Orissa University of Agriculture and Technology, Bhubaneswar 751003

(Received: October 25, 1994; accepted: December 15, 1994)

ABSTRACT

Forty genotypes of rice comprising 34 mutants of Kalakeri, the parent variety, and five standard varieties were evaluated in two seasons with a view to group the entries into different clusters following various methods of analysis. Based on 13 characters, the genotypes were grouped into different clusters following D^2 analysis, canonical analysis, metroglyph analysis and numerical classificatory analysis. Clustering pattern in different methods indicated that out of 40 entries, 30 tall genotypes were grouped into 3 clusters in D^2 analysis, 6 clusters in metroglyph analysis, and 5 and 9 clusters in UPGMA method of numerical taxonomic approach at 80% and 85% phenons, respectively. The remaining 10 semidwarf strains were placed in 4 clusters by D^2 analysis, 6 clusters in metroglyph analysis, and 5 and 9 clusters in numerical taxonomic approach at 80% and 85% phenon levels, respectively. The study of clusters formed by different methods indicated that the numerical taxonomic approach was more potent for classificatory analysis of biological populations compared to other methods.

Key words: *Oryza sativa*, rice, multivariate analysis, D^2 static, canonical, metroglyph, and numerical classificatory analyses.

Genetic diversity is generally considered as an important criterion for choosing the parents for recombination breeding. A number of techniques involving multivariate analysis are used to measure it quantitatively. Mahalanobis D^2 statistic has been used extensively to quantify genetic diversity in a number of crop plants with diverse breeding systems [1–5]. In most of these studies, both D^2 statistic and canonical analysis were concurrently used and, in general, there was good agreement between the two methods of analysis in grouping the populations. The classification using generalized distance is workable when the number of entries is not very large [6]. While classifying large number of germplasm collections in rice, Vairavan et al. [3] used canonical analysis for initial grouping. But simple two-dimensional representation of multidimensional disposition of varieties cannot be as exact as the Tocher's method of grouping which scans the full

multidimensional space, even when the first two canonical vectors account for high proportion of variation [6]. On the other hand, as a simple method for initial grouping the metroglyph analysis of Anderson [7] has been used in many crop plants, such as, finger millet [8], potato [9], linseed [10], pearlmillet [11], rice [12–14], and sugarcane [15]. Here the groups are formed on the basis of two variables, which are plotted as X and Y coordinates in the scatter diagram. As the characters on X and Y axes determine grouping of the genotypes, it poses problem for judicious selection of these two characters. In the numerical classificatory analysis, the general similarity coefficient of Gower [16] has been used for clustering of populations at appropriate phenon level [17].

The classification of genotypes into different clusters following these methods are compared for determination of genetic closeness or divergence among the genotypes, more particularly the mutants with subtle differences. The efficient method(s) would be preferred by plant breeders for an efficient classificatory analysis of various crops. An attempt has been made in the present study to classify a number of mutant lines of Kalakeri, a traditional upland rice cultivar of Orissa, and to compare the aforementioned methods of classification for assessing diversity among the mutant lines, especially micromutants.

MATERIALS AND METHODS

The experimental material consisted of 40 rice varieties of which 34 were true breeding mutant lines of Kalakeri and 6 standard varieties including the parent. Besides Kalakeri, the tall standards were N 22 and Kalinga 3 and semidwarfs were Annapurna, Parijat and CR 143-2-2. Twenty two mutant lines (No. 1 to 22) were derived from 1% EMS and six each from 20 kR-gamma-ray (No. 23 to 28), and 30 kR X-ray (No. 29 to 34) treatments. The field trial was conducted in randomized block design with three replications at the Central Research Station, O.U.A.T., Bhubaneswar both in wet (kharif) and dry (rabi) seasons of 1988–89. The plot size was 3.75 x 0.75 m in both seasons. Recommended package of practices were followed to raise the crop.

Observations were recorded on five random competitive plants from each plot for days to heading, plant height, panicle length, panicle exsertion, panicle number, flag leaf length, breadth and area, spikelets and grains/panicle, 1000-grain weight, harvest index, and yield per plot. D^2 analysis and clustering by Tocher's method were done following [18]. Canonical analysis was carried out according to Anderson [19].

For metroglyph analysis, a method was devised to delineate the performance of the entries into three classes in respect of yield and six other characters influencing yield. The grand mean (m) and standard deviation (s) were calculated for each character. The performance of the entries for the seven characters was scored as high with mean more than

m + s, low with mean less than m-s, and entries with mean performance between m-s and m+s as medium. For grouping through metroglyph analysis, two variables plotted as X and Y coordinates in a scatter diagram, were chosen on the basis of higher contribution of the characters to divergence and high magnitude of significance in analysis of variance.

In the numerical classificatory analysis, the general similarity coefficient (S_G) of Gower [16] was used as a measure of resemblance between different operational taxonomic units or OTUs (entries included in the study). The S_G values were calculated following [17]. Based on the matrix of the S_G values, phenograms (dendograms) were constructed using the UPGMA (unweighted pair-group method using arithmetic average) technique in one of the SAHN (sequential, agglomerative, heirarchic, nonoverlapping) clustering methods [17], and the clusters were identified at appropriate phenon levels.

RESULTS AND DISCUSSION

Analysis of variance showed the existence of distinct genetic difference among the mutants as well as standard varieties in respect of all the 13 characters (Table 1). The differential response of genotypes to change in season was revealed from the significance of interaction components. The significant differences between mutant lines and their interaction with the season warranted grouping of the entries to identify the genetically diverse ones to ensure success in recombination breeding. Further, the differences due to seasons widened the spectrum of environmental conditions for assessing genetic divergence among the entries. The bias in clustering of genotypes in different environments, presumably due to differential response of different characters, could be eliminated by pooled analysis.

In the multivariate analysis, 13 characters contributed differently to the total D^2 for each pair of varieties. The major contributors to genetic diversity for plant height (24.6%) and 1000-grain weight (18.3%), accounting for 42.9% of the total divergence. Following the Tocher's method, the 40 entries were grouped into seven clusters and the clustering was in board agreement with the groupings obtained by using the first two canonical vectors (Fig. 1). Semidwarf entries formed four clusters of which cluster I comprised seven entries, five of which were dwarf mutants and the other two standards, while clusters IV, V and VI had single variety each. The tall genotypes constituted the remaining three clusters of which cluster II included 28 entries (25 mutants, 2 standards and the parent) and the other two (clusters III and VII) were monotypic.

As breeding programme aims at developing varieties with higher level of production, the performance of the entries in respect of seven important characters including yield was scored as high, medium and low. Delimitation of the three classes was done on the basis of

grand mean of the entries and its standard deviation. Groups are formed on the basis of two variables, which are plotted as X and Y coordinates in a scatter diagram. The selection of these two variables is very important. In the present study plant height and 1000-grain weight were choosen owing to their higher contribution in D^2 as well as principal component analyses. Contribution of such variables to total divergence has also been reported in rice as well as other crops [3, 20–23]. In spite of the limitation of choosing these two variables, this technique has its own importance in initial grouping. Means of such preliminary groups can further be used for grouping through multivariate analysis. The grouping through metroglyph analysis showed overall similarity with those of D^2 analysis (Fig. 2). The scatter diagram resulted in 12 groups, most of which are superimposed in the configuration on the basis of D^2 analysis. In this classification, clusters I and II of D^2 analysis were further divided into three and four subclusters, respectively, while variety 14 of cluster I and variety 9 of cluster II entered cluster VI and cluster VII, respectively. On the other hand, the monotypic clusters III, IV and V remained unchanged. Variety 6 of cluster I variety 39 of cluster II formed single-variety subclusters.

In the numerical taxonomic approach, the genotypes were

Table 1. Analysis of variance (MSS) for 13 characters in 40 rice genotypes pooled over two seasons

Source	d.f.	Days to heading	Plant height	Panicle length	Panicle exsertion	Panicle number	Flag leaf			Spikelet grain number	Grain number	1000-grain weight	Yield/ha	Harvest index
							length	breadth	area					
Replications/season	4	4.7	17.7	2.5	1.5	11524.0	0.1	0.02	3.0	11.0	6.2	1.1	10.9	2.8
Seasons (S)	1	19135.1**	33644.4**	462.9**	235.8**	32646190.0**	7243.5**	1.48**	7908.6**	42246.3**	10609.5**	7.9**	4322.3**	251.1**
Genotypes (G)	39	66.3**	1314.0**	26.1**	33.5**	72833.9**	75.1**	0.07**	86.9**	501.3**	340.5**	79.8**	99.2**	93.3**
Mutants (M)	33	59.9**	1435.8**	25.2**	33.7**	75727.0**	80.9**	0.06**	92.1**	556.7**	378.6**	83.0**	1103.9**	104.2**
Standards (V)	5	119.1**	763.0**	14.6**	21.1**	17320.0**	50.6**	0.08**	56.4**	206.9**	154.1**	53.4**	83.6**	39.5**
M vs V	1	12.5**	49.6**	114.1**	87.1**	254928.0**	7.1*	0.07**	67.0**	143.8*	15.5**	107.8**	24.7**	3.1
G x S	39	41.2**	148.4**	4.4**	5.8**	55618.5**	36.9**	0.03**	48.6**	500.8**	299.1**	7.5**	34.7**	38.5**
M x S	33	42.5**	120.9**	3.5**	6.3**	55540.8**	29.9**	0.04**	44.8**	525.1**	325.1**	6.7**	30.4**	43.1**
V x S	5	21.5**	348.5**	11.0**	3.2**	10199.2	84.9**	0.01**	67.7**	439.4**	186.7**	4.2**	55.2**	11.6**
M vs V x S	1	96.9**	53.4**	2.0	0.04	285276.0**	26.4**	0.05**	80.6**	8.2	5.9	52.1**	74.7**	21.4**
Error	156	1.1	6.2	0.6	0.5	5853.7	1.8	0.003	2.1	29.1	18.1	0.3	2.0	2.4

*,** Significant at the 5% and 1% levels, respectively.

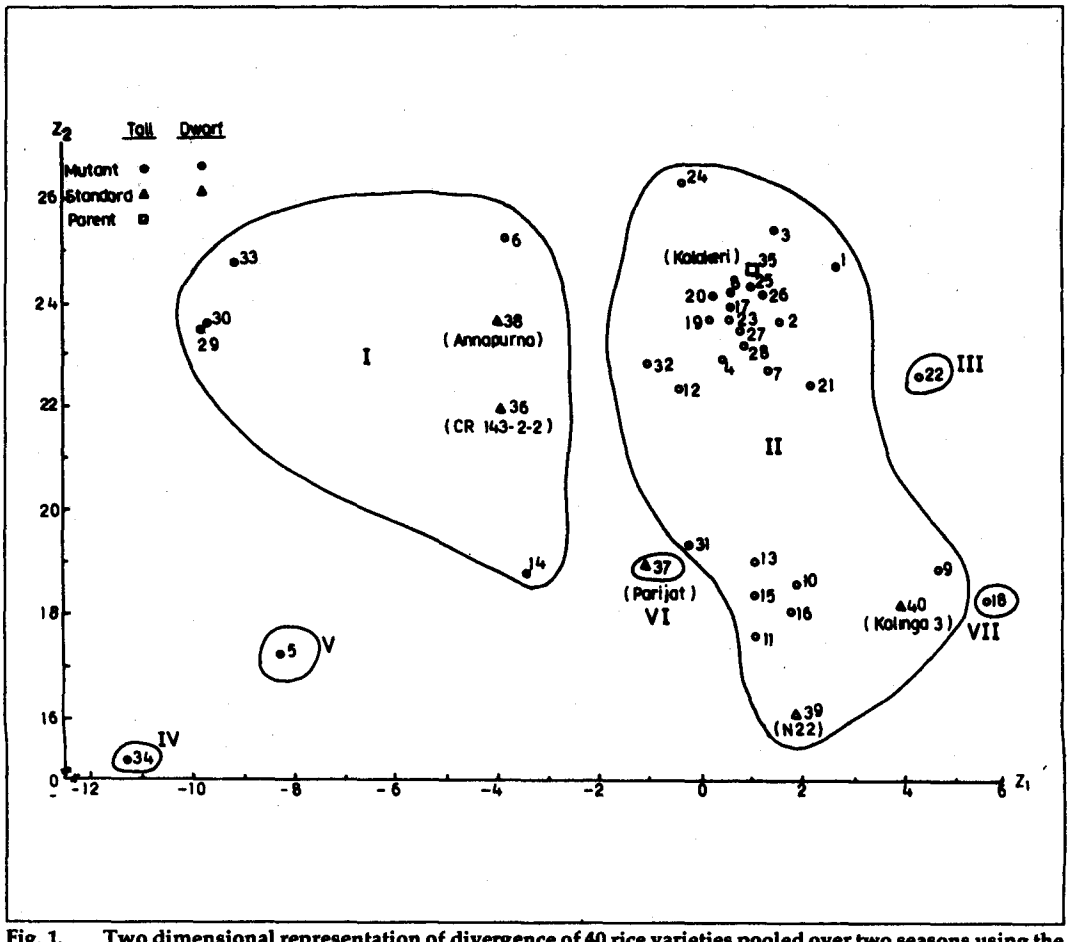

Fig. 1. **Two dimensional representation of divergence of 40 rice varieties pooled over two seasons using the first two canonical vectors (Z_1 & Z_2) as coordinates, the grouping obtained by D^2 analysis super imposed.**

classified using all the 13 characters to calculate similarity coefficients (S_G). The dendrogram showing clusters at the appropriate phenon levels indicates that at 70% phenon all the entries could be broadly classified into three clusters, viz. cluster I comprising five entries, cluster II consisting of 34, and cluster III with one (Fig. 3). When the phenon line was drawn at 75% the two multivariety clusters were further classified into two groups each, i.e., IA consisting of 4, IB having 1, IIA with 10, and IIB with 24 genotypes, thus increasing the number of genotypic constellations from three to five. Increase in the phenon level to 80% and 85% further dissociated the multivariety clusters into several subclusters, thus enhancing the number of clusters to 10 and 18, respectively (Table 2). Clusters IA, IIA and IIB were divided into subclusters and groups at 80% and 85% phenon levels, respectively.

Four genotypes in cluster IA were grouped into two subclusters, i.e. IA_1 consisting of three and IA_2 with one genotype at 80% phenon. Subclusters IA_1 was further divided into two groups $IA_{1(a)}$ having two genotypes and $IA_{1(b)}$ with one at 85% phenon. Clusters IIA and IIB were split into three subclusters each, i.e., IIA_1 with 7, IIA_2 with 2, IIA_3 with 1, IIB_1 with 19, IIB_2 with 3 and IIB_3 with 2 genotypes. Subclusters IIA_1 and IIA_2 were further divided into two groups each at 85% phenon, thus producing three monotypic clusters subcluster IIB_1 produced four groups one of which was a monotypic cluster and IIB_2 comprising three genotypes was dissociated into three monotypic clusters. It was thus possible to discern subtle differences between genotypes grouped in different clusters and/or subclusters at different phenon levels. It can be seen from Fig. 3 and Table 2 that the number of monotypic clusters identified was 1, 2, 4 and 12 at 70%, 75%, 80% and 85% phenon levels, respectively. The dwarf mutants of Kalakeri Nos. 5, 6, 14, 33 and 34 formed monotypic clusters while

Table 2. Composition of clusters/subclusters identified from the dendrogram at different phenon levels

No. of clusters/subclusters at different phenon levels				Variety (S. No.)
70%	75%	80%	85%	
I (5)	IA	IA_1	IA_{1a}	29,30
	(4)	(3)	IA_{1b}	33
		IA_2	IA_2	5
		(1)		
	IB	IB	IB	34
	(1)			
II (34)	IIA	IIA_1	IIA_{1a}	11, 13, 10, 15, 16, 31
	(10)	(7)	IIA_{1b}	39
		IIA_2	IIA_{2a}	14
		(2)	IIA_{2b}	37
		IIA_3	IIA_3	40
		(1)		
	IIB	IIB_1	IIB_{1a}	20, 23, 8, 27, 19, 28, 4, 17
	(24)	(19)		
			IIB_{1b}	3
			IIB_{1c}	7, 35, 26, 1, 2, 12, 25, 21
			IIB_{1d}	24, 32
		IIB_2	IIB_{2a}	6
		(3)	IIB_{2b}	36
			IIB_{2c}	38
		IIB_3	IIB_3	9, 18
		(2)		
III (1)	III	III	III	22

*Figures in parentheses indicate number of entries in the respective clusters/subclusters.

Nos. 29 and 30 remained together in a single cluster because of their close similarity for most of the characters studied. Among the 27 tall mutants, Nos. 3 and 22 formed monotypic clusters while Nos. 9 and 18 as well as 24 and 32 remained together in two separate clusters. Among the remaining 21 tall mutants forming three multi- variety clusters only seven

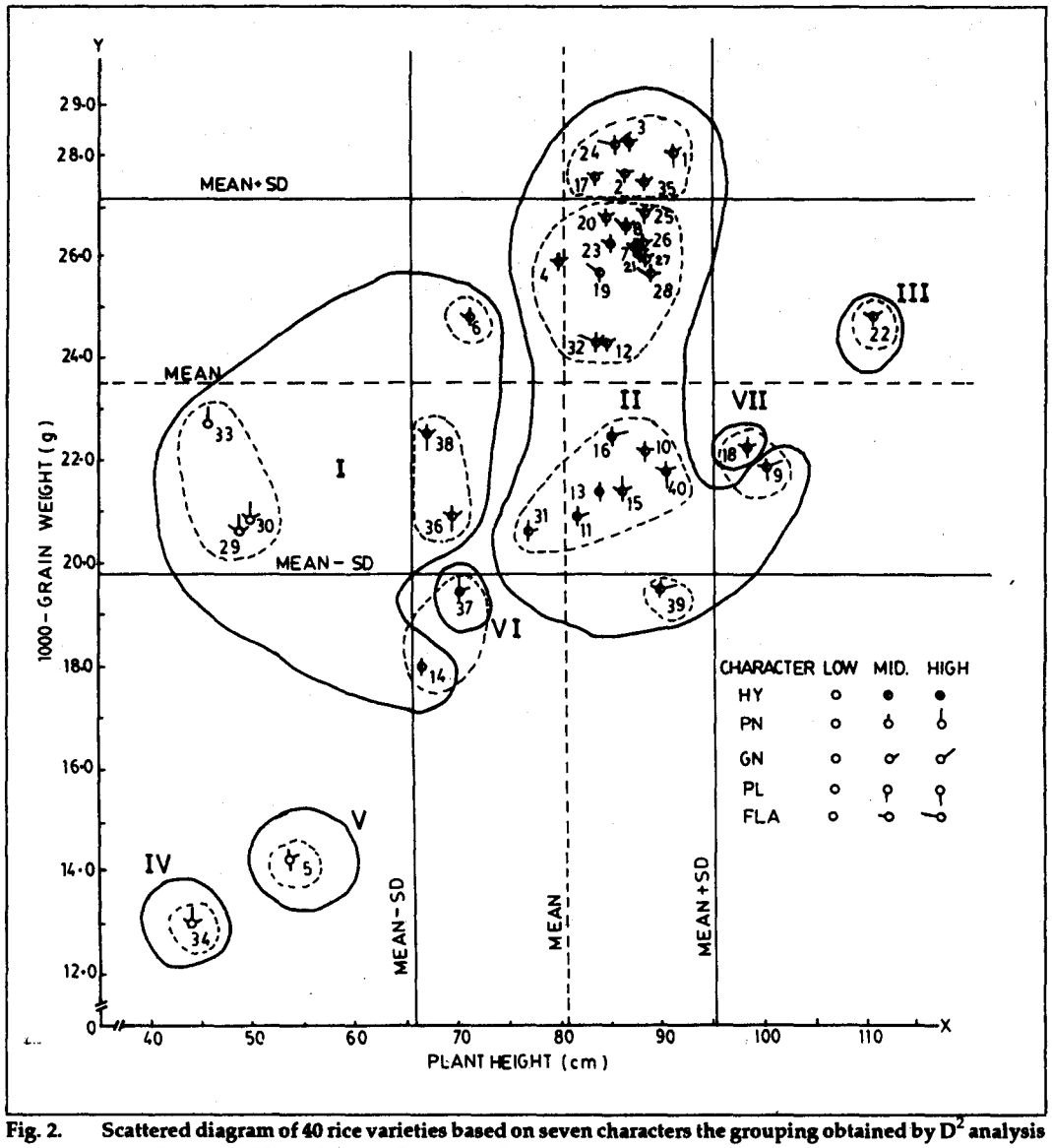

Fig. 2. **Scattered diagram of 40 rice varieties based on seven characters the grouping obtained by D^2 analysis superimposed (HY—yield/ha, PN—panicle number, GN—grain number, PL—panicle length, and FLA—flag leaf area).**

mutants were included in one cluster, i.e. $IIB_{1(c)}$ along with the parent (No. 35). This indicates genetic differences among the tall mutants and also distinction of 20 tall mutants from their parent.

Table 3. Number of clusters formed by different methods

Varieties (Nos.)	D^2	Metro-glyph	N T at phenon level	
			80%	85%
Tall (30)	3(2)	6(2)	5(2)	9(4)
Dwarf (10)	4(3)	6(3)	5(2)	9(8)
Total (40)	7(5)	12(5)	10(4)	18(12)

*Number of monotypic clusters given in parentheses.

Plant height being the major contributor to total divergence, the dwarf mutant (Nos. 5, 29, 30, 33 and 34) were found to be different from the dwarf varieties used as standards (Nos. 36, 37 and 38). This indicates that the genes responsible for dwarfism in these mutants presumably differ from those present in the semidwarf standard varieties. Similarly, three tall mutants (Nos. 9, 18 and 22) were quite different from the other tall mutants as well as the parent variety (No. 35) and the standard tall genotypes (Nos. 39 and 40). This differentiation may be due to subtle changes in quantitative characters like plant height, 1000-grain weight etc. resulting from the micromutations affecting character expression.

A comparison of the clustering patterns based on different methods shows that the 30 tall entries were grouped into 3 clusters in D^2 analysis, 6 in metroglyph analysis, and 5 and 9 clusters in UPGMA method of numerical taxonomic approach at 80% and 85% phenons, respectively (Table 3). Similarly, the 10 semidwarfs fell in 4 clusters by D^2 analysis and 6 by

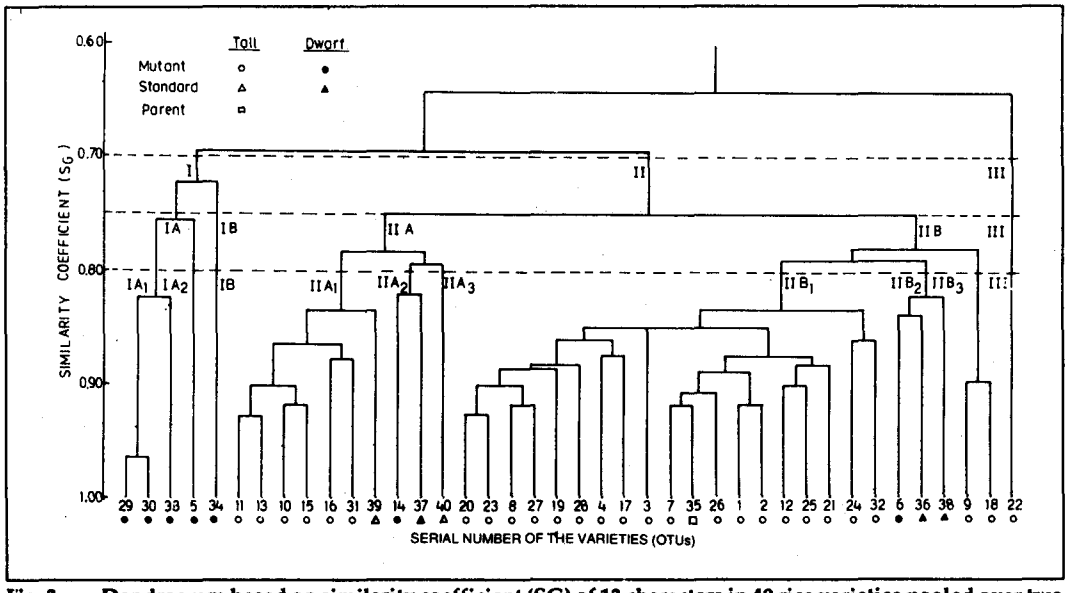

Fig. 3. Dendrogram based on similarity coefficient (SG) of 13 characters in 40 rice varieties pooled over two seasons.

metroglyph analysis, while in the numerical classificatory analysis they formed 5 and 9 clusters at 80 and 85% phenon levels, respectively. It is, therefore, evident that the numerical taxonomic approach for classification of biological population into different groups is more potent to distinctly discriminate mutant lines as well as the standard varieties for their use in recombination breeding programme. However, metroglyph analysis seems to be easy and simple, and can be used for initial grouping when the number of collections is large.

REFERENCES

1. B. R. Murty and V. Arunachalam. 1966. The nature of divergence in relation to breeding system in some crop plants. Indian J. Genet., **26A**: 188–198.

2. G. M. Bhatt. 1970. Multivariate analysis approach to selection of parents aiming at improvement in self-pollinated crops. Aust. J. Agric. Res., **21**: 1–7.

3. S. Vairavan, E. A. Siddiq, V. Arunachalam and M. S. Swaminathan. 1973. A study on the genetic divergence in rice from North-East Himalaya. Theor. Appl. Genet., **43**: 213–221.

4. J. Lee and P. J. Kaltsikes. 1973. The application of Mahalanobis' generalized distance to measure genetic divergence in durum wheat. Euphytica, **22**: 124–131.

5. S. Ramanujam, A. S. Tiwari and R. B. Mehra. 1974. Genetic divergence and hybrid performance in mungbean. Theor. Appl. Genet., **45**: 211–214.

6. V. Arunachalam. 1981. Genetic divergence in plant breeding. Indian J. Genet., **41**: 226–236.

7. E. Anderson. 1957. A semigraphical method for the analysis of complex problems. Proc. Natl. Acad. Sci. Wash., **43**: 923–927.

8. G. P. Ranga Rao. 1963. Studies on Variation, Heritability and Correlation of Components of Yield in Ragi. M.Sc (Ag.) Thesis. Orissa University of Agriculture & Technology, Bhuabneswar.

9. S. K. Sinha and Pushkarnath. 1964. The relationship of Indian potato varieties to *Solanum tuberosum* subspecies *andigena*. Indian Potato J., **6**: 24–29.

10. S. Chandra. 1977. Comparison of Mahalanobis' method and metroglyph technique in the study of genetic divergence in *Linum usitatissimum* L. germplasm collection. Euphytica, **26**: 141–148.

11. M. B. L. Saxena, G. V. Subba Rao, G. K. Gupta and R. C. Verma. 1978. Bajra germplasm of arid areas of Rajasthan. Indian J. Genet., **38**: 390–395.

12. K. C. Kotaiah, C. P. Rao, N. S. Reddi and Y. R. B. Sarma. 1986. Mahalanobis' D^2 and metroglyph analysis, in mid-duration genotypes of rice. Indian J. agric. Sci., **50**: 151–156.

13. J. S. Chauhan, V. S. Chauhan, P. K. Sinha and K. Prasad. 1989. Analysis of in situ variabiloity for some panicle and grain characters in native germplasm of rice. Oryza, **26**: 243–249.

14. P. K. Sinha, K. Prasad, V. S. Chauhan and J. S. Chauhan. 1990. Studies on Gora rices of Bihar. II. Variability pattern in agro-morphological characters. Indian J. Genet., **50**: 111–115.

15. R. B. Katiyar, M. L. Sharma and M. M. S. Saxena. 1988. Metroglyph analysis of geographical complexes in sugarcane. Indian J. Genet., **48**: 25–30.

16. G. C. Gower. 1971. A general coefficient of similarity and some of its properties. Biometrics, **27**: 857–871.

17. P. H. A. Sneath and R. R. Sokal. 1973. Numerical Taxonomy. The Principles and Practice of Numerical Classification. W. H. Freeman and Company, San Francisco–New York.

18. Rao C. R. 1952. Advanced Statistical Methods in Biometrical Research. 1st edn. (reprinted with corrections 1970). Hafner Press, New York.

19. T. W. Anderson. 1958. An Introduction to Multivariate Statistical Analysis. John Willey and Sons, New York.

20. G. R. Das and D. N. Borthakur. 1973. Genetic divergence in rice. Indian J. Genet., **33**: 436–443.

21. V. P. Singh, M. S. Swaminathan, R. B. Mehra and E. A. Siddiq. 1979. Divergence among dwarfs of cultivated rice. Indian J. Genet., **39**: 315–322.

22. K. S. Kanwal, R. M. Singh, J. Singh and R. B. Singh. 1983. Divergent gene pools in rice improvement. Theor. Appl. Genet., **65**: 263–267.

23. K. C. Mohapatra. 1988. Genetics of Agronomic Traits in Rice: Studies on Stability, Genetic Divergence, Combining Ability and Heterosis in Early Rice Varieties. Ph. D. Thesis. Orissa University of Agriculture and Technology, Bhubaneswar.