Indian J. Pl. Genet. Resources 12(1): 50-55, 1999

CHARACTERIZATION OF CITRONS GROWING IN NORTH EAST INDIA USING HIERARCHICAL CLUSTER ANALYSIS

B. K. RAY AND P. C. DEKA, Department of Agricultural Biotechnology, Assam Agricultural University, Jorhat 785 013 (Assam)

Eight different citron biotypes collected from different parts of north east were utilised for hierarchial agglomerative cluster analysis using sixty-five morphological characters. Euclidean distance between the genotypes were measured by re-scaled reversed absolute squared Euclidean similarity coefficient matrix. The cluster analysis of the Euclidean similarity co- efficient matrix was performed by five different methods. Three clusters were formed in all the five methods utilised for dendrogram construction. The first cluster includes the genotypes Seuli (Barapani), Gondharaj (Nowgaon and Agartala collections). The genotypes Haijange (Imphal), Bira Jora (Titabor) and Jora Tenga. (Margherita) falls in the second cluster. The third cluster includes the genotypes Bira Jora (Jorhat and Teok collections). Existence of variability in the morphological characters confirms that north east India is one of the centre of origin of the citrons.

Key words : Citron, *Citrus medica*, cluster analysis, germplasm evaluation, numerical taxonomy, origin

The north eastern parts of India which falls under the sub- Himalayan ranges are the principal areas of citrus diversification. Citron (*Citrus medica* L.) is one of the ancient species of this region. It is monoembryonic and available in diverse forms (Scora, 1988). The citron was first introduced into the Mediterranean areas and brought under cultivation (Tolkowsky, 1938). Very little is known about the cross and self fertility in the citron. Limited observations suggest that the citron produces vigorous selfed seedlings and appears to be highly homozygous (Barret and Rhodes, 1976).

The taxonomy of citrus species are very much complicated and confusing (Swingle, 1943, Tanaka, 1954). Based on numerical taxonomy (Barret and Rhodes, 1976), isozyme analysis (Torres *et al.*, 1978), RFLP analysis (Green *et al.*, 1986) and RAPD markers (Suguwara *et al.*, 1995), it has been proposed that citron (*Citrus medica* L.), Pummelo (*Citrus grandis* Osbeck) and mandarin (*Citrus reticulata* L. lanco) are the basic species of citrus. Though north east India is one of the major centre of origin for *Citrus* species, unfortunately very little attention has been given for the genetical characterization of citrus species growing in this area.

Numerical taxonomic analysis (Sneath and Sokal, 1973) utilising morpholoical characters can generate important information on citrus germplasm variability and can effectively be utilised alongwith the modern biochemical and molecular biological techniques.

Among the techniques available for numerical taxonomic study, heirarchical cluster analysis can readily be used to asses relatedness and distance of any type of sample characterised by any type of descriptors (Peeters and Martinelli, 1989). It may be used to asses the genetic similarity and

dissimilarity in germplasm collection and the technique could also have application for the selection of parental lines for which varying degrees of segregation are sought.

In this paper, genetic relatedness among the eight different citron biotypes utilising hierarchical agglomerative cluster analysis has been reported.

MATERIALS AND METHODS

Eight different citron genotypes were collected from the different parts of north east India and a list of the collected genotyppes are presented in the Table 1. Sixty five morphological characters (Table 2) of the tree, leaves, flowers, fruits (both external and internal characters) and seeds were studied.

Table 1. List of the collected Citrus germplasm

Germplasm		Place of Collection
1.	Haijange	Imphal, Manipur
2.	Jora Tenga	Margherita, Assam
3.	Gondharaj	Agartala, Tripura
4.	Seuli	Barapani, Meghalaya
5.	Bira Jora	Jorhat, Assam
6.	Bira Jora	Teok, Assam
7.	Bira Jora	Titabor, Assam
8.	Gondharaj	Nowgaon, Assam

Numerical taxonomic studies were carried out using morphological characters. Weightage of the morphological characters were given according to the recommendations of IBPGR (1988). Morphological characters and their status (weightage) used for numerical taxonomic study are given in Table 2. Hierarchical Agglomerative Cluster Analysis were performed and dendrograms were prepared. A breif account of the procedures followed are described below.

Hierarchical Agglomerative Cluster Analysis: It is a special class of log-linear model. If a term for the interaction of a set of variables exists, there must be lower-order terms for all possible combinations of these variables. For example, if the term A by B by C is in the model, then the terms A, B, C, A by B, A by C, and B by C must also be in the model.

For performing hierarchical cluster analysis, first Euclidean distance is measured and clusters are made by several methods using corresponding agglomeration schedule. This schedule are being prepared on the basis of Euclidean similarity or dissimilarity distance coefficient matrix. The values of Euclidean dissimilarity matrix were found to be in meaningless scales. Hence, the values were transformed using the following procedure:

(i) Absolute values were taken for the values of the distances. (This transformation generally used when the sign indicates the direction of the relationship, but only the magnitude of the relationship is of interest).

(ii) Dissimilarities values were changed to similarity values. (This transformation generally used to reverse the ordering of the distances by negating the values.)

(iii) The distance values were rescaled to 0 to 1 range.

By transforming the values of the dissimilarity matrix, "Rescaled Reversed Absolute Squared Euclidean Similarity Coefficient Matrix" were prepared. Using this matrix, values the clustering analysis were performed by five methods *viz.*, average linkage (Between groups), average linkage (within group), single linkage, complete linkage and median method.

Table 2. List of citrus characters and their weightage used for numerical taxonomic study

I. TREE

i. **Tree size** : 1. Small 2. Medium 3. Large

ii. **Spines** : 1. Absent 2. Few 3. Many

iii. **Tree habitat** : 1. Upright 2. Spreading 3. Drooping 4. Weeping

iv. **Shoot tip surface** : 1. Glabrous 2. Pubescent 3. Densely pubescent

v. **Shoot tip colour** : 1. Green 2. Purple

II. LEAF

vi. **Vegetative life cycle** : 1. Evergreen 2. Semi-deciduous 3. Deciduous

vii. **Type of leaf** : 1. Simple 2. Trifoliate

viii. **Leaf form** : 1. Sessile 2. Brevipetiolate 3. Longipetiolate

ix. **Petiole wing** : 0. Absent 3. Narrow 7. Broad

x. **Shape of petiel wing** : 1. Cordiform 2. Deltoid 3. Obovate

xi. **Leaf margin** : 1. Crenate 2. Dentate 3. Entire 4. Wavy

xii. **Length of spine at leaf axil** (mm)

xiii. **Shape of spine** : 1. Curved 2. Straight

xiv. **Leaf length** (mm)

xv. **Leaf width** (mm)

xvi. **Leaf shape** : 1. Elliptic 2. Ovate 3. Obovate 4. Lanceolate

xvii. **Leaf apex** : 1. Acute 2. Obtuse 3. Round

xviii. **Leaf colour** : 1. Light green 2. Green 3. Dark green

III. FLOWER

xix. **Arrangement of flowers** : 1. Solitary 2. In an inflorescence

xx. **Position of flowers or inflorescence** : 1. Axillary 2. Terminal

xxi. **Type of inflorescence** : 1. Panucle 2. Raceme 3. Cormb

xxii. **Number of flower buds per inflorescence**

xxiii. **Colour of flower buds** : 1. Greenish 2. hite 3. Yellow 4. Purple 5. Pink

xxiv. **Length of Colour of flower buds pedicel** (mm)

xxv. **Length of flower bud** (mm)

xxvi. **Colour of open flowers**: 1. White 2. Yellow 3. Purple

xxvii. **Length of petal** (mm)

xxviii. **Width of petal** (mm)

xxix. **Number of stamens**

xxx. **Length of anther** (mm)

IV. FRUIT : External characters

xxxi. **Fruit shape** : 1. Spheroid 2. Ellipsoid 3. Pyriform 4. Oblique 5. Oblate 6. Ovoid-oblique 7. Ovoid

xxxii. **Fruit weight** (g)

xxxiii. **Fruit height** (mm)

xxxiv. **Fruit diameter** (mm)

xxxv. **Shape of base of fruit** : 1. Necked 2. Convex 3. Truncate 4. Concave 5. Concave collared 6. Concave with neck

xxxvi. **Shape of apex of fruit** : 1. Mammiform 2. Angular 3. Convex 4. Truncate 5. Depressed

xxxvii. **Epicarp colour** : 1. Green 2. Yellow 3. Orange

xxxviii. **Surface of epicarp** : 1. Smooth 2. Rugose 3. Papillate 4. Pitted 5. Bumpy

V. FRUIT : Internal characters

xxxix. **Adherence of epicarp to mesocarp** : 3. Slight 5. Moderate 7. Strong

xL. **Nature of oil glands** : 1. Inconspicuous 5. Conspicuous 9. Very conspicuous

xLi. **Thickness of mesocarp** (mm) :

xLii. **Colour of mesocarp** : 1. White 2. Yellow

xLiii. **Number of segments per fruit**

xLiv. **Adherence of segments to each other** : 3. Slight 5. Moderate 7. Strong

xLv. **Toughness of skin around segments** : 3. Very delicate 5. Delicate 7. Tough

xLvi. **Fruit axis** : 1. Solid 2. Semi-hollow 3. Hollow

xLvii. **Cross-section of fruit axis** (mm)

xLviii. **Colour of pulp** : 1. Yellow 2. Orange 3. Pink 4. Red 5. Green

xLix. **Uniformity of colour of pulp** : 3. Uniform 7. Streaked

L. **Texture of pulp** : 3. Tender 5. Firm 7. Tough

Li. **Size of vesicles** : 3. Small 7. Large

Lii. **Shape of vesicles** : 3. Thin 7. Thick

Liii. **Juice in endocarp** : 3. Low 5. Medium 7. High

Liv. **Colour of juice** : 1. Greenish 2. White 3. Pale Yellow 4. Yellow 5. Orange 6. Reddish

Lv. **Taste of juice** : 1. Very poor 3. Poor 5. Fair 7. Good 9. Excellent

Lvi. **Aroma of juice** : 3. Weak 7. Strong

VI. SEED

Lvii. **Average number of seeds per fruit**

Lviii. **Average length of seeds** (mm)

Lix. **Average width of seeds** (mm)

Lx. **Shape of seeds** : 1. Fusiform 2. Clavate 3. Cuneiform 4. Ovoid 5. Deltoid 6. Globose 7. Semi-spheroid

Lxi. **Texture of seed surface** : 1. Smooth 2. Wrinkled 3. Hairy

Lxii. **Seed colour** : 1. White 2. Cream 3. Yellowish 4. Green 5. Brown

Lxiii. **Cotyledon colour** : 1. White 2. Light green 3. Green

Lxiv. **Chalazal spot colour** : 1. White 2. Ivory 3. Cream 4. Yellow 5. Beige 6. Brown 7. Reddish 8. Purple

Lxv. **Average number of embryos per seed**

All the statistical computation were performed using a Statistical Software Package, "SPSS for MS Windows Release 6.0".

RESULTS AND DISCUSSION

Fruit weights of the different biotypes were found to be highly variable. Bira Jora biotype of Jorhat collection showed highest fruit weight (3.5 kg), whereas for the Gondharaj biotypes of Agartala collection it was 0.75 kg. The length and diameter of the fruits were varied from 13.5-31 cm and 8-16 cm, respectively. Citron fruits are generally characterized by large in size and of irregular shape (Swingle, 1943).

The "Rescaled Reversed Absolute Squared Euclidean Similarity Coefficient Matrix" has been presented in the Table 3. The lowest value 0.000 was observed between Gondharaj (Agartala) and Bira Jora (Jorhat) and highest value between Seuli (Barapani) and Gondharaj (Nowgaon).

Dendrogram utilizing different methods has been presented in the Fig. 1. Utilizing any of the five methods for dendogram construction, the data can be divided into 3 different clusters. Except the average linkage (between group) method, in all the other four methods, [average linkage (within group), single linkage, complete linkage and median method], the first cluster includes Seuli (Barapani), Gondharaj (Nowgaon) and Gondharaj (Agartala) collections. Whereas Haijange (Imphal), Bira Jora (Titabor) and Jora Tenga (Margherita) falls in the second cluster. In the average linkage (within group) method, Jora Tenga (Margherita) was included in the first cluster. In all the cluster methods the composition of third cluster was found to be same. Bira Jora (Jorhat) and Bira Jora (Teok) collections fall in this cluster (Fig. 1.)

Table 3. Rescaled Reversed Absolute Squared Euclidean Similarity Coefficient Matrix of different biotypes

		1	2	3	4	5	6	7	8
1.	Haijange (Imphal)	0							
2.	Jora Tenga (Margherita)	0.9762	0						
3.	Gondharaj (Agartala)	0.7718	0.9009	0					
4.	Seuli (Barapani)	0.8827	0.9671	0.9894	0				
5.	Bira Jora (Jorhat)	0.6084	0.5257	0.0000	0.2168	0			
6.	Bira Jora (Teok)	0.8213	0.7459	0.3168	0.5013	0.9754	0		
7.	Bira Jora (Titabor)	0.9849	0.9224	0.6482	0.7859	0.7629	0.9226	0	
8.	Gondharaj (Nowgaon)	0.8787	0.9706	0.9945	1.0000	0.2094	0.4970	0.7743	0

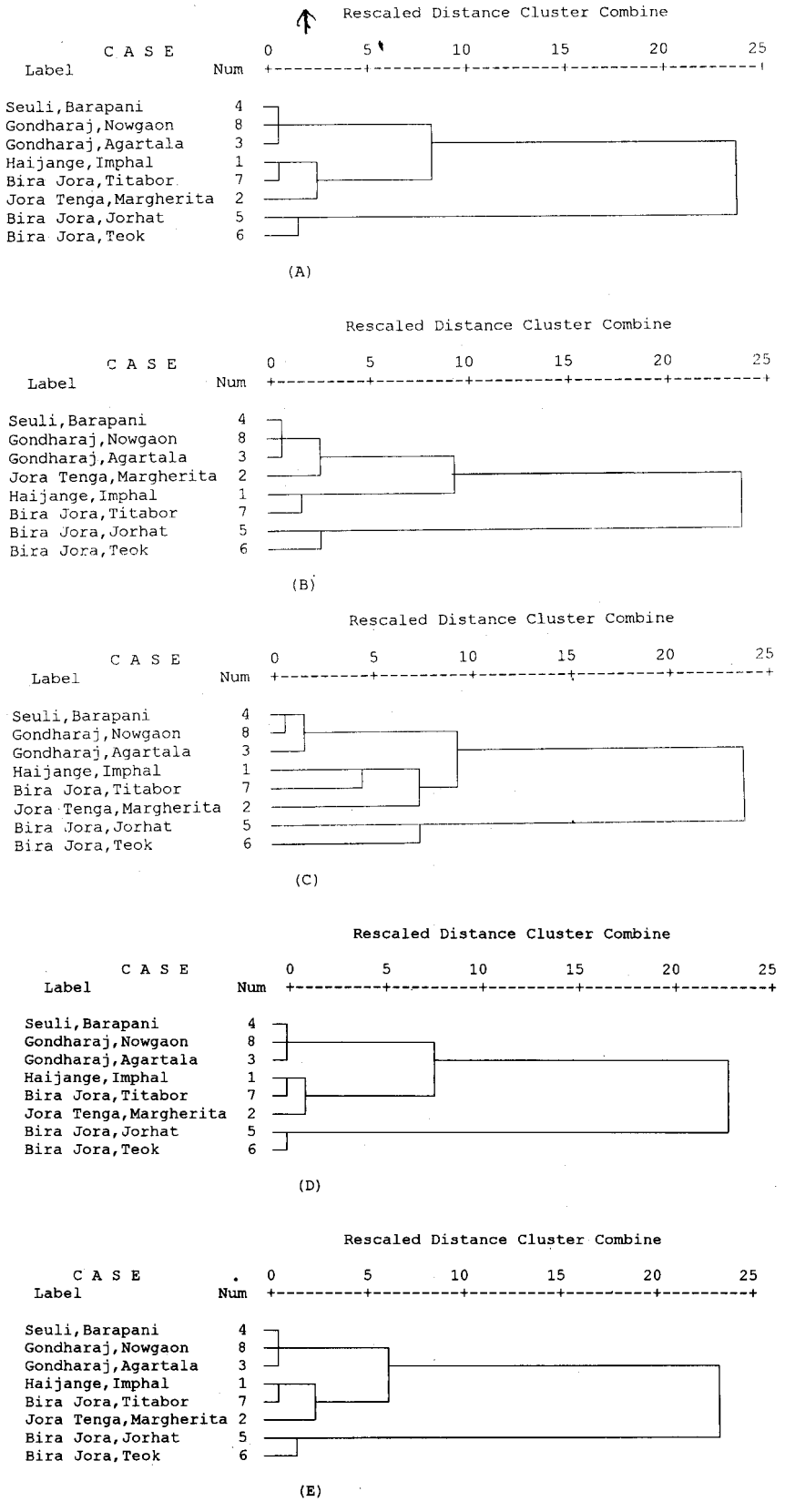

Fig. 1. Dendrogram using (A) Average linkage (between group); (B) Average linkage (within group); (C) Single linkage; (D) complete linkage and (E) Median method

Thus for the utilized procedures for cluster formation the results of the clustering information are nearly identical. Lebeda and Jendrulek (1987) allso found in their analysis of host-parasite interaction, that for six of the used procedures, results were nearly identical.

The genotypes which are included into a cluster are having less variation within themselves, whereas genotypes of one cluster should have wide variation between individuals of another cluster. *C. medica* is a monoembryonic species and commonly propagated through seeds which creates free gene exchange and recombination during their sexual reproduction. But this species has maintained its unique phenotypic characters. *C. medica* was considered to be a basic species by several workers (Swingle 1943; Barret and Rhodes, 1976; Green *et al.*, 1986 and Sugawara *et al.*, 1995). The native home of this species has not been determined with certainity. The citron is commonly supposed to be indigenous to India (Scora, 1988). The present results strongly support the above view that citron is indigenous to the north-east India.

ACKNOWLEDGEMENTS

Thanks are due to Dr. H. Changmai, Associate Professor, Assam Agricultural University, Jorhat (Assam) for his help during data analysis. This work was supported by Dr. K. S. Krishnan (DAE) Research Fellowship offered by Department of Atomic Energy, Government of India to B. K. Ray.

REFERENCES

Barret, H. C. and A. M. Rhodes. 1976. A numerical taxonomic study of affinity relationships in cultivated citrus and its close relatives. *Syst. Bot* 1: 105-136.

Green, R. M.; A. Vardi and E. Galun. 1986. The plastome of Citrus: physical map, variation among *Citrus* cultivars and species and comparison with related genera. *Theor. Appl. Genet.* 72: 170-177.

IBPGR. 1988. Descriptors for Citrus. International Board for Plant Genetic Resources, Rome.

Lebeda, A. and T. Jendrulek. 1987. Cluster analysis as a method for evaluation of genetic similarity in specific host-parasite interaction (*Lactuca sativa - Bremia lactucae*). *Theor. Appl. Genet.* 75 : 194-199.

Peters, J. P. and J. A. Martinelli. 1989. Hierarchical cluster analysis as a tool to manage variation in germplasm collections. *Theor. Appl. Genet.* 78 : 42-48.

Scora, R. W. 1988. Biochemistry, taxonomy and evolution of modern cultivated citrus. *In* : Goren, R. and Mendel, K. (eds.) Proceedings of the Sixth International Citrus Congress, p 277-289, Tel Aviv, Israel.

Sneath, P. H. A. and R. R. Sokal. 1973. Numerical taxonomy. W. H. Freeman & Co., San Francisco.

Sugawara, K., A. Oowada, T. Moriguchim and M. Omura. 1995. Identification of citrus chiemeras by RAPD markers. *Hor Sci.* 30 : 1276-1278.

Swingle, W. T. 1943. The botany of Citrus and wild relatives of the orange subfamily (family Rutaceae, subfamily Aurantioideae). *In*: Webber, J. H. and Batchelor, L. D. (eds.) The Citrus Industry. Vol. 1, p. 129-474. Univ. Calif. Press. Berkeley.

Tanaka, T. (1954). Species problem in Citrus (Revisio aurantiacearum IX) *Jpn. Soc. Prom. Sci., Ueno, Tokyo*, 152.

Tolkowsky, S. 1938. Hesperides: a history of the culture and of citrus fruits. John Bale Sons & Curnow, London.

Torres, A. M., R. K. Soost and U. Diedenhofen. 1978. Leaf isozymes as genetic markers in citrus. *Amer. J. Bot.*, 65 : 869-881.