# MULTIVARIATE CLUSTERING TECHNIQUES – A COMPARISON BASED ON ROSE ( *Rosa* spp.)

*by*

**ARYA V CHANDRAN**

**(2016-19-002)**

**THESIS**

**Submitted in partial fulfilment of the**
**requirement for the degree of**

## MASTER OF SCIENCE IN AGRICULTURE

**Faculty of Agriculture**
**Kerala Agricultural University**

**DEPARTMENT OF AGRICULTURAL STATISTICS**
**COLLEGE OF AGRICULTURE**
**VELLAYANI, THIRUVANANTHAPURAM – 695 522**
**KERALA, INDIA**
**2018**

# DECLARATION

I, hereby declare that this thesis entitled "**MULTIVARIATE CLUSTERING TECHNIQUES – A COMPARISON BASED ON ROSE ( *Rosa* spp.)** " is a bonafide record of research work done by me during the course of research and the thesis has not previously formed the basis for the award to me of any degree, diploma, associateship, fellowship or other similar title, of any other University or Society.

Vellayani,

30-08-2018

**ARYA V CHANDRAN**

(2016-19-002)

3

# CERTIFICATE

Certified that this thesis entitled **"MULTIVARIATE CLUSTERING TECHNIQUES – A COMPARISON BASED ON ROSE ( *Rosa* spp.) "** is a record of research work done independently by Ms. **ARYA V CHANDRAN** (2016-19-002) under my guidance and supervision and that it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to her.

Vellayani
30-08-2018

**Dr. Vijayaraghava Kumar**
(Major Advisor, Advisory Committee)
Professor and Head
Department of Agricultural Statistics,
College of Agriculture, Vellayani,
Thiruvananthapuram -695 522

## CERTIFICATE

We, the undersigned members of the advisory committee of Ms. **Arya V Chandran** (2016-19-002), a candidate for the degree of **Master of Science in Agriculture** with major in Agricultural Statistics, agree that the thesis entitled **"MULTIVARIATE CLUSTERING TECHNIQUES – A COMPARISON BASED ON ROSE ( *Rosa* spp.) "** may be submitted by Ms. **Arya V Chandran**., in partial fulfillment of the requirement for the degree.

Kumar
30/8/18

**Dr. Vijayaraghava Kumar**
(Major Advisor, Advisory Committee)
Professor and Head
Department of Agricultural Statistics,
College of Agriculture, Vellayani,
Thiruvananthapuram -695 522

**Smt. Brigit Joseph**
(Member, Advisory Committee)
Associate Professor
Department of Agricultural Statistics,
College of Agriculture, Vellayani,
Thiruvananthapuram -695 522

30.8.18

**Dr. C. Lekha Rani**
(Member, Advisory Committee)
Professor
Department of Plant Breeding and Genetics,
College of Agriculture, Vellayani,

**Dr. G. S. Sreedaya**
(Member, Advisory Committee)
Assistant Professor (Sr. Scale)
Department of Agricultural Extension,
College of Agriculture, Vellayani,

**EXTERNAL EXAMINER**

Professor
School of Management & Entrepreneurship
Kerala University of Fisheries & Ocean Studies,
Panangad, Kochi - 682 506

# ACKNOWLEDGEMENT

*With great pleasure, I wish to express my deepest gratitude and indebtedness to* ***Dr. Vijayaraghava Kumar****, Professor and Head, Department of Agricultural Statistics, College of Agriculture, Vellayani and Chairperson of my Advisory Committee for his valuable suggestions, constant support, extreme patience and co-operation throughout the investigation. It was his dedication and perfectionism helped me in all the time of research and writing of this thesis I feel proud of myself in confessing that it has been a unique privilege for me being one of his students.*

*My esteem sense of gratitude and indebtedness to* ***Smt. Brigit Joseph****, Associate Professor, Department of Agricultural Statistics, College of Agriculture, Vellayani and member of Advisory Committee, for her prudent suggestions, valuable advice and whole hearted approach for the successful completion of the thesis.*

*I would also like to express my gratitude towards* ***Dr. C. Lekha Rani, Professor****, Department of Plant Breeding and Genetics, College of Agriculture, Vellayani and member of Advisory Committee for the constant support, and valuable suggestions rendered throughout the period of research work.*

*I express my gratitude to* ***Dr. G. S. Sreedaya****, Assistant Professor, Department of Agricultural Extension, College of Agriculture, Vellayani, for her encouragement, wholehearted help and support throughout the period of research work.*

*I gratefully acknowledge.* ***Susan madam*** *who has helped me to have an idea about computer application.*

*I am extremely thankful to Meenu mam for valueable suggestions, encouragement and constant support. I extend my thankfulness and respect to Suprej sir, Sajina mam, Nisha mam and Asha mam for their support throughout my course work.*

*I extend my thankfulness to my seniors Rakhi chechi, Neethu chechy, BR chettan and Jaslam ikka for their affection, help and support. Words are scarce to express my deep sense of gratitude to my classmates Muhsi, Fallu, Stephy and Abhinav for their encouragement, help and support. I am also indebted to express my thanks to my juniors Dhanya, Athulya, Neethu, Priyanka, Ajmal, Murukesh and Suresh for their support and help.*

Words fail to express my sincere thanks to Anjaly and Amritha who were there with me during all hard times, for their love, care and for the happiest moments we cherished together. Mere words cannot express my profound indebtness to my dearest friend Sree for her affection and constant support. Words are inadequate to express my thanks to my beloved friends Geethu Jacob, Pooja, Elzu, Avinash for their help, support and love.

I am most indebted to my beloved father **Sri. K.R. Karmachandran pillai** and my dearest mother **Smt. K. Vijayamma** for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. I will forever be grateful to my best friend, my love, Ajay for his unfailing support, care, love and encouragement. My heartful thanks to Kannettan, Chandu chettan, Remya chechi, Ambika chichi, Kalyani, Sujith chettan, Vallyachan and Vallyammachi for all the support in every way, spiritual, moral and physical. I have no words to express my gratitude and love towards my grandmother for her unconditional love and prayers.

I once again express my sincere gratitude to all those who helped me in one way or another in the successful completion of this venture.

Arya.

## TABLE OF CONTENTS

8

## LIST OF TABLES

11

12

## LIST OF FIGURES

14

18

## LIST OF APPENDICES

## LIST OF ABBREVIATIONS

**ANOVA : Analysis of variance**

**cm : centimeter**

**d.f : degrees of freedom**

***et al.*** **: Co-workers**

**etc. : et cetera**

**F / f : Floribunda**

**Fig. : Figure**

**g : gram**

**H / h : Hybrid Tea**

***i.e.*** **: that is**

**MANOVA : Multivariate analysis of variance**

**PCA : Principal component analysis**

**UPGMA : Unweighted Pair Group Average Method**

**UPGMC : Unweighted Pair Group Centroid Method**

**WPGMA : Weighted Pair Group Average Method**

**% : Percent**

# *Introduction*

# 1. INTRODUCTION

Multivariate analysis is very effective tool to study objects characterized by a number of traits. Data of this type arise in all branches of science and methods of analyzing multivariate data constitute an increasingly important area in statistics. It is a ever expanding set of techniques for data analysis which includes a wide variety of techniques such as Principal Component Analysis (PCA), Factor analysis, Discriminant function analysis, Canonical Correlation Analysis (CCA), Cluster analysis and Path analysis etc. Selection of appropriate techniques depends on the objective of study. Multivariate techniques give a more realistic picture than looking at single variable.

Multivariate cluster analysis includes many diverse techniques for discovering association structure within complex bodies of data. It is a technique used for combining observations into groups such that each group is homogenous or compact with respect to certain characteristics and each group should be different from other. Grouping is done on the basis of similarities or distances. Cluster analysis stand out from other multivariate technique as it does not require an initial hypothesis with respect to the population and gives an easy interpretation. The need for cluster analysis arises in natural ways in many fields such as life science, medicine, engineering, agriculture, social science, etc.

Due to the size and complexity of the underlying data sets multivariate analysis requires much computational effort. With the continued growth of computational power, multivariate techniques play an increasingly important role in data analysis. Clustering and classification helps to make sense of and extract value from large sets of structured and unstructured data. When we are working with huge volumes of unstructured data, it only makes sense to try to partition the data into some sort of logical groupings before attempting to analyze it. Clustering techniques differ from normal classification in the sense that there is no group information in cluster analysis while classification is based on prior group information. Cluster analysis is based on the degree of correspondence among objects across all of the characteristics used in the analysis.

The cluster analysis involves measure of similarity, selection of clustering technique, carrying out clustering based on the selected technique, making decision on number of clusters and finally interpretation of results. The data used in cluster analysis can be interval, ordinal or categorical. However, having a mixture of various types of variables will make the analysis more complicated because we need to measure the distance between observations and the type of measure used will depend on what type of data we have. It is the only multivariate technique on the comparison of objects based on variates, not on the estimation of the variate itself.

No generalization about cluster analysis is possible as a large number of clustering methods have been developed in different fields with different similarity measures. Association measures have a great impact on clustering results. There are different association measures for different types of data. Clustering methods are applied to these measures to obtain the cluster. Clustering methods range from those that are largely heuristic to more formal procedures based on statistical models. The results obtained by using different association measures and different clustering methods are not unique. Usually we follow either a hierarchical strategy or one in which observations are relocated among tentative clusters for ease of computation. There is no superior technique, but performance of each technique seemed to depend heavily on the nature of dataset. The present study is conducted with the following objectives:

1. To compare different clustering techniques.
2. To identify the suitable technique for different types of qualitative and quantitative data.
3. To illustrate the procedures using data based on a field experiment on rose (*Rosa* spp.).

# Review of Literature

# 2. REVIEW OF LITERATURE

Critical review of literatures related to the research area helped to understand different methodologies for conducting the study. The review of literature is presented in the following titles:

2.1 Multivariate analysis

2.2 Cluster analysis

2.3 Applied studies

2.4 Comparison of association measures

2.5 Comparison of clustering methods

## 2.1 MULTIVARIATE ANALYSIS

Multivariate data consist of observations on several variables for a number of individuals or objects. Simplification, analysis of dependence, analysis of interdependence, reducing dimensionality, clustering and testing of multiple variables are the objectives of multivariate techniques. The selection of the most appropriate method depends on the type of data, type of problem, and the sort of objectives which are envisaged for the analysis (Chatfield and Collins,1980).

Cluster analysis is different from discriminant analysis as groups are predetermined in discriminant analysis while in cluster analysis groups are not predetermined. There is a well established procedure for discriminant analysis and in cluster analysis different methods are there which may result in different grouping structure (Rangaswamy, 1995).

Score obtained by taking sum of products of principal component loadings and corresponding mean value can be used to plot the objects in graph. Such graph of scores (score plot) helps to identify the natural groupings visually (Rencher,2002).

The principal component analysis is done to derive a small number of linear combinations (principal components) of a set of variables that retain as much information in the original variables as possible. The original correlated variables are transforms to a set of new uncorrelated random variables. These new variables are linear combinations of the original variables and are derived in decreasing order of importance so that the first principal component accounts for as much as possible of the variation in the original data so that the effective dimensionality of the data can be reduced (Parsad,2007).

Multivariate analysis is a statistical tool that simultaneously analyzes multiple measurements on objects under study. Proper application of these techniques reveals the relationships that otherwise would not be identified and these are extension of univariate and bivariate analysis. (Hair *et al.*, 2015).

## 2.2 CLUSTER ANALYSIS

Sneath and Sokal (1973) categorized the clustering methods into hierarchical or non-hierarchical, divisive or agglomerative and polythetic or monothetic.

Hardle and Simar (2007) defined cluster analysis as a set of tools for building groups from multivariate data and it is divided into two fundamental steps, which include choice of proximity measures and choice of group building algorithm.

Hair *et al.*(2015) stated that cluster analysis is a multivariate technique which classifies the objects based on a set of characteristics in such a way that the resulting clusters have high internal homogeneity and high external heterogeneity.

### 2.2.1 Association measures

Jaccard (1901) introduced a similarity coefficient for binary data often referred as Jaccard's coefficient in which all terms have equal weight.

Mahalanobis (1936) in his paper on 'generalized distance' mentioned the $D^2$ statistics which has become the standard measure of distance between two populations when all observed variables are quantitative.

Dice (1945) developed another method for qualitative data which do not consider negative matches and gives double weight to unmatches.

Sokal and Michener (1958) introduced the simple matching coefficient which assumes that there is no difference between double-0 and double-1. In this case any one of the two states of each descriptor could be coded 0 or 1 indifferently.

Rogers and Tanimoto (1960) found another measure in which differences are given more weight than resemblances.

Hair *et al.*(2015) listed four different distance or dissimilarity measures such as Euclidean distance, Squared Euclidean, City Block distance and Mahalanobis for continuous variables

### 2.2.2 Clustering method

Sorensen (1948) proposed complete linkage agglomeration which is opposite to the single linkage. The fusion of two clusters depends on the most distant pair of objects.

Rao (1952) described Tocher method of clustering based on $D^2$ statistics which is widely used for grouping of quantitative data.

Ward (1963) proposed a clustering method which is related to the centroid method. It is a method which minimizes an objective by using squared error criterion as that used in multivariate analysis of variance.

MacQueen (1967) introduced k- means clustering which is a portioning approach towards the grouping of objects.

Sneath and Sokal (1973) developed four different clustering methods which rely instead on average similarities among objects or on centroids of clusters. Unweighted pair group average clustering (UPGMA) and Weighted pair group average clustering (WPGMA) depends on the arithmetic average while Unweighted pair group centriod clustering (UPGMA) and Weighted pair group centroid clustering (WPGMA) depends on the centroid.

Hartigen (1979) found that the k-means algorithm produces a clustering which is only locally optimum. The within-cluster sum of squares may not be decreased by transferring a object from one cluster to another, but different partitions may have the same or smaller within cluster sum of squares. Usually less than 10 iterations are required to attain local optimality.

### 2.3 Applied studies

Sharma *et al.* (2006) used non hierarchical Euclidean clustering approach to estimate the genetic divergence in tomato. 60 genotypes were grouped into 10 clusters.

Sheela *et al.*(2006) studied molecular characterization of *Heliconia* by RAPD assay. Seventeen *Heliconia* species and varieties were analyzed using RAPD markers and the genetic similarity matrix constructed with Jaccard's coefficient using RAPD marker scores. Nine distinct clusters were identified using UPGMA method.

Ali *et al.* (2008) conducted a study on genetic variability, association and diversity studies in wheat (*Triticum aestivum* L.) germplasm using seventy genotypes. The genetic diversity analysis was done through the cluster analysis using Euclidean dissimilarity and complete linkage method. The varieties were grouped into 4 clusters at 30% linkage distance.

Singh *et al.*(2008) used Mahalanobis $D^2$ statistics for divergence analysis for quality traits in forty five indigenous basmati rice. The genotypes were grouped into 8 clusters depending on $D^2$ estimates and it was observed that the character grain length was the major contributing factors towards genetic diversity.

Akbar *et al.* (2011) studied genetic divergence in sesame (*Sesamum indicum* L.) landraces based on qualitative and quantitative traits. The hierarchical cluster analysis based on Euclidian distance resulted in seven clusters. The relationship among sesame accessions were observed by plotting the scores obtained from first two principal components.

Bharathi *et al.* (2012) estimated the genetic divergence among 41 genotypes by $D^2$ statistics of Mahalanobis followed by clustering by Tocher's method. The study revealed that the tender fruit yield per vine contributed maximum to the genetic divergence.

Shathi *et al.* (2012) determined the genetic divergence in mustard. Six clusters were formed based on Mahalanobis $D^2$ values among 25 genotypes. Principal component analysis (PCA) used to verify the grouping obtained through $D^2$ statistics. The study revealed that the first two principal components were sufficient for explaining the variation. Group constellations were developed independently by using score obtained through first two components and confirmed the results obtained through $D^2$ analysis.

Longkumar and Kabir (2014) studied genetic divergence of mandarin genotypes in Nagaland using Mahalanobis $D^2$ technique. Fifty genotypes of orange were grouped according to Tocher's method. The character leaf length among morphological characters and equatorial diameter among physico-chemical characters were found to be important for the expression of genetic divergence.

Vasanthi *et al.* (2014) measured the magnitude of genetic divergence among 29 genotypes of groundnut using Mahalanobis $D^2$ statistics for a set of ten characters. $\lambda$ statistic from analysis of dispersion was significant indicating significant difference among the genotypes when all the characters were considered simultaneously. Twenty nine genotypes were grouped into eight clusters.

Kuswardhani *et al.* (2014) performed cluster analysis for classification of farm household based on socio-economic characteristics for technology adoption in agriculture. Survey was done in 580 farm households. Ward's hierarchical procedure and k-means clustering was done for clustering farm households separately for three different regions.

Toma *et al.* (2015) used cluster algorithm for analyzing organic farming patterns. Farm type, farming type and certified area share were criteria under study. Hierarchical cluster analysis based on Squared Euclidean distance and Centroid linkage method permitted them to group localities with organic agriculture in to clusters.

Aswathi *et al.*(2015) conducted a study on genetic divergence in cowpea (*Vigna* spp.) varieties for seed quality of ten cowpea varieties. The seeds were evaluated for four seed quality parameters *viz.*, germination, speed of germination, seedling vigour index I and seedling vigour index II. On the basis of $D^2$ values the genotypes were grouped into four clusters.

Priya *et al.* (2017) assessed the genetic divergence in rice on 11 morphological and yield determining characters of 40 genotypes. Genetic divergence was estimated using Mahalanobis $D^2$ statistics and varieties were grouped into 7 clusters by Tocher's method.

Mahesh *et al.* (2017) utilized the Mahalanobis $D^2$ statistics for Genetic divergence studies in soybean (*Glycine max* L. Merrill). Forty genotypes got grouped into six clusters based on $D^2$ estimates for eleven quantitative traits.

Sajeevkumar *et al.* (2017) performed Diversity analysis of KAU released cocoa (*Theobroma cacao* L.) varieties based on morphological parameters. Agglomerative hierarchical clustering (UPGMA) was done based Jaccard's similarity. Five clusters were formed from ten varieties at 68 percent similarity level. Morphological observations on distinguishable eight quantitative and six qualitative characters were used and distribution pattern of varieties based on qualitative and quantitative clustering are found to be varied.

Sunayana *et al.* (2017) evaluated genetic divergence in mungbean (*Vigna radiata* L. Wilczek) using morpho-physio and molecular markers to identify drought tolerant genotype.Clustering of 60 mungbean genotypes were carried out using Unweighted pair group average method( UPGMA) with City Block distance and Principal component analysis (PCA).Clustering was done for phenotypic and molecular markers.

## 2.4 Comparison of association measures

Nei and Li (1979) in thier study on mathematical model for studying genetic variation in terms of restriction endonucleases found that the Jaccard and Dice coeffcients are very similar and there was no difference in dendrogram topology but in branch length. Generally, there is a slight preference for the Dice coefficient as it is same as that of the Nei and Li coefficient, which is most suitable to determine genetic relation based upon DNA restriction fragment patterns.

Jackson *et al.*(1989) studied different similarity coefficients that whether they are measures for cooccurrence and association or simply measures of co-occurrence. It was found that the choice of association measures has high impact on clustering results.

Meyar *et al.* (2004) compared different similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L) using 18 inbred lines. They compared eight association measures for qualitative data. The measures were differentiated with the help of Spearman correlation, cluster analysis with dendrogram (UPGMA, WPGMA, single linkage, complete linkage and Neighbour-Joining methods), the consensus fork index of dendrograms, groups created from Tocher optimization procedure and projection efficiency in a two-dimensional space. They used different measures like Jaccard, Dice, Anderberg, Ochiai, Simple-matching, Rogers and Tanimoto, Ochiai II and Russel and Rao. For almost all methodologies and marker systems, the Jaccard, Dice, Anderberg and Ochiai measures gave similar results since all of them exclude negative co- ccurrences. Simple Matching, Rogers and Tanimoto, and Ochiai II coefficients were also showed similar results may be due to the fact that they all include negative co-occurrences. The Russel and Rao coefficient showed very different results from the other measures.

Kosman and Leonard (2005) in their study on similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid and polyploidy species mentioned that the selection of appropriate association measure is crucial since different methods may give conflicting results. They found that Dice is suitable for haploid with codominant markers and Jaccard for distantly related haploid.

Dalirsefat *et al.* (2009) conducted study on 'Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silkworm, *Bombyx mori*'. They assessed the variation caused by Jaccard, Dice and Simple matching coefficient by the visual inspection and consensus fork index

- *CI* of dendrogram produced and by comparing Spearman correlation, projection efficiency in a two-dimensional space, and clusters created with Tocher optimization procedure. The study revealed that in almost all methods Jaccard and Dice gave same results as both methods exclude negative co-occurrences.

Ojurongbe (2012) compared proximity measures such as Jaccrd, Dice Simple matching coefficient and classification methods for binary data. Clustering with single linkage, complete linkage, UPGMA, WPGMA and Neighbour-Joining (NJ) method. The result showed that Jaccard and Dice measure gave similar result under different method. It also suggested that the single linkage method is not an appropriate one since it has low consensus fork index value. It was suggested that UPGMA method gives consistent results with respect to grouping irrespective of the similarity measure/coefficient based on the cophenetic correlation value.

Alves *et al.*(2012) conducted a study on comparison of efficiency of distance measurement methodologies in mango (*Mangifera indica*) progenies based on physicochemical descriptors. The study was based on twenty five mango variety using six characters. The measures like Coler-Rodgers distance, Euclidean distance, average Euclidean distance, Gower distance, Mahalanobis' generalized distance, weighted distance by squared residuals, Euclidean distance squared were compared along nearest neighbor, farthest neighbor, Ward, Gower, UPGMA methods. Spearman correlation indicated that the distances are correlated except the Mahanalobis distance. Euclidean distance, average Euclidean distance and Euclidean distance squared showed the same distance ranking between the genotypes. The clustering based on Tocher's modified method showed variations form other different distance measures. It was found that UPGMA was the most efficient among the grouping methods assessed.

Dahal (2015) studied the effect of different distance measures in result of cluster analysis and found that the difference between clusters created by Euclidean, squared Euclidean and Manhattan distance is small. These measures gave different result from

Mahalanobis distance. Analysis was carried out with the help of Kappa coefficient and cluster mapping. Comparison were done with k-medoid clustering method and cluster map from clustering using the dataset on slope information, total cross sectional areas of trees from 1 to 3 meters in height and soil type of the given

terrain.

Shirkhorshid *et al.* (2015) compared different distance measures for clustering of continuous data. Clustering obtained from k- means , k- median and hierarchical clustering under distance measures like Euclidean, Average, Cosine, Chord, Mahalanobis, Canberra, Coefficient of Divergence, Czekanowski Coefficient, Index of Association, Manhattan, Mean Character Difference and Pearson coefficient were compared using Rand index. It was found that among these measures Average Distance is the top most accurate measures for all clustering.

## 2.5 Comparison of clustering methods

Cunningham and Ogilvie (1972) compared seven hierarchical methods based on the association between the input dissimilarity values and corresponding distance values obtained from the final clustering hierarchy.

Kuiper and Fisher (1975) analysed six hierarchical clustering procedures (single linkage, complete linkage, median, average linkage, centroid and Ward's method) for multivariate normal data. In their study with equal cluster size, Ward's method and complete linkage method, and with unequal cluster sizes centroid and average linkage method were found to be the best respectively.

Blashfield (1976) reported that Ward's method performed significantly better than the other clustering procedures and average linkage gave relatively poor results. It was based on his comparative study on four types of hierarchical clustering methods (single linkage, complete linkage, average linkage and Ward's method) for accuracy in recovery of original population clusters.

Milligan (1980) found that complete linkage and Ward's method reacted badly when outliers were introduced into the simulated data.

Hands and Everitt (1987) found that Ward's method was the best overall than other hierarchical methods. They compared five hierarchical clustering techniques (single linkage, complete linkage, average, centroid, and Ward's method) on multivariate binary data.

Peeters and Martinelli (1989) compared five clustering methods, namely UPGMA, UPGMC (Unweighted Paired Group Method using centroids), single linkage, complete linkage and median, for their utility in revealing genotype associations in barley germplasm collections. UPGMA and UPGMC were found to be almost comparable with a relatively high level of accuracy, in accordance with pedigrees. Single linkage and median clustering methods led to "chaining effect," which gave poor resolution of individual groups and complicate the interpretation of results.

Milligan (1996) shown that results of single linkage clustering are sensitive to noise in the data because noise changes the similarity values and may thus easily modify the order in which objects cluster. Study revealed that it is due the chaining effect occur in single linkage method.

Rincon *et al.* (1996) compared several clustering methods in grouping maize accessions on the basis of agronomic morphological characters. UPGMA method was generally consistent with regard to the allocation of clusters, when different types and number of characters were used.

Lombard *et al.* (2000) analysed genetic relationships in rapeseed (*Brassica* spp.) cultivars on the basis of amplified fragment length polymorphisms (AFLP) by means of UPGMA and Ward's method in combination with Jaccard, simple matching, and modified simple matching coefficients. Despite very high correlations between distance matrices obtained through the use of different coefficients, and derivation of the same patterns with both clustering methods, Ward's method was found more suitable as it avoids the chaining effects that are often observed with UPGMA.

Tarpey (2007) analysed several clustering methods for functional data. The focused on k-means clustering and examined the effect on the clustering outcomes based on how the observed data were smoothed. The result of the analysis suspect that clustering on functional data depend on how well the smooth curves fit the raw data, but that the choice of best smoothing method depends on the true mean curve of each cluster.

Ferreira and Hitchcock (2009) compared the performance of four major hierarchical methods such as single linkage, complete linkage, average linkage and Ward's method for clustering functional data. They used the Rand index to compare the performance of each clustering method. Based on their study, Ward's method was usually the best, while average linkage performed best in some special situations, in particular, when the number of clusters is over specified.

Cluster validation gives the quantitative evaluation of the result of clustering algorithm. Validation techniques are categorized into internal, external and realative cluster validation techniques. Internal validation measures rely on the compactness, the connectedness and the separation of the cluster while the external validation compares the clusters to an external reference (Theodoridis & Koutroumbas , 2003).

Halkidi *et al.* (2002) listed out several cluster validity measures. They also mentioned that evaluating the clustering algorithm is an important aspect as it is a unsupervised process. As there are no predefined classes it is difficult to find out the appropriate method for clustering of objects.

Legany *et al.*(2006) compared different cluster validity measures by using runtime comparison. Measures like Dunn index, Davies – Bouldin index, SD validity index, S- Dbw validity index were compared by using various sets of data. The comparison showed that Dunn and S- Dbw are able to find out well separated clusters. Dunn index is time consuming while SD index is the fastest. The surface diagram of DB contained a lot of false peaks.

# Materials and Methods

# 3. MATERIALS AND METHODS

Cluster analysis is a multivariate technique to find the groupings (identical types) of a set of individuals. It attempts to maximize the homogeneity of objects within the clusters while maximizing the heterogeneity between the clusters. Cluster analysis is based on the degree of correspondence among objects across all the characteristics used in the analysis. It is measured in terms of similarity or dissimilarity (distance). Different types of similarity measures and clustering algorithms are available for quantitative and qualitative data. Results obtained from different methods are not unique. The cluster analysis involves measure of similarity, selection of clustering technique, carrying out clustering based on the selected technique, making decision on number of clusters and finally interpretation of results. The data used in cluster analysis can be interval, ordinal or categorical. Results obtained from different distance methods and clustering procedures are not unique. Selection of appropriate distance methods and clustering techniques seemed to depend heavily on the nature of dataset. In cluster analysis there are no predefined classes therefore evaluating the results of clustering algorithms is important. Comparisons of different methods were done using cluster validation techniques. Different distance measures and clustering methods were studied and results are illustrated using a field experiment on rose (*Rosa* spp.).

## 3.1 MATERIALS

The quantitative and qualitative traits used for the study collected from a field experiment on rose (*Rosa* spp.) entitled 'Characterization and genetic improvement in Rose (*Rosa* spp.) through mutagenesis' (2014-2017) at College of Agriculture, Vellayani and Regional Agriculture Research Station (RARS), Ambalavayal, Wayanad. Twenty five cultivars of rose each coming under the Hybrid Tea (single flowering) and Floribunda (group flowering) groups were evaluated for the study. These genotypes were selected, budded, established and maintained at Rose garden, Regional Agriculture

Research Station , Ambalavayal. The varieties were grown in completely randomized design with six replications. The first part of the study were the evaluation natural variability among different genotypes with the help of morphological characters and the second part were genetic improvement of genotypes by induced mutagenesis using physical and chemical mutagens. The varieties taken for the study are listed in the Table 1.

Data on the following characters were collected from the plant.

1. Number of leaves at first flower
2. Number of days to first flower
3. Prickle density (per five cm)
4. Flower size (cm)
5. Flower weight (g)
6. Pedicel length (cm)
7. Number of petals flower$^{-1}$
8. Size of petals (cm)
9. Number of flower plant $^{-1}$/bunch $^{-1}$
10. Fragrance
11. Flower colour
12. Vase life/ longevity

Among this traits number of leaves at first flower, number of days to first flower, prickle density, flower size, flower weight, pedicel length, number of petals flower$^{-1}$ ,Size of petals, number of flower plant $^{-1}$ bunch $^{-1}$ are quantitative character and fragrance, flower colour, vase life/ longevity are qualitative character. Details of scoring pattern for qualitative data are given in the Table 2. ( multistage coding) and Table 3. ( binary coding).

Table 1. List of cultivars under Hybrid Tea and Floribunda genotype.

| Hybrid tea genotypes (H1 to H25) | | | Floribunda genotypes (F1 to F25) | | |
|---|---|---|---|---|---|
| Sl. No. | | Genotype | Sl. No. | | Genotype |
| 1 | H1 | Madame George Delbard | 1 | F1 | Versailles |
| 2 | H2 | Aiswarya | 2 | F2 | Tickled Pink |
| 3 | H3 | Christ of Colomb | 3 | F3 | Rosarale de Chateau |
| 4 | H4 | Pink Panther | 4 | F4 | Rose Mary Gandhi |
| 5 | H5 | Roughe Miland | 5 | F5 | Princess de Monaco |
| 6 | H6 | Shrewsbury show | 6 | F6 | Ochi di Fita |
| 7 | H7 | Alaine Souchen | 7 | F7 | Carry Free Beauty |
| 8 | H8 | Amara | 8 | F8 | Sterntaler |
| 9 | H9 | Fryat | 9 | F9 | Orange N Lemon |
| 10 | H10 | Perfume Perfect | 10 | F10 | Lisa |
| 11 | H11 | Silver Star | 11 | F11 | The Mccartney Rose |
| 12 | H12 | Lincoln Cathedral | 12 | F12 | Chesire |
| 13 | H13 | A tago | 13 | F13 | Monnalisa |
| 14 | H14 | Demestra | 14 | F14 | Carolanne |
| 15 | H15 | Golden Fairy Sport | 15 | F15 | City of Glasgow |
| 16 | H16 | Mary Jean | 16 | F16 | Messara |
| 17 | H17 | Toplesse | 17 | F17 | Michel Fish |
| 18 | H18 | Priority Pride | 18 | F18 | Mini Pink |
| 19 | H19 | Majestic | 19 | F19 | Sans Souci |
| 20 | H20 | Prince Jardiner | 20 | F20 | Schloss Elutin |
| 21 | H21 | Cel b Lau | 21 | F21 | Lasting Piece |
| 22 | H22 | Lois Wilson | 22 | F22 | Plantein on Blumen |
| 23 | H23 | Mom's Rose | 23 | F23 | Winchester Cathedral |
| 24 | H24 | Alabama | 24 | F24 | Golden Fairy |
| 25 | H25 | Josepha | 25 | F25 | Prosperity |

Table 2. Scoring pattern (multistage coding) for qualitative data.

| Character | Scoring (Hybrid Tea Type) | | Scoring (Floribunda Type | |
|---|---|---|---|---|
| 1. Fragrance | Low | 1 | Low | 1 |
| | Medium | 2 | Medium | 2 |
| | High | 3 | High | 3 |
| 2. Flower colour | Yellow, Cream,White | 1 | Yellow, Cream | 1 |
| | Pink | 2 | Orange, Pink | 2 |
| | Red, Rose | 3 | Red , Rose | 3 |
| 3. Vase life/ longevity | 1 day | 1 | 1 day | 1 |
| | 2 days | 2 | 2 days | 2 |
| | 3 days | 3 | 3 days | 3 |
| 4. Prickle dendity (per five cm) | Above 7.31 | 1 | Above 9.39 | 1 |
| | 2.44 – 7.31 | 2 | 2.18 – 9.39 | 2 |
| | Upto 2.44 | 3 | Upto 2.18 | 3 |
| 5. Flower size (cm) | Upto 22.62 | 1 | Upto 16.02 | 1 |
| | 22.62 – 41.28 | 2 | 16.02- 29.31 | 2 |
| | Above 41.28 | 3 | Above 29.31 | 3 |
| 6. Size of petals (cm) | Upto 3.38 | 1 | Upto 2.95 | 1 |
| | 3.38- 11.09 | 2 | 2.95-8.40 | 2 |
| | Above 11.09 | 3 | Above 8.40 | 3 |
| 7. Flower size (cm) | Upto 1.3 | 1 | Upto 1.41 | 1 |
| | 1.3 – 2.13 | 2 | 1.41 – 2.58 | 2 |
| | Above 2.13 | 3 | Above 2.58 | 3 |

Table 3. Scoring pattern (binary coding) for qualitative data.

| Character | Scoring (Hybrid Tea Type) | | Scoring (Floribunda Type | |
|---|---|---|---|---|
| 1. Fragrance | Low | 0 | Low | 0 |
| | Medium, High | 1 | Medium, High | 1 |
| 2. Flower colour | Yellow,Cream, White, Pink | 0 | Yellow,Cream, Pink | 0 |
| | Red, Rose | 1 | Red , Rose | 1 |
| 3. Vase life/ longevity | 1 day | 0 | 1 day | 0 |
| | 2 / 3 days | 1 | 2 /3 days | 1 |
| 4. Prickle dendity (per five cm) | Above 6 | 0 | Above 6 | 0 |
| | Upto 6 | 1 | Upto 6 | 1 |
| 5. Flower size (cm) | Upto 36 | 0 | Upto 24 | 0 |
| | Above 36 | 1 | Above 24 | 1 |
| 6. Size of petals (cm) | Upto 6 | 0 | Upto 5 | 0 |
| | Above 6 | 1 | Above 5 | 1 |
| 7. Flower size (cm | Upto 1.3 | 0 | Upto 1.9 | 0 |
| | Above 1.3 | 1 | Above 1.9 | 1 |

## 3.2 PRELIMINARY STATISTICAL ANALYSIS

### 3.2.1 Multivariate analysis of variance (MANOVA)

The genotypes under study were tested for homogeneity prior to any attempt to form clusters of genotypes based on a set of characters. There is no necessity to form different cluters if they are homogenous as they form a single group. Multivariate analysis of variance was first developed by Wilks (1932).

Multivariate analysis of variance is the technique used for testing the homogeneity of a given set of genotypes with respect to a number of characters. MANOVA is an extension of ANOVA with several dependent variables. The process involves the technique of analysing variance and covariances of variables in multivariate case and partioning of these variances into different components (Rao, 1952). The model for each observation vector is

$$Y = \mu + \alpha + \varepsilon$$

Y – vector of individual responses

$\mu$ - vector of general mean efffect

$\alpha$ -vector of treatment effect

$\varepsilon$ - random error vector $\sim N(0,\Sigma)$

The total dispersion is split up into various components as given in Table 4.

Table 4: Multivariate analysis of variation

| **Source of variation** | **d.f** | **Dispersion matrix** |
|---|---|---|
| Replications | (r-1) | R |
| Between genotypes | (v-1) | B |
| Within genotypes | (r-1)(v-1) | W |
| **Total** | **(n-1)** | **C** |

Wilks (1932) developed a criteria through generalized likelihood ratio principle and is given as Wilk's lamda(Λ) criterion,

$$\Lambda = \frac{|W|}{|W + B|}$$

Where W is the within dispersion matrix

B is the between dispersion matrix

The statistic used for testing the homogeneity of a given set of genotypes with respect to a number of characters is given by

$V_{(stat)} = -m\log\Lambda$

where $V_{(stat)}$ is distributed as $\chi^2$ with *pq* degrees of freedom and

$m = n-(p+q+1)/2$,

p = number of variables

q = d.f. for variety

n = d.f. for error + variety

Significance of $V_{(stat)}$ shows that the differences between the populations with respect to means of 'p' characters are significant.

### 3.2.2 Analysis of variance (ANOVA)

The data were subjected to analysis of a completely randomized design (CRD) with ANOVA model as

$$X_{ij} = \mu + t_i + e_{ij}, \quad i = 1,2,\ldots,p$$

where $\mu$ is the general mean, $t_i$ is the effect of $i^{th}$ treatment and $e_{ij}$ is the error component with respect to $i^{th}$ character and $e_{ij}$ are normally distributed with mean zero and constant variance. The analysis of variance for completely randomized design is given in the Table 5.

Table 5: ANOVA for completely randomized design

| Source of variation | d.f. | M.S. |
|---|---|---|
| Between groups | (v-1) | SS(T) |
| Within groups | (n-v) | SS(E) |
| Total | (n-1) | |

### 3.2.3 Structure of multivariate observations

Multivariate analysis refers to all statistical techniques that simultaneously analyze observations on several variables for a number of individuals or objects (Chatfield and Collins, 1980). At a time it deals with 'p' measures on 'n' objects jointly. Multivariate procedures rely on the assumptions of multivariate normal distribution. Measurement on 'p' variables for 'n' individuals can be denoted as $X_{ij}$, where i =

(1,2,.....,p) and j = (1,2,......,n). A p dimensional random variable X is said to follow the multivariate normal distribution if its joint p.d.f. is of the form

$$f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

where Σ is p x p symmetric positive definite matrix and μ is the mean vector. The assumption of multivariate normal distribution for multiple measures can be justified by the central limit theorem.

### 3.2.4 Discriminant function analysis

Discriminant function analysis is an appropriate analytical technique for elucidating the differences between two or more groups. Discriminant function is a linear combination of two or more independent variables that used to predict membership in naturally occurring groups. Discrimination is accomplished by identifying the relative contribution of number of variables to separation of the groups and denoting it as the weighted coefficients corresponding to each variable. The linear discriminant function for 'p' variables can be written as

$$\boldsymbol{Z = a_1X_1 + a_2X_2 + a_3X_3 + \cdots \ldots \ldots + a_pX_p}$$

where $a_1,a_2,a_3,\ldots.a_p$ are the weighted coefficients such that the ratio of variance between groups to within groups will be maximum.

Discriminant score for each object is the summation of the values obtained by multiplying each independent variable by its discriminant weight. Group mean is obtained by averaging the discriminant scores for all the individuals within a particular group. There will be two group averages for analysis involving two groups. Comparison of group average reveals how far the groups are in terms of Discriminent function.The average of group average provides overall average which is the basis for the allocation of new objects into groups. The overall average is given by

$$\bar{Z} = \frac{\overline{Z_1} + \overline{Z_2}}{2}$$

Where $\bar{Z}$ is the overall average,

$\overline{Z_1}$ is the average of first group,

$\overline{Z_2}$ is the average of second group.

Discriminant score of each object is compared with the overall average. If the discriminant score of an object is greater than the overall average, it will assign to the group whose group average is greater than the overall average and vice versa. Effectiveness of Discriminent analysis is obtained by counting number of objects following the grouping rule. More number of objects following the grouping rule indicates better discrimination.

## 3.3 CLUSTER ANALYSIS

It is a multivariate technique of data analysis that divides the data into groups. The aim is to construct groups with homogeneous properties out of heterogeneous large samples. The groups or clusters should be as homogeneous as possible and the differences among the various groups as large as possible. Steps in cluster analysis include:

1. Choice of proximity measures: A similarity (proximity) measure is defined to measure the "closeness" of the objects. The "closer" they are, the more homogeneous they are.
2. Choice of group building algorithm: On the basis of proximity measures the objects are assigned to groups so that differences between groups become large and observations in a group become as close as possible (Hardle and Simar, 2007).

### 3.3.1 Similarity measures

A measure of closeness is necessary for performing cluster analysis. Similarity coefficient or dissimilarity coefficient are used to measure the proximity between two objects. Measure of similarity and dissimilarity are closely related in an inverse way,

i.e. measure of dissimilarity = (constant – measure of similarity)

The dissimilarity coefficients are required to satisfy the following conditions:

$d_{xy} \geq 0$ , for every x,y

$d_{xx} = 0$ , for every x

$d_{xy} = d_{yx}$ , for every x,y

where $d_{xy}$ denotes the dissimilarity of x and y

Similarity measures obtained by using different methods are not unique and selection of appropriate measure depends on the nature of variables. Different methods are used for binary, continuous and mixed data.

#### *3.3.1.1 Similarity measure for qualitative data*

Data that cannot be expressed as numbers are known as the qualitative data. They give the information that can't actually be measured like colour, fragrance etc. Generally they are measured as presence or absence of character. If they are measured in terms of presence or absence the data form binary structure.Objects with binary structure are data whose unit can take on only two possible states i.e. 0 and 1.In order to measure the similarity between objects we always compare pairs of observation $(x_i,x_j)$

where $x_i^T = (x_{i1},\ldots\ldots x_{ip})$, $x_j^T = (x_{j1},\ldots,x_{jp})$, and $x_{ik},x_{jk} \varepsilon \{0,1\}$. Obviously there are four classes:

$x_{ik} = x_{jk} = 1$; $x_{ik} = 0$, $x_{jk} = 1$; $x_{ik} = 1$, $x_{jk} = 0$; $x_{ik} = x_{jk} = 0$

Define $a_1 = \sum_{k=1}^{p} I(x_{ik} = x_{jk} = 1)$,

$a_2 = \sum_{k=1}^{p} I(x_{ik} = 0, x_{jk} = 1)$,

$a_3 = \sum_{k=1}^{p} I(x_{ik} = 1, x_{jk} = 0)$,

$a_4 = \sum_{k=1}^{p} I(x_{ik} = 0, x_{jk} = 0)$

p = no. of observations

Qualitative character can also represent in the form of categorical variable. All similarity measures used in binary data cannot use in such cases. Different similarity measures for binary data were listed by Gower (1985).

#### *3.3.1.1.1 Jaccard coefficient*

Coefficient which regards positive matches only as indicating similarity, with negative matches treated as missing. It is used in case of binary data ( Jaccard, 1908).

$$S = \frac{a_1}{a_1 + a_2 + a_3}$$

#### *3.3.1.1.2 Dice coefficient*

Coefficient used in case of binary data which give double weights for positive matches and zero weights for negative matches.

$$S = \frac{2a_1}{2a_1 + (a_2 a_3)}$$

### 3.3.1.1.3 Hamann's coefficient

It is a measure of similarity in multi- state data using number of matching and unmatching between objects.

$$C = \frac{(m - u)}{n}$$

where m is the number of matched observation,

u is the number of unmatched observation,

n = m+u

### 3.3.1.1.4 Simple matching coefficient

Used to measure similarity in case of multi- state data by considering number of matches and total number of observation

$$C = \frac{m}{n}$$

where m is the number of matched observation,

u is the number of unmatched observation,

n = m+u

### 3.3.1.2 Similarity measures for quantitative data

Quantifiable data are called quantitative data and which can be expressed in numbers. Similarity measures for quantitative data are different from measures of qualitative data. Important similarity measures for quantitative data generally used are given below. Gower (1985) listed different similarity measures for continuous data.

#### *3.3.1.2.1 Euclidian distance*

Most commonly used method which is referred to as straight line distance. The Euclidian distance measure between i<sup>th</sup> and j<sup>th</sup> individual is given as

$$D(X,Y) = \sqrt{\sum_{i=1}^{p}(X_i - Y_i)^2}$$

Euclidean distance between two vectors X= $(x_1, x_2, \ldots, .x_p)'$ and Y= $(y_1, y_2, \ldots\ldots, y_p)'$, defined as

$$D(X,Y) = \sqrt{(X-Y)'(X-Y)}$$

#### *3.3.1.2.2 Squared Euclidian distance*

It gives greater weight on the objects that are farther apart. The distance measure between i<sup>th</sup> and j<sup>th</sup> individual is given as

$$D^2(X,Y) = \sum_{i=1}^{p}(X_i - Y_i)^2$$

For two vectors X and Y

$$D^2(X,Y) = (X-Y)'(X-Y)$$

### 3.3.1.2.3 *Minkowski metric*

The Minkowski metric is a measure of distance such that

$$D(X,Y) = \left\{\sum_{i=1}^{p} |X_i - Y_i|^r\right\}^{\frac{1}{r}}$$

where r≥ 1

When r = 2, this reduces to Euclidian distance. For p = 2 and r =1, measures the "City block" distance between two observations.

### 3.3.1.2.4 *City block (Manhattan) distance*

It uses the sum of the absolute differences of variable. The distance measure is given by

$$D(X,Y) = \sum_{i=1}^{p} |X_i - Y_i|$$

### 3.3.1.2.5 *Chebychev's distance*

Distance measure in which dissimilarity is the major point of interest and measure is based on the assumption that two objects are different if they differ in any one of the characteristics.

$$D(X,Y) = Maximum|X_i - Y_i|$$

#### *3.3.1.2.6 Mahalanobis $D^2$ statistics*

A measure of distance between two population based on multiple characters was given by Mahalanobis, 1936. With $x_1, x_2, x_3, \ldots, x_p$ as multiple measurements available on each individual, Mahalanobis $D^2$ statistics is defined as follows:

$$D^2 = (\bar{X}_1 - \bar{X}_2)' W^{-1} (\bar{X}_1 - \bar{X}_2)$$

Where $W^{-1}$ is the inverse of variance covariance matrix, $\bar{X}_1$ is the mean of first population, $\bar{X}_2$ is the mean of second population. It is used for quantitative data.

#### *3.3.1.3 Similarity measures for mixed data*

Mixed data indicate data that contain both quantitative and qualitative characters.

#### *3.3.1.3.1 Gower's measure*

Gower (1971) suggests the similarity measure for mixed data as

$$S_{ij} = \frac{\sum_{k=1}^{p} w_{ijk} s_{ijk}}{\sum_{k=1}^{p} w_{ijk}}$$

where $S_{ijk}$ is a measure of similarity based on the variable $x_k$.

$$S_{ijk} = \begin{cases} 0 \; if \; X_{ik} = X_{jk} \\ 1 \; if \; X_{ik} \neq X_{jk} \end{cases} \text{ for binary or categorical data,}$$

$$S_{ijk} = \frac{|X_{ik} - X_{jk}|}{r_k} \text{ for continous data , where } r_k \text{ is the range of } x_k.$$

$w_{ijk} = 0$ when $x_k$ is missing on both objects

$w_{ijk} = 1$ , otherwise

### 3.3.2 Clustering methods

The frequently used methods of clustering classified into two general categories, Hierarchical clustering and Non hierarchical clustering method (Johnson and Wichern, 2006).

#### *3.3.2.1 Hierarchical clustering methods*

It is done by either a series of merges or a series of successive divisions. They are classified under two categories namely agglomerative hierarchical method and divisive hierarchical method.

#### *3.3.2.1.1 Divisive hierarchical method*

It works in the reverse direction of agglomerative clustering method. Initial stage contains only one group of all objects. Single group is divided into two sub groups in a way that the objects in one subgroup are more dissimilar than the objects in the other. Further division of these subgroups results in more dissimilar subgroup. This process continues until each object form a group.

#### *3.3.2.1.2 Agglomerative hierarchical method*

It is a method which starts clustering with single objects. Thus the number of clusters at initial stage will be equal to number of objects to cluster. The most similar objects forms the groups and these initial groups merged according to the similarities between groups. All the subgroups will cluster together as the similarity increases and finally there will be a single cluster consisting of all objects.

Steps in agglomerative clustering for groups of N objects

i. Start with N clusters, each containing a single entity and form an N*N symmetric matrix of distance or similarities

ii. Group the objects having smallest dissimilarity (highest similarity). Let the distance between most similar clusters A and B be $d_{AB}$ .

iii. Merge clusters A and B. Label the newly formed cluster (AB). Update the entries in the distance matrix by

  (1) Deleting rows and columns corresponding to clusters A and B.

  (2) Adding row and column giving distance between cluster (AB) and remaining clusters

iv. Repeat steps (ii) and (iii) a total of N-1 times.

### *3.3.2.1.3 Single linkage clustering technique*

It is also known as nearest neighbour clustering technique. In this method, the first cluster is formed by joining two objects having minimum distance (nearest neighbour). In the succeeding step, either a third object will join them or another two closest unclustered objects are joined to form another cluster. This depends on whether the distance from one of the unclustered object to the first cluster is short than the distance between the two closest unclustered or not. The major drawback of this method is that it tends to produce long thin clusters in which nearby elements of the same clusters have small distances, but elements at opposite ends of a cluster may be much farther than to elements of other clusters.

$$\boldsymbol{d\,(A\,,B) = \;min\,d\,(x,y), x \in A, y \in B}$$

where d (A ,B ) is the distance between cluster A and B and d (x, y) is the distance between x and y.

### *3.3.2.1.4 Complete linkage clustering technique*

It is also an agglomerative method. In this method, the object having maximum distance (farthest distance) between them constitutes two groups. Next object either join one of the previous two clusters or form its own group following the above rule as in the case of first two objects and the process continues till all objects are covered. This method is also known as farthest neighbour clustering. It avoids the drawback of single linkage method i.e. chaining phenomenon.

$$d\ (A, B) = \ max\ d\ (x, y), x \in A, y \in B$$

where d (A ,B ) is the distance between cluster A and B and d (x, y) is the distance between x and y.

*3.3.2.1.5 UPGMA (Unweighted Pair Group Average Method)*

This agglomerative hierarchical clustering method was given by Sokal and Michener (1973). It is a method which uses average similarity of all pairs of objects instead of of taking only a distance between the closest or the farthest neighbour. Distance between a cluster and an object is calculated as the average distance between all the objects in the cluster and the objects suppose to enter in to the cluster.

$$d\ (A\ , B) = \frac{\sum_{x \in A} \sum_{y \in B} d(x, y)}{(n_A * n_B)}, x \in A, y \in B$$

where d (A ,B ) is the distance between cluster A and B

d (x, y) is the distance between x and y.

$n_A$ is the number of objects in cluster A

$n_B$ is the number of objects in cluster B

*3.3.2.1.6 WPGMA (Weighted pair group average method*

It is a modified form of Unweighted pair group average method used when the cluster sizes are suspected to be greatly uneven. In this method similarity between two clusters equals the mean similarity of previously existing clusters when they are grouped and average always involves only two terms and does not weight clusters by their size.

Distance of an unclustered object k to the cluster AB is given as

$$d\ (AB), k = \frac{d_{Ak} + d_{Bk}}{2}$$

where $d_{Ak}$ is the distance between A and k and $d_{Bk}$ is the distance between B and k.

### *3.3.2.1.7 UPGMC (Unweighted Pair Group Centroid Method)*

In a cluster of points, the centroid is the point that has the average coordinates of all the objects of the cluster. UPGMC joins the objects or groups that have the smallest distance and by replacing all the objects of the group by the centroid of that group. This centroid is considered as a single object at the next clustering step. A simple manner to achieve this is to replace, in the similarity matrix, the two rows and columns corresponding to the two objects about to join by a single series obtained by computing the averages of the similarities of the two objects with all the others.

$$\boldsymbol{d\ (A\ ,B) = \ d\ (A_c\ ,B_c)}$$

where d (A ,B ) is the distance between cluster A and B

$A_c$ is the centroid of cluster A and

$B_c$ is the centroid of cluster B

### *3.3.2.1.8 Tocher's method of clustering and its modification*

This method is widely used in clustering quantitative data based on Mahalanobis Distance. The $D^2$ values are arranged in ascending order. First cluster is formed by joining objects having smallest distance. A third object having smallest average $D^2$ value from the first two objects is added. The procedure repeats until an abrupt change in the average $D^2$ value is noticed after the addition of particular object in the particular cluster. Tocher fixed the level of abrupt as the max among min $D^2$ values. A new cluster is started at this stage, starting with the discarded object from the previous cluster. Thus the process of clustering continues till the objects are included in one or the other cluster. Cluster configuration can be improved by a modification over Tocher method. In this genotypes are re-locate from the clusters in which they are placed based on Tocher method. Take out each genotype one by one from the cluster to which it was allotted and calculate the average $D^2$ values between these genotypes and each cluster. Allocate each genotype into that cluster where the average $D^2$ value is found minimum. The iteration has to be continued till two successive iterations end up with same cluster configuration (Suresh, 1986).

### *3.3.2.1.9 Ward's clustering technique*

This method is different from linkage methods as it uses an analysis of variance approach to calculate the similarity of clusters. This procedure is based on minimizing the loss of information from joining two groups. It is also called the incremental sum of squares method, uses the within cluster distance and the between cluster distances (Ward, 1963). First for a given cluster k, let $ESS_k$ be sum of squares due deviation. If there are currently k clusters, the total $ESS = ESS_1 + ESS_2 + \ldots\ldots + ESS_k$. At each step in the analysis the union of every possible pair of cluster is considered and the two clusters whose combination results in the smallest increase in ESS are joined. Initially each cluster consist of a single item, and if there are N items, $ESS_k = 0$, $k = 1,2,\ldots.,N$ so ESS= 0 at the extreme, when all the clusters are combined in a single group of N items, the value of ESS is

$$ESS = \sum_{j=1}^{N} (X_j - \bar{X})'(X_j - \bar{X})$$

where $X_j$ is the multivariate associated with $j^{th}$ item and $\bar{X}$ is the mean of all the items. The results of Ward's method can be displayed by a Dendrogram. The vertical axis gives the value of ESS at which the merges occur.

### *3.2.5.2.1.10 Hierarchical trees*

Tree is a family of clusters for which any two clusters are either disjoint or one includes the other (Hartigan, 1975). The hierarchical structure is often represented by a two dimensional diagram called as tree diagram or dendrogram. The tree is often presented upside down so that the branches are at the bottom and the roots of the tree is at the top. We can illustrate the merges or divisions that have been made at successive levels using dendrogram.

### *3.3.2.2 Non Hierarchical clustering method*

In this case the number of clusters, k, may either be specified in advance or determined as part of the clustering procedure. Here matrix of distance does not have to be determined and the basic data do not have to be stored during the entire process.

### *3.3.2.2.1 k means approach*

This is a partioning approach and it allows the items to be moved from one cluster to another, a reallocation that is not available in the hierarchical methods.

i. Divide the data into k initial clusters.
ii. Calculate the means or centroid of the k clusters.
iii. For a given case, calculate its distance to each centroid. If the case is closest to the centroid of its own cluster, leave it in that cluster; otherwise reassign it to the cluster whose centroid is closest to it
iv. Repeat step (iii) for each case.
v. Repeat steps (ii), (iii) and (iv) until no cases are reassigned.

### *3.3.2.3 Intra cluster distance*

Average distance of objects within a cluster is called as intra cluster distance. It can be measured using the formula $\sum \frac{D_i^2}{n}$ , where $\sum D_i^2$ is the sum of distances between all possible combinations (n) of the populations included in a cluster.

### *3.3.2.4 Inter – Cluster distance*

Average distance of objects between two cluster is called as inter cluster distance. To find inter- cluster distance we have to find out distances between all the populations of considering clusters.

## 3.3.3 Cluster validity

### *3.3.3.1 Pseudo $t^2$ statistic*

The pseudo t-square statistic for the clustering of two clusters A and B is given by:

$$Pseudo\ t^2 = \frac{B_{AB}}{((W_A + W_B)/(N_A + N_B - 2))}$$

Where $N_A$ and $N_B$ are the number of observations in clusters A and B,

$W_A$ and $W_B$ are within cluster sum of squares of clusters A and B

$B_{AB}$ is the between-cluster sum of squares.

It measures the difference between two clusters merged at a given step. If the pseudo *t*-square statistic has a distinct jump at step *k* of the hierarchical clustering, then the clustering in step *k+1* is selected as the optimal cluster. This index can be use for hierarchical clustering if the methods used are UPGMA, UPGMC or Ward's method.

### *3.3.3.2 Pseudo F statistic*

The pseudo *F* statistic narrates the ratio of between-cluster variance to within cluster variance (Calinski and Harabasz, 1974):

$$Pseudo\ F = \frac{(GSS)/(K-1)}{WSS/(N-K)}$$

where N is the number of observations,

K is the number of clusters at any step in the hierarchical clustering,

GSS is the between-group sum of squares,

WSS is the within group sum of squares.

Large values of Pseudo F indicate optimum number of clusters and index can use for k-means clustering.

### *3.3.3.3 SD validity index*

The SD validity index measures the average scattering and total separation of clusters. Scattering is obtained by calculating variance of the clusters and the variance of the complete dataset. For a compact cluster, variance of the cluster will be smaller than the variance of dataset and a low Scatt measure. The average scattering for a cluster is defined as:

$$Scatt = \frac{1}{n_c}\sum_{i=1}^{n_c}\frac{\|\sigma(v_i)\|}{\|\sigma(x)\|}$$

The separation of clusters based on the distance of cluster centre points is given as

$$Dist = \frac{\max\left(\left\|v_j - v_i\right\|\right)}{\min\left(\left\|v_j - v_i\right\|\right)}\sum_{i=1}^{k}\left(\sum_{z=1,z\neq t}^{k}\left(\left\|v_j - v_i\right\|\right)^{-1}\right.$$

SD index is defined as:

$$SD = \propto Scatt + Dist$$

Where $\alpha$ is a weighing factor. Lower SD index indicates better cluster configuration.

## 3.4 PRINCIPAL COMPONENT ANALYSIS

The principal component analysis is a multivariate technique that is used to get a small number of linear combinations (principal components) of a set of variables that retain as much information in the original variables as possible. This technique helps to transform the original set of variables to a new set of uncorrelated mutually orthogonal random variables which are linear combinations of the original variables. The principal components derived in decreasing order of importance so that the first principal component accounts for the maximum variation in the original data.

For a random vector $X = (X_1, X_2, \ldots, X_p)'$ with symmetric and non negative definite dispersion matrix $\Sigma$, the measured variables $X_1, X_2, \ldots, X_p$ can be transformed into principal components $Y_1, Y_2, \ldots, Y_p$ by means of the linear transformations

$$Y_i = a_{i1}X_1 + a_{i2}X_2 + \ldots + a_{ip}X_p \quad (i = 1, 2, \ldots, p)$$

$$Y_i = a_i^T X$$

so that variance of $Y_i$ is as large as possible subject to the condition that $a_1^T a_1 = 1$.

where $a_i^T = (a_{i1}, \ldots, a_{ip})$ is a vector of constant.

$$Var\,(Y_I) = Var\,(a_1^T X)$$

$$= a_1^T \Sigma a_1$$

To find out the coefficient vector that maximizes the variance of principal component subject to the constraint $a_1^T a_1 = 1$, we use Lagrange multiplier method with $a_1^T \Sigma a_1$ as objective function.

Lagrange function $L(x) = f(x) - \lambda (g(x) - c)$

i.e. $L(a_1) = a_1^T \Sigma a_1 - \lambda (a_1^T a_1 - 1)$

On differentiation with $a_1$ we will get,

$$\frac{\partial L}{\partial a_1} = 2\Sigma a_1 - 2\lambda a_1$$

Equating to zero we get,

$$(\Sigma - \lambda I)a_1 = 0$$

where I is the unit matrix. The equation will have a solution for a1 if $(\Sigma - \lambda I)$ is a singular matrix. The value of λ must be chosen so that

$$|\Sigma - \lambda I| = 0$$

Thus the non zero solution for the equation $(\Sigma - \lambda I)a_1 = 0$ exists if and only if λ is an eigen value of Σ. There will be 'p' eigen values denoted as $\lambda_1, \lambda_2, \ldots\ldots\ldots\ldots, \lambda_p$ and $\lambda_1 > \lambda_2 > \ldots\ldots\ldots\ldots > \lambda_p \geq 0$. As we want to maximize the variance we choose largest eigen value for first principal component and $a_1^T$ is the normalised eigenvector corresponding to the largest eigen value $\lambda_1$ of Σ. The variance explained by principal components will be equal to the corresponding eigen values. Total variance is given as

$\sum_{i=1}^{p} \lambda_i$ = trace of Σ = sum of variance = $\sum_{i=1}^{p} V(x_i)$

The variance of i<sup>th</sup> principal component is given by

$V(PC_i) = \lambda_i$ and percent of information given by $PC_i = \frac{\lambda_i * 100}{\sum_{i=1}^{p} \lambda_i}$

The importance of the component is evaluated by means of the percentage of the total variation it explains. The first few principal components accounts for most of the variation in the original data and these principal components can then replace the variables in subsequent analysis. Thus principal component analysis effectively reduces the dimensionality. It also removes the multicollinearity in the data by deriving uncorrelated principal components.

The ability of PCA to reveal the structure of data helps to identify the clustering in data. Plotting of scores of the first two or three components for each genotype is a useful way of finding clusters in the data. Graph plotted using components which accounts for maximum variation reveals distinct groups of objects.

### *3.4.1 Score plot*

Score of each object obtained by taking sum of products of component loadings and corresponding mean value are used for score plot. Score corresponding to two or three components are plotted in 2 or3 dimension scale to find out natural groupings.

# Results and Discussion

# 4. RESULTS AND DISCUSSION

The results acquired by the application of suitable statistical techniques on the secondary data collected from the field experiment entitled 'Characterization and genetic improvement in Rose (*Rosa* spp.) through mutagenesis' (2014-2017) at College of Agriculture, Vellayani and Regional Agricultural Research Station (RARS), Ambalavayal, Wayanad are given below. Twenty five cultivars each coming under the Hybrid Tea and Floribunda groups were evaluated in the present study.

## 4.1 THE PRELIMINARY STATISTICAL ANALYSIS

Analysis of variance was done for each of the character under study which showed significant difference among different genotypes with respect to each character. Analysis of variance of different characters for Hybrid Tea type are given in Table 6 and for Floribunda Type are given in Table 7 along with their F values. The mean values of various characters corresponding to different genotypes are shown in the Table 8 and Table 9.

## 4.2 ANALYSIS OF DISPERSION

The total variation was split up into variation between groups and within groups by analysis of dispersion method. The Wilk's lamda value obtained was 0.004 (Hybrid tea) and 0.003(Floribunda). The corresponding $V_{(stat)}$ is 728.833for Hybrid tea type and 766.807 for Floribunda type which is distributed as chi-square with 216 degrees of freedom and this was significant at one percent level. The results showed that there was significant difference between the varietal means with respect to character under study.

Table 6. Mean values of characters of Hybrid Tea type

| Sl. No | Hybrid Tea Genotypes | No. of leaves at first flower | No. of days to first flower | Prickle density/5 cm | Flower size (cm) | Flower weight (g) | Pedicel length (cm) | No. of petals/ flower | Size of petals (cm) | No. of flower per plant/bunch |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Madame George Delbard | 25.67 | 195.83 | 6.50 | 36.12 | 3.17 | 4.72 | 18.00 | 5.35 | 1.50 |
| 2 | Aiswarya | 42.50 | 191.67 | 9.33 | 42.38 | 6.87 | 4.47 | 25.33 | 12.92 | 1.33 |
| 3 | Christ of Colomb | 28.50 | 182.50 | 6.00 | 20.14 | 4.98 | 6.92 | 28.67 | 5.70 | 1.33 |
| 4 | Pink Panther | 28.50 | 193.83 | 2.83 | 22.67 | 4.63 | 6.47 | 27.83 | 3.68 | 1.50 |
| 5 | Roughe Miland | 52.67 | 204.67 | 2.33 | 23.97 | 4.51 | 5.84 | 33.00 | 3.40 | 1.50 |
| 6 | Shrewsbury show | 61.67 | 203.17 | 6.50 | 34.63 | 7.62 | 6.67 | 55.00 | 3.66 | 2.00 |
| 7 | Alaine Souchen | 24.33 | 203.17 | 6.33 | 27.46 | 6.09 | 5.85 | 45.17 | 4.37 | 1.67 |
| 8 | Amara | 59.67 | 207.50 | 3.50 | 53.04 | 7.17 | 7.30 | 17.33 | 6.66 | 2.00 |
| 9 | Fryat | 54.17 | 160.00 | 6.33 | 23.81 | 4.89 | 5.85 | 33.67 | 6.79 | 2.00 |
| 10 | Perfume Perfect | 61.50 | 192.33 | 2.00 | 26.41 | 5.00 | 8.07 | 32.50 | 7.41 | 2.00 |
| 11 | Silver Star | 24.67 | 182.83 | 1.83 | 37.86 | 5.60 | 4.92 | 25.50 | 11.60 | 1.50 |
| 12 | Lincoln Cathedral | 57.67 | 211.00 | 3.67 | 23.32 | 4.80 | 4.98 | 23.67 | 3.56 | 1.50 |
| 13 | A tago | 53.33 | 204.67 | 3.17 | 32.65 | 6.99 | 6.25 | 62.33 | 5.15 | 1.67 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 14 | Demestra | 46.00 | 164.00 | 4.00 | 41.65 | 1.76 | 4.12 | 18.17 | 1.43 | 2.83 |
| 15 | Golden Fairy Sport | 31.17 | 182.00 | 5.83 | 32.89 | 2.29 | 6.15 | 13.67 | 14.46 | 2.33 |
| 16 | Mary Jean | 82.67 | 197.00 | 11.67 | 32.51 | 6.13 | 5.52 | 23.83 | 7.25 | 2.00 |
| 17 | Toplesse | 17.67 | 170.00 | 2.33 | 22.60 | 5.81 | 5.80 | 22.83 | 9.73 | 1.50 |
| 18 | Priority Pride | 66.17 | 178.17 | 3.33 | 22.26 | 5.19 | 5.72 | 43.50 | 5.40 | 1.33 |
| 19 | Majestic | 26.50 | 190.33 | 4.33 | 23.36 | 7.09 | 5.53 | 14.83 | 6.64 | 2.33 |
| 20 | Prince Jardiner | 55.83 | 191.83 | 3.17 | 33.38 | 3.80 | 5.69 | 18.00 | 8.99 | 2.00 |
| 21 | Cel b Lau | 53.00 | 205.67 | 8.50 | 23.82 | 6.36 | 5.85 | 31.00 | 5.30 | 1.83 |
| 22 | Lois Wilson | 34.33 | 165.83 | 4.33 | 49.08 | 5.98 | 6.48 | 26.00 | 10.41 | 1.00 |
| 23 | Mom’s Rose | 43.67 | 166.50 | 4.00 | 44.00 | 5.06 | 4.93 | 23.00 | 8.97 | 1.33 |
| 24 | Alabama | 65.67 | 110.00 | 3.83 | 42.32 | 5.83 | 5.98 | 14.67 | 17.69 | 1.33 |
| 25 | Josepha | 25.33 | 203.17 | 6.33 | 26.61 | 6.06 | 6.02 | 45.17 | 4.43 | 1.50 |

Table 7. Mean values of characters of Floribunda type

| Sl. No | Floribunda Genotypes | No. of leaves at first flower | No. of days to first flower | Prickle density/5cm | Flower size (cm) | Flower weight (g) | Pedicel length (cm) | No. of petals/ flower | Size of petals (cm) | No. of flower per plant/bunch |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Versailles | 70.67 | 194.00 | 2.67 | 22.25 | 1.83 | 3.60 | 25.00 | 3.79 | 1.83 |
| 2 | Tickled Pink | 49.17 | 205.67 | 3.50 | 21.74 | 3.75 | 4.45 | 35.00 | 3.91 | 1.83 |
| 3 | Rosarale de Chateau | 65.33 | 171.33 | 3.17 | 25.03 | 4.90 | 4.25 | 37.33 | 4.85 | 1.17 |
| 4 | Rose Mary Gandhi | 62.50 | 207.17 | 9.17 | 14.11 | 1.42 | 4.42 | 24.50 | 2.76 | 1.33 |
| 5 | Princess de Monaco | 46.00 | 195.00 | 2.00 | 21.19 | 4.43 | 5.23 | 33.67 | 6.91 | 2.00 |
| 6 | Ochi di Fita | 12.83 | 198.67 | 8.00 | 13.31 | 1.61 | 4.68 | 30.83 | 3.48 | 2.00 |
| 7 | Carry Free Beauty | 35.17 | 171.67 | 2.33 | 15.04 | 3.07 | 3.53 | 45.50 | 3.75 | 1.50 |
| 8 | Sterntaler | 34.17 | 183.00 | 6.83 | 23.26 | 5.87 | 6.93 | 33.67 | 5.06 | 1.67 |
| 9 | Orange N Lemon | 53.67 | 181.67 | 8.50 | 19.98 | 5.40 | 6.40 | 18.83 | 5.55 | 3.83 |
| 10 | Lisa | 83.83 | 196.33 | 3.00 | 18.60 | 2.14 | 4.52 | 22.83 | 3.91 | 2.33 |
| 11 | The Mccartney Rose | 54.00 | 198.83 | 12.17 | 22.79 | 6.21 | 5.73 | 16.00 | 6.79 | 1.67 |
| 12 | Chesire | 53.50 | 198.83 | 12.17 | 11.19 | 1.60 | 4.00 | 16.67 | 5.02 | 1.83 |
| 13 | Monnalisa | 52.67 | 203.50 | 3.17 | 25.49 | 6.50 | 4.35 | 54.67 | 3.70 | 2.17 |
| 14 | Carolanne | 45.83 | 196.33 | 2.33 | 37.47 | 5.49 | 4.38 | 14.67 | 5.03 | 2.50 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 15 | City of Glasgow | 32.17 | 168.50 | 2.33 | 23.72 | 3.46 | 4.45 | 22.83 | 5.77 | 1.50 |
| 16 | Messara | 32.17 | 211.50 | 3.17 | 19.99 | 3.49 | 4.97 | 11.00 | 4.77 | 2.33 |
| 17 | Michel Fish | 50.17 | 174.00 | 5.33 | 25.20 | 3.38 | 2.62 | 14.50 | 5.82 | 3.33 |
| 18 | Mini Pink | 22.83 | 193.50 | 3.00 | 23.84 | 4.43 | 3.60 | 25.17 | 5.06 | 1.67 |
| 19 | Sans Souci | 93.17 | 170.00 | 3.83 | 35.16 | 6.22 | 6.33 | 33.83 | 9.77 | 1.67 |
| 20 | Schloss Elutin | 83.83 | 203.17 | 7.17 | 22.74 | 5.61 | 4.62 | 34.67 | 5.18 | 2.33 |
| 21 | Lasting Piece | 17.17 | 186.00 | 11.50 | 32.86 | 5.58 | 6.78 | 16.17 | 14.82 | 1.67 |
| 22 | Plantein on Blumen | 27.50 | 193.83 | 6.17 | 15.89 | 5.08 | 4.72 | 23.67 | 4.96 | 1.83 |
| 23 | Winchester Cathedral | 57.50 | 172.50 | 13.17 | 21.09 | 5.48 | 5.06 | 24.83 | 4.47 | 2.33 |
| 24 | Golden Fairy | 74.83 | 191.17 | 2.67 | 33.72 | 3.20 | 3.80 | 12.17 | 11.89 | 2.00 |
| 25 | Prosperity | 56.67 | 197.50 | 7.33 | 21.05 | 5.08 | 3.60 | 16.17 | 4.96 | 1.67 |

Table 8. Analysis of variance for nine characters in 25 genotypes of Hybrid Tea Roses

| Sources of variation | No. of leaves at first flower | No. of days to first flower | Prickle density/5 cm | Flower size (cm) | Flower weight (g) | Pedicel length (cm) | No. of petals/ flower | Size of petals (cm) | No. of flower per plant/bunch |
|---|---|---|---|---|---|---|---|---|---|
| MSS Treatment | 1790.52** | 2883.83** | 35.99** | 522.02** | 12.88** | 4.71** | 968.99** | 89.22** | 1.02** |
| MSS Error | 5.1 | 2.16 | 1.09 | 1.93 | 0.60 | 0.22 | 2.79 | 0.30 | 0.47 |
| F value | 351.08** | 1335.11** | 32.79** | 270.31** | 21.33** | 21.20** | 347.06** | 296.62** | 2.19** |

**Significant at 1% level of significance.

Table 9. Analysis of variance for nine characters in 25 genotypes of Floribunda Roses

| Sources of variation | No. of leaves at first flower | No. of days to first flower | Prickle density/5cm | Flower size (cm) | Flower weight (g) | Pedicel length (cm) | No. of petals/ flower | Size of petals (cm) | No. of flower per plant/bunch |
|---|---|---|---|---|---|---|---|---|---|
| MSS Treatment | 2642.48** | 1007.74** | 78.17** | 265.02** | 15.55** | 6.93** | 706.81** | 44.60** | 2.07** |
| MSS Error | 4.51 | 1.95 | 1.10 | 1.26 | 0.21 | 0.09 | 3.44 | 0.30 | 0.67 |
| F value | 585.31 | 517.68 | 71.33 | 210.37 | 74.28 | 78.11 | 205.71 | 147.17 | 3.07 |

**Significant at 1% level of significance.

## 4.3 DISCRIMINANT FUNCTION ANALYSIS

The linear functions of the variable were used to elucidate the differences between two groups i.e. Hybrid Tea and Floribunda genotypes. Discriminant linear function were developed using nine quantitative characters from both of the groups. The linear discriminant function for the data were

$$y = 0.3673x_1 + 0.1047x_2 + 0.1232\,x_3 - 1.0680x_4 + 0.1064x_5 - 0.9331x_6 - 0.1189x_7 + 0.3920x_8 + 0.3951x_9$$

The high coefficient corresponding to $x_4$ revealed the highest contribution of the character flower size (-1.0680) to discriminate between two groups. Pedicel length (-0.9331) had the second most relative contribution followed by number of leaves at first flower (0.3673), size of petals (0.3920), number of flower per plant/bunch(0.3951). All other characters have relatively small contribution towards the Discriminant linear function. The discriminant score obtained for different genotypes are presented in the Table 10.

The average value for the Hybrid Tea type obtained was 11.09 and – 2.34 for Floribunda type. The mid- point of these averages gave the overall average value i.e. 4.38. We can allocate a new genotype to Hybrid Tea group if its mean value is greater than overall average value and else can add to Floribunda group. This suggests that 80% of Hybrid Tea and 72% of Floribunda were correctly classified. Discriminant function analysis reassured the difference between two groups under study.

## 4.4 ASSOCIATION MEASURES

Different association measures such as Euclidean, Squared Euclidean, Chebychev, City Block distance and $D^2$ statistics were used for quantitative data. Jaccard, Dice, Simple matching and Hamann's coefficient were used with qualitative data. Gower's measure was used for mixed data. Distance matrix obtained for Hybrid Tea and Floribunda genotypes under different distance measures are given in Appendix I to Appendix XVIII.

Table 10. Discriminant score of genotypes

| Hybrid Tea genotypes | Discriminant score | Floribunda genotypes | Discriminant score |
|---|---:|---|---:|
| H1 | -9.09 | F1 | -25.95 |
| H2 | -4.64 | F2 | -17.29 |
| H3 | 0.31 | F3 | -15.80 |
| H4 | 0.49 | F4 | -14.06 |
| H5 | 0.76 | F5 | -13.90 |
| H6 | 6.25 | F6 | -11.37 |
| H7 | 6.56 | F7 | -9.30 |
| H8 | 8.80 | F8 | -8.34 |
| H9 | 10.01 | F9 | -6.16 |
| H10 | 10.07 | F10 | -5.07 |
| H11 | 10.16 | F11 | -4.44 |
| H12 | 10.17 | F12 | -3.29 |
| H13 | 10.17 | F13 | -2.67 |
| H14 | 10.21 | F14 | -1.99 |
| H15 | 11.14 | F15 | 0.07 |
| H16 | 12.61 | F16 | 2.23 |
| H17 | 13.85 | F17 | 2.59 |
| H18 | 14.49 | F18 | 2.59 |
| H19 | 14.53 | F19 | 6.51 |
| H20 | 17.77 | F20 | 7.60 |
| H21 | 18.91 | F21 | 8.50 |
| H22 | 23.77 | F22 | 10.93 |
| H23 | 25.43 | F23 | 12.28 |
| H24 | 27.16 | F24 | 13.85 |
| H25 | 27.60 | F25 | 14.01 |
| AVERAGE | 11.09 | AVERAGE | -2.34 |

## 4.5 CLUSTERING TECHNIQUES

Clustering technique was separately carried out for Hybrid Tea and Floribunda groups as it is evident that the two groups were remarkably different from each other. Cluster analysis was done for quantitative data, qualitative data and for mixed data.

### 4.5.1 Clustering of Hybrid Tea genotypes

### *4.5.1.1 Clustering of Hybrid Tea type based on Quantitative data*

Clustering based on quantitative character used the data corresponding to the characters number of leaves at first flower, number of days to first flower, prickle density, flower size, flower weight, pedicel length, number of petals flower$^{-1}$, size of petals and number of flower plant $^{-1}$/bunch $^{-1}$. Clustering techniques such as single linkage, complete linkage, UPGMA, WPGMA, Ward's method, modified Tocher method and k-means clustering were performed using different measures of distance.

### *4.5.1.1.1 Single linkage clustering of Hybrid Tea genotypes based on quantitative data*

Single linkage clustering of 25 genotypes coming under Hybrid Tea group was carried out using the nine quantitative data. Similarity measures used were Squared Euclidean distance, Euclidean distance, Chebychev distance and City block distance. Clusters were formed by combining objects having smallest distance. Distance between two clusters was measured as the distance between nearest objects. Clustering were done using SPSS statistical package. Clustering schedule got some differences with different distance measures except for Euclidean and Squared Euclidean distance. The result of different clustering techniques based on Squared Euclidean results gave approximately same result as that of Euclidean distance. So the result corresponding to Euclidean distance is not presented separately to avoid duplication. Dendrogram of single linkage method using different distance measures are shown in Figure 1, Figure 2 and Figure 3. Cluster membership of different genotypes under single linkage method is given in Table 11.

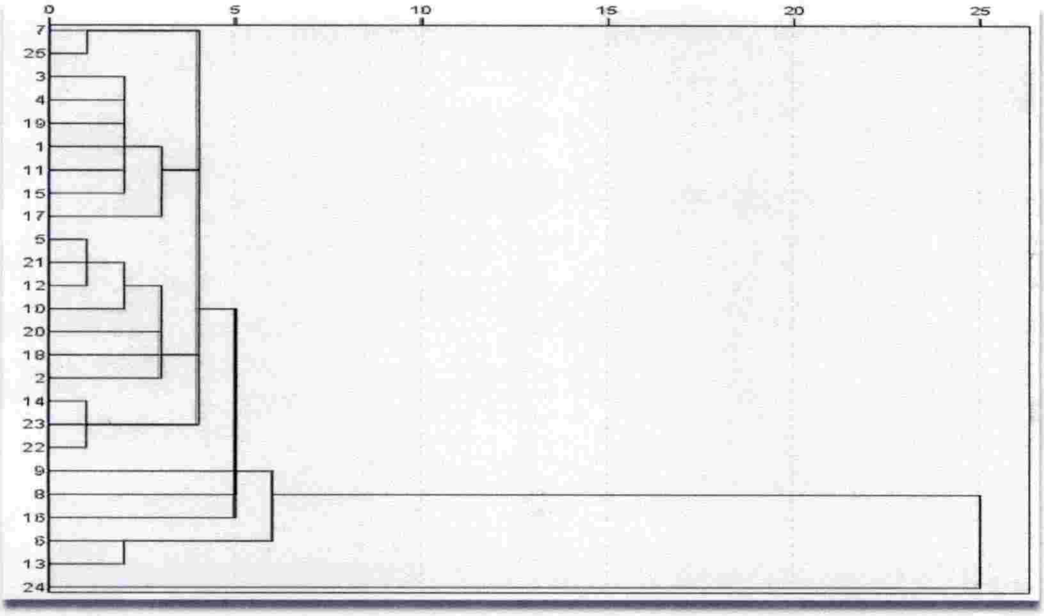


Fig. 1: Dendrogram of Hybrid Tea genotypes under single linkage clustering based on Squared Euclidean distance

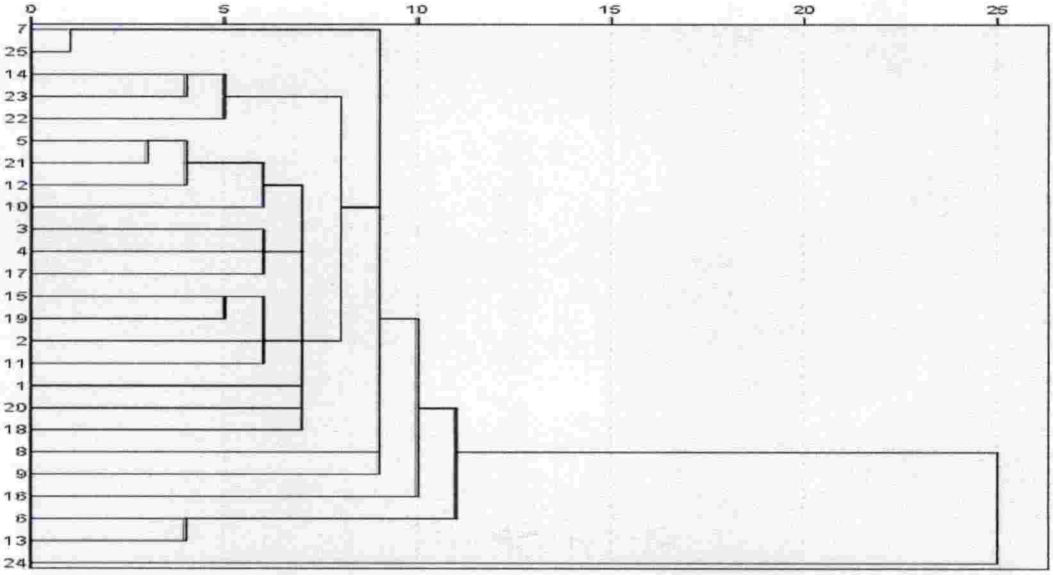


Fig. 2: Dendrogram of Hybrid Tea genotypes under single linkage clustering based on Chebychev distance.

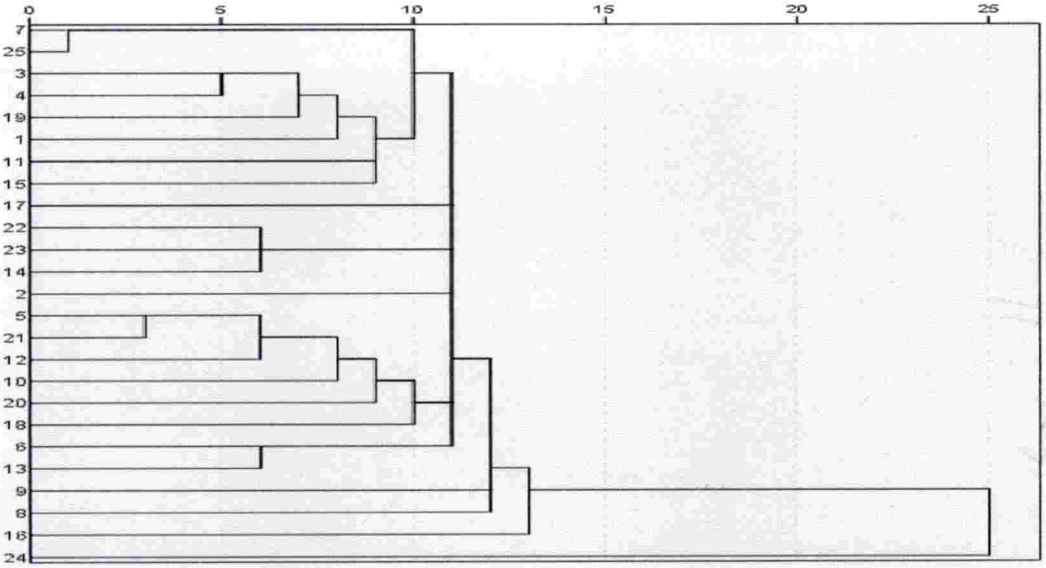


Fig. 3: Dendrogram of Hybrid Tea genotypes under single linkage clustering based on City block distance.

Table 11. Cluster membership of Hybrid Tea types under single linkage based on quantitative characters.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | H3, H4, H1, H19, H11, H15,H17 | H17, H15, H19, H2, H11, H1, H20, H18, H5, H21, H12, H10, H3, H4 | H7, H25, H3, H4., H1, H19, H11, H15 |
| 2 | H5,H21, H12, H10,H20, H18, H2 | H14, H23, H22 | H5, H21, H12, H10, H20, H18 |
| 3 | H14, H23, H22 | H6,H13 | H14, H23, H22 |
|  | H6,H13 | H7, H25 | H6, H13 |
| 5 | H8 | H16 | H2 |
| 6 | H7, H25 | H9 | H9 |
| 7 | H9 | H8 | H8 |
| 8 | H 16 | H24 | H16 |
| 9 | H24 |  | H24 |
| 10 |  |  | H17 |

Single linkage clustering of Hybrid Tea genotypes resulted in the formation of nine clusters under Squared Euclidean distance, eight clusters under Chebychev distance and ten clusters under City Block distance. Under Squared Euclidean distance the largest cluster included eight members and the second largest cluster included seven members. Cluster with fifteen members formed the largest cluster under Chebychev distance. Under City Block distance the largest cluster included eight members followed by the second largest cluster with six members. In all the distance measures H8 (Amara), H9 (Fryat), H16 (Mary Jean), H24 (Alabama) formed individual clusters. Shrewsbury Show, A Tago and Demestra, Mom's Rose and Lois Wilson came under single cluster under all the three distance measures.

#### *4.5.1.1.2 Complete linkage clustering of Hybrid Tea genotypes based on quantitative characters*

Complete clustering was performed using SPSS statistical package. Dendrograms were drawn using different distance measures. Distance between two clusters was measured as the distance between furthest objects. Clusters obtained from different distance measures were different from each other. Dengrograms obtained from different distance measures are shown below (Figure 4, Figure 5, Figure 6). Clustering pattern of genotypes under complete linkage method are given in the Table 12.

Ten clusters were formed under Squared Euclidean distance, eight under Chebychev distance and seven under City Block distance. Largest cluster contained five members under Squared Euclidean distance while under Chebychev and City Block distance largest cluster included seven members. H24 (Alabama) formed a single cluster under all the three distance measures. Clustering of all other genotypes showed some variations. H6 and H13 came under same cluster in Squared Euclidean and Chebychev distance. Similarly H7 and H25 also formed separate clusters under these association measures. But these four came under a single cluster with City Block distance.

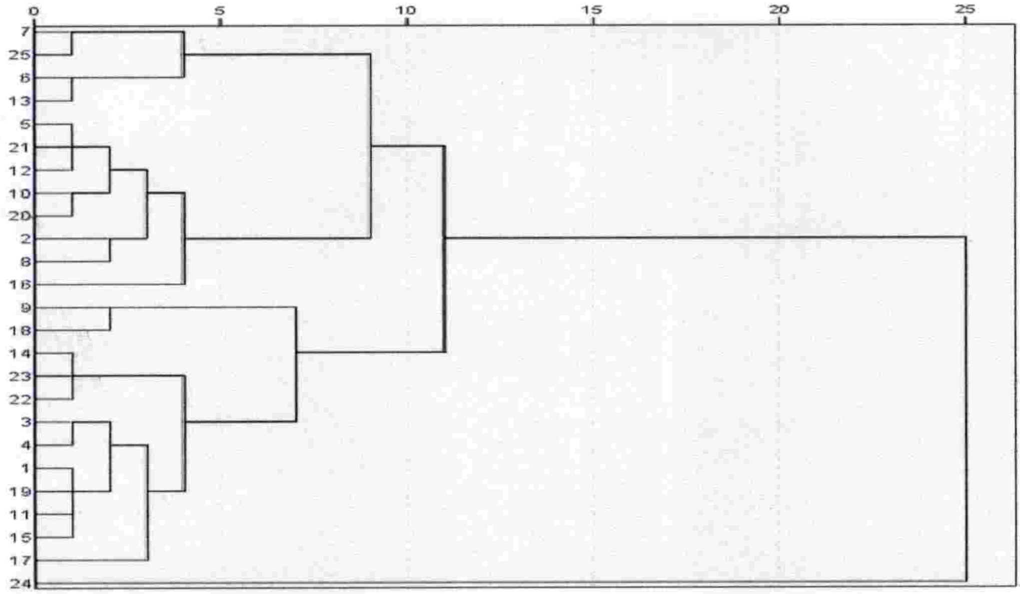


Fig. 4: Dendrogram of Hybrid Tea genotypes under complete linkage clustering based on Squared Euclidean distance.

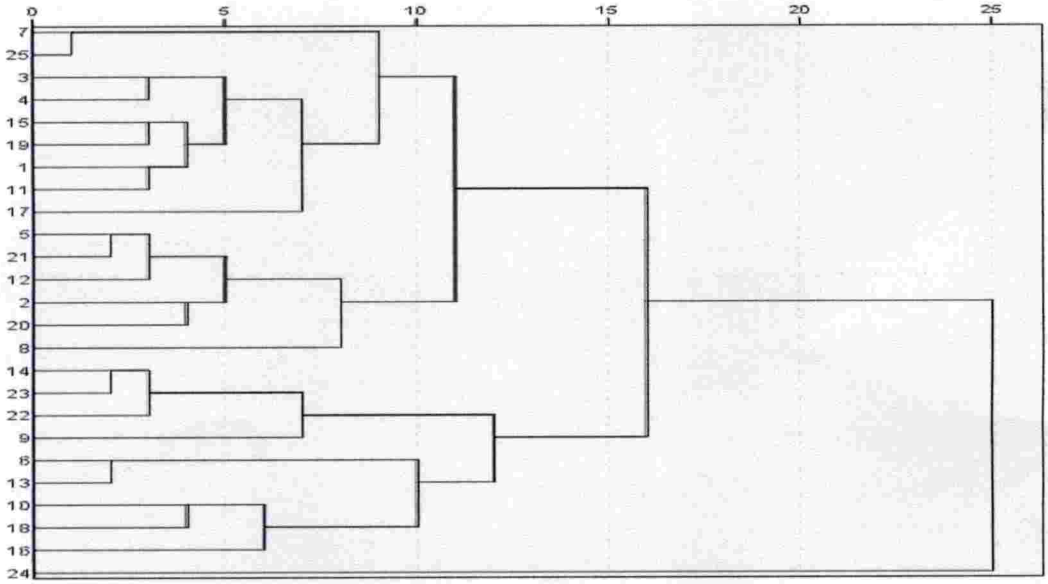


Fig. 5: Dendrogram of Hybrid Tea genotypes under complete linkage clustering based on Chebychev distance.

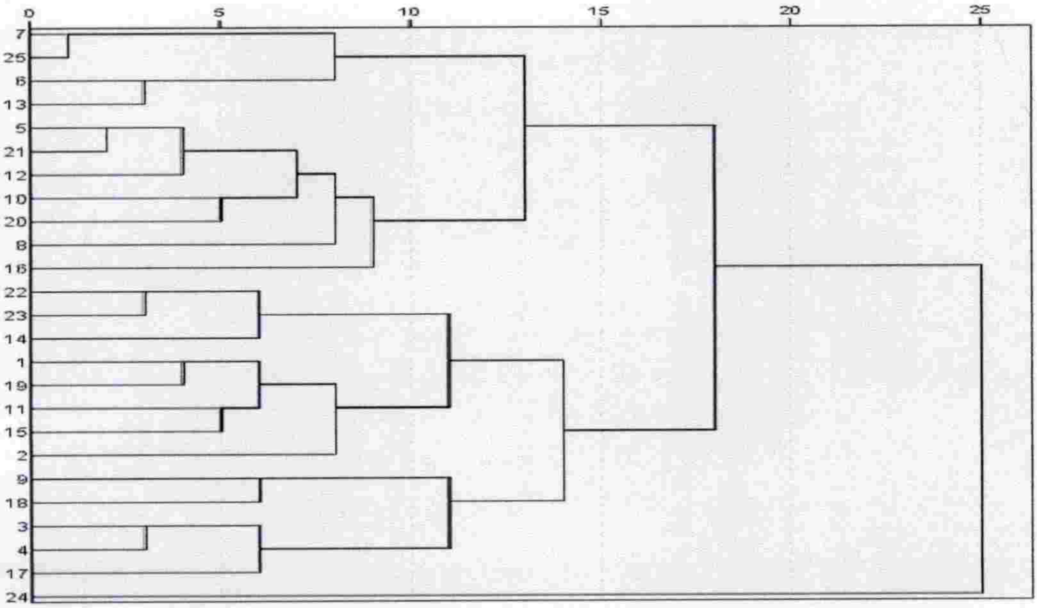


Fig. 6: Dendrogram of Hybrid Tea genotypes under complete linkage clustering based on City block distance.

Table 12. Cluster membership of Hybrid Tea types under complete linkage based on quantitative characters.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | H3, H4, H1, H19, H11, H15 | H3, H4, H1, H19, H11, H15, H17 | H5,H21, H12, H10,H20, H8, H16 |
| 2 | H5,H21, H12, H10,H20 | H5,H21, H12, H20, H2 | H1, H19, H11, H15, H2 |
| 3 | H14, H23, H22 | H14, H23, H9 | H7, H25, H6,H13 |
| 4 | H6,H13 | H10, H18, H16 | H14, H23, H22 |
| 5 | H2,H8 | H6,H13 | H3, H4, H17 |
| 6 | H7, H25 | H7, H25 | H9, H18 |
| 7 | H9, H18 | H8 | H24 |
| 8 | H 16 | H24 | |
| 9 | H17 | | |
| 10 | H24 | | |

#### *4.5.1.1.3 UPGMA of Hybrid Tea genotypes based on quantitative characters*

Unweighted pair group average method of clustering was done with quantitative characters and different measures of distances. Composition of different clusters obtained by applying different distance measures showed some variation. Distances between clusters were taken as the average distance between objects. Analysis was carried out with the help of statistical packages SPSS and SAS. Pseudo $t^2$ statistics was used to find out the optimum number of clusters under UPGMA method. It was found that the optimum number of cluster is six, ten, twelve and fifteen for Chevbychev distance, eight, ten, thirteen and eighteen for Squared Euclidean distance, seven, nine and eleven for City Block distance. Plot of Pseudo $t^2$ statistics are in Figure 7, Figure 8 and Figure 9. Dendrograms of clustering method is given in Figure 10, Figure 11 and Figure 12. Membership of different genotypes under various distance measures are displayed in the Table 13.

UPGMA divided the genotypes into ten clusters under Squared Euclidean and Chebychev distance and seven clusters under City Block distance. H16 (Mary Jean) and H24 (Alabama) formed individual clusters under all the three distance methods. Cluster of H3 (Christ of Colomb), H4 (Pink Panther), H19 (Majestic), H11 (Silver Star), H15 (Golden Fairy Sport), H17 (Toplesse), H1 (Madame George Delbard) formed the largest cluster in all the methods. H9 (Fryat), H18 (Priority Pride) came under single cluster in City Block distance while they became individual clusters in other methods. H8 (Amara), H18 (Priority Pride), H9 (Fryat) formed single member clusters under Squared Euclidean and Chebychev distances. Grouping of other genotypes showed slight variation among the different distance measures.

#### *4.5.1.1.4 WPGMA of Hybrid Tea genotypes based on quantitative characters*

Clustering was completed by calculating the similarity between two clusters as the mean similarity of previously existing clusters. Cluster analysis was done with the help of

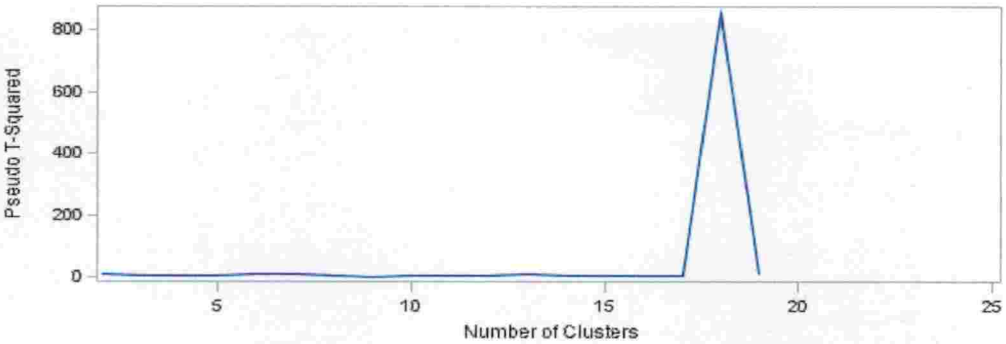


Fig.7: Pseudo $t^2$ statistic plot of Hybrid tea genotypes for Squared Euclidean distance under UPGMA.

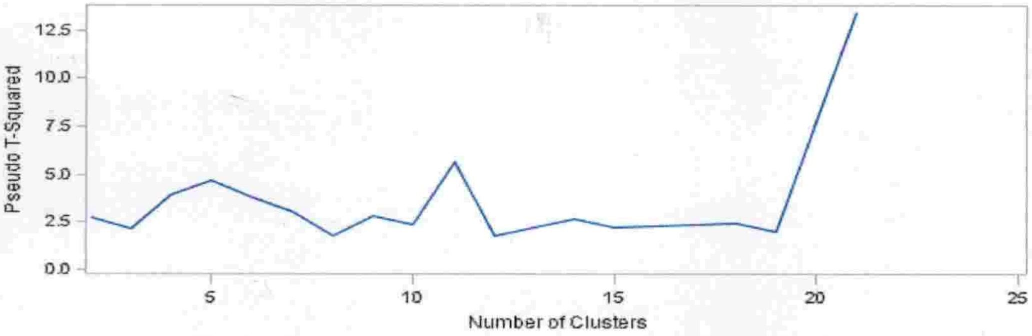


Fig.8: Pseudo $t^2$ statistic plot of Hybrid tea genotypes for Chebychev distance under UPGMA.

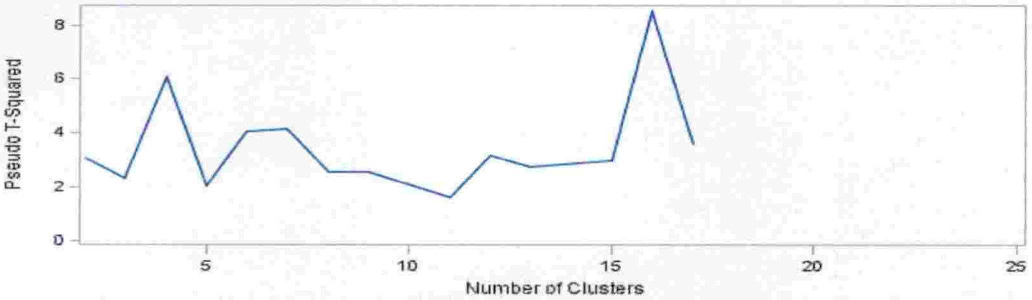


Fig.9: Pseudo $t^2$ statistic plot of Hybrid tea genotypes for City Block distance under UPGMA.

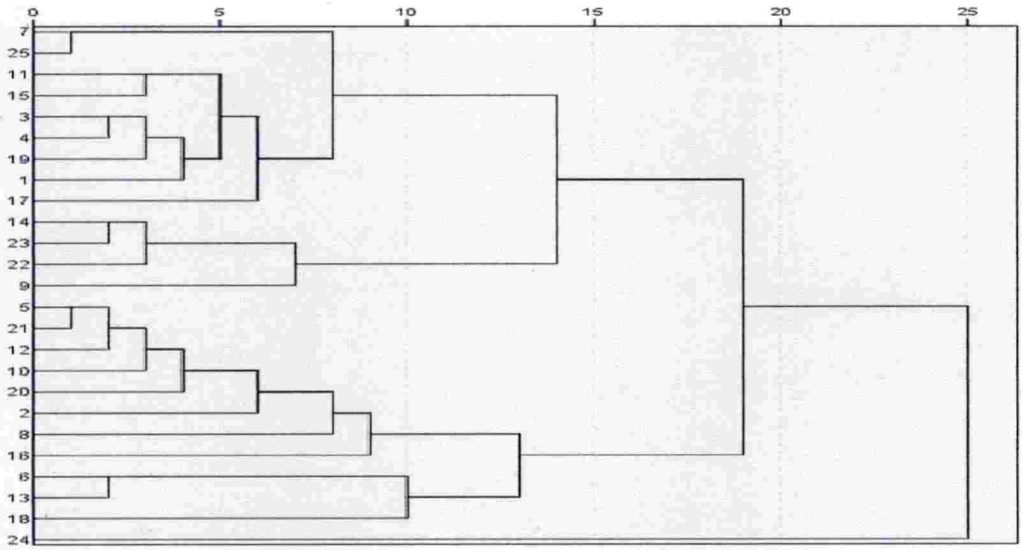


Fig.10: Dendrogram of Hybrid tea genotypes under UPGMA method based on Squared Euclidean distance.

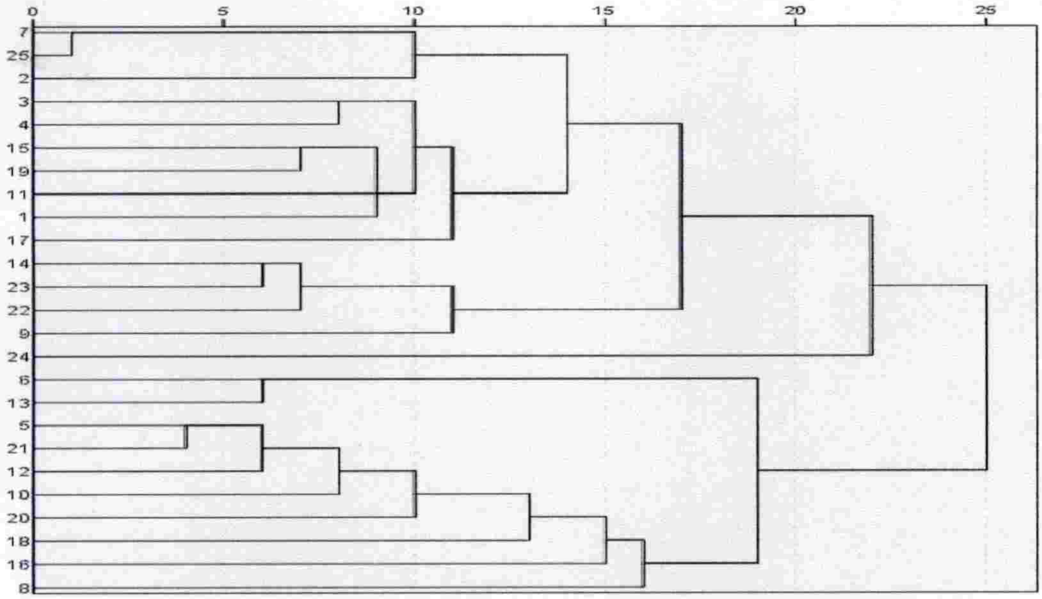


Fig.11: Dendrogram of Hybrid tea genotypes under UPGMA method based on Chebychev distance.

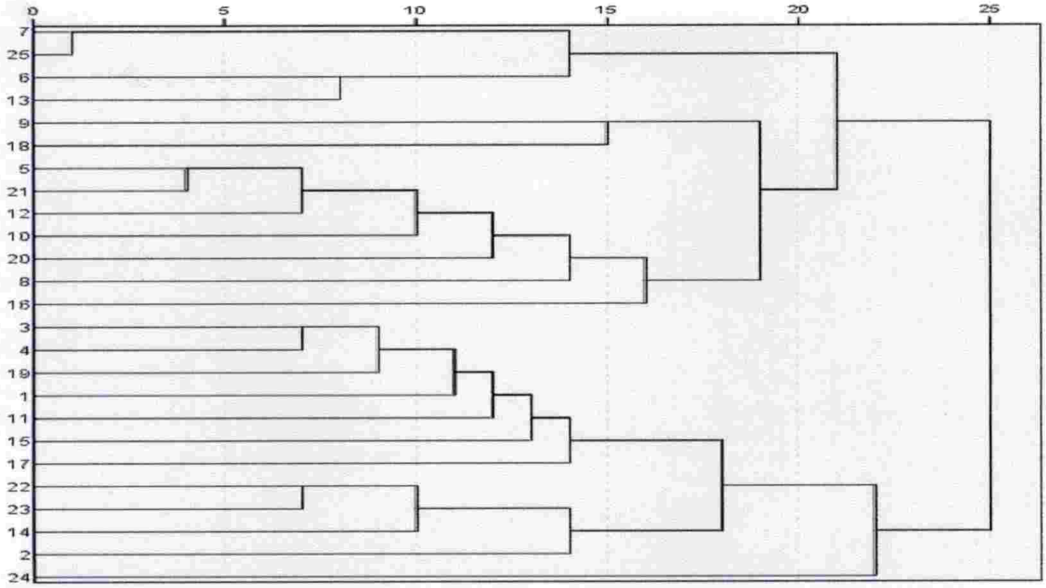


Fig.12: Dendrogram of Hybrid tea genotypes under UPGMA method based on City block distance.

Table 13. Cluster membership of Hybrid Tea types under UPGMA based on quantitative characters.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | H1, H3, H4, H11, H15, H17, H19 | H1, H3, H4, H11, H15, H17, H19 | H1, H3, H4, H11, H15, H17, H19 |
| 2 | H5,H10, H12, H20,H21, H2 | H5,H10, H12, H20,H21 | H5,H10,H12, H20, H21, H8, |
| 3 | H14, H23, H22, H9 | H14, H23,H22, H9 | H7, H25, H6,H13 |
| 4 | H6,H13 | H6,H13 | H14, H23, H22,H2 |
| 5 | H7, H25 | H7, H25, H2 | H9 |
| 6 | H18 | H18 | H18 |
| 7 | H24 | H24 | H16 |
| 8 | H 16 | H16 | H24 |
| 9 | H8 | H8 | |

SPSS package. Clustering result as dendrograms is shown in Figure 13, Figure 14 and Figure 15. Clustering pattern of Hybrid Tea genotypes under WPGMA method are presented in Table 14.

Genotypes were clustered into 10, 8 and 7 numbers of clusters under Squared Euclidean, Chebychev and City Block distance respectively. Cluster of H3 (Christ of Colomb), H4 (Pink Panther), H1 (Madame George Delbard), H19 (Majestic), H11 (Silver Star), H15 (Golden Fairy Sport),H17 (Toplesse) formed the largest cluster in all these distance methods. H24 (Alabama) formed a single member cluster in all the methods. Grouping of other genotypes showed slight variations among different methods. H6 and H13 came under the same cluster in Squared Euclidean and Chebychev distance. Similarly H7 and H25 also formed separate cluster under these association measures. But all these four genotypes came under a single cluster with City Block distance. H14, H22 and H23 came under single cluster in all association measures.

#### *4.5.1.1.5 UPGMC of Hybrid Tea genotypes based on quantitative characters*

Unweighted pair group centroid method was performed only using the Squared Euclidean measure as this method gave valid result only for that distance measure. Objects in each cluster were replaced with the centroid of that particular cluster. Clustering was done with the help of statistical packages SPSS and SAS. Optimum number of clusters obtained from the plot of Pseudo $t^2$ statistics (Figure 16) as six, eight and eighteen. Dendrogram of clustering method is presented in the Figure 17. Clustering pattern of genotypes under UPGMC method is given in the Table 15.

Six clusters were formed under UPGMC method. Second cluster formed the largest cluster with eleven members followed by fifth cluster with seven members. First, third and fourth clusters contain two members each and H24 (Alabama) formed a single cluster.

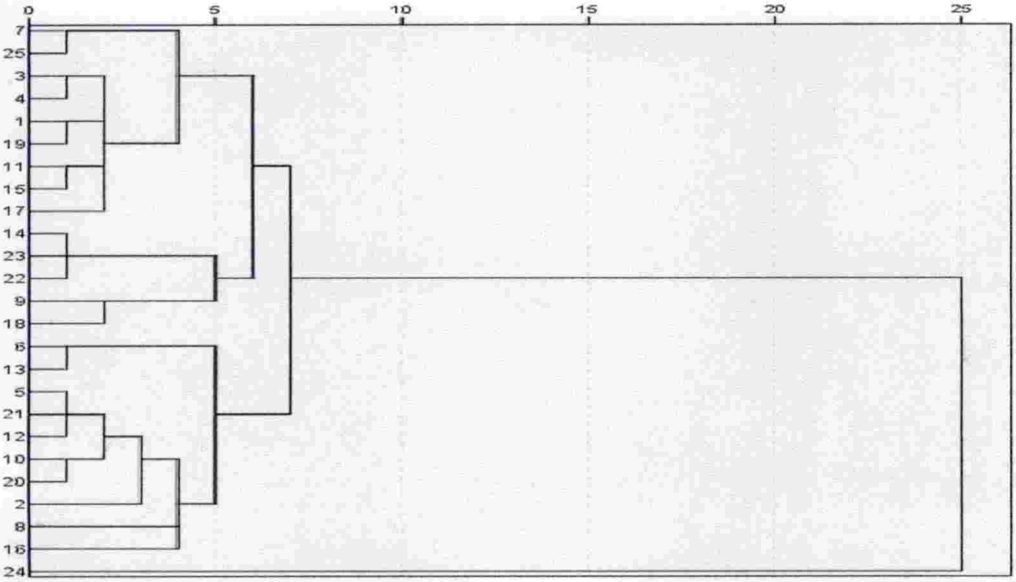


Fig.13: Dendrogram of Hybrid Tea genotypes under WPGMA method based on Squared Euclidean distance.

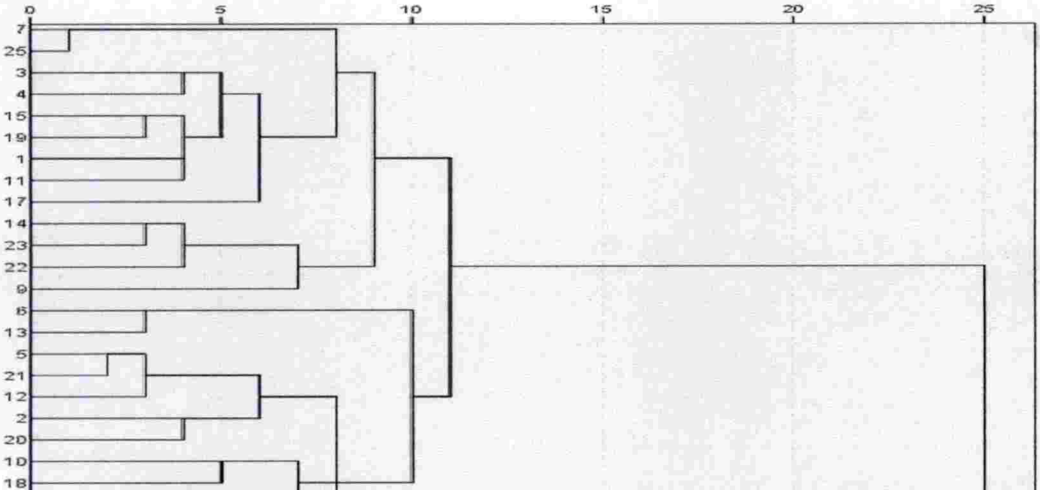


Fig.14: Dendrogram of Hybrid Tea genotypes under WPGMA method based on Chebychev distance.

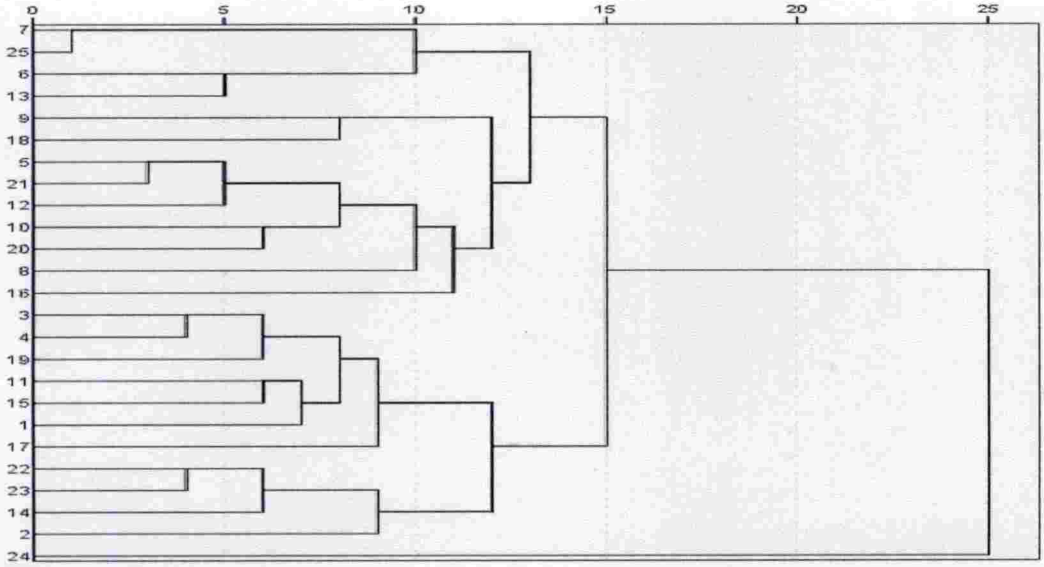


Fig.15: Dendrogram of Hybrid Tea genotypes under WPGMA method based on City block distance.

Table 14. Cluster membership of Hybrid Tea types under WPGMA based on quantitative data.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | H1, H3, H4, H11, H15, H17, H19 | H1, H3, H4, H11, H15, H17, H19 | H1, H3, H4, H11, H15, H17, H19 |
| 2 | H5,H10, H12, H20,H21 | H5, H12, H20,H21, H2 | H5,H10,H12, H20, H21, H8, |
| 3 | H14, H23, H22 | H14, H23,H22, H9 | H7, H25, H6,H13 |
| 4 | H6,H13 | H18, H10,H16 | H14, H23, H22,H2 |
| 5 | H7, H25 | H7, H25 | H9,H18 |
| 6 | H9, H18 | H6,H13 | H16 |
| 7 | H8 | H8 | H24 |
| 8 | H 16 | H24 | |
| 9 | H24 | | |
| 10 | H2 | | |

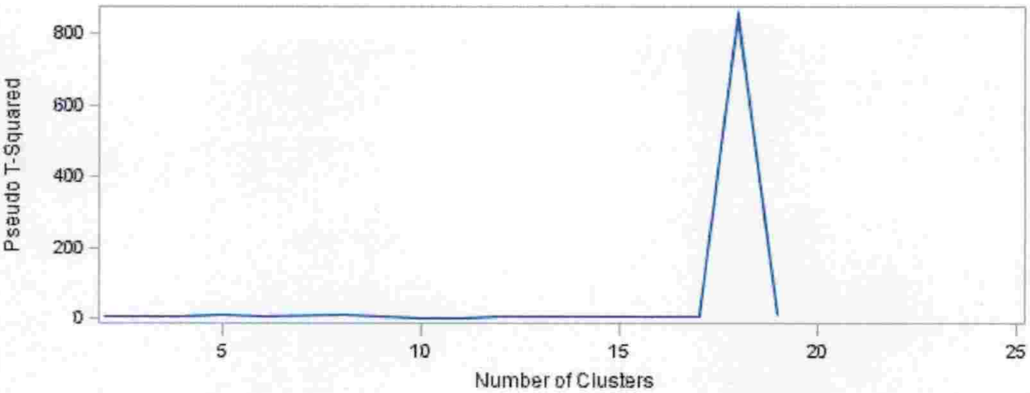


Fig.16: Pseudo $t^2$ statistic plot of Hybrid Tea genotypes for Squared Euclidean distance under UPGMC.

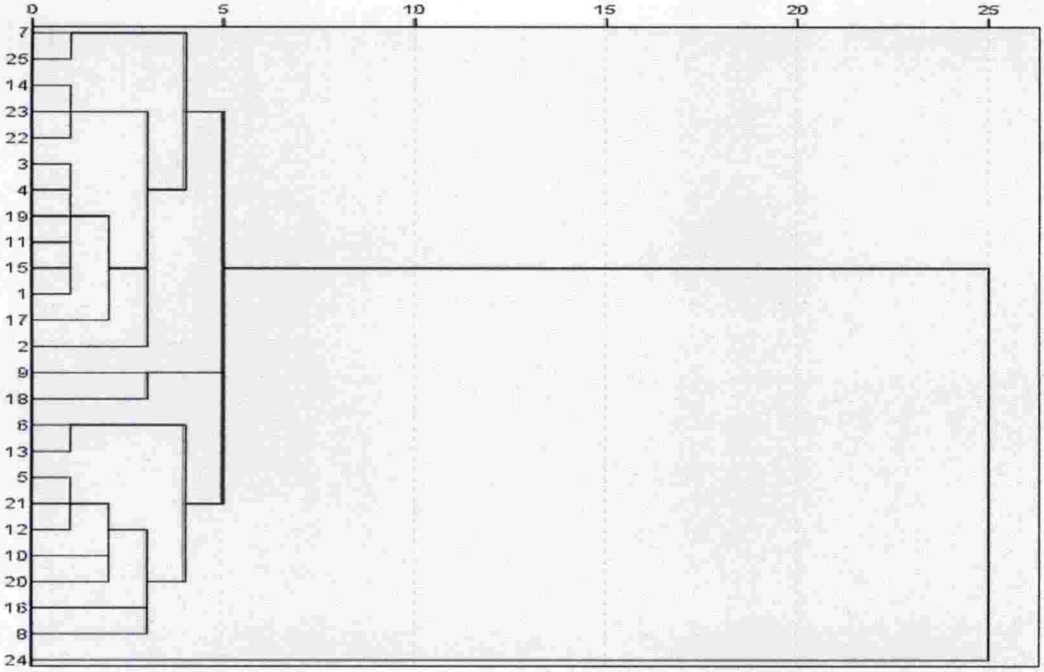


Fig.17: Dendrogram of Hybrid Tea genotypes under UPGMC method based on Squared Euclidean distance.

Table 15. Cluster membership of Hybrid Tea types under UPGMC method.

| Cluster No | UPGMC |
|---|---|
| 1 | H7, H25 |
| 2 | H14, H22, H23, H1, H3, H4, H11, H15, H17, H19, H2 |
| 3 | H9, H18 |
| 4 | H6, H13 |
| 5 | H5,H10, H12, H20,H21, H16, H8 |
| 6 | H24 |

### *4.5.1.1.6 Ward's method of Hybrid Tea genotypes based on quantitative characters*

Ward's method of clustering based on minimizing sum of square was performed using Squared Euclidean distance. Clustering was performed with the help of statistical packages SPSS and SAS. Optimum number of clusters was found to be seven, fourteen and eighteen from plot of Pseudo $t^2$ statistics (Figure 18). Dendrogram of clustering are given in Figure 19. Table of cluster membership showed that second and fifth clusters were the largest with seven members which followed by the sixth cluster with three members and first, third and fourth clusters with two members each. Alabama formed a single cluster (Table 16).

### *4.5.1.1.7 $D^2$ analysis of Hybrid Tea genotypes*

The genetic distance among varieties were estimated based on nine quantitative characters and the values are presented in the Appendix XIX. The genotypes were arranged in the ascending order of $D^2$ values and clustering was done by modified Tocher method. The 25 genotypes corresponding to Hybrid Tea were clustered into eight clusters and cluster membership corresponding to each cluster is shown in Table 17. The first cluster was the largest with six members. There were five members in cluster two, four members in cluster four and five, two members in clusters three and seven and single member in clusters six and eight. The intra and inter cluster distances are given in Table 18. The genetic divergence was maximum between cluster I and VIII (6005.51) followed by cluster III and VIII (5935.55) and cluster VI and VIII (5083.19). Cluster VI and VIII are having zero intra cluster as they contain single object.

Table 19 shows the relative contribution of different characters towards divergence. Number of days to first flower (40.67%) had highest contribution towards divergence followed by number of leaves at first flower (19.00%), number of petals/ flower (16.33%), flower size (12.00%), size of petals (11.67%), prickle density/5cm (0.33%) and the remaining characters did not have any contribution to the divergence.

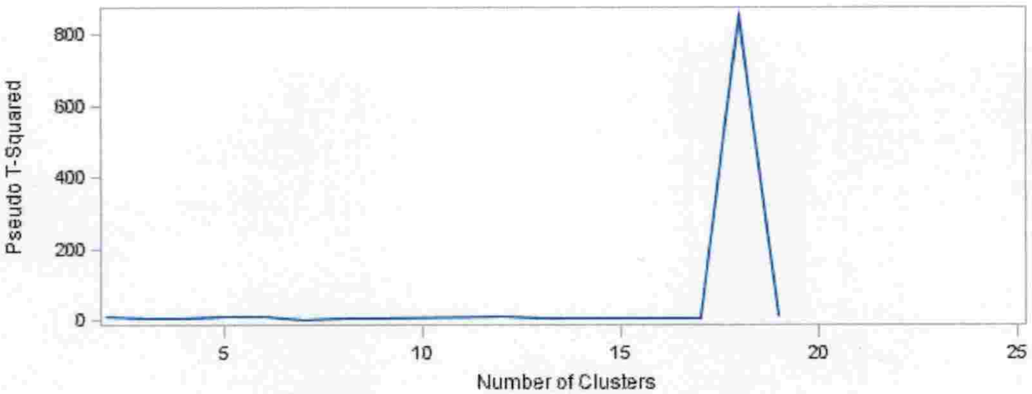


Fig.18: Pseudo t$^2$ statistic plot of Hybrid Tea genotypes for Squared Euclidean distance under Ward's method.

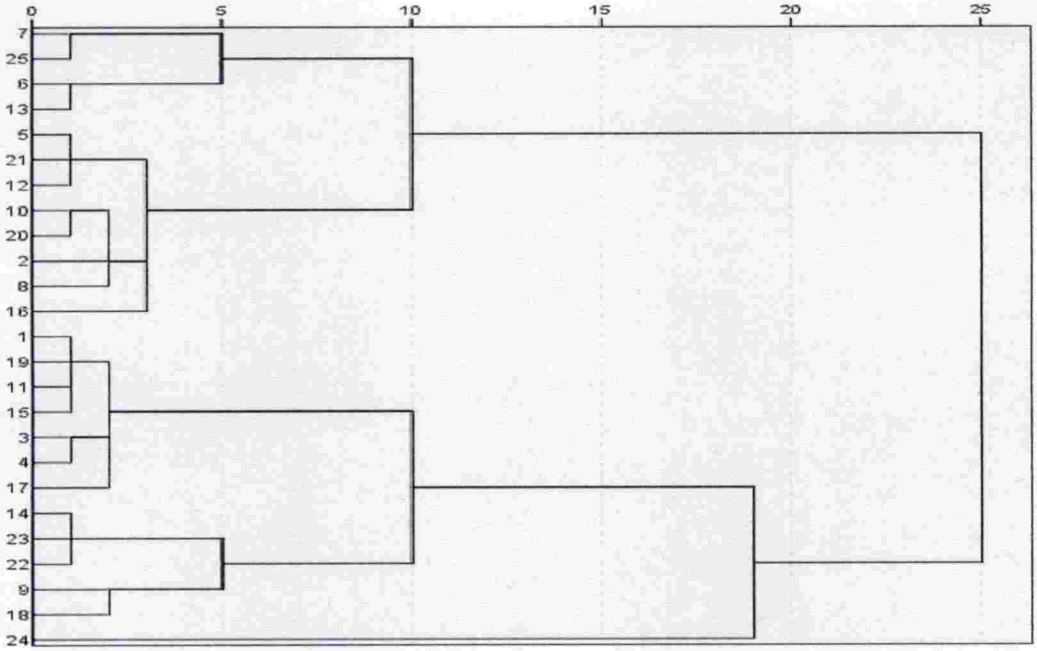


Fig.19: Dendrogram of Hybrid Tea genotypes under Ward's method based on Squared Euclidean distance.

Table 16. Cluster membership of Hybrid Tea types under Ward's method.

| Cluster no | Ward's Method |
|---|---|
| 1 | H7, H25 |
| 2 | H1, H3, H4, H11, H15, H17, H19 |
| 3 | H9, H18 |
| 4 | H6, H13 |
| 5 | H5,H10, H12, H20,H21, H16, H8, H2 |
| 6 | H14, H22, H23 |
| 7 | H24 |

Table 17. Cluster membership of Hybrid Tea Genotypes based on $D^2$ analysis

| Cluster | Name of the Genotypes | No. of Genotypes |
|---|---|---|
| I | H7, H25, H5, H21, H12, H4 | 6 |
| II | H22, H23, H11, H15, H2 | 5 |
| III | H10, H18, H20, H16 | 4 |
| IV | H1, H 19, H3, H17 | 4 |
| V | H6, H13 | 2 |
| VI | H9, H14 | 2 |
| VII | H8 | 1 |
| VIII | H24 | 1 |

Table 18. Inter and intra cluster distances of Hybrid Tea Genotypes based on D$^2$ values

| | I | II | III | IV | V | VI | VII | VII |
|---|---|---|---|---|---|---|---|---|
| I | 232.43 | 1012.56 | 433.45 | 495.72 | 565.47 | 850.52 | 1209.75 | 6005.51 |
| II | | 281.94 | 1213.13 | 546.45 | 667.18 | 778.29 | 660.02 | 2954.43 |
| III | | | 58.36 | 1007.40 | 310.85 | 943.58 | 1342.19 | 5935.55 |
| IV | | | | 283.68 | 642.57 | 1025.49 | 745.14 | 4303.37 |
| V | | | | | 303.91 | 628.32 | 669.13 | 3969.84 |
| VI | | | | | | 0 | 1358.00 | 5083.19 |
| VII | | | | | | | 402.41 | 2431.26 |
| VII | | | | | | | | 0 |

Table 19. Percentage contribution of characters divergence - Hybrid Tea genotypes.

| Character | Percentage contribution to divergence |
|---|---|
| No. of leaves at first flower | 19.00 |
| No. of days to first flower | 40.67 |
| Prickle density/5cm | 0.33 |
| Flower size (cm) | 12.00 |
| Flower weight (g) | 0.00 |
| Pedicel length (cm) | 0.00 |
| No. of petals/ flower | 16.33 |
| Size of petals (cm) | 11.67 |
| No. of flower per plant/bunch | 0.00 |

#### *4.5.1.2 Clustering of Hybrid Tea genotypes based on qualitative characters*

Cluster analysis of Hybrid Tea genotypes were carried out using seven characters. Among them three were qualitative traits and four were quantitative characters converted to qualitative characters. The qualitative characters included fragrance, flower colour, vase life/ longevity and the characters prickle density (per five cm), number of petals flower$^{-1}$, size of petals (cm) and number of flower plant $^{-1}$/bunch $^{-1}$ were converted to qualitative characters. Different clustering methods were adopted with distance methods like Jaccard and Dice coefficient for binary data and Simple matching coefficient and Hamann's coefficient for multistage data. Clustering pattern obtained from Simple matching coefficient did not make any valuable results for Hybrid Tea type. So it was excluded from further results.

##### *4.5.1.2.1 Single linkage clustering of Hybrid Tea genotypes based on qualitative data*

Single linkage clustering of qualitative data was performed with the help of statistical packages SPSS, NTSYS and SAS. Clustering was done using Jaccard coefficient, Dice coefficient and Hamann's coefficient. Dendrogram of clustering are given in Figure 20, Figure 21 and Figure 22. Clustering pattern of genotypes is presented in Table 20. It was found that single linkage clustering using Jaccard and Dice distance have same clustering pattern with one large cluster with nineteen members and all other genotypes formed individual cluster. Under Hamann's coefficient largest cluster included twenty one members and all other genotypes formed single member cluster. There were seven clusters under Jaccard and Dice distance and five clusters under Hamann's coefficient.

##### *4.5.1.2.2 Complete linkage clustering of Hybrid Tea genotypes based on qualitative data*

Complete linkage clustering was performed using Jaccard, Dice And Hamann's coefficient with the help of statistical packages SPSS, NTSYS and SAS. Dendrograms are given in Figure 23, Figure 24 and Figure 25. Clustering pattern of genotypes are presented in Table 21. Here clustering under Jaccard and Dice showed similar clustering pattern while clustering under Hamann's coefficient showed some variation. H3 (Christ of Colomb)

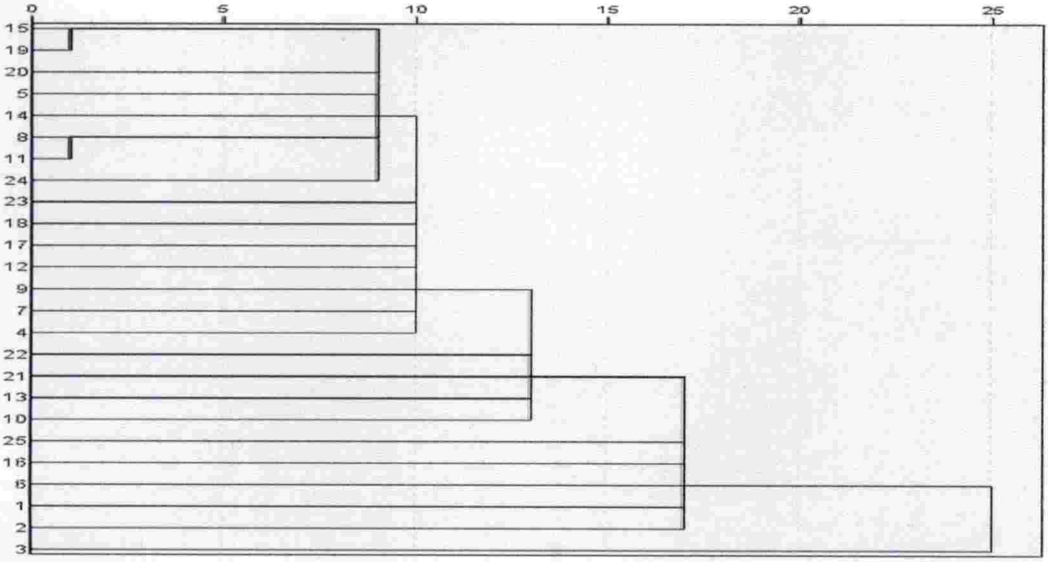


Fig. 20: Dendrogram of single linkage clustering of Hybrid Tea genotypes based on Jaccard distance.

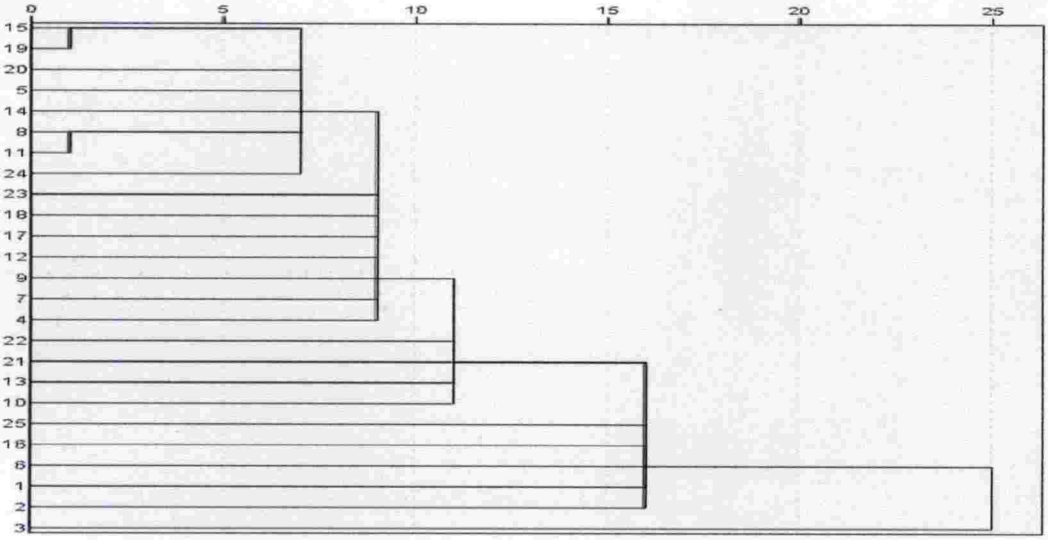


Fig.21: Dendrogram of single linkage clustering of Hybrid Tea genotypes based on Dice distance

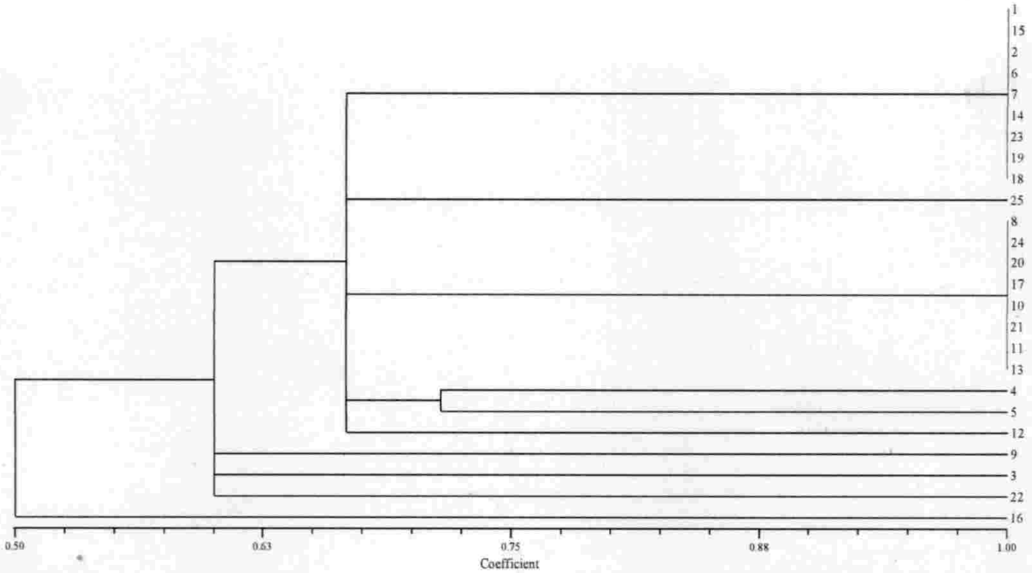


Fig. 22: Dendrogram of single linkage clustering of Hybrid Tea genotypes based on Hamann's coefficient.

Table 20. Cluster membership of Hybrid Tea type under single linkage for qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient |
|---|---|---|---|
| 1 | H15, H19, H20, H5, H14, H8, H24, H23,H18, H17, H12, H11,H9, H7, H4, H22, H21, H13, H10 | H15, H19, H20, H5, H14, H8, H24, H23,H18, H17, H12, H11,H9, H7, H4, H22, H21, H13, H10 | H1, H2,H15,H6, H7, H14, H19, H20, H23, H25, H24, H8, H18, H17, H21, H10, H11, H13, H4, H5, H12 |
| 2 | H25 | H25 | H9 |
| 3 | H16 | H16 | H16 |
| 4 | H6 | H6 | H22 |
| 5 | H1 | H1 | H3 |
| 6 | H2 | H2 | |
| 7 | H3 | H3 | |

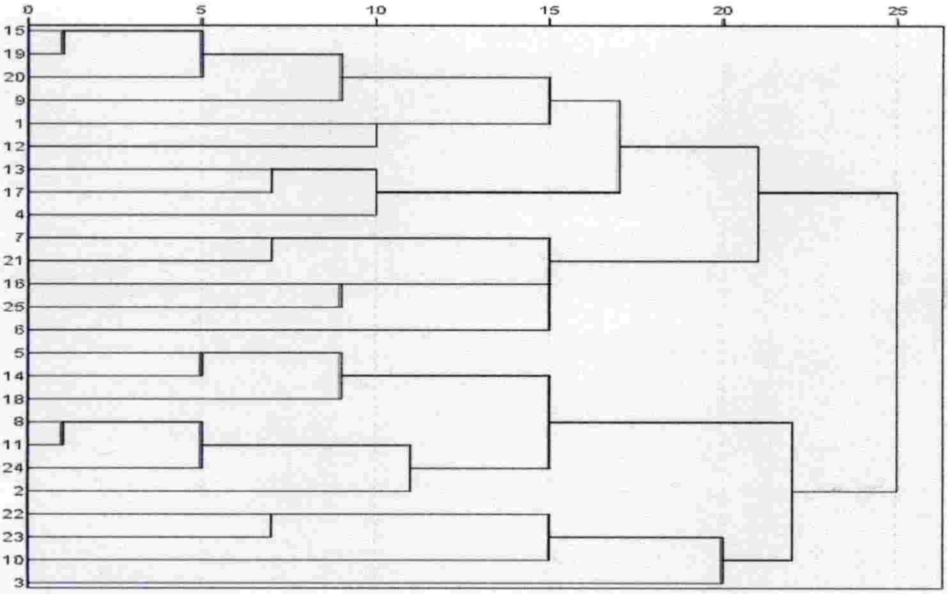


Fig. 23: Dendrogram of complete linkage clustering of Hybrid Tea genotypes based n Jaccard coefficient.

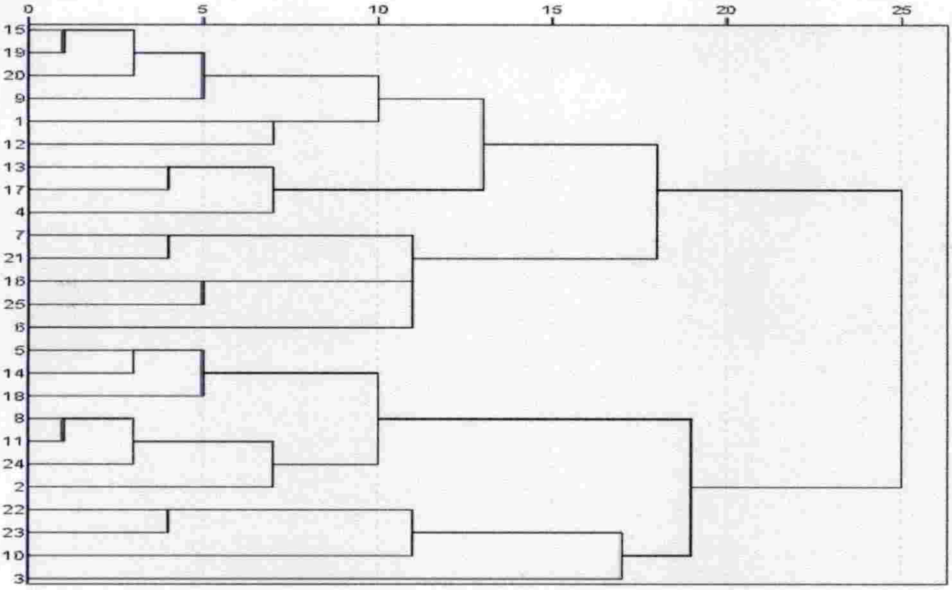


Fig. 24: Dendrogram of complete linkage clustering of Hybrid Tea genotypes based on Dice coefficient.

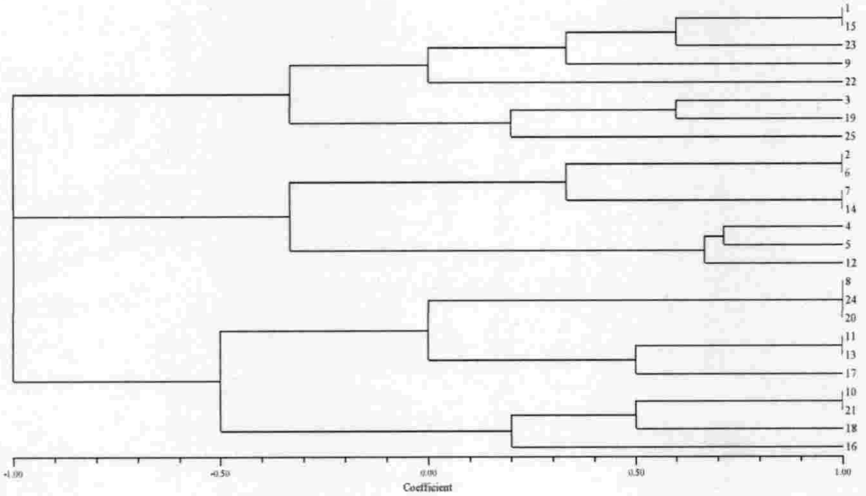


Fig. 25: Dendrogram of complete linkage clustering of Hybrid Tea genotypes based on Hamann's coefficient.

Table 21. Cluster membership of Hybrid Tea type under complete linkage for qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient |
|---|---|---|---|
| 1 | H5, H14, H18, H8, H11, H24, H2 | H5, H14, H18, H8, H11, H24, H2 | H1, H15, H9, H23 |
| 2 | H15, H19, H20, H9, H1, H12 | H15, H19, H20, H9, H1, H12 | H10, H21, H18, H16 |
| 3 | H7, H21, H16, H25, H6 | H7, H21, H16, H25, H6 | H2, H6, H7, H14 |
| 4 | H13, H17, H4 | H13, H17, H4 | H4, H5, H12 |
| 5 | H22, H23, H10 | H22, H23, H10 | H8, H24, H20 |
| 6 | H3 | H3 | H11, H13, H17 |
| 7 | | | H25, H19, H3 |
| | | | H22 |

formed a single cluster under Jaccard and Dice distance while H22 (Lois Wilson) formed a single member cluster under Hamann' s coefficient. Genotypes were divided into six clusters by Jaccard and Dice and into eight clusters by Hamann,s coefficient.

#### *4.5.1.2.3 UPGMA clustering of Hybrid Tea genotypes based on qualitative data*

UPGMA method of clustering was performed with Jaccard, Dice and Hamann's coefficient with the help of statistical packages SPSS, SAS and NTSYS. Optimum number of clusters was obtained from the plot of Pseudo $t^2$ statistics (Figure 26, Figure 27 and Figure 28). It was observed that optimum number of cluster is eight and fourteen for Dice distance, seven and eleven for Jaccard distance and seven for Hamann's coefficient. Dendrograms of clustering are given in Figure 29, Figure 30 and Figure 31. Clustering pattern of genotypes is presented in Table 22. Clustering under Jacaard and Dice distance showed almost same clustering pattern. H7 (Alaine Souchen), H21 (Cel b Lau), H16 (Mary Jean) and H25 (Josepha) came under the same cluster in Jaccard distance while they formed two clusters each containing two members in Dice distance. Clustering under Hamann's coefficient showed variation from the other methods.

#### *4.5.1.2.4 WPGMA clustering of Hybrid Tea genotypes based on qualitative data*

WPGMA clustering was done under different distance methods. Clustering results as dendrogram are shown in Figure 32, Figure 33 and Figure 34. Cluster membership of genotypes is given in Table 23. Here also Jaccard and Dice distance have almost the same clustering pattern. Six clusters were formed under Jaccard distance, seven under Dice distance and nine under Hamann's coefficient. H1 (Madame George Delbard) formed an individual cluster under Dice distance while it formed cluster with other genotypes under Jaccard distance. Clustering pattern of Hamann's coefficient is different from the other two.

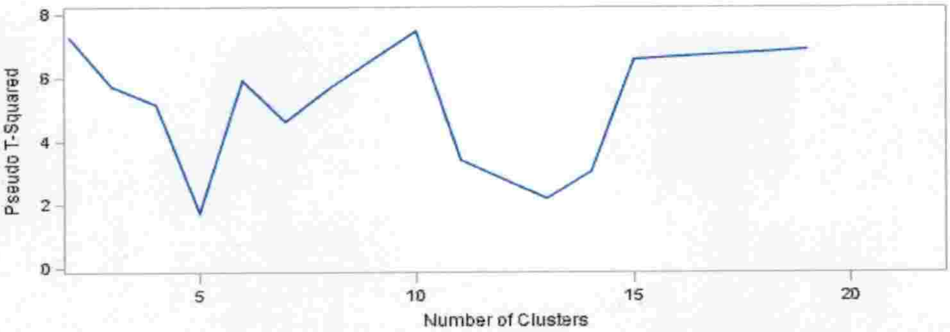


Fig.26: Pseudo t$^2$ statistic plot of Hybrid Tea genotypes for Jaccard coefficient under UPGMA.

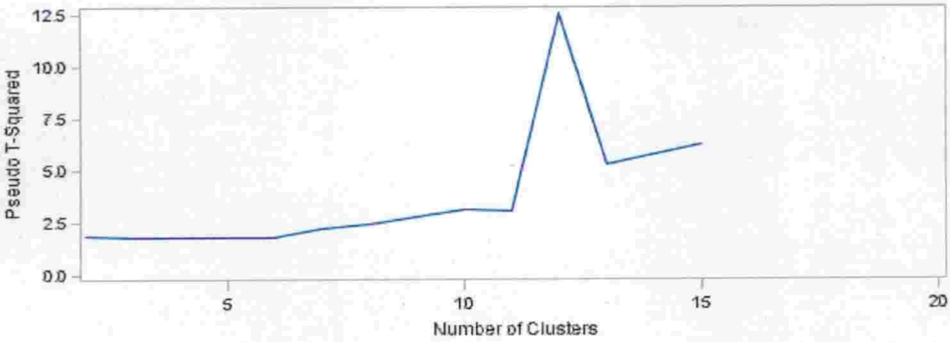


Fig.27: Pseudo t$^2$ statistic plot of Hybrid Tea genotypes for Dice coefficient under UPGMA.

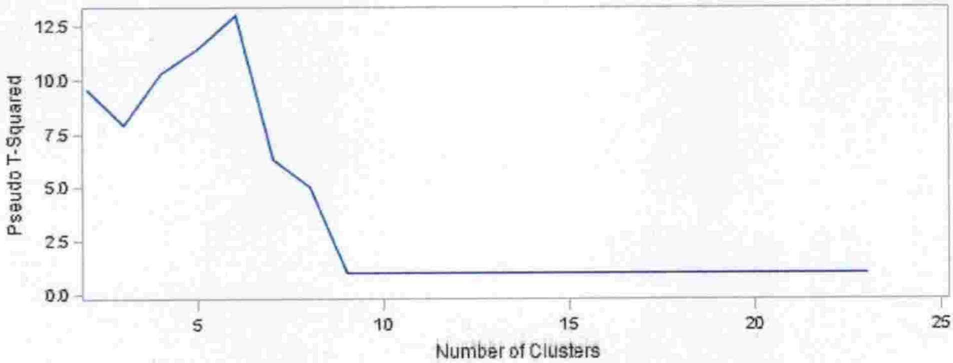


Fig.28: Pseudo t$^2$ statistic plot of Hybrid tea genotypes for Hamann's coefficient under UPGMA.

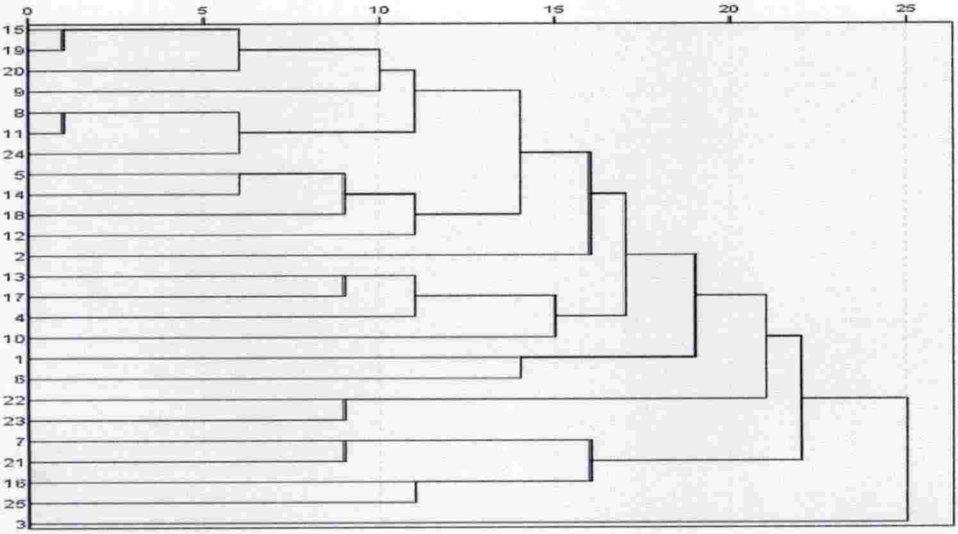


Fig. 29: Dendrogram of UPGMA of Hybrid Tea genotypes based on Jaccard coefficient

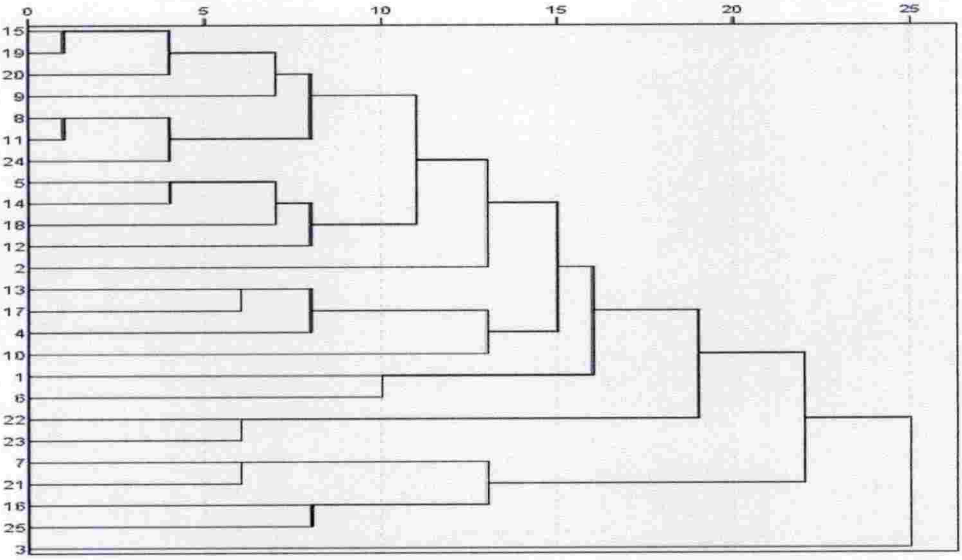


Fig. 30: Dendrogram of UPGMA of Hybrid Tea genotypes based on Dice coefficient

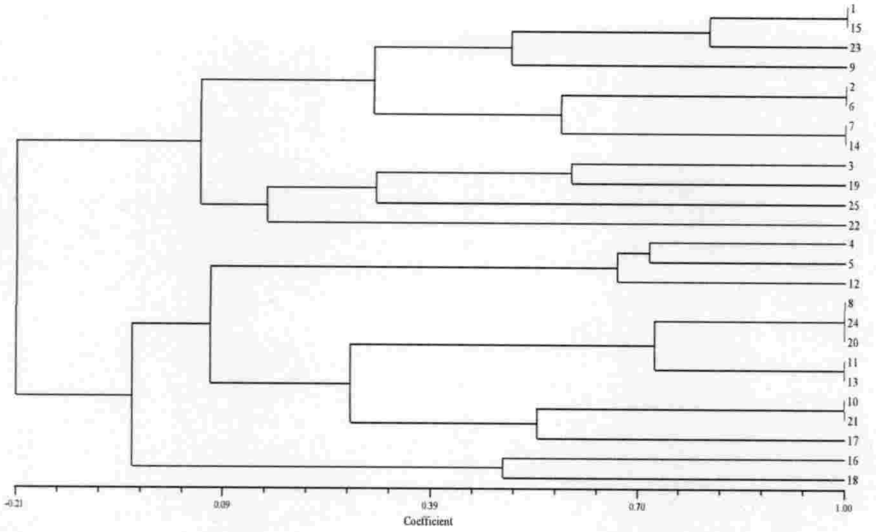


Fig. 31: Dendrogram of UPGMA of Hybrid Tea genotypes based on Hamann's coefficient

Table 22. Cluster membership of Hybrid Tea type under UPGMA for qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient |
|---|---|---|---|
| 1 | H15, H19, H20, H9, H8, H11, H24, H5,H14, H18, H12 | H15, H19, H20, H9, H8, H11, H24, H5,H14, H18, H12 | H1, H15, H23, H9, H2, H6, H7, H14 |
| 2 | H13, H17, H 4, H10 | H13, H17, H 4, H10 | H8, H24, H20, H11, H13 |
| 3 | H7, H21, H16, H25 | H7, H21 | H3, H19, H25 |
| 4 | H1, H6 | H1, H6 | H4, H5, H12 |
| 5 | H22, H23 | H22, H23 | H10, H21, H17 |
| 6 | H2 | H16, H25 | H16, H18 |
| 7 | H3 | H3 | H22 |
| 8 | | H2 | |

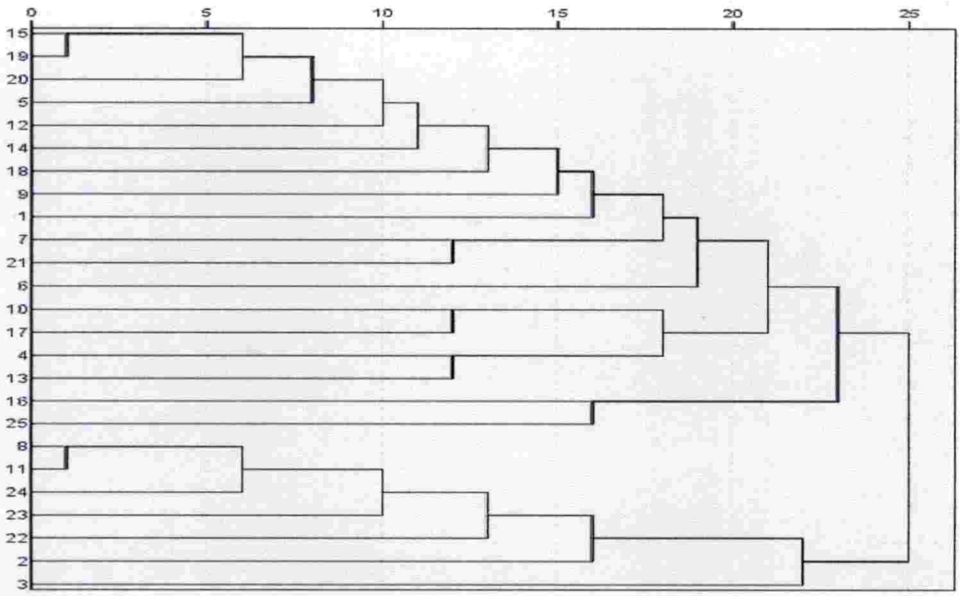


Fig. 32: Dendrogram of WPGMA of Hybrid Tea genotypes based on Jaccard coefficient

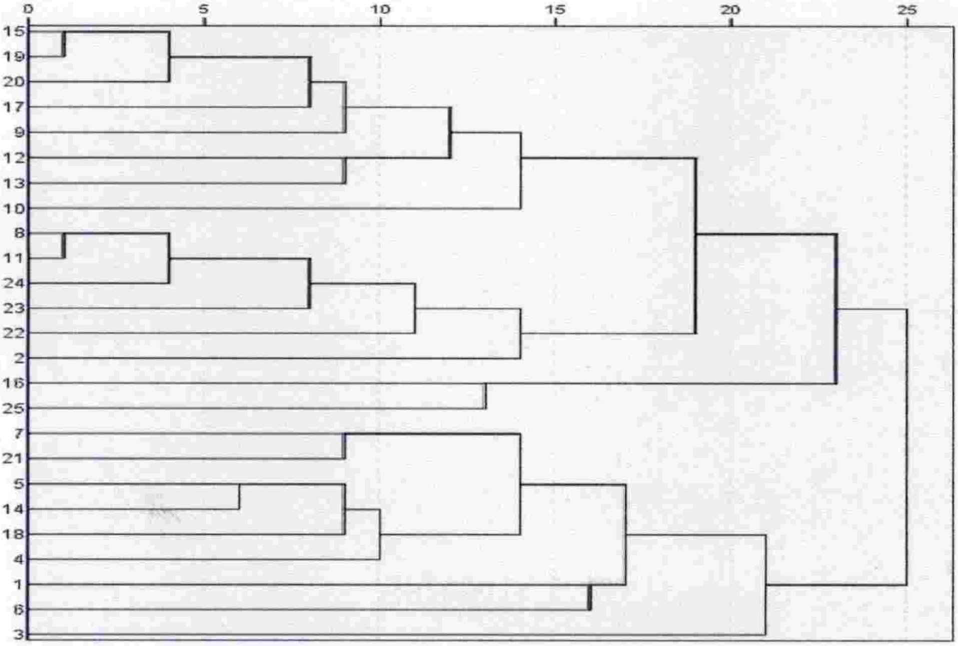


Fig. 33: Dendrogram of WPGMA of Hybrid Tea genotypes based on Dice coefficient

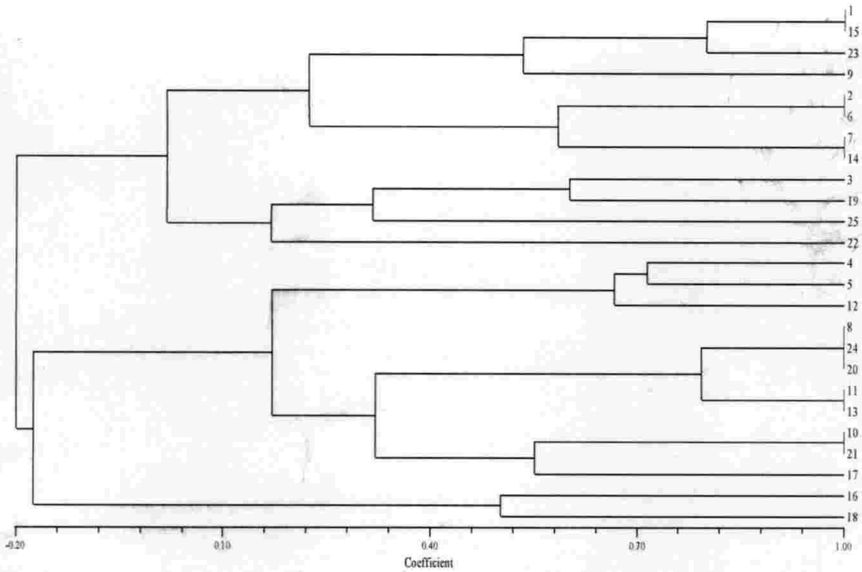


Fig. 34: Dendrogram of WPGMA of Hybrid Tea genotypes based on Hamann's coefficient

Table 23. Cluster membership of Hybrid Tea type under WPGMA for qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient |
|---|---|---|---|
| 1 | H15, H20, H5, H12, H14, H19, H18, H9, H1, H7, H21 | H15, H19, H20, H17, H9, H12, H13, H10 | H8, H24, H20, H11, H13 |
| 2 | H8, H11, H24, H23, H22, H2 | H8, H11, H24, H23, H22, H2 | H1, H15, H23, H9 |
| 3 | H10, H17 H4, H13 | H7, H21, H5, H14, H18, H4 | H2, H6, H7, H14 |
| 4 | H16, H25 | H16, H25 | H4, H5, H12 |
| 5 | H3 | H3 | H10, H21, H17 |
| 6 | H6 | H6 | H16, H18 |
| 7 | | H1 | H3, H19 |
| | | | H22 |
| | | | H25 |

### *4.5.1.3 Clustering of Hybrid Tea genotypes based on quantitative and qualitative characters*

Cluster analysis was performed for combined data which contains both quantitative and qualitative characters. Among the twelve characters number of leaves at first flower, number of days to first flower, prickle density (per five cm), flower size (cm), flower weight (g), pedicel length (cm), number of petals flower$^{-1}$, size of petals (cm) and number of flower plant$^{-1}$/ bunch $^{-1}$ were taken as quantitative characters and Fragrance, Flower colour and Vase life/ longevity were taken as qualitative characters. Clustering was performed under different clustering algorithm with Gower's measure. Dendrograms of clustering are shown in Figure 35, Figure 36, Figure 37 and Figure 38. Cluster memberships of genotypes are given in the Table 24. Optimum number of clusters under UPGMA method was found to be eight, eleven and fourteen from the plot of Pseudo $t^2$ statistics (Figure 39). Clustering was done using STATA and SAS.

Seven clusters were formed under single linkage method. One large cluster contained nineteen members and all the other clusters contained one or two members only. Eight clusters were formed under complete linkage, UPGMA and WPGMA methods. Clustering under UPGMA and WPGMA are almost the same but it was different from complete linkage method.

### *4.5.1.4. Non Hierarchical clustering of Hybrid Tea genotypes*

Non hierarchical clustering of twenty five Hybrid Tea genotypes were done by k-means clustering technique. It is a technique in which the number of clusters is predetermined. The optimum number of clusters was obtained from Pseudo F statistics. From the table of Pseudo F statistics (Table 25) optimum number of clusters was obtained as nine with Pseudo F statistics value 14.49. Cluster memberships of genotypes in nine clusters are given in the Table 26.Nine clusters were formed under k- means clustering. First cluster with six members formed the largest cluster followed by the second cluster with five members,

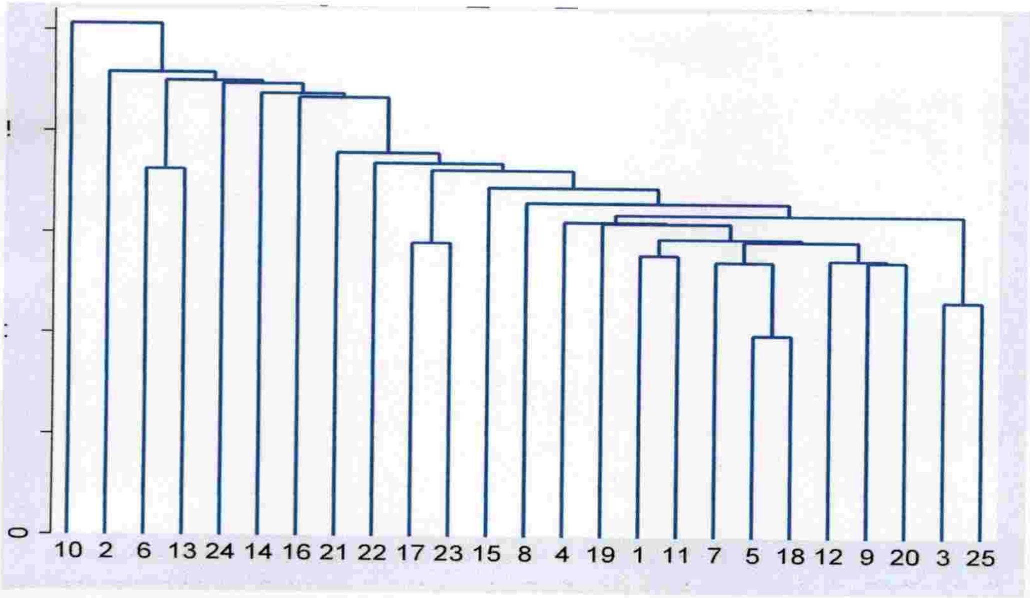


Fig. 35: Dendrogram of single linkage clustering of Hybrid Tea genotypes based on Gower's Measure

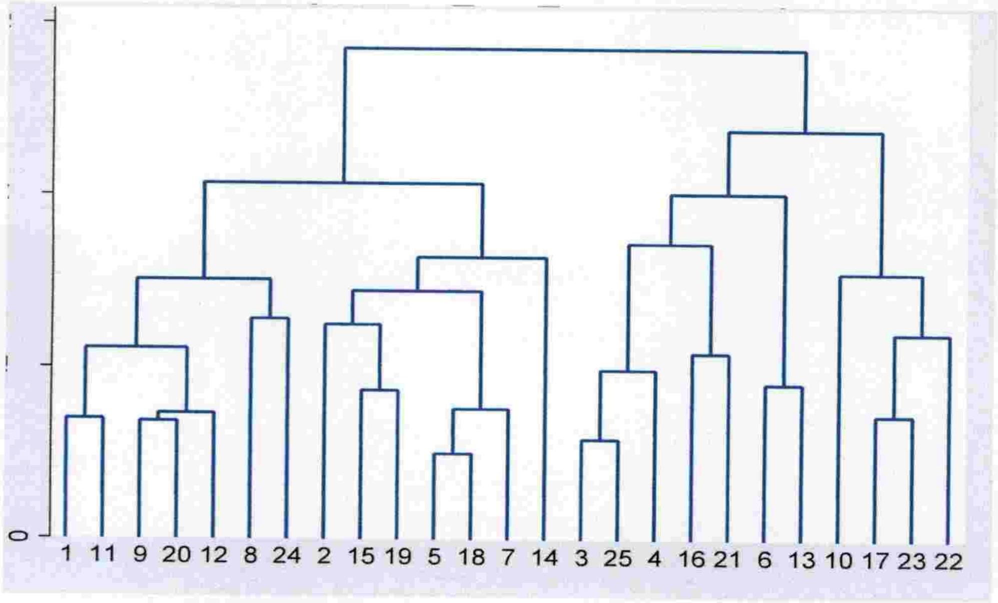


Fig. 36: Dendrogram of complete linkage clustering of Hybrid Tea genotypes based on Gower's measure.

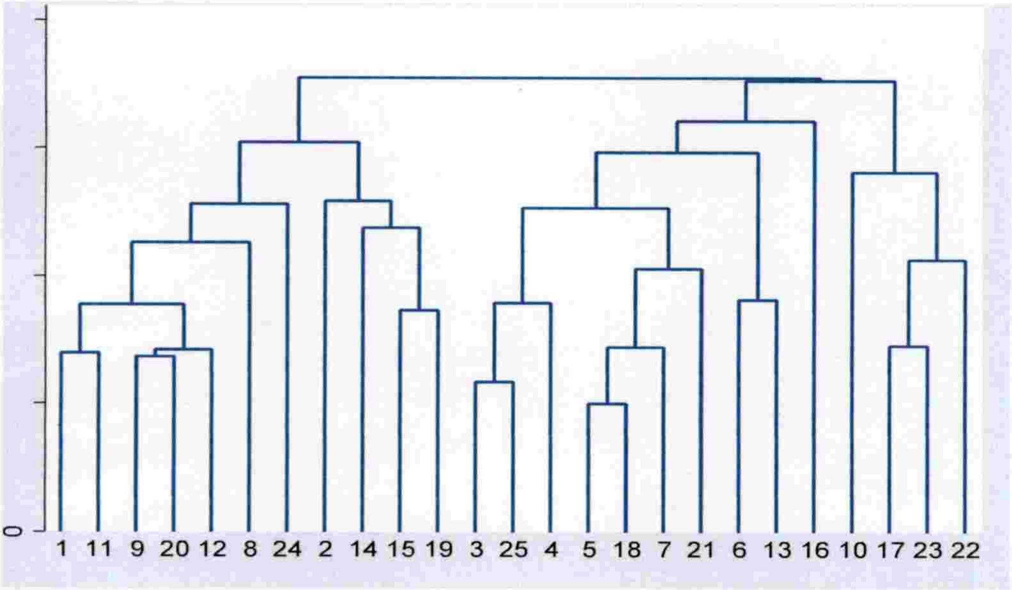


Fig. 37: Dendrogram of UPGMA of Hybrid Tea genotypes based on Gower's measure.

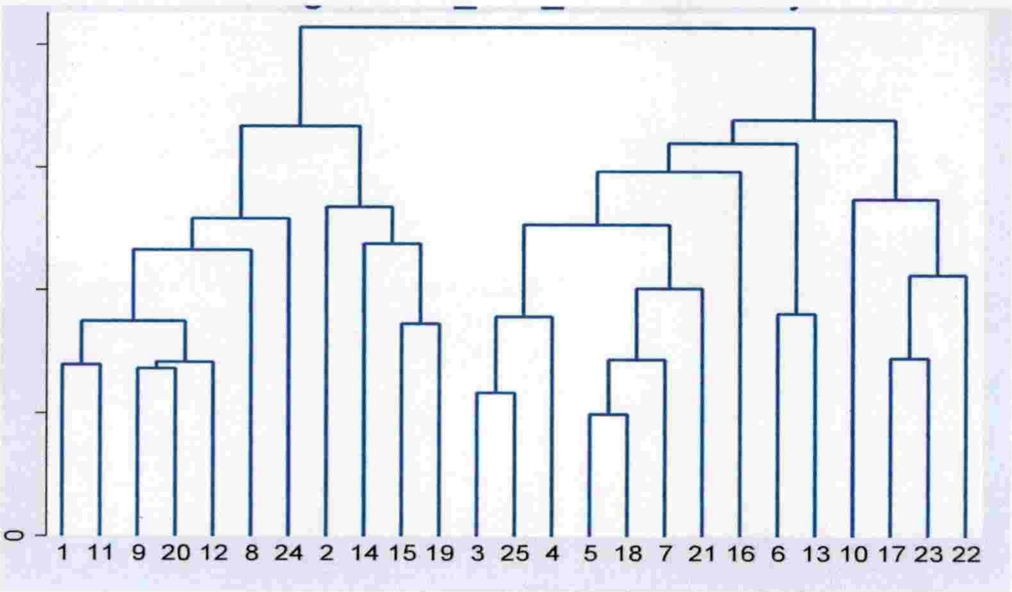


Fig. 38: Dendrogram of WPGMA of Hybrid Tea genotypes based on Gower's measure.

117

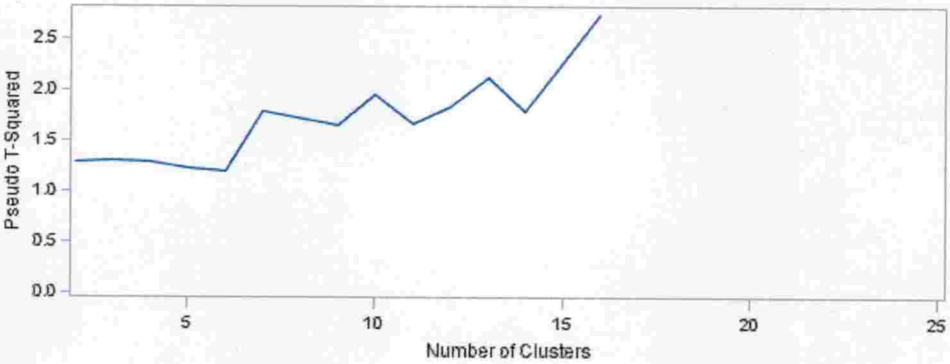


Fig.39: Pseudo $t^2$ statistic plot of Hybrid tea genotypes for Gower's measure under UPGMA.

Table 24. Cluster membership of Hybrid Tea type under Gower's measure.

| Cluster number | Single linkage | Complete linkage | UPGMA | WPGMA |
|---|---|---|---|---|
| 1 | H25, H3, H20, H9, H12, H18, H5, H7, H11, H1, H19, H4, H8, H15, H23, H17, H22, H21 | H2, H15, H19, H5, H18, H7 | H1, H11, H9, H20, H12, H8, 24 | H1, H11, H9, H20, H12, H8, 24 |
| 2 | H6, H13 | H1, H11, H9, H12, H20 | H3, H4, H5, H7, H18, H21, H25 | H3, H4, H5, H7, H18, H21, H25 |
| 3 | H16 | H10, H17, H22, H23 | H17, H22,H23 | H17, H22,H23 |
| 4 | H10 | H3, H4, H25 | H14, H15, H19 | H14, H15, H19 |
| 5 | H24 | H16, H21 | H6, H13 | H6, H13 |
| 6 | H2 | H6, H13 | H16 | H16 |
| 7 | H14 | H8, H24 | H10 | H10 |
| 8 |  | H14 | H2 | H2 |

Table 25.Table of Pseudo F statistics for Hybrid Tea genotypes under k- means clustering.

| No of clusters | Pseudo F Statistic |
|---|---|
| 5 | 12.12 |
| 6 | 10.66 |
| 7 | 13.46 |
| 8 | 13.37 |
| 9 | 14.49 |
| 10 | 13.15 |

Table 26. Cluster membership of Hybrid Tea genotypes under k-means Clustering.

| Cluster number | Members |
|---|---|
| 1 | H1,H3, H11, H15, H17, H19 |
| 2 | H5, H10, H12, H20, H21 |
| 3 | H14, H22, H23 |
| 4 | H4, H7, H25 |
| 5 | H9, H18 |
| 6 | H6, H13 |
| 7 | H2,H8 |
| 8 | H16 |
| 9 | H24 |

third and fourth cluster with three members and fifth, sixth and seventh with two members. H16 (Mary Jean) and H24 (Alabama) formed single member clusters.

### 4.5.2 Clustering of Floribunda genotypes

#### *4.5.2.1 Clustering of Floribunda type based on Quantitative data*

Data corresponding to the characters number of leaves at first flower, number of days to first flower, prickle density, flower size, flower weight, pedicel length, number of petals flower$^{-1}$, size of petals and number of flowers plant $^{-1}$/bunch $^{-1}$ were used for clustering based on quantitative character. Clustering techniques such as single linkage, complete linkage, UPGMA, WPGMA, Ward's method, modified Tocher method and k- means clustering were performed using different measures of distance.

#### *4.5.2.1.1 Single linkage clustering of Floribunda genotypes based on quantitative characters*

Single linkage clustering of 25 genotypes coming under Floribunda group was carried out using the nine quantitative data. Similarity measures used were Squared Euclidean distance, Euclidean distance, Chebychev distance and City block distance. Clusters were formed by combining objects having smallest distance. Distance between two clusters was measured as the distance between nearest objects. Clustering was done with the help of SPSS and SAS statistical package. Clustering schedule got some differences with different distance measures except for Euclidean and Squared Euclidean distance. The result of different clustering techniques based on Squared Euclidean results gave approximately same result as that of Euclidean distance. So the result corresponding to Euclidean distance is not presented separately to avoid duplication. Dendrogram of single linkage method using different distance measures are shown in Figure 40, Figure 41 and Figure 42. Clustering membership of different genotypes under various distance measures using single linkage method are given in Table 27.

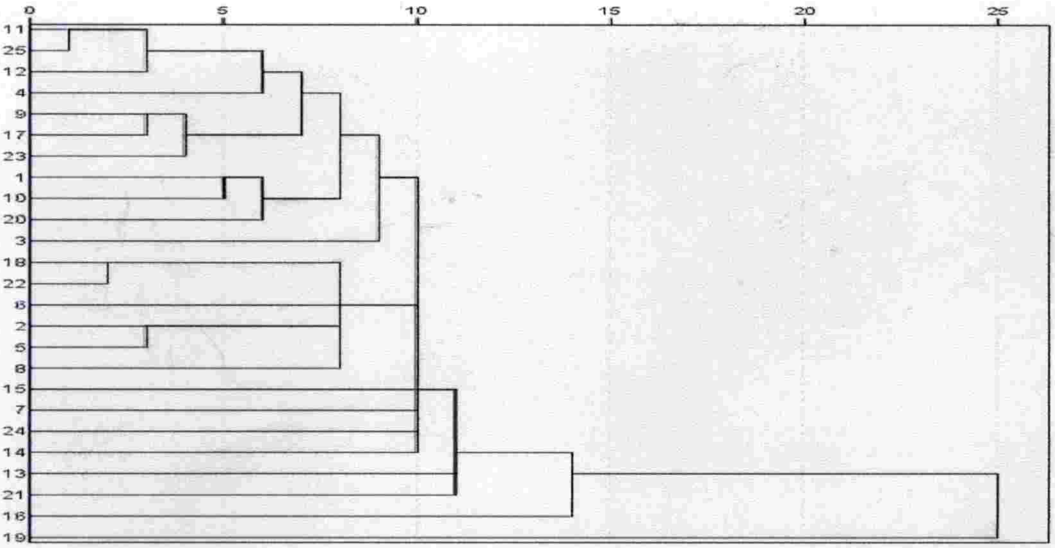


Fig. 40: Dendrogram of single linkage clustering of Floribunda genotypes based on Squared Euclidean distance.

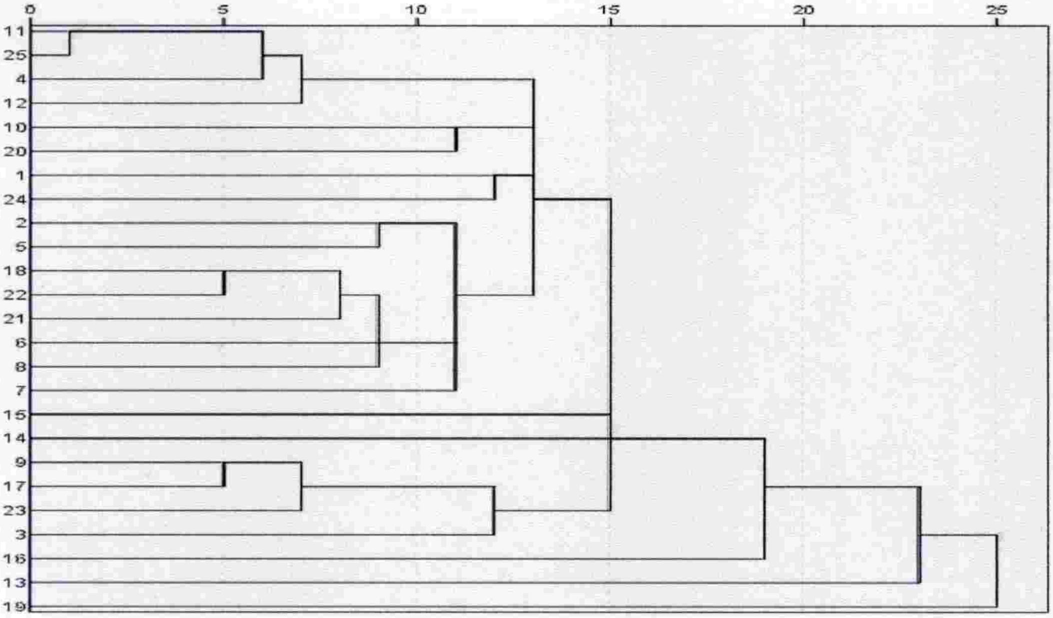


Fig. 41: Dendrogram of single linkage clustering of Floribunda genotypes based on Chebychev distance.

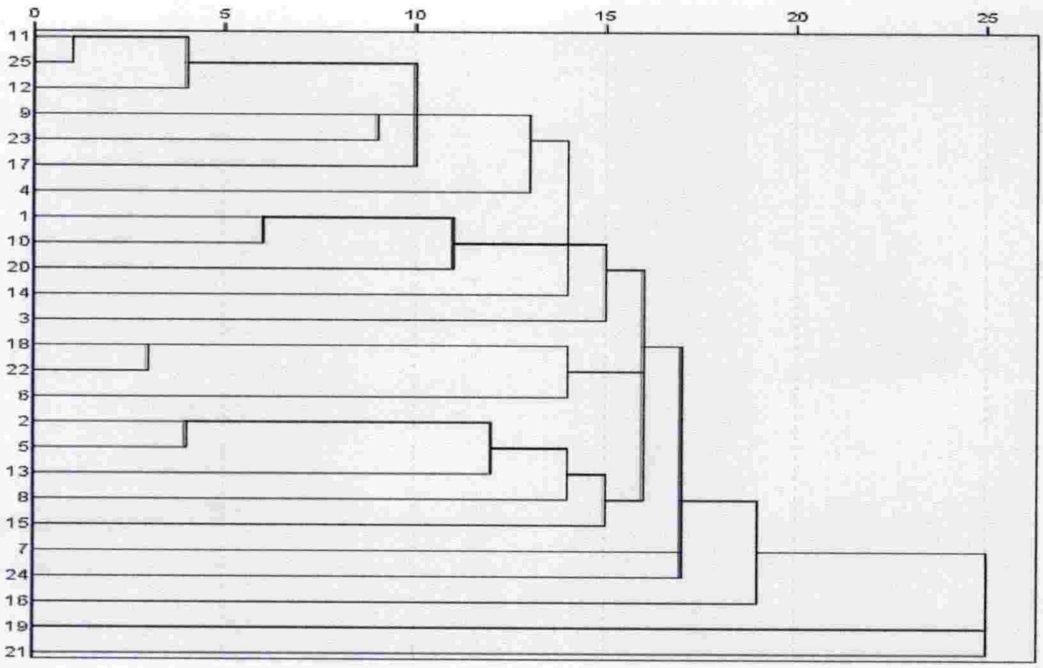


Fig. 42: Dendrogram of single linkage clustering of Floribunda genotypes based on City block distance.

Table 27. Cluster membership of Floribunda genotypes under single linkage based on quantitative data.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | F1, F3, F4, F9, F10, F11, F12, F17, F20, F23, F25 | F1, F2, F4, F5, F6, F7, F8, F10, F11 | F1, F3, F4, F9, F10, F11, F12, F14, F17, F20, F23, F25 |
| 2 | F2, F5, F6, F8, F18, F22 | F12, F18, F20, F21, F22, F 24, F25 | F2, F5, F8, F13, F15 |
| 3 | F15 | F3, F9, F17, F23 | F6,F 18, F22 |
| 4 | F7 | F14, F15 | F7 |
| 5 | F24 | F16 | F24 |
| 6 | F14 | F13 | F16 |
| 7 | F13 | F19 | F19 |
| 8 | F21 |  | F21 |
| 9 | F16 |  |  |
| 10 | F19 |  |  |

Grouping based on Squared Euclidean distance resulted in ten clusters. The largest cluster included 12 members followed by the second largest cluster with 6 members. All the other clusters using Squared Euclidean distance are single member clusters. Seven clusters were formed when Chebychev distance was used as distance measure. Cluster with nine members formed the largest cluster followed by clusters with 7 and 4 members. All the others are clusters with single members. Clustering based on City block distance resulted in 8 clusters with 5 single member clusters. Largest cluster included 12 members. Genotypes F19 (Sans Souci) and F16 (Messara) formed single member clusters in all the three distance measures. F21 (Lasting Piece) and F7 (Carry Free Beauty) also formed single member clusters under Squared Euclidean and City Block distance. F13 (Monnalisa) was found to form a single member cluster under Squared Euclidean and Chebychev distances.

#### *4.5.2.1.2 Complete linkage clustering of Floribunda genotypes based on quantitative characters*

Complete linkage clustering was performed using SPSS and SAS statistical package. Dendrograms were drawn using different distance measures. Distance between two clusters was measured as the distance between furthest objects. Clusters obtained from different distance measures were different from each other. Dengrograms obtained from different distance measures are shown in Figure 43, Figure 44 and Figure 45. Cluster membership of different genotypes under different distance measures is given the Table 28.

Genotypes were divided into 7 clusters under Squared Euclidean distance, 6 clusters under Chebychev distance and 8 clusters under City Block distance using complete linkage method. Among them F19 (Sans Souci) formed a single member cluster under Squared Euclidean and Chebychev distances but it formed a cluster with F3 (Rosarale de Chateau) under City Block distance. The largest cluster contained six members under Squared Euclidean, eight members under Chebychev distance and five members under City Block distance. F3 (Rosarale de Chateau), F9 (Orange N Lemon), F17 (Michel Fish), F23 (Winchester Cathedral) and F1 (Versailles), F10 (Lisa), F20 (Schloss Elutin) and F24 (

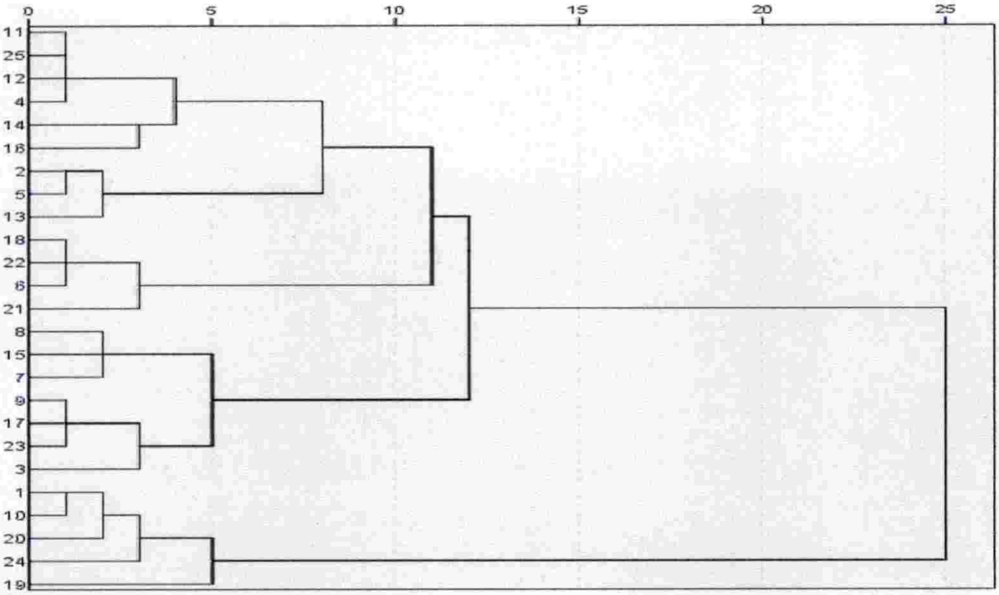


Fig. 43: Dendrogram of complete linkage clustering of Floribunda genotypes based on Squared Euclidean distance.

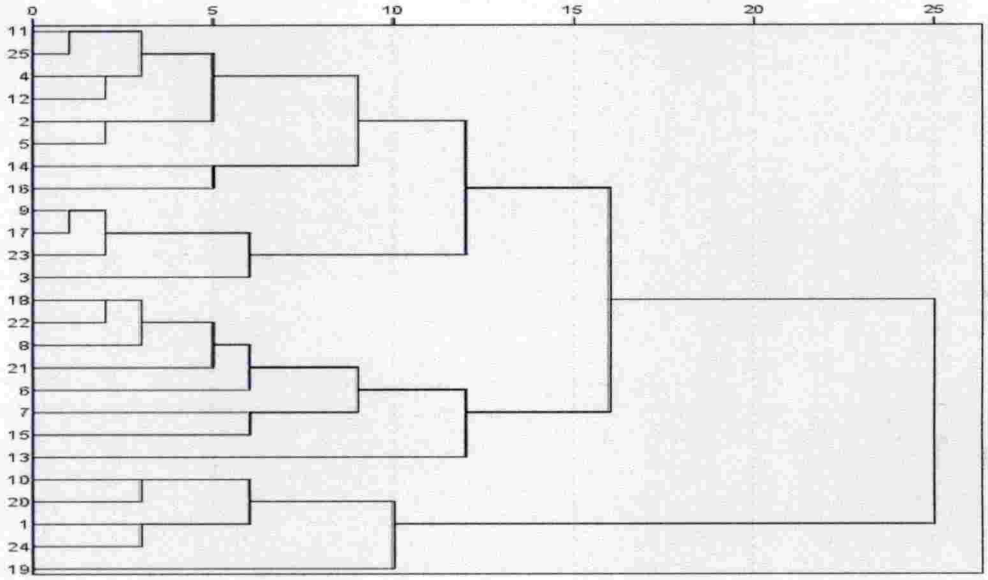


Fig 44: Dendrogram of complete linkage clustering of Floribunda genotypes based on Chebychev distance.

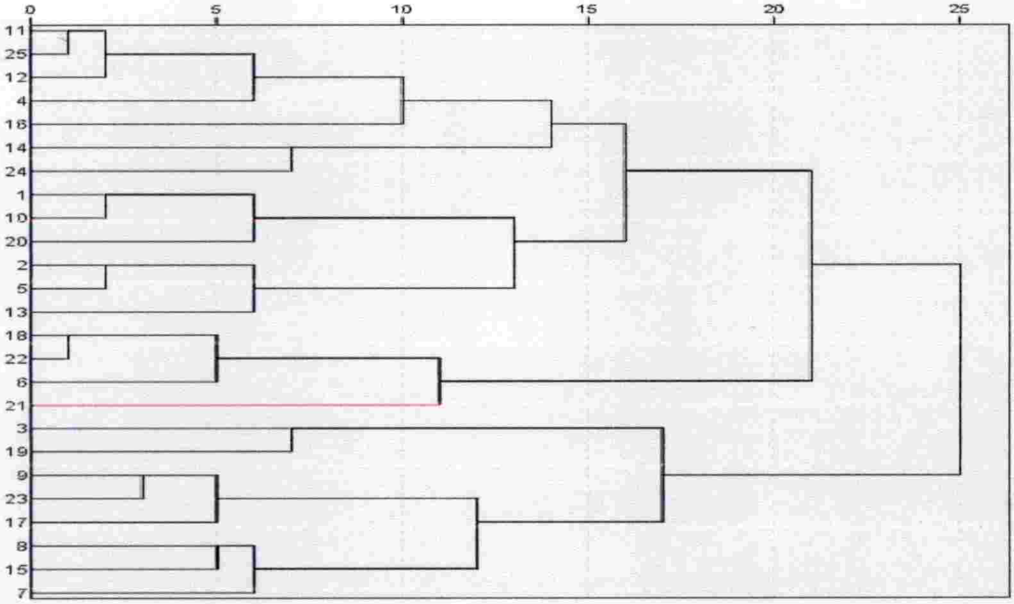


Fig. 45: Dendrogram of complete linkage clustering of Floribunda genotypes based on City block distance

Table 28. Cluster membership of Floribunda type under complete linkage based on quantitative data.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | F4, F11, F12, F14, F16, F25 | F2, F4, F5, F11, F12, F14, F16, F25 | F4, F11, F12, F 16, F25 |
| 2 | F3, F9, F17, F23 | F6, F7, F8, F15, F18, F21, F22 | F6, F18, F22, F21 |
| 3 | F6, F18, F21, F22 | F3, F9, F17, F23 | F1, F10, F20 |
| 4 | F1, F10, F20, F24 | F1, F10, F20, F24 | F2, F5, F13 |
| 5 | F7, F8, F15 | F19 | F9, F18, F23 |
| 6 | F2, F5, F13 | F13 | F7, F8, F15 |
| 7 | F19 |  | F14, F24 |
| 8 |  |  | F3, F19 |

Golden Fairy) came under the same cluster in Squared Euclidean and Chebychev distances. Under Chebychev distance F13 (Monnalisa) formed a single cluster.

#### *4.5.2.1.3 UPGMA of Floribunda genotypes based on quantitative characters*

Unweighted pair group average method of clustering was done with quantitative characters and different measures of distances. Composition of different clusters obtained by applying different distance measures showed variation. Distances between clusters were taken as the average distance between objects. Pseudo $t^2$ statistics were used to find out the optimum number of clusters with the help of SAS statistical package.

Figure 46 shows possibly good clustering levels at five clusters, eight clusters, and thirteen clusters and at nineteen clusters. Optimum numbers of cluster are three, six, eight and seventeen under Chebybychev distance (Figure 47). For City Block distance it comes as six, seven, thirteen and nineteen clusters. (Figure 48). Dendrogram for UPGMA method using different distance measures are shown in Figure 49, Figure 50 and Figure 51.Clustering pattern of different genotypes under UPGMA method given in the Table 29.

UPGMA method grouped the 25 Floribunda genotypes into eight clusters under Squared Euclidean and Chebychev distances and six clusters under City Block distance. Under Squared Euclidean, cluster with seven members formed the largest cluster. F13 (Monnalisa), F19 (Sans Souci) and F16 (Messara) formed individual cluster. Under Chebychev distance cluster with seven clusters formed the largest clusters. Cluster with six members formed the second largest cluster followed by cluster with four members, cluster with three members, cluster with two members and all others formed single member clusters. Under City Block distance, cluster with eight members formed the largest cluster, followed by cluster with six members, cluster with five members and cluster with four members. F21 (Lasting Piece) and F19 (Sans Souci) formed individual clusters. F19 (San Souci) formed single cluster in all the distance methods

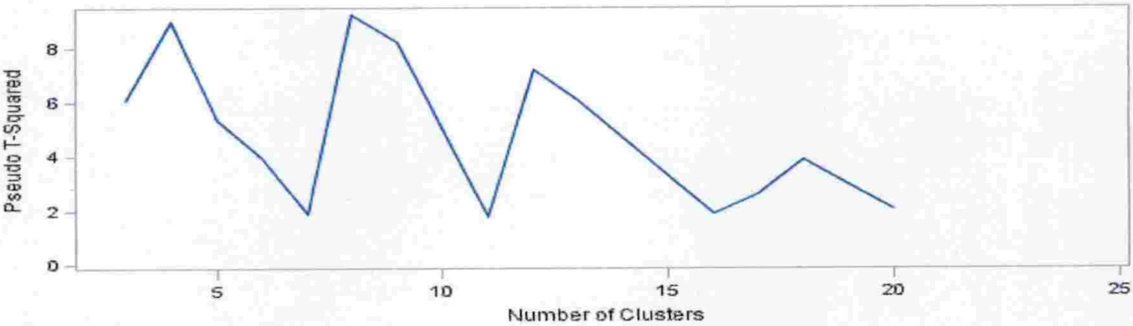


Fig. 46: Pseudo $t^2$ statistic plot of Floribunda genotypes for Squared Euclidean distance under UPGMA

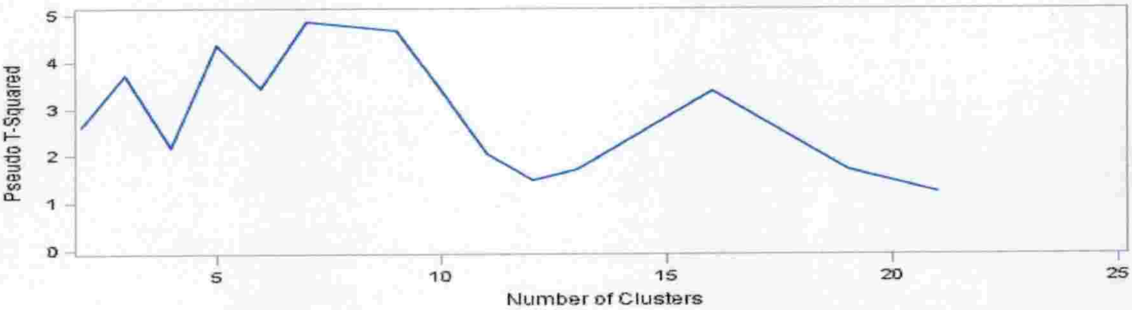


Fig. 47: Pseudo $t^2$ statistic plot of Floribunda genotypes based on Chebychev distance under UPGMA

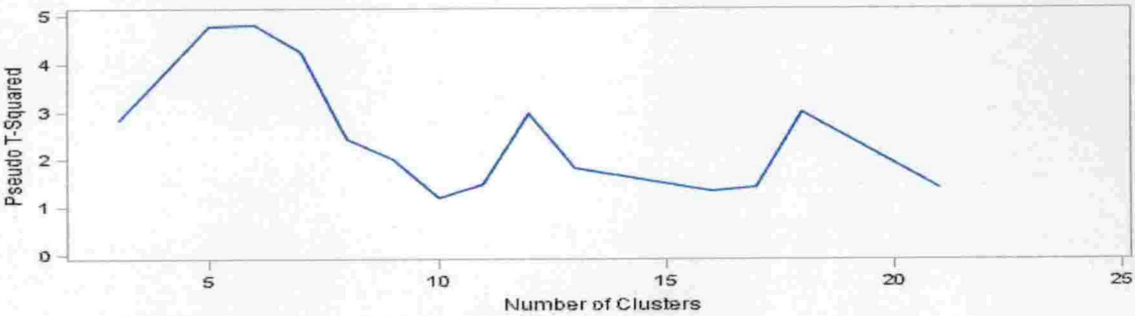


Fig. 48: Pseudo $t^2$ statistic plot of Floribunda genotypes based on City Block distance under UPGMA

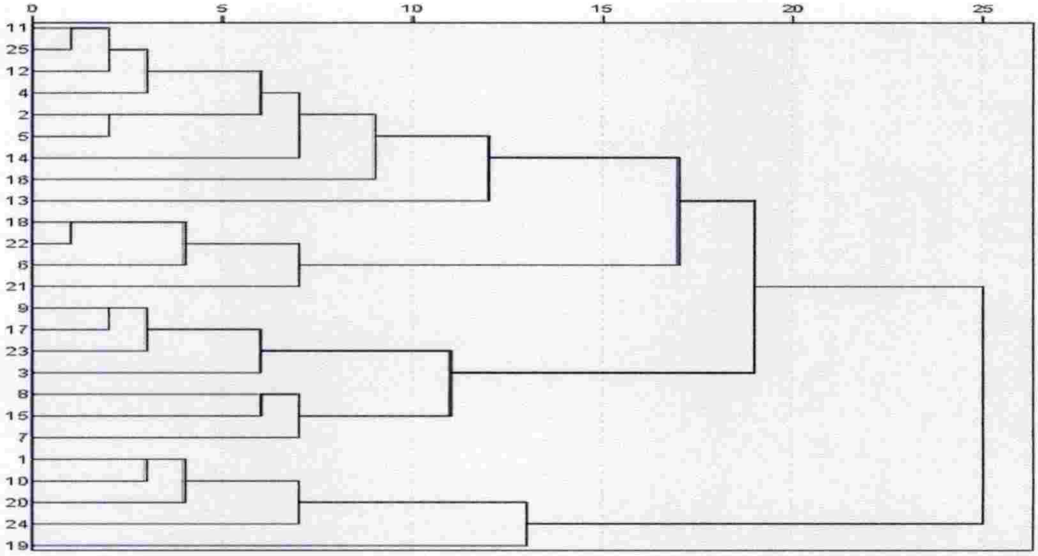


Fig. 49: Dendrogram of Floribunda genotypes based on UPGMA method under Squared Euclidean distance.

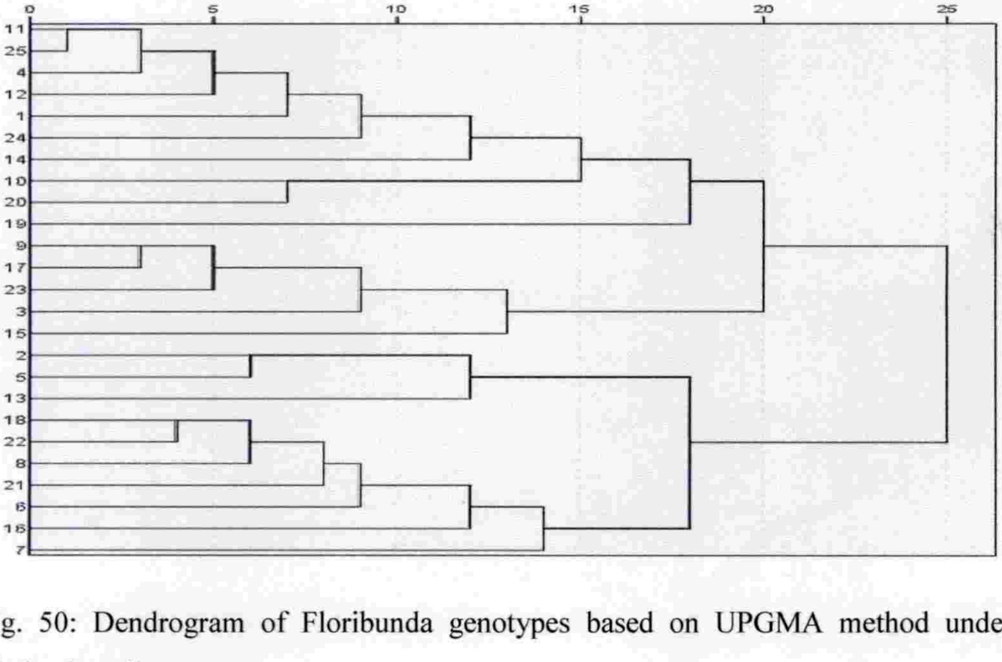


Fig. 50: Dendrogram of Floribunda genotypes based on UPGMA method under Chebychev distance.

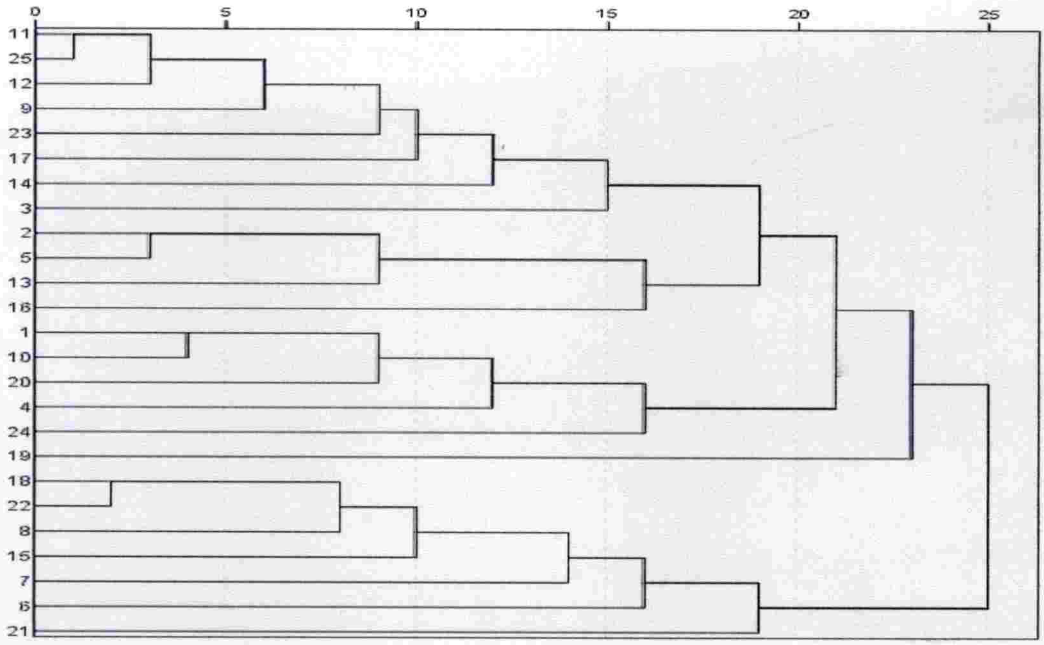


Fig. 51: Dendrogram of Floribunda genotypes based on UPGMA method under City Block distance.

Table 29. Clustering pattern of Floribunda genotypes under UPGMA based on quantitative data.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | F2, F4, F5, F11, F12, F14, F16, F25 | F1, F4, F11, F12, F14, F24, F25 | F3, F9, F11, F12, F14, F18, F23, F25 |
| 2 | F3, F9, F17, F23 | F6, F8, F16, F18, F21, F22 | F6, F7, F8, F15, F18, F22 |
| 3 | F6, F18, F21, F22 | F3, F9, F17, F23 | F2, F5, F13, F16 |
| 4 | F7, F8 F15 | F2, F5, F13 | F1, F4, F20, F24, F10 |
| 5 | F1, F10, F20, F24 | F10, F20 | F21 |
| 6 | F13 | F15 | F19 |
| 7 | F19 | F7 | |
| 8 | F16 | F19 | |

### *4.5.2.1.4 WPGMA of Floribunda genotypes based on quantitative characters*

Clustering was completed by calculating the similarity between two clusters as the mean similarity of previously existing clusters. Clustering was done with the help of SPSS statistical package. Clustering results are presented in Figure 52, Figure 53 and Figure 54. Clustering pattern of genotypes under WPGMA method is shown below in Table 30.

Genotypes were grouped into eight clusters under Squared Euclidean and City Block distance method while they were grouped into six clusters under Chebychev distance. F19 (Sans Souci) formed a single cluster in all the distance methods. F21 (Lasting Piece) formed a single cluster under City Block distance and F16 (Messara) under Squared Euclidean distance. Largest cluster contained five, eight and seven members under Squared Euclidean, Chebychev and City Block distance respectively. F3 (Rosarale de Chateau), F9 (Orange N Lemon), F17 (Michel Fish), F23 (Winchester Cathedral) and F7 (Carry Free Beauty), F8 (Sterntaler) and F15 (City of Glasgow) came under single member clusters in Squared Euclidean and City Block distance. F3 (Rosarale de Chateau), F9 (Orange N Lemon), F17 (Michel Fish) and F23 (Winchester Cathedral) came under single cluster in all the distance methods. Slight variations are there with member respect to the other genotypes under different distance methods.

### *4.5.2.1.5 UPGMC of Floribunda genotypes based on quantitative characters*

Clustering was performed using Squared Euclidean measure with Unweighted Pair Group Centroid Method. This clustering was performed using Squared Euclidean measure alone as it give valid result only for that distance measure. Objects in each cluster were replaced with the centroid of that particular cluster. Clustering was done with the help of statistical packages SPSS and SAS. Optimum number of clusters from Figure 55 is five, nine, eleven, thirteen and fifteen. Dendrogram of clusteing is as follows (Figure 56). Cluster membership of genotypes under UPGMC method is shown in Table 31.

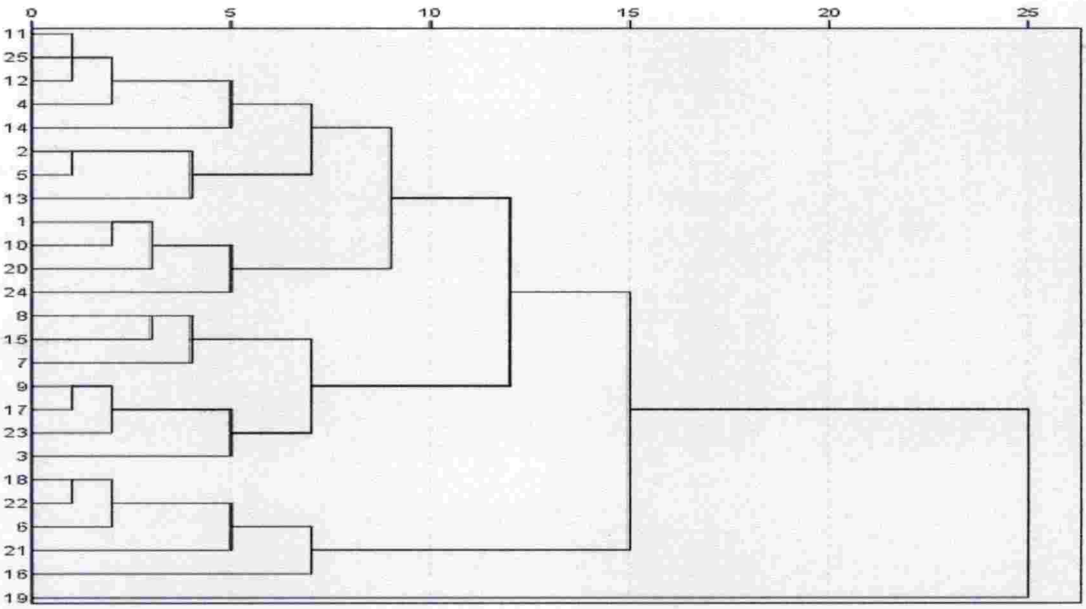


Fig. 52 Dendrogram of Floribunda genotypes under WPGMA method based on Squared Euclidean distance.

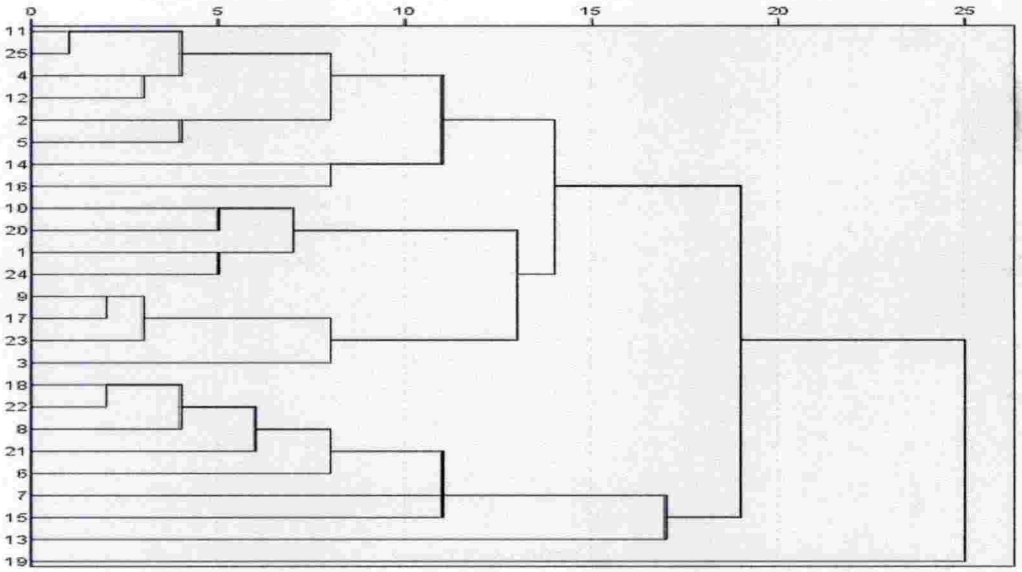


Fig. 53 Dendrogram of Floribunda genotypes under WPGMA method based on Chebychev distance.

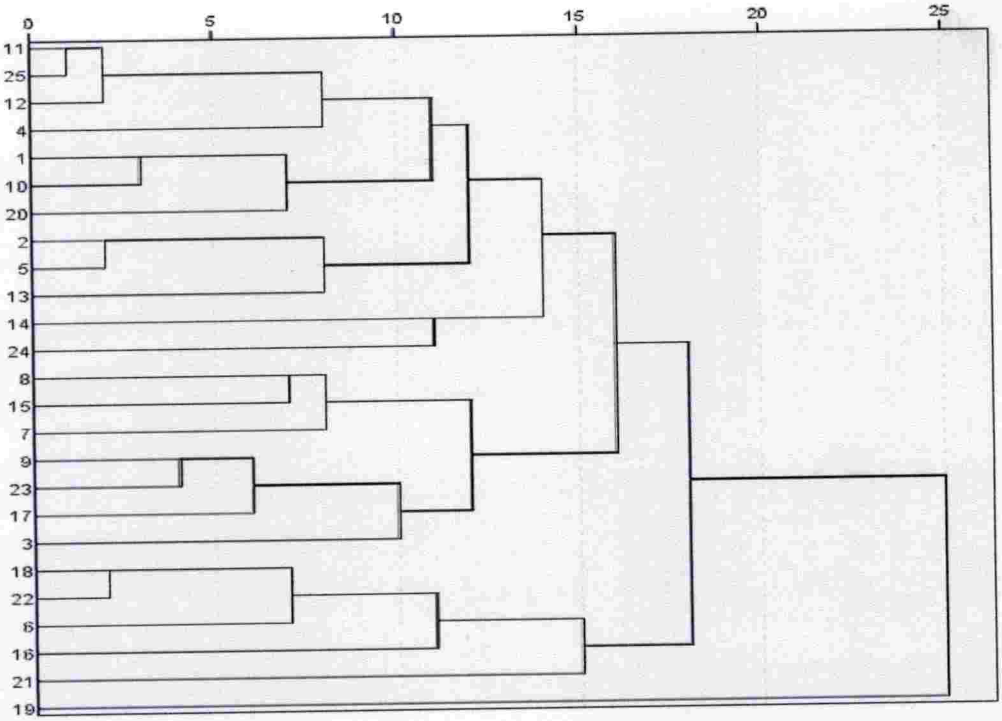


Fig. 54: Dendrogram of Floribunda genotypes under WPGMA method based on City block distance.

Table 30. Cluster membership of Floribunda type under WPGMA based on quantitative characters.

| Cluster number | Squared Euclidean | Chebychev distance | City Block distance |
|---|---|---|---|
| 1 | F4, F11, F12, 14, F25 | F2, F4, F5, F11, F12, F14, F16, F25 | F1, F4, F11, F10, F12, F20, F25 |
| 2 | F3, F9, F17, F23 | F6, F7, F8, F15, F18, F21, F22 | F3, F9, F17, F23 |
| 3 | F1, F10, F20, F24 | F3, F9, F17, F23 | F2, F5, F13 |
| 4 | F6, F18, F21, F22 | F1, F10, F20, F24 | F6, F16, F18, F22 |
| 5 | F7, F8, F15 | F16 | F7, F8, F15 |
| 6 | F2, F5, F13 | F19 | F14, F24 |
| 7 | F16 |  | F21 |
| 8 | F19 |  | F19 |

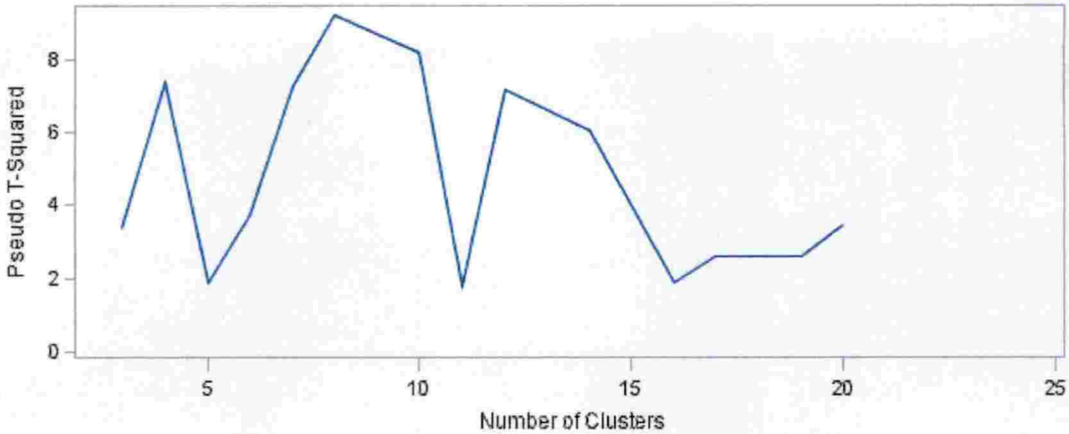


Fig. 55: Pseudo $t^2$ statistic plot of Floribunda genotypes for Squared Euclidean distance under UPGMC method.

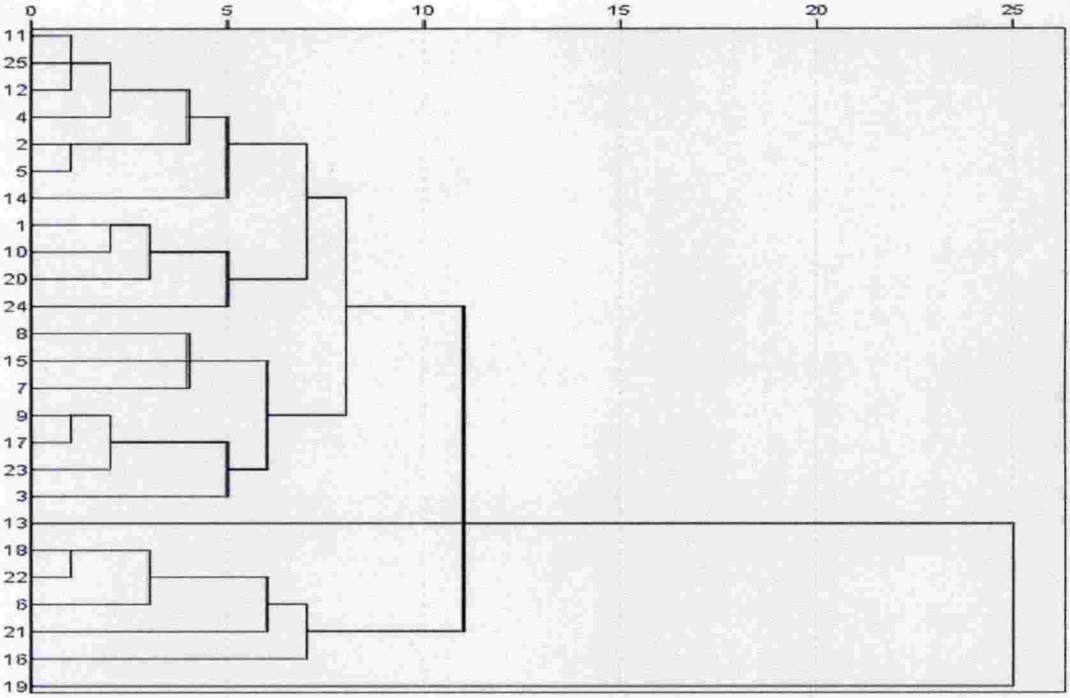


Fig. 56: Dendrogram of Floribunda genotypes under UPGMC method based on Squared Euclidean distance.

Table 31. Cluster membership of Floribunda type under UPGMC

| Cluster number | Squared Euclidean |
|---|---|
| 1 | F2, F4, F5, F11, F12, F14, F25 |
| 2 | F3, F9, F17, F23 |
| 3 | F1, F20, F10, F24 |
| 4 | F8, F7, F15 |
| 5 | F6, F18, F22 |
| 6 | F13 |
| 7 | F21 |
| 8 | F16 |
| 9 | F19 |

Twenty five genotypes were grouped into nine clusters. First cluster formed the largest cluster with seven members followed by second and fourth clusters with four members and third and sixth clusters with three members. F13 (Monnalisa), F21 (Lasting Piece), F16 (Messara) and F19 (Sans Souci) formed single clusters.

#### *4.5.2.1.6 Ward's method of Floribunda genotypes based on quantitative characters*

Ward's method of clustering based on minimizing sum of square was performed using Squared Euclidean distance with the help of statistical packages SPSS and SAS. Optimum number of clusters was obtained from the Pseudo $t^2$ statistic plot (Figure 57) as seven, eleven, fourteen, sixteen and eighteen. Dendrogram of Ward's method with Squared Euclidean distance (Figure 58) at a scale three gave seven clusters. Table of cluster membership (Table 32) showed that first cluster with six members forms the largest cluster followed by third, fifth and sixth clusters with four members each and second and fourth clusters with three members each. F19 (Sans Souci) formed a cluster with single member.

#### *4.5.2.1.7 $D^2$ analysis of Floribunda genotypes*

The $D^2$ distances among varieties were estimated based on quantitative characters and the values are presented in the Appendix XX. The $D^2$ values were arranged in the ascending order and clustering was done by modified Tocher method. The 25 genotypes corresponding to Floribunda type were clustered into seven clusters and cluster membership corresponding to each cluster is shown in Table 33. There were 10 members in the first cluster, five members in clusters two and three, two members in cluster four and clusters five, six and seven are single member clusters. The intra and inter cluster distances are given in Table 34. Genetic divergence was maximum between clusters II and VI (1864.17) followed by clusters V and VI (1785.08), clusters V and VII (1682.83) and clusters VI and VII (1476.95). Cluster V, VI and VII were clusters with single objects and having zero intracluster distance.Table 35 shows the relative contribution of different characters towards the divergence. Number of leaves at first flower (36.33%) had the highest contribution towards divergence followed by

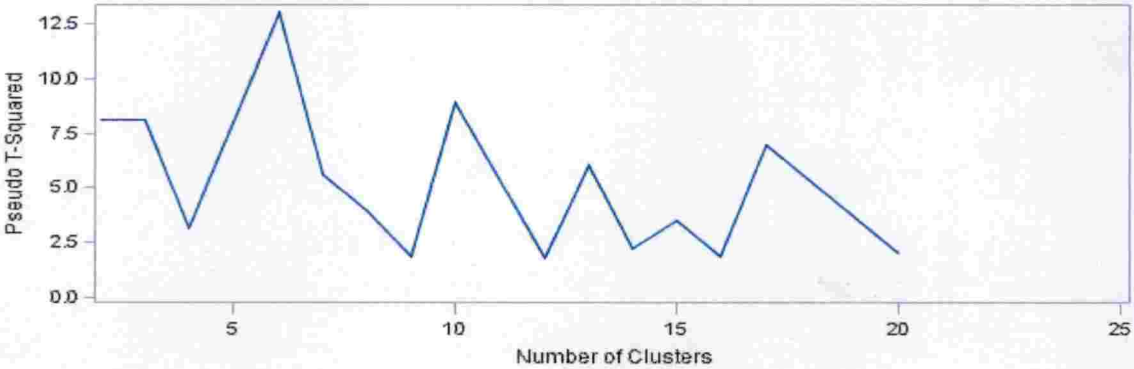


Fig. 57: Pseudo $t^2$ statistic plot of Floribunda genotypes based on Squared Euclidean distance under Ward's method

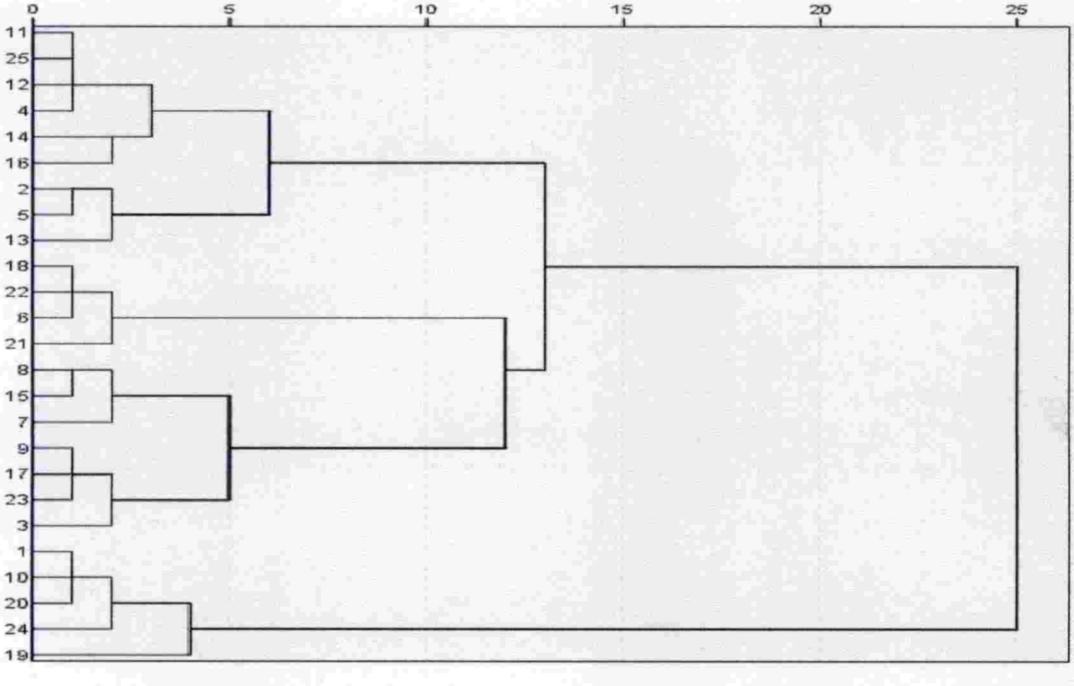


Fig.58: Dendrogram of Floribunda genotypes under UPGMC method based on Squared Euclidean distance.

Table 32. Cluster membership of Floribunda type under Ward's method

| Cluster number | Squared Euclidean distance |
|---|---|
| 1 | F4, F11, F12, F14, F16, F25 |
| 2 | F3, F9, F17, F23 |
| 3 | F6, F18, F21, F22 |
| 4 | F1, F10, F20, F24 |
| 5 | F2, F5, F13 |
| 6 | F7, F8, F15 |
| 7 | F19 |

Table 33. Cluster membership of Floribunda genotypes based on $D^2$ statistics.

| Cluster number | Name of the Genotypes | No. of Genotypes |
|---|---|---|
| I | F1, F10, F25, F20, F4, F2, F5, F12, F9, F11 | 10 |
| II | F18, F22, F6, F8, F16 | 5 |
| III | F15, F17, F23, F3, F7 | 5 |
| IV | F14, F24 | 2 |
| V | F21 | 1 |
| VI | F19 | 1 |
| VII | F13 | 1 |

Table 34. Inter and intra cluster distances of Floribunda Genotypes based on $D^2$ values.

| | I | II | III | IV | V | VI | VII |
|---|---|---|---|---|---|---|---|
| **I** | 289 | 552.57 | 730.43 | 597.02 | 1353.88 | 1099.56 | 532.66 |
| **II** | | 312.67 | 737.40 | 849.28 | 917.58 | 1864.17 | 600.19 |
| **III** | | | 322.39 | 927.87 | 1218.79 | 930.88 | 910.74 |
| **IV** | | | | 400.79 | 841.82 | 873.49 | 1001.88 |
| **V** | | | | | 0 | 1785.08 | 1682.83 |
| **VI** | | | | | | 0 | 1476.95 |
| **VII** | | | | | | | 0 |

Table 35. Relative contribution of characters towards divergence for Floribunda genotypes.

| Character | Percentage contribution to variance |
|---|---|
| No. of leaves at first flower | 36.33 |
| No. of days to first flower | 32.00 |
| Prickle density/5cm | 0.67 |
| Flower size (cm) | 7.00 |
| Flower weight (g) | 2.33 |
| Pedicel length (cm) | 2.00 |
| No. of petals/ flower | 11.33 |
| Size of petals (cm) | 8.33 |
| No. of flower per plant/bunch | 0.00 |

no. of days to first flower (32.00%), number of petals/ flower (11.33%), flower size (7.00%), size of petals (8.33%), flower weight (2.33%), flower weight (2.00%),prickle density (0.67%) and no. of flower per plant/bunch did not contribute to divergence.

#### *4.5.2.2 Clustering of Floribunda genotypes based on qualitative characters*

Cluster analysis of Floribunda genotypes was carried out using seven traits. Among them three were qualitative traits and four were quantitative characters converted to qualitative characters. Fragrance, flower colour, vase life/ longevity are the qualitative characters included and prickle density (per five cm), number of petals flower$^{-1}$, size of petals (cm), number of flower plant $^{-1}$/bunch $^{-1}$ are the quantitative characters which were converted to qualitative characters. Different clustering methods were adopted with distance methods like Jaccard coefficient and Dice coefficient for binary data and Simple matching coefficient and Hamann's coefficient for multistage data.

##### *4.5.2.2.1 Single linkage clustering of Floribunda genotypes based on qualitative data*

Single linkage clustering of qualitative data was performed with the help of statistical packages SPSS, SAS and NTSYS. Clustering was done using Jaccard coefficient, Dice coefficient, Simple matching and Hamann's coefficient. Dendrogram of clustering are given in Figure 59, Figure 60, Figure 61 and Figure 62. Clustering pattern of genotypes is presented in Table 36. It was found that single linkage clustering using Jaccard and Dice distances have the same clustering pattern with one large cluster with nineteen members and all other genotypes forming individual clusters. Identical clustering was observed under Simple matching and Hamann's coefficient with the largest cluster including seventeen members. There were eight clusters under Jaccard and Dice distance and six clusters under Simple matching and Hamann's coefficient.

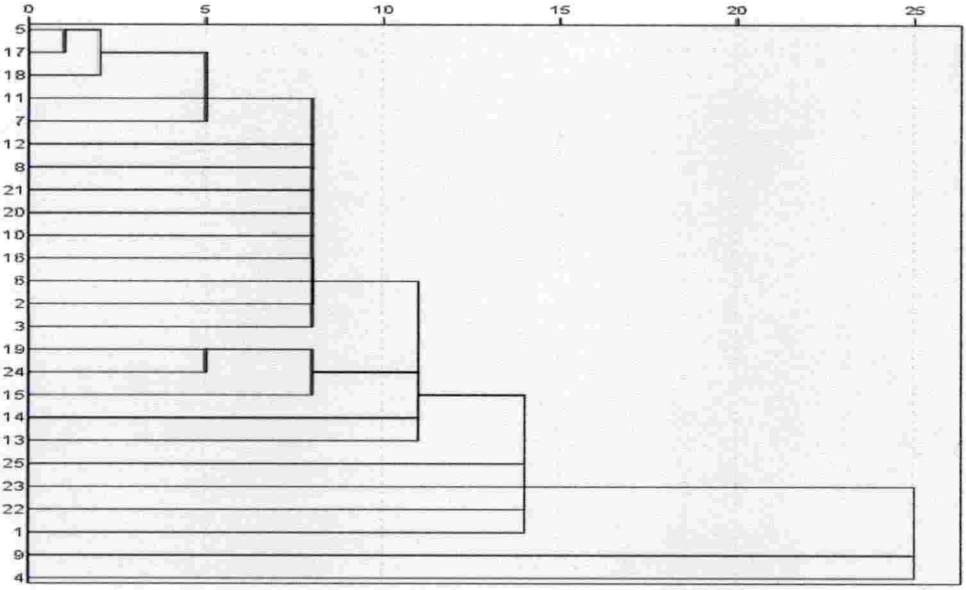


Fig.59: Dendrogram of Floribunda genotypes under single linkage method based on Jaccard coefficient

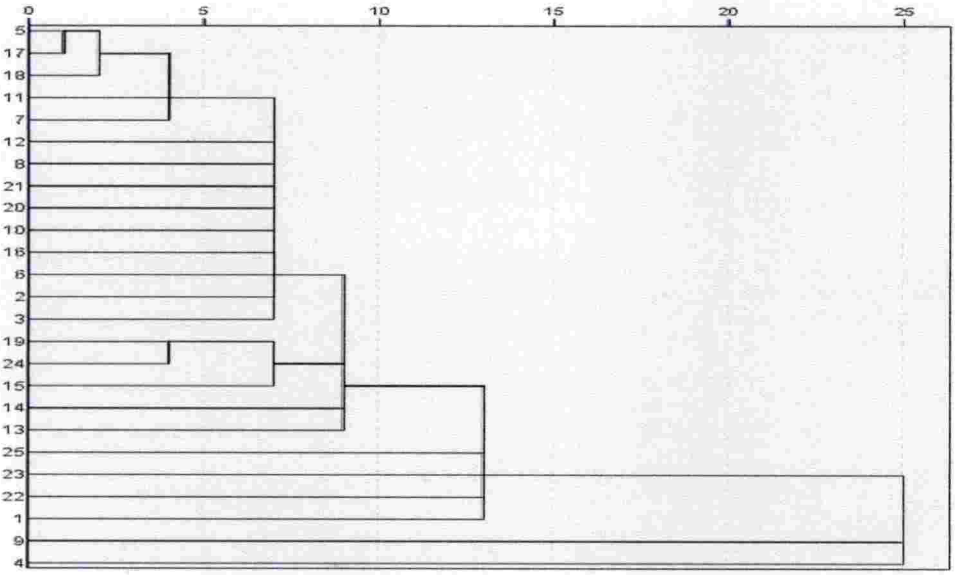


Fig.60: Dendrogram of Floribunda genotypes under single linkage method based on Dice coefficient

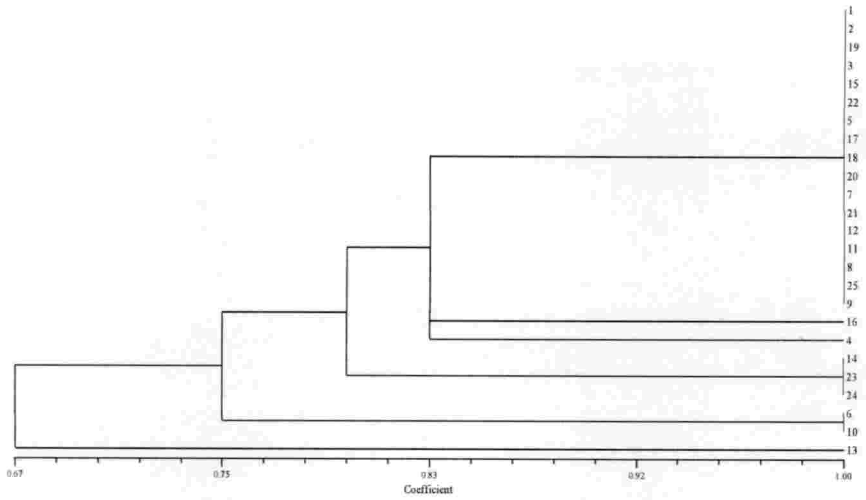


Fig.61: Dendrogram of Floribunda genotypes under single linkage method based on Simple matching coefficient

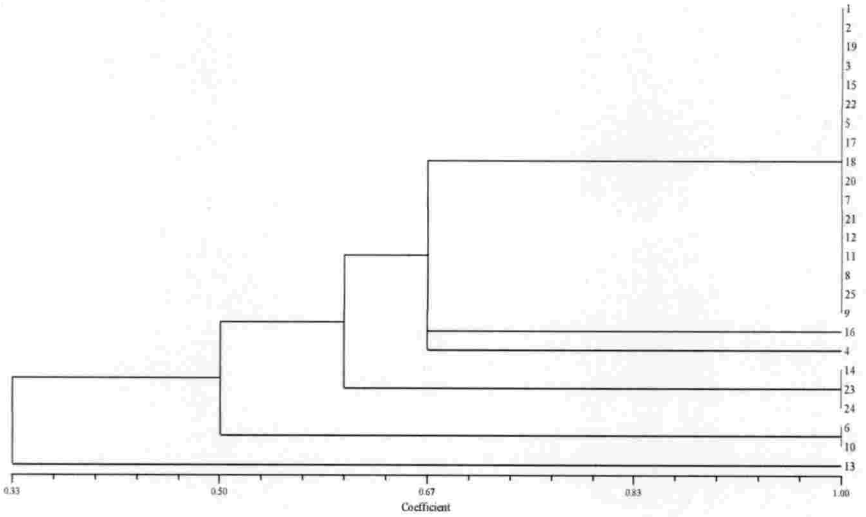


Fig.62: Dendrogram of Floribunda genotypes under single linkage method based on Hamann's coefficient

Table 36. Cluster membership of Floribunda type under single linkage for qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient | Simple matching |
|---|---|---|---|---|
| 1 | F5, F17, F18, F11, F7, F12, F8, F21, F20, F10, F6, F2, F3, F19, F24, F15, F13, F14 | F5, F17, F18, F11, F7, F12, F8, F21, F20, F10, F6, F2, F3, F19, F24, F15, F13, F14 | F1, F2, F19, F3, F15, F22, F5, F17, F18. F20, F7, F21, F12, F11, F8, F9, F25 | F1, F2, F19, F3, F15, F22, F5, F17, F18. F20, F7, F21, F12, F11, F8, F9, F25 |
| 2 | F25 | F25 | F14, F23, F24 | F14, F23, F24 |
| 3 | F23 | F23 | F6, F10 | F6, F10 |
| 4 | F22 | F22 | F16 | F16 |
| 5 | F1 | F1 | F4 | F4 |
| 6 | F9 | F9 | F13 | F13 |
| 7 | F4 | F4 | | |

#### *4.5.2.2.2 Complete linkage clustering of Floribunda genotypes based on qualitative data*

Complete linkage clustering was performed using Jaccard, Dice, Simple matching and Hamann's coefficients. Dendrograms are given in Figure 63, Figure 64 Figure 65 and Figure 66. Clustering pattern of genotypes is presented in Table 37. Here clustering under Jaccard and Dice showed similar clustering pattern whereas clustering under Hamann's coefficient and simple matching showed same clustering pattern. Genotypes were divided into eight clusters by Jaccard and Dice and into six clusters by Hamann,s coefficient and simple matching coefficient.

#### *4.5.2.2.3 UPGMA clustering of Floribunda genotypes based on qualitative data*

UPGMA method of clustering was performed with Jaccard, Dice, Simple matching and Hamann's coefficient with the help of statistical packages SPSS, SAS and NYSYS. Optimum number of clusters was obtained from the plot of Pseudo $t^2$ statistics (Figure 67, Figure 68, Figure 69, and Figure 70). It was observed that optimum number of cluster is seven for Dice distance, five, seven and ten for Jaccard distance, three and six for Hamann's coefficient and five eleven and thirteen for Simple matching coefficient. Dendrograms of clustering are given in Figure 71, Figure 72, Figure 73 and Figure 74. Clustering pattern of genotypes is presented in Table 38. Clustering under Jacaard and Dice distance showed the same clustering pattern. Clustering pattern under Simple matching and Hamann's are also same even though there is a variation in number of clusters.

#### *4.5.2.2.4 WPGMA clustering of Floribunda genotypes based on qualitative data*

WPGMA clustering was done under different distance methods. Clustering result as dendrogram is shown in Figure 75, Figure 76, Figure 77 and Figure 78. Cluster membership of genotypes is given in Table 39. Here also Jaccard & Dice and Simple matching and Hamann's coefficients have same clustering pattern when grouped into equal number of clusters. Seven clusters each were formed under Jaccard distance and Dice distance and six clusters each under Simple matching and Hamann's coefficient.

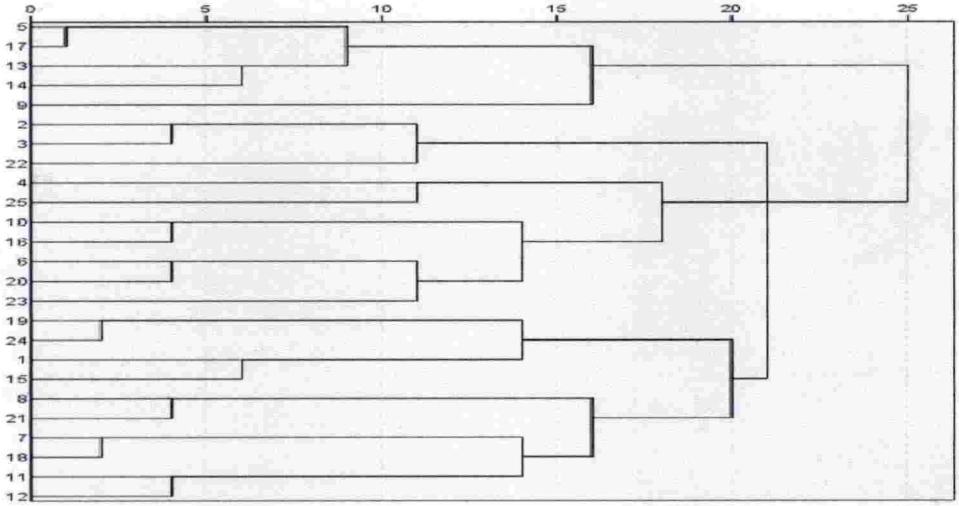


Fig.63: Dendrogram of Floribunda genotypes under complete linkage method based on Jaccard coefficient

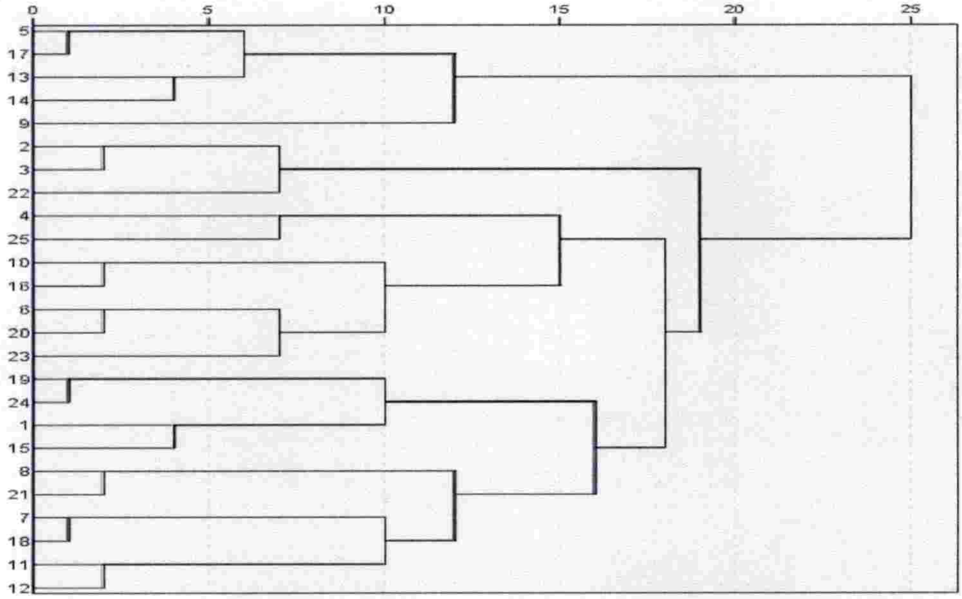


Fig.64: Dendrogram of Floribunda genotypes under complete linkage method based on Dice coefficient

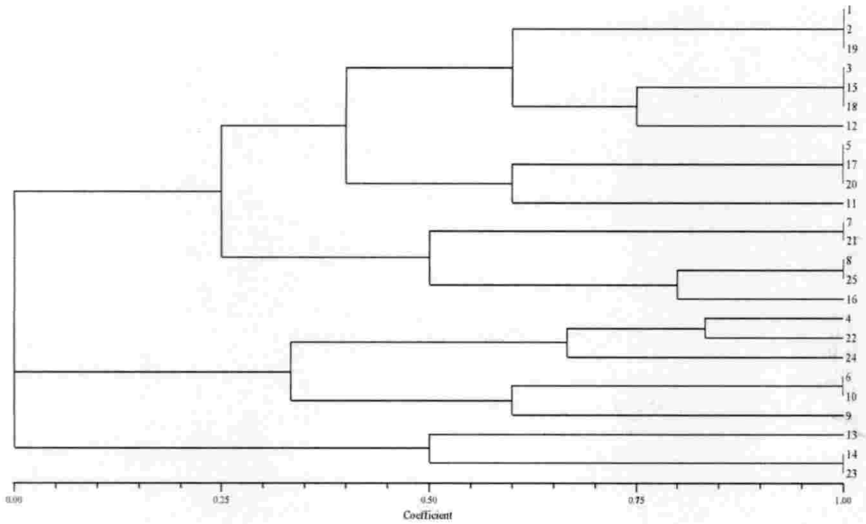


Fig.65: Dendrogram of Floribunda genotypes under complete linkage method based on Simple matching coefficient

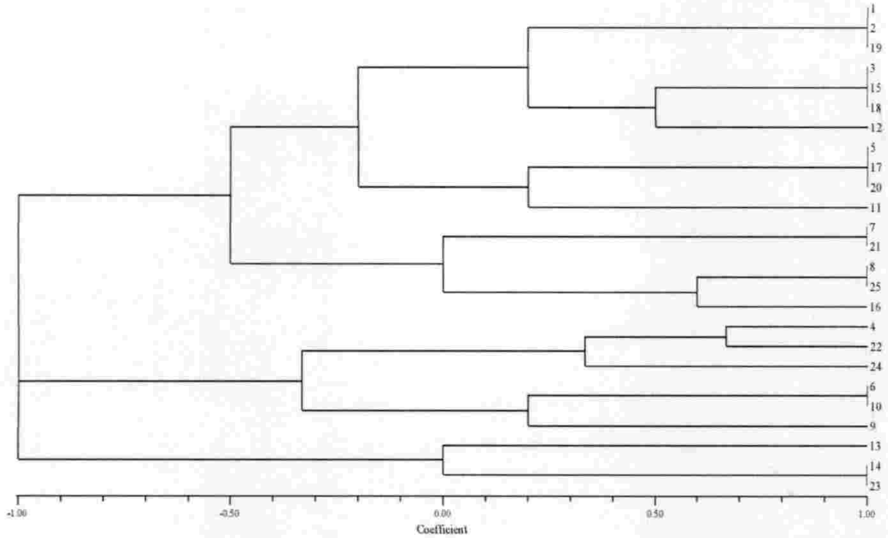


Fig.66: Dendrogram of Floribunda genotypes under complete linkage method based on Hamann's coefficient

Table 37. Cluster membership of Floribunda type under complete linkage based on qualitative data.

| Cluster number. | Jaccard distance | Dice distance | Hamann's coefficient | Simple matching |
|---|---|---|---|---|
| 1 | F10, F16, F6, F20, F23 | F10, F16, F6, F20, F23 | F1, F2, F19, F3, F15, F18, F12 | F1, F2, F19, F3, F15, F18, F12 |
| 2 | F19, F24, FF15, F1 | F19, F24, FF15, F1 | F8, F21, F25, F7, F16 | F8, F21, F25, F7, F16 |
| 3 | F7, F18, F11, F12 | F7, F18, F11, F12 | F5, F17, F20, F11 | F5, F17, F20, F11 |
| 4 | F5, F13, F14, F17 | F5, F13, F14, F17 | F4, F22, F24 | F4, F22, F24 |
| 5 | F2, F3, F22 | F2, F3, F22 | F6, F10, F9 | F6, F10, F9 |
| 6 | F4, F25 | F4, F25 | F13, F14, F23 | F13, F14, F23 |
| 7 | F8, F21 | F8, F21 | | |
| 8 | F9 | F9 | | |

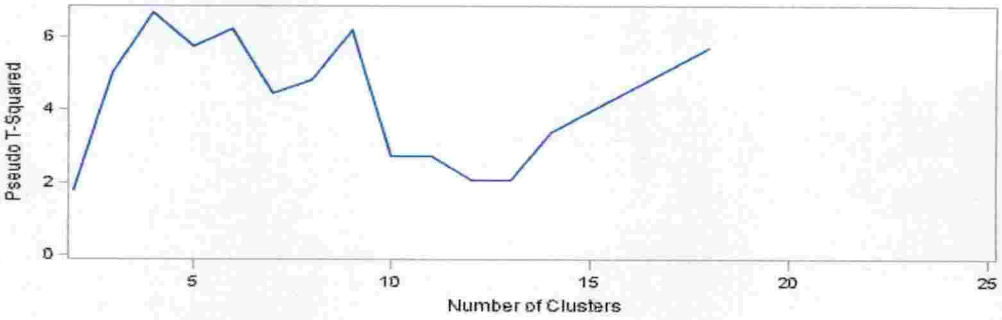


Fig. 67: Pseudo $t^2$ statistic plot for Floribunda genotype based on Jaccard coefficient under UPGMA

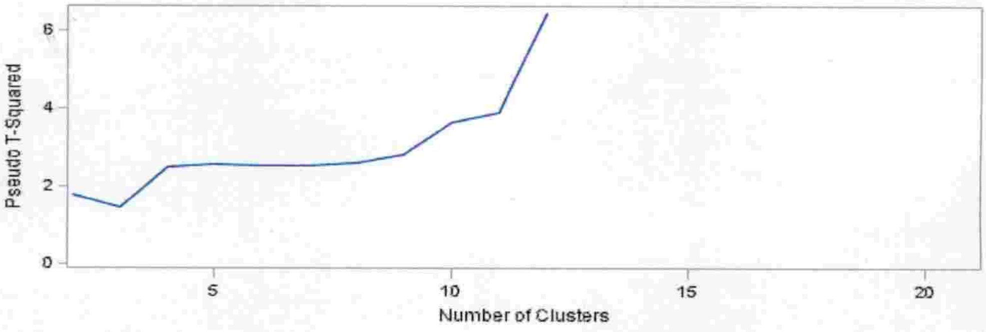


Fig. 68: Pseudo $t^2$ statistic plot for Floribunda genotype based on Dice coefficient under UPGMA

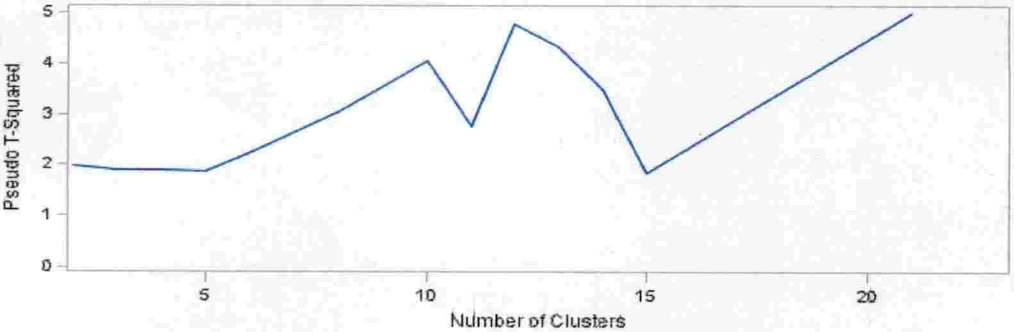
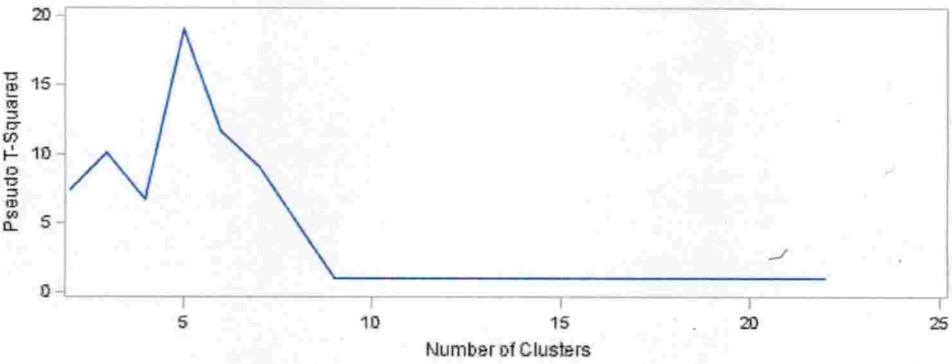


Fig. 69: Pseudo $t^2$ statistic plot for Floribunda genotype based on Simple matching coefficient under UPGMA



ig.

70: Pseudo $t^2$ statistic plot for Hamann's coefficient under UPGMA

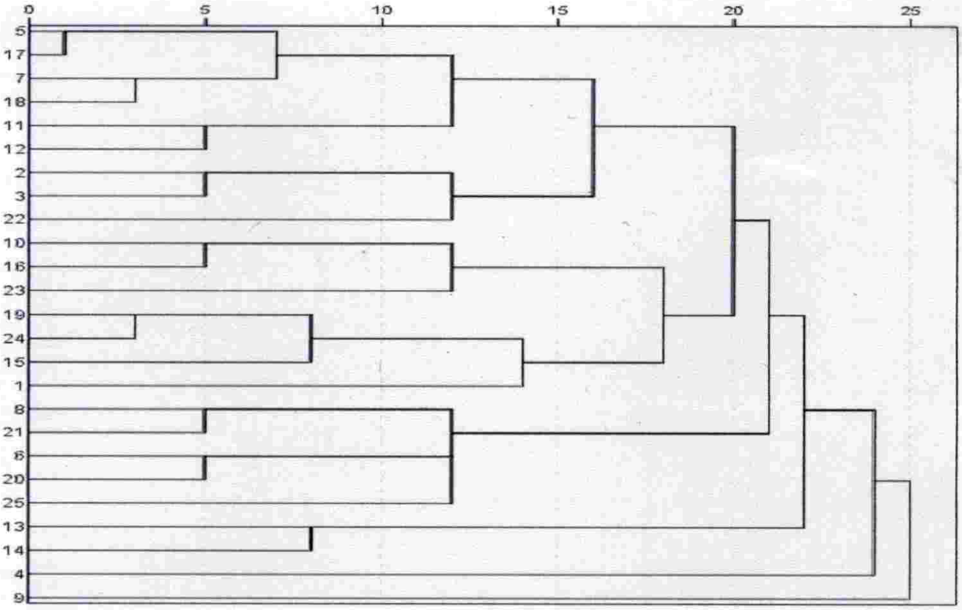


Fig.71: Dendrogram of Floribunda genotypes under UPGMA based on Jaccard coefficient

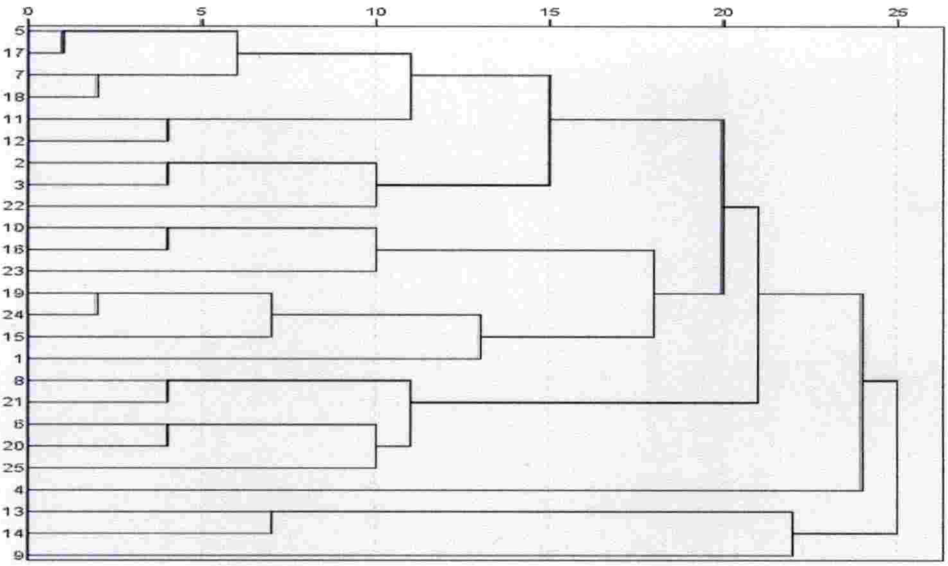


Fig.72: Dendrogram of Floribunda genotypes under UPGMA based on Dice coefficient

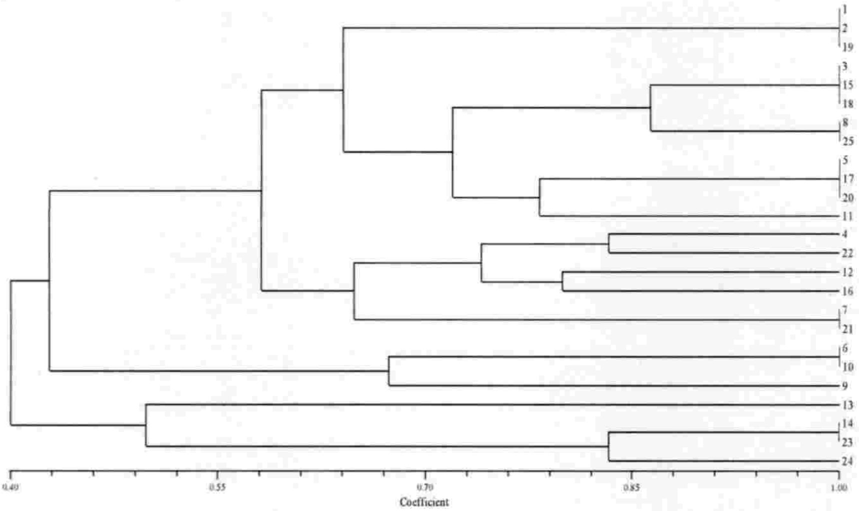


Fig.73: Dendrogram of Floribunda genotypes under UPGMA based on Simple matching coefficient

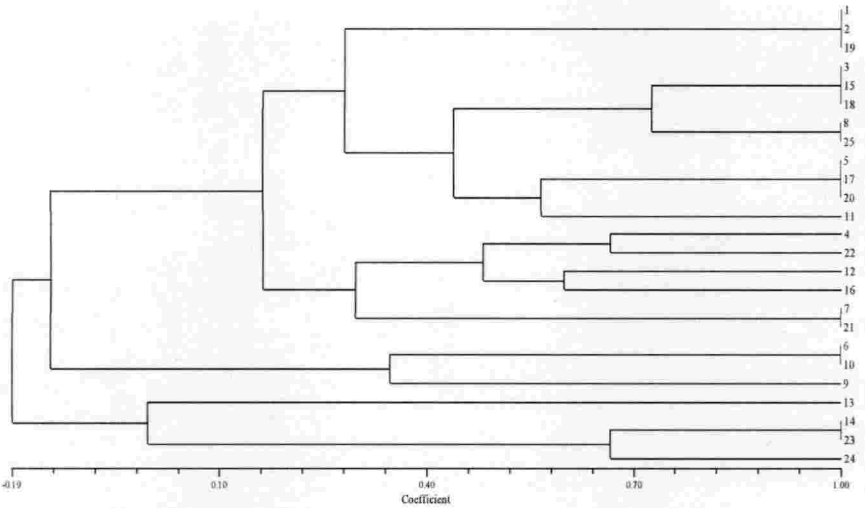


Fig.74: Dendrogram of Floribunda genotypes under UPGMA based on Hamann's coefficient

Table 38. Cluster membership of Floribunda type under UPGMA based on qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient | Simple matching |
|---|---|---|---|---|
| 1 | F5, F17, F7, F18, F11, F12, F22, F2, F3 | F5, F17, F7, F18, F11, F12, F22, F2, F3 | F3, F15, F18, F8, F5, F17, FF20, F25, F11 | F3, F15, F18, F8, F5, F17, FF20, F25, F11, F1, F2, F19 |
| 2 | F8, F6, F20, F21, F25 | F8, F6, F20, F21, F25 | F4, F22, F12, F16, F7, F21 | F4, F22, F12, F16, F7, F21 |
| 3 | F1, F15, F19, F24 | F1, F15, F19, F24 | F1, F2, F19 | F14, F23, F24 |
| 4 | F10, F16, F23 | F10, F16, F23 | F6,F10, F9 | F6,F10, F9 |
| 5 | F13, F14 | F13, F14 | F14, F23, F24 | F13 |
| 6 | F4 | F4 | F13 | |
| 7 | F9 | F9 | | |

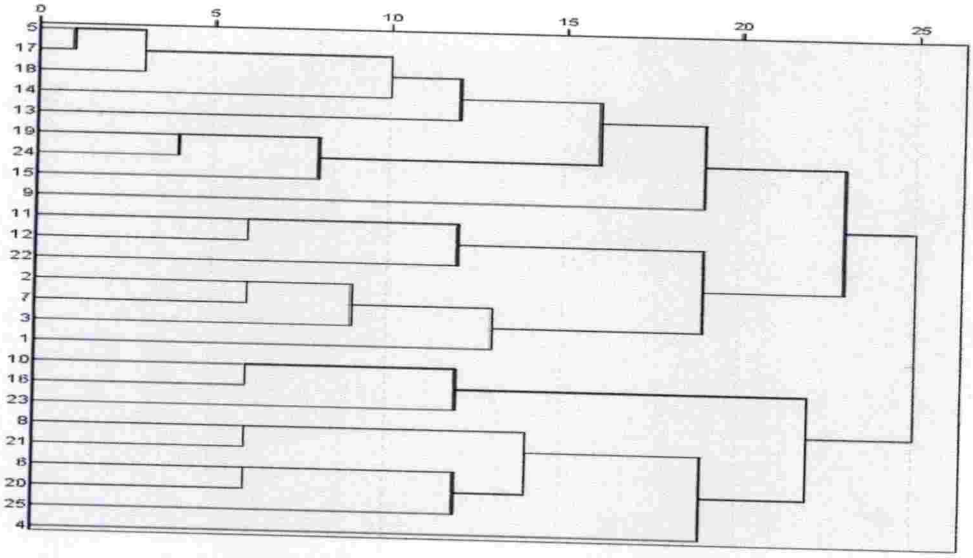


Fig.75: Dendrogram of Floribunda genotypes under WPGMA based on Jaccard coefficient

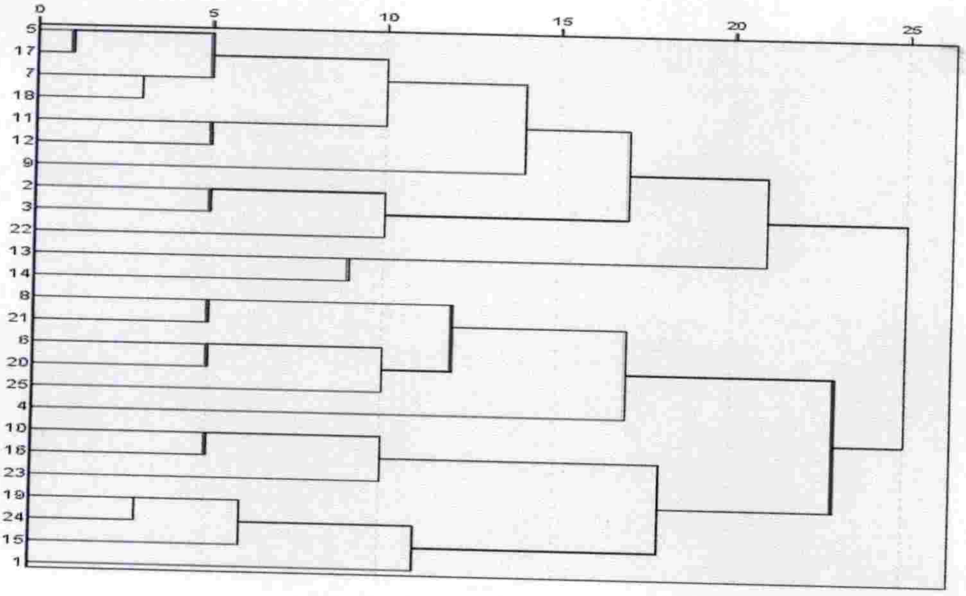


Fig.76: Dendrogram of Floribunda genotypes under WPGMA based on Dice coefficient

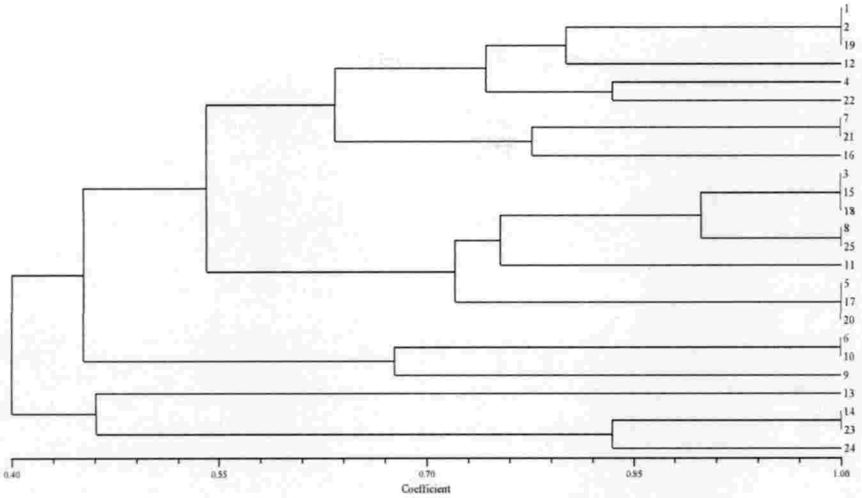


Fig.77: Dendrogram of Floribunda genotypes under WPGMA based on Simple matching coefficient

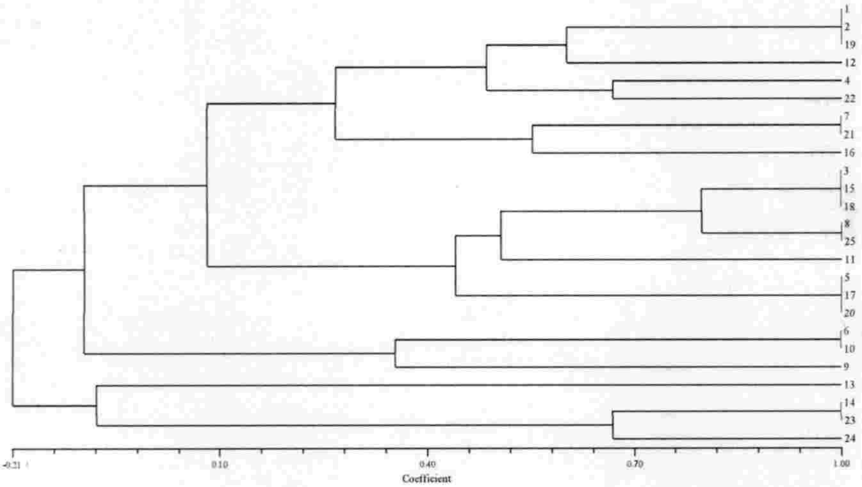


Fig.78: Dendrogram of Floribunda genotypes under WPGMA based on Hamann's coefficient

Table 39. Cluster membership of Floribunda type under WPGMA based on qualitative data.

| Cluster number | Jaccard distance | Dice distance | Hamann's coefficient | Simple matching |
|---|---|---|---|---|
| 1 | F5, F17, F18, F14, F13, F19, F15, F24 | F5, F17, F18, F14, F13, F19, F15, F24 | F3, F15, F18, F8, F11, F25, F5, F17, F20 | F3, F15, F18, F8, F11, F25, F5, F17, F20 |
| 2 | F8, F6, F20, F21, F25 | F8, F6, F20, F21, F25 | F1, F2, F19, F12, F4, F22 | F1, F2, F19, F12, F4, F22 |
| 3 | F1, F2, F3, F7 | F1, F2, F3, F7 | F7, F21, F16 | F7, F21, F16 |
| 4 | F11, F12, F22 | F11, F12, F22 | F14, F23, F24 | F14, F23, F24 |
| 5 | F10, F12, F23 | F10, F12, F23 | F6, F9, F10 | F6, F9, F10 |
| 6 | F4 | F4 | F13 | F13 |
| 7 | F9 | F9 | | |

### *4.5.2.3 Clustering of Floribunda genotypes based on quantitative and qualitative characters*

Cluster analysis of Floribunda genotypes were performed for combined data which contains both quantitative and qualitative characters. Among the twelve characters number of leaves at first flower, number of days to first flower, prickle density (per five cm), flower size (cm), flower weight (g), pedicel length (cm), number of petals flower$^{-1}$, size of petals (cm) and number of flower plant$^{-1}$/bunch $^{-1}$ were taken as quantitative characters and fragrance, flower colour and vase life/ longevity were taken as qualitative characters. Clustering was performed under different clustering algorithms with Gower's measure as distance measures with statistical packages STATA and SAS. Dendrogram of clustering are shown in Figure 79, Figure 80, Figure 81and Figure 82. Cluster memberships of genotypes are given in Table 40. Optimum number of clusters Under UPGMA method was found to be five, seven, thirteen and seventeen from the plot of Pseudo $t^2$ statistics (Figure 83).

Seven clusters were formed in all the clustering methods. Under single linkage first cluster containing nineteen members formed the largest cluster. All the other clusters were single membership clusters. Clustering under UPGMA and WPGMA showed similar results.

### *4.5.2.4 Non Hierarchical clustering of Floribunda genotypes.*

Non hierarchical clustering of Floribunda genotypes were done by k- means clustering technique. It is a technique in which the numbers of clusters are predetermined. The optimum number of clusters was obtained from Pseudo F statistics. From the table of Pseudo F statistics (Table 41) optimum number of clusters was obtained as 10 with Pseudo F statistics value 10.47. Cluster memberships of genotypes in ten clusters are given in the Table 42.

Ten clusters were formed under k- means clustering. First and second clusters with four members formed the largest clusters followed by third, fourth, fifth and sixth clusters with three members and seventh cluster with two members. F21 (Lasting Piece), F19 (Sans

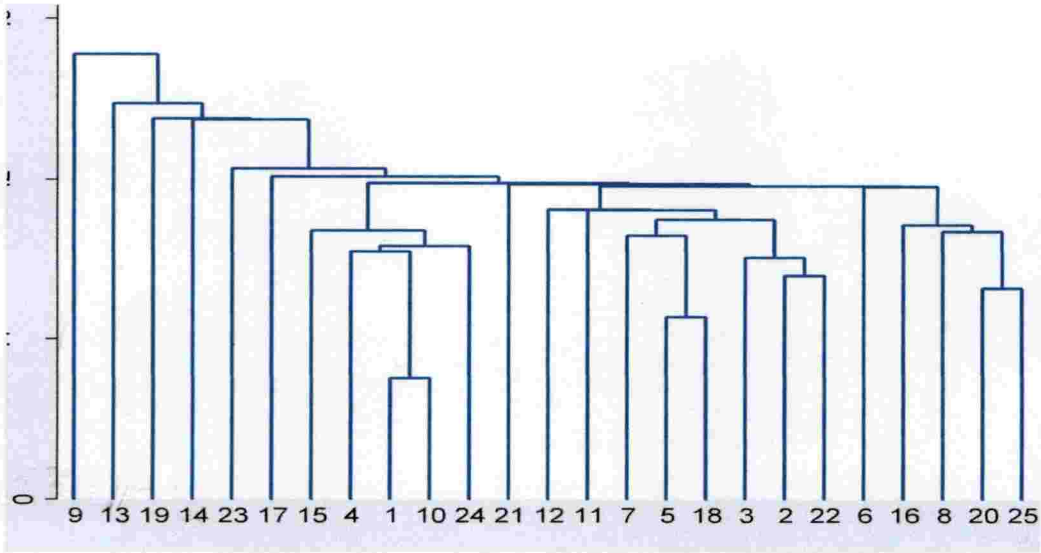


Fig. 79: Dendrogram of single linkage clustering of Floribunda genotypes based on Gower's measure

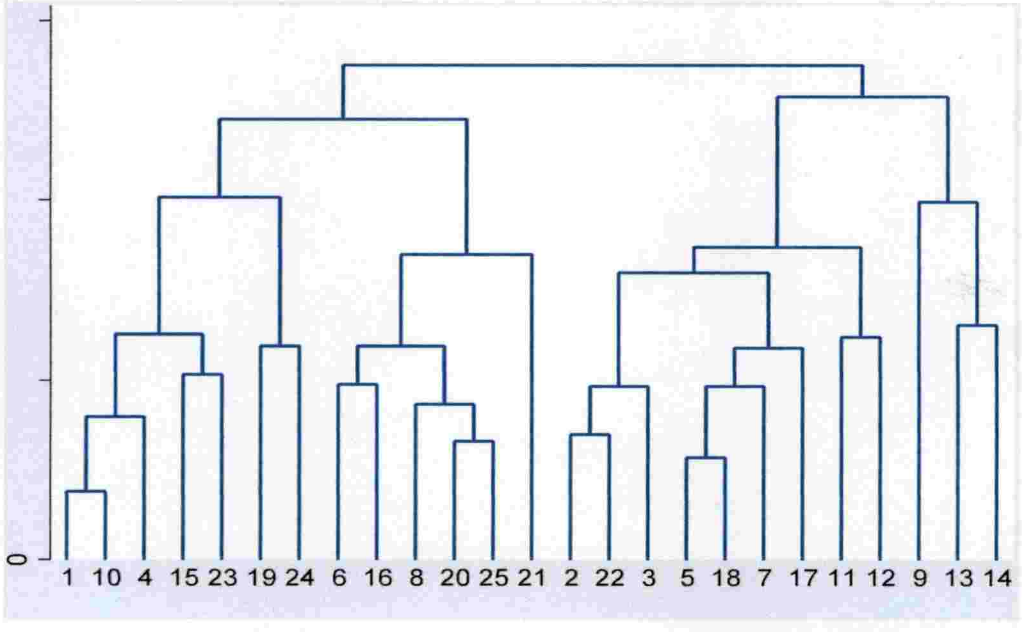


Fig. 80 Dendrogram of complete linkage clustering of Floribunda genotypes based on Gower's measure

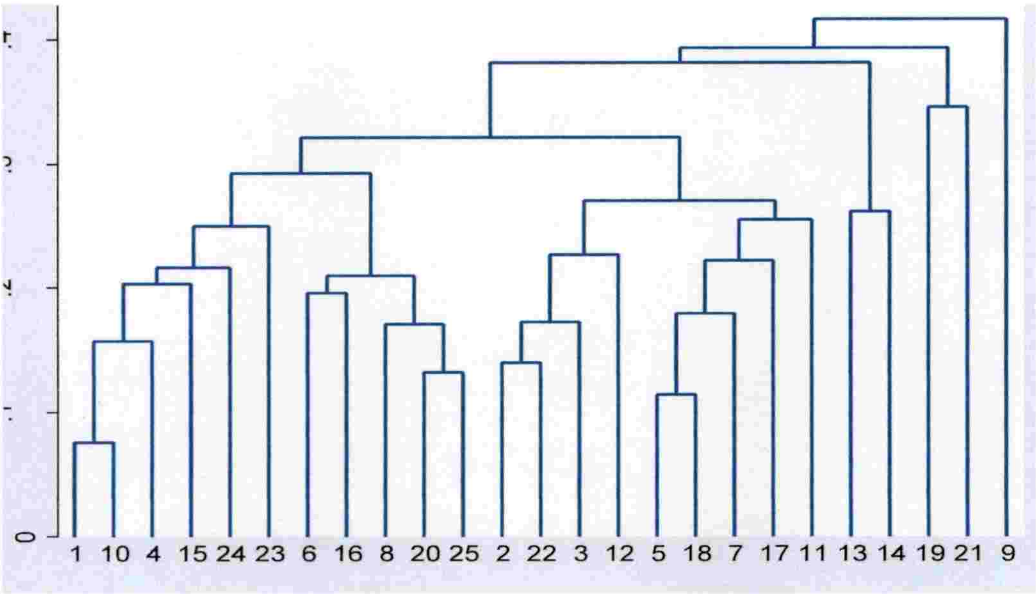


Fig. 81: Dendrogram of UPGMA clustering of Floribunda genotypes based on Gower's measure

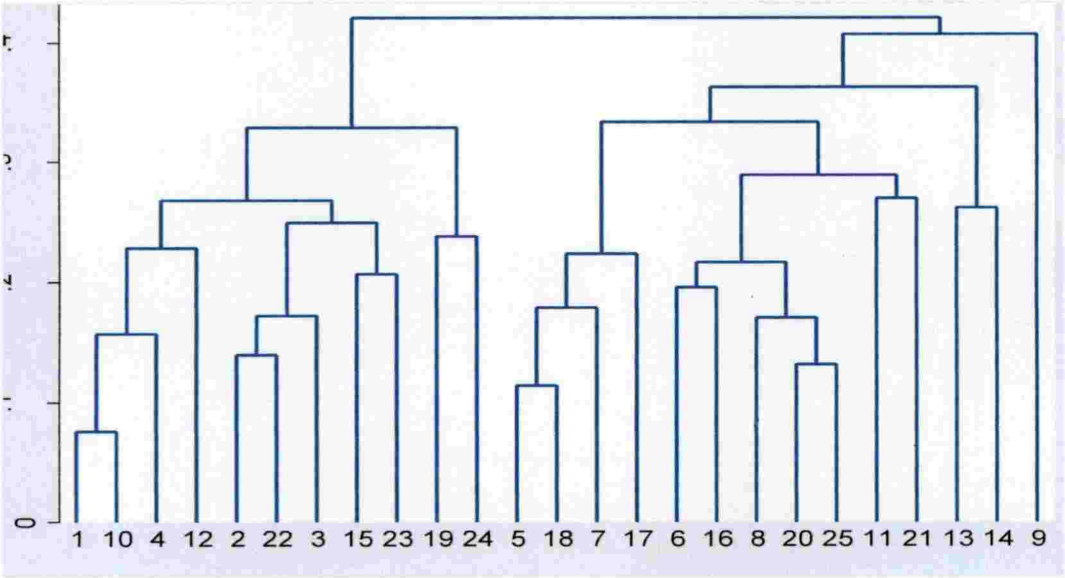


Fig. 82: Dendrogram of WPGMA clustering of Floribunda genotypes based on Gower's measure

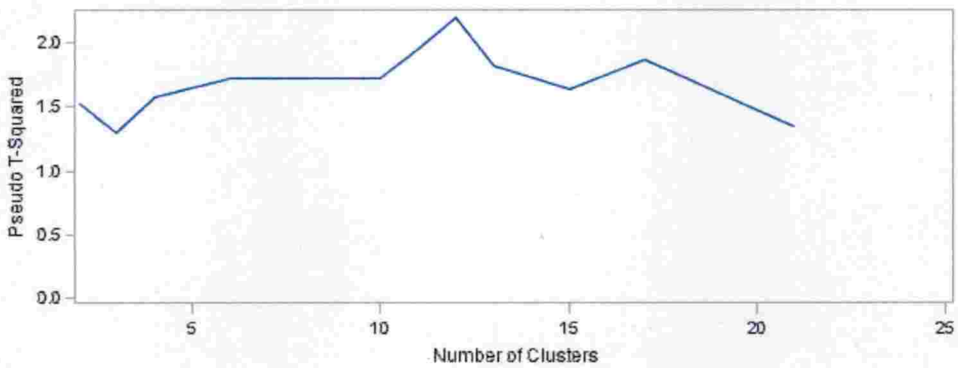


Fig.83: Pseudo $t^2$ statistic plot of Floribunda genotypes for Gower's measure under UPGMA

Table 40. Cluster membership of Floribunda type under Gower's measure.

| Cluster number | Single linkage | Complete linkage | UPGMA | WPGMA |
|---|---|---|---|---|
| 1 | F15, F1, F10, F24, F21, F12, F11, F7, F5, F18, F22, F2, F3, F4, F6, F8, F16, F20, F25 | F2, F3, F5, F7, F17, F18, F22 | F2, F3, F5, F7, F12, F11, F17, F18, F22 | F1, F10, F4, F12, F22, F3, F2, F15, F23 |
| 2 | F13 | F6, F8, F16, F20, F21, F25 | F1, F10, F4, F15, F24, F23 | F6, F8, F16, F20, F25 |
| 3 | F19 | F1, F10, F4, F15, F23 | F6, F8, F16, F20, F25 | F5, F18, F7, F17 |
| 4 | F14 | F19, F24 | F13, F14 | F19, F24 |
| 5 | F23 | F13, F14 | F19 | F11, F21 |
| 6 | F17 | F11, F12 | F21 | F13, F14 |
| 7 | F9 | F9 | F9 | F9 |

Table 41. Table of Pseudo F statistics for Floribunda genotypes under k- means clustering.

| No of clusters | Pseudo F Statistic |
|---|---|
| 5 | 10.18 |
| 6 | 9.76 |
| 7 | 8.59 |
| 8 | 9.84 |
| 9 | 8.97 |
| 10 | 10.47 |

Table 42. Cluster membership of Floribunda genotypes under k- means Clustering.

| Cluster no. | Members |
|---|---|
| 1 | F3, F9, F17, F23 |
| 2 | F4, F11, F12, F25 |
| 3 | F2, F5, F13 |
| 4 | F6, F18, F22 |
| 5 | F7, F8, F15 |
| 6 | F1, F10, F20 |
| 7 | F14, F16 |
| 8 | F21 |
| 9 | F19 |
| 10 | F24 |

Souci) and F24 (Golden Fairy) formed single member clusters.

## 4.6 PRINCIPAL COMPONENT ANALYSIS

Principal component analysis was carried out using the data corresponding to mean of qualitative characters under study. Principal components, percentage variance and cumulative variances are given in Table 43 and Table 45. The first 3 principal components accounted for 91.91% of total variance in case of Hybrid Tea genotypes, and 92.83% in case of Floribunda genotypes. For Hybrid Tea group number of days to first flower, number of leaves at first flower and number of petals/flower got highest loading in first, second and third component respectively (Table 44). In case of Floribunda genotypes, number of leaves at first flower, number of days to first flower and number of petals/flower got highest loading in first, second and third component respectively (Table 46). Score plot of first three principal components helped to find out natural clustering in each of the two sets of genotypes (Figure 84 and Figure 85).

Table 43. Principal component analysis of Hybrid Tea genotypes

| Principal component | Percentage variance | Cumulative variance |
|---|---|---|
| Component 1 | 50.19 | 50.19 |
| Component 2 | 29.012 | 79.21 |
| Component 3 | 12.71 | 91.91 |

Table 44. Principal component loadings of different characters of Hybrid Tea genotypes.

| Characters | PC 1 | PC 2 | PC 3 |
|---|---|---|---|
| No. of leaves at first flower | -0.028 | 0.983 | -0.123 |
| No. of days to first flower | 0.940 | -0.011 | -0.328 |
| Prickle density/5cm | 0.017 | 0.026 | -0.024 |
| Flower size (cm) | -0.146 | 0.076 | -0.335 |
| Flower weight (g) | 0.016 | 0.017 | 0.030 |
| Pedicel length (cm) | 0.006 | 0.009 | 0.015 |
| No. of petals/ flower | 0.286 | 0.163 | 0.872 |
| Size of petals (cm) | -0.110 | -0.013 | -0.059 |
| No. of flower per plant/bunch | 0.001 | 0.003 | -0.010 |

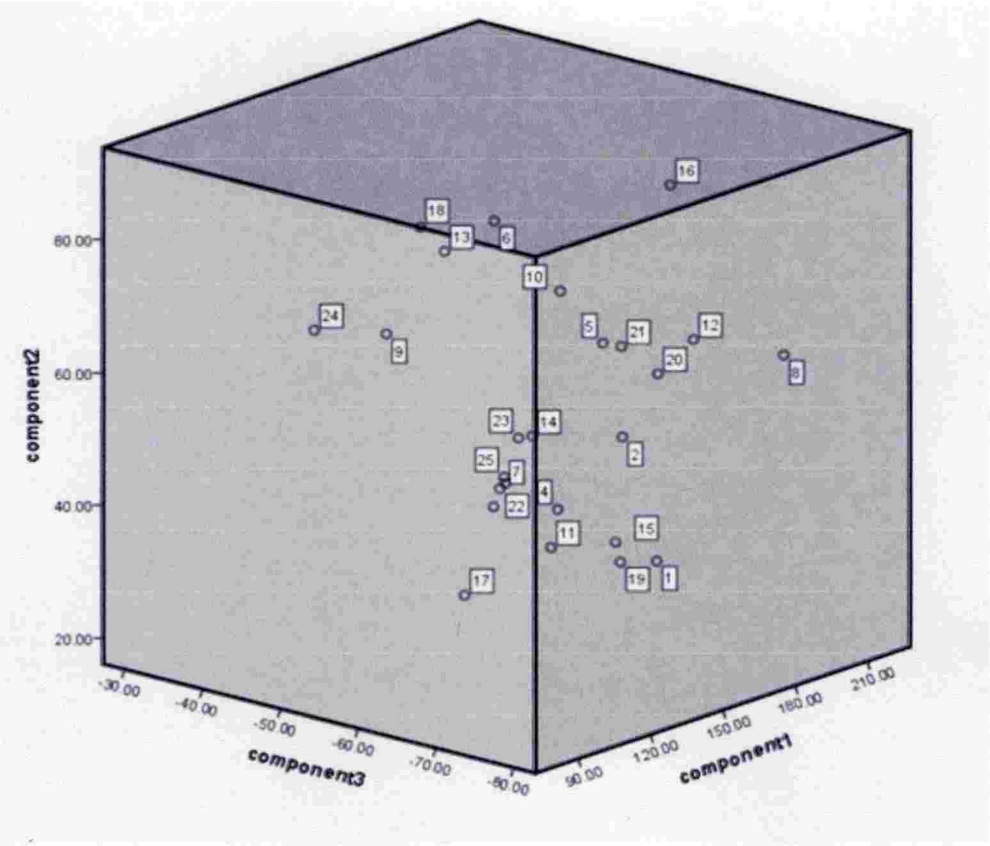


Fig. 84: Score plot of Principal components for Hybrid Tea genotypes

168

Table 45. Principal component analysis of Floribunda genotypes.

| Principal component | Percentage variance | Cumulative variance |
|---|---|---|
| Component 1 | 55.88 | 55.89 |
| Component 2 | 22.05 | 77.93 |
| Component 3 | 14.89 | 92.83 |

Table 46. Principal component loadings of different characters of Floribunda genotypes.

| Characters | PC 1 | PC 2 | PC 3 |
|---|---|---|---|
| No. of leaves at first flower | 0.995 | 0.066 | -0.020 |
| No. of days to first flower | -0.045 | 0.957 | 0.204 |
| Prickle density/5cm | -0.020 | 0.034 | -0.088 |
| Flower size (cm) | 0.079 | -0.137 | -0.204 |
| Flower weight (g) | 0.004 | -0.034 | 0.008 |
| Pedicel length (cm) | -0.004 | -0.010 | -0.005 |
| No. of petals/ flower | 0.046 | -0.238 | 0.945 |
| Size of petals (cm) | 0.003 | -0.046 | -0.123 |
| No. of flower per plant/bunch | 0.002 | 0.003 | -0.015 |

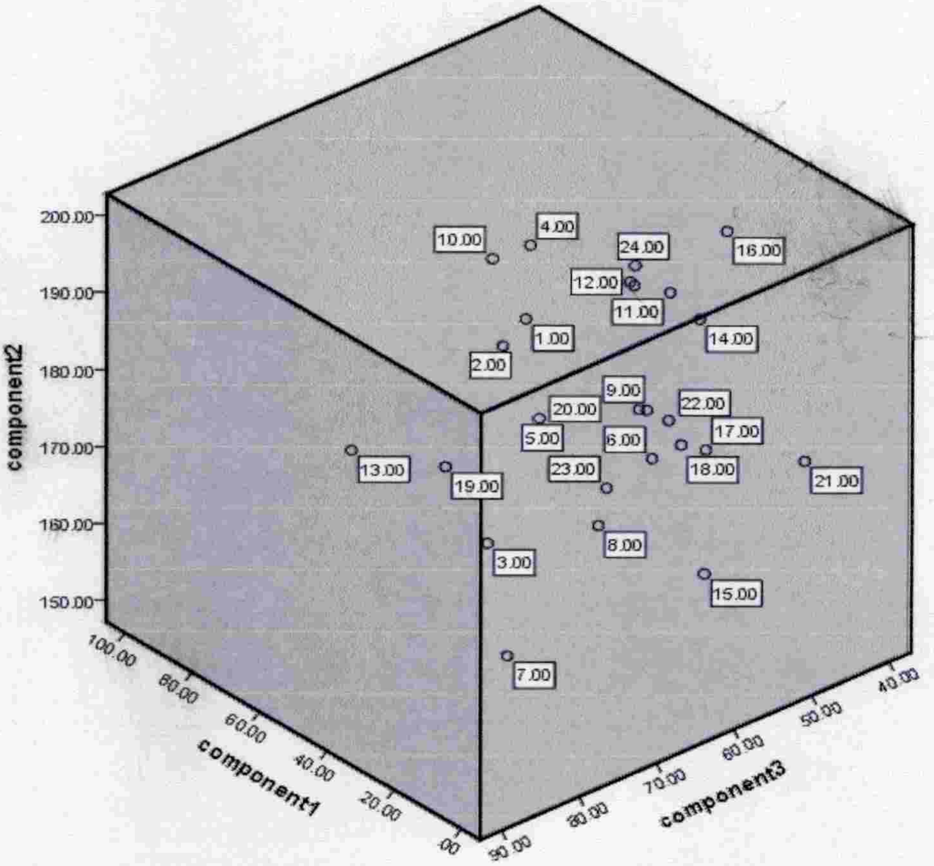


Fig. 85: Score plot of Principal components for Floribunda genotypes

## 4.7 COMPARISON AMONG DIFFERENT ASSOCIATION MEASURES

Different clustering patterns were obtained for different association measures when grouping was done with the same clustering algorithm. While considering the clustering result using different association measures it was revealed that clustering can be influenced by the choice of similarity and better understanding of different coefficient leads to efficient clustering. These results are in accordance with Jackson *et al.* (1989), Duarte *et al.* (1999) and Meyer *et al.* (2004). Comparison of different association measures are given in Table 47 and Table 48.

Among different association measures used for quantitative data, the result of different clustering techniques based on Squared Euclidean distance gave approximately the same result as that of Euclidean distance. There is only a small difference between clusters created by these two distance measures. It indicates that the calculated distances are highly correlated and showed few changes in genotype ranking (Dahal, 2015, Alves *et al.,* 2012). The Jaccard and Dice coefficients were found to be very similar, so that there was no difference in topology of dendrogram but only in branch length. Simple matching coefficient and Hamann's coefficient showed some distinct differences from these two. These results support what was reported in earlier studies stating the high correlation between the Dice and Jaccard measures (Nei and Li, 1979; Duarte *et al.*, 1999; Meyer *et al.*, 2004; Ojurongbe, 2012). Visual inspection of the dendrograms revealed a high level of similarity among those generated using the Dice and Jaccard measures. Simple matching and Hamann's coefficient also gave similar results.

## 4.8 COMPARISON OF CLUSTERING TECHNIQUES

Comparison among different clustering techniques based on quantitative character revealed that, among different methods studied single linkage clustering under different distance measures create a set of one or two clusters including the majority of the genotypes and the remaining genotypes were in single or two member clusters. Single linkage

171

Table 47. Comparison of different association measures for Hybrid Tea

| Distance | Minimum distance between | Maximum distance between |
|---|---|---|
| Squared Euclidean | (7,25), (21,5),(14,23) | (7,24), (24,25),(14,23) |
| Euclidean | (7,25),(21,5),(14,23) | (7,24),(24,25),(14,23 |
| Chebychev | (7,25),(21,5),(21,12) | (12,24),(24,8), (24,21) |
| City block | (7,25), (21,5), (4,3) | (7,24),(24,25),(24,13) |
| $D^2$ | (7,25), (6,13), (5,21) | (12,24),(24,7), (24,13) |
| Dice | (11,8), (19,15) | (1,3),(9,3),(20,22),(20,23),<br>(22,25),(23,25),(24,25) |
| Jaccard | (11,8), (19,15) | (1,3),(9,3),(20,22),(20,23),<br>(22,25),(23,25),(24,25) |
| Simple matching | (1,12),(1,20),(20,12) | (14,11), (17,14),(17,15) |
| Hamann's | (1,12),(1,20),(20,12) | (14,11), (17,14),(17,15) |
| Gower's measure | (18,5) | (10,4) |

Table 48. Comparison of different association measures for Floribunda

| Distance | Minimum distance between | Maximum distance between |
|---|---|---|
| Squared Euclidean | (11,25),(22,18), (17,9) | (6,19),(19,21),(19,16) |
| Euclidean | (11,25),(22,18),(2,5) | (6,19),(19,21),(19,16) |
| Chebychev | (11,25),(17,9),(22,18) | (6,19),(19,21),(19,18) |
| City block | (11,25),(22,18),(25,12) | (16,19),(19,6),(19,12) |
| $D^2$ | (1,10),(18,22),(11,25) | (6,19),(19,16),(,21,10) |
| Dice | (5,17) | (1,9), (9,4),(13,4),(14,4),(22,14) |
| Jaccard | (5,17) | (1,9), (9,4),(13,4),(14,4),(22,14) |
| Simple matching | (1,2),(20,5), (8,16) | (11, 4) |
| Hamann's | (1,2),(20,5), (8,16) | (11, 4) |
| Gower's measure | (1,10) | (21,13) |

clustering tends to produce long chain type clusters as opposed to bunch clusters has been reported Kupier and Fisher, 1975. In other terms, the single linkage algorithm suffers chaining effect. Since the chaining effect is a common problem with respect to clustering quality, avoiding it implies that high clustering quality is achieved (Oyang *et al.*, 2002).

Among other clustering algorithms complete linkage method and Ward's clustering method showed similar results under Squared Euclidean distance. UPGMA, WPGMA and UPGMC methods under Squared Euclidean method gave comparable results (Kupier and Fisher, 1975). Clustering using UPGMA and WPGMA methods gave almost the same clustering pattern under different distance measures. SD indices were calculated for clustering based on quantitative data. SD index of single linkage, complete linkage and UPGMA method under different distance measures were evaluated for both groups separately. Under different distance measures single linkage have minimum SD index, but it was suffering from chaining effect. Therefore UPGMA method under Squared Euclidean was found to be the best with SD index 0.651 for Hybrid Tea (Table 49) and 0.689for Floribunda group (Table 50).

Results obtained from k- means clustering are comparable with the results obtained from hierarchical clustering except for single linkage clustering. There is some similarity between k-means and $D^2$ analysis but not to up that of other clustering methods.

Comparison of clustering technique for qualitative data revealed that here also single linkage clustering produced long chain dendrograms with lots of singletons and small clusters under different distance measures. The same result was observed by Stuetzle and Nuggent, 2007. Some level of closeness was observed with dendrograms produced using the UPGMA and WPGMA. However, the dendrograms constructed using the single linkage was quite different. These findings are in line with the report of Ojurongbe, 2012.

Under Hybrid Tea genotypes, H16 (Mary Jean) formed a single cluster under single linkage method using different distance measures for quantitative, qualitative and mixed data analysis. Under complete linkage method H7 (Alaine Souchen) and H25 (Josepha) came

Table 49. SD index of clustering for Hybrid Tea genotypes.

| Clustering method | SD Index |
|---|---|
| Single linkage - Squared Euclidean | 0.587 |
| UPGMA- Squared Euclidean | 0.651 |
| Single linkage - Chebychev | 0.684 |
| Single linkage - City Block | 0.722 |
| UPGMC - Squared Euclidean | 0.784 |
| UPGMA- Chebychev | 0.784 |
| Ward's method - Squared Euclidean | 0.79 |
| Complete linkage - Squared Euclidean | 0.803 |
| Complete linkage - Chebychev | 0.881 |
| UPGMA - City Block | 1.38 |
| Complete linkage - City Block | 1.41 |

Table 50. SD index of clustering for Floribunda genotypes.

| Clustering method | SD Index |
|---|---|
| Single linkage - Chebychev | 0.659 |
| UPGMA- Squared Euclidean | 0.689 |
| UPGMC - Squared Euclidean | 0.689 |
| Single linkage - City Block | 0.821 |
| Complete linkage - Squared Euclidean | 0.886 |
| Single linkage - Squared Euclidean | 0.903 |
| Ward's method - Squared Euclidean | 1.02 |
| Complete linkage - Chebychev | 1.13 |
| UPGMA- Chebychev | 1.34 |
| Complete linkage - City Block | 1.5 |
| UPGMA - City Block | 1.9 |

under the same cluster, in clustering based on quantitative and qualitative characters. H22 (Mom's Rose) and H23 (Lois Wilson) came under the same cluster in clustering based on complete linkage, UPGMA and WPGMA except under Hamann's coefficient. These came under the same cluster under $D^2$ analysis also. Among Floribunda genotypes, F2 (Tickled Pink) and F5 (Princess de Monaco) were included in the same cluster under UPGMA method for both quantitative and qualitative data. F1 (Versailles) and F24 (Golden Fairy) also came under the same cluster except for multistage distances under UPGMA. UPGMA under Jaccard or Dice coefficient gave better clustering.

Clustering based on mixed data gave approximately the same results as that of quantitative data under different clustering algorithms except for single linkage clustering.

Clustering pattern observed from score plot of PCA is comparable with clusters obtained from quantitative data especially with $D^2$ analysis. Contribution of characters towards variance obtained $D^2$ analysis and PCA showed similar results.

From the study it is possible to compare different methods and exclude inappropriate methods. Groups formed from modified Tocher method and PCA are different from other methods. SD index indicated that UPGMA under Squared Euclidean distance is the best for quantitative data.

# *Summary*

# 5. SUMMARY

The present study entitled 'Multivariate clustering techniques – a comparison based on rose (*Rosa* spp.)' was undertaken to compare different clustering techniques, to identify the suitable technique for different types of qualitative and quantitative data and to illustrate the procedures using data based on a field experiment on rose (*Rosa* spp.). Data on quantitative and qualitative traits collected from a field experiment on "Characterization and genetic improvement in Rose (*Rosa* spp.) through mutagenesis" done during 2014-2017 at College of Agriculture, Vellayani and Regional Agriculture Research Station (RARS), Ambalavayal, Wayanad was used for the study. Twenty five cultivars each coming under the Hybrid Tea and Floribunda groups of rose were evaluated for the study. The varieties were grown in completely randomized design with six replications. There were nine quantitative characters and three qualitative characters. Statistical studies were carried out with the help of statistical packages STATA, SPSS, SAS, R and NTSYS.

Analysis of variance was done for each of the quantitative characters under study which showed significant difference among different genotypes with respect to each character. The Wilk's lamda value obtained from multivariate analysis of variance was 0.004 for Hybrid Tea group and 0.003 for Floribunda group. The corresponding $V_{(stat)}$ is 728.833 and 766.807 respectively which is distributed as chi-square with 216 degrees of freedom and indicated significant at one percent level. The results showed that there was significant difference between the varietal means with respect to all characters under study.

Linear discriminant analysis revealed that the characters flower size, pedicel length had highest contribution to discriminate between groups with coefficients -1.068 and -0.933 respectively. The average value obtained for the Hybrid Tea was 11.09 and – 2.34 for Floribunda type with an overall average of 4.38. 80% of Hybrid

Tea genotypes had discriminant score above 4.38 and 72% Floribunda group had discriminant score below 4.38. Thus the discriminant function analysis reassured the difference between two groups under study.

Cluster analysis was carried out separately for Hybrid Tea type and Floribunda type. Cluster analyses were performed for quantitative, qualitative and mixed data. Association measures used were Euclidean distance, Squared Euclidean, Chebychev distance, City Block distance and Mahalanobis $D^2$ for quantitative data, Jaccard, Dice, Simple matching and Hamann's coefficient for qualitative data and Gower's measure for mixed data. Different methods such as single linkage, complete linkage, Unweighted Pair Group Average Method (UPGMA), Weighted Pair Group Average Method (WPGMA), Unweighted Pair Group Centroid Method (UPGMC), Ward's method, modified Tocher method, k-means clustering and Principal Component Analysis (PCA) were adopted for the clustering of cultivars. UPGMC and ward's method were performed only using Squared Euclidean distance as these methods gave valid results only for that measure. $D^2$ statistics were used for modified Tocher method. Optimum numbers of clusters were determined by Pseudo $t^2$ statistics for hierarchical clustering and by Pesudo F statistics for k-means clustering. SD (Scatterness- Distance) index was used to test validity of clustering based on quantitative data.

Comparison among different association measures showed that Jaccard and Dice coefficients gave similar results and it may due to fact that both are not considering the negative matches. Simple matching gave valid results in case Floribunda group but not for Hybrid Tea type. In case of Floribunda group, Simple matching and Hamann's coefficient gave same results. Both these measures are different from Jaccard and Dice coefficient. Among association measures for quantitative data Euclidean and Squared Euclidean gave approximately same results.

For Hybrid Tea genotypes among all the clustering methods, single linkage clustering under different distance measures tends to create a set of one or two clusters including majority of the genotypes and the remaining genotypes remains singletons and single linkage algorithm suffers chaining effect for qualitative and quantitative data. Ward's and complete linkage method showed similar clustering pattern under Squared Euclidean distance. Under Squared Euclidean distance UPGMA, WPGMA and UPGMC were showed similar clustering. UPGMA and WPGMA gave approximately same clustering results. k – means clustering also have almost similar clustering pattern as that of other methods based on quantitative data except for modified Tocher method. Clustering based on $D^2$ gave different results from other methods. Clustering based on quantitative data are different from qualitative data while clustering based on quantitative and mixed data have some sort of similarity except for single linkage method. Under Hybrid Tea genotypes, H16 (Mary Jean) formed a single cluster under single linkage method using different distance measures for quantitative, qualitative and mixed data analysis. Under complete linkage method H7 (Alaine Souchen) and H25 (Josepha) came under same cluster, in clustering based on quantitative and qualitative characters. H22 (Mom's Rose) and H23 (Lois Wilson) came under same cluster in clustering based on complete linkage, UPGMA and WPGMA except under Hamann's coefficient. These came under the same cluster under $D^2$ analysis also.

For Floribunda genotypes also single linkage clustering suffering from chaining effect for different types of data. Here also UPGMA and WPGMA gave similar clustering under different measures. Clustering based on $D^2$ analysis showed some variation from other methods. Similar to Hybrid Tea groups Ward's and complete linkage found to be similar and UPGMA, WPGMA and UPGMC under Squared Euclidean also found to be similar. Clustering based on k- means clustering gave approximately similar result as that of hierarchical methods. Clustering with quantitative and qualitative data gave different result while clustering with quantitative and mixed data showed some similarity. Comparison using SD index indicated high index value for clustering based on Gower's measure.

Comparison among single linkage, complete linkage and average linkage under different association measures using SD index were carried out. Average linkage method under Squared Euclidean was found to be the best for both type with SD index 0.651 for Hybrid Tea and 0.659 for Floribunda type.

Clustering pattern observed from score plot of PCA is comparable with the pattern obtained from quantitative data especially with $D^2$ analysis. Contribution of characters towards divergence was studied using $D^2$ statistics. In case of Hybrid Tea groups it was found that number of days to first flower has highest contribution followed by number of leaves at first flower and number of petals/ flower. In case of Floribunda group number of leaves at first flower has highest contribution towards divergence followed by number of days to first flower and number of petals / flower. Similar trend was observed under PCA also.

# *References*

## 6. REFERENCES

Akbar, F., Rabbani, M. A., Shinwari, Z. K., and Khan, S. J. 2011. Genetic divergence in sesame (*Sesamum indicum* L.) landraces based on qualitative and quantitative traits, *Pak. J. Bot.*, 43(6): 2737-2744.

Ali, Y., Atta, B.M., Akhter, J., Monneveux, P., and Lateef, Z. 2008. Genetic variability, association and diversity studies in wheat (*Triticum aestivum* L.) germplasm. *Pak. J. Bot.* 40:2087-2097.

Alves, E. O. S., Silva, C. B. M. C., Souza, A. M., Santos, C. A. F., Neto, F. P. L., and Correa, R. X. 2012. Comparison of efficiency of distance measurement methodologies in mango (*Mangifera indica*) progenies based on physicochemical descriptors. *Genet. Mol. Res.* 11 (1): 591-596.

Anderson, T. W. 2004. *An Introduction to Multivariate Analysis.*( 3rd Ed.). John Wiley & Sons, Inc, New York, 721p.

Aswathi, C., Devadas, V. S., Francies, R. M., and Bastian, D. 2015. Genetic divergence in cowpea (*Vigna* spp.) varieties for seed quality. *J. Trop. Agric.* 53: 197-199.

Bharathi, A., Punitha, A., and Devi, S. 2012. Studies on genetic divergence on cucumber (*Cucumber sativum* L.). *Asian J. Biosci.* 7:169-173.

Blashfield, R. K. 1976. Mixture model tests of cluster analysis : accuracy of four agglomerative hierarchical methods. *Psychol. Bull.* 83:377-388.

Calinski, T. and Harabasz, J. 1974. A Dendrite Method for Cluster Analysis: Communications in Statistics. Theory and Methods, 3: 1-27.

Chatfield, C. and Collins, A. J. 1980. *Introduction to Multivariate Analysis* (2$^{nd}$ Ed). Chapman and Hall, USA, 246p.

Cunningham, K. M. and Ogilvie, J. C. 1972. Evaluation of hierarchical grouping techniques: A preliminary study. *Comput. J.* 15:209–213.

Dahal S. 2015. Effect of different distance measures in result of cluster analysis, M.Sc. thesis, Aalto University School of Engineering, Finland, 77p.

Dalirsefat, S., Meyer, A. and Mirhoseini, S. (2009). Comparison of similarity coefficients used for cluster analysis with amplified fragment length polymorphism markers in the silkworm *Bombyx mori*. *J. Insect Sci.* 9 (71) 1-8.

Dash, S. 2008. Some investigations on classificatory techniques in agriculture. M.Sc. (Ag) thesis, IARI, New Delhi, 158p.

Dice, L.R. 1945. Measures of the amount of ecologic association between species. *Ecology 26*: 297-302.

Ferreira, L. and Hitchcock, D.B. 2009. A comparison of hierarchical methods for clustering functional data. *Commun. Stat. Simul. Comput.* 38:1925- 1949.

Gower, J. C. 1971. A general coefficients of similarity and some of its properties. *Biometrics*. 27: 857-872.

Gower, J. C. 1985. Measures of similarity, dissimilarity and distance. In: S. Kotz and N. L. Johnson (eds), *Encyclopedia of Statistical Sciences,* Wiley, New York, USA, pp. 397-405.

Hair J. F., Black W. C., Babin B. J., Anderson R. E. 2015. *Multivariate Data Analysis* (7th Ed.), Pearson Education Limited, New York, 734p.

Halkidi, M., Batistakis, Y., and Vazirgiannis, M. 2002. Cluster validity methods: part II, *SIGMOD Rec.*, 31 (3): 19-27.

Hands, S. and Everitt, B. 1987. A Monte Carlo study of the recovery of cluster structure in binary data by hierarchical clustering techniques. *Multivar. Behav. Res.* 22:235-243.

Hardle, W. and Simar, L. 2007. *Applied Multivariate Statistical Analysis* (2nd Ed). Springer International Edition, Germany, 352p.

Hartigan, J. A. 1975. *Clustering algorithms* (2nd Ed.). John Wiley & Sons, Inc., Canada, 420p.

Jaccard, P. 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull. Soc. Vaudoise Sci. Nat. 37*: 547-579.

Jackson, A.A., Somers, K.M. and Harvey, H.H. (1989). Similarity coefficients: measures for co-occurrence and association or simply measures of co-occurrence? *Am Nat* .133: 436-453.

Johnson, R. A. and Wichern, D. W. 2006. *Applied Multivariate Statistical Analysis* (5th Ed.). Inc. Pearson Prentice Hall, London, 455p.

Kosman, E. and Leonard, K.J. (2005). Similarity coefficients for molecular markers in studies of genetic relationships between individuals for haploid, diploid and polyploidy species. *Mol. Ecol.* 14: 415 - 424.

Kuiper, F. K. and Fisher, L.A. 1975. A Monte Carlo comparison of six clustering procedures. *Biometrics* 31:777-783.

Kuswardhani, N., Soni, P., and Shivakoti, G. P. 2014. Cluster analysis for classification of farm households based on socio-economic characteristics for technology adoption in agriculture: A case study of West Java province, Indonesia. *J. Food. Agric. Environ.* 12: 238-247.

Legany, C., Juhasz, S., and Babos, A. 2006. Cluster validity measurement techniques. *Proceedings of the 5th WAEAS International; Conference on Artificial Intelligence, Knowledge Engineering and Data Bases,* 15-17 February 2006, Madrid, Spain, pp. 388-393.

Lombard, V., Baril, C. P., Dubreuil, P., Blouet, F., and Zhang, D. 2000. Genetic relationships and fingerprinting of rapeseed cultivars by AFLP: Consequences for varietal registration. *Crop Sci.* 40:1417–1425.

Longkumer, M. and Kabir, J. 2014. Genetic divergence of mandarin genotypes in Nagaland. *Indian J. Genet.* 74: 115-118.

MacQueen, J. B. 1967. Some methods for classification and analysis of multivariate observations, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley, University of California Press, 1:281-297.

Mahalanobis, P. C. 1936. On the Generalized distance in Statistics. *Proc. Nat. Inst. Sci. India.,* 2: 49-55.

Mahesh, M., Sudhanshu, J., and Bhadoria, V.S. 2017. Genetic divergence studies in soybean (*Glycine max* L. Merrill). *Int. J. Agric. Sci.* 9(19): 4196-4197.

Meyer, S. A., Garcia, A. A. F., Souza, A. P., and Souza, C. L. 1999. Comparison of similarity coefficients used for cluster analysis with dominant markers in maize (*Zea mays* L). *Genet. Mol. Biol.* 27(1): 83-91.

Milligan, G. W. 1980. An examination of the effect of six types of error perturbation on fifteen clustering algorithms. *Psychometrika* 45:325-342.

Milligan, G. W. 1996. Clustering validation – Results and implications for applied analyses. In: P. Arabie, L. J. Hubert & G. De Soete (eds.), *Clustering and Classification*. World Scientific Publ. Co., River Edge, New Jersey, pp. 341-375.

Nei, M., and Li, W. H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA*. 76: 5269-5273.

Ojurongbe, T. A. 2012. Comparison of different proximity measures and classification methods for binary data. Ph.D. thesis, Justus Liebig University Giesen, Hesse, Germany, 118p.

Parsad, R. 2007. *Multivariate techniques: an overview* [Lecture notes]. IASRI, New Delhi, 60p.

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

Priya, S., Suneetha, Y., Babu, D.R., and Rao, V.S. 2017. Assessment of genetic divergence in rice for grain yield and quality traits. *Environ. Ecol.* 35: 2007-2011.

Rangaswamy, R.1995. *A Textbook of Agricultural Statistics* (2nd Ed.). New Age International (p) Ltd. 524p.

Rao, C. R. 1952. *Advanced Statistical Methods in Biometric Research,* John Wiley and Sons, Inc. New York, 390p.

Rencher, A. C. 2002. *Methods of Multivariate Analysis*. (2nd Ed.). John Wiley & Sons, Inc, New York, 707p.

188

Rincon, F., Johnson, B., Crossa, J., and Taba, S. 1996. Cluster analysis an approach to sampling variability in maize accessions. *Maydica* 41:307–316.

Rogers, D. J. and Tanimoto, T. T. 1960. A computer program for classifying plants. *Sci.132*: 1115-1118.

Sajeevkumar, S. S., Minimol, J. S., Ajmal, P. M., and Suma, B. 2017. Diversity analysis of KAU released cocoa (*Theobroma cacao* L.) varieties based on morphological parameters. *J. Trop. Agric.* 55: 81-86.

Sharma, H.R., Sharma, D., and Thakur, A.K. 2006. Genetic divergence in mustard (*Brassica* spp. L). *J. Hort. Sci.* 1:52-54.

Shathi, D.A., Arifuzzaman, M., Biswas, B.K., Hasauzzaman, M., and Azad, A.K. 2012. Genetic divergence in mustard (*Brassica* spp. L). *Afr. J. Plant Sci.* 6: 239-243.

Sheela, V. L., Lekshmi, P. R. G., Nair, C. S. J., and Rajmohan, K. 2006. Molecular characterization of *Heliconia* by RAPD assay. *J. Trop. Agric.* 44: 37-41.

Shirkhorshidi, A. S., Aghabozorgi, S., and Wah, T. Y. 2015. A comparison study on similarity and dissimilarity measures in clustering continuous data. *PLoS ONE*. 10(12): e0144059. doi:10.1371/ journal.pone.0144059.

Singh, R. K. and Chaudhary, B. D. 1977. *Biometrical Methods in Quantitative Genetic Analysis* (2nd Ed.). Kalyani publishers, New Delhi, 318p.

Singh, Y., Pani, D.R., Pradhan, S.K., Bajpai, A., and Singh, U.S. 2008. Divergence analysis for quality traits in some indigenous Basmati rice genotypes. *Oryza* 45: 263-267.

Sneath, P. H. A. and Sokal, R. R.1973. *Numerical Taxonomy*, WH Freeman, San Francisco. 573p.

Sokal, R.R. and Michener, C.D.1958. A statistical method for evaluating systematic relationships. *Univ. Kans. Sci. Bull. 38*: 1409-1438.

Sorensen, T. 1948. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *K. Dan. Vidensk. Selsk. Biol. Skr. 5*: 1-34.

Sunayana., Yadav, R., Punia, M. S., and Ravika. 2017. Genetic divergence studies in mungbean (*Vigna radiata* L. Wilczek) using morpho-physio and molecular markers to identify drought tolerant genotypes. *Indian. J. Genet.* 77: 574-578.

Suresh, K. M. 1986. Standardisation of techniques of clustering genotypes using Mahalanobis $D^2$ and Wilks' Λ criterion. M.Sc. (Ag) thesis, Kerala Agricultural University, Thrissur, 162p.

Tarpey, T. 2007. Linear transformations and the k-means clustering algorithm. *Am. Stat.* 61:34–40.

Theodoridis, S., and Koutroumbas, K. 2003. *Pattern Recognition* (2nd Ed). Academic Press., USA, 984p.

Toma, E., Balan, A. V., Dobre, E., and Soare, E. 2015. Organic farming patterns analysis based on clustering methods. *Agric. Agric. Sci. Proc.* 6: 639-646.

Toma, E., Balan, A. V., Dobre, E., and Soare, E. 2015. Organic farming patterns analysis based on clustering methods. *Agric. Agric. Sci. Proc.* 6: 639-646.

190

Vasanthi, R.P., Suneetha, N., and Sudhakar, P. 2014. Genetic diversity based on physiological attributes among released and pre-release cultures of groundnut (*Arachis hypogaea* L.). *Legume Res.* 38:47-50.

Ward, J. H. 1963. Hierarchical grouping to optimize an objective function. *J. Am. Statist. Assoc.* 58: 69-78.

Wilks, S.S. 1932. Certain generalizations in the analysis of variance. *Biometrrika.* 24:471.

# *Appendices*

# APPENDIX – I

## Euclidean distance between Hybrid Tea genotypes

| Genotypes | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 | H17 | H18 | H19 | H20 | H21 | H22 | H23 | H24 | H25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H1 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H2 | 3.68642 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H3 | 3.45469 | 4.61492 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H4 | 3.17230 | 5.06268 | 1.66900 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H5 | 3.36315 | 4.94572 | 2.76423 | 1.71719 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H6 | 5.35357 | 4.87645 | 4.18840 | 4.24417 | 3.78794 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H7 | 3.37256 | 4.02231 | 2.46296 | 2.46887 | 2.79860 | 2.91370 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H8 | 5.14905 | 4.94666 | 4.89504 | 4.57829 | 4.45729 | 3.93213 | 4.70404 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H9 | 3.64490 | 4.22547 | 2.78117 | 3.08470 | 3.02811 | 3.63204 | 3.08866 | 4.76700 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H10 | 5.26715 | 6.02836 | 3.40398 | 3.13035 | 3.11146 | 3.76995 | 4.14004 | 3.66642 | 3.45020 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H11 | 3.13737 | 3.49747 | 3.80070 | 3.29874 | 3.51997 | 5.11160 | 3.57916 | 4.51670 | 3.80428 | 4.69313 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H12 | 2.98150 | 4.42107 | 3.31366 | 2.56212 | 1.40603 | 4.22552 | 3.16933 | 4.69323 | 3.22579 | 4.05525 | 3.60252 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H13 | 5.18154 | 5.04957 | 4.04733 | 3.71361 | 3.12309 | 1.92218 | 2.69622 | 4.43981 | 3.86106 | 3.71089 | 4.34553 | 3.87984 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| H14 | 4.21502 | 6.41666 | 6.13748 | 5.43094 | 5.19754 | 6.53564 | 5.80011 | 6.21187 | 4.55362 | 6.19299 | 5.35298 | 5.01391 | 6.64860 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . |
| H15 | 3.66819 | 4.89757 | 4.29370 | 4.34188 | 4.75928 | 6.06403 | 4.88776 | 5.22138 | 3.75858 | 4.53258 | 3.95826 | 4.82838 | 6.10682 | 4.61256 | 0.00000 | . | . | . | . | . | . | . | . | . | . |
| H16 | 4.72754 | 3.73557 | 4.84563 | 5.36072 | 4.79539 | 3.96255 | 4.55384 | 4.78944 | 3.56387 | 5.20437 | 5.62784 | 4.22254 | 5.13246 | 5.96535 | 5.22096 | 0.00000 | . | . | . | . | . | . | . | . | . |
| H17 | 3.57719 | 4.41342 | 2.53733 | 2.33391 | 3.27878 | 5.11872 | 3.26042 | 5.07402 | 3.24927 | 4.11447 | 2.12592 | 3.61320 | 4.49285 | 5.81912 | 4.02572 | 5.79057 | 0.00000 | . | . | . | . | . | . | . | . |
| H18 | 4.17091 | 4.70492 | 3.04453 | 2.83756 | 1.90246 | 3.51467 | 3.18369 | 4.98219 | 2.47299 | 3.41756 | 3.82723 | 2.46955 | 2.82603 | 5.75865 | 6.24718 | 4.50176 | 3.52838 | 0.00000 | . | . | . | . | . | . | . |
| H19 | 3.87054 | 4.46265 | 3.51784 | 3.17262 | 3.66654 | 4.44314 | 3.23681 | 4.38374 | 3.22217 | 4.24158 | 3.39949 | 3.49938 | 4.63211 | 5.07294 | 4.11983 | 4.72240 | 2.78959 | 4.30854 | 0.00000 | . | . | . | . | . | . |
| H20 | 2.96405 | 4.26983 | 3.55970 | 2.98504 | 2.59708 | 4.54691 | 3.84957 | 3.74297 | 2.70231 | 3.20154 | 2.91781 | 2.59599 | 4.39620 | 3.97972 | 2.74771 | 4.20088 | 3.29822 | 3.25223 | 3.23200 | 0.00000 | . | . | . | . | . |
| H21 | 3.62690 | 3.60335 | 2.86157 | 3.23881 | 2.98598 | 2.78979 | 2.28205 | 4.31525 | 2.54641 | 3.92603 | 4.29153 | 2.70998 | 3.51127 | 5.69285 | 4.65597 | 2.56192 | 3.97337 | 3.12712 | 3.05932 | 3.38781 | 0.00000 | . | . | . | . |
| H22 | 4.00927 | 3.60303 | 3.69968 | 3.93923 | 4.33623 | 4.90230 | 4.13912 | 3.85479 | 4.19337 | 4.59870 | 2.85949 | 4.68447 | 4.56785 | 6.48761 | 4.78834 | 5.46376 | 3.44560 | 4.17442 | 4.79169 | 3.93038 | 4.61467 | 0.00000 | . | . | . |
| H23 | 2.74905 | 3.00685 | 3.80292 | 3.61308 | 3.52566 | 4.86490 | 3.84568 | 4.21026 | 3.28362 | 4.74781 | 1.94894 | 3.44690 | 4.48312 | 4.81624 | 4.11189 | 4.71628 | 3.07342 | 3.42570 | 4.01506 | 2.73516 | 4.02490 | 2.21648 | 0.00000 | . | . |
| H24 | 6.17568 | 5.14258 | 5.86082 | 6.26559 | 6.34685 | 6.92673 | 6.76462 | 5.92573 | 4.88938 | 5.93111 | 4.73564 | 6.43849 | 6.80492 | 7.11390 | 5.37812 | 6.24101 | 4.99322 | 5.46980 | 6.14991 | 5.00628 | 6.38505 | 3.96039 | 3.94092 | 0.00000 | . |
| H25 | 3.43638 | 4.05610 | 2.22600 | 2.35758 | 2.72503 | 2.99140 | 0.45953 | 4.75995 | 3.17445 | 4.09085 | 3.63623 | 3.15615 | 2.69040 | 6.07497 | 5.01474 | 4.62924 | 3.23010 | 3.05497 | 3.47049 | 3.94371 | 2.34132 | 4.00950 | 3.85572 | 6.72033 | 0 |

# APPENDIX – II

## Squared Euclidean distance between Hybrid Tea genotypes

| Genotypes | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 | H17 | H18 | H19 | H20 | H21 | H22 | H23 | H24 | H25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H1 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H2 | 13.6897 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H3 | 11.9349 | 21.2975 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H4 | 10.0635 | 25.6307 | 2.7856 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H5 | 11.3108 | 24.4601 | 7.6410 | 2.9487 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H6 | 28.6607 | 23.7993 | 17.5427 | 18.0130 | 14.3485 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H7 | 11.3742 | 16.1789 | 8.0662 | 6.0953 | 7.8322 | 8.4896 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H8 | 26.5127 | 24.4694 | 23.9614 | 20.9608 | 19.8675 | 15.4617 | 22.1280 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H9 | 13.2853 | 17.8546 | 7.7349 | 9.5153 | 9.1694 | 13.1917 | 9.5398 | 22.7243 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H10 | 27.7429 | 36.3412 | 11.5871 | 9.7991 | 9.6812 | 14.2126 | 17.1399 | 13.4426 | 11.9039 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H11 | 9.8431 | 12.2323 | 14.4454 | 10.8817 | 12.3902 | 26.1286 | 12.8104 | 20.4006 | 14.4725 | 22.0254 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H12 | 8.8894 | 19.5458 | 10.9804 | 6.5133 | 1.9769 | 17.8550 | 10.0447 | 22.0264 | 10.4057 | 16.4450 | 12.9782 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H13 | 26.8484 | 25.4982 | 16.3809 | 13.7909 | 9.7537 | 3.6948 | 7.2696 | 19.7119 | 14.9077 | 13.7707 | 18.8837 | 15.0532 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . |
| H14 | 17.7664 | 41.1736 | 37.6687 | 29.4951 | 27.0144 | 42.7146 | 33.6413 | 38.5873 | 20.7355 | 38.3531 | 28.6544 | 25.1393 | 44.2039 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . |
| H15 | 13.4556 | 23.9862 | 18.4359 | 18.8519 | 22.6508 | 36.7724 | 23.8902 | 27.2628 | 14.1269 | 20.5443 | 15.6678 | 23.3132 | 37.2932 | 21.2757 | 0.0000 | . | . | . | . | . | . | . | . | . | . |
| H16 | 22.3497 | 13.9545 | 23.4801 | 28.7373 | 22.9958 | 15.7018 | 20.7376 | 22.9388 | 12.7012 | 27.0854 | 31.6726 | 17.8298 | 26.3422 | 35.5854 | 27.2584 | 0.0000 | . | . | . | . | . | . | . | . | . |
| H17 | 12.7963 | 19.4783 | 6.4380 | 5.4471 | 10.7504 | 26.2013 | 10.6303 | 25.7456 | 10.5577 | 16.9288 | 4.5195 | 13.0552 | 20.1857 | 33.8521 | 16.2064 | 33.5307 | 0.0000 | . | . | . | . | . | . | . | . |
| H18 | 17.3965 | 22.1363 | 9.2691 | 8.0517 | 3.6194 | 12.3529 | 10.1359 | 24.8222 | 6.1157 | 11.6797 | 14.6477 | 6.0987 | 7.9865 | 33.1620 | 27.5329 | 20.2659 | 12.4495 | 0.0000 | . | . | . | . | . | . | . |
| H19 | 14.9811 | 19.9152 | 12.3752 | 10.0655 | 13.4436 | 19.7415 | 10.4769 | 19.2172 | 10.3824 | 17.9910 | 11.5565 | 12.2457 | 21.4565 | 25.7348 | 16.9730 | 22.3010 | 7.7818 | 18.5636 | 0.0000 | . | . | . | . | . | . |
| H20 | 8.7856 | 18.2314 | 12.6714 | 8.9105 | 6.7448 | 20.6744 | 14.8192 | 14.0098 | 7.3025 | 10.2499 | 8.5136 | 6.7392 | 19.3266 | 15.8382 | 7.5499 | 17.6474 | 10.8783 | 10.5770 | 10.4458 | 0.0000 | . | . | . | . | . |
| H21 | 13.1689 | 12.9841 | 8.1886 | 10.4899 | 8.9161 | 7.7829 | 5.2077 | 18.6213 | 6.4842 | 15.4137 | 18.4172 | 7.3440 | 12.3290 | 32.4085 | 21.6780 | 6.5634 | 16.7877 | 9.7789 | 9.3594 | 11.4773 | 0.0000 | . | . | . | . |
| H22 | 16.0742 | 12.9818 | 13.6875 | 15.6175 | 18.8029 | 24.0325 | 17.1323 | 14.8594 | 17.5843 | 21.1480 | 8.1767 | 21.9442 | 20.8653 | 42.0891 | 22.9282 | 29.8527 | 11.8722 | 17.4258 | 22.9603 | 15.4479 | 21.2952 | 0.0000 | . | . | . |
| H23 | 7.5573 | 9.0411 | 14.4622 | 13.0543 | 12.4302 | 23.6673 | 14.7893 | 17.7263 | 10.7822 | 22.5417 | 3.7984 | 11.8811 | 20.0984 | 23.1962 | 16.9076 | 22.2433 | 9.4459 | 11.7354 | 16.1207 | 7.4811 | 16.1998 | 4.9128 | 0.0000 | . | . |
| H24 | 38.1390 | 26.4461 | 34.3493 | 39.2576 | 40.2825 | 47.9796 | 45.7601 | 35.1143 | 23.9060 | 35.1780 | 22.4263 | 41.4541 | 46.3069 | 50.6075 | 28.9241 | 38.9602 | 24.9323 | 29.9187 | 37.8214 | 26.0628 | 40.7689 | 15.6847 | 15.5309 | 0.0000 | . |
| H25 | 11.8087 | 16.4519 | 4.9551 | 5.5582 | 7.4258 | 8.9485 | 0.2112 | 22.6571 | 10.0771 | 16.7351 | 13.2222 | 9.9613 | 7.2383 | 36.9052 | 25.1476 | 21.4299 | 10.4335 | 9.3329 | 12.0443 | 15.5529 | 5.4818 | 16.0761 | 14.8666 | 45.1629 | 0 |

# APPENDIX – III

## Chebychev distance between Hybrid Tea genotypes

| Genotypes | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 | H17 | H18 | H19 | H20 | H21 | H22 | H23 | H24 | H25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H1 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H2 | 2.52448 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H3 | 2.48432 | 2.76663 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H4 | 1.97617 | 2.66883 | 1.30020 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H5 | 1.71079 | 2.87413 | 1.50550 | 1.39896 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H6 | 3.03224 | 2.48432 | 2.07216 | 2.13773 | 2.11818 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H7 | 2.13773 | 2.21551 | 1.29838 | 1.43706 | 1.64236 | 2.16114 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H8 | 2.91720 | 3.19951 | 3.52772 | 3.25577 | 3.11729 | 2.96397 | 2.74223 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H9 | 1.64960 | 1.99070 | 1.61602 | 1.54326 | 2.03739 | 1.96897 | 1.96897 | 3.13408 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H10 | 3.78295 | 4.06525 | 1.91030 | 1.91030 | 2.52008 | 1.84765 | 2.50314 | 2.85570 | 2.50314 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H11 | 1.91609 | 3.07942 | 2.25847 | 2.05516 | 2.12864 | 2.32134 | 1.67493 | 2.69135 | 1.84765 | 3.55710 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H12 | 1.85241 | 2.42730 | 2.18319 | 1.68839 | 0.96173 | 2.46560 | 1.92959 | 3.18680 | 2.32628 | 3.48182 | 2.08671 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H13 | 3.48857 | 2.91151 | 2.64921 | 2.71479 | 2.30822 | 1.36863 | 1.67874 | 3.54103 | 2.25576 | 2.34757 | 2.89840 | 3.04266 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| H14 | 3.23205 | 3.63605 | 3.63605 | 3.23205 | 3.23205 | 3.99635 | 2.95285 | 3.69356 | 2.13831 | 4.46049 | 3.23205 | 3.23205 | 3.56582 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . |
| H15 | 2.36290 | 3.12574 | 2.42404 | 2.79597 | 2.86945 | 3.63350 | 2.61574 | 3.33071 | 1.98990 | 2.16437 | 2.26002 | 2.82752 | 3.82955 | 3.37903 | 0.00000 | . | . | . | . | . | . | . | . | . | . |
| H16 | 3.29960 | 2.32516 | 3.13559 | 3.62688 | 3.83217 | 2.45249 | 3.37678 | 3.35315 | 2.18981 | 3.96904 | 4.03747 | 3.28472 | 3.49001 | 3.14786 | 2.98122 | 0.00000 | . | . | . | . | . | . | . | . | . |
| H17 | 1.79741 | 2.87413 | 1.50550 | 1.57022 | 2.02607 | 2.54706 | 1.75740 | 3.26363 | 2.11290 | 2.65960 | 1.63619 | 2.31561 | 3.10823 | 3.23205 | 2.39868 | 3.83217 | 0.00000 | . | . | . | . | . | . | . | . |
| H18 | 2.34445 | 2.46354 | 2.18044 | 2.18044 | 1.20875 | 1.65887 | 2.42164 | 3.30026 | 1.61602 | 2.65371 | 2.40234 | 1.56067 | 1.46198 | 3.63605 | 2.42404 | 3.42158 | 2.80756 | 0.00000 | . | . | . | . | . | . | . |
| H19 | 2.67349 | 2.42404 | 2.42404 | 2.02003 | 2.02003 | 3.16069 | 2.38691 | 3.18215 | 1.60156 | 2.86073 | 2.02003 | 2.02003 | 3.73775 | 3.63760 | 3.27475 | 3.25136 | 2.02003 | 2.42404 | 0.00000 | . | . | . | . | . | . |
| H20 | 1.74628 | 2.53197 | 1.61602 | 1.58226 | 1.45007 | 2.91151 | 2.13773 | 2.30472 | 1.45202 | 2.68570 | 1.80417 | 1.40814 | 3.48857 | 2.02003 | 1.42790 | 3.49001 | 2.20938 | 2.00658 | 2.24876 | 0.00000 | . | . | . | . | . |
| H21 | 2.17107 | 1.98927 | 1.41825 | 2.32668 | 2.53197 | 1.88855 | 1.65945 | 3.13266 | 2.08300 | 2.66883 | 2.73727 | 1.98452 | 2.46560 | 3.13518 | 2.77233 | 1.71734 | 2.53197 | 2.12138 | 1.71079 | 2.18981 | 0.00000 | . | . | . | . |
| H22 | 1.99499 | 2.27730 | 3.10317 | 2.83122 | 2.69274 | 2.42404 | 2.31789 | 2.42404 | 2.70954 | 2.43115 | 1.76914 | 2.76225 | 2.86905 | 4.44407 | 3.23205 | 3.01099 | 2.83908 | 2.87571 | 3.23205 | 2.42404 | 2.70811 | 0.00000 | . | . | . |
| H23 | 1.33799 | 2.18981 | 2.55873 | 2.28677 | 2.14830 | 2.51606 | 1.77324 | 2.67253 | 2.16509 | 3.53828 | 1.09987 | 2.21780 | 3.09512 | 3.63605 | 2.42404 | 3.14786 | 2.29464 | 2.33127 | 2.42404 | 1.61602 | 2.16366 | 1.75032 | 0.00000 | . | . |
| H24 | 3.91514 | 3.72508 | 3.30696 | 3.82391 | 4.31805 | 4.24963 | 4.24963 | 4.44729 | 2.82666 | 3.75549 | 3.32216 | 4.60694 | 4.31805 | 4.21579 | 3.28415 | 3.96835 | 2.77861 | 3.18583 | 3.66426 | 3.73268 | 4.36367 | 2.54674 | 2.57715 | 0.00000 | . |
| H25 | 2.13773 | 2.19995 | 1.29838 | 1.43706 | 1.64236 | 2.10325 | 0.40401 | 2.83426 | 1.96897 | 2.31494 | 1.85937 | 1.87170 | 1.62086 | 3.23205 | 2.60018 | 3.31890 | 1.75740 | 2.36375 | 2.38691 | 2.13773 | 1.60156 | 2.40971 | 1.86526 | 4.24963 | 0 |

# APPENDIX – IV

## City Block distance between Hybrid Tea genotypes

| Genotypes | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 | H17 | H18 | H19 | H20 | H21 | H22 | H23 | H24 | H25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H1 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H2 | 8.7490 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H3 | 7.7382 | 11.1783 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H4 | 7.3813 | 12.4742 | 3.8235 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H5 | 8.8436 | 12.6616 | 7.2042 | 3.5119 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H6 | 12.3737 | 12.9707 | 10.9214 | 10.7485 | 9.5132 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H7 | 7.4715 | 11.0889 | 6.5192 | 6.2551 | 6.4327 | 6.7290 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H8 | 12.7982 | 12.5261 | 12.1906 | 11.4441 | 10.8929 | 8.2796 | 12.1186 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H9 | 9.9446 | 12.1161 | 6.5948 | 7.9397 | 6.2045 | 8.9001 | 7.2411 | 11.2923 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H10 | 13.0387 | 14.8065 | 8.3497 | 7.3237 | 6.6157 | 9.3449 | 10.3794 | 8.0087 | 6.7841 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H11 | 6.8504 | 7.1303 | 8.7214 | 7.4147 | 8.8124 | 14.2745 | 9.1196 | 12.3607 | 10.7457 | 10.6836 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H12 | 7.4038 | 11.0885 | 7.9401 | 5.0330 | 3.1311 | 10.4912 | 7.9896 | 10.2604 | 7.5542 | 8.6100 | 8.3604 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| H13 | 12.0242 | 12.6212 | 10.6756 | 8.4875 | 6.6376 | 4.8008 | 6.2224 | 8.8716 | 9.7032 | 9.3339 | 11.1878 | 8.5279 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . |
| H14 | 10.1514 | 14.7947 | 15.9083 | 14.0768 | 13.5488 | 16.8467 | 16.1461 | 14.9328 | 12.2507 | 16.0135 | 13.3635 | 12.1926 | 16.8314 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . |
| H15 | 8.5130 | 12.3197 | 10.1870 | 10.9069 | 12.9546 | 14.2677 | 11.7877 | 13.6788 | 10.0001 | 12.6237 | 9.9340 | 13.4380 | 14.6156 | 10.9755 | 0.0000 | . | . | . | . | . | . | . | . | . | . |
| H16 | 10.9458 | 9.4819 | 12.2175 | 12.5076 | 11.2291 | 9.5398 | 10.0254 | 10.7529 | 8.5724 | 10.4403 | 12.1252 | 10.0456 | 11.3481 | 16.2928 | 12.9152 | 0.0000 | . | . | . | . | . | . | . | . | . |
| H17 | 9.3401 | 11.0792 | 6.7062 | 5.4456 | 7.1203 | 14.5907 | 7.8617 | 12.9513 | 7.8493 | 9.7883 | 4.6630 | 8.0852 | 11.0809 | 15.0556 | 10.6305 | 12.3648 | 0.0000 | . | . | . | . | . | . | . | . |
| H18 | 10.8638 | 12.5472 | 6.4428 | 6.4531 | 4.9274 | 9.7302 | 7.3235 | 12.2292 | 6.0199 | 7.6948 | 9.5214 | 5.7729 | 7.3221 | 14.9702 | 13.9522 | 10.8463 | 7.2990 | 0.0000 | . | . | . | . | . | . | . |
| H19 | 8.7563 | 11.3107 | 8.2664 | 7.5072 | 9.4453 | 11.0972 | 7.6046 | 9.2894 | 8.0397 | 10.0913 | 8.8874 | 8.7223 | 10.3050 | 12.4902 | 8.3795 | 9.8952 | 6.9708 | 9.8875 | 0.0000 | . | . | . | . | . | . |
| H20 | 7.2684 | 10.9670 | 10.0923 | 7.7710 | 6.6180 | 10.3606 | 10.1641 | 7.9605 | 6.6107 | 6.6320 | 7.8284 | 6.8142 | 8.9133 | 10.2237 | 7.1397 | 8.0706 | 7.9884 | 8.0157 | 7.4101 | 0.0000 | . | . | . | . | . |
| H21 | 9.4646 | 9.1261 | 7.5403 | 7.7577 | 5.3460 | 7.0959 | 4.9954 | 9.6835 | 5.0374 | 8.5460 | 11.3654 | 6.4277 | 6.9945 | 15.5954 | 12.4729 | 6.3456 | 9.3652 | 7.4754 | 7.6845 | 8.3215 | 0.0000 | . | . | . | . |
| H22 | 11.2111 | 8.8224 | 8.2987 | 9.0998 | 11.6594 | 13.5050 | 10.8997 | 9.9485 | 10.3656 | 12.3400 | 7.1601 | 12.1147 | 11.8996 | 14.6311 | 11.4172 | 13.6176 | 7.3388 | 10.5807 | 11.2492 | 10.4009 | 12.0263 | 0.0000 | . | . | . |
| H23 | 7.5184 | 6.5526 | 8.5788 | 9.0685 | 9.1209 | 13.9580 | 11.0029 | 10.7446 | 8.1956 | 11.2658 | 5.0685 | 7.2855 | 11.6556 | 9.6449 | 11.2200 | 11.5464 | 6.7492 | 7.9478 | 10.1650 | 7.0703 | 11.0489 | 5.0470 | 0.0000 | . | . |
| H24 | 15.1009 | 11.8294 | 14.5768 | 14.9297 | 14.2751 | 16.8165 | 16.2465 | 13.1672 | 12.6861 | 15.0623 | 11.1890 | 13.7843 | 15.2421 | 16.7555 | 13.0566 | 14.9958 | 11.5787 | 11.6913 | 14.8386 | 11.3910 | 15.2120 | 9.5422 | 8.7307 | 0.0000 | . |
| H25 | 7.2526 | 10.9132 | 5.7309 | 5.5069 | 6.0609 | 7.0161 | 0.7793 | 12.3748 | 7.6463 | 10.4081 | 8.9780 | 7.6179 | 6.4784 | 16.7663 | 12.0005 | 10.6578 | 7.5746 | 6.9206 | 8.0529 | 10.7532 | 5.4437 | 10.3045 | 10.7841 | 15.7266 | 0 |

# APPENDIX – V

## Jaccard coefficient between Hybrid Tea genotypes

| accn | h1 | h2 | h3 | h4 | h5 | h6 | h7 | h8 | h9 | h10 | h11 | h12 | h13 | h14 | h15 | h16 | h17 | h18 | h19 | h20 | h21 | h22 | h23 | h24 | h25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h1 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h2 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h3 | 1.00000 | 0.83333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h4 | 0.66667 | 0.71429 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h5 | 0.50000 | 0.57143 | 0.60000 | 0.20000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h6 | 0.40000 | 0.50000 | 0.80000 | 0.40000 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h7 | 0.40000 | 0.50000 | 0.80000 | 0.40000 | 0.20000 | 0.40000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h8 | 0.33333 | 0.42857 | 0.85714 | 0.57143 | 0.42857 | 0.57143 | 0.57143 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h9 | 0.40000 | 0.50000 | 1.00000 | 0.66667 | 0.50000 | 0.66667 | 0.40000 | 0.33333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h10 | 0.83333 | 0.85714 | 0.75000 | 0.60000 | 0.66667 | 0.83333 | 0.83333 | 0.50000 | 0.60000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h11 | 0.33333 | 0.42857 | 0.85714 | 0.57143 | 0.42857 | 0.57143 | 0.57143 | 0.00000 | 0.33333 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h12 | 0.40000 | 0.71429 | 0.80000 | 0.40000 | 0.20000 | 0.66667 | 0.40000 | 0.33333 | 0.40000 | 0.60000 | 0.33333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h13 | 0.60000 | 0.85714 | 0.75000 | 0.25000 | 0.40000 | 0.60000 | 0.60000 | 0.50000 | 0.60000 | 0.50000 | 0.50000 | 0.25000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| h14 | 0.33333 | 0.42857 | 0.66667 | 0.33333 | 0.16667 | 0.33333 | 0.33333 | 0.28571 | 0.57143 | 0.71429 | 0.28571 | 0.33333 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . |
| h15 | 0.57143 | 0.42857 | 0.66667 | 0.33333 | 0.16667 | 0.57143 | 0.33333 | 0.28571 | 0.33333 | 0.50000 | 0.28571 | 0.33333 | 0.50000 | 0.28571 | 0.00000 | . | . | . | . | . | . | . | . | . | . |
| h16 | 0.83333 | 0.66667 | 0.75000 | 0.60000 | 0.66667 | 0.60000 | 0.60000 | 0.71429 | 0.60000 | 0.50000 | 0.71429 | 0.83333 | 0.80000 | 0.71429 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . |
| h17 | 0.66667 | 0.71429 | 0.80000 | 0.40000 | 0.50000 | 0.66667 | 0.66667 | 0.33333 | 0.40000 | 0.25000 | 0.33333 | 0.40000 | 0.25000 | 0.57143 | 0.33333 | 0.60000 | 0.00000 | . | . | . | . | . | . | . | . |
| h18 | 0.66667 | 0.50000 | 0.50000 | 0.40000 | 0.20000 | 0.66667 | 0.40000 | 0.57143 | 0.66667 | 0.83333 | 0.57143 | 0.40000 | 0.60000 | 0.33333 | 0.33333 | 0.83333 | 0.66667 | 0.00000 | . | . | . | . | . | . | . |
| h19 | 0.57143 | 0.42857 | 0.66667 | 0.33333 | 0.16667 | 0.57143 | 0.33333 | 0.28571 | 0.33333 | 0.50000 | 0.28571 | 0.33333 | 0.50000 | 0.28571 | 0.00000 | 0.50000 | 0.33333 | 0.33333 | 0.00000 | . | . | . | . | . | . |
| h20 | 0.50000 | 0.57143 | 0.83333 | 0.50000 | 0.33333 | 0.71429 | 0.50000 | 0.16667 | 0.20000 | 0.40000 | 0.16667 | 0.20000 | 0.40000 | 0.42857 | 0.16667 | 0.66667 | 0.20000 | 0.50000 | 0.16667 | 0.00000 | . | . | . | . | . |
| h21 | 0.60000 | 0.66667 | 0.75000 | 0.60000 | 0.40000 | 0.60000 | 0.25000 | 0.71429 | 0.60000 | 0.80000 | 0.71429 | 0.60000 | 0.80000 | 0.50000 | 0.50000 | 0.50000 | 0.83333 | 0.60000 | 0.50000 | 0.66667 | 0.00000 | . | . | . | . |
| h22 | 0.83333 | 0.66667 | 0.75000 | 0.83333 | 0.85714 | 0.83333 | 1.00000 | 0.50000 | 0.83333 | 0.50000 | 0.50000 | 0.83333 | 0.80000 | 0.71429 | 0.71429 | 0.80000 | 0.60000 | 0.83333 | 0.71429 | 0.66667 | 1.00000 | 0.00 | . | . | . |
| h23 | 0.66667 | 0.50000 | 0.80000 | 0.66667 | 0.71429 | 0.66667 | 0.85714 | 0.33333 | 0.66667 | 0.60000 | 0.33333 | 0.66667 | 0.60000 | 0.57143 | 0.57143 | 0.83333 | 0.40000 | 0.66667 | 0.57143 | 0.50000 | 1.00000 | 0.25 | 0.0 | . | . |
| h24 | 0.50000 | 0.33333 | 0.83333 | 0.71429 | 0.57143 | 0.71429 | 0.71429 | 0.16667 | 0.50000 | 0.66667 | 0.16667 | 0.50000 | 0.66667 | 0.42857 | 0.42857 | 0.85714 | 0.50000 | 0.50000 | 0.42857 | 0.33333 | 0.85714 | 0.40 | 0.2 | 0 | . |
| h25 | 0.80000 | 0.83333 | 0.66667 | 0.50000 | 0.60000 | 0.50000 | 0.50000 | 0.85714 | 0.80000 | 0.75000 | 0.85714 | 0.80000 | 0.75000 | 0.66667 | 0.66667 | 0.33333 | 0.80000 | 0.80000 | 0.66667 | 0.83333 | 0.33333 | 1.00 | 1.0 | 1 | 0 |

196

# APPENDIX – VI

## Dice coefficient between Hybrid Tea genotypes

| acon | h1 | h2 | h3 | h4 | h5 | h6 | h7 | h8 | h9 | h10 | h11 | h12 | h13 | h14 | h15 | h16 | h17 | h18 | h19 | h20 | h21 | h22 | h23 | h24 | h25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h1 | 1.00000 | | | | | | | | | | | | | | | | | | | | | | | | |
| h2 | 0.66667 | 1.00000 | | | | | | | | | | | | | | | | | | | | | | | |
| h3 | 0.00000 | 0.28571 | 1.00000 | | | | | | | | | | | | | | | | | | | | | | |
| h4 | 0.50000 | 0.44444 | 0.66667 | 1.00000 | | | | | | | | | | | | | | | | | | | | | |
| h5 | 0.66667 | 0.60000 | 0.57143 | 0.88889 | 1.00000 | | | | | | | | | | | | | | | | | | | | |
| h6 | 0.75000 | 0.66667 | 0.33333 | 0.75000 | 0.66667 | 1.00000 | | | | | | | | | | | | | | | | | | | |
| h7 | 0.75000 | 0.66667 | 0.33333 | 0.75000 | 0.88889 | 0.75000 | 1.00000 | | | | | | | | | | | | | | | | | | |
| h8 | 0.80000 | 0.72727 | 0.25000 | 0.60000 | 0.72727 | 0.60000 | 0.60000 | 1.00000 | | | | | | | | | | | | | | | | | |
| h9 | 0.75000 | 0.66667 | 0.00000 | 0.50000 | 0.66667 | 0.50000 | 0.75000 | 0.80000 | 1.00000 | | | | | | | | | | | | | | | | |
| h10 | 0.28571 | 0.25000 | 0.40000 | 0.57143 | 0.50000 | 0.28571 | 0.28571 | 0.66667 | 0.57143 | 1.00000 | | | | | | | | | | | | | | | |
| h11 | 0.80000 | 0.72727 | 0.25000 | 0.60000 | 0.72727 | 0.60000 | 0.60000 | 1.00000 | 0.80000 | 0.66667 | 1.00000 | | | | | | | | | | | | | | |
| h12 | 0.75000 | 0.44444 | 0.33333 | 0.75000 | 0.88889 | 0.50000 | 0.75000 | 0.80000 | 0.75000 | 0.57143 | 0.80000 | 1.00000 | | | | | | | | | | | | | |
| h13 | 0.57143 | 0.25000 | 0.40000 | 0.85714 | 0.75000 | 0.57143 | 0.57143 | 0.66667 | 0.57143 | 0.66667 | 0.66667 | 0.85714 | 1.00000 | | | | | | | | | | | | |
| h14 | 0.80000 | 0.72727 | 0.50000 | 0.80000 | 0.90909 | 0.80000 | 0.80000 | 0.83333 | 0.60000 | 0.44444 | 0.83333 | 0.60000 | 0.66667 | 1.00000 | | | | | | | | | | | |
| h15 | 0.60000 | 0.72727 | 0.50000 | 0.80000 | 0.90909 | 0.80000 | 0.80000 | 0.83333 | 0.80000 | 0.66667 | 0.83333 | 0.80000 | 0.66667 | 0.83333 | 1.00000 | | | | | | | | | | |
| h16 | 0.28571 | 0.50000 | 0.40000 | 0.57143 | 0.50000 | 0.57143 | 0.57143 | 0.44444 | 0.57143 | 0.66667 | 0.44444 | 0.28571 | 0.33333 | 0.44444 | 0.66667 | 1.00000 | | | | | | | | | |
| h17 | 0.50000 | 0.44444 | 0.33333 | 0.75000 | 0.66667 | 0.50000 | 0.50000 | 0.80000 | 0.75000 | 0.85714 | 0.80000 | 0.75000 | 0.85714 | 0.80000 | 0.80000 | 0.57143 | 1.00000 | | | | | | | | |
| h18 | 0.50000 | 0.66667 | 0.66667 | 0.75000 | 0.88889 | 0.50000 | 0.75000 | 0.60000 | 0.50000 | 0.28571 | 0.60000 | 0.75000 | 0.57143 | 0.80000 | 0.80000 | 0.28571 | 0.50000 | 1.00000 | | | | | | | |
| h19 | 0.60000 | 0.72727 | 0.50000 | 0.60000 | 0.90909 | 0.60000 | 0.80000 | 0.83333 | 0.60000 | 0.66667 | 0.83333 | 0.60000 | 0.66667 | 0.83333 | 1.00000 | 0.66667 | 0.80000 | 0.80000 | 1.00000 | | | | | | |
| h20 | 0.66667 | 0.60000 | 0.28571 | 0.66667 | 0.80000 | 0.44444 | 0.66667 | 0.90909 | 0.88889 | 0.75000 | 0.90909 | 0.88889 | 0.75000 | 0.72727 | 0.90909 | 0.50000 | 0.88889 | 0.66667 | 0.90909 | 1.00000 | | | | | |
| h21 | 0.57143 | 0.50000 | 0.40000 | 0.57143 | 0.75000 | 0.57143 | 0.85714 | 0.44444 | 0.57143 | 0.33333 | 0.44444 | 0.57143 | 0.33333 | 0.66667 | 0.66667 | 0.66667 | 0.28571 | 0.57143 | 0.66667 | 0.50000 | 1.00 | | | | |
| h22 | 0.28571 | 0.50000 | 0.40000 | 0.28571 | 0.25000 | 0.28571 | 0.00000 | 0.66667 | 0.28571 | 0.66667 | 0.66667 | 0.28571 | 0.33333 | 0.44444 | 0.44444 | 0.33333 | 0.57143 | 0.28571 | 0.44444 | 0.50000 | 0.00 | 1.00000 | | | |
| h23 | 0.50000 | 0.66667 | 0.33333 | 0.50000 | 0.44444 | 0.50000 | 0.25000 | 0.80000 | 0.50000 | 0.57143 | 0.80000 | 0.50000 | 0.57143 | 0.60000 | 0.60000 | 0.28571 | 0.75000 | 0.50000 | 0.60000 | 0.66667 | 0.00 | 0.85714 | 1.00000 | | |
| h24 | 0.66667 | 0.80000 | 0.28571 | 0.44444 | 0.60000 | 0.44444 | 0.44444 | 0.90909 | 0.66667 | 0.60000 | 0.90909 | 0.66667 | 0.50000 | 0.72727 | 0.72727 | 0.25000 | 0.66667 | 0.66667 | 0.72727 | 0.80000 | 0.25 | 0.75000 | 0.88889 | 1 | |
| h25 | 0.33333 | 0.28571 | 0.50000 | 0.66667 | 0.57143 | 0.66667 | 0.66667 | 0.25000 | 0.33333 | 0.40000 | 0.25000 | 0.33333 | 0.40000 | 0.50000 | 0.50000 | 0.80000 | 0.33333 | 0.33333 | 0.60000 | 0.28571 | 0.80 | 0.00000 | 0.00000 | 0 | 1 |

# APPENDIX – VII

## Hamann's coefficient between Hybrid Tea genotypes

| accn | h1 | h2 | h3 | h4 | h5 | h6 | h7 | h8 | h9 | h10 | h11 | h12 | h13 | h14 | h15 | h16 | h17 | h18 | h19 | h20 | h21 | h22 | h23 | h24 | h25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h2 | -0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h3 | -0.14286 | -0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h4 | 0.42857 | -0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h5 | 0.14286 | -0.14286 | -0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h6 | 0.42857 | -0.42857 | 0.14286 | 0.71429 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h7 | 0.71429 | 0.14286 | 0.14286 | 0.71429 | 0.42857 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h8 | 0.71429 | 0.14286 | -0.14286 | 0.14286 | -0.14286 | 0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h9 | 0.71429 | -0.42857 | -0.14286 | 0.42857 | 0.42857 | 0.42857 | 0.42857 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h10 | 0.14286 | -0.71429 | -0.14286 | -0.14286 | 0.42857 | -0.14286 | -0.14286 | -0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h11 | 0.14286 | -0.14286 | -0.71429 | -0.42857 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | -0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h12 | 1.00000 | -0.14286 | -0.14286 | 0.42857 | 0.14286 | 0.42857 | 0.71429 | 0.71429 | 0.71429 | 0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h13 | 0.71429 | -0.42857 | 0.14286 | 0.71429 | 0.14286 | 0.71429 | 0.42857 | 0.42857 | 0.71429 | 0.14286 | -0.14286 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| h14 | -0.42857 | -0.42857 | -0.71429 | -0.42857 | -0.42857 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | -0.71429 | -1.00000 | -0.42857 | -0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| h15 | 0.14286 | -0.14286 | -0.71429 | -0.14286 | -0.42857 | 0.14286 | 0.14286 | -0.14286 | -0.14286 | -0.71429 | -0.14286 | 0.14286 | -0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| h16 | -0.14286 | -0.14286 | 0.14286 | 0.14286 | 0.42857 | -0.14286 | 0.14286 | -0.42857 | 0.14286 | 0.42857 | -0.42857 | -0.14286 | -0.14286 | -0.71429 | -0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . |
| h17 | -0.14286 | -0.71429 | 0.14286 | -0.14286 | -0.14286 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | 0.14286 | 0.14286 | -0.14286 | 0.14286 | -1.00000 | -1.00000 | -0.42857 | 1.00000 | . | . | . | . | . | . | . | . |
| h18 | -0.14286 | -0.42857 | 0.42857 | 0.14286 | 0.14286 | -0.14286 | 0.14286 | -0.14286 | 0.14286 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | -0.42857 | -0.71429 | 0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . |
| h19 | 0.42857 | -0.14286 | -0.14286 | 0.42857 | 0.14286 | 0.14286 | 0.71429 | 0.14286 | 0.14286 | -0.42857 | -0.42857 | 0.42857 | 0.14286 | -0.14286 | 0.42857 | -0.14286 | -0.71429 | -0.14286 | 1.00000 | . | . | . | . | . | . |
| h20 | 1.00000 | -0.14286 | -0.14286 | 0.42857 | 0.14286 | 0.42857 | 0.71429 | 0.71429 | 0.71429 | 0.14286 | 0.14286 | 1.00000 | 0.71429 | -0.42857 | 0.14286 | -0.14286 | -0.14286 | -0.14286 | 0.42857 | 1.00000 | . | . | . | . | . |
| h21 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | 0.14286 | 0.14286 | -0.14286 | -0.42857 | 0.14286 | 0.42857 | -0.42857 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | 0.71429 | -0.42857 | -0.14286 | -0.42857 | -0.14286 | 1.00000 | . | . | . | . |
| h22 | -0.14286 | -0.71429 | 0.14286 | -0.14286 | -0.71429 | -0.14286 | -0.42857 | 0.14286 | -0.14286 | -0.14286 | -0.71429 | -0.14286 | 0.14286 | -0.42857 | -0.71429 | -0.42857 | -0.14286 | -0.42857 | -0.42857 | -0.14286 | -0.42857 | 1.00000 | . | . | . |
| h23 | 0.42857 | -0.14286 | 0.14286 | 0.42857 | -0.14286 | 0.42857 | 0.14286 | 0.71429 | 0.42857 | -0.14286 | -0.42857 | 0.42857 | 0.71429 | -0.14286 | -0.42857 | -0.42857 | 0.14286 | -0.14286 | -0.14286 | 0.42857 | -0.42857 | 0.42857 | 1.00000 | . | . |
| h24 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | -0.42857 | -0.42857 | -0.42857 | 0.14286 | 0.14286 | -0.14286 | 0.14286 | -0.14286 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | -0.14286 | 0.14286 | -0.71429 | -0.14286 | -0.42857 | -0.14286 | 0.14286 | 1.00000 | . |
| h25 | 0.14286 | -0.42857 | 0.71429 | 0.71429 | 0.14286 | 0.42857 | 0.42857 | -0.14286 | 0.14286 | 0.14286 | -0.42857 | 0.14286 | 0.42857 | -0.71429 | -0.42857 | 0.42857 | -0.14286 | 0.14286 | 0.14286 | 0.14286 | 0.14286 | 0.14286 | 0.14286 | -0.42857 | 1 |

# APPENDIX – VIII

## Simple matching coefficient between Hybrid Tea genotypes

| accn | h1 | h2 | h3 | h4 | h5 | h6 | h7 | h8 | h9 | h10 | h11 | h12 | h13 | h14 | h15 | h16 | h17 | h18 | h19 | h20 | h21 | h22 | h23 | h24 | h25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h2 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h3 | 0.42857 | 0.28571 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h4 | 0.71429 | 0.42857 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h5 | 0.57143 | 0.42857 | 0.42857 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h6 | 0.71429 | 0.28571 | 0.57143 | 0.85714 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h7 | 0.85714 | 0.57143 | 0.57143 | 0.85714 | 0.71429 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h8 | 0.85714 | 0.57143 | 0.42857 | 0.57143 | 0.42857 | 0.57143 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h9 | 0.85714 | 0.28571 | 0.42857 | 0.71429 | 0.71429 | 0.71429 | 0.71429 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h10 | 0.57143 | 0.14286 | 0.42857 | 0.42857 | 0.71429 | 0.42857 | 0.42857 | 0.42857 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h11 | 0.57143 | 0.42857 | 0.14286 | 0.28571 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 0.42857 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h12 | 1.00000 | 0.42857 | 0.42857 | 0.71429 | 0.57143 | 0.71429 | 0.85714 | 0.85714 | 0.85714 | 0.57143 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h13 | 0.85714 | 0.28571 | 0.57143 | 0.85714 | 0.57143 | 0.85714 | 0.71429 | 0.71429 | 0.85714 | 0.57143 | 0.42857 | 0.85714 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| h14 | 0.28571 | 0.28571 | 0.14286 | 0.28571 | 0.28571 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 0.14286 | 0.00000 | 0.28571 | 0.28571 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| h15 | 0.57143 | 0.42857 | 0.14286 | 0.42857 | 0.28571 | 0.57143 | 0.57143 | 0.42857 | 0.42857 | 0.14286 | 0.42857 | 0.57143 | 0.42857 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| h16 | 0.42857 | 0.42857 | 0.57143 | 0.57143 | 0.71429 | 0.42857 | 0.57143 | 0.28571 | 0.57143 | 0.71429 | 0.28571 | 0.42857 | 0.42857 | 0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . |
| h17 | 0.42857 | 0.14286 | 0.57143 | 0.42857 | 0.42857 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.00000 | 0.00000 | 0.28571 | 1.00000 | . | . | . | . | . | . | . | . |
| h18 | 0.42857 | 0.28571 | 0.71429 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.42857 | 0.57143 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 0.28571 | 0.14286 | 0.57143 | 0.57143 | 1.00000 | . | . | . | . | . | . | . |
| h19 | 0.71429 | 0.42857 | 0.42857 | 0.71429 | 0.57143 | 0.57143 | 0.85714 | 0.57143 | 0.57143 | 0.28571 | 0.28571 | 0.71429 | 0.57143 | 0.42857 | 0.71429 | 0.42857 | 0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . |
| h20 | 1.00000 | 0.42857 | 0.42857 | 0.71429 | 0.57143 | 0.71429 | 0.85714 | 0.85714 | 0.85714 | 0.57143 | 0.57143 | 1.00000 | 0.85714 | 0.28571 | 0.57143 | 0.42857 | 0.42857 | 0.42857 | 0.71429 | 1.00000 | . | . | . | . | . |
| h21 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 0.57143 | 0.57143 | 0.42857 | 0.28571 | 0.57143 | 0.71429 | 0.28571 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.85714 | 0.28571 | 0.42857 | 0.28571 | 0.42857 | 1.00000 | . | . | . | . |
| h22 | 0.42857 | 0.14286 | 0.57143 | 0.42857 | 0.14286 | 0.42857 | 0.28571 | 0.57143 | 0.42857 | 0.42857 | 0.14286 | 0.42857 | 0.57143 | 0.28571 | 0.14286 | 0.28571 | 0.42857 | 0.28571 | 0.28571 | 0.42857 | 0.28571 | 1.00000 | . | . | . |
| h23 | 0.71429 | 0.42857 | 0.57143 | 0.71429 | 0.42857 | 0.71429 | 0.57143 | 0.85714 | 0.71429 | 0.42857 | 0.28571 | 0.71429 | 0.85714 | 0.42857 | 0.28571 | 0.28571 | 0.57143 | 0.42857 | 0.42857 | 0.71429 | 0.28571 | 0.71429 | 1.00000 | . | . |
| h24 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.28571 | 0.28571 | 0.28571 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.42857 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.42857 | 0.57143 | 0.14286 | 0.42857 | 0.28571 | 0.42857 | 0.57143 | 1.00000 | . |
| h25 | 0.57143 | 0.28571 | 0.85714 | 0.85714 | 0.57143 | 0.71429 | 0.71429 | 0.42857 | 0.57143 | 0.57143 | 0.28571 | 0.57143 | 0.71429 | 0.14286 | 0.28571 | 0.71429 | 0.42857 | 0.57143 | 0.57143 | 0.57143 | 0.57143 | 0.57143 | 0.57143 | 0.28571 | 1 |

# APPENDIX – IX

## Gower's measure between Hybrid Tea genotypes

| accn | h1 | h2 | h3 | h4 | h5 | h6 | h7 | h8 | h9 | h10 | h11 | h12 | h13 | h14 | h15 | h16 | h17 | h18 | h19 | h20 | h21 | h22 | h23 | h24 | h25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| h1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h2 | 0.73495 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h3 | 0.59068 | 0.50183 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h4 | 0.67984 | 0.55886 | 0.84035 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h5 | 0.73031 | 0.73833 | 0.68711 | 0.84391 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h6 | 0.57825 | 0.85343 | 0.68749 | 0.77243 | 0.71809 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h7 | 0.76109 | 0.77077 | 0.70104 | 0.78730 | 0.85353 | 0.77464 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h8 | 0.73648 | 0.56504 | 0.49086 | 0.58980 | 0.68676 | 0.65558 | 0.55842 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h9 | 0.79598 | 0.56080 | 0.61829 | 0.67460 | 0.79660 | 0.65078 | 0.76938 | 0.76409 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h10 | 0.56701 | 0.64906 | 0.74623 | 0.68448 | 0.62135 | 0.64065 | 0.53929 | 0.66133 | 0.70053 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h11 | 0.66032 | 0.76938 | 0.57018 | 0.68126 | 0.73350 | 0.54187 | 0.73011 | 0.74757 | 0.77842 | 0.61655 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h12 | 0.64232 | 0.56621 | 0.59093 | 0.73161 | 0.85319 | 0.61060 | 0.74922 | 0.78575 | 0.85360 | 0.66564 | 0.82488 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| h13 | 0.66573 | 0.57424 | 0.60926 | 0.73558 | 0.69263 | 0.81844 | 0.70031 | 0.72700 | 0.71864 | 0.72687 | 0.68565 | 0.73823 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| h14 | 0.71540 | 0.70619 | 0.51136 | 0.63264 | 0.72749 | 0.57576 | 0.67171 | 0.61986 | 0.86490 | 0.42872 | 0.64893 | 0.67118 | 0.49340 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| h15 | 0.74959 | 0.74978 | 0.62537 | 0.69479 | 0.73551 | 0.62212 | 0.75950 | 0.63318 | 0.76970 | 0.49123 | 0.71523 | 0.64492 | 0.53271 | 0.78072 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| h16 | 0.62163 | 0.53800 | 0.74674 | 0.65563 | 0.59993 | 0.71836 | 0.62146 | 0.52405 | 0.57190 | 0.70267 | 0.49899 | 0.54326 | 0.59750 | 0.50186 | 0.56700 | 1.00000 | . | . | . | . | . | . | . | . | . |
| h17 | 0.72445 | 0.60275 | 0.69615 | 0.80831 | 0.68734 | 0.61417 | 0.57287 | 0.64594 | 0.75423 | 0.71905 | 0.81354 | 0.75383 | 0.76814 | 0.52543 | 0.61619 | 0.57577 | 1.00000 | . | . | . | . | . | . | . | . |
| h18 | 0.68697 | 0.73672 | 0.69627 | 0.78323 | 0.90012 | 0.71958 | 0.84821 | 0.66150 | 0.79574 | 0.59680 | 0.71437 | 0.80125 | 0.68169 | 0.59328 | 0.71055 | 0.60898 | 0.67931 | 1.00000 | . | . | . | . | . | . | . |
| h19 | 0.73749 | 0.76663 | 0.66874 | 0.76678 | 0.80748 | 0.68383 | 0.84448 | 0.71975 | 0.75198 | 0.54336 | 0.73445 | 0.74093 | 0.61777 | 0.74325 | 0.82690 | 0.62282 | 0.69502 | 0.79598 | 1.00000 | . | . | . | . | . | . |
| h20 | 0.85211 | 0.59194 | 0.54156 | 0.67159 | 0.78073 | 0.61946 | 0.70376 | 0.83444 | 0.86337 | 0.69794 | 0.83898 | 0.88264 | 0.73301 | 0.71408 | 0.77132 | 0.58132 | 0.74954 | 0.74963 | 0.75977 | 1.00000 | . | . | . | . | . |
| h21 | 0.63524 | 0.72958 | 0.76474 | 0.67482 | 0.80791 | 0.68359 | 0.80987 | 0.62671 | 0.73522 | 0.66274 | 0.59887 | 0.70368 | 0.60173 | 0.60080 | 0.66008 | 0.78261 | 0.55922 | 0.76526 | 0.75914 | 0.66032 | 1.00000 | . | . | . | . |
| h22 | 0.60676 | 0.57442 | 0.74073 | 0.64331 | 0.50872 | 0.55564 | 0.52570 | 0.63566 | 0.61622 | 0.74680 | 0.68923 | 0.58506 | 0.67007 | 0.45919 | 0.51691 | 0.63693 | 0.75940 | 0.52439 | 0.51967 | 0.62042 | 0.58591 | 1.00000 | . | . | . |
| h23 | 0.76216 | 0.70185 | 0.65379 | 0.72939 | 0.64538 | 0.63316 | 0.50601 | 0.70338 | 0.74650 | 0.68945 | 0.81143 | 0.76664 | 0.76302 | 0.64307 | 0.60620 | 0.59599 | 0.85472 | 0.66244 | 0.62382 | 0.77333 | 0.52349 | 0.81602 | 1.00000 | . | . |
| h24 | 0.69651 | 0.58331 | 0.45065 | 0.52908 | 0.62856 | 0.49820 | 0.58593 | 0.74365 | 0.74106 | 0.53487 | 0.77504 | 0.72570 | 0.60951 | 0.58896 | 0.65461 | 0.45159 | 0.67642 | 0.67757 | 0.61815 | 0.77578 | 0.52734 | 0.64124 | 0.74403 | 1.00000 | . |
| h25 | 0.59635 | 0.50712 | 0.88301 | 0.80199 | 0.70452 | 0.76893 | 0.81803 | 0.48666 | 0.59580 | 0.70611 | 0.56566 | 0.59021 | 0.69521 | 0.48320 | 0.58864 | 0.77599 | 0.67858 | 0.68982 | 0.67003 | 0.52587 | 0.80209 | 0.70333 | 0.60994 | 0.42882 | 1 |

# APPENDIX – X

## Euclidean distance between Floribunda genotypes

| accn | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 2.19473 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 3.17065 | 3.14077 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 2.72737 | 2.85357 | 4.46136 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 2.89574 | 1.70107 | 2.87137 | 3.86903 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 3.82388 | 2.91983 | 4.75218 | 2.81486 | 3.42479 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 3.39157 | 3.25817 | 2.87721 | 4.30019 | 3.29217 | 3.54086 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 4.68409 | 3.43721 | 3.39925 | 4.74223 | 2.67499 | 4.11756 | 4.28643 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 5.32243 | 4.89711 | 5.60270 | 5.81633 | 4.22806 | 5.11650 | 6.02524 | 4.14406 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 1.50805 | 2.54248 | 3.77887 | 2.90183 | 2.91098 | 3.88012 | 4.10685 | 4.59829 | 4.45990 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 4.58569 | 3.76074 | 4.33408 | 4.05841 | 3.55818 | 4.47437 | 5.53956 | 2.98844 | 4.18232 | 4.40124 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 3.38439 | 3.68526 | 4.95458 | 1.89084 | 4.22178 | 2.77454 | 4.64019 | 4.98312 | 5.17348 | 3.37291 | 3.78963 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 4.26805 | 2.62965 | 3.62646 | 5.08149 | 2.93288 | 4.80917 | 4.04559 | 3.88183 | 5.40624 | 4.42855 | 4.78585 | 5.75786 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 3.84581 | 3.51986 | 4.22038 | 5.39769 | 3.36926 | 5.22926 | 5.41709 | 4.27219 | 4.49848 | 4.06520 | 4.08779 | 5.52407 | 4.27189 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 3.13692 | 3.33898 | 2.38515 | 4.42829 | 2.74845 | 3.81294 | 2.73549 | 3.39586 | 5.09740 | 3.80236 | 4.35399 | 4.44126 | 4.73290 | 3.80046 | 0.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 3.24273 | 2.62836 | 4.87116 | 3.55729 | 2.80967 | 3.15469 | 4.91714 | 4.16132 | 4.34159 | 3.10405 | 3.72242 | 3.62573 | 4.71925 | 3.28778 | 3.86573 | 0.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 3.72293 | 4.47682 | 4.72062 | 5.36763 | 4.31996 | 4.85851 | 4.77649 | 5.57461 | 4.07641 | 3.82591 | 5.20707 | 4.61062 | 5.49050 | 3.59047 | 3.86035 | 4.24460 | 0.00000 | . | . | . | . | . | . | . | . |
| f18 | 2.85775 | 2.11094 | 3.06209 | 3.85551 | 2.27483 | 3.13607 | 3.08330 | 3.62081 | 5.17530 | 3.65587 | 3.88868 | 4.03700 | 3.63014 | 3.06860 | 2.26576 | 2.72527 | 3.84405 | 0.00000 | . | . | . | . | . | . | . |
| f19 | 5.28347 | 5.11761 | 3.56536 | 6.50640 | 4.14271 | 6.92676 | 5.80639 | 4.01979 | 6.42511 | 5.17631 | 4.62969 | 6.73981 | 5.03653 | 4.80057 | 4.86898 | 5.12546 | 6.92432 | 5.36704 | 0.00000 | . | . | . | . | . | . |
| f20 | 3.26010 | 2.47320 | 3.53362 | 3.80778 | 2.71146 | 4.54400 | 4.60469 | 3.74262 | 4.07020 | 2.83511 | 3.14104 | 4.14023 | 2.78307 | 3.71367 | 4.48050 | 3.80688 | 4.45060 | 3.72160 | 4.23290 | 0.00000 | . | . | . | . | . |
| f21 | 6.83383 | 6.11961 | 6.07439 | 6.90968 | 5.12984 | 6.32242 | 7.14446 | 4.46601 | 5.74202 | 6.89769 | 4.02219 | 6.36265 | 6.82958 | 5.41781 | 5.34766 | 5.75453 | 6.59179 | 5.53097 | 5.06722 | 6.03021 | 0.00000 | . | . | . | . |
| f22 | 3.38067 | 2.27665 | 3.46286 | 3.35104 | 2.14235 | 2.54809 | 3.38568 | 2.73683 | 4.19128 | 3.62222 | 2.82966 | 3.23529 | 3.79002 | 3.82280 | 2.80469 | 2.47003 | 4.21705 | 1.89518 | 5.37034 | 3.27731 | 5.17169 | 0.00000 | . | . | . |
| f23 | 4.40479 | 4.12390 | 3.76315 | 4.42298 | 3.89889 | 4.32336 | 4.57551 | 3.18910 | 3.28234 | 4.20090 | 2.75163 | 3.81644 | 4.73736 | 4.49087 | 3.85701 | 4.62497 | 4.01677 | 4.13080 | 4.63279 | 3.32616 | 5.19943 | 3.17177 | 0.00000 | . | . |
| f24 | 3.75884 | 4.42631 | 4.42841 | 5.40057 | 3.89547 | 5.89979 | 5.73353 | 5.45170 | 5.62933 | 4.03232 | 4.65195 | 5.23731 | 5.69726 | 3.43680 | 4.07170 | 4.38538 | 4.08708 | 4.12770 | 4.18955 | 4.33584 | 5.02230 | 4.79681 | 5.22283 | 0.00000 | . |
| f25 | 2.68558 | 2.49118 | 3.29332 | 3.04772 | 2.89588 | 3.72313 | 4.11026 | 3.87176 | 4.71023 | 3.01933 | 2.60307 | 2.98442 | 4.10115 | 3.30628 | 3.25958 | 2.82827 | 3.76835 | 2.27205 | 5.07102 | 2.67171 | 5.54763 | 2.08693 | 3.19367 | 3.62096 | 0 |

# APPENDIX – XI

## Squared Euclidean distance between Floribunda genotypes

| accn | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 4.8169 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 10.0530 | 9.8644 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 7.4386 | 8.1428 | 19.9038 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 8.3863 | 2.8936 | 8.2448 | 14.9694 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 13.1309 | 8.5242 | 22.5632 | 7.9234 | 11.7292 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 11.8028 | 10.5157 | 7.1674 | 18.4916 | 10.8384 | 12.6377 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 21.9407 | 11.8144 | 11.5549 | 22.4888 | 7.1555 | 16.9543 | 18.3735 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 28.3282 | 23.9816 | 31.3909 | 33.8297 | 17.8765 | 28.1798 | 36.3035 | 17.1732 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 2.2682 | 6.4693 | 14.2798 | 8.4206 | 8.4737 | 15.0863 | 16.8662 | 21.1443 | 19.8907 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 21.0286 | 14.1432 | 18.7842 | 16.4707 | 12.6463 | 20.0200 | 30.6867 | 8.9188 | 17.4918 | 19.3709 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 11.4541 | 13.5811 | 24.5479 | 3.5753 | 17.8234 | 7.6980 | 21.5314 | 24.8314 | 28.7649 | 11.3765 | 14.3613 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 18.2163 | 6.9162 | 13.1512 | 25.8215 | 8.8018 | 23.1281 | 16.3868 | 15.0686 | 29.2274 | 19.6120 | 22.9043 | 33.1530 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 14.7887 | 12.3880 | 17.8116 | 29.1350 | 11.3519 | 27.3451 | 29.3448 | 18.2816 | 20.1285 | 16.5259 | 18.7100 | 30.5184 | 18.2483 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 9.8402 | 11.1488 | 5.6660 | 19.6096 | 7.5540 | 14.6385 | 7.4829 | 11.5318 | 25.9834 | 14.4579 | 18.8672 | 19.7247 | 22.4003 | 14.4436 | 0.0000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 10.5153 | 6.9083 | 23.7282 | 12.6643 | 7.8942 | 9.8521 | 24.1783 | 17.3186 | 18.8494 | 9.8351 | 13.8584 | 13.1459 | 22.2713 | 10.8095 | 14.9439 | 0.0000 | . | . | . | . | . | . | . | . | . |
| f17 | 13.8602 | 20.0419 | 22.2843 | 28.8116 | 18.6620 | 23.6051 | 22.8148 | 31.0798 | 16.6172 | 14.8452 | 27.1136 | 21.2598 | 30.1458 | 13.6198 | 14.9023 | 18.0166 | 0.0000 | . | . | . | . | . | . | . | . |
| f18 | 8.1668 | 4.4581 | 9.4993 | 14.8850 | 5.1753 | 9.8349 | 9.5087 | 13.1103 | 28.7837 | 13.3727 | 15.1298 | 16.2974 | 13.1779 | 9.4040 | 5.1337 | 7.4271 | 14.7787 | 0.0000 | . | . | . | . | . | . | . |
| f19 | 27.9151 | 26.1906 | 12.6406 | 42.3333 | 17.1621 | 47.9799 | 33.7142 | 16.1587 | 29.4318 | 26.7942 | 21.4389 | 45.4250 | 25.3687 | 21.1661 | 20.8673 | 37.8212 | 36.0975 | 28.8061 | 0.0000 | . | . | . | . | . | . |
| f20 | 10.6283 | 6.1167 | 12.4966 | 14.4992 | 7.3520 | 20.6479 | 21.2032 | 14.0072 | 18.5665 | 8.0378 | 9.8662 | 17.1415 | 7.7465 | 13.7913 | 20.0749 | 14.4923 | 19.8078 | 13.6603 | 17.9175 | 0.0000 | . | . | . | . | . |
| f21 | 46.7422 | 37.4497 | 36.8982 | 47.7437 | 26.3153 | 39.9730 | 51.0433 | 19.9463 | 32.9708 | 47.5781 | 18.1780 | 40.4834 | 46.6432 | 29.3527 | 28.5975 | 33.1146 | 43.4517 | 30.6917 | 25.6767 | 36.3834 | 0.0000 | . | . | . | . |
| f22 | 11.4282 | 5.1982 | 11.9222 | 11.2295 | 4.5897 | 6.4928 | 11.5307 | 7.4902 | 17.5668 | 12.4060 | 8.0081 | 10.4671 | 14.3642 | 14.6123 | 7.8674 | 6.1010 | 17.7835 | 3.5917 | 28.8405 | 10.7406 | 26.7485 | 0.0000 | . | . | . |
| f23 | 19.4022 | 17.0065 | 14.1613 | 19.5628 | 15.1236 | 16.6915 | 20.9352 | 10.1704 | 10.7738 | 17.6475 | 7.5714 | 14.5652 | 22.4425 | 20.1661 | 14.8765 | 21.3903 | 16.1345 | 17.0619 | 21.4627 | 11.0633 | 27.0341 | 10.0601 | 0.0000 | . | . |
| f24 | 14.1288 | 16.5922 | 19.6108 | 29.1662 | 15.1747 | 34.6075 | 32.8733 | 29.7210 | 31.6894 | 16.2596 | 21.6407 | 27.4294 | 32.4688 | 11.8116 | 16.5795 | 19.2316 | 16.7043 | 17.0379 | 17.5523 | 16.7996 | 25.2235 | 23.0094 | 27.2780 | 0.0000 | . |
| f25 | 7.2129 | 6.2080 | 10.6458 | 9.2886 | 8.3851 | 13.6617 | 16.8942 | 14.9906 | 22.1863 | 9.1184 | 8.7760 | 8.9067 | 16.6194 | 10.9315 | 10.6249 | 7.9981 | 14.2005 | 6.1622 | 25.7153 | 7.1380 | 30.7762 | 4.3553 | 10.1986 | 14.5997 | 0 |

# APPENDIX – XII

## Chebychev distance between Floribunda genotypes

| accn | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 1.19232 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 1.80717 | 2.64921 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 1.80077 | 1.66990 | 2.76496 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 1.61212 | 1.10041 | 1.82616 | 1.98546 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 2.75580 | 1.73131 | 2.50166 | 2.30665 | 1.75115 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 1.88877 | 2.62349 | 1.50315 | 2.73923 | 1.80043 | 2.08336 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 3.10089 | 2.31017 | 2.49622 | 2.76540 | 1.58146 | 2.64697 | 3.18291 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 3.40548 | 3.40548 | 4.54064 | 4.25685 | 3.12169 | 3.12169 | 3.97306 | 3.68927 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 0.85275 | 1.65189 | 1.98853 | 1.70942 | 1.80279 | 3.38320 | 2.31900 | 2.36665 | 2.55411 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 2.72192 | 2.40103 | 2.49337 | 2.97939 | 2.81659 | 2.86096 | 2.72424 | 1.62772 | 3.68927 | 2.53956 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 2.63189 | 2.40103 | 2.49337 | 0.85137 | 2.81659 | 1.93780 | 2.72424 | 2.72879 | 3.40548 | 2.53956 | 2.86458 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 2.89305 | 1.81199 | 2.48203 | 3.15642 | 1.93484 | 3.03799 | 2.45631 | 2.40319 | 3.30151 | 2.93297 | 3.56258 | 3.50114 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 2.29084 | 2.36682 | 2.27032 | 3.51537 | 2.46033 | 3.63474 | 3.37519 | 2.37218 | 2.63238 | 2.84003 | 2.72424 | 3.95448 | 3.68541 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 1.96762 | 2.95784 | 1.58042 | 2.98358 | 2.04478 | 2.32771 | 2.08840 | 2.31017 | 3.97306 | 2.46196 | 2.72424 | 2.72424 | 2.93297 | 2.14768 | 0.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 1.63455 | 2.21125 | 3.09932 | 1.70274 | 2.08840 | 1.82735 | 3.17865 | 2.19910 | 2.55411 | 2.46196 | 2.49337 | 2.49337 | 4.02324 | 2.62963 | 3.31794 | 0.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 2.55411 | 2.55411 | 3.68927 | 3.60548 | 2.43420 | 2.27032 | 3.12169 | 4.01566 | 3.51952 | 1.76751 | 2.89934 | 2.55411 | 3.70078 | 1.84571 | 3.12169 | 2.89356 | 0.00000 | . | . | . | . | . | . | . | . |
| f18 | 2.27929 | 1.25480 | 2.02516 | 1.89015 | 1.51944 | 1.75364 | 1.87342 | 3.10089 | 3.68927 | 2.90870 | 2.53955 | 2.53955 | 2.71799 | 2.05109 | 1.92904 | 1.38891 | 2.83790 | 0.00000 | . | . | . | . | . | . | . |
| f19 | 2.72555 | 2.75209 | 1.93806 | 3.16755 | 2.24753 | 3.82794 | 3.02736 | 2.81139 | 3.68927 | 2.52996 | 2.30869 | 3.50665 | 2.58491 | 2.25547 | 2.90670 | 3.20220 | 3.45750 | 3.35144 | 0.00000 | . | . | . | . | . | . |
| f20 | 2.34550 | 1.65189 | 2.45631 | 2.60297 | 1.80279 | 3.38320 | 2.43059 | 2.36665 | 2.55411 | 2.14994 | 1.71996 | 2.48816 | 1.84270 | 2.21585 | 2.67493 | 2.46195 | 2.25054 | 2.90670 | 2.55919 | 0.00000 | . | . | . | . | . |
| f21 | 4.04643 | 4.00120 | 3.65640 | 4.42363 | 2.90079 | 4.18137 | 4.05989 | 3.58182 | 3.68927 | 4.00303 | 2.94684 | 3.59588 | 4.08008 | 3.59160 | 3.31956 | 3.68575 | 3.87612 | 3.58121 | 3.82146 | 3.53597 | 0.00000 | . | . | . | . |
| f22 | 2.05592 | 1.04420 | 1.80279 | 2.27334 | 1.15434 | 2.15491 | 2.01162 | 2.06210 | 3.40548 | 2.66432 | 1.66225 | 2.15853 | 2.85619 | 3.24729 | 1.85476 | 1.36319 | 2.65411 | 1.19620 | 3.12807 | 2.68432 | 3.51728 | 0.00000 | . | . | . |
| f23 | 2.90894 | 2.67807 | 2.77041 | 2.87493 | 3.00363 | 2.40441 | 3.00128 | 1.75460 | 2.55411 | 2.81659 | 2.03192 | 2.40803 | 2.77041 | 3.00128 | 3.00128 | 3.00930 | 2.27296 | 2.81659 | 2.58672 | 2.36629 | 3.79579 | 1.93929 | 0.00000 | . | . |
| f24 | 2.96987 | 2.92483 | 2.57984 | 3.34707 | 1.98091 | 3.08480 | 3.07117 | 2.91484 | 3.12169 | 2.92647 | 2.63189 | 3.39049 | 3.91675 | 2.51504 | 2.24299 | 2.60918 | 2.27032 | 2.50464 | 2.35688 | 2.45941 | 2.77530 | 2.88330 | 2.90894 | 0.00000 | . |
| f25 | 2.01587 | 1.73521 | 2.01906 | 2.27334 | 1.61237 | 2.15491 | 2.70263 | 3.10089 | 3.68927 | 1.82031 | 1.98457 | 2.15853 | 3.54721 | 2.47139 | 2.23768 | 1.27137 | 2.83790 | 1.61216 | 2.54273 | 1.70450 | 3.51728 | 1.38981 | 1.92904 | 2.54071 | 0 |

203

# APPENDIX – XIII

## City Block distance between Floribunda genotypes

| accn | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 5.1812 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 7.7338 | 6.6025 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 6.7222 | 7.1916 | 10.7586 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 6.9562 | 4.1286 | 7.3711 | 10.9074 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 8.0204 | 7.1623 | 12.6846 | 5.8126 | 8.4944 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 7.8938 | 7.4905 | 6.7241 | 10.5070 | 8.8577 | 9.2605 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 11.0513 | 8.0682 | 8.0276 | 12.5996 | 6.8674 | 10.1940 | 10.1589 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 13.1597 | 12.0506 | 12.4012 | 13.5769 | 10.0575 | 13.3398 | 16.3061 | 8.0107 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 3.5917 | 6.0168 | 9.4535 | 7.0582 | 7.3236 | 7.6858 | 8.8960 | 12.2749 | 11.5962 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 10.8006 | 9.1295 | 11.2130 | 10.2182 | 7.5517 | 11.5420 | 14.9978 | 7.3083 | 8.3106 | 11.2651 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 7.2196 | 8.5706 | 12.6438 | 5.2454 | 10.4212 | 6.2186 | 11.3087 | 12.6809 | 12.0623 | 7.9379 | 7.2415 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 9.1492 | 5.2492 | 7.9639 | 11.8874 | 7.4515 | 11.3518 | 9.9823 | 9.8934 | 13.4460 | 9.2745 | 10.3345 | 12.9657 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 9.2804 | 8.1300 | 9.8746 | 13.3247 | 7.3979 | 12.5590 | 13.0920 | 10.7554 | 10.6209 | 8.4846 | 9.4061 | 11.3330 | 8.3099 | 0.0000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 7.4160 | 6.8468 | 5.5456 | 10.6132 | 6.7721 | 10.1858 | 5.6194 | 7.8771 | 11.7534 | 8.5491 | 10.4319 | 10.9551 | 10.9810 | 8.9153 | 0.0000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 8.4637 | 5.6381 | 11.4261 | 9.8096 | 6.7077 | 8.5570 | 10.7559 | 10.4390 | 10.8891 | 6.5476 | 9.6743 | 9.3456 | 9.5674 | 7.1862 | 7.4791 | 0.0000 | . | . | . | . | . | . | . | . | . |
| f17 | 9.8490 | 10.5948 | 10.1639 | 14.2233 | 10.3665 | 13.8706 | 11.0379 | 12.8064 | 6.5450 | 10.6761 | 12.3430 | 11.5184 | 13.0909 | 9.2818 | 7.9985 | 9.7936 | 0.0000 | . | . | . | . | . | . | . | . |
| f18 | 5.0289 | 5.4706 | 6.9056 | 10.2206 | 5.4490 | 8.2753 | 7.3232 | 7.2787 | 12.1470 | 7.9580 | 9.1645 | 9.5267 | 8.5387 | 7.3343 | 4.7273 | 6.8611 | 9.3230 | 0.0000 | . | . | . | . | . | . | . |
| f19 | 13.7515 | 12.7867 | 8.8736 | 17.7403 | 10.5563 | 18.0349 | 14.4626 | 8.9572 | 13.5520 | 13.7203 | 11.5607 | 18.6793 | 13.1719 | 12.2427 | 11.3876 | 16.6176 | 15.5598 | 13.2811 | 0.0000 | . | . | . | . | . | . |
| f20 | 8.2001 | 5.6671 | 7.9219 | 9.4949 | 6.7007 | 9.4729 | 12.7718 | 7.6840 | 9.8150 | 6.1061 | 8.0082 | 10.5351 | 6.5049 | 8.3646 | 10.8343 | 8.6001 | 11.7684 | 8.7040 | 10.9592 | 0.0000 | . | . | . | . | . |
| f21 | 17.6479 | 16.2818 | 16.1629 | 17.8112 | 13.6990 | 18.5990 | 18.8156 | 9.2981 | 12.6510 | 18.4722 | 8.7735 | 15.1562 | 17.7683 | 12.8369 | 13.6848 | 18.2172 | 16.8266 | 12.6497 | 11.6174 | 16.6298 | 0.0000 | . | . | . | . |
| f22 | 7.6026 | 6.0645 | 8.7216 | 8.0627 | 5.7289 | 5.6413 | 8.6337 | 6.2427 | 9.2807 | 7.4821 | 7.6624 | 7.4857 | 9.3322 | 7.9331 | 6.6088 | 6.3414 | 10.9611 | 4.2197 | 13.7142 | 7.0818 | 11.9583 | 0.0000 | . | . | . |
| f23 | 10.1094 | 9.3656 | 8.2160 | 10.5553 | 8.4824 | 10.9903 | 11.5433 | 8.1513 | 7.1461 | 9.2532 | 6.6125 | 9.1879 | 10.6666 | 9.9221 | 8.8097 | 9.8610 | 9.9764 | 9.8912 | 12.1481 | 7.1899 | 12.5944 | 7.5045 | 0.0000 | . | . |
| f24 | 7.6184 | 10.6325 | 11.2207 | 13.8764 | 9.5391 | 14.6990 | 13.5341 | 14.9202 | 15.2717 | 8.9961 | 12.3196 | 12.0292 | 13.1459 | 7.9971 | 10.2237 | 10.3618 | 10.4450 | 9.4557 | 10.8146 | 11.5054 | 11.9911 | 12.0146 | 14.1144 | 0.00000 | . |
| f25 | 5.9533 | 6.1713 | 7.7388 | 7.7505 | 7.0194 | 9.1524 | 10.0433 | 7.9041 | 8.8058 | 7.7619 | 5.2087 | 5.9578 | 8.8194 | 7.0315 | 8.1804 | 7.4990 | 8.5750 | 4.8082 | 13.8007 | 6.2298 | 12.5959 | 4.7855 | 7.3121 | 9.38113 | 0 |

# APPENDIX – XIV

## Jaccard coefficient between Floribunda genotypes

| accn | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f23 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 0.33333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 0.50000 | 0.25000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 0.50000 | 0.66667 | 0.75000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 0.66667 | 0.50000 | 0.57143 | 0.83333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 0.75000 | 0.80000 | 0.83333 | 0.66667 | 0.50000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 0.50000 | 0.25000 | 0.40000 | 0.75000 | 0.33333 | 0.60000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 0.75000 | 0.80000 | 0.83333 | 0.66667 | 0.50000 | 0.50000 | 0.60000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 1.00000 | 0.80000 | 0.83333 | 1.00000 | 0.50000 | 0.80000 | 0.83333 | 0.80000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 0.33333 | 0.50000 | 0.80000 | 0.66667 | 0.50000 | 0.50000 | 0.60000 | 0.80000 | 0.80000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 0.80000 | 0.60000 | 0.66667 | 0.75000 | 0.33333 | 0.60000 | 0.40000 | 0.25000 | 0.80000 | 0.83333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 0.75000 | 0.50000 | 0.60000 | 0.66667 | 0.50000 | 0.80000 | 0.60000 | 0.50000 | 0.50000 | 0.80000 | 0.25000 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 0.83333 | 0.66667 | 0.50000 | 1.00000 | 0.42857 | 0.66667 | 0.50000 | 0.85714 | 0.66667 | 0.66667 | 0.71429 | 0.85714 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 0.83333 | 0.85714 | 0.71429 | 1.00000 | 0.42857 | 0.66667 | 0.71429 | 0.66667 | 0.66667 | 0.66667 | 0.71429 | 0.85714 | 0.33333 | 0.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 0.33333 | 0.50000 | 0.80000 | 0.66667 | 0.50000 | 0.80000 | 0.60000 | 0.50000 | 0.80000 | 0.50000 | 0.60000 | 0.50000 | 0.85714 | 0.66667 | 0.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 0.50000 | 0.60000 | 0.66667 | 0.75000 | 0.33333 | 0.25000 | 0.40000 | 0.60000 | 0.83333 | 0.25000 | 0.66667 | 0.83333 | 0.50000 | 0.50000 | 0.60000 | 0.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 0.71429 | 0.57143 | 0.42857 | 0.85714 | 0.14286 | 0.57143 | 0.42857 | 0.57143 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.28571 | 0.28571 | 0.57143 | 0.42857 | 0.00000 | . | . | . | . | . | . | . | . |
| f18 | 0.60000 | 0.40000 | 0.50000 | 0.80000 | 0.16667 | 0.66667 | 0.20000 | 0.40000 | 0.66667 | 0.66667 | 0.20000 | 0.40000 | 0.57143 | 0.57143 | 0.40000 | 0.50000 | 0.28571 | 0.00000 | . | . | . | . | . | . | . |
| f19 | 0.50000 | 0.60000 | 0.40000 | 0.75000 | 0.57143 | 0.83333 | 0.66667 | 0.60000 | 0.83333 | 0.60000 | 0.66667 | 0.60000 | 0.71429 | 0.50000 | 0.25000 | 0.66667 | 0.42857 | 0.50000 | 0.00000 | . | . | . | . | . | . |
| f20 | 0.80000 | 0.83333 | 0.85714 | 0.75000 | 0.33333 | 0.25000 | 0.66667 | 0.25000 | 0.60000 | 0.60000 | 0.40000 | 0.60000 | 0.71429 | 0.50000 | 0.60000 | 0.40000 | 0.42857 | 0.50000 | 0.66667 | 0.0 | . | . | . | . | . |
| f21 | 0.80000 | 0.83333 | 0.66667 | 0.75000 | 0.57143 | 0.60000 | 0.66667 | 0.25000 | 0.83333 | 0.83333 | 0.40000 | 0.60000 | 0.71429 | 0.50000 | 0.60000 | 0.66667 | 0.42857 | 0.50000 | 0.40000 | 0.4 | 0.0 | . | . | . | . |
| f22 | 0.66667 | 0.33333 | 0.50000 | 0.50000 | 0.66667 | 0.75000 | 0.50000 | 0.75000 | 0.75000 | 0.75000 | 0.50000 | 0.33333 | 0.83333 | 1.00000 | 0.75000 | 0.80000 | 0.71429 | 0.60000 | 0.80000 | 0.8 | 0.8 | 0.00000 | . | . | . |
| f23 | 0.66667 | 0.75000 | 0.80000 | 0.50000 | 0.66667 | 0.33333 | 0.80000 | 0.75000 | 0.75000 | 0.33333 | 0.80000 | 0.75000 | 0.83333 | 0.83333 | 0.75000 | 0.50000 | 0.71429 | 0.83333 | 0.80000 | 0.5 | 0.8 | 0.66667 | 0.00000 | . | . |
| f24 | 0.60000 | 0.66667 | 0.50000 | 0.80000 | 0.42857 | 0.66667 | 0.71429 | 0.66667 | 0.66667 | 0.40000 | 0.71429 | 0.66667 | 0.57143 | 0.33333 | 0.40000 | 0.50000 | 0.28571 | 0.57143 | 0.20000 | 0.5 | 0.5 | 0.83333 | 0.80000 | 0.00000 |  |
| f25 | 0.66667 | 0.75000 | 0.80000 | 0.50000 | 0.66667 | 0.33333 | 0.50000 | 0.33333 | 1.00000 | 0.75000 | 0.50000 | 0.75000 | 0.83333 | 0.83333 | 0.75000 | 0.50000 | 0.71429 | 0.80000 | 0.80000 | 0.5 | 0.5 | 0.66667 | 0.66667 | 0.83333 | 0 |

205

206

# APPENDIX – XV

## Dice coefficient between Floribunda genotypes

| acon | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 0.80000 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 0.66667 | 0.85714 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 0.66667 | 0.50000 | 0.40000 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 0.50000 | 0.66667 | 0.60000 | 0.28571 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 0.40000 | 0.33333 | 0.28571 | 0.50000 | 0.66667 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 0.66667 | 0.85714 | 0.75000 | 0.40000 | 0.80000 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 0.40000 | 0.33333 | 0.28571 | 0.50000 | 0.66667 | 0.66667 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 0.00000 | 0.33333 | 0.28571 | 0.00000 | 0.66667 | 0.33333 | 0.28571 | 0.33333 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 0.80000 | 0.66667 | 0.57143 | 0.50000 | 0.66667 | 0.66667 | 0.57143 | 0.33333 | 0.33333 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 0.33333 | 0.57143 | 0.50000 | 0.40000 | 0.80000 | 0.57143 | 0.75000 | 0.85714 | 0.57143 | 0.28571 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 0.40000 | 0.66667 | 0.57143 | 0.50000 | 0.66667 | 0.33333 | 0.57143 | 0.66667 | 0.66667 | 0.33333 | 0.85714 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 0.28571 | 0.50000 | 0.66667 | 0.00000 | 0.72727 | 0.50000 | 0.66667 | 0.25000 | 0.50000 | 0.50000 | 0.44444 | 0.25000 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 0.28571 | 0.25000 | 0.44444 | 0.00000 | 0.72727 | 0.50000 | 0.44444 | 0.50000 | 0.50000 | 0.50000 | 0.44444 | 0.25000 | 0.80000 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 0.80000 | 0.66667 | 0.57143 | 0.50000 | 0.66667 | 0.33333 | 0.57143 | 0.66667 | 0.33333 | 0.66667 | 0.57143 | 0.66667 | 0.25000 | 0.60000 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 0.66667 | 0.57143 | 0.50000 | 0.40000 | 0.80000 | 0.85714 | 0.75000 | 0.57143 | 0.28571 | 0.85714 | 0.50000 | 0.28571 | 0.66667 | 0.66667 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 0.44444 | 0.60000 | 0.72727 | 0.25000 | 0.92308 | 0.60000 | 0.72727 | 0.60000 | 0.60000 | 0.60000 | 0.72727 | 0.60000 | 0.83333 | 0.83333 | 0.60000 | 0.72727 | 1.00000 | . | . | . | . | . | . | . | . |
| f18 | 0.57143 | 0.75000 | 0.66667 | 0.33333 | 0.90909 | 0.50000 | 0.88889 | 0.75000 | 0.50000 | 0.50000 | 0.88889 | 0.75000 | 0.60000 | 0.60000 | 0.75000 | 0.66667 | 0.83333 | 1.00000 | . | . | . | . | . | . | . |
| f19 | 0.66667 | 0.57143 | 0.75000 | 0.40000 | 0.80000 | 0.28571 | 0.50000 | 0.57143 | 0.28571 | 0.57143 | 0.50000 | 0.57143 | 0.44444 | 0.66667 | 0.85714 | 0.50000 | 0.72727 | 0.66667 | 1.00000 | . | . | . | . | . | . |
| f20 | 0.33333 | 0.28571 | 0.25000 | 0.40000 | 0.80000 | 0.85714 | 0.50000 | 0.85714 | 0.57143 | 0.57143 | 0.75000 | 0.57143 | 0.44444 | 0.66667 | 0.57143 | 0.75000 | 0.72727 | 0.66667 | 0.50000 | 1.00000 | . | . | . | . | . |
| f21 | 0.33333 | 0.28571 | 0.50000 | 0.40000 | 0.60000 | 0.57143 | 0.50000 | 0.85714 | 0.28571 | 0.28571 | 0.75000 | 0.57143 | 0.44444 | 0.66667 | 0.57143 | 0.50000 | 0.72727 | 0.66667 | 0.75000 | 0.75000 | 1.00000 | . | . | . | . |
| f22 | 0.50000 | 0.80000 | 0.66667 | 0.66667 | 0.50000 | 0.40000 | 0.66667 | 0.40000 | 0.40000 | 0.40000 | 0.66667 | 0.60000 | 0.28571 | 0.00000 | 0.40000 | 0.33333 | 0.44444 | 0.57143 | 0.33333 | 0.33333 | 0.33333 | 1.00000 | . | . | . |
| f23 | 0.50000 | 0.40000 | 0.33333 | 0.66667 | 0.50000 | 0.80000 | 0.33333 | 0.40000 | 0.40000 | 0.80000 | 0.33333 | 0.40000 | 0.28571 | 0.28571 | 0.40000 | 0.66667 | 0.44444 | 0.28571 | 0.33333 | 0.66667 | 0.33333 | 0.50000 | 1.00000 | . | . |
| f24 | 0.57143 | 0.50000 | 0.66667 | 0.33333 | 0.72727 | 0.50000 | 0.44444 | 0.50000 | 0.50000 | 0.75000 | 0.44444 | 0.50000 | 0.60000 | 0.80000 | 0.75000 | 0.66667 | 0.83333 | 0.60000 | 0.88889 | 0.66667 | 0.66667 | 0.28571 | 0.57143 | 1.00000 | . |
| f25 | 0.50000 | 0.40000 | 0.33333 | 0.66667 | 0.50000 | 0.80000 | 0.66667 | 0.80000 | 0.00000 | 0.40000 | 0.66667 | 0.40000 | 0.28571 | 0.28571 | 0.40000 | 0.66667 | 0.44444 | 0.57143 | 0.33333 | 0.66667 | 0.66667 | 0.50000 | 0.50000 | 0.28571 | 1 |

207

# APPENDIX – XVI

## Hamann's coefficient between Floribunda genotypes

| accn | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 1.00000 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 0.42857 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | -0.14286 | -0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 0.71429 | 0.71429 | 0.14286 | -0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 0.42857 | 0.42857 | -0.14286 | -0.14286 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 0.14286 | 0.14286 | 0.14286 | -0.14286 | 0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 0.42857 | 0.42857 | 0.14286 | -0.14286 | 0.42857 | 0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 0.42857 | 0.42857 | 0.14286 | -0.42857 | 0.14286 | -0.14286 | -0.42857 | -0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 0.71429 | 0.71429 | 0.14286 | -0.42857 | 0.42857 | 0.14286 | -0.14286 | 0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 0.14286 | 0.14286 | -0.42857 | -1.00000 | 0.14286 | -0.14286 | -0.14286 | 0.14286 | -0.14286 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 0.14286 | 0.14286 | 0.14286 | -0.14286 | -0.14286 | 0.14286 | 0.42857 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 0.14286 | 0.14286 | 0.14286 | -0.71429 | 0.14286 | -0.14286 | 0.42857 | 0.42857 | 0.14286 | 0.14286 | 0.14286 | -0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | 0.14286 | 0.14286 | -0.14286 | 0.14286 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 0.71429 | 0.71429 | 0.42857 | 0.14286 | 0.42857 | 0.14286 | 0.14286 | 0.71429 | 0.14286 | 0.42857 | -0.14286 | 0.14286 | 0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 0.42857 | 0.42857 | 0.14286 | -0.14286 | 0.42857 | 0.14286 | 0.42857 | 1.00000 | -0.14286 | 0.14286 | 0.14286 | -0.14286 | 0.42857 | 0.14286 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 0.42857 | 0.42857 | 0.14286 | -0.42857 | 0.71429 | 0.42857 | -0.14286 | 0.14286 | 0.42857 | 0.14286 | -0.14286 | -0.42857 | -0.14286 | -0.14286 | 0.14286 | 0.14286 | 1.00000 | . | . | . | . | . | . | . | . |
| f18 | 0.42857 | 0.42857 | 0.42857 | -0.42857 | 0.71429 | 0.42857 | 0.42857 | 0.42857 | -0.14286 | 0.14286 | -0.14286 | 0.14286 | 0.42857 | 0.14286 | 0.42857 | 0.42857 | 0.42857 | 1.00000 | . | . | . | . | . | . | . |
| f19 | 0.14286 | 0.14286 | -0.42857 | -0.42857 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | -0.14286 | 0.42857 | -0.14286 | -0.42857 | -0.42857 | -0.14286 | -0.14286 | -0.42857 | -0.42857 | -0.42857 | 1.00000 | . | . | . | . | . | . |
| f20 | 0.71429 | 0.71429 | 0.14286 | -0.42857 | 1.00000 | 0.71429 | 0.14286 | 0.42857 | 0.14286 | 0.42857 | 0.14286 | -0.14286 | 0.14286 | 0.14286 | 0.42857 | 0.42857 | 0.71429 | 0.71429 | -0.14286 | 1.00000 | . | . | . | . | . |
| f21 | -0.14286 | -0.14286 | -0.71429 | -0.71429 | -0.14286 | -0.14286 | -0.14286 | -0.14286 | -0.71429 | -0.42857 | 0.14286 | -0.14286 | -0.42857 | -0.42857 | -0.42857 | -0.14286 | -0.42857 | -0.42857 | 0.14286 | -0.14286 | 1.00000 | . | . | . | . |
| f22 | 0.42857 | 0.42857 | 0.42857 | 0.14286 | 0.14286 | 0.42857 | 0.71429 | 0.14286 | -0.14286 | 0.14286 | -0.42857 | 0.71429 | 0.14286 | -0.14286 | 0.42857 | 0.14286 | -0.14286 | 0.42857 | -0.14286 | 0.14286 | -0.42857 | 1.00000 | . | . | . |
| f23 | 0.14286 | 0.14286 | -0.14286 | -0.42857 | -0.14286 | -0.42857 | -0.42857 | 0.14286 | -0.14286 | 0.42857 | 0.42857 | 0.14286 | -0.14286 | -0.14286 | 0.42857 | 0.14286 | -0.42857 | -0.14286 | -0.14286 | -0.14286 | -0.42857 | -0.14286 | 1.00000 | . | . |
| f24 | 0.14286 | 0.14286 | -0.14286 | 0.14286 | -0.14286 | -0.14286 | -0.14286 | 0.14286 | -0.42857 | -0.14286 | -0.71429 | -0.14286 | -0.42857 | 0.14286 | 0.42857 | 0.14286 | -0.42857 | -0.14286 | 0.42857 | -0.14286 | 0.14286 | 0.14286 | -0.14286 | 1.00000 | . |
| f25 | 0.14286 | 0.14286 | -0.14286 | -0.42857 | 0.14286 | -0.14286 | 0.14286 | 0.71429 | -0.14286 | 0.42857 | 0.42857 | -0.42857 | 0.42857 | 0.14286 | 0.42857 | 0.71429 | -0.14286 | 0.14286 | -0.14286 | 0.14286 | -0.42857 | -0.14286 | 0.42857 | -0.14286 | 1 |

# APPENDIX – XVII

## Simple matching coefficient between Floribunda genotypes

| acen | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 1.00000 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 0.71429 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 0.42857 | 0.42857 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 0.85714 | 0.85714 | 0.57143 | 0.28571 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 0.71429 | 0.71429 | 0.42857 | 0.42857 | 0.85714 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 0.57143 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 0.71429 | 0.71429 | 0.57143 | 0.42857 | 0.71429 | 0.57143 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 0.71429 | 0.71429 | 0.57143 | 0.28571 | 0.57143 | 0.42857 | 0.28571 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 0.85714 | 0.85714 | 0.57143 | 0.28571 | 0.71429 | 0.57143 | 0.42857 | 0.57143 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 0.57143 | 0.57143 | 0.28571 | 0.00000 | 0.57143 | 0.42857 | 0.42857 | 0.57143 | 0.42857 | 0.71429 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 0.57143 | 0.57143 | 0.57143 | 0.42857 | 0.42857 | 0.57143 | 0.71429 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 0.57143 | 0.57143 | 0.57143 | 0.14286 | 0.57143 | 0.42857 | 0.71429 | 0.71429 | 0.57143 | 0.57143 | 0.57143 | 0.42857 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 0.85714 | 0.85714 | 0.71429 | 0.57143 | 0.71429 | 0.57143 | 0.57143 | 0.85714 | 0.57143 | 0.71429 | 0.42857 | 0.57143 | 0.57143 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 0.71429 | 0.71429 | 0.57143 | 0.42857 | 0.71429 | 0.57143 | 0.71429 | 1.00000 | 0.42857 | 0.57143 | 0.57143 | 0.42857 | 0.71429 | 0.57143 | 0.85714 | 1.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 0.71429 | 0.71429 | 0.57143 | 0.28571 | 0.85714 | 0.71429 | 0.42857 | 0.57143 | 0.71429 | 0.57143 | 0.42857 | 0.28571 | 0.42857 | 0.42857 | 0.57143 | 0.57143 | 1.00000 | . | . | . | . | . | . | . | . |
| f18 | 0.71429 | 0.71429 | 0.71429 | 0.28571 | 0.85714 | 0.71429 | 0.71429 | 0.71429 | 0.42857 | 0.57143 | 0.42857 | 0.57143 | 0.71429 | 0.57143 | 0.71429 | 0.71429 | 0.71429 | 1.00000 | . | . | . | . | . | . | . |
| f19 | 0.57143 | 0.57143 | 0.28571 | 0.28571 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.42857 | 0.71429 | 0.42857 | 0.28571 | 0.28571 | 0.42857 | 0.42857 | 0.28571 | 0.28571 | 0.28571 | 1.00000 | . | . | . | . | . | . |
| f20 | 0.85714 | 0.85714 | 0.57143 | 0.28571 | 1.00000 | 0.85714 | 0.57143 | 0.71429 | 0.57143 | 0.71429 | 0.57143 | 0.42857 | 0.57143 | 0.57143 | 0.71429 | 0.71429 | 0.85714 | 0.85714 | 0.42857 | 1.00000 | . | . | . | . | . |
| f21 | 0.42857 | 0.42857 | 0.14286 | 0.14286 | 0.42857 | 0.42857 | 0.42857 | 0.42857 | 0.14286 | 0.28571 | 0.57143 | 0.42857 | 0.28571 | 0.28571 | 0.28571 | 0.42857 | 0.28571 | 0.28571 | 0.57143 | 0.42857 | 1.00000 | . | . | . | . |
| f22 | 0.71429 | 0.71429 | 0.71429 | 0.57143 | 0.57143 | 0.71429 | 0.85714 | 0.57143 | 0.42857 | 0.57143 | 0.28571 | 0.85714 | 0.57143 | 0.42857 | 0.71429 | 0.57143 | 0.42857 | 0.71429 | 0.42857 | 0.57143 | 0.28571 | 1.00000 | . | . | . |
| f23 | 0.57143 | 0.57143 | 0.42857 | 0.28571 | 0.42857 | 0.28571 | 0.28571 | 0.57143 | 0.42857 | 0.71429 | 0.71429 | 0.57143 | 0.42857 | 0.42857 | 0.71429 | 0.57143 | 0.28571 | 0.42857 | 0.42857 | 0.42857 | 0.28571 | 0.42857 | 1.00000 | . | . |
| f24 | 0.57143 | 0.57143 | 0.42857 | 0.57143 | 0.42857 | 0.42857 | 0.42857 | 0.57143 | 0.28571 | 0.42857 | 0.14286 | 0.42857 | 0.28571 | 0.57143 | 0.71429 | 0.57143 | 0.28571 | 0.42857 | 0.71429 | 0.42857 | 0.57143 | 0.57143 | 0.42857 | 1.00000 | . |
| f25 | 0.57143 | 0.57143 | 0.42857 | 0.28571 | 0.57143 | 0.42857 | 0.57143 | 0.85714 | 0.42857 | 0.71429 | 0.71429 | 0.28571 | 0.71429 | 0.57143 | 0.71429 | 0.85714 | 0.42857 | 0.57143 | 0.42857 | 0.57143 | 0.28571 | 0.42857 | 0.71429 | 0.42857 | 1 |

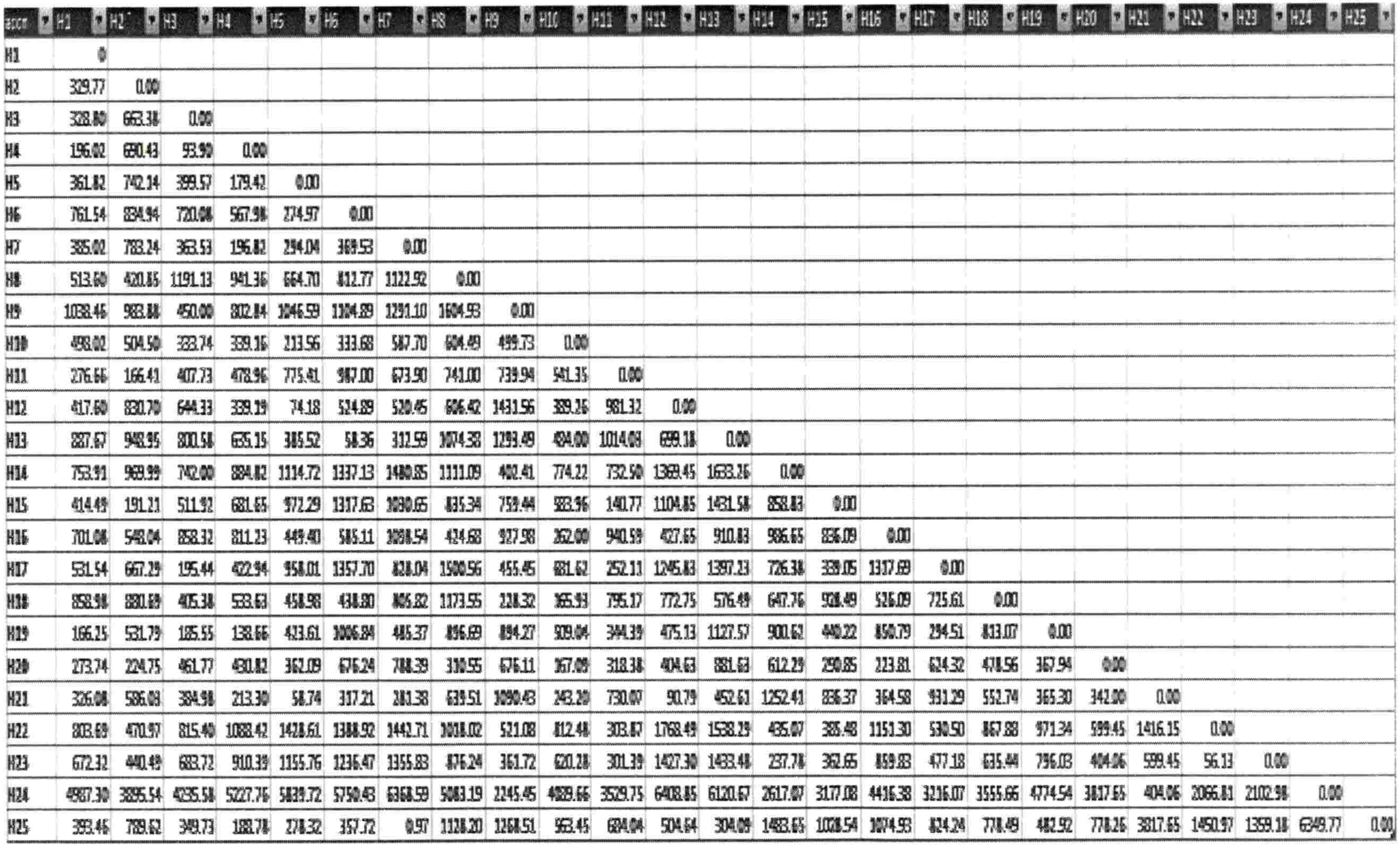
# APPENDIX – XVIII

## Gower's measure between Floribunda genotypes

| scon | f1 | f2 | f3 | f4 | f5 | f6 | f7 | f8 | f9 | f10 | f11 | f12 | f13 | f14 | f15 | f16 | f17 | f18 | f19 | f20 | f21 | f22 | f23 | f24 | f25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| f1 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f2 | 0.79811 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f3 | 0.74003 | 0.84926 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f4 | 0.84545 | 0.75090 | 0.66036 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f5 | 0.68005 | 0.82507 | 0.75581 | 0.58457 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f6 | 0.73642 | 0.66594 | 0.54234 | 0.79132 | 0.71942 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f7 | 0.65547 | 0.74690 | 0.76638 | 0.58580 | 0.80648 | 0.70188 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f8 | 0.66210 | 0.64581 | 0.65374 | 0.62863 | 0.75878 | 0.76719 | 0.68570 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f9 | 0.54206 | 0.65322 | 0.65558 | 0.53635 | 0.61364 | 0.45999 | 0.50334 | 0.58696 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f10 | 0.92432 | 0.78243 | 0.70639 | 0.84038 | 0.67274 | 0.74371 | 0.61426 | 0.63876 | 0.57262 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f11 | 0.57790 | 0.70390 | 0.65751 | 0.59969 | 0.81635 | 0.65508 | 0.65337 | 0.74962 | 0.65554 | 0.57186 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f12 | 0.74766 | 0.79757 | 0.70302 | 0.80230 | 0.67301 | 0.69681 | 0.65211 | 0.53903 | 0.64907 | 0.73450 | 0.75155 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . | . |
| f13 | 0.54021 | 0.71469 | 0.65596 | 0.47711 | 0.75238 | 0.58909 | 0.69678 | 0.61470 | 0.62522 | 0.53868 | 0.68837 | 0.53179 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . | . |
| f14 | 0.62803 | 0.57270 | 0.53726 | 0.53162 | 0.67770 | 0.63234 | 0.54835 | 0.68218 | 0.60171 | 0.64312 | 0.62145 | 0.48458 | 0.73837 | 1.00000 | . | . | . | . | . | . | . | . | . | . | . |
| f15 | 0.83203 | 0.76085 | 0.79455 | 0.74870 | 0.68202 | 0.68115 | 0.71521 | 0.73328 | 0.57611 | 0.80941 | 0.68046 | 0.65800 | 0.50148 | 0.83515 | 1.00000 | . | . | . | . | . | . | . | . | . | . |
| f16 | 0.72928 | 0.71469 | 0.58081 | 0.69789 | 0.76883 | 0.80431 | 0.68102 | 0.76183 | 0.50378 | 0.76877 | 0.68961 | 0.61387 | 0.62186 | 0.75336 | 0.74946 | 1.00000 | . | . | . | . | . | . | . | . | . |
| f17 | 0.81864 | 0.68904 | 0.70468 | 0.52275 | 0.76395 | 0.61448 | 0.77175 | 0.84728 | 0.84275 | 0.59942 | 0.72105 | 0.65717 | 0.82753 | 0.62382 | 0.66668 | 0.70044 | 1.00000 | . | . | . | . | . | . | . | . |
| f18 | 0.71829 | 0.79594 | 0.76407 | 0.59607 | 0.88613 | 0.72622 | 0.83541 | 0.74996 | 0.57198 | 0.65956 | 0.78575 | 0.69018 | 0.73014 | 0.67700 | 0.72302 | 0.76696 | 0.79828 | 1.00000 | . | . | . | . | . | . | . |
| f19 | 0.69236 | 0.63209 | 0.72998 | 0.59716 | 0.59683 | 0.51027 | 0.52019 | 0.71992 | 0.54653 | 0.69632 | 0.56908 | 0.48984 | 0.48721 | 0.56968 | 0.75096 | 0.54974 | 0.60386 | 0.54149 | 1.00000 | . | . | . | . | . | . |
| f20 | 0.72484 | 0.70329 | 0.65289 | 0.70462 | 0.76254 | 0.78511 | 0.62565 | 0.83264 | 0.54090 | 0.76910 | 0.74080 | 0.59082 | 0.68840 | 0.73040 | 0.66567 | 0.80216 | 0.65454 | 0.72135 | 0.68113 | 1.00000 | . | . | . | . | . |
| f21 | 0.52478 | 0.47542 | 0.48291 | 0.62806 | 0.61402 | 0.66029 | 0.50368 | 0.80288 | 0.49320 | 0.50970 | 0.73030 | 0.50412 | 0.44914 | 0.63943 | 0.60953 | 0.66243 | 0.55488 | 0.63019 | 0.65464 | 0.66822 | 1.00000 | . | . | . | . |
| f22 | 0.73968 | 0.86038 | 0.80625 | 0.72827 | 0.76890 | 0.70024 | 0.71831 | 0.69632 | 0.72197 | 0.74477 | 0.74206 | 0.81908 | 0.62596 | 0.57738 | 0.76010 | 0.68686 | 0.68037 | 0.82144 | 0.61602 | 0.67883 | 0.57792 | 1.00000 | . | . | . |
| f23 | 0.75401 | 0.89056 | 0.73114 | 0.75288 | 0.62945 | 0.65689 | 0.58911 | 0.73968 | 0.68035 | 0.76984 | 0.68708 | 0.70380 | 0.49840 | 0.80055 | 0.79317 | 0.67375 | 0.60958 | 0.60122 | 0.73254 | 0.74558 | 0.65502 | 0.74103 | 1.00000 | . | . |
| f24 | 0.84179 | 0.69233 | 0.67787 | 0.69687 | 0.62814 | 0.59635 | 0.54992 | 0.59156 | 0.50578 | 0.81313 | 0.55167 | 0.64595 | 0.45573 | 0.66232 | 0.78254 | 0.69678 | 0.61370 | 0.63465 | 0.78198 | 0.86081 | 0.64319 | 0.65630 | 0.68000 | 1.00000 | . |
| f25 | 0.77336 | 0.89238 | 0.65773 | 0.73891 | 0.76101 | 0.79632 | 0.68425 | 0.82639 | 0.56787 | 0.73954 | 0.79693 | 0.68887 | 0.83875 | 0.76247 | 0.72288 | 0.82858 | 0.72834 | 0.80417 | 0.81615 | 0.88821 | 0.73025 | 0.72984 | 0.74818 | 0.71003 | 1 |

# APPENDIX – XIX

## D2 distance between Hybrid Tea genotypes



| accn | H1 | H2 | H3 | H4 | H5 | H6 | H7 | H8 | H9 | H10 | H11 | H12 | H13 | H14 | H15 | H16 | H17 | H18 | H19 | H20 | H21 | H22 | H23 | H24 | H25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| H1 | 0 | | | | | | | | | | | | | | | | | | | | | | | | |
| H2 | 329.77 | 0.00 | | | | | | | | | | | | | | | | | | | | | | | |
| H3 | 328.80 | 663.38 | 0.00 | | | | | | | | | | | | | | | | | | | | | | |
| H4 | 196.02 | 690.43 | 93.90 | 0.00 | | | | | | | | | | | | | | | | | | | | | |
| H5 | 361.82 | 742.14 | 399.57 | 179.42 | 0.00 | | | | | | | | | | | | | | | | | | | | |
| H6 | 761.54 | 834.94 | 720.08 | 567.98 | 274.97 | 0.00 | | | | | | | | | | | | | | | | | | | |
| H7 | 385.02 | 783.24 | 363.53 | 196.82 | 294.04 | 369.53 | 0.00 | | | | | | | | | | | | | | | | | | |
| H8 | 513.60 | 420.85 | 1191.13 | 941.36 | 664.70 | 812.77 | 1122.92 | 0.00 | | | | | | | | | | | | | | | | | |
| H9 | 1038.46 | 983.88 | 450.00 | 802.84 | 1046.59 | 1104.89 | 1291.10 | 1604.93 | 0.00 | | | | | | | | | | | | | | | | |
| H10 | 498.02 | 504.50 | 323.74 | 339.16 | 213.56 | 333.68 | 587.70 | 604.49 | 499.73 | 0.00 | | | | | | | | | | | | | | | |
| H11 | 276.66 | 166.41 | 407.73 | 478.96 | 775.41 | 987.00 | 673.90 | 741.00 | 739.94 | 541.35 | 0.00 | | | | | | | | | | | | | | |
| H12 | 417.60 | 830.70 | 644.33 | 339.19 | 74.18 | 524.89 | 520.45 | 606.42 | 1431.56 | 389.26 | 981.32 | 0.00 | | | | | | | | | | | | | |
| H13 | 887.67 | 948.95 | 800.58 | 635.15 | 385.52 | 58.36 | 312.59 | 1074.38 | 1293.49 | 494.00 | 1014.08 | 699.18 | 0.00 | | | | | | | | | | | | |
| H14 | 753.91 | 969.99 | 742.00 | 884.82 | 1114.72 | 1337.13 | 1480.85 | 1111.09 | 402.41 | 774.22 | 732.50 | 1369.45 | 1633.26 | 0.00 | | | | | | | | | | | |
| H15 | 414.49 | 191.21 | 511.92 | 681.65 | 972.29 | 1317.63 | 1090.65 | 835.34 | 758.44 | 923.96 | 140.77 | 1104.85 | 1431.58 | 858.83 | 0.00 | | | | | | | | | | |
| H16 | 701.06 | 548.04 | 858.32 | 811.23 | 449.40 | 585.11 | 1098.54 | 424.68 | 927.98 | 262.00 | 940.59 | 427.65 | 910.83 | 986.65 | 836.09 | 0.00 | | | | | | | | | |
| H17 | 531.54 | 667.29 | 195.44 | 422.94 | 958.01 | 1357.70 | 828.04 | 1500.56 | 455.45 | 681.62 | 252.11 | 1245.83 | 1397.23 | 726.38 | 339.05 | 1317.69 | 0.00 | | | | | | | | |
| H18 | 858.98 | 880.69 | 405.38 | 533.63 | 458.98 | 438.80 | 805.82 | 1173.55 | 228.32 | 365.93 | 795.17 | 772.75 | 576.49 | 647.76 | 928.49 | 526.09 | 725.61 | 0.00 | | | | | | | |
| H19 | 166.25 | 531.79 | 185.55 | 138.66 | 423.61 | 1006.84 | 485.37 | 896.69 | 894.27 | 909.04 | 344.39 | 475.13 | 1127.57 | 900.62 | 440.22 | 850.79 | 294.51 | 813.07 | 0.00 | | | | | | |
| H20 | 273.74 | 224.75 | 461.77 | 430.82 | 362.09 | 676.24 | 788.39 | 330.55 | 676.11 | 167.09 | 318.38 | 404.63 | 881.63 | 612.29 | 290.85 | 223.81 | 624.32 | 478.56 | 367.94 | 0.00 | | | | | |
| H21 | 326.08 | 586.03 | 384.98 | 213.30 | 58.74 | 317.21 | 281.38 | 639.51 | 1090.43 | 243.20 | 730.07 | 90.79 | 452.61 | 1252.41 | 836.37 | 364.58 | 931.29 | 552.74 | 365.30 | 342.00 | 0.00 | | | | |
| H22 | 803.69 | 470.97 | 815.40 | 1088.42 | 1428.61 | 1388.92 | 1442.71 | 1018.02 | 521.08 | 812.48 | 303.87 | 1768.49 | 1538.29 | 435.07 | 385.48 | 1151.30 | 530.50 | 867.88 | 971.34 | 599.45 | 1416.15 | 0.00 | | | |
| H23 | 672.32 | 440.49 | 683.72 | 910.39 | 1155.76 | 1236.47 | 1355.83 | 876.24 | 361.72 | 620.28 | 301.39 | 1427.30 | 1433.48 | 237.78 | 362.65 | 859.83 | 477.18 | 635.44 | 796.03 | 404.06 | 599.45 | 56.13 | 0.00 | | |
| H24 | 4987.30 | 3895.54 | 4235.58 | 5227.76 | 5839.72 | 5750.43 | 6368.59 | 5083.19 | 2245.45 | 4829.66 | 3529.75 | 6408.85 | 6120.67 | 2617.07 | 3177.08 | 4416.38 | 3216.07 | 3555.66 | 4774.54 | 3817.65 | 404.06 | 2066.81 | 2102.98 | 0.00 | |
| H25 | 393.46 | 789.62 | 349.73 | 188.78 | 278.32 | 357.72 | 0.97 | 1128.20 | 1268.51 | 963.45 | 684.04 | 504.64 | 304.09 | 1483.65 | 1028.54 | 1074.93 | 824.24 | 778.49 | 482.92 | 778.26 | 3817.65 | 1450.97 | 1359.18 | 6349.77 | 0.00 |

# APPENDIX – XX

## D2 distance between Floribunda genotypes

| accn | F1 | F2 | F3 | F4 | F5 | F6 | F7 | F8 | F9 | F10 | F11 | F12 | F13 | F14 | F15 | F16 | F17 | F18 | F19 | F20 | F21 | F22 | F23 | F24 | F25 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| F1 | 0.00 | | | | | | | | | | | | | | | | | | | | | | | | |
| F2 | 237.09 | 0.00 | | | | | | | | | | | | | | | | | | | | | | | |
| F3 | 396.94 | 740.42 | 0.00 | | | | | | | | | | | | | | | | | | | | | | |
| F4 | 204.00 | 183.83 | 953.47 | 0.00 | | | | | | | | | | | | | | | | | | | | | |
| F5 | 242.00 | 106.09 | 436.89 | 361.52 | 0.00 | | | | | | | | | | | | | | | | | | | | |
| F6 | 809.97 | 422.05 | 1211.54 | 673.34 | 408.52 | 0.00 | | | | | | | | | | | | | | | | | | | |
| F7 | 881.80 | 887.04 | 386.29 | 1235.60 | 577.30 | 676.26 | 0.00 | | | | | | | | | | | | | | | | | | |
| F8 | 519.01 | 381.37 | 379.67 | 707.02 | 150.15 | 450.98 | 455.33 | 0.00 | | | | | | | | | | | | | | | | | |
| F9 | 323.81 | 502.62 | 301.39 | 563.90 | 250.52 | 798.34 | 704.73 | 172.11 | 0.00 | | | | | | | | | | | | | | | | |
| F10 | 75.84 | 421.36 | 590.20 | 224.94 | 455.52 | 1288.14 | 1270.86 | 796.19 | 448.49 | 0.00 | | | | | | | | | | | | | | | |
| F11 | 349.25 | 314.76 | 780.22 | 352.27 | 262.66 | 788.22 | 1287.43 | 376.50 | 218.40 | 452.04 | 0.00 | | | | | | | | | | | | | | |
| F12 | 260.13 | 297.39 | 784.69 | 126.15 | 305.77 | 490.64 | 938.82 | 523.80 | 339.36 | 359.40 | 256.27 | 0.00 | | | | | | | | | | | | | |
| F13 | 513.01 | 170.95 | 705.42 | 562.64 | 258.67 | 710.91 | 815.20 | 481.02 | 767.05 | 742.12 | 650.49 | 735.57 | 0.00 | | | | | | | | | | | | |
| F14 | 499.61 | 453.74 | 864.63 | 793.70 | 405.02 | 992.98 | 1439.56 | 515.21 | 510.26 | 784.91 | 308.56 | 751.62 | 760.22 | 0.00 | | | | | | | | | | | |
| F15 | 721.61 | 873.49 | 326.06 | 1216.58 | 463.94 | 739.99 | 313.65 | 263.79 | 340.07 | 1117.30 | 824.00 | 807.74 | 1068.06 | 748.60 | 0.00 | | | | | | | | | | |
| F16 | 607.98 | 270.23 | 1434.40 | 404.42 | 384.20 | 460.45 | 1547.42 | 661.84 | 724.22 | 824.78 | 351.28 | 431.84 | 785.83 | 415.25 | 1129.06 | 0.00 | | | | | | | | | |
| F17 | 477.98 | 800.61 | 278.29 | 985.05 | 477.09 | 950.70 | 560.39 | 420.01 | 280.69 | 793.40 | 593.74 | 602.57 | 1059.77 | 561.78 | 143.13 | 1045.73 | 0.00 | | | | | | | | |
| F18 | 605.77 | 306.60 | 762.46 | 718.11 | 302.80 | 219.36 | 606.23 | 248.24 | 502.95 | 1030.82 | 488.95 | 516.85 | 522.58 | 389.47 | 381.25 | 355.52 | 463.56 | 0.00 | | | | | | | |
| F19 | 835.66 | 1484.13 | 452.91 | 1613.30 | 1077.04 | 2579.36 | 1601.67 | 1063.44 | 688.80 | 833.43 | 1052.48 | 1509.70 | 1476.95 | 1166.11 | 1119.02 | 2204.57 | 848.40 | 1770.87 | 0.00 | | | | | | |
| F20 | 203.97 | 324.10 | 657.62 | 286.31 | 409.18 | 1316.17 | 1382.40 | 751.31 | 528.83 | 156.86 | 346.46 | 472.13 | 422.45 | 718.28 | 1318.76 | 842.46 | 966.71 | 996.08 | 829.57 | 0.00 | | | | | |
| F21 | 1569.29 | 1337.26 | 1496.44 | 1824.42 | 904.85 | 1191.42 | 1714.89 | 740.34 | 874.79 | 2013.53 | 790.79 | 1313.28 | 1682.83 | 744.09 | 844.38 | 1091.02 | 903.33 | 724.30 | 1785.08 | 1861.46 | 0.00 | | | | |
| F22 | 518.45 | 249.98 | 705.05 | 496.97 | 144.55 | 154.95 | 549.60 | 175.50 | 334.59 | 826.67 | 350.41 | 292.61 | 500.58 | 544.52 | 420.60 | 317.81 | 498.71 | 82.12 | 1700.59 | 834.86 | 840.79 | 0.00 | | | |
| F23 | 480.97 | 781.04 | 167.87 | 844.36 | 475.06 | 1027.02 | 535.60 | 294.76 | 117.94 | 658.28 | 488.24 | 533.97 | 905.25 | 811.39 | 309.12 | 1234.73 | 203.45 | 697.00 | 632.42 | 712.25 | 1134.95 | 541.35 | 0.00 | | |
| F24 | 476.76 | 858.31 | 793.46 | 949.16 | 645.43 | 1707.16 | 1735.86 | 921.34 | 561.08 | 587.76 | 449.19 | 811.70 | 1243.54 | 400.79 | 968.43 | 974.07 | 564.46 | 986.97 | 580.88 | 632.29 | 939.55 | 1045.78 | 790.58 | 0.00 | |
| F25 | 153.50 | 191.75 | 557.96 | 235.60 | 178.38 | 655.89 | 962.26 | 378.09 | 224.98 | 296.34 | 92.84 | 161.35 | 503.73 | 291.84 | 651.68 | 321.89 | 402.13 | 345.37 | 1071.52 | 283.88 | 1049.10 | 255.02 | 426.23 | 451.27 | 0.00 |

# MULTIVARIATE CLUSTERING TECHNIQUES – A COMPARISON BASED ON ROSE (*Rosa* spp.)

*by*

**ARYA V CHANDRAN**
**(2016-19-002)**

**Abstract of the thesis**

**Submitted in partial fulfilment of the**
**requirement for the degree of**

## MASTER OF SCIENCE IN AGRICULTURE

**Faculty of Agriculture**
**Kerala Agricultural University**

**DEPARTMENT OF AGRICULTURAL STATISTICS**
**COLLEGE OF AGRICULTURE**
**VELLAYANI, THIRUVANANTHAPURAM – 695 522**
**KERALA, INDIA**
**2018**

# ABSTRACT

The study entitled "Multivariate clustering techniques – a comparison based on rose (*Rosa* spp.)" was undertaken to compare different clustering techniques, to identify the suitable technique for different types of qualitative and quantitative data and to illustrate the procedures using data based on a field experiment on rose (*Rosa* spp.). Data on quantitative and qualitative traits collected from a field experiment on "Characterization and genetic improvement in Rose (*Rosa* spp.) through mutagenesis" done during 2014-2017 at College of Agriculture, Vellayani and Regional Agriculture Research Station (RARS), Ambalavayal, Wayanad was used for the study. Twenty five cultivars each coming under the Hybrid Tea and Floribunda groups of rose were evaluated for the study. There were nine quantitative characters and three qualitative characters. Statistical studies were carried out with the help of statistical packages SPSS, STATA, SAS, R and NTSYS.

Preliminary statistical analysis by applying Analysis of variance (ANOVA) for all quantitative characters under study revealed significant difference among different genotypes with respect to each character. Multivariate analysis of variance (MANOVA) was carried out to test the significance of varietal means for each group. The results indicated difference among the cultivar means for both groups with respect to all quantitative characters.

Linear discriminant function developed using nine quantitative characters for each of the groups were used to elucidate the differences between them. The average score obtained was 11.01 for the Hybrid Tea type and – 2.34 for Floribunda type with an overall average of 4.38. Discriminant function analysis reassured the difference between the two groups under study.

Cluster analysis on Hybrid Tea type and Floribunda type were performed for quantitative, qualitative and mixed data. Association measures used were Euclidean distance, Squared Euclidean, Chebychev distance, City Block distance and Mahalanobis $D^2$ for quantitative data, Jaccard, Dice, Simple matching and Hamann's coefficient for qualitative data and Gower's measure for mixed data. Different methods such as single linkage, complete linkage, Unweighted Pair Group Average Method (UPGMA), Weighted Pair Group Average Method (WPGMA), Unweighted Pair Group Centroid Method (UPGMC), Ward's method, modified Tocher method, k means clustering and Principal Component Analysis (PCA) were adopted for the clustering of cultivars. Optimum numbers of clusters were determined by Pseudo $t^2$ statistics for hierarchical clustering and by Pesudo F statistics for k means clustering. SD ( Scatterness- Distance) index was used to test validity of clustering based on quantitative data.

Clustering based on qualitative data was carried out using seven characters, three of which are qualitative traits and all others are quantitative characters converted to qualitative traits. Jaccard and Dice coefficient were used for binary data while Simple matching and Hamann's were used for multi-state data. The result of different clustering techniques based on Squared Euclidean distance gave approximately the same result as that of Euclidean distance. The Jaccard and Dice coefficients were found to be very similar, so that there was no difference in topology of dendrogram but only in branch length. Clustering pattern under Simple matching and Hamann's coefficient provided were of similar type.

For both groups among all the clustering methods, single linkage clustering under different distance measures tends to create a set of one or two clusters including majority of the genotypes and the remaining genotypes are single or two member clusters. Single linkage clustering tends to produce long chain types clusters as opposed to bunched clusters. On the other hand, the single linkage algorithm suffers chaining effect. Among other clustering algorithms, complete linkage method and Ward's clustering method showed similar results under Squared Euclidean distance. UPGMA, WPGMA and UPGMC methods under Squared Euclidean method gave comparable results. Clustering using UPGMA and WPGMA method gives almost same clustering pattern under different distance measures for qualitative and quantitative data. Results obtained from k means clustering are comparable with results obtained from hierarchical clustering except for single linkage clustering. A certain degree of similarity was observed between k means and $D^2$ analysis but not to up that between other clustering methods.

Under Hybrid Tea genotypes, H16 (Mary Jean) formed a single cluster under single linkage method using different distance measures for quantitative, qualitative and mixed data analysis. Under complete linkage method H7 (Alaine Souchen) and H25 (Josepha) came under same cluster, in clustering based on quantitative and qualitative characters. H22 (Mom's Rose) and H23 (Lois Wilson) came under same cluster in clustering based on complete linkage, UPGMA and WPGMA except under Hamann's coefficient. These came under the same cluster under $D^2$ analysis also. Among Floribunda genotypes F2 (Tickled Pink) and F5 (Princess de Monaco) were included in the same cluster under UPGMA method for both quantitative and qualitative data. F1 (Versailles) and F24 (Golden Fairy) also came under the same cluster except for multistage distances under UPGMA.

Clustering based on mixed data gave approximately the same results as that of quantitative data under different clustering algorithms except for single linkage clustering. Comparison using SD index indicated high index value for clustering based on Gower's measure.

Comparison among single linkage, complete linkage and Average linkage under different association measures using SD index were carried out. Average linkage method under Squared Euclidean was found to be the best for both type with SD index 0.651 for Hybrid Tea and 0.659 for Floribunda type.

Clustering pattern observed from score plot of PCA is comparable with the pattern obtained from quantitative data especially with $D^2$ analysis. Contribution of characters towards variance obtained $D^2$ analysis and PCA showed similar results.

From the study it is possible to compare different methods and exclude inappropriate methods. Groups formed from modified Tocher method and PCA are different from other methods. SD index indicated that UPGMA under Squared Euclidean distance is the best for quantitative data.