TAXONOMIC CHARACTERS

Concept of characters

Any expressed attribute of the organism that can be measured, counted or otherwise assessed may be called a 'character'. Devis and Heywood (1963) have attempted to give a general definition of character, which is considered as the most suitable definition of character. According to them a character is 'Any attribute of form, structure, physiology or behavior which the taxonomist separates from the whole organism for a particular purpose such a comparison, identification or interpretation'.

The terms 'character', 'character-state' and 'characteristics' have often been used loosely or interchangeably. Strictly in true sense, they denote different things. Character is the particular attribute that the taxonomists consider for classification purposes, whereas the various patterns of their expressions are called character-states. Thus, the same character may have several character-states. For example leaf arrangement is a character while it's phyllotaxy like alternate, opposite and whorled are character-states of leaf arrangement. Thus, ovary position is a character, while it is hypogynous, epigynous and perigynous are character-states of leaf arrangement. It may therefore be said that for practical purposes a character may be defined as any feature whose expression can be measured, counted or otherwise assessed.

When a particular character-state is exclusive to a particular taxon in a given assemblage, it is said to be 'characteristic' of the group. For example, pitchers are 'characteristic' of the insectivorous genera like *Nepenthes*, *Darlingtonia*, *Sarracenia*, *Heliamphora* and *Cephalotus*. Bladders are 'characteristic' of the insectivorous taxon *Urticularia*, pappus is characteristic of the family Asteraceae, gynobasic style is characteristic of the family Lamiaceae, papilionaceous corolla and vexillum aestivation are characteristics of the sub-family Papilionoideae, cruciferous corolla and siliqua fruit are characteristics of the family Brassicaceae, etc.

Unit character

Sneath and Sokal (1962) defined 'unit character' as a taxonomic character of two or more

states, which within the study at hand cannot be further divided logically. But in practical purposes, it is rather difficult to divide characters into units. In contrast, 'multiple characters' can be broken down into unit characters. For example, leaf length, flower diameter, seed number, etc are unit characters; but leaf shape can be subdivided according to the length-breadth ratio and hence is not a unit character.

In numerical taxonomy only 'unit characters' can be used.

Analytical and Synthetic characters

Analysis and synthesis are two aspects of taxonomic procedures. On the other hand, in taxonomic procedure, characters may be employed for two main activities: (i) identification, characterization and delimitation, and (ii) classification of these species into higher taxa. Corresponding with these activities, characters may be regarded as 'analytic' and 'synthetic'.

Analytical characters are mostly diagnostic ones and are useful in identification, characterization and delimitation of lower taxa. On the other hand, synthetic characters are useful in grouping these taxa into higher groups. Synthetic characters gradually decrease in number as one goes from lower to higher groups. This is obvious because only a few characters become increasingly constant at higher position of the group in the hierarchy. Analytical characters of a given may turn out to be synthetic characters in another. For example, parallel venation is the analytical character of Angiosperms but synthetic character for Monocotyledone. Likewise, Head inflorescence is the analytical character of Rubiaceae but synthetic character for Asteraceae, leguminous fruit is the analytical character of Polypetale but synthetic character for Fabaceae, etc.

The characters like sympetalous vs. polypetalous condition of calyx and corolla, superior vs. inferior gynoecium, etc are quite constant and do not vary within smaller groups or within populatios of the same species. Such characters are referred as '***constitutive***' o r '***organizational***' characters.

Qualitative and Quantitative characters

The characters relating to form and structure are termed as 'qualitative' characters, e.g. leaf shape and arrangement, flower color, relative length of stamens, etc. The characters such as presence or absence of corolla, position of ovary, union of stamens, union of stamens with corolla and gynoecium, etc. can be clearly said to be qualitative. Qualitative characters can not be measured.

Characters those can be measured or counted are termed as 'quantitative' characters, e.g. leaf size, number of calyx, corolla, stamens, etc.

Both these characters are used in taxonomic descriptions. In quantitative description, the number of parts and their range, size as expressed as length and other similar features are used. Qualitative characters are generally more useful to distinguish taxa of specific or higher ranks while quantitative ones are often useful to separate lower taxonomic categories at infra-specific levels/ranks. Although both the qualitative and quantitative characters are used in classical taxonomy but in taximetrics, quantitative characters are preferred.

Character weighting

The characters which are more constant have more weight to these. Characters are weighted by taxonomists for four main reasons – (i) because they are most easily observed, (ii) because they show highest correlation with others in a natural group, (iii) because they are believe to have some known or inferred phylogenetic importance or they are shown to be more conservative characters, and (iv) because they were left out characters while constructing a natural group.

In organisms, there are innumerable characters, and all of them can not obviously be considered for taxonomy. When taxonomists select characters, they are giving a greater importance (weight) to those selected. Heywood (1967) called it 'selection weighting'. In turn, many other characters are rejected by 'rejection weighing' or 'residual weighing'.

'Weighing of characters' is a technique in evolutionary taxonomy. Traditional taxonomists also weighted in an intuitive manner and without even being aware of it, in constructing classifications. Selection or rejection of characters is usually an intuitive process. Taxonomists select characters that are supposed to be greater taxonomic importance, those which are genetically more stable and conservative, those that are of greater diagnostic value and those thought to be evolutionary markers.

Character variation

It is accepted that individuals of a population, like the fingers on our palm, are not alike. They are variable. Consequently, the study of variation has acquired the greatest importance in systematic and evolutionary biology.

Variations are universal and form the basis of evolutionary and systematic studies. Without variation, there would not have array of organisms that now have. Detailed study of variation patterns helps one to speculate on the possible evolutionary processes and pathways along which a given group has evolved. Variations are of the utmost importance in taxonomy because they help one to discriminate between the different entities in nature. However, the taxonomic value of variations is not dependent on their evolutionary significance or adaptive value. Variations may be genetically fixed or environmentally controlled. While the former

collateral open vascular bundles arranged in the form of broken ring, four to five-merous flowers, whereas, in Monocotyledons the single cotyledon is associated with isobilateral, parallel veined leaves, closed and scattered vascular bundles and trimerous flower.

NUMERICAL TAXONOMY

Numerical taxonomy is a branch of modern approach to taxonomy or synthetic taxonomy. The analysis of various types of taxonomic data by mathematical or computerized methods is termed as Numerical taxonomy. It is also referred as **'Taximetrics'**. This approach of modern taxonomy involves the numerical evaluation of similarities or affinities between taxonomic units, and then arrangement of these units into taxa on the basis of their affinities.

Adenson (1763), a French botanist, was the first to put forward a plan for assigning numerical values to the similarity between organisms. He tried to use as many as characters as possible for classification. And such classification is known as **'Adensonian classification'**. Adensonian principles have developed several new methods in taxonomy during recent years, particularly after 1960, and all these methods are included in numerical taxonomy.

The use of modern electronic-data processing technique and several other methods of the numerical taxonomy have helped in the evolution of several new classifications of plants. Several objectives and qualitative methods are employed in their evolution.

PRINCIPLES OF NUMERICAL TAXONOMY

Numerical taxonomy is developed on the basis of certain principles which are also called **'Neo-Adensonian principles'**. Sneath and Sokal (1973) have enumerated following seven principles of Numerical taxonomy:

1. The greater the content of information in the taxa of a classification system and more the characters on which it is based, the better a given classification will be.
2. Every character has equal weightage in creating new taxa.
3. Between any two entities, the overall similarity 'is a function of their individual similarities in each of the many characters in which they are being compared'.
4. Distinct taxa can be recognized because correlations of characters differ in the groups of organisms under study.
5. Assuming some evolutionary mechanisms and pathways, the phylogenetic conclusions may be drawn from the taxonomic structure of a group and form character correlations.
6. The science of taxonomy is viewed and practiced as an empirical science.
7. Phenetic similarity is the base of classifications.

STEPS OF NUMERICAL TAXONOMY

The successive steps used in the construction of taxonomic groups are as follows:

Operational Taxonomic Units (OTUs)

The lowest rank of taxa or the basic units in any specific study are called Operational

Taxonomic Units or OTUs. An OTU may be an individual, a species, a genus, or a taxon of any higher rank. Therefore, the rank of OTUs may very from study to study. A species is generally treated as an OTU.

Unit Taxonomic Characters/Attribute

Sneath and Sokal (1973) defined unit characters as a 'taxonomic character of two or more states, which within the study in general, cannot be subdivided logically'. Usually the phenotypic characters are used as unit characters, e.g. the presence or absence of pappus in the seed of members of Asteraceae.

Maximum number of unit characters, and certainly not less than 50 in any case, should be used to obtain a fairly stable and reliable classification. Sneath and Sokal (1973) have given a list of disqualified characters which are not useful in numerical taxonomy. Such characters are called 'inadmissible characters'.

Regarding the proper selection of the unit characters, Sneath and Sokal (1973) suggested that (i) they should come from all the parts of the organisms, (ii) they should belong to all the stages of the life cycle of the organisms, (iii) all variable characters within the group should also be used and (iv) due attention should be given to characters related to morphology, physiology, ecology and distribution of the organisms.

TYPES OF BASIC CHARACTERS/ATTRIBUTES

Following are the major types of characters used in numerical taxonomy:

1. **Binary characters:** These characters contain two contrasting states, such as the presence or absence of some features, e.g. presence or absence of stipules/hairs, etc.

2. **Multi-state characters:** These are of two types –

   (a) **Qualitative Multi-state characters:** These characters contain three or more contrasting forms, e.g. corolla colour blue, orange or red.

   (b) **Quantitative Multi-state characters:** These characters represent measures of the size on a continuous scale such as length, breadth, height, etc. e.g. length of stigma, breadth of calyx, etc.

Coding of characters

The selected characters are assigned as symbol or mark, e.g. 1, 2, 3, ..., or + and -, etc. This is called 'coding of characters'. The symbol NC is used for 'no comparison'. Coding of characters may be following types –

1. **Two-state coding:** The characters in this type of coding are divided as 1 and 0, or as + and –. The characters recorded as 1 or + are the positive characters while those recorded as 0 or – are negative. The word NC is used if an organism not possessing a given character present in other organisms.

2. **Multi-state coding:** The characters in this type of coding are expressed by a single numerical value, such as 1, 2, 3, ... corresponding to the range of the variation of the characters. On the other hand these characters cannot be arranged in a reliable sequence or definite order.

Estimation of Resemblance

Between any two OTUs, the resemblance is estimated or measured in terms of similarity, *i.e.* percentage of characters in which they similar or percentage of characters in which they do not similar.

Phenetic resemblance between taxonomic groups is usually estimated by 3 methods – (*i*) Coefficients of association, (*ii*) Coefficients of correlation and (*iii*) Measurement of taxonomic distance between OTUs. These all 3 methods have specific formula to detect the phenetic resemblance between taxonomic groups.

Cluster Analysis

Clusters are groups of OTUs. With the help of estimations of resemblances, the affinities of different OTUs are determined. OTUs of similar affinities are grouped together in different taxa. A taxonomic system is constructed on the basis of resemblances and differences between groups of OTUs or clusters. The main features of the cluster analysis are arranged in the form of a 'dendrogram', *i.e.* cluster analyses are often displayed graphically using a tree-like diagram called '*dendrogram*'.

Fig. 6.1. Dendrogram
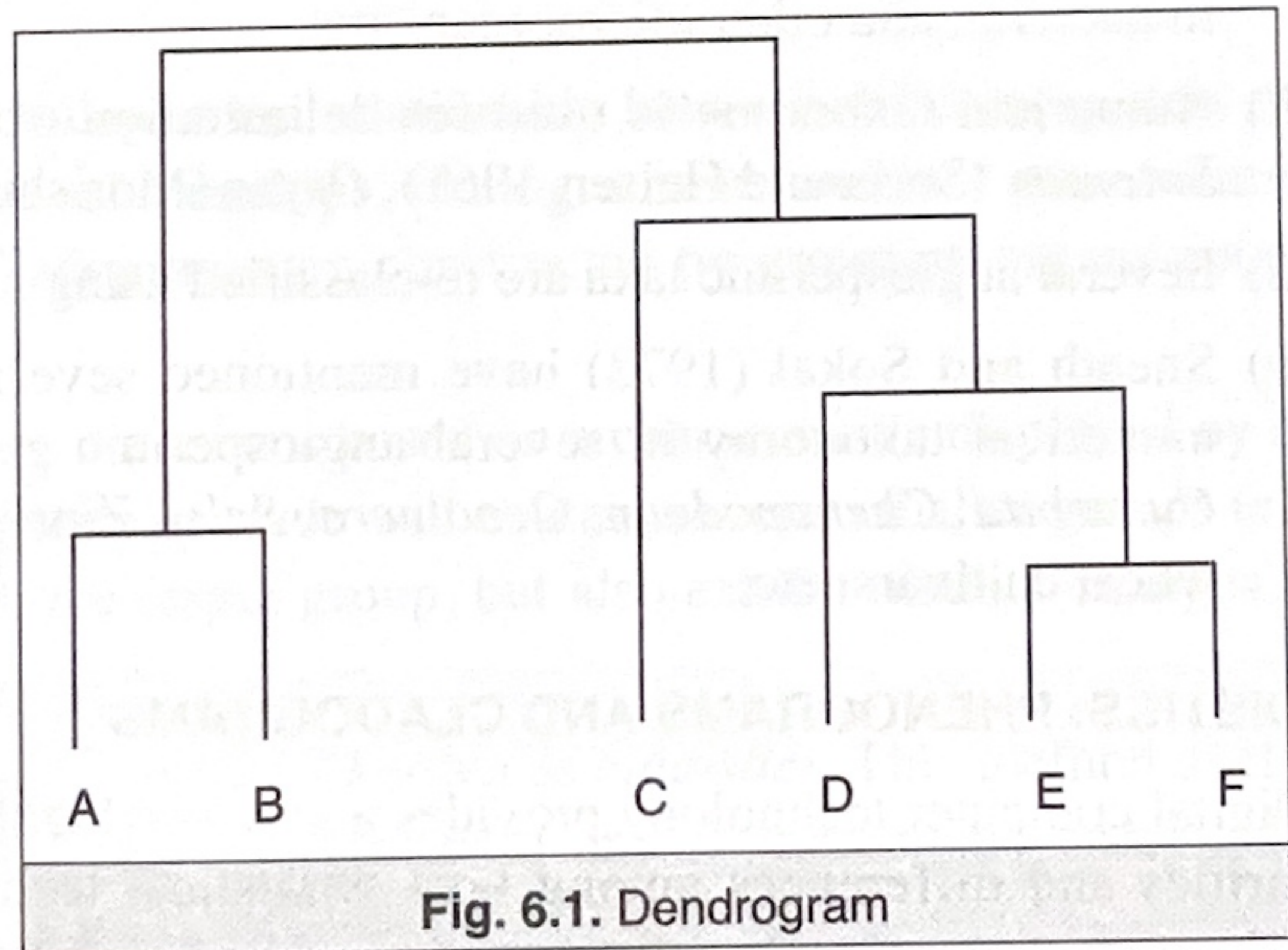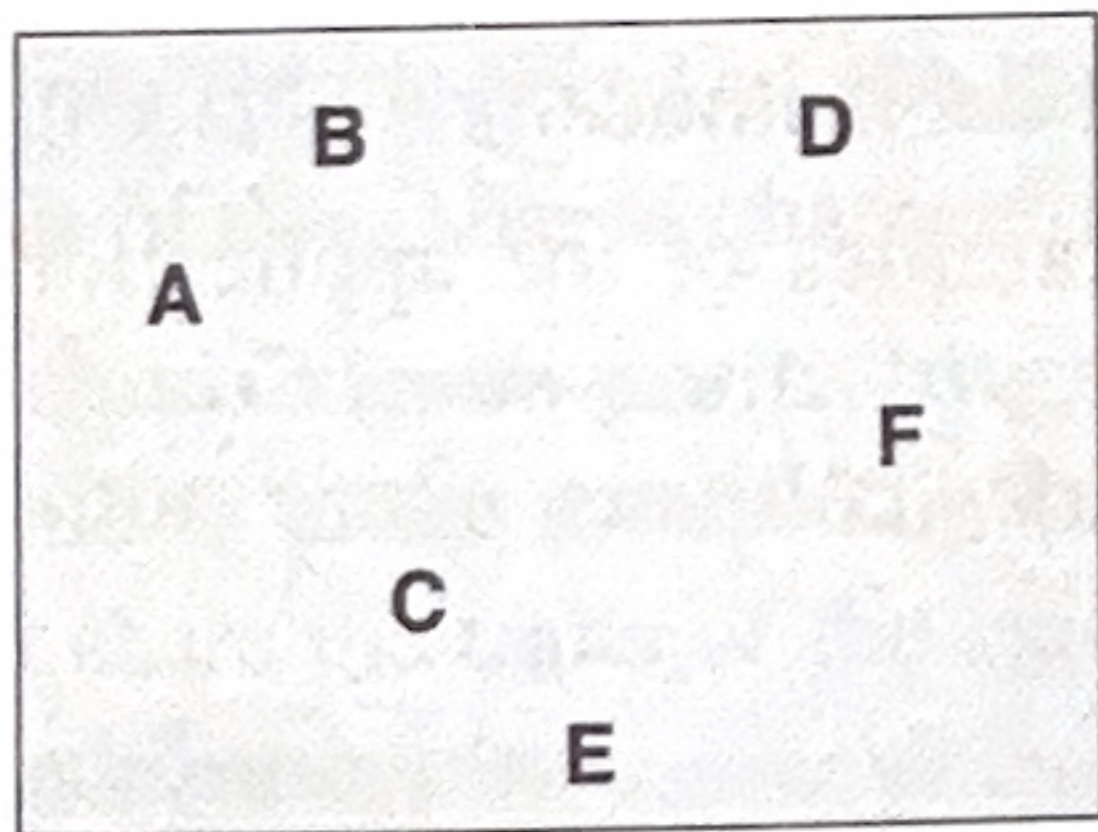
Phenons and Ranks

Groups of similar organisms recognized by numerical methods are called '*phenons*'. The phenon is however, not a synonym of the term 'taxon'. A phenon may or may not be equivalent to the ranks of the classical taxonomy, such as species, genus, etc.

Advantages of Numerical taxonomy

Sneath and Sokal (1973) have mentioned the following advantages of numerical taxonomy:

1. In comparison with the classical taxonomy, the data in numerical taxonomy are collected from more variety of sources, such as morphology, physiology, phytochemistry, cytology, proteins, etc.

2. Through numerical taxonomy, much taxonomic work can be done by less highly skilled workers.

3. Numerically coded data may be used easily for the creation of keys, maps, descriptions, catalogues, etc. with the help of existing electronic data processing system in herbaria and other taxonomic institutions.

4. Because the numerical methods are more sensitive in delimiting taxa, they provide better keys and classification systems in comparison to classical taxonomic methods.

5. The quality of classical taxonomy is improved by numerical taxonomy as more and better described characters are used in the latter.

6. Numerical taxonomy has suggested several fundamental changes in the classical principles of taxonomy in general and classification system in particular.

7. Numbers of the existing biological concepts have been reinterpreted in the light of numerical taxonomy.

ROLE/APPLICATION OF NUMERICAL TAXONOMY

Some of the examples of the application of numerical taxonomy are cited below –

(i) Numerical methods are used considerably to study the similarities and dissimilarities in bacteria, and other microorganisms.

(ii) Numerical taxonomy is used for delimitation of several angiospermic genera like *Solanum* (Soria and Heiser, 1961), *Oryza* (Morishima and Oka, 1960), etc.

(iii) Several angiospermic taxa are re-classified using numerical methods.

(iv) Sneath and Sokal (1973) have mentioned several examples of the application of numerical taxonomy in several angiospermic genera, *viz. Apocynum, Crotolaria, Cucurbita, Chenopodium, Oenothera, Salix, Zinnia,* barley cultivars, maize cultivars, wheat cultivars, etc.

CLADISTICS: PHENOGRAMS AND CLADOGRAMS

Digital computer technology provides a new and highly objective approach to evaluating similarities and differences among taxa. Statistical techniques like principal components analysis, multidimensional scaling, can be applied methodically to large data sets containing morphological, behavioural or biochemical information about taxonomic groups of interests.

These statistical methods allow taxonomists to quantify the degree of overall similarity among groups of organisms. The calculations may employ both 'primitive' and 'advanced' character states (the **phenetic approach**) or rely exclusively on 'advanced' characteristics (the **cladistic approach**). In both case, the resulting 'phylogenetic trees', variously known as **phenograms, cladograms** or **dendrograms**, are visual representations of perceived biological affinity (**Fig. 7.7A and 7.7B**).

Phenogram

Numerical phenetic methods organize taxa (OTUs) into groups based on overall similarity. OTUs are grouped in such a way that there is a higher degree of similarity within groups than between groups. It means two OTUs within a group share more character states than two OTUs in different groups. The least inclusive (smallest) groups represent the highest

level of similarity. More inclusive groups share successively fewer character states, and the most inclusive groups is the most heterogeneous, including all the OTUs in the study.

A clustering method is usually used to group OTUs that are most similar. The agglomerative clustering algorithms are most popular in this regard.

A measure of similarity or dissimilarity (or both) is computed for each pair of OTUs. This is usually done by **similarity matrix.** A similarity matrix consists of pair-wise comparisons between all the OTUs. One way to express the similarity of two OTUs is the **simple matching coefficient** (SMC). The value of SMC for two OUTs (A and B) is represented as follows:

$$\text{SMC } A, B = \frac{\text{Number of characters for which (A, B) share the same state}}{\text{Total Number of characters}}$$

This hierarchical pattern of grouping is expressed in a kind of branching diagram known as a *phenogram.* The relationship of OTUs and groups in a phenogram reflect only similarity. No evolutionary relationships are implied.

Cladogram

Cladograms are diagrams which depict the relationship between different groups of taxa called '*clades*'. By depicting these relationships, cladograms reconstruct the evolutionary history (phylogeny) of the taxa. Cladograms are constructed by grouping organisms on the basis of their shared characteristics.

A cladogram begins by grouping organisms based on a characteristic displayed by all the members of the group. The larger group or clade will contain several smaller groups (clades) that share the traits of the clades of the larger group, but also exhibit distinct changes as the organisms evolved.

Cladograms are constructed using a method known as *cladistics.* This method analyses a collections of heritable character data (morphological and/or molecular) compiled by a researcher.

The first step in basic cladistic analysis is to determine which character states are primitive and which are derived.

Though Hennig (1966) and others have proposed several methods for determining the primitive and derived character states, but the **out-group comparison method** is being widely used by the taxonomists today.

What is clade/cladistics/cladograms/phylogenetic tree?

A **clade** derived from the Greek word 'klados' means 'branch', also known as a **monophyletic group** or **natural group**, is a group of organisms that are monophyletic, *i.e.* composed of a common ancestor and all its lineal descendants. Rather than the English term, the equivalent Latin term *cladus* (plural *cladi*) is often used in taxonomical literature.

The science that tries to reconstruct phylogenetic trees and thus discover clades is called

phylogenetics or *cladistics*, the latter term coined by Ernst Mayr (1965), derived from "clade". The results of phylogenetic/cladistic analyses are tree-shaped diagrams called *cladograms* (Fig. 7.8) or *phylogenetic tree* (Fig. 7.7B).

Thus, a clade is by definition monophyletic, meaning that it contains one ancestor (which can be an organism, a population, or a species) and all its descendants. The ancestor can be known or unknown; any and all members of a clade can be extant or extinct.

The common ancestor may be an individual, a population, a species (extinct or extant), and so on right up to a kingdom and further. Clades are nested, one in another, as each branch in turn splits into smaller branches. These splits reflect evolutionary history as populations diverged and evolved independently. Clades are termed monophyletic groups.

Over the last few decades, the cladistic approach has revolutionized biological classification and revealed surprising evolutionary relationships among organisms. Increasingly, taxonomists try to avoid naming taxa that are not clades; that is, taxa that are not monophyletic. Some of the relationships between organisms that the molecular biology arm of cladistics has revealed are that fungi are closer relatives to animals than they are to plants, archaea are now considered different from bacteria, and multicellular organisms may have evolved from archaea.