Classification Systems

- **Classification**: is a very broad term with applications far beyond the biological sciences that simply refers to the act of systematically arranging objects (ideas) into categories according to specific criteria.

¬ The differences in approach to comparison of **character states** have resulted in the recognition of **five** **systems of classification.** They are:

1. Artificial Classification system

- With this system only one or at most a few characters are selected for use in making comparison among objects and organisms.
- The selection of characters is usually made

  - Because a few characters are involved, the difficulties encountered in describing, measuring and comparing the character states usually are minimal.
  - Ranking is done subjectively with certain character states being regarded as subordinate to others.
  - Artificial system is **monothetic** i.e. possession of unique set of characters/features is both sufficient and necessary for membership in the group thus defined.
- The artificial system was the first to be used in the classification of organisms and these origins begin with the **ancient Greeks**:

2. Natural (Adansonian) System/Approach

- It is one based upon states of several to many characters selected a **posteriori** for their value in positively correlating with states if other characters to form a hierarchical structure of groups in ranks containing high information content and predictive value.
- The characters selected are in effect weighted by their selection and employment over those features not selected
- The taxonomist does the selection and comparison, and eventual evaluation

- The process occurs rapidly in the mind of the maker (taxonomist).
- **Natural classification system** is therefore **polythetic** i.e. it places together organisms that have the greatest number of shared features, and no single feature is either essential to group membership or is sufficient to make an organism a member of the group.
- Examples of Natural system are those of Adanson, John Ray and de Jussieu
- After **de Jussieu,** nearly all-subsequent systems of classification were natural until the development of evolutionary perspective brought about by Charles Darwin's the Origin of Species in 1859

3. Phyletic (Evolutionary) System/Approach

- It is a classification that aims at:
- Reconstructing a sequence starting with the most primitive or least advanced or least specialised and ending with the most advanced or derived.
- Ensuring that each taxon recognised is monophyletic i.e. has arisen by the diversification of a single ancestor – as opposed to polyphyletic – arising from more than one ancestral group.

Phyletic Principle

- One method of inferring phylogeny is to pin point:
- Primitive (ancient) as opposed to advanced (recent) characters and to assign primitiveness to taxa which possess high proportions of the primitive characters.
- Examples of phyletic systems of classification are those of Bessey, Hutchinson, Sporne, etc.

a. Besseys (1911)

Table.1 Comparison of characters alleged to be primitive/ advanced by three plant taxonomists/phyleticists

Bessey	Hutchinson	Sporne
Chlorophyll/no chlorophyll	Chlorophyll/no chlorophyll	–
Woody/herbaceous	Woody/herbaceous	Woody/herbaceous
Bundles collateral/scattered	Bundles collateral/scattered	–
Leaves simple/compound	Leaves simple/compound	–
Leaves opposite/spiral	Leaves spiral/ opposite	Leaves spiral/ opposite
Flowers polymerous/oligomerous	Flowers polymerous/oligomerous	–
Flowers bisexual/ unisexual	Flowers bisexual/ unisexual	Flowers unisexual/ bisexual
Flowers monoecious/dioecious	Flowers monoecious/dioecious	–
Flowers petalous/apetalous	Flowers petalous/apetalous	–
Flowers actinomorphic/zygomorphic	Flowers actinomorphic/zygomorphic	Flowers actinomorphic/zygomorphic
Flowers polypetalous/gamopetalous	Flowers polypetalous/gamopetalous	–
Stamens many/few	Stamens many/few	Stamens many/few
Stamens separate/fused	Stamens separate/fused	Stamens separate/fused
Flowers hypogynous/epigynous	Flowers hypogynous/epigynous	–
Gynoecium polycarpous/oligocarpous	Gynoecium polycarpous/oligocarpous	Gynoecium polycarpous/oligocarpous
Gynoecium apocarpous/syncarpous	Gynoecium apocarpous/syncarpous	Gynoecium apocarpous/syncarpous
Seeds with/without endosperm	Seeds with/without endosperm	–
Stems simple/branched	–	–
Leaves evergreen/deciduous	–	–
–	–	Stipules present/absent
–	–	Secretary cells present/absent
Venation reticulate/parallel	–	–
–	Flowers solitary/clustered	–
Floral parts spiral/whorled	–	–
–	–	Pollen 1- or 2-/3-nucleate
–	–	Pollen few/many pored
–	–	Placentation axile/parietal
–	–	Ovules crassinucellate/tenuinucellate
–	Fruits single/aggregate	–
–	Fruits capsule/drupe or berry	–
–	–	Seeds arillate/not
–	–	Endosperm nuclear/cellular

Note that there are agreements and some disagreements among the three taxonomists.

4. Phenetics (Numerical/Mathematical/Taxonometrics/taxometrics/taximetrics, Statistical/ Modern Phenetics)

- A classification based on numerous characters of equal weight and their comparison using computer programmes.
- The word "**phenetic**" was introduced by **Cain** and Harrison in 1960 to mean a relationship by overall similarity, based on all available characters without any weighting.

Sneath & Sokal (1963)

- Used phenetic to refer to a relationship between taxa "evaluated purely on the basis of the resemblance arising now in the material at hand and the overall similarity as judged by the characters of

Sokal & Sneath 1973 defined Phenetic as

- Similarity/resemblance based on a set of phenotypic characteristics of the objects or organisms under study.
- Equal weight is explicit in the definition of phenetics.
- Phenetic as numerical Taxonomy was defined by Sokal & Sneath (1973) as
- The grouping by numerical methods of taxonomic units into taxa on the basis of their character states.

Methodology of Phenetics

1. Selection of taxa (individuals) for study. Those are usually called Operational Taxonomic Units (OTUs).
    - Which is the basic unit of numerical taxonomy or the term given to the lowest taxon being studied in a particular investigation.

- They are simply the starting point units of phenetic classification.
- They might be:
    - Individual organisms
    - Populations
    - Species
    - Genera, etc.

2. Selection of character states. Ideally more than 100 unit characters – usually a mixture of qualitative and quantitative characters.
3. Description and/or measurement of character states
4. Comparison of states to:
    a) determine a measure of overall similarity (phenetic relationship) between each pair of OTUs, and
    b) determine the taxonomic structure, i.e. the detection of possible groups and subgroups among all OTUs

5. Construct a phenogram that expresses the relationships among all the OTUs in a conventional form.

-The phenogram reveals the taxonomic structure of groups and subgroups among the OTUs.

6. Ranking of all OTUs into the categories of the taxonomic hierarchy.

-But critics argue that formal taxonomic ranking from the phenogram is meaningless because the initial level of the OTUs is instated (they could be individuals, populations, species, genera

An example of phenetic classification

1. First select as many characters as possible

- These initially will probably be morphological features but any other data can be used, such as:
  - Anatomical characters
  - Palynological characters
  - Chemical aspects

Here ten characters are selected for simplicity.

- The characters are usually a mixture of qualitative and quantitative characters.
- In this example only characters with two states are used and are given arbitrary numerical values of 0 for one state and 1 for the other.
- The numerical values (0 and 1) have no reference to primitive or derived conditions – they are simply arbitrary numerical descriptions.
- Multiple states for characters also can be used and compared.
- The character states then are compared from one OTU to the next.
- A number of statistics (or coefficients) exist for making such comparisons, and one of the simplest is called the Simple Matching

- **Simple Matching Coefficient** is calculated by counting the number of character states for each character in common between two OTUs (i.e., the positive matches) and then dividing this number by the total number of characters used positive plus negative matches).

Table 2. Basic data Matrix of Characters Selected and Numerical Values Accorded to the two Character States (either 1 or 0) in OTUs S–Z for Phenetic Analysis. For Simplicity, the States of Only Ten Characters are Shown in the Table, but the Actual Number Might be Well Over 100.

	Characters									
OTU	1	2	3	4	5	6	7	8	9	10
S	0	0	0	0	0	0	0	0	1	1
T	0	0	0	1	0	0	1	0	1	1
W	0	1	0	1	0	1	1	0	1	0
X	0	1	0	1	0	1	0	0	1	0
Y	1	1	1	0	1	0	0	1	0	0
Z	1	1	1	0	1	1	0	1	0	0

Table 3. Data Matrix of Simple Matching Coefficients Among OTUs S-Z.

	S	T	W	X	Y	Z
S	1.0					
T	0.8	1.0				
W	0.5	0.7	1.0			
X	0.6	0.6	0.9	1.0		
Y	0.3	0.1	0.2	0.2	1.0	
Z	0.2	0.0	0.3	0.4	0.9	1.

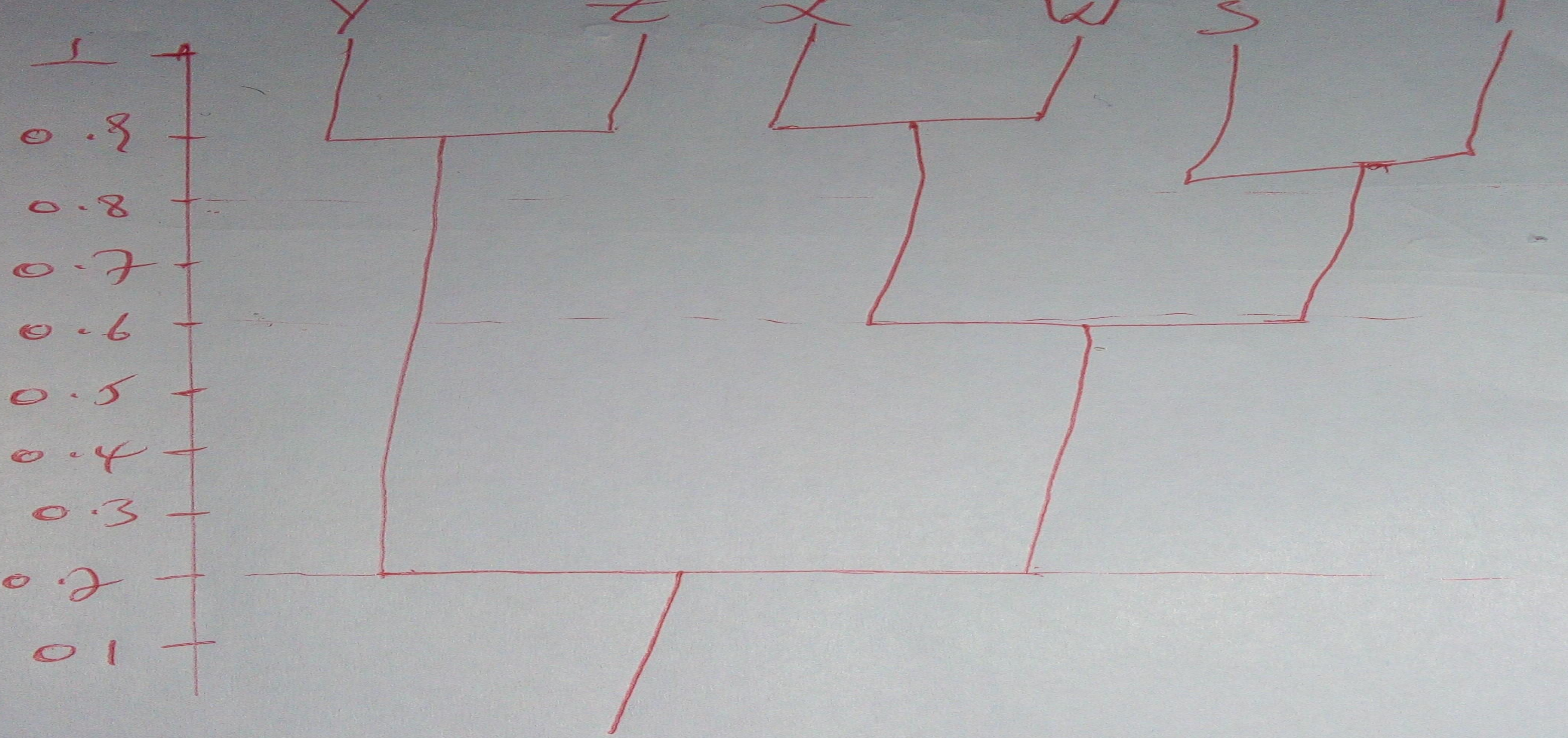
- The higher the coefficient of association (i.e. as the value approaches 1.0), the more closely related are the OTUs. For example, the coefficient between **S and T is 8 character states** in common divided **by 10 characters equals 0.8.**

Simple Matching Coefficient between

$$\textbf{S \& T} = \frac{\textbf{character states in common}}{\textbf{Total no. of characters}}$$

$$\frac{8}{10} = \mathbf{0.8}$$

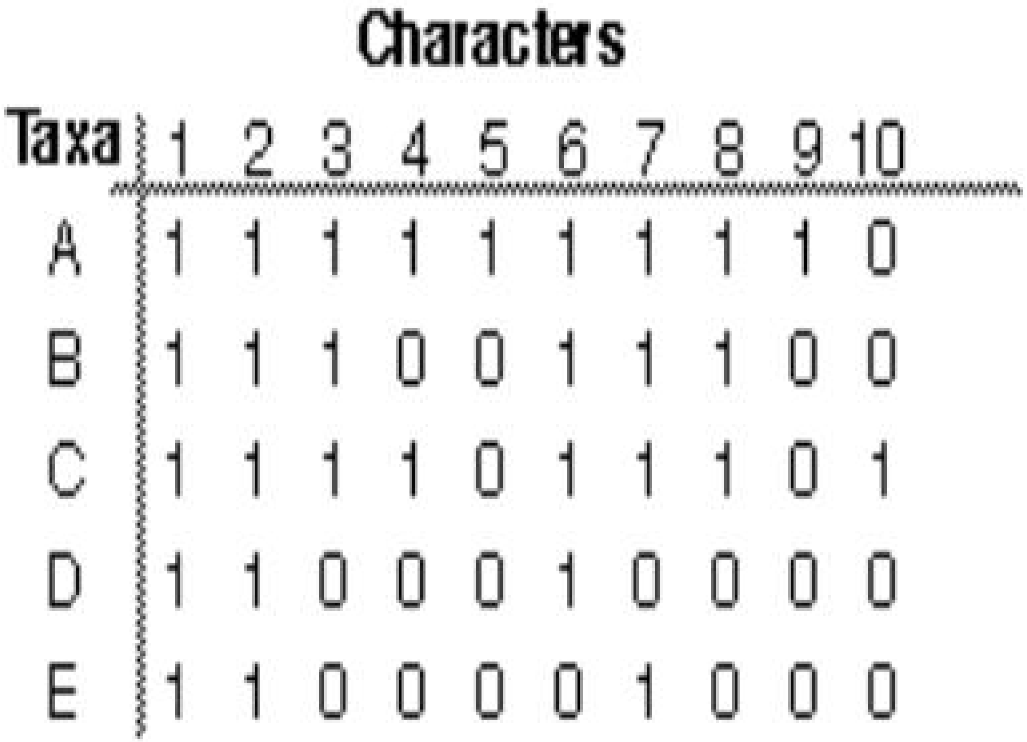
- Such calculations are done between all 15 pairs of OTUs and the resultant figures are displayed in a data matrix of Coefficients of Association (Table 4.2 above)
- If many characters and groups are involved the calculations are usually completed with the aid


Y
Z
X
W
S
T
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0 1

- Fig. 2. Phenograms of relationships among OTUs S_Z based Simple Matching Coefficients and unweighted arithmetic averages
  - The vertical lines delineating groups or clusters of OTUs are based on the Coefficient of Association.
  - S & T are similar at the 0.8 level.
  - W & X, and Y & Z are similar at the 0.9 level.
  - The former OTUs (S & T) are more similar to each other than either pair is to the latter (W & X and Y & Z) closely related.
  - The level at which they are connected is based on an average (unweighted arithmetic) of values of the pairs S – W, S – X & T-X,T-W (= 0.6)
  - The relationship of Y & Z to the other four OTUs

EXCERSISE

- consider five species of plants (A-E) and construct a phenogram



	Characters									
Taxa	1	2	3	4	5	6	7	8	9	10
A	1	1	1	1	1	1	1	1	1	0
B	1	1	1	0	0	1	1	1	0	0
C	1	1	1	1	0	1	1	1	0	1
D	1	1	0	0	0	1	0	0	0	0
E	1	1	0	0	0	0	1	0	0	0

Important Points About the Phenogram

- The diagram attempts to show only phenetic similarity, which is based on the comparison of character states.
- Evolutionary pathways are not implied here.
- The phenogram is a slight distortion of the relationships seen in the matrix of Coefficients of Association.
- Formal taxonomic ranking from the phenogram is meaningless because the initial level of the OTUs is un-stated (they could be individuals, populations, species, genera, families, etc.)

5. Cladistics (Modern Phylogenetic Systematics)

- **Cladistics** is a methodology that attempts to analyze phylogenetic data objectively.
- is the concepts and methods for the determination of branching patterns of evolution.
- In cladistics much emphasis is placed on distinguishing primitive versus derived character states,
- Willi Henning (1950) was the pioneer worker in this field.
- Several terms are unique to cladistics

Definition of terms in cladistics

1. **Plesiomorphy (plesiomorphous) –** means primitive character state.
2. **Apomorphy** (apomorphous) – refers to derived or advanced character state.
3. **Symplesiomorphies** (symplesiomorphous) means plesiomorphies that are shared by two or more taxa or lineages.
4. **Synapomorphies** (synapomorphic) – refers to shared apomorphies.
5. **Autapomorphies** - derived character states (apomorphies) that are found in only one evolutionary line or one taxon.
6. **Cladogram** – a branching diagram/dendrogram constructed by cladistics principles and methods.
7. **EU's (or OTU's)** – are taxa compared and evaluated in a cladistic study. They are also sometimes called OUT's following phenetic terminology.
8. **Sister group** – the most closely related group

9. **Semaphoront** – the organism or the individual at a particular point of time or during a certain theoretically infinitely small, period of its life.

-It is also called the character-bearing semaphoront.

10. **Holomorphy** – the totality of all characters of the semaphoront, including morphology, physiology, chemistry, etc.

11. **Homology** – resemblance due to inheritance from a common ancestry.

12. **Homoplasy** – resemblance not due to inheritance from a common ancestry,

- *Parallelism* – the development of similar characters (or character states) separately in two or more lineages of common ancestry, and on the basis or channelled by characteristics of that ancestry.
- *Convergence* – the development of similar characters or character states in different lineages but without a direct common ancestry.

13. **Dichotomy** – two branches diverging on dendrograms of any type including cladograms.

    -*Trichotomy* - three branches arising from the same point

    -*Tetrachotomy* – four branches arising from the same point.

14. ***Monophyletic*** – a group that includes all the descendants of a common ancestor.

15. Paraphyletic a group that has a common ancestor but which does not include all the descendants of the ancestor.

- Paraphyletic groups to some cladists are not useful for classification.

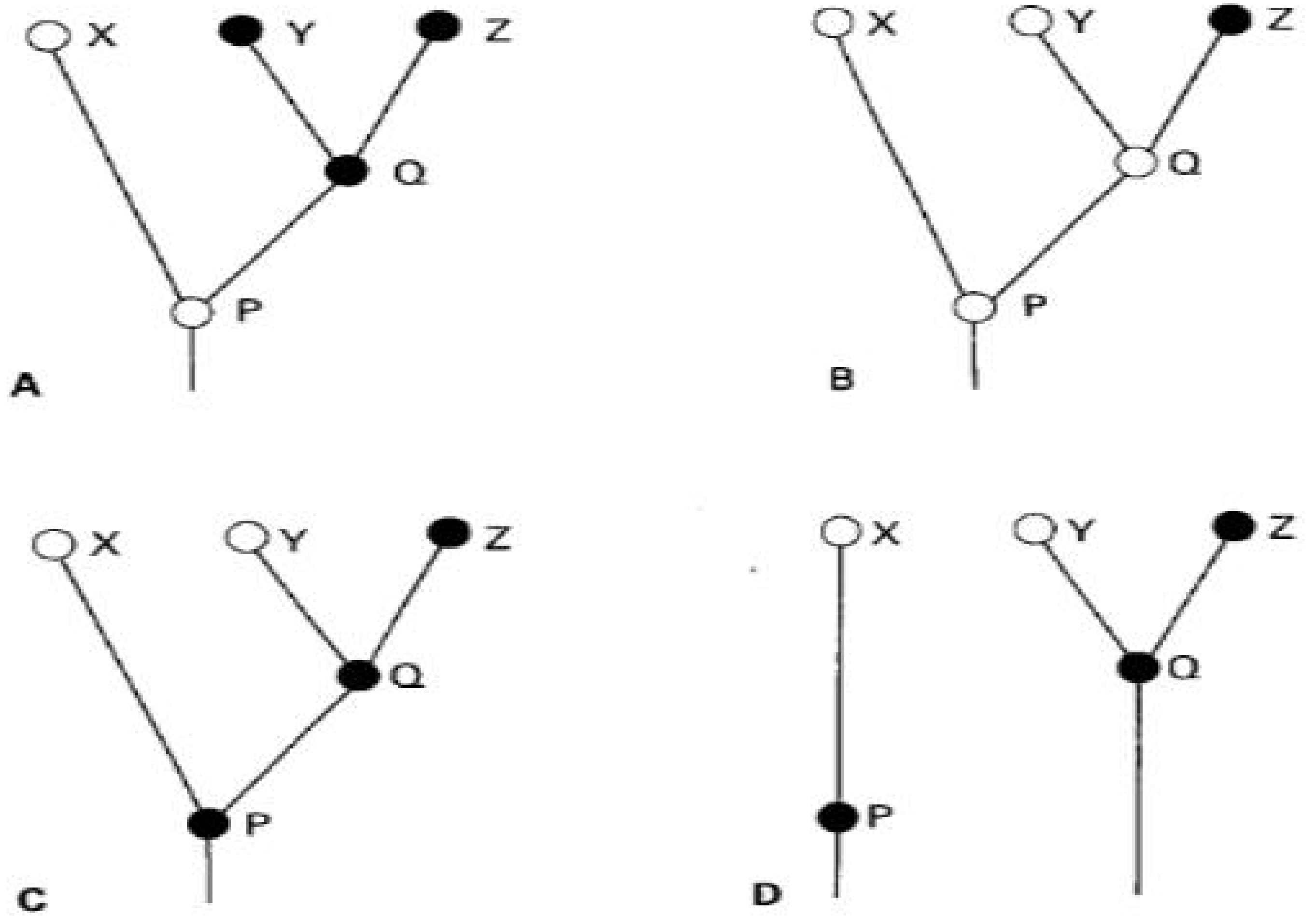


Fig. 2.6 Four diagrams showing different origins of three species (X, Y, Z) from the ancestral taxa P and Q in order to illustrate the concepts of monophyly, paraphyly, polyphyly, parallelism and convergence. The possession of one or other of two contrasting character-states by each of the five taxa is indicated by an open or closed circle respectively. **A.** Groups YZ and XYZ are both monophyletic; the similarity between Y and Z is a synapomorphy; the difference between X and YZ is due to divergence. **B.** Group XY is paraphyletic; group XYZ is monophyletic; the similarity between X and Y is a symplesiomorphy; the difference between Y and Z is due to divergence. **C.** Group XY is polyphyletic; group XYZ is monophyletic; the similarity between X and Y is a false synapomorphy caused by parallelism. **D.** Groups XY and XYZ are both polyphyletic; group YZ is monophyletic; the similarity between X and Y is a false synapomorphy caused by convergence.

Methodology of Cladistics

- Although the procedures for cladistic analysis and classification vary, there is a general set of procedures that are used by most cladists with conventional types of data, especially morphology.

1. Make evolutionary assumptions (select EU's; determine monophyly i.e. do all the taxa involved in the study form a monophyletic unit and determine initial rank of each of the EUs i.e. whether they are populations, species, genera, etc.)
2. Select characters of evolutionary interest.
    - The characters that are selected should be of evolutionary interest and minimally the characters must be one character less than the number of EUs for all branches of the tree to be

3. Describe and/or measure character states.
   - The characters must be divided into states, and usually these are qualitative, although this does not have to be the case.

4. Ascertain homologies of characters and character states.
   - A check on the homologies of characters and states at this point is essential to be certain that the same kinds of features are being compared.

5. Construct character state network.
   - The next step is to place the character states in a logical sequence based primarily

- This is obvious with only twos states, but with three or more states, it becomes more complicated and more discerning judgement is needed.
- For example, if three states of corolla colour (e.g. white, red and pink) occur in the study group, the simplest way to connect the states logically is red – pink – red rather than white – red - pink.
- The degree of colour change between each state is the smallest in the first approach and more drastic in the second.

6. Determine polarity of character state networks (primitive vs. derived conditions), i.e. "root" the character state networks to form character state trees.

- The character state network must be "rooted"

7. Construct basic data matrix.
   - There are 9 criteria to determine primitive conditions of character states.
   - Once the polarities of the character state network have been determined, an appropriate coding (letters or numbers) must be assigned to the states so that they can be placed conveniently in the basic data matrix.
   - Usually 0 is used for the primitive state and 1 for the derived condition.
8. Select algorithm and generate trees (cladograms). Once the above (7) has been done, the character states can be compared and evolutionary trees constructed, using one of the algorithms, e.g. Parsimony algorithm - Farris Parsimony

9. Construct classification based upon the cladograms.

- Formal classification can be done once trees have been generated.
- There are two viewpoints:
- The cladogram should be used directly in constructing the classification and this should be done explicitly so that the cladogram can be derived from the classification and vice versa
- The cladogram should be used as a guide to construct the classification but other aspects of phylogeny (such as autapomorphies or phenetic divergence)

Table 4. Basic data matrix of binary (two state) characters 1-10 in EUs A-F for cladistic analysis. 1 = derived; 0 = primitive state.

	Character									
EU	1	2	3	4	5	6	7	8	9	10
A	0	1	0	1	0	1	1	0	0	0
B	0	1	1	1	0	0	0	0	0	0
C	1	0	0	1	0	0	0	0	0	1
D	1	0	0	1	0	0	0	1	0	0
E	0	0	0	0	1	0	0	0	0	0
F	0	0	0	0	1	0	0	0	1	0

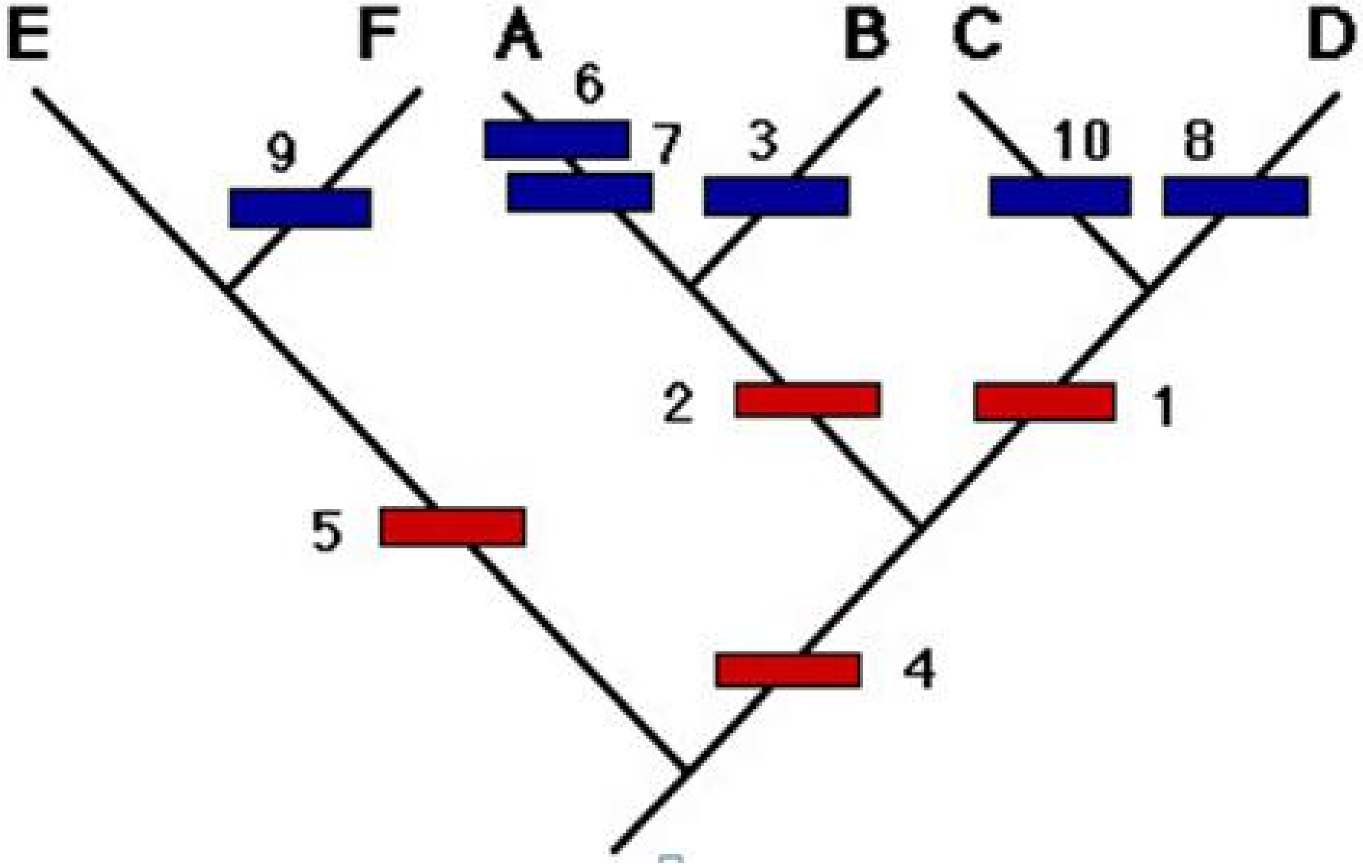


Fig. Cladogram illustrating relationship among EUs A-F based on binary characters 1-10. Bars indicate apomorohies. (See attached figure). Note.

Autapomorphies: 3, 6, 7, 8, 9 & 10.

Synapomorphies: 1, 2, 4 & 5.

The CRITERIA

1. **Fossil evidence.** In a series of fossils the stratigraphically oldest fossil will be regarded as the most primitive of the lineages and all of its states will be treated as primitive ones.

2. **Common is primitive:**
   - *In-group analysis* – involves looking at the distribution of character states among EUs of the study group, and those features which are most prevalent are judged to be most primitive.
   - *Out-group analysis* – involves looking at the distribution of character states in the most closely related group (or groups but not necessarily restricted to the sister group) and the common (ideally the uniform) conditions there are regarded as the most primitive for the study group.

EXERCISE

Feature	**A**	**B**	**C**	**D**	**E**
1	0	+	+	+	0
2	0	+	+	+	+
3	0	0	+	0	0
4	0	0	+	+	0
5	0	+	0	0	0