**Course: Advanced Bioinformatics**

**Module title: Phylogeny Tree Basic**

**Module no. : 45**

This module provides the information about “Phylogentics – Basic Concepts”

Classical phylogenetic analysis was based on morphological

Features which are visible to human eye, i.e. they have some explicit presence; can be seen and touched such as number of legs, lengths of legs, etc.

Modern biological methods allow to use molecular features which are not obvious but can be gathered by biological methods. Examples of these features include

**–**  Gene sequences

**–**  Protein sequences

Now we give overview of some of the approaches for phylogeny tree construction. Details of these approaches are provided in coming modules.

Distance: find tree that accounts for estimated evolutionary distances

Parsimony: find the tree that requires minimum number of changes to explain the data

Maximum likelihood: find the tree that maximizes the likelihood of the data

Leaves: things (genes, individuals, strains, species) being compared.

Taxon (taxa plural) is used to refer to this.

Internal nodes: hypothetical ancestral units

Rooted tree: path from root represents an evolutionary path (root

represents the common ancestor)

Unrooted tree: specifies relationships among things, but no evolutionary path.

Example

Seq A: A A C C G G T T

Seq B: A A C C G G T G

Seq C: A C C C G G T C

Seq D: A C C C G G T A

Which Sequences? DNA: very sensitive, non uniform mutation rates

CDNA/RNA: useful for remote homologies

Protein Sequences: useful for most remote homologies, deep phylogenies, more uniform mutation rates more character states

Ribosomal RNA 16S Sequences:

* Exist in all organisms
* Highly conserved
* Suitable for broad, very deep phylogeny studies
* Complied for 10s-1000s of organisms, mostly microbial
* Unsuited to fine grained phylogeny

Computational Process of Phylogenic analysis:

For generation of phylogeny tree following steps should be followed.(Note: details of these steps are provided in subsequent lectures)

1. Get DNA/RNA/PS
2. Construct multiple sequence alignment
3. Compute pair wise distanced
4. Build tree: topology + branch lengths
5. Estimate reliability
6. Visualize