

Phylogenetic Tree Construction- Character Based Methods- Maximum Parsimony

Manu Madhavan

Lecture 10

- Phylogenetic Trees - importance
- Methods- Distance based and Character based
- UPGMA, Neighbor joining

- Character Based methods - Maximum Parsimony

Character Based Methods

- Character-based methods (also called discrete methods) are **based directly on the sequence characters** rather than on pairwise distances
- They count mutational events accumulated on the sequences
- Avoid the loss of information when characters are converted to distances
- This preservation of character information means that evolutionary dynamics of each character can be studied
- Popular algorithms- Maximum Parsimony and Maximum Likelihood approach

Maximum Parsimony method

- *Parsimony* n. extreme unwillingness to spend money or use resources
- Biological parsimony: 'Parsimonious in allowing mutational changes between sequences
- The **parsimony** method chooses a tree that has the fewest evolutionary changes or shortest overall branch lengths
- The two premises that underlie biological parsimony are quite simple:
 - Mutations are exceedingly rare events and
 - The more unlikely events a model invokes, the less likely the model is to be correct.
- The relationship that requires the fewest number of mutations to explain the current state of the sequences being considered is the relationship that is most likely to be correct

Maximum Parsimony method

- *Parsimony* n. extreme unwillingness to spend money or use resources
- Biological parsimony: 'Parsimonious in allowing mutational changes between sequences
- The **parsimony** method chooses a tree that has the fewest evolutionary changes or shortest overall branch lengths
- The two premises that underlie biological parsimony are quite simple:
 - Mutations are exceedingly rare events and
 - The more unlikely events a model invokes, the less likely the model is to be correct.
- The relationship that requires the fewest number of mutations to explain the current state of the sequences being considered is the relationship that is most likely to be correct

Maximum Parsimony method

- Occams razor: the simplest explanation (with fewest number of assumptions) is probably the correct one.
- A tree with the least number of substitutions is probably the best to explain the differences among the taxa under study
- Parsimony tree building works by searching for all possible tree topologies and reconstructing ancestral sequences that require the minimum number of changes to evolve to the current sequences

Maximum Parsimony method

Informative and uninformative sites

- Only a small number of sites that have the richest phylogenetic information are used
- **Informative sites**, which are defined as sites that have at least two different kinds of characters
- **uninformative site**, constant sites or sites that have changes occurring only once
- The uninformative sites are thus discarded in parsimony tree construction.

Sequence	Site								
	1	2	3	4	5	6	7	8	9
1	A	A	G	A	G	T	T	C	A
2	A	G	C	C	G	T	T	C	T
3	A	G	A	T	A	T	C	C	A
4	A	G	A	G	A	T	C	C	T
					*		*		*

Two types of variable sites:

Informative: favors a subset of trees over other possible trees.

Uninformative: a character that contains no grouping

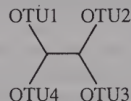
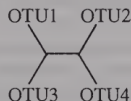
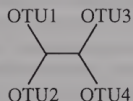
Computing Parsimony at each informative site

- The minimum number of substitutions at each informative site is computed for a given tree topology
- The total number of changes at all informative sites are summed up for each possible tree topology
- The tree that has the smallest number of changes is chosen as the best tree

Maximum Parsimony method-Example

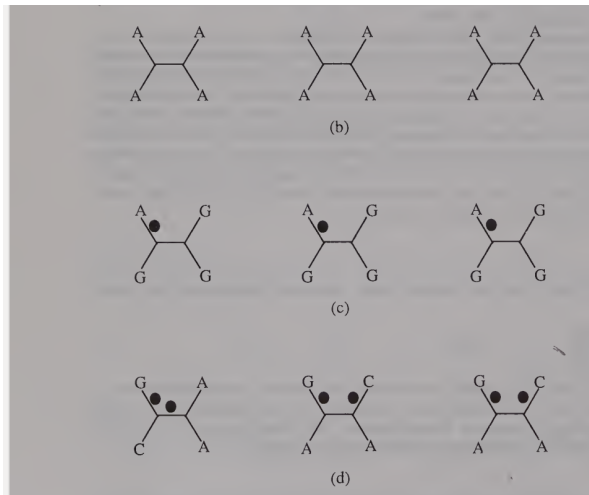
```
OTU1:  A A G A G T G C A
OTU2:  A G C C G T G C G
OTU3:  A G A T A T C C A
OTU4:  A G A G A T C C G
```

This multiple sequence alignment of the four DNA sequences is the sta

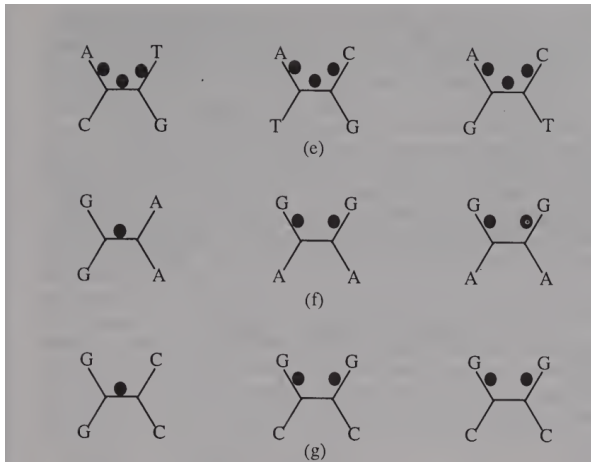


- Identify all possible un-rooted trees
- Identify informative/uninformative sites
- Compute total parsimony from informative sites of each tree
- Select the one tree with minimum value (maximum parsimony) as optimal tree

Maximum Parsimony method-Example



Maximum Parsimony method-Example



Maximum Parsimony method-Example

	Tree I	Tree II	Tree III
Site 5	1	2	2
Site 7	1	2	2
Site 9	2	1	2
Total (of all informative sites)	4	5	6

- Character based methods- Weighted Parsimony, Maximum Likelihood method