

Assignments:

1. Motif finding:
- (a) n -upstream of starting codon of a gene
 - (b) n -downstream of stop codon of a gene
 - (c) $n+m$ -midstream of donor sites
 - (d) $n+m$ - " of acceptor sites
 - (e) n -upstream of SNP sites
 - (f) n -downstream of SNP sites
2. Analysis of restriction fragments: length distribution
- (g) n -upstream of LTR sites.
 - (h) n -downstream of LTR sites
 - (i) n -upstream of DNA transposon sites.
 - (j) n -downstream of DNA transposon sites.

3. Local seq. alignment \rightarrow given a gene of one org. & genome of another find similar gene (use cDNA database)
4. global seq. alignment \rightarrow tandem
5. Annotation and description of repeat sequences.
6. Webcrawler for downloading genome files & species and database.
7. generation of theoretical tandem mass spectrum given a polypeptide and reconstructing it by forming the spectrum graph.
8. SNP based STS mapping.
9. generation of alternatively spliced protein given a gene
10. given ORF's (defined by start & stop codon intervals) of bacterial genome \rightarrow predict genes by searching a database.
11. Restriction length polymorphism: \rightarrow
12. Clustering of n feature vectors using Cluster Affinity Search Technique (CAST).
13. Splicing based sequence alignment.
I/p: ORF, a Target Gene, set of candidate exon blocks.
14. Clustering of 1-D feature vectors using K-Means clustering Alg.