|  |
| --- |
| Course Outcomes:  1. Learning fundamental concepts in computational biology and its applications.  2. To learn application of networks, algorithms, and models in biology.  3. To learn manipulation of DNA and proteins to generate information from sequences and whole genomes  4. To learn modeling of biological processes by application of computer programming  5. To learn techniques and concepts for investigating interactions among biomolecules  6. To develop skills in applications of computational biology. |

|  |  |  |  |
| --- | --- | --- | --- |
| **S.No.** | **Topic** | **No. of lectures** | **Teacher (SB/SS)** |
|  | **Introduction and overview** - Mathematics, statistics and computer science. | Lecture 1 | SS |
|  | Introduction to NCBI: sequence databases, sequence retrieval, sequence file formats | **Practical-1** | SS |
|  | Cloning and clone libraries | Lecture 2 | SS |
|  | libraries by complete and partial digestion | Lecture 3 | SS |
|  | Writing a sequence assembly program | **Practical-2** | SS |
|  | Physical genome maps – mapping by fingerprinting and mapping by anchoring | Lecture 4 &  Lecture 5 | SS |
|  | clone overlap and sequence assembly | Lecture 6 & Lecture 7 | SS |
|  | Introduction to NCBI: Literature and other resources  Introduction to NCBI: Structure and other resources | **Practical-3** | SS |
|  | Shotgun sequencing | Lecture 8 | SS |
|  | Sequencing by hybridization. | Lecture 9 | SS |
|  | **Simple problems in Computational Biology:** Double digest problem, | Lecture 10 | SS |
|  | Algorithms for DDP, | **Lecture 11** | **SB** |
|  | Exact string matching – classical comparison-based methods, semi numerical string matching | **Lecture 1** | **SB** |
|  | suffix trees – construction and application | **Lecture 2** | **SB** |
|  | **Pairwise sequence comparisons:** | **Lecture 3** | **SB** |
|  | Databases and rapid sequence analysis – | **Lecture 4** | **SB** |
|  | Pairwise comparisons: - effect of different substitution matrices, change in gap penalties | **Practical-4** | **SB** |
|  | Tree representation of a sequence- hashing a sequence, repeats in a sequence | **Lecture 5** | **SB** |
|  | sequence comparison by hashing | **Lecture 6** | **SB** |
|  | sequence comparison by at most *l* mismatches, sequence comparison by statistical content | **Lecture 7 & 8** | **SB** |
|  | Dynamic programming alignment | **Lecture 9** | **SB** |
|  | The number of alignments- shortest and longest paths in a network | **Lecture 10** | **SB** |
|  | global distance and similarity alignments, scoring matrices and gap penalties, Filtering position specific scoring matrices | **Lecture 12** | **SB** |
|  | Database search: Evaluate the statistical significance of the match with a web program. Effect of presence of low complexity regions in the sequence and filtering | **Practical-5** | **SB** |
|  | **Database search techniques-** parametric sequence comparison, statistical significance of alignments | **Lecture 13** | **SB** |
|  | Probability and statistics for sequence alignment, extreme value distributions | **Lecture 14** | **SB** |
|  | database searching algorithms and artifacts  sequence alignment with scores. | **Lecture 15** | **SB** |
|  | Multiple sequence alignment  uses of multiple sequence alignment | **Lecture 16** | **SB** |
|  | Multiple sequence alignment: Tools and resources | **Practical-6** | **SB** |
|  | programs and methods for multiple sequence alignment  pattern searching programs | **Lecture 17** | **SB** |
|  | family and superfamily representation  structural inference | **Lecture 18** | **SB** |
|  | dynamic programming in r-dimensions  weighted average sequences  profile analysis | **Lecture 19 & 20** | **SB** |
|  | Alignment by hidden Markov models | **Lecture 21** | **SB** |
|  | **consensus word analysis**  **more complex scoring** | **Lecture 22** | **SB** |
|  | **Phylogenetic prediction** | **Lecture 11 & 12** | **SS** |
|  | Trees-splits and metrices on tress | **Lecture 13** | **SS** |
|  | tree interpretation | **Lecture 14** | **SS** |
|  | Distance - additive, ultrameric and nonadditive distances | **Lecture 15 & Lecture 16** | **SS** |
|  | Phylogenetic analysis: Tools and resources | **Practical-7** | **SB** |
|  | tree building methods | **Lecture 17** | **SS** |
|  | phylogenetic analysis | **Lecture 18** | **SS** |
|  | estimating the rate of change | **Lecture 19** | **SS** |
|  | likelihood and trees | **Lecture 20 & 21** | **SS** |
|  | analysis software. | **Lecture 22** | **SS** |
| **SUGGESTED READINGS:**  1 D. Gusfield, “Algorithms on Strings, Trees, and Sequences: Computer Science and Computational Biology”, Cambridge University Press.  2. Biocomputing Hypertext Coursebook at http://www.techfak.unibielefeld.de/bcd/Curric/welcome.html/  3. A.D. Baxevanis and B.F.F. Ouellette, “Bioinformatics: A Practical Guide to the Analysis of Genes and Proteins”, Wiley-Interscience.  4. D.W.Mount, “Bioinformatics: Sequence and Genome Analysis”, Cold Spring Harbor Laboratory Press. | | | |
| 1. Writing a sequence assembly program  2. Implementation of a selected sequence alignment algorithm  3. Introduction to NCBI: sequence databases, sequence retrieval, sequence file formats  4. Introduction to NCBI: Literature and other resources  5. Introduction to NCBI: Structure and other resources  6. Pairwise comparisons: - effect of different substitution matrices, change in gap penalties  7. Database search: Evaluate the statistical significance of the match with a web program. Effect of presence of low complexity regions in the sequence and filtering.  8. Multiple sequence alignment: Tools and resources  9. Phylogenetic analysis: Tools and resources | | | |