

## **BIOINFORMATICA (BIO/11) 6 CFU**

Dr. Ausiello *Programma*

Matrici per gli allineamenti (Rappresentazione di un allineamento con una matrice, Matrice di punti, Ricerca del migliore allineamento su matrice, Regioni di identità, duplicazioni e inserzioni sulle matrici). Programmazione dinamica (Algoritmo di programmazione dinamica, Programmazione dinamica con gaps, Needleman & Wunsch). Allineamenti locali e globali (Differenza fra allineamenti locale e globale, Significato biologico di locale e globale, Allineamenti locali e globali su di una matrice, Smith & Waterman). Matrici di sostituzione (Frequenze degli aminoacidi e composizione Amminoacidica, Probabilità di sequenze e allineamenti casuale, Misura delle sostituzioni casuali, Misura delle sostituzioni osservate). Costruzione di una matrice di punteggio (Rapporto fra sostituzioni osservate/casuali, Matrice di probabilità e probabilità di un allineamento, Logaritmo del rapporto osservate/casuali e matrice di punteggio). Tipi di matrici di sostituzione (Distanza di una matrice di sostituzione, Matrici BLOSUM, Matrici PAM). Parole di aminoacidi (Composizione in parole di una sequenza, Allineamenti con parole, Matrici di parole, Parole simili). Ricerche in banche dati (Generalità, Algoritmi euristici, Allineamenti euristici, Descrizione dell'algoritmo FastA, Descrizione dell'algoritmo Blast, Significatività statistica del punteggio). Distanza fra sequenze (Similarità e distanza, Distanza genetica, Distanza Jukes & Cantor, Matrici delle distanze). Alberi filogenetici (Descrizione di un albero, Ortologi e paraloghi, Tipi di alberi, Ipotesi dell'orologio molecolare, Test distanze ultrametriche). Costruzione di un albero filogenetico (Algoritmi di clustering, Clustering gerarchico e additivo, Distanza fra clusters, Metodo UMPGMA). Allineamenti multipli (Descrizione, Utilità, Colorazione, Consensus, Punteggio di una colonna, Punteggi alternativi, Somma delle coppie, Peso delle sequenze, Somma pesata delle coppie). Costruzione di un allineamento multiplo (Programmazione dinamica in più dimensioni, Albero guida, Allineamento progressivo (CLUSTALW), Allineare fra di loro due allineamenti, Metodi iterativi). Qualità di motivi funzionali (Veri/falsi e Positivi/Negativi, Sensitività, Selettività, Curva ROC, Coefficiente di correlazione). Matrici posizionali di peso (PSSM, Ricerche con PSSM, Logo di una sequenza). Profili (Profilo di un allineamento multiplo, Ricerca con un profilo, PSI- BLAST). Ricerca di geni (Struttura di un gene, Analisi Schemi di lettura aperti, Metodi estrinseci). Analisi di contenuto (Frequenze esanucleotidi, Altri contenuti, Punteggio delle regioni codificanti). Analisi di segnale (Siti di splicing, Altri segnali, Integrazione con analisi di contenuto). Modello dei geni (Ordinamento corretto degli esoni, Allenamento dei metodi, Valutazione della prestazioni dei metodi). Ricerca dei promotori (Analisi del contenuto, Analisi dei segnali, Impronta filogenetica, Moduli cis regolativi). Metodi di predizione delle interazioni molecolari. Analisi dei microRNA. Text mining. Ontologie.