

Testi del Syllabus

Resp. Did. **GERDOL MARCO** **Matricola: 014720**

Docenti **GERDOL MARCO, 4,5 CFU**
PALLAVICINI ALBERTO, 1,5 CFU

Anno offerta: **2018/2019**

Insegnamento: **644SM - BIOINFORMATICA**

Corso di studio: **SM51 - SCIENZE E TECNOLOGIE BIOLOGICHE**

Anno regolamento: **2016**

CFU: **6**

Settore: **BIO/18**

Tipo Attività: **B - Caratterizzante**

Anno corso: **3**

Periodo: **Secondo Semestre**



Testi in italiano

Lingua insegnamento ITALIANO

Contenuti (Dipl.Sup.) Introduzione alla Bioinformatica. Gestione ed analisi delle informazioni biologiche. Database campi e livelli di specializzazione. Information retrieval e interrogazioni complesse. Allineamento di sequenze di acidi nucleici e proteine. Allineamento dinamico ed euristico. Matrici di sostituzione e FASTA. Applicazioni principali di BLAST. Applicazioni speciali di BLAST. Allineamenti multipli, algoritmi iterativi e progressivi. Allineamento di profili (PSI-BLAST). Pattern e motivi funzionali nel DNA. Pattern e motivi funzionali in proteine. Modeling strutturale di sequenze proteiche. Basi di analisi filogenetica, metodi basati su matrici di distanze e caratteri.

Testi di riferimento Pascarella, Paiardini. Bioinformatica: dalla sequenza alla struttura delle proteine. Zanichelli (2011)

Valle, Helmer-Citterich, Attimonelli, Pesole. Introduzione alla Bioinformatica. Zanichelli

Lesk. Introduzione alla Bioinformatica. McGraw-Hill

Krane, Raymer. Fondamenti di Bioinformatica. Pearson

Mount D. Bioinformatics. Sequence and Genome Analysis. Second Edition. CSHL Press

Baxevanis, Oullette. Bioinformatics. A practical guide to the analysis of genes and proteins. Wiley

Obiettivi formativi Il corso si prefigge di permettere agli studenti di acquisire le abilità per utilizzare in modo razionale ed efficiente i metodi bioinformatici e per interpretare correttamente i risultati.

Conoscenza e comprensione

- conoscere i potenziali campi di applicazione delle bioinformatica
- conoscere l'organizzazione dei principali database bioinformatici ed apprendere il tipo di contenuti che è possibile ritrovarvi
- apprendere quali sono gli strumenti più utili per ricercare una sequenza

all'interno di un database a seconda dell'applicazione di interesse

- conoscere i principali genome browser e comprendere la loro struttura ed il loro utilizzo
- comprendere il significato e le implicazioni biologiche di un allineamento multiplo di sequenze
- conoscere quale è lo stato dell'arte dei metodi di predizione strutturale delle proteine
- apprendere come interpretare un albero filogenetico

Capacità di applicare conoscenza e comprensione

Gli studenti sapranno effettuare ricerche semplici e complesse di letteratura e di sequenza nei principali database bioinformatici. Acquisiranno competenze nell'utilizzo di BLAST per ricercare similarità tra sequenze in genomi, trascrittomi e vari database di sequenza, nella consultazione di genome browser e nella caratterizzazione funzionale di sequenze proteiche. Inoltre, saranno capaci di generare allineamenti multipli da analizzarsi per la ricerca di motivi conservati, siti sotto pressione selettiva e generazione di alberi filogenetici.

Autonomia di giudizio

L'autonomia di giudizio viene sviluppata tramite la preparazione all'esame, che necessita della rielaborazione e assimilazione individuale del materiale presentato in aula, oltre che tramite il coinvolgimento diretto degli studenti nelle dimostrazioni pratiche dell'utilizzo di alcuni tools bioinformatici descritti nelle lezioni precedenti.

Abilità comunicative

Gli studenti saranno stimolati ad interagire con il docente durante le lezioni, al fine di sviluppare un corretto lessico scientifico e favorire un confronto tra gli studenti stessi. Le ultime lezioni saranno principalmente caratterizzate da dimostrazioni pratiche di quanto affrontato nelle lezioni precedenti, di modo da dare un ruolo di primo piano agli studenti stessi nella risoluzione di problemi.

Capacità di apprendimento

La capacità di apprendimento è stimolata dall'approfondimento delle conoscenze apprese durante le lezioni frontali tramite 'interazione con il docente e tra gli studenti stessi nelle ultime lezioni del corso, che prevedono l'applicazione pratica di concetti visti teoricamente in precedenza, che daranno modo di approfondire punti poco chiari e consentire una migliore comprensione di argomenti ostici per quali le conoscenze pregresse acquisite da altri corso non dovessero essere sufficienti a garantire una completa comprensione.

Prerequisiti

Genetica

Metodi didattici

Lezione frontale coadiuvata da presentazioni ed esercitazioni.

Altre informazioni

Il materiale didattico verrà messo a disposizione sulla piattaforma moodle.

Modalità di verifica dell'apprendimento

Test scritto con domande a risposta multipla

Programma esteso

Definizione di bioinformatica e potenziali campi di applicazione, con particolare attenzione agli ambiti genomici. Database ed archivi di dati biologici; struttura e tipi, campi principali di una entry; EBI, NCBI e DDBJ. Approfondimento sui database di letteratura Pubmed e Scopus. Ricerche semplici ed avanzate nei database. Funzione e struttura dei genome browser; UCSC ed Ensembl, NCBI genomes. Analisi di sintenia. Human protein atlas. Concetti di similarità ed omologia, geni ortologhi e paraloghi. Algoritmi di allineamento dinamico, allineamenti locali e globali. Algoritmo di Needleman-Wunsch a di Smith-Waterman. Matrici di

sostituzione PAM e BLOSUM e loro utilizzo. Da FASTA a BLAST e valutazione della significatività di un allineamento. BLAST, funzionamento e parametri. BLASTn, BLASTp, tBLASTn, BLASTx e tBLASTx, applicazioni particolari. Interfaccia web ed interpretazione dei risultati. BLAST e BLAST locali ed integrati in genome browsers. Concetto di profilo proteico ed Hidden Markov Models: definizione e possibili utilizzi. PSI-BLAST, DELTA-BLAST. Ricerca di domini conservati con Interproscan, organizzazione di Interpro. Pattern matching/recognition e pattern discovery; applicazione nello studio di promotori. PHI-BLAST, ELM e MEME suite. Caratterizzazione funzionale di proteine: signalP, Phobius, TargetP, SecretomeP, proP, TMHMM. Predizione di struttura secondaria: metodi GOR, Chou-Fassman e machine-learning. Predizione di strutture 3D di proteine. Metodi ab initio e metodi knowledge-based. Homology-modeling e threading. ROSETTA, CASP e CAPRI. Predizione di struttura di RNA con ViennaRNA. Allineamento multiplo di sequenze; penalità dovute a gaps, metodi di allineamento progressivo: CLUSTALW e T-COFFEE. Metodi di allineamento iterativo, MUSCLE. Metodi basati su blocchi, metodi probabilistici e metodi consensus-based. Allineamenti profilo-profilo. Valutazione della qualità di un allineamento multiplo. Identificazione di siti sotto selezione positiva e negativa. Definizione di filogenetica e filogenesi, terminologia relativa ad alberi filogenetici. Concetti di monofilia, polifilia e parafilia. Tipi di albero, radicazione e scelta degli outgroup. Scelta dei marker molecolari e del modello di evoluzione molecolare. Metodi clustering-based: NJ ed UPGMA. Alberi optimality-based. Metodi basati sui caratteri: maximum parsimony e maximum likelihood. Metodi Bayesiani. Valutazione della solidità di un albero: bootstrap e jackknife. Alberi time-calibrated e BEAST. Cenni sugli approcci filogeneomici.