UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Bioinformatic Methodologies

2526-1-F0803Q054

Obiettivi

Obiettivi formativi
Conoscenze:
• Conoscere le principali metodologie computazionali per l'analisi di dati biologici, in particolare dati omici.
• Comprendere le basi teoriche e gli algoritmi sottostanti agli strumenti di analisi bioinformatica.

Competenze:
• Saper utilizzare strumenti bioinformatici per l'elaborazione e l'interpretazione di dati biologici.
• Saper selezionare e applicare metodologie appropriate a seconda del problema biologico posto.

Risultati di apprendimento attesi
Al termine del corso lo studente sarà in grado di:

- *Conoscenza e capacità di comprensione*: descrivere i principi fondamentali delle principali metodologie bioinformatiche, con riferimento ad applicazioni in genomica, trascrittomica e analisi di dati omici.
- *Capacità di applicare conoscenza e comprensione*: utilizzare ambienti e strumenti software per l'analisi bioinformatica di dataset reali, interpretando criticamente i risultati ottenuti.
- *Autonomia di giudizio:* valutare l'idoneità di approcci metodologici differenti rispetto a specifici quesiti biologici o clinici, proponendo soluzioni alternative se necessario. Questa abilità viene sviluppata attraverso attività interattive in aula, tra cui discussione guidata di casi studio, esercizi interpretativi basati su dati reali ed esempi pratici svolti collettivamente.
- *Abilità comunicative:* comunicare in modo chiaro e coerente le scelte metodologiche e i risultati dell'analisi, anche in forma scritta o attraverso presentazioni orali.
- *Capacità di apprendimento*: acquisire autonomia nell'approfondimento di strumenti e risorse bioinformatiche non trattate durante il corso, aggiornandosi con la letteratura scientifica e la documentazione tecnica.

Contenuti sintetici

• Introduzione alla bioinformatica
• Pillole di Stiatistica e Intelligenza Artificiale
• La generazione dei dati: dalle piattaforme di sequenziamento all'assemblaggio e annotazione del genoma, fino alla multi-omica
• Organizzazione e gestione dei dati biologici
• Analisi di sequenze
• Analisi del trascrittoma
• Integrazione dati multi-omici

Programma esteso

1. Introduzione alla bioinformatica
2. Pillole di Intelligenza Artificiale e statistica
   a Statistica per dati ad alta dimensionalità: test multipli, p-value, FDR
   b. Machine learning supervisionato
   c. Machine learning non supervisionato
3. La generazione dei dati
   a. Piattaforme di sequenziamento degli acidi nucleici
   b. Dal cromatogramma alle reads: il "base calling"
   c. Coverage, qualità delle read, formato dei dati
   d. Dalle read alla sequenza: algoritmi di assemblaggio
   e. Annotazione del genoma
4. Analisi di sequenze
   a. Allineamento locale e globale
   b. Algoritmi esatti ed euristici
5. Organizzazione e gestione dei dati:
   a. Database e DBMS: database relazionali e flat file
   b. Banche di dati biologiche
   i. Banche dati genomiche (GenBank - ENA – DDBJ)
   ii. Banche dati proteomiche (UniprotKB, Swiss-Prot, TrEMBL – PDB)
   iii. Genome browsers: ENSEMBL, UCSC
6. Analisi del trascrittoma
   a. Dal sequenziamento dell'RNA ai valori di espressione
   b. Pre-processamento e analisi dati di single-cell RNA sequencing
7. Integrazione dati multi-omici
   a. Analisi di arricchimento e visualizzazione integrata di dati trascrittomici, proteomici e metabolomici
   b. Integrazione di dati omici in modelli di rete metabolica

Prerequisiti

Le conoscenze di Biologia cellulare, molecolare e biochimica, acquisite nei corsi base di una Laurea triennale in Scienze Biologiche o in Biotecnologie saranno date per assodate.

Conoscenze pregresse utili che saranno solo brevemente riprese durante il corso:
• Elementi di architettura dei calcolatori e algoritmi
• Elementi di probabilità e statistica

Modalità didattica

42 ore di lezioni frontali, così suddivise:
– 36 ore in modalità erogativa in presenza, supportate da presentazioni in aula;
– 6 ore in modalità interattiva, dedicate all’elaborazione e discussione guidata in aula delle proposte di progetto finale da parte degli studenti.

10 ore di esercitazioni pratiche in presenza, finalizzate all’utilizzo di banche dati biologiche e piattaforme di analisi bioinformatiche.

Materiale didattico

Slides e videoregistrazioni delle lezioni e delle esercitazioni reperibili sulla pagina e-learning dell'insegnamento.
Libro di testo suggerito: Citterich, Ferré, Pavesi, Romualdi, Pesole. Fondamenti di Informatica. BIOLOGIA ZANICHELLI
Articoli specialistici e di rassegna e capitoli di libro verranno consigliati a lezione e caricati sulla piattaforma e-learning del corso.

Periodo di erogazione dell'insegnamento

Primo semestre

Modalità di verifica del profitto e valutazione

Progetto di ricerca individuale o di gruppo su un argomento a scelta dello studente che preveda l'utilizzo critico di alcuni strumenti visti a lezione
+
Esame orale in cui lo studente presenta il progetto di ricerca e dimostra la padronanza dei concetti utilizzati nel progetto. Durante l'esame orale saranno inoltre proposti esercizi (da svolgere su carta) di allineamento di sequenze.

Non sono previste prove intermedie per i frequentanti.

Orario di ricevimento

Gli studenti sono invitati a contattare il docente per email per accordarsi su data e giorno (eventualmente via WebEx)

Sustainable Development Goals