UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

SYLLABUS DEL CORSO

Bioinformatica

2526-1-F0602Q114

Obiettivi

L’insegnamento è focalizzato sulle tecnologie genomiche, su come queste sono utilizzate per la generazione di dati di tipo “-omico”, e sui principali concetti e insidie nell’analisi di questi dati. Il corso si focalizza sulle tecnologie basate su approcci di sequenziamento massivo. Particolare enfasi sarà dedicata allo studio dei programmi di espressione genica, di fattori che legano il DNA e delle modificazioni epigenetiche della cromatina.

Conoscenza e capacità di comprensione.
Gli studenti familiarizzeranno con le principali tecnologie genomiche e i principali tipi di dati. Impareranno le migliori pratiche e le principali modalità di analisi nel campo della genomica.
Una serie di lezioni interattive saranno dedicate a Galaxy, una suite di strumenti di analisi di dati omics gratuitamente disponibile online. Al termine del corso, gli studenti acquisiranno i concetti chiave nel campo della bioinformatica e della genomica, e familiarizzeranno e con risorse di riferimento alle quali potranno tornare durante future attività di ricerca.

Capacità di applicare conoscenza e comprensione.
Lo studente sarà in grado di applicare le conoscenze acquisite nei successivi insegnamenti ed in esperienze di laboratorio e di utilizzare la capacità di comprensione ai fini di successive attività di studio e/o di ricerca.

Autonomia di giudizio.
Lo studente sarà in grado di elaborare quanto appreso e saprà riconoscere le situazioni e i problemi in cui le conoscenze apprese possano essere utilizzate.

Abilità comunicative.
Alla fine dell'insegnamento, lo studente saprà descrivere con proprietà di linguaggio e sicurezza di esposizione argomenti inerenti la genomica e la bioinformatica.

Capacità di apprendimento.
Alla fine dell'insegnamento, lo studente sarà in grado di consultare la letteratura sugli argomenti trattati e saprà analizzare, applicare, integrare e collegare le conoscenze acquisite con quanto verrà appreso in insegnamenti

correlati.

Contenuti sintetici

1. I programmi di espressione genica e I loro determinanti
2. Dalle cellule alla libreria di sequenziamento
3. Tecnologie per il sequenziamento massivo di acidi nucleici
4. Processamento di dati di sequenziamento
5. Analisi di dati ChIP-seq
6. Analisi di dati RNA-seq
7. Genome-Browsers
8. Clustering e heatmaps
9. Introduzione a Galaxy
10. Analisi di dati ChIP-seq tramite Galaxy
11. Analisi di dati RNA-seq tramite Galaxy
12. Dinamiche di regolazione trascrizionale
13. Studio dei programmi di espressione genica a livello della singola cellula
14. Studio della conformazione della cromatina

Programma esteso

1. I programmi di espressione genica e I loro determinanti: (i) programmi di espressione genica, (ii) legame di fattori regolativi alla cromatina, (iii) modificazioni epigenetiche, (iv) RNA polimerasi, (v) modificazioni dell'RNA
2. Dalle cellule alla libreria di sequenziamento: ChIP-seq, RIP-seq e RNA-seq
3. Tecnologie per il sequenziamento massivo di acidi nucleici: (i) sequenziamento a reads corte (Ion Torrent, Illumina), (ii) sequenziamento a reads lunghe (Pacific Biosciences, Oxford Nanopore Technologies)
4. Processamento di dati di sequenziamento: (i) FASTQ file, (ii) controlli di qualità, (iii) allineamento delle reads, (iv) SAM/BAM file, (v) ispezione degli allineamenti tramite un genome-browser
5. Analisi di dati ChIP-seq: (i) identificazione dei picchi, (ii) falsi positivi e falsi negativi, (iii) arricchimento di motivi, (iv) GeneOntology e arricchimento di insiemi di geni, (v) identificazione dei picchi differenziali
6. Analisi di dati RNA-seq: (i) espressione assoluta e normalizzazione, (ii) espressione differenziale, (iii) varianti di splicing, (iv) modificazioni dell'RNA
7. Genome-Browsers: IGV/IGB, UCSC genome-browser, caricamento dei dati, ottenimento di immagini, integrazione con dati ENCODE, UCSC Table
8. Clustering e heatmaps
9. Introduzione a Galaxy: dai picchi ai geni
10. Analisi di dati ChIP-seq tramite Galaxy
11. Analisi di dati RNA-seq tramite Galaxy
12. Dinamiche di regolazione trascrizionale: (i) RNA nascente, (ii) regolazione post-trascrizionale, (iii) ciclo vitale della RNA polimerasi
13. Studio dei programmi di espressione genica a livello della singola cellula, scRNA-seq
14. Studio della conformazione della cromatina

Prerequisiti

Prerequisiti: nozioni fondamentali di Biologia Molecolare.
Propedeuticità: nessuna.

Modalità didattica

- 12 lezioni da 2 ore svolte in modalità erogativa in presenza;
- 9 lezioni da 2 ore in presenza svolte in modalità erogativa nella parte iniziale che è volta a coinvolgere gli studenti in modo interattivo nella parte successiva. Le lezioni svolte in modalità interattiva prevedono l'utilizzo del sistema di analisi dati Galaxy. Questa parte include un *project work* facoltativo da parte degli studenti, individuale o di gruppo, riguardante analisi svolte in autonomia con Galaxy. Il *project work* si conclude con una breve presentazione da parte degli studenti e costituisce una prova parziale d'esame.

Materiale didattico

Slide, articoli scientifici e registrazioni delle lezioni reperibili sulla piattaforma e-learning dell'insegnamento.
Testi consigliati:

- M.H. Citterich et al. "Fondamenti di Bioinformatica" Zanichelli

Periodo di erogazione dell'insegnamento

Secondo semestre

Modalità di verifica del profitto e valutazione

Colloquio sugli argomenti svolti a lezione. La discussione può, a scelta dello studente, includere la discussione di un articolo scientifico pertinente agli argomenti trattati durante il corso, valutando la capacità e l'analisi critica di interpretazione dei risultati.
La valutazione integra il risultato della prova parziale, per gli studenti che hanno scelto di svolgerla, un *project work* facoltativo finalizzato al controllo delle competenze di *problem solving* disciplinare.

Orario di ricevimento

Ricevimento su appuntamento via e-mail con il docente.

Sustainable Development Goals

SALUTE E BENESSERE