UNIVERSITÀ DI PISA

ANALISI GENETICHE E GENOMICHE

ROBERTO MARANGONI

Academic year	2022/23
Course	BIOLOGIA MOLECOLARE E CELLULARE
Code	176EE
Credits	6

Modules	Area	Type	Hours	Teacher(s)
ANALISI GENETICHE E GENOMICHE	BIO/18	LEZIONI	48	STEFANO LANDI ROBERTO MARANGONI

Obiettivi di apprendimento

Conoscenze
Il corso di Analisi Genetiche e Genomiche si pone l'obiettivo di formare laureati con una specifica competenza nelle tematiche e nelle metodologie caratterizzanti la Genetica, la Genomica e la Trascrittomica nelle loro applicazioni per la ricerca di base e applicata.

Modalità di verifica delle conoscenze
La verifica delle conoscenze avviene con esame scritto e orale

Capacità
Al termine del corso lo studente sarà in grado di imparare un metodo per allineare sequenze derivanti dai sequenziatori di nuova generazione (Next Generation Sequencing) e disegnare alcuni test molecolari di analisi del DNA, nonchè di contestualizzare l'utilizzo e l'applicazione di una determinata metodica genomica, genetica e di biologia molecolare, con particolare riferimento al campo della biomedicina. Lo studente apprenderà a usare l'ambiente R e in particolare la suite Bioconductor per l'elaborazione di pipeline di analisi su dati genome-wide.

Modalità di verifica delle capacità
La verifica delle capacità acquisite avverrà in sede di esame tramite domande aperte o esercizi mirati

Comportamenti
Al termine del corso lo studente avrà le basi necessarie (da approfondire) per progettare uno studio di indagine genetica e molecolare su malattie mendeliane e complesse .

Modalità di verifica dei comportamenti
I comportamenti acquisiti verranno verificati in sede di esame.

Prerequisiti (conoscenze iniziali)
Le conoscenze della genetica di base, biologia molecolare e alcune nozioni di statistica sono essenziali. E' utile, ma non indispensabile, la conoscenza dell'ambiente R.

Indicazioni metodologiche
Le lezioni sono tutte di tipo frontale.

Programma (contenuti dell'insegnamento)
Introduzione al corso. Aggiornamento delle basi della Genetica Tipi di polimorfismi nei genomi. Minisatelliti, microsatelliti. DNA fingerprinting, Instabilità dei microsatelliti. Il modello Slippage-disallineamento. Neurodisturbi causati dall'espansione aberrante dei microsatelliti tripletti. SNP e INS/DEL. Metodi di scoperta: Polimorfismo di conformazione a filamento singolo (SSCP). Fusione ad alta risoluzione (HRM). La reazione di sequenziamento di Sanger. Genotipizzazione: DOT BLOT. Genotipizzazione mediante RFLP (polimorfismo della lunghezza dei frammenti di restrizione) e PCR-RFLP, ASO-PCR (o ARMS-PCR), OLA (Oligonucleotide Ligation Assay), MALDI-TOF, TaqMan Allelic Discrimination Assay.

Microarray per la genotipizzazione: la Single Base Extension (principio generale). Illumina® BeadArray. Microarray per la genotipizzazione: ibridazione con Affymetrix Un breve tour in dbSNPs, gnomAD, UCSC Genome Browser, i database più utilizzati per sfogliare le variazioni genetiche. Duplicazioni segmentali. Obbligatorio e facoltativo. Meccanismi di formazione: crossing-over ineguale, duplicazioni dell'intero genoma. riarrangiamenti cromosomici. I loci del CYP2D6 come esempio di duplicazioni obbligatorie e polimorfiche e delezioni interstiziali nel genoma umano. L'esempio del CYP2D6 nel metabolismo degli antidepressivi e di altri farmaci. I loci di GSTM1, GSTT1 e TP53 come esempio di duplicazioni obbligatorie e polimorfiche e delezioni interstiziali nel genoma umano. Genotipizzazione di delezioni interstiziali polimorfe e piccole inserzioni mediante elettroforesi su gel di prodotti di PCR. Analisi mediante multiplex Ligation-dependent Probe Amplification (chiamato anche Multiplex Oligonucleotide Ligation Assay), analisi mediante saggio TaqMan (PCR quantitativa in tempo reale). In che modo i polimorfismi sono correlati tra loro all'interno dei genomi? Aplotipi e Linkage Disequilibrium (LD), definizioni. La storia delle migrazioni umane, il concetto di blocchi LD. Siamo un mosaico di cromosomi ancestrali. Calcolo della forza LD tra due polimorfismi utilizzando i parametri D e r2. Le forze che modellano la variabilità aplotipica e modulano l'intensità del linkage disequilibrium. Metodi molecolari e non molecolari per definire un aplotipo. Il progetto HAPMAP (www.hapmap.org). Gli SNP che codificano gli aplotipi. Un esempio visivo. Perché non abbiamo bisogno di genotipizzare tutti i polimorfismi nel genoma. Il database HaploReg, che elenca gli SNP con il loro reciproco LD. Un uso pratico degli array SNP per rilevare delezioni/duplicazioni nei genomi. L'esempio della distrofia muscolare di Duchenne. Dalle prime sonde (Southern blot), ai microsatelliti, analisi MLPA e array SNP. Utilizzo di array SNP per la mappatura di delezioni ereditarie e de novo in probandi affetti da tratti mendeliani/oligogeni. Combinazione di diversi pazienti con lo stesso fenotipo per affinare la regione critica minima per una malattia. Altri esempi sulle delezioni rilevate dagli array SNP. Introduzione all'analisi di linkage per la mappatura dei tratti mendeliani. Storia (malattia di Huntington). Il contenuto informativo di un polimorfismo (PIC) e l'eterozigosi come proxy per PIC. Perché questo parametro è importante per gli studi di linkage. Rilevamento di famiglie ricombinanti per il calcolo dei punteggi LOD nei tratti dominanti. Esempio di calcolo di un punteggio LOD in due famiglie. Il punteggio lod multipunto (MLS). Introduzione all'analisi di linkage per i caratteri recessivi. Mappatura dell'autozigosi che sfrutta i segmenti di identità per discendenza (IBD) nelle famiglie. Esempio di mappatura del gene per la fibrosi cistica sfruttando i segmenti IBD a livello di popolazione. Introduzione al sequenziamento di nuova generazione. La reazione di pirosequenziamento e il sistema 454. 454/GS_FLX – PCR a base di emulsione. Torrente ionico. Illumina®, l'amplificazione a ponte e il sequenziamento per sintesi. Metodi NGS a singola molecola: Oxford Nanopore e Pacific-Biosciences. Applicazioni: sequenziamento dell'intero genoma (WGS), sequenziamento dell'intero esoma (WES), sequenziamento dell'RNA (RNA-seq). Illustrazione grafica della mappatura delle "letture" in un browser genoma. Copertura irregolare del genoma. Relazione tra "profondità" e frazione di genoma non sequenziato. Relazione tra numero di letture e frazione del genoma non sequenziato. Relazione tra profondità e copertura. Relazione tra profondità e numero di "variazioni genetiche" rilevabili in un genoma. Possibile uso "quantitativo" di RNA-seq. RNA-seq per la rilevazione delle isoforme dell'mRNA. NGS per rilevare CNV. I progetti 1000Genomes, ExAC e GnomAD. Il progetto di sequenziamento dell'esoma del progetto NHLBI. Esempio di mappatura per l'autozigosi usando NGS. Sequenziamento dell'esoma vs sequenziamento dell'intero genoma. Quali informazioni possiamo comprendere? Il progetto ENCODE.

prof. Marangoni: i 3 crediti sono focalizzati sulle metodiche bioinformatiche di analisi ed elaborazione dei dati di NGS. Lo strumento fondamentale è rappresentato dalla suite Bioconductor che permette di implementare la quasi totalità delle pipeline normalmente impiegate per l'analisi di dati molecolari. I principali punti che verranno coperti sono:
- Introduzione all'ambiente R (elementi fondamentali del linguaggio, tipi di variabili, costrutti)
- L'ambiente "tidy R": caratteristiche e usabilità per le applicazioni bioinformatiche
- La suite Bioconductor: istallazione e caratteristiche fondamentali
- Acquisizione di dati NGS
- Un esempio di workflow di mappatura e genotipizzazione
- Esempi di workflow per analisi di dati NGS genome-wide

Bibliografia e materiale didattico

1) Libri di testo consigliati:
"Genetica molecolare umana ", by Tom Strachan & Andrew P. Read (Zanichelli)
Powepoint forniti dai docenti

Indicazioni per non frequentanti

Tutto il materiale si puo' trovare su e-learning e sulla pagina
http://www.stefanolandi.eu

Modalità d'esame

Prova d'esame scritta e prova oral

Pagina web del corso

http://www.stefanolandi.eu

Altri riferimenti web

cercare su e-learning e Moodle

Ultimo aggiornamento 29/07/2022 11:26