



# UNIVERSITÀ DI PISA

---

## ANALISI GENETICHE E GENOMICHE

**STEFANO LANDI**

Academic year 2016/17  
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 |

### Obiettivi di apprendimento

#### *Conoscenze*

Gli studenti dovranno arrivare a conoscere

- (1) le principali metodiche di indagine molecolare del DNA
- (2) come affrontare lo studio di caratteri mendeliani e non mendeliani (caratteri complessi) al fine di identificare i geni implicati

#### *Modalità di verifica delle conoscenze*

La verifica delle conoscenze avviene con esami scritti (orale opzionale)

#### *Capacità*

Al termine del corso lo studente sarà in grado di disegnare alcuni test molecolari di analisi del DNA e sarà in grado di interpretare la letteratura scientifica più recente riguardo agli studi di genetica delle malattie mendeliane e complesse

#### *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 su qualsiasi tipo di tratto, mendeliano e non.

#### *Modalità di verifica dei comportamenti*

In sede dei comportamenti acquisiti avviene in sede di esame tramite esercizi mirati

#### *Prerequisiti (conoscenze iniziali)*

Le conoscenze della genetica di base (del corso triennale) sono essenziali. Alcune nozioni di statistica, come i test del chi-quadro, l'analisi della varianza e la regressione lineare semplice sarebbero preferibili.

#### *Corequisiti*

Biochimica

#### *Prerequisiti per studi successivi*

GLi argomenti trattati nel corso aiuteranno l'approfondimento qualora lo studente intraprendesse un Dottorato di Ricerca

#### *Indicazioni metodologiche*

Le lezioni sono tutte di tipo frontale ed includono anche ore di esercitazione basate sugli esami proposti negli anni passati.

#### *Programma (contenuti dell'insegnamento)*

Presentazione del corso.



## UNIVERSITÀ DI PISA

Libri di testi consigliati, illustrazione del programma e del sito internet ad esso collegato.

La PCR. Introduzione sul principio di funzionamento. Visualizzazione del prodotto di PCR. Il disegno dei primers. La  $T_m$  dei primers. Multiplex PCR. La marcatura dei prodotti di PCR. Aggiunta di code al 5'.

La PCR asimmetrica. Il metodo diretto del sequenziamento di DNA (reazione di Sanger). La nested PCR. La semi-nested. La ASO-PCR. La PCR inversa. PCR hotstart. Touch-down PCR. RT-PCR. Long range PCR.

Primers mutagenetici. Un esempio di mutagenesi sito-specifica con primers mutagenetici: il "QuickChange Lightning Site-Directed Mutagenesis".

Whole-genome amplification (WGA). DOP-PCR. Adaptor-ligation PCR. Inter-Alu PCR. Phi29 DNA-polymerase.

La quantificazione dell'espressione genica. Il Northern Blot. La Real time quantitativa (qPCR).

Molecular Beacons e Sonde Taqman. Utilizzo delle giunzioni esone-esone per il posizionamento dei primers in qPCR.

Ultime considerazioni:

1) i primers devono essere unici, verificati con BLAST

2) Origine di contaminazione delle PCR

Calcolo della quantità di espressione con metodo qPCR. Quantificazione relativa e assoluta (metodo del DeltaDelta-Ct, approssimato).

Esempio di utilizzo di qPCR.

Introduzione alla digital droplet PCR (ddPCR). Principio di funzionamento della ddPCR.

Esempi di utilizzo. Aspetti matematici della ddPCR.

I numeri notevoli dell'RNA. Micro-array per analisi di trascrittomiche. Tipi di sonde: a cDNA, 65-mers e 25-mers. Metodo competitivo a due colori.

Metodi a singolo colore. Sintesi di sonde nell'array Affymetrix (metodo di stampa fotolitografica). Cenni sul processamento dell'immagine:

sottrazione del background. Normalizzazione dei dati (Z-score). Pitfalls: specificità delle sonde, condizioni di stringenza dell'ibridazione, strutture secondarie delle sonde, uniformità delle condizioni di ibridazione, retro-trascrizione dell'mRNA senza amplificazione (PCR) al fine di non distorcere il contenuto relativo dei trascritti.

Cenni sui metodi di raggruppamento (hierarchical clustering) a due vie (per geni e per campioni).

Esempi di applicazioni del gene profiling: lettura consigliata da Nature,403(6769):503-511.

Tipologie di polimorfismi genetici. Frequenze alleliche, differenza tra mutazioni e polimorfismi. Clini di frequenze alleliche. Minisatelliti.

Microsatelliti. Micro-insersioni. Micro-delezioni.

Metodi di indagine dei minisatelliti. Metodi di indagine dei microsatelliti. DNA fingerprinting. DNA profiling in genetica forense. Evoluzione delle sequenze ripetute in tandem (mini e micro satelliti).

Micro-satellite instability. Slippage-misalignment. Polimorfismi a singolo nucleotide, micro-insersioni, micro-delezioni. Metodi di rilevamento (discovery e genotyping). Reazione di sequenziamento di Sanger. Rilevazione di mutazioni tramite dHPLC. Single strand conformation polymorphisms (SSCP).

Genotipizzazione mediante PCR-RFLP (restriction fragment length polymorphism), ASO-PCR/ARMS (amplification refractory mutation system), oligonucleotide ligation assay, MALDI-TOF, TaqMan Allelic Discrimination Assay, DOT-BLOT. Gli array di genotipizzazione.

Metodica Arrayed Primer Extension Assay (APEX). Genotipizzazione tramite Illumina BeadArray. Il Bead decoding. Genotipizzazione tramite Affymetrix GeneChip. Il pattern delle sonde MM e PM.

Le duplicazioni segmentali. Definizione. Duplicazioni segmentali costitutive. Evoluzione tramite duplicazione di intero genoma, crossing-over ineguale e riarrangiamento cromosomico. Pseudogeni.

Esempi di delezioni e duplicazioni polimorfiche: CYP2D6, GSTM1, GSTT1.

Metodi per rilevare delezioni/duplicazioni segmentali polimorfiche o mutanti. TaqMan, MLPA, analisi delle delezioni nella DMD nei maschi.

Comparative Genomic Hybridization (Classical CGH). BAC arrayCGH, tiling BAC arrayCGH. SNP array per CGH.

Le inversioni segmentali. Un polimorfismo di inversione in 17q21.31

Le principali forze che modellano le frequenze alleliche nelle popolazioni.

Digital Droplet PCR e biopsie liquide. Applicazioni per ricerca di rare copie di DNA mutante entro miscela di DNA wild-type.

Organizzazione del genoma e altre sequenze ripetute: Il DNA satellite, l'eterocromatina centromerica, bandeggio cromosomico (G e Q).

Il DNA minisatellite telomerico, "protein coding-genes" multicopia, geni duplicati, geni multicopia degli RNA (tRNA, rRNA, snoRNA, snRNA, miRNA), retro-trasposoni (LINE-1, LINES).

Meccanismo di retrotrasposizione delle LINES. Le SINES (la sequenza Alu),

Pseudogeni processati, virus-like elements (retro-trasposoni LTR), trasposoni a DNA e meccanismo di trasposizione.

Dal fenotipo al gene. I tratti mendeliani.

Step 1a: identificazione della regione candidata tramite osservazioni citogenetiche (Esempio: Distrofia muscolare di Duchenne e clonaggio per sottrazione di Kunkel). Casi di patologie "de novo" ed analisi di delezioni interstiziali con metodiche ad alta risoluzione (SNP array).

Esempi di patologie legate a delezioni interstiziali di un singolo gene o di pochi geni.

Step 1b: Identificazione della regione candidata tramite analisi di linkage.

Primo esempio di Analisi di linkage: la Corea di Huntington.

Misurare il "contenuto di informazione di un marcatore genetico" (PIC) tramite calcolo della eterozigosità'.

Esempi di alberi genealogici dove capire la "fase gametica" e la prole "ricombinante" e "parentale".

Calcolo del LOD SCORE (mappatura a due punti).

Da un singolo marcatore a più marcatori. Il multipoint LOD score. La prima mappa con densità' elevate dei marcatori umani (CHLC, CEPH). Il multipoint LOD score.

La mappatura per i tratti recessive sfruttando l'autozigosità'. Segmenti cromosomici con "Identity by Descent" (IBD) rispetto alle omozigosità' di segmenti cromosomici con "Identity by State" (IBS). Esempio di calcolo di LOD score in figli di secondi cugini. Relazione tra numero di generazioni e segmenti IBD. Relazione tra densità' dei marcatori e capacità' di rilevare i segmenti IBD. Esempio della mappatura del gene della Fibrosi cistica sfruttando gli aplotipi ancestrali. Aplotipi ancestrali e Linkage disequilibrium nella mappatura della Fibrosi Cistica.

Riassunto della lezione precedente (multipoint LOD score). Regione candidata identificata tramite mappatura per autozigosità'. IBD e IBS.

Possibilità' di restringere la regione candidata sfruttando l'ancestore comune (esempio della fibrosi cistica e della Nijmegen Breakage Syndrome) in popolazioni chiuse. Esempi dalla letteratura di mappature per autozigosità' (vedere diapositive).



## UNIVERSITÀ DI PISA

Esempio di fine mapping in alberi genealogici (dominanza e recessività).

Identificazione della regione candidata tramite l'uso di animali modello. ENU-mutagenized mice. Cenni sul clonaggio posizionale, ordinamento dei contigui di cloni e predizione dei loci genici.

L'era post-genomica: uso delle banche dati per l'identificazione dei geni nella regione candidata. Quali dei geni della regione candidate sono da sottoporre a screening di mutazione? Prioritizzazione sulla base del pattern di espressione, della funzione, delle informazioni provenienti da ortologi e paraloghi. Esempi: Retinite pigmentosa, Sindrome di Marfan, Sindrome di Beals, Malattia di Wilson e Malattia di Menkes. Prioritizzazione possibile anche sulla base di predizioni strutturali della proteina (domini funzionali).

Prioritizzazione sulla base di informazioni parziali dei domini proteici.

Sequenziamento degli esoni dei geni prioritizzati, ma non solo: il cDNA. Esempi della fibrosi cistica e dell'emofilia grave (fattore VIII). Verifiche all'ipotesi che le mutazioni riscontrate in un probando siano effettivamente causative della patologia: la corrispondenza del modello genetico con la segregazione della mutazione; mutazioni o polimorfismi?

Sequenziamento di probandi di famiglie diverse: cosa mi aspetto? Tre possibili situazioni (stessa mutazione, stesso gene diversa mutazione, nessuna mutazione). Inferire l'effetto funzionale della mutazione (predizione in silico).

Esempi di patologie il cui gene è stato rinvenuto tramite clonaggio posizionale e esempi di patologie dove il gene candidato non era immediatamente prevedibile dalla funzione.

The HGP (Human Genome Project).

Metodo clone-by-clone, metodo "shotgun".

Il sequenziamento e il ri-sequenziamento dei genomi oggi: le metodiche di "Next generation sequencing" (NGS). Sistema "454" (Roche).

Questa lezione non verrà tenuta

L'importanza del "Depth" e del "Coverage".

Illustrazione grafica della mappatura delle "reads" in un genome browser. Coverage irregolare del genoma. Relazione tra "depth" e frazione di genoma non sequenziato. Relazione tra numero di reads e frazione del genoma non sequenziato. Relazione tra Depth e Coverage. Relazione tra Depth e numero di "variazioni genetiche" rilevabili in un genoma.

Metodo NGS Solexa-Illumina. Ion-Torrent. Metodi NGS su singola molecola: Pacific-Biosciences e Helicos Biosciences. Cenni sulle applicazioni dell'NGS: whole genome sequencing (WGS), Exome sequencing, RNA sequencing (RNA-seq). Possibile utilizzo "quantitativo" dell'RNA-seq. NGS: qualità delle reads. RNA-seq per l'individuazione delle isoforme degli mRNA. NGS per rilevare le CNV. Il progetto 1000Genomes. Il progetto NHLBI exome sequencing project. Esempio di mappatura per autozigosità mediante NGS. Exome sequencing e Whole genome sequencing. Quali informazioni possiamo capire? Il progetto ENCODE.

Gli studi di associazione caso-controllo per lo studio dei tratti multifattoriali. Studi guidati da una ipotesi: scelta del gene e scelta del polimorfismo. Come scegliere i casi, come scegliere i controlli. Esempi di "bias" ("recall bias", "selection bias", "stratification bias").

Errori con l'uso della prevalenza invece che dell'incidenza. Stratificazioni della popolazione.

Errori nel campionamento dei controlli.

Calcolo del rischio tramite stima dell'OR (Odd Ratio) e dei suoi intervalli di confidenza.

Possibili modelli di "trend" del rischio: co-dominanza, recessività, dominanza.

Limiti degli studi di associazione.

La meta-analisi come strumento statistico.

Saper interpretare uno studio caso-controllo: associazione non vuol dire causa; l'importanza di sapere il significato dell'errore di tipo I (alpha).

ApoE4 come esempio di variante allelica associata a patologia umana.

Questa lezione non verrà tenuta

Scelta degli SNPs da genotipizzare in uno studio di associazione guidato da ipotesi (entro un gene candidato). Utilizzo del linkage disequilibrium. Calcolo della forza di LD tra due polimorfismi tramite i parametri D e r<sup>2</sup>. Le forze che plasmano la variabilità aplo-tipica e modulano l'intensità del linkage disequilibrium. Metodi molecolari e non per definire un aplotipo. Il progetto HAPMAP ([www.hapmap.org](http://www.hapmap.org)).

Come estrarre gli haplotype-tagging SNPs.

Dai geni candidate (hypothesis driven studies) ai genome-wide associations studies (GWAS, hypothesis generating studies).

I Manhattan plot. Esempio di GWAS ad identificare locus di suscettibilità al cancro, nella regione 8q24.

### Bibliografia e materiale didattico

#### 1) Libri di testo consigliati:

"Genetica molecolare umana", by Tom Strachan & Andrew P. Read (Zanichelli)

"Introduzione alla Genomica", by Greg Gibson & Spencer Muse (Zanichelli)

### Indicazioni per non frequentanti

Tutte le informazioni possono essere trovate presso il sito

[www.stefanolandi.eu](http://www.stefanolandi.eu)

### Modalità d'esame

#### 6) L'esame, prova scritta:

È prevista una prova scritta della durata di due ore.

#### L'esame, prova orale opzionale:

È data facoltà di sostenere anche una prova orale facoltativa (permette di modificare la votazione dello scritto di - / + 1 punto)



## UNIVERSITÀ DI PISA

---

Questa puo' essere sostenuta in qualsiasi momento dopo la correzione della prova scritta (anche fuori dal periodo delle date degli appelli). Per accedere alla prova orale basta prendere un appuntamento inviando una e-mail all'indirizzo: [stefano.landi@unipi.it](mailto:stefano.landi@unipi.it)

Stage e tirocini  
Non previsti

Pagina web del corso  
<http://www.stefanolandi.eu>

Altri riferimenti web  
cercare su e-learning e Moodle

*Ultimo aggiornamento 30/03/2017 17:26*