



UNIVERSITÀ DI PAVIA

Anno Accademico 2021/2022

PRINCIPI DI BIOINFORMATICA E METODOLOGIE OMICHE

Anno immatricolazione	2020/2021
Anno offerta	2021/2022
Normativa	DM270
SSD	ING-INF/06 (BIOINGEGNERIA ELETTRONICA E INFORMATICA)
Dipartimento	DIPARTIMENTO DI BIOLOGIA E BIOTECNOLOGIE "LAZZARO SPALLANZANI"
Corso di studio	BIOTECNOLOGIE
Curriculum	PERCORSO COMUNE
Anno di corso	2°
Periodo didattico	Secondo Semestre (01/03/2022 - 14/06/2022)
Crediti	6
Ore	48 ore di attività frontale
Lingua insegnamento	Italiano
Tipo esame	SCRITTO E ORALE CONGIUNTI
Docente	PEVERALI ANTONIO FIORENZO (titolare) - 6 CFU
Prerequisiti	Sono consigliate conoscenze di base di biologia della cellula, genetica, biologia molecolare e biochimica.
Obiettivi formativi	L'integrazione di strumenti bioinformatici con discipline del settore biologico, biomedico, biotecnologico e farmacologico permette di sviluppare una visione integrata e interdisciplinare delle scienze della vita. Obiettivi dell'insegnamento sono : 1 - apprendere un approccio metodologico per raggiungere una sufficiente autonomia nella disciplina; 2 – ottenere un rapido e ampio apprendimento di “tools” bioinformatici dedicati ai principali "database" per migliorare la conoscenza in vari ambiti biologici; 3 - acquisire una moderna e interdisciplinare formazione biotecnologica;

- 4 - sviluppare strumenti di studio e professionali mirati a integrare dati per:
 - chiarire basi molecolari della diversità genetica e di eventuali patologie associate;
 - comprendere le interazioni della materia vivente con piccole molecole ad attività farmacologica/terapeutica o nociva per la salute e/o l'ambiente;
 - progettare molecole ricombinanti di DNA, RNA o proteine;
 - manipolare genomi in vitro e in vivo;
- 5 - acquisire conoscenze di base sui principali metodi "omici" e il loro ruolo nella moderna visione delle biotecnologie e della medicina personalizzata.

Programma e contenuti

Il programma dell'insegnamento è sostanzialmente suddiviso in due principali contenitori (eventuali variazioni al presente saranno comunicate a lezione):

Il PRIMO verte sulla descrizione e utilizzo di principali tools bioinformatici disponibili in rete. Una particolare attenzione è posta alla comprensione e utilizzo di genome browser quali ENSEMBL, NCBI e UCSC e altri strumenti e database dedicati alle proteine, alle interazioni tra molecole, pathway, allineamento di sequenze.

Il SECONDO copre argomenti di carattere OMICO. Saranno pertanto affrontate tematiche su: principi e metodi di sequenziamento massivo parallelo (Next Generation sequencing), analisi dell'esoma (Exome), analisi del trascrittoma (RNA sequencing, microarray), analisi dell'interazione proteine-cromatina (ChIP sequencing), ruolo delle omiche nella produzione di farmaci biotecnologici e riposizionamento dei farmaci (Drug discovery and repositioning), omiche nelle terapie oncologiche (immunoterapia, analisi di espressione differenziale di marcatori genetici, oncoarray).

DETTAGLIO SUI CONTENUTI:

1. INTRODUZIONE alla BIOINFORMATICA:

La bioinformatica: strumento per la divulgazione libera della conoscenza per il progresso dell'umanità.

Dai mattoni della vita ad una visione integrata delle scienze della vita. Accordi internazionali, consorzi, regole e convenzioni.

L'esempio del Progetto genoma: <https://www.genome.gov/>

Dalla lettura della "Natura" alla deduzione della "Regola": il Sistema Internazionale IUPAC per gli acidi nucleici e proteine.

<http://www.chem.qmul.ac.uk/iupac/>

<http://www.chem.qmul.ac.uk/iubmb/misc/naseq.html#100>

2. STRUMENTI E DATABASE DEDICATI ALLA LETTERATURA SCIENTIFICA E BREVETTI;

Consultazione di database mediante l'impiego di "filtri", di strumenti di "Ricerca avanzata" e di personalizzazione (MyNCBI) nel portale di NCBI:

<https://www.ncbi.nlm.nih.gov/>

nel portale di Europe PMC:

<https://europepmc.org/>

nel portale dedicato ai brevetti: <http://www.epo.org/>

Cenni a:

- Peer reviewing; Open access journals; parametri bibliometrici.
- Identificativi personali unici digitalizzati.
- Strumenti per la gestione informatica delle referenze bibliografiche e per l'allestimento di documenti scientifici.
- plagio e strumenti informatici di analisi.

3. BANCHE DATI: DNA; GENI; e NOMENCLATURA

Consorzio internazionale dei nucleotidi (<http://www.insdc.org/>),
Annotazione genica: formati GenBank e Fasta
Banche dati (<https://www.ncbi.nlm.nih.gov/>): Nucleotide Database, Reference Sequence (RefSeq), Consensus CDS (CCDS), GenBank, Database of Expressed Sequence Tags (dbEST), Gene, GeneCards <http://www.genecards.org/>
Nomenclatura ufficiale dei geni: <http://www.genenames.org/>
Manipolazione di DNA ricombinante e progettazione di proteine ricombinanti:
http://www.snapgene.com/products/snapgene_viewer/

5. GENOME BROWSER: ENSEMBL, NCBI, UCSC.

Progetto genoma. Sequenziamento e approcci: Top-down vs bottom-up;
Mappe fisiche, citogenetiche e genetiche. Coordinate genomiche vs genetiche. Concetto di sintenia. Marcatori molecolari: STS.
Genome Assembly e Processo di aggiornamento di un genoma (Release Number).
https://en.wikipedia.org/wiki/Human_Genome_Project
ENSEMBL browser: <http://www.ensembl.org/index.html>
NCBI genome browser: <http://www.ncbi.nlm.nih.gov/mapview/>
UCSC genome browser: <https://genome.ucsc.edu/index.html>

6. ESPRESSIONE GENICA

Gene expression in Ensembl
<http://www.ensembl.org/info/website/tutorials/expression.html>
Array express <https://www.ebi.ac.uk/arrayexpress/>
Expression Atlas <https://www.ebi.ac.uk/gxa/home>
Unigene, EST profile; GEO profiles
<http://www.ncbi.nlm.nih.gov/unigene>

7. STRUMENTI E DATABASE PER L'ANALISI DELLE VARIANTI POLIMORFICHE E PATOLOGICHE. Variabilit  genetica normale e patologica mediante analisi di SNP, Ins/Del e/o riarrangiamenti citogenetici. Definizione genetica e molecolare di SNP, definizione di MAF.

Ensembl SNPs and other variants for my gene:
http://www.ensembl.org/info/website/tutorials/gene_snps.html
Database of Short Genetic Variations (dbSNP)
<http://www.ncbi.nlm.nih.gov/snp/>
Malattie genetiche a trasmissione mendeliana. Malattie rare a base genetica. Relazione genotipo fenotipo.
<http://omim.org/>
<http://www.orpha.net/consor/cgi-bin/index.php>
<http://www.ncbi.nlm.nih.gov/medgen/>
<http://www.ncbi.nlm.nih.gov/variation/>

<http://www.ncbi.nlm.nih.gov/books/NBK21088/>
<http://www.ncbi.nlm.nih.gov/dbvar>
<http://www.ncbi.nlm.nih.gov/clinvar/>
cenni al progetto 1000 Genomes
<http://www.1000genomes.org/>
genome aggregation database (Gnomad)
<https://gnomad.broadinstitute.org/>

8. PROTEINE, STRUTTURE ed ENZIMI

Proteine, enzimi, strutture 3-D, interazioni proteiche

<http://www.uniprot.org/>
<http://www.ebi.ac.uk/interpro/>
<http://www.ebi.ac.uk/enzymeportal/>
<http://www.brenda-enzymes.info/>
<http://www.rcsb.org/pdb/home/home.do>
<https://www.ebi.ac.uk/pdbe/pdbe-kb>
cenni a interazione DNA-proteine ed epigenoma. Cenni al progetto ENCODE. Ricerca nel genoma di: promotori, CpG island, modificazioni epigenetiche di istoni, siti di ipersensibilità alla DNasi, siti di legame DNA-proteine, insulators
mediante l' UCSC e l' ENSEMBL genome browser.

9. ALLINEAMENTO DI SEQUENZE.

Descrizione di algoritmi BLAST per allineamenti tra sequenze nucleotidiche; tra sequenze proteiche; tra sequenze nucleotidiche e proteiche e viceversa. Allineamenti multipli di sequenze; Significato dei parametri di allineamento: "word size", "expect threshold", "Match/Mismatch scores", "Gap cost".
<http://blast.ncbi.nlm.nih.gov/Blast.cgi>
<http://www.ebi.ac.uk/Tools/msa/tcoffee/>

10. INTERAZIONI PROTEICHE, METABOLISMI e PATHWAYS

Definizione ed analisi di pathways e metabolismi.
Ricerca di interattori tra proteine e proteine-piccole molecole.
<http://www.genome.jp/kegg/pathway.html>
<http://www.reactome.org/>
http://string-db.org/newstring.cgi/show_input_page.pl?UserId=bfrMLUg_TgzQ&sessionId=JRVOp7Hh7JFy&input_page_type=single_identifier
<http://stitch.embl.de/>
<https://www.ebi.ac.uk/intact/>
<https://www.ebi.ac.uk/complexportal/home>

10. PICCOLE MOLECOLE, FARMACI, TOSSICOLOGIA - SMALL MOLECULES

Siti dedicati a piccole molecole con attività biologica, composti chimici ed effetti sul vivente:
<http://toxnet.nlm.nih.gov/newtoxnet/toxnetallsearch.html>
<https://www.ebi.ac.uk/chembl/index.php/compound/results>
<https://www.ebi.ac.uk/chebi/>
<http://www.drugbank.ca/>
<https://pubchem.ncbi.nlm.nih.gov/>
CAS number. Material Safety Data Sheet. Descrizione di alcuni Pittogrammi sulla sicurezza

<http://www.sigmaaldrich.com/help-welcome/hazard-and-precautionary-statements.html#pictogram>

11. RISVOLTI BIOTECNOLOGICI DELL'IMPIEGO DI TOOLS BIOINFORMATICI:

Esempi di simulazione in silico per la costruzione di acidi nucleici ricombinanti, di proteine ricombinanti e di manipolazioni del genoma in vitro e in vivo "genome editing". Esempi di modelli animali transgenici.

12. Argomenti "omici" trattati:

- Next Generation Sequencing
- Whole Exome sequencing
- RNA sequencing
- ChIP sequencing
- Microarray e oncoarray
- Drug discovery e drug repositioning
- Immunoterapia e omica

Metodi didattici

Lezioni frontali e seminari per un totale di 48 ore si svolgeranno in aule dotate di collegamento wi-fi per consentire con dispositivi personali la consultazione contestuale dello strumento bioinformatico presentato. I tutorati per circa 20 ore si svolgeranno in aule dotate di personal computer in cui gli studenti iscritti all'insegnamento si eserciteranno sugli strumenti bioinformatici presentati a lezione.

Saranno inoltre allestiti quiz test di autovalutazione dell'apprendimento sulla piattaforma di e-learning di UniPV – KIRO:

<https://elearning2.unipv.it/bio/login/index.php>

I quiz test saranno esercitati durante i tutorati e/o lezione e successivamente resi disponibili anche in remoto agli iscritti all'insegnamento.

Testi di riferimento

In generale, alle voci "HELP", "GETTING HELP", "EDUCATION", "TRAINING AND TUTORIALS" e demo su You Tube sono disponibili risorse, spiegazioni e chiarimenti per il rapido apprendimento dei siti bioinformatici presi in esame. Di seguito sono riportate alcune voci come esempio:

- EBI training online;
- Ensembl help & Documentations;
- Uniprot documentations and tutorials;
- UCSC genome bioinformatics help;
- NCBI Training and Tutorials;
- NCBI Handbook,
- NCBI shelves;

Articoli specifici discussi durante le lezioni sono resi disponibili agli iscritti all'insegnamento sulla piattaforma di e-learning di unipv (kIRO).

Altri siti ove reperire materiale sono:

- Nucleic Acids Research Database Issue and Molecular Biology Database Collection

<https://www.oxfordjournals.org/nar/database/c/>

dettagliata collezione di articoli che descrivono "Database" suddivisi per argomenti

- Database resources of the National Center for Biotechnology Information:
<https://www.ncbi.nlm.nih.gov/home/learn.shtml>
<https://www.ncbi.nlm.nih.gov/books/NBK3831/>
<https://www.ncbi.nlm.nih.gov/books/NBK143764/>
The NCBI Handbook, 2nd edition. Il PDF è scaricabile da questo sito ed è inoltre disponibile nella piattaforma di KIRO.

Nel corso dell'insegnamento saranno resi disponibili sulla piattaforma didattica KIRO (Moodle 2.7) di UniPV (<http://elearning2.unipv.it/bio/>):

- Programma d'esame aggiornato sugli argomenti trattati;
- Elenco dei siti web consultati;
- Esempi di esercitazioni e quiz test di autovalutazione.
- Materiale didattico, articoli, review e altra documentazione.

Un libro di testo consigliato è:

Fondamenti di Bioinformatica, di Citterich et al.; Zanichelli 2018, ISBN: 9788808621122

Modalità verifica apprendimento

Per un rapido e proficuo apprendimento dell'insegnamento è consigliata la frequenza delle lezioni e del tutorato. La verifica dell'apprendimento del programma sarà costantemente monitorata durante l'insegnamento mediante quiz test di autovalutazione sulla piattaforma KIRO.

Le domande dei quiz saranno in formato:

- Vero/Falso;
- Risposta chiusa multipla;
- Risposta breve;
- Risposta sequenziale;
- Risposta a corrispondenza.

Istruzioni ed informazioni dettagliate su modalità di svolgimento dei quiz di apprendimento, quiz d'esame conclusivo, argomenti trattati durante l'insegnamento, quiz di autovalutazione saranno resi disponibili per gli iscritti all'insegnamento sulla piattaforma KIRO.

Lo Studente sosterrà la prova finale d'esame scritta mediante "quiz " sulla piattaforma didattica "KIRO" (Moodle 2.7) di UniPV (<http://elearning2.unipv.it/bio/>) in aula informatizzata dell'Università. Eventuali disabilità verranno prese in considerazione singolarmente previo contatto con il docente.

Per il superamento dell'esame, è richiesto allo studente l'acquisizione di conoscenze teoriche e dimestichezza con gli strumenti bioinformatici "on-line" presentati ed esercitati sia a lezione sia al tutorato. La durata del quiz d'esame avrà un tempo di circa 2 ore.

Durante l'insegnamento verrà valutata e comunicata l'eventuale aggiunta di una prova orale da effettuarsi solo dopo il superamento dello scritto.

Altre informazioni

Il Docente è disponibile per chiarimenti e ulteriori informazioni, previa richiesta di colloquio esclusivamente all'indirizzo e-mail dell'università.

