

**Insegnamento di Biologia - Prof. Stefania Bortoluzzi I Esercitazione – 27 novembre 2023**

Seguire il percorso indicato nel testo sottostante e stilare una breve relazione con i contenuti recuperati e elaborati man mano, usando diverse risorse informatiche.

**A) Data retrieval da database dell'NCBI**

NCBI (<http://www.ncbi.nlm.nih.gov/>) è un'importante risorsa da cui reperire informazioni biologiche. Fornisce un'interfaccia unificata per accedere a molti database differenti per contenuto e organizzazione, quali database di conoscenza (letteratura, geni, malattie), biosequenze (nucleotidiche, proteiche) e strutture (3D di proteine, domini proteici conservati).

Partire da <http://www.ncbi.nlm.nih.gov/guide/all/>, il portale d'entrata per la ricerca in diversi database, divisi per tipo di contenuto.

**1. Ricerca bibliografica in NCBI Bookshelf (Literature):**

- Considerare “Book List”.
- Cercare Cadherins, scegliere un capitolo di libro relativo a questo argomento e salvarsi in formato breve il contenuto raggiunto (ad es. un estratto del capitolo “Adhesion” di Cancer Medicine).

**2. Ricerca bibliografica in NCBI PubMed (Literature):**

- Ricerca semplice: quanti articoli ci sono che trattano di “macular degeneration”?
- Usare “advanced search” con operatori Booleani e parole chiave cercate in specifici campi:
  1. Quanti articoli ci sono con “macular degeneration” nel titolo?
  2. Quanti pubblicati sulla rivista “Ophthalmology”?
- Usare i filtri: quanti articoli ci sono con “macular degeneration” nel titolo, pubblicati sulla rivista “Ophthalmology negli ultimi tre anni.
- Trovare l'articolo più recente sulla rivista Science riguardante il cristallino (parole chiave nel titolo “eye” “lens” e “proteins”). Quale dato porta?

**3. Retrieval di sequenze con NCBI Gene database ():**

- Cercare il gene del citocromo b (cytochrome; cytb) nel genoma mitocondriale umano (*Homo sapiens*), consultare la pagina del database di geni. In fondo nel campo “NCBI Reference Sequences (RefSeq)” dove sono fornite le sequenze di riferimento, scaricare la sequenza del gene nei formati GenBank e FASTA. Quanto è lunga?

**4. Retrieval di sequenze con NCBI Protein:**

- Cercare la sequenza della PROTEINA umana alpha crystallin-A (usare il filtro “solo Refseq database” per trovare quella di riferimento), annotarsi la lunghezza e copiare la sequenza in formato FASTA in un file Word. Quali sono le funzioni della cristallina? Fare una breve

ricerca, o vedere

[https://www.pianetachimica.it/mol\\_mese/mol\\_mese\\_2010/07\\_Cristalline/Cristalline.htm](https://www.pianetachimica.it/mol_mese/mol_mese_2010/07_Cristalline/Cristalline.htm)

## 5. NCBI Genome:

Consultare le pagine raggiungibili da NCBI Genome (<http://www.ncbi.nlm.nih.gov/genome>), quali “Human genome”, “Viruses”, ecc.

Provare a raggiungere le informazioni relative ad uno specifico genoma completo:

*Cutibacterium acnes*, ceppo (strain). Quanti geni protein coding contiene questo genoma batterico?

### **B) Dalla malattia genetica al gene, attraverso il genoma.**

Prenderemo in considerazione una malattia genetica e cercheremo di utilizzare i dati contenuti in diversi database (di malattie, geni, proteine e genomi) per reperire informazioni sul fenotipo, sui geni, sulle mutazioni che lo causano, ecc.

1) Collegarsi a OMIM e cercare “RETINITIS PIGMENTOSA”

- Analizzare i risultati e commentarli (entries di geni; entries di malattie; ci sono più malattie diverse?).

2) Scegliere “RETINITIS PIGMENTOSA 9”

- Analizzare le informazioni sulla patologia (pagina principale, clinical synopsis etc.) per capire di che malattia si tratta, come si eredita e in quale cromosoma e regione genomica si trova il locus associato.

3) Seguire il link al locus e quindi alla “location” per arrivare alla visualizzazione grafica del contenuto di informazione della regione del geno a umano che contiene il gene. Con l’aiuto dell’esercitatore cercare di orientarsi nell’UCSC Genome Browser.

Cliccare sulla track del gene (colore blu).

- Di che gene si tratta? Che funzione ha? Che altre informazioni sul gene possono essere reperite in questo database?
- Di quanti esoni è composto il gene?
- Quali sono i due geni “vicini” sul DNA genomico?
- A partire dal campo “Sequence and Links to Tools and Databases” scaricare la sequenza della proteina codificata.