**Programma di Bioinformatica: Fondamenti e Applicazioni nella Gestione dei Dati Biologici**

Il corso è suddiviso in moduli teorici e pratici. Le sessioni pratiche (Lab) sono integrate per applicare immediatamente i concetti teorici, con un focus specifico sull'utilizzo della riga di comando in un ambiente Linux.

**Modulo 1: Fondamenti di Bioinformatica e Biologia Molecolare**

Questo modulo fornisce le basi teoriche indispensabili per affrontare il resto del corso, introducendo la disciplina e ripassando i concetti biologici chiave.

* **Lezione 1.1: Cos'è la Bioinformatica?**
  + Definizione, storia e importanza della bioinformatica.
  + Casi di studio: scoperta di farmaci, medicina personalizzata.
* **Lezione 1.2: Basi di Biologia Molecolare**
  + Revisione di DNA, RNA e proteine.
  + Il dogma centrale della biologia molecolare.
  + Concetti di gene, genoma, trascrittoma, proteoma.
  + Introduzione alle tecnologie "omiche" (genomica, trascrittomica, proteomica).

**Modulo 2: Banche Dati Biologiche e Formati di File**

In questo modulo gli studenti imparano a navigare e a utilizzare le risorse online fondamentali e a riconoscere i formati di file più comuni, prerequisito per l'analisi successiva.

* **Lezione 2.1: Banche Dati di Acidi Nucleici e Proteiche**
  + Panoramica di **GenBank**, **ENA**, **UniProtKB** e database specializzati (es. Ensembl).
  + **Lab 2.1:** Esercitazioni guidate sui portali web di GenBank e UniProt per la ricerca e l'estrazione di sequenze.
* **Lezione 2.2: Banche Dati Specializzate e Bibliografiche**
  + Introduzione a **KEGG**, **STRING**, **GEO** e ai sistemi di integrazione dati (**Entrez**).
  + Ricerca avanzata su **PubMed** e gestione dei riferimenti bibliografici.
* **Lezione 2.3: Formati di File Comuni in Bioinformatica**
  + Struttura e utilizzo dei formati **FASTA**, **FASTQ**, **GenBank Flat File**, **GFF/GTF** e **PDB**.

**Modulo 3: La Riga di Comando Linux per la Gestione dei Dati**

Questo modulo è cruciale per la parte pratica. Qui gli studenti acquisiscono le competenze fondamentali per operare in un ambiente Linux, il sistema operativo standard in bioinformatica.

* **Lezione 3.1: Fondamenti della Riga di Comando Linux**
  + Navigazione nel filesystem: pwd, ls, cd.
  + Creazione e gestione di file e directory: mkdir, touch, cp, mv, rm.
  + Visualizzazione e manipolazione di file: cat, less, head, tail.
* **Lab 3.1: La tua prima sessione in Linux**
  + Esercitazioni pratiche per muoversi nel filesystem, creare e manipolare file e directory.
* **Lezione 3.2: Lavorare con i Dati di Sequenza dalla Riga di Comando**
  + Introduzione ai comandi di base per l'elaborazione dei file: grep, wc, sort, uniq.
  + **Redirezione dell'output** (>) e **piping** (|).
* **Lab 3.2: Esercitazione sui File di Dati**
  + Scaricare un file FASTA di esempio.
  + Usare grep e wc per contare le sequenze.
  + Combinare comandi con il pipe (|) per filtrare e analizzare i dati.
  + **Gestione dei file con editor di testo** e **manipolazione di stringhe** con espressioni regolari.

**Modulo 4: Gestione dei Dati con i Database Relazionali e Strumenti Online**

Questo modulo collega i concetti di base con l'uso di strumenti più avanzati, introducendo i database relazionali come mezzo per una gestione più strutturata e scalabile dei dati.

* **Lezione 4.1: Introduzione ai Sistemi di Gestione di Database (DBMS)**
  + Vantaggi dei **DBMS** rispetto ai file di testo: efficienza, coerenza, scalabilità.
  + Concetti di base: **tabelle**, **campi**, **record** e **chiavi**.
* **Lezione 4.2: Interrogazione dei Dati con SQL (Structured Query Language)**
  + Introduzione a **SQL** (SELECT, FROM, WHERE).
  + Esempi pratici: SELECT \* FROM Genes WHERE species = 'Homo sapiens';
* **Lab 4.1: Lab sui Database**
  + Installazione e configurazione di **SQLite**.
  + Creare tabelle e importare dati da un file FASTA.
  + Eseguire query per analizzare i dati importati.
* **Lezione 4.3: Strumenti Bioinformatici Online**
  + **BLAST**: Principi e utilizzo del Basic Local Alignment Search Tool.
  + **Clustal Omega**: Allineamento multiplo di sequenze.
* **Lab 4.2: Esercitazione con BLAST e Clustal**
  + Eseguire un'allineamento di sequenza con BLAST e un allineamento multiplo con Clustal.
* **Lezione 4.4: Strategie di Download e Gestione dei Dati su Linux**
  + Uso di wget e curl per scaricare dati.
  + Sistemi di compressione (gzip, tar).
  + Esempio pratico: scaricare e decomprimere un dataset.

**Modulo 5: Analisi Bioinformatiche e Programmazione**

Questo modulo introduce argomenti più avanzati che richiedono una base solida. Qui gli studenti iniziano a comprendere gli algoritmi di base e a scrivere i loro primi script per automatizzare le analisi.

* **Lezione 5.1: Annotazione, Curatela e Programmazione**
  + L'importanza dell'annotazione accurata e del controllo di qualità dei dati.
  + Introduzione al **parsing** dei file con semplici script.
  + Concetti base di programmazione: variabili, tipi di dati, cicli, condizionali.
* **Lab 5.1: Il tuo primo script**
  + Scrivere uno script **Python** o **Perl** per leggere un file FASTA e stampare le intestazioni di sequenza.
* **Lezione 5.2: Allineamenti di Sequenza e Filogenesi**
  + Allineamento a coppie e allineamenti multipli.
  + Concetti di albero filogenetico e metodi di ricostruzione.
* **Lezione 5.3: Nuovi Orizzonti in Bioinformatica**
  + Bioinformatica del sequenziamento di nuova generazione (**NGS**).
  + **Machine Learning** e **Intelligenza Artificiale** in bioinformatica.