

# QUESITI PROVA PRATICA "A VISTA" - SELEZIONE N. 2024N49

## PROVA N. 1

Si richiede di sviluppare una soluzione per il controllo di qualità delle reads generate da un sequenziamento genomico, adatto ad essere eseguito su un sistema di calcolo condiviso. Tale soluzione richiede:

### **Installazione del software**

Completare l'installazione del software necessario all'analisi. Questo può includere il download, manuale o automatizzato tramite programmi terzi, di pacchetti; l'installazione delle dipendenze; la configurazione del software stesso e di eventuali variabili di ambiente.

### **Esecuzione dell'analisi**

Eeguire l'analisi delle reads che verranno rese disponibili in un file opportuno. Il programma utilizzato per l'analisi dovrà fornire elementi che indichino la presenza di anomalie, quali una bassa qualità o la presenza di contaminanti (tecnici oppure biologici). Dovrà inoltre produrre un report che riassume i risultati ottenuti e che possa essere stoccato indipendentemente dal dato grezzo analizzato.

## PROVA N. 2

Realizzi uno script per l'analisi di varianti genomiche, in un linguaggio di programmazione a sua scelta. Il programma dovrà svolgere le seguenti operazioni:

### **Filtraggio delle varianti**

Leggere una tabella di varianti genomiche da un file di input fornito.

Filtrare le varianti in base ai seguenti criteri:

- Cromosoma: 5
- Posizione: compresa tra 10200 e 10300 (estremi inclusi)

Scrivere le varianti filtrate in un nuovo file di output, mantenendo tutte le informazioni originali per ciascuna variante selezionata.

### **Esecuzione su un cluster**

Creare un job file per sottomettere lo script precedentemente realizzato ad un sistema di gestione dei job (ad esempio, SLURM o PBS), specificando opportunamente le risorse richieste.

Il job file deve essere correttamente formattato e includere tutti i comandi necessari per l'esecuzione dello script in un ambiente di calcolo distribuito.

## PROVA N. 3

Realizzi un container software per un'analisi di allineamento di sequenza, seguendo le specifiche sotto riportate.

### Creazione del container

Utilizzando un sistema di containerizzazione quale Docker oppure Singularity/Apptainer, crei un container che includa:

- Un sistema operativo base (ad esempio, Ubuntu o RHEL).
- Un software di allineamento di sequenze a sua scelta (ad esempio, BWA, Bowtie2, HISAT2).
- Tutte le dipendenze necessarie per il funzionamento del software scelto.

### Dimostrazione dell'esecuzione

Utilizzando il container creato, dimostri come eseguire il software di allineamento per allineare una serie di reads, fornite in un file opportuno, con il cromosoma 21 del genoma umano.

Fornisca i comandi necessari per:

- Avviare il container.
- Eseguire l'allineamento all'interno del container
- Salvare i risultati in un percorso al di fuori del container per la conservazione dopo il completamento dell'analisi.