



FSC
Fondo per lo Sviluppo
e la Coesione



Progetto LIFEMap

Seminario scientifico

30 maggio 2025

12:00 - 13:00

Analisi genomica scalabile: tutorial GWAS con Hail

Relatori: Francesco Versaci e Giovanni Busonera - CRS4

Abstract

In questa presentazione sarà illustrato un workflow completo per l'esecuzione di uno studio di associazione genome-wide (GWAS) utilizzando Hail, una libreria scalabile ed efficiente per l'analisi genetica. Il tutorial mostrerà le principali strutture dati e funzionalità fornite dalla libreria, per poter manipolare i dati genomici al fine di applicare le principali operazioni necessarie in uno studio GWAS.

La presentazione inizierà con una breve panoramica sul calcolo distribuito e sull'architettura e configurazione di Hail (cluster Spark + object storage S3), per poi proseguire con una dimostrazione pratica del suo utilizzo. In dettaglio vedremo:

- il caricamento di dati genomici e metadati dei sample;
- il controllo di qualità;
- l'esecuzione dell'associazione con la gestione delle covariate.

L'obiettivo è fornire ai partecipanti una visione pratica del processo, con esempi di codice.

Link - <https://meet.google.com/vaa-oczm-yjy>