Association Mapping: GWAS and Sequencing Data

Instructors: Joelle Mbatchou and Loic Yengo

Summer Institute in Statistical Genetics (SISG)

July 2023

Introduction: Course Goals

This is a course on statistical methods and software for genetic association studies of complex traits. We aim to cover:

- Genetic Association Testing with Case-Control & Quantitative Traits
- Population Structure/Ancestry Inference
- Genetic Association Testing in Samples with Structure
- Conditional analyses, Colocalization, Fine-mapping & Gene and Pathway Level Analysis
- Polygenic Risk Scores
- Association Testing for Rare Variant Analysis
- Interaction Analysis, GWAX, Time-to-event & Multi-trait Analysis
- Power and Sample Size, Design Considerations and Emerging Issues

Introduction: Resources

Importantly, the class site is:
https://joellembatchou.github.io/SISG2023_Association_Mapping/

Contains (or will contain):

- Link to PDF copies of slides
- Practical exercises for you to try
- Link to datasets used in exercises
- Our solutions to exercises (later!)
- Links to software packages

Introduction: About Loic

- Statistical Geneticist,
  the University of Queensland, Australia
- Research interests:
  - GWAS of anthropometric traits
  - GWAS transferability
  - Non-random mating in humans

Introduction: About Joelle

- Statistical Geneticist,
  Regeneron Genetics Center
- Research interests:
  - Genetic Association Studies
  - Genetic Data with Structure
  - Mixed Models methods
  - Association methods for large-scale datasets

Slack Channel

Slack channel for class:
https://uwbiostatisticssisg.slack.com/archives/C05EB2ZDDSN

Will contain:

- Key annoucements
- Link to the class website
- PDF copies of slides

Cloud server

We will use a cloud server to do the practical exercises. For more info on getting set up on the server:
https://joellembatchou.github.io/SISG2023_Association_Mapping/using_server.html

Things to note:

- **Let us know if you cannot access the server**
- We will run exercises on the server
- Datasets used in practicals are at:
  `/data/SISG2023M15/data/`
- RStudio server can be accessed at:
  `http://si2023-compute.biostat.washington.edu:8787/`

Introduction: Course Structure

- Full schedule on class website (**Seattle time, PDT**)
- 10 sessions, 60-90 minutes each, over 2.5 days
- What to expect in a typical session;
  - 45 mins teaching/lecture
  - 30 mins hands-on exercises
  - 15 mins summary/discussion
- For the practicals:
  - Students will be split into smaller groups
  - You can run them on the server or your own machine