

Association Mapping: GWAS and Sequencing Data

Instructors: Joelle Mbatchou and Loic Yengo

Summer Institute in Statistical Genetics (SISG)

July 2022

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 & Polygenic Risk Scores
- ▶ Gene and Pathway Level Analysis
- ▶ 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/SISG2022_Association_Mapping/

Contains (or will contain):

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

Introduction: About Loic



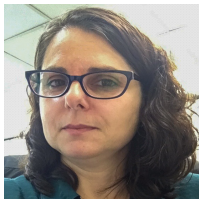
- ▶ A
- ▶ B
- ▶ C

Introduction: About Joelle



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

UW department's finest – here to help you!



Andrea Horimoto



Amarise Little



Anya Mikhaylova



Seth Temple

Slack Channel

Expect to 'see' them on Zoom chat, and our Slack channel:

<https://uwbiostatisticssig.slack.com/archives/C03KLGA4HE3>

Contains (or will contain);

- ▶ Link to class website
- ▶ PDF copies of slides
- ▶ Practical exercises for you to try
- ▶ Our solutions to exercises (later!)

Cloud server

We will use a cloud server to do the practical exercises. To access `ssh -XY <username>@si2022-m15.biostat.washington.edu`

Things to note:

- ▶ **Let us know if you cannot access the server**
- ▶ We will run command line exercises on the server
- ▶ Check the course website for details on where the datasets are for each session

Introduction: Course Structure

- ▶ 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