

Genetic Association Course

With Application to Sequence and Genotype Data

September 13 - 17, 2021

Max Delbrück Center for Molecular Medicine–Berlin, Germany

Each session will begin with a theoretical introduction followed by practical exercises. The instructors for the course are Suzanne Leal (Columbia University) and Michael Nothnagel (University of Cologne and Max Delbrück Center). Please feel free to bring your own data sets for discussion and/or analysis.

The course will be held daily from 9:00 a.m. to 5:00 p.m., except for Wednesday, when the course will end at 1:00 pm to have free time in the afternoon for sightseeing. On Monday, registration will be held from 8:30 to 9:00 am. A wine and cheese get together will be held for students and faculty directly after the course on Monday.

MONDAY September 13 th	Morning	<i>Lecture</i> Introduction to genetic epidemiology, population genetics and statistical testing. Introduction to PLINK and R; file formats <i>Computer Exercises</i> ; PLINK and R – manipulating data
	Afternoon	<i>Lecture</i> Basic statistical test for the analysis of genotype and sequence data Introduction to PLINK and R; file formats <i>Computer Exercises</i> ; PLINK and R – simple tests of association
	17:00-18:30	Wine and Cheese Party
TUESDAY September 14 th	Morning	<i>Lecture</i> Linkage disequilibrium (LD), pairwise measures of LD Data quality control, cleaning genotype data <i>Computer Exercises</i> PLINK – data quality control <i>Pencil and Paper Exercises</i> r ² , D', etc.
	Afternoon	<i>Lecture</i> Analysis of quantitative and qualitative traits using linear and logistic regression; confounding and how to control for it in the analysis Haplotype reconstruction and estimation, testing for associations using haplotype data <i>Computer Exercises</i> UNPHASED – Haplotype and LD estimation & association testing PLINK & R – Logistic and linear regression – adjusting for covariates
WEDNESDAY	Morning	<i>Lecture</i>

September 15 th		<p>Population substructure/admixture detection and control of confounding due to population substructure (structure, principal components analysis, etc.)</p> <p><i>Computer Exercises</i></p> <p>PLINK – Multidimensionality scaling (MDS) and principal components analysis (PCA)</p>
THURSDAY September 16 th	Afternoon	Free for sightseeing
	Morning	<p><i>Lecture</i></p> <p>Generalized linear mixed models and linear mixed models; Data quality control for rare variant data obtained from next generation sequencing;</p> <p><i>Computer Exercises</i></p> <p>FAST-GWA, REGENIE, Variant Association Tools (VAT)-part I</p>
FRIDAY September 17 th	Afternoon	<p><i>Lecture</i></p> <p>Complex trait rare variant association analysis of sequence data. population and family-based aggregate tests.</p> <p><i>Computer Exercises</i></p> <p>VAT-part II and PSEQ</p>
	Morning	<p><i>Lecture</i></p> <p>Sample size estimation and power calculations (for Rare Variant Aggregation Tests) and Genome-Wide Association Studies (GWAS); the multiple testing problem; controlling the family wise error rate (FWER); and permutation and false discovery rate (FDR)</p> <p><i>Computer Exercises</i></p> <p>Cochran-Armitage test for trend power tool, GAS, Genetic Power Calculator R-permutation, FDR</p>
	Afternoon	<p><i>Lecture</i></p> <p>Imputing genotype data from sequence and genotype data; analyzing imputed genotype data; polygenic risk scores; detecting gene x gene and gene x environment interactions</p> <p><i>Computer Exercises</i></p> <p>PLINK & R -Testing for gene x gene interactions</p>