Course

Genetic Association and Beyond: Statistical Methods to Elucidate Complex Trait Etiology

September 14 - 18, 2026
Max Delbrück Center for Molecular Medicine (MDC), Berlin, Germany

The emphasis of this course is analysis of human complex trait data using a variety of statistical and bioinformatic approaches. The course instructors are Suzanne Leal (Columbia University) and Michael Nothnagel (University of Cologne).

Sessions will begin with a theoretical introduction followed by practical exercises. 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 12:30 pm to have free time in the afternoon for sightseeing. Questions you have about the analysis of your own data are welcome. On Monday, registration will be held from 8:30 to 9:00 am. A dinner at a local restaurant will be held for students and faculty directly following the course on Monday. There is no fee to attend the dinner.

MONDAY Sep 14th	Morning	*Lectures* (1) Introduction to genetic epidemiology and population genetics (2) Introduction to PLINK and R; File formats *Computer Exercises* PLINK and R – Manipulating data *Pencil and Paper Exercises* Hardy-Weinberg equilibrium, F_{ST}, etc.
	Afternoon	*Lectures* (1) Data quality control for genotype array and sequence (NGS) data (2) Linkage disequilibrium (LD): pairwise measures, haplotype reconstruction and estimation *Computer Exercises* BCFtools, ANNOVAR *Pencil and Paper Exercises* r2, D', etc.
	17:30 -21:30	Dinner at *Il Castello* – Alt-Buch Karower Str. 1, 13125 Berlin
TUESDAY Sep15th	Morning	*Lectures* Genetic analysis of quantitative and qualitative traits using linear and logistic regression; Confounding and how to control for it in the analysis; Linear mixed models (LMM) and generalized LMM (GLMM)

		Computer Exercises PLINK & R – Logistic and linear regression – adjusting for covariates, FAST-LMM
	Afternoon	*Lectures* (1) Detecting population substructure/admixture and controlling potential confounding (structure-like approaches, principal components analysis, G/LMM, etc.) (2) Heritability and its estimation; Meta-analysis; and Fine mapping *Computer Exercises* PLINK – Principal components analysis (PCA), GCTA, LDSC regression
WEDNESDAY Sep 16th	Morning	*Lectures* Rare-variant association analysis of complex traits using sequence data *Computer Exercises* REGENIE
	Afternoon	*Free for sightseeing*
THURSDAY Sep 17th	Morning	*Lectures* (1) Detecting GxG and GxE interactions (2) Pleiotropy analysis; Mediation analysis *Computer Exercises* R, PLINK, LDClumping, mediation
	Afternoon	*Lectures* 1) Mendelian randomization (MR) 2) Genetic correlations *Computer Exercises* TwoSampleMR LDSC Regression
FRIDAY Sep 18th	Morning	*Lectures* (1) Polygenic scores (PGS) (2) Imputation of genetic variants and analyzing imputed genotype data (3) The multiple-testing problem; Controlling the family-wise error rate (FWER) and the false discovery rate (FDR) *Computer Exercises* R-permutation, FDR
	Afternoon	*Lectures* (1) Sample size estimation and power calculations for rare-variant aggregation tests and genome-wide association studies (GWAS) (2) Imputation of genetic variants and analyzing imputed genotype data (3) Proteomic-wide association analysis (PWAS) and direct-PWAS: predicting protein expression levels using *cis* and *trans* variants and analyzing measured and predicted protein expression levels *Computer Exercises* GAS, Genetic Power Calculator