Human Linkage and Association M21-5483/ L41-5483
Fall, 2017,
Tuesday, Thursday, 9:30-11:00
John Rice, Course master
Lectures: Becker Library, Room 502; Labs 501

This is a survey course on contemporary approaches to the statistical analysis of human genetic data. Topics include assessment of familial resemblance, determination of mode of inheritance, human gene maps, statistical and computational approaches to assessing linkage for gene discovery, and fine mapping approaches using linkage disequilibrium. Approaches to test and estimate the effect of specific genetic variants on disease risk or trait variability also will be covered. The course is a combination of didactic lectures and hands-on computing exercises.

Course Description:
Basic Genetic concepts: meiosis, inheritance, Hardy-Weinberg Equilibrium, Linkage, segregation analysis; Linkage analysis: definition, crossing over, map functions, phase, LOD scores, penetrance, phenocopies, liability classes, multi-point analysis, non-parametric analysis (sibpairs and pedigrees), quantitative trait analysis, determination of power for mendelian and complex trait analysis; Linkage Disequilibrium analyses: allelic association (case control designs and family bases studies), QQ and Manhattan plots, whole genome association analysis; population stratification; Quantitative Trait Analysis; measured genotypes and variance components. Hands-on computer lab experience doing parametric linkage analysis with the program LINKAGE, model free linkage analyses with Genehunter and Merlin, power computations with SLINK, quantitative trait analyses with SOLAR, LD computations with Haploview, and family-based and case-control association analyses with PLINK and SAS. The methods and exercises are coordinated with the lectures, and students are expected to understand underlying assumptions and limitations, and the basic calculations performed by these computer programs. Auditors will not have access to the computer lab sessions.

Learning Objectives:

- Define and understand basic genetic concepts
- Compute LOD scores and perform linkage analysis
- Determine power for Mendelian and complex traits
- Perform GWAS analysis (case control and family based designs)
- Understand sequencing technologies and perform rare variant association

Grading: There will be 3 quizzes (a total of 35%) given after each block of lectures. Homework will count 35% and the final exam will count 30%.

Date	Topic	Lecturer
Tu 8/29	**Basics:** Hardy-Weinberg, meiosis and mitosis overview, crossing over, comparison of genetic and physical maps, genetic disease models	Nancy Saccone
Th 8/31	**Linkage I:** Ascertainment. Definition of linkage, phase, recombination. Map functions. Likelihoods and maximum likelihood estimation. Definition of LOD score. Phase known and phase unknown examples.	Nancy Saccone
Tu 9/5	**Linkage II:** LOD score examples, continued. Interpretation of LOD scores. Model specification issues. Genetic heterogeneity, penetrance, phenocopies, liability classes	Nancy Saccone
Th 9/7	**Linkage III:** Multipoint analysis and map functions (revisited). Algorithms for likelihood calculations.	Nancy Saccone
Tu 9/12	**QUIZ, followed by Computer Lab:** LINKAGE (MLINK)	Nancy Saccone
Th 9/14	**Non-parametric analysis I:** Measures of relatedness; IBD and IBS; recurrence risk to relatives; familial recurrence patterns; affected sib pair and affected relative pair analyses.	Arpana Agrawal
Tu 9/19	**Non-parametric analyses II**: Definition of inheritance vectors and the NPL score. Variance components approaches.	Arpana Agrawal
Th 9/21	**QUIZ, followed by Computer Lab: Merlin, PREST-plus**	Arpana Agrawal
Tu 9/26	**Determination of Power for Mendelian and Complex Trait Analysis:** Power and Type I error, multiple testing, prior probability of linkage, ELOD and EMLOD, power for a Mendelian qualitative trait through simulation, power for a complex trait in terms of the risk ratio lambda, number of affected sib pairs needed to achieve a given power under an additive model with from 1 to 10 genes, with different heritability, and with different prevalences.	John Rice
Th 9/28	**Computer Lab:** SLINK	John Rice
Tu 10/3	**Linkage Disequilibrium Analyses:** Haplotype vs. genotype, measures of disequilibrium, decrease of linkage disequilibrium over generations, sources of disequilibrium, problems due to population stratification, use of the EM algorithm to estimate haplotype frequencies, discussion of standard measures (r-squared, D and D-prime), relationship of disequilibrium with physical distance, and a discussion of haplotype	John Rice

	block structure.	
Th 10/5	**Computer Lab:** HAPLOVIEW	John Rice
Tu 10/10	**Allelic Association I -- Case-control designs:** Derivation of allele differences in cases and controls for the generalized single locus model, testing for haplotype differences under the single locus model, general test for haplotype differences, analysis using contingency tables, logistic regression and conditional logistic regression.	John Rice
Th 10/12	**Computer Lab:** SAS (PROC FREQ, PROC LOGISTIC), PLINK (10/17 – Fall Break)	John Rice
Th 10/19	**Allelic Association II -- Family Based Studies:** TDT test, haplotype relative risk (HRR) test, joint distribution of transmitted and non-transmitted alleles from a parent, use of multiple affected sibs in a family, association using pedigrees. QTDT	John Rice
Tu 10/24	**Computer Lab:** TDT in PLINK	John Rice
Th 10/26	**Sample GWAS Analysis/Quiz**	John Rice

Tu 10/31	**Whole Genome Association Analysis I:** Empirical LD (The HapMap Project & 1000 Genomes), Tag SNPs, Haplotype Blocks	Nancy Saccone
Th 11/2	**Whole Genome Association Analysis II**: Imputation, , multiple testing issues, meta-analysis	Nancy Saccone
Tu 11/7	**Computer Lab:** Imputation	Nancy Saccone
Th 11/9	**Introduction** Review of sequencing technologies	Ira Hall
Tu 11/14	**Analysis of sequence data. The C-alpha test**	John Rice
Th 11/16	Rare variant detection, association analysis, annotation (VEP), and functional categorization (GEMINI).	Karyn Meltz Steinberg
Tu 11/21	**Cancer Genome Sequence Analysis**: reference alignment, variant calling, germline vs somatic, SNVS, indels, CNVs, SVs, LOH, visualization, and manual review of events. (11/23 – Thanksgiving)	Obi Griffith
Tu 11/28	**EPACTS and meta-analysis methods I**	Adam Locke
Th 11/30	**EPACTS and meta-analysis methods II**	Adam Locke
Tu 12/5	**Analysis of RNA.Seq Data**	Eli Robertson

Th 12/7	**TBA**	TA
Tu 12/12	Review Session	
Tu 12/19	Final Exam	