M21-515 Fundamentals of Genetic Epidemiology  
Summer 2017

Revised: 6/25/2017

Course Masters  
Treva Rice, Ph.D. (treva@wustl.edu)  
Yun Ju Sung, Ph.D. (yunju@wustl.edu)

Grading  
Computer Lab Assignments  30%  
Daily Quiz  30%  
Midterm Exam (covers lectures / homework 1-5)  20%  
Final Project  20%  
Final Grade (+/– letter grades)

Software  
Software packages include:  
(1) R (http://www.r-project.org/)  
(2) PEDSTATS, Merlin and QTDT (http://www.sph.umich.edu/csg/abecasis/software.html)  
(3) PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/)

Textbook  
Austin, MA. Genetic Epidemiology: Methods & Applications  
2013, CABI: Oxfordshire, UK.

Format  
2-week INTENSIVE course  
Morning lecture: 9:00 am – 12:00 noon  
Afternoon computer lab: 1:30 pm – 4:00 pm  
10-minute break ~ every hour

1st Homework Assignment  
Due first day of class, Textbook ( Chapters 1-2)  
On-line genetics tutorial, Chapters 3, 4, and 6 (http://anthro.palomar.edu/tutorials/biological.htm)

Prerequisite  
1) Knowledge of R programming, either having taken the MSIBS summer R-course or experience in R-programming; 2) experience in Linux/Unix operating system.

The core competency for the Fundamentals of Genetic Epidemiology course (M21-515) is for students to understand basic concepts, methods and analytical approaches in genetic epidemiology.

Learning objectives are to

- Understand familial resemblance, heritability and family study designs  
- Appreciate maximum likelihood methods and hypothesis testing  
- Be aware of selected molecular and population genetics principles, including Hardy-Weinberg Equilibrium  
- Grasp the basic concepts and principles underlying genetic linkage and association
• Be able to perform analysis in heritability, linkage and association using selected software and critically evaluate and interpret the corresponding results

Additional Information
1. Quiz based on lecture/assignment from previous day.
2. No make-up quizzes unless pre-arranged before day of quiz.
3. Homework (reading) due BEFORE day of assignment
4. Computer lab (practicum) due on day following assignment BEFORE lecture.
5. Five points per day deducted for lateness in submitting practicum assignment.
6. Final project consists of oral presentation and one-page written report.

Additional Reading

1. Heritability:

2. Molecular Genetics:

3. Markers:

4. Huntington disease:
# Course Syllabus

<table>
<thead>
<tr>
<th>Day</th>
<th>Lecture</th>
<th>Instructor</th>
<th>Reading Homework</th>
<th>Computer Lab Practicum</th>
<th>Instructor</th>
</tr>
</thead>
</table>
| 1     | 7/13/Th
Overview of Course
Overview of GE
Heritability, Family Designs | Rice       | On-line Tutorial Ch 1 – 2
Additional Reading 1 | Heritability, Merlin | Rice         |
| 2     | 7/14/F
MLE
Hypothesis testing
Molecular Genetics (DNA) | Rice       | Ch 3
Additional Reading 2 | Phenotype Data QC,
PEDSTATS, R | Rice         |
| 3     | 7/17/M
Population Genetics
(Mendel, Segregation, HWE)
Genetic Markers | Rice       | Ch 3
Additional Reading 3 | Phenotype Data QC,
Rice
Review Midterm Exam | Rice         |
| 4     | 7/18/T
GxE, Non-Mendelian Genetics
Ethics, Public Health | Rice       | Ch 8, 10, 11 | 1:30 pm – 4 pm (Lecture Room) | Rice         |
| 5     | 7/19/W
Gene-mapping Example
Introduction to Linkage | Sung       | Ch 3, 4
Additional Reading 4 | Genotype Data QC using
PEDSTATS
Introduction to Merlin | Sung         |
| 6     | 7/20/Th
Model-based Linkage
Model-free Linkage | Sung       | Ch 4 | Linkage Analysis using
Merlin | Sung         |
| 7     | 7/21/F
Association Studies | Sung       | Ch 5 | GWAS using PLINK (Part 1) | Sung         |
| 8     | 7/24/M
Population Stratification | Sung       | Ch 6 | GWAS using PLINK (Part 2) | Sung         |
| 9     | 7/25/T
Analysis of Rare Variants
Data Resources | Sung       | Ch 5, 9 | Practice Presentation
(Lecture Room) |               |
| 10    | 7/26/W
Final Project Presentation
9:00 - noon and 1:30 - 4 pm (Lecture Room) |               |               | Note that the schedule and topics are subject to change. |               |
Final Project

Each student will choose one topic from the list provided below (from page 4 to page 7). After doing research, reading suggested (and additional) papers on the topic, you will give a 10-minutes oral presentation and submit a one-page written report.

You need to give a

- Practice presentation (5 minutes) on July 25, Tuesday afternoon (1:30 – 4 pm in the lecture room) to show us the rough idea of what you will present the next day. You will have a chance to see how your classmates are doing and will receive some feedback and suggestions from your instructor and TA.
- Final presentation (10 minutes) on July 26, Wednesday (in the lecture room). You may have questions at the end of your presentation. Both instructors and TA will provide scores on your presentation. Also Karen Schwander has kindly agreed to listen to your presentation and give a score.
- One-page written report (by midnight on July 26, Wednesday). I will be responsible for grading your written report.

As we want each topic to be covered by only one student, no two students should choose the same topic. Choose the topic that you want to present at your earliest convenience before others choose.

Here are some guidelines for choosing a topic and preparing a presentation:

- The order of presentation will follow the order of topics that I have created.
- Some topics will be covered during my lecture. You need to present new information, rather than repeating what you have learned from my lecture.
- We expect that
  - Your voice should be loud and clear.
  - Your presentation should be well organized. A 10-minute presentation is very short compared to the many hours you spend researching the topic.
  - You only include the materials that you understand or consider important.
  - Your slides should not be too pretty (Please use simple style to avoid distractions!)

Please educate your classmates and instructors. We have a high expectation and confidence that you will do well!
1. Statistical Methodologies for GWAS Era

1a. Genotype Imputation


1b. Population Stratification


1c. Meta-analysis


2. Beyond GWAS

2a. Missing Heritability


2b. 1000 Genomes Project

- [http://www.1000genomes.org/](http://www.1000genomes.org/)
- The 1000 Genomes Project. A map of human genome variation from population-scale sequencing. Nature 467:1061-1073. October 2010
2c. Rare Variants


2d. PheWAS (reverse of GWAS)


2e. Machine Learning in Genetics

- Machine learning applications in genetics and genomics: http://www.nature.com/nrg/journal/v16/n6/abs/nrg3920.html
- From Statistical Genetics to Predictive Models in Personalized Medicine: http://videolectures.net/nipsworkshops2011_personalized_medicine/
- Regularized Machine Learning in the Genetic Prediction of Complex Traits: http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1004754

3. From Association to Function

3a. ENCODE Project

- https://www.encodeproject.org/
- http://www.nature.com/encode/#/threads

3b. UCSC Genome Browser

- http://genome.ucsc.edu/

3c. HaploReg

4. Genetics Software/Programs

4a. ABEL Suite
• [http://www.genabel.org/](http://www.genabel.org/)

4b. METAL
• [http://genome.sph.umich.edu/wiki/METAL_Documentation](http://genome.sph.umich.edu/wiki/METAL_Documentation)

4c. RAREMETAL
• [http://genome.sph.umich.edu/wiki/RAREMETAL_Documentation](http://genome.sph.umich.edu/wiki/RAREMETAL_Documentation)