NEW YORK UNIVERSITY
DEPARTMENT OF BIOLOGY

QUANTITATIVE METHODS IN HUMAN GENETICS
(V23.0045)

SPRING 2010 SEMESTER

Sessions: Two 75-minute lecture classes per week
One 60-minute recitation per week

Location: Computer laboratory classroom

Required texts: *Human Molecular Genetics, 4th edition*
Tom Strachan and Andrew Read (2010, Garland Science)

Introduction to Probability and Statistics for Engineers and Scientists
Sheldon Ross (2009, Academic Press)

Recommended text: *Introductory Statistics with R*
Peter Dalgaard (Springer)

Instructor: Professor David Gresham

Teaching Assistant: TBD

Course Aims: To provide an integrated study of the biological basis of human heredity and statistical approaches to identifying and studying human genes. Students will learn fundamentals of statistical analysis and basic computer programming skills.

Prerequisites: Principles of Biology I
Principles of Biology II
Molecular and Cellular Biology I
Molecular and Cellular Biology II

Grading:

Weekly assignments:	50%
Quizzes:	10%
Midterm:	15%
Final Exam:	25%
Attendance/Participation:	5%

The quizzes are intended to test memorization of key facts and will be held during the first five to ten minutes of each recitation.

Assignments will typically be problem based and require the use of R. With each assignment you should submit a .R file containing the computer code used to generate your results. The code should include comments describing what each step in the code is doing.

Course Description:

Deciphering the information encoded in the human genome is one of the great challenges of the 21st century. This course will provide an introduction to the human genome and the statistical methods that are required for its study. Fundamental concepts in human genetics will be introduced including inheritance of mendelian disease, population genetics, multifactorial disease and functional genomics. Accompanying each topic will be an introduction to the statistical concepts and tools that are required to study inheritance, genes and gene function including probability, hypothesis testing, ANOVA, regression, correlation, likelihood and principal component analysis. Hands on experience will be provided through weekly exercises using the statistical programming language, R. Prior experience with statistics is beneficial, but not required.

Policy on missed tests:
Tests will be excused only for medical or family emergencies. I need to be notified by phone or email before the exam time. An **unexcused** absence from an exam will be calculated as 0% for that particular test! If you miss an exam and present a legitimate excuse, a make-up test will be made available to you. There will be only one opportunity for such an exam; it could be an essay test, and the appropriate instructors will grade it. This situation will be dealt with partly on an individual basis.

Assignments:
Must be handed in on time. Late assignments will be penalized 25% if they are handed in during the next class meeting and will not be accepted after that.

Course Syllabus

Each class will address a topic in human genetics (**Genetics**) and statistics (**Statistics**). The relevant chapters of the two text books are indicated. Additional readings will be provided in class. Please complete the readings **before** class.

Part I: Mendelian inheritance in humans

Lecture 1: **Genetics:** Distributions of human phenotypes; *Strachan and Read Chapter 3*
Statistics: Descriptive Statistics; *Ross Chapter 1*

Lecture 2: **Genetics:** Mendelian inheritance; *Strachan and Read Chapter 3*
Statistics: Standard error, counting error; *Ross Chapter 2*

Lecture 3: **Genetics:** Chromosomal inheritance, recombination; *Strachan and Read Chapter 2*
Statistics: Introduction to probability; *Ross Chapter 3*

Lecture 4: **Genetics:** Mutations in the human genome; *Strachan and Read Chapter 2*
Statistics: Theoretical distributions: binomial, poisson and normal distributions,

Lecture 5: **Genetics:** Pedigree analysis, modes of inheritance; *Strachan and Read Chapter 3*
Statistics: Rules of probability, confidence intervals

Lecture 6: **Genetics:** Mendelian inheritance and disease risk; *Strachan and Read Chapter 3*

Statistics: Conditional probability, Bayes theorem

Lecture 7: **Genetics:** Non-mendelian single gene traits; *Strachan and Read Chapter 3*
Statistics: Advanced Probability

Lecture 8: **Genetics:** Genetic markers and gene mapping in families
Statistics: Likelihood, maximum likelihood estimation; *Ross Chapter 7*

Lecture 9: **Genetics:** Genetic mapping of mendelian traits; *Strachan and Read Chapter 14*
Statistics: Likelihood ratio tests, LOD scores

Lecture 10: **Genetics:** Genetic mapping II
Statistics: Statistical power

Part II: Genes in populations

Lecture 11: **Genetics:** Gene frequencies in populations; *Strachan and Read Chapter 13*
Statistics: Allele frequencies, Hardy-Weinberg equilibrium

Lecture 12: **Genetics:** Comparison of phenotypes between populations
Statistics: Comparing two means, t-tests, ANOVA; *Ross Chapter 10*

Lecture 13: **Genetics:** Genetic structure of the human population
Statistics: Hypothesis testing, p-values; *Ross Chapter 8*

Lecture 14: **Genetics:** Population isolates
Statistics: F_{ST} statistics, Regression; *Ross Chapter 9*

Lecture 15: Midterm exam

Lecture 16: **Genetics:** Linkage disequilibrium
Statistics: D statistics, correlation coefficients

Lecture 17: **Genetics:** Forensic applications
Statistics: Joint probability; *Ross Chapter 4*

Part III: Complex traits

Lecture 18: **Genetics:** Inheritance of multifactorial human traits, Dominance, Epistasis; *Strachan and Read Chapter 15*
Statistics: Heritability

Lecture 19: **Genetics:** Linkage analysis of complex traits
Statistics: LOD scores and likelihood II

Lecture 20: **Genetics:** Association studies I; *Strachan and Read Chapter 16*
Statistics: Chi-square tests, Fisher's exact test; *Ross Chapter 11*

Lecture 21: **Genetics:** Disease risk in populations; *Strachan and Read Chapter 18*
Statistics: Odds ratio

Part IV: Studying the function of human genes

Lecture 22: **Genetics:** Sequencing the human genome; *Strachan and Read Chapter 9*
Statistics: Hidden Markov Models

Lecture 23: **Genetics:** Whole genome expression analysis; *Strachan and Read Chapter 11*
Statistics: Principal component analysis

Lecture 24: **Genetics:** Model organisms; *Strachan and Read Chapter 10*
Statistics: Coalescence

Lecture 25: **Genetics:** Pathways and Networks
Statistics: Cluster analysis, Introduction to statistical algorithms

Part V: The future of human genetics

Lecture 26: **Genetics:** Cancer Genetics; *Strachan and Read Chapter 17*
Statistics: Permutation testing; *Ross Chapter 15*

Lecture 27: **Genetics:** Identifying genes underlying humanness
Statistics: Multiple hypothesis testing

Lecture 28: **Genetics:** Personalized medicine, pharmacogenomics and consumer genomics; *Strachan and Read Chapter 19*
Statistics: Interpretation of statistics