BTRY 6820/4820 (concurrent)  
Statistical Genomics: Coalescent Theory and Human Population Genomics  

"Theories have four stages of acceptance. i) this is worthless nonsense; ii) this is an interesting, but perverse, point of view, iii) this is true, but quite unimportant; iv) I always said so."  


Spring 2016  
Room: Comstock B108  
LEC: TR 10:10-11:25am  
DIS: R 12:20-1:10pm  

Instructor: Prof. Alon Keinan, ak735  
Office Hours: T 4-5pm, 102C Weill Hall  

TA: Andrea Slavney, ajs592  
Office Hours: T 1-2pm, 102 Weill Hall conference room  

Course website: On blackboard (blackboard.cornell.edu – automatically added if signed up for class)  

Grading: S/U or Graded (4 credits)  

I. Rationale:  
Introduction to quantitative molecular population genetics through a coalescent theory derived framework. We will discuss inference within and outside the coalescent framework, with a focus on methods for human population genomics. We will also cover statistical tests for natural selection utilizing the coalescent theory and other frameworks. Throughout the course, we will connect the theory to current research in the field of human population genomics, including studies on modern human adaptation, genetic history, admixture with archaic hominids, the recent explosive growth of human populations, human population structure, and tools employed by direct-to-consumer services such as offered by 23andMe and the Genographic Project.  

II. Specific topics:  
selection evidence in human populations, Diploid and two sexes coalescent, Ancestral recombination graph, Linkage disequilibrium in human populations, Coalescent simulations, Statistical inference and the coalescent.

III. Format and Procedures:
Two “lectures” and one discussion weekly. “Lectures” will be interactive, will include in-class exercises and assignments, and will require active student participation, including completion of readings before class to facilitate in-class discussions, and including student presentations (see below for additional details regarding student presentations). Active participation in both “lectures” and discussion will constitute a component of the final grade (see below).

IV. Prerequisites:
The course is intended for graduate and advanced undergraduate students in fields/majors such as computational biology, biometry, statistics, applied math, genetics and development, ecology and evolutionary biology, and computer science. While there are no strict prerequisites for this course, statistical, mathematical, and computational skills are highly encouraged. Homework will include programming (preferably in C/C++, Python, or Java, but R, Matlab, Perl and others can also be considered). A background in genetics will also be helpful. If in any doubt about your background, please do not hesitate to contact the instructor or the TA.


VI. Requirements and Grading
1. Participation in class and section (20%)
2. 4 homework assignments throughout the first part of the semester (before spring break), including theoretical and programming assignments. (8.75% each assignment for graduate (35% total), 11.25% for undergraduate students (45% total))
3. Final research project, including a written plan and a written final report (35%). Permissible research projects include, but are not limited to
   1. Substantial analysis of genetic data, either from publicly available databases, generated in the laboratory by the student, or provided by the instructor
   2. New theory development as proposed by the student, group of students, or instructor
   3. Research project of a quantitative nature as proposed by the student, group of students, or instructor
   4. (Undergraduate and S/U students only) Literature review in a subject related to the class
   5. (Undergraduate and S/U students only) Implementation of a published method

Several possible project suggestions will be introduced in class. Project will be graded differently for graduate and undergraduate students.
4. (10% for graduate; undergraduates can do for extra credit)
   30-minute class presentation of (1) journal paper pertaining to material covered in class, (2)
   final project proposal, or (3) progress with final project. An extensive list of suggested papers
   for (1) above is provided in Suggested Reading List below, and recommendation out of these or
   of other papers will be made by the instructor or TA.

VII. Academic Integrity
   Each student in this course is expected to abide by the Cornell University Code of Academic
   Integrity. Any work submitted by a student in this course for academic credit will be the
   student's own work, excluding cases in which the final research project is conducted jointly by a
   group of students. You are, however, highly encouraged to study together and to discuss
   information and concepts covered in lecture and sections with other students. You can give
   "consulting" help to or receive "consulting" help from other students.

VIII. Accommodations for students with disabilities
   In compliance with the Cornell University policy and equal access laws, instructors are
   available to discuss appropriate academic accommodations that may be required for student
   with disabilities. Requests for academic accommodations are to be made during the first three
   weeks of the semester, except for unusual circumstances, so arrangements can be made.
   Students are encouraged to register with Student Disability Services to verify their eligibility for
   appropriate accommodations.

VIII. Tentative Course Schedule (May change to accommodate student needs & guest presenters)

<table>
<thead>
<tr>
<th>Class</th>
<th>Topic</th>
<th>Suggested reading beyond main textbook*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (1/28)</td>
<td>Overview</td>
<td></td>
</tr>
<tr>
<td>2 (2/2)</td>
<td>Introduction to Genetics and Genealogies</td>
<td></td>
</tr>
<tr>
<td>3 (2/4)</td>
<td>Intro to Coalescent &amp; The Moran and Wright-Fisher models</td>
<td></td>
</tr>
<tr>
<td>4 (2/9)</td>
<td>Standard Coalescent</td>
<td></td>
</tr>
<tr>
<td>5 (2/11)</td>
<td>Properties of Coalescent Genealogies: Distribution of branch lengths</td>
<td></td>
</tr>
<tr>
<td>6 (2/18)</td>
<td>Properties of Coalescent Genealogies: Branching Patterns &amp; Polymorphisms (Infinite-Sites Model)</td>
<td></td>
</tr>
<tr>
<td>7 (2/23)</td>
<td>Ancient DNA, archaic human populations, and Neandertal DNA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(guest presentation by Aaron Sams, PhD)</td>
<td></td>
</tr>
<tr>
<td>8 (2/25)</td>
<td>Student presentations: Evidence of Neandertal-Human and Denisova-</td>
<td>[1-14]</td>
</tr>
<tr>
<td></td>
<td>human admixture</td>
<td></td>
</tr>
<tr>
<td>9 (3/1)</td>
<td>Allele Frequency-based Tests of “Neutrality”</td>
<td></td>
</tr>
<tr>
<td>Date</td>
<td>Event</td>
<td>References</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
</tbody>
</table>
| 10 (3/3) | - Allele Frequency-based Tests of “Neutrality” – cont’d  
- Introduction of Additional Types of Natural Selection Tests  
- (First Discussion of Final Project) |                                                                                           |
| 11 (3/8) | Linkage Disequilibrium-based Tests of Natural Selection  
guest presentation by Kaixiong Ye, PhD) | [15, 16: sections 2.5 & 2.6; 17-20]                                       |
| 12 (3/10) | Tests of Natural Selection based on Contrasting Polymorphism and Interspecies Divergence  
guest presentation by Leonardo Arbiza, PhD) | [16: sections 2.5 & 2.6; 21]                                               |
| 13 (3/15) | Student presentations:  
Evidence of Natural Selection in Human Populations |                                                                                           |
| 14 (3/17) | Demographic History & Recent Explosive Growth of Human Populations  
guest presentation by Feng Gao) | [33-49]                                                                          |
| 15 (3/22) | Population Structure inference  
(via Principal Component Analysis & STRUCTURE) |                                                                                           |
| 16 (3/24) | Population Structure inference – cont’d |                                                                                           |
| (3/29) | **Happy Spring Break!** |                                                                                           |
| (3/31) | 17 (4/5) | Student presentations: Project Proposal Ideas |                                                                                           |
| 18 (4/7) | Structured Coalescent |                                                                                           |
| 19 (4/12) | Student presentations:  
Methods and Applications for Inference of Population Structure, Admixture, and Local Ancestry Deconvolution of Human Populations | [50-75]                                                                          |
| 20 (4/14) | Diploid and Two Sexes Coalescent |                                                                                           |
| 21 (4/17) | Ancestral Recombination Graph |                                                                                           |
| 22 (4/21) | - Additional Views of Linkage Disequilibrium measures  
- More on Inference in the Coalescent |                                                                                           |
| 23 (4/26) | Student presentations: Linkage Disequilibrium in Human Populations | [54, 76-86]                                                                  |
| 24 (4/28) | Student presentations: Final Research Project |                                                                                           |
| 25 (5/3) | Age of Mutation Inference |                                                                                           |
| 26 (5/5) | Identity-by-Descent Inference |                                                                                           |
| 27 (5/10) | Summary, conclusions, Q&A, and future research directions |                                                                                           |

**Suggested Reading List**

*Note: Only a small number of the following papers will be covered in class, and mostly in student presentations and guest presentations as outlined in course schedule above.*


