

Basic Association and Linkage Course (formerly the Rockefeller University Basic Association and Linkage Course)

Rutgers University
Nelson Biological Labs
Rutgers University, Piscataway, New Jersey
July 11-15, 2011
Document updated: 03/17/2011

Course Description

This course is mainly focused on strategies for genetic mapping of human disease genes. The course is for researchers with little or no experience in using statistical genetic analysis programs. A basic knowledge of association and linkage analyses is, however, required. The primary focus of the course will be association analysis methods and next-generation sequencing statistical analysis methods, with approximately 1/2 day spent discussing linkage.

Topics include: Introduction to association and linkage analysis; practical aspects of data collection and analysis; basic statistics for performing population-based and family-based association analysis and linkage analysis; association and linkage analyses with real data; genetic heterogeneity; power and sample size calculations for population-based and family-based association; population stratification; genomic resources for gene mapping; introduction to computer simulation; and next-generation sequencing methods and practice analyses. The major part of the course will consist of carrying out exercises using various real and simulated data sets.

The course will take place in the computer lab room of the Nelson Biological Labs at Rutgers University, which is equipped with desktop PCs. The number of participants is limited to 20. We may increase that number to 25 if 5 people bring their own laptops. For additional information, please contact Katherine Montague.

Contact information:

Email: katherine.montague@rockefeller.edu

(P): +1 (212) 327-7979,

(F): +1 (212) 327-7996.

Applications are accepted any time after the course announcement, however, participants are generally accepted on a "first come, first served" basis. We usually receive many more applications that we can accept, so you should apply as soon as possible. Since we sometimes have cancellations, we encourage you to apply at any time, but your chances are reduced if you apply late.

Operating systems for course

We will be using PCs running Windows XP for the course. We will also be using Cygwin (www.cygwin.org) (a Windows-based interface for running Unix/Linux commands) to run a number of programs. For that reason, it is most helpful to have at least a rudimentary knowledge of such commands. You can find free documentation on an Introduction to Linux by going to the URL:

http://www.linux-books.us/linux_general_0003.php

In this download, you only need focus on Chapters 1 and 2. This introduction is one such example of freely available documentation.

An EXCELLENT book on teaching yourself Unix (which is the predecessor of Linux and for which many of the commands are the same) is: "Teach Yourself Unix in 24 Hours" (2005) by Dave Taylor.

The internet is replete with resources on introductions to Linux, and Linux basic commands "cheat sheets". With a little searching, you may find exactly what you are looking for! **Please familiarize yourself with the basic Linux commands before coming to class.** It will greatly facilitate your ability to use the computer programs when doing the exercises.

Finally, text editing will be performed using common programs such as Notepad and Wordpad.

Accommodations

We will forward information for reserving a hotel room at a nearby hotel that has been selected to provide good value at reasonable cost. Shuttle service will be arranged for transportation between the hotel and the course computer lab at Rutgers University. Although New York City is nearby, it is not practical to plan to stay in NYC during the course. Travel time from Penn Station (train) in NYC to the New Brunswick station (closest to the university) is 1 hour, plus an additional 30 minutes would be required to take the bus from the train station to the course computer lab.

Course Fee

The fee for the 5-day-course is \$1200 for researchers at an academic institution, and \$1600 for individuals from private (for profit) companies. Send no money now. Applicants accepted for the course will receive payment instructions. As there is presently no support for this course from sources other than the course fee, no reduction of the cost to applicants is possible. This fee covers tuition and course related expenses (handouts, etc.) but not room, board or meals. Course participants may find reasonably priced hotels in the the New Brunswick/Piscataway area by using software such as <http://www.kayak.com> or <http://www.priceline.com> (In Priceline, you can bid on prices for hotels). That is, participants will have to make their own travel and hotel arrangements (except foreign participants).

Course Instructors and Schedule

Each session will start with a theoretical introduction followed by practical exercises with several different computer programs. Instructors for the course will be (in alphabetical order): Marcella Devoto, Derek Gordon, Robert Klein, Mark Levenstien, Douglas Londono, and Tara Matisse. Be sure to bring a calculator, and a flash drive in case you wish to copy programs and course data files. Also, you may want to bring data sets for discussion and/or analysis.

The course starts at 9:00 am in the 1st Floor Computer Lab of the Nelson Biological Building on the Rutgers University Busch campus. This building is located at the corner of Bevier Road and Allison Road in Piscataway, NJ. You can locate the building using Google Maps, for example. **If you plan to drive to campus each day, we need to know in advance to arrange sufficient guest parking.**

Morning sessions will usually go from 9:00 am to 12:00 pm, and afternoon from 1:30 pm to 5:00 pm, but there may be slight variations (+/- 1/2 hour) in the time the morning session ends and the afternoon session begins. We'll have a 1 to 1-1/2 hour break for lunch every day. We'll also take a 10-15 minutes coffee break every day in the morning usually between 10:15 and 11:15, and in the afternoon between 3:00 and 4:00. Below is a **tentative** schedule of the course. Days and times for different lectures may be shifted around. We will make every effort to keep you up to date of any such changes.

MONDAY, July 11

AM

Introduction to gene mapping (D. Gordon)

File format for family-based and population-based data (D. Gordon)

PM

Data quality and data cleaning (D. Gordon)

TUESDAY, July 12

AM

Introduction to basic association statistics, including next-generation sequencing statistics (D. Gordon)

PM

Introduction to association study analyses, including next-generation sequencing analysis (D. Gordon)

Study-design for association analysis (D. Gordon)

WEDNESDAY, July 13

AM

Genomic resources (T. Matisse)

PM
Simulation Analysis with SLINK and SUP (D. Gordon)

THURSDAY, July 14

AM
Genetic Heterogeneity (M. Levenstien)
Population Stratification (D. Londono)

PM
Linkage Analysis with genome-wide data using MERLIN (M. Devoto)

FRIDAY, July 23

AM
Case-control association studies (R. Klein)

PM
General discussion (D. Gordon)
Looking forward to seeing you!!

Sincerely,

A handwritten signature in blue ink, appearing to read "Derek Gordon".

Course Organizer and Instructor