

Elective Course Proposal

Course Title: Population Genetics Modeling and Inference in R

Proposer/Lead Instructor Name: Cheryl Hayashi; Marcelo Gehara

Other Course Instructor(s): Ariadna Morales, Arianna Kuhn, Sean Harrington

Course Description: This course is an introduction to population genetics modeling and inference in R. The course will cover an introduction to the R statistical platform, the coalescent model, simulation-based inference methods such as approximate Bayesian computation and supervised machine-learning. Additionally, there will be an introduction to population assignment methods. By the end of the course attendees will have all the tools to perform population genetics model-inference after data assembly. The course will cover the following software and r-packages: r-base, adegenet, LEA, conStruct, Phrapl, Pipe Master, abc, and caret.

Duration in weeks: 1

Number of lectures/seminars/meetings per week: 5

Lecture/seminar/meeting length per session (min.): 50 min

Number of labs per week: 5

Lab length per session (min.): 100 min

Credits: 1

Location: RGGS, AMNH

Prerequisites: none

Learning Objectives:

- Basic knowledge of R statistical platform (first day)
- Basic knowledge of population assignment methods (second day)
- Basic Understanding of the coalescent model and its properties (third day)
- Basic knowledge of phrapl r-package (third day)
- Basic knowledge of PipeMaster r-package (fourth day)
- Basic knowledge of abc r-package (fifth day)
- Basic knowledge of machine-learning in r with caret r-package (fifth day)

Bibliography:

An introduction to R (<https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>)

An introduction to statistical learning with applications in R

Coalescent Theory; An Introduction by John Wakeley

www.github.com/ariadnamorales/phrapl-manual

www.github.com/gehara/PipeMaster

Grading/Evaluation Basis: None

Final Exam/Project Description: None

Statement on Academic Integrity: Each graduate student bears the responsibility to observe traditional canons of scholarly discourse, scientific research, and academic honesty. Plagiarism, cheating, and fraud in research will not be tolerated. Accordingly, it is expected that students work individually unless specifically instructed to work in groups. The full Academic Integrity policy is in the student handbook.

Course Evaluations: Each student is required to complete an anonymous course evaluation at the end of the term. The course evaluation is a tool for faculty and administrators to improve the student learning experience.

Course Calendar

Date	Lecture Topics	packages	Lecturer
Jan 13	R statistical platform basics	r-base	Marcelo Gehara
Jan 14	population assignment	r-package LEA/adegenet/con Struct	Sean Harrington/Arianna Kuhn
Jan 15	the coalescent model & properties/phylogeographic inference using approximate Likelihood	r-package PHRAPL	Marcelo Gehara/Ariadna Morales
Jan 16	demographic model construction & coalescent data simulation	r-package PipeMaster	Marcelo Gehara
Jan 17	approximate Bayesian computation overview & application	r-package abc	Marcelo Gehara