

# Syllabus

Course No.: ***GMS6233***

## Course Title: **Quantitative models of protein evolution and phylogenetics**

Spring 2017

2 credit hours

Modules 1-2 (Jan 4 – Mar 9)

Wednesday 10:40am – 12:00 noon

Thursday 10:40am – 12:00 noon

Location: Rm. CGRC 336

**Instructor: Luciano Brocchieri, Ph.D., Assistant Professor.**

### Contact Information

Genetics Institute, Cancer and Genetics Research Complex, Room 114. 2033 Mowry Rd, Gainesville, FL 32610.

Phone: 352.273.8131

Email: [lucianob@ufl.edu](mailto:lucianob@ufl.edu)

Office Hours: By appointment.

### Course Objectives

Methods of phylogenetic analysis from molecular data are widely used but the intricacies and assumptions of these methods are not always appreciated. This two-unit course emphasizes the study of the theoretical principles underlying models of protein evolution, as used in phylogenetic analysis. Through this course students will develop a critical understanding of the biological principles and mathematical tools used to model protein evolution in inferring phylogenetic relations. The students will be guided towards understanding algorithmic applications of quantitative models of protein evolution. The course will consist of interactive lectures in a socratic format and will be tailored to the progress of the students. This formal course is targeted at graduate students interested in mathematical descriptions of protein evolution (e.g., Genetics and Genomics, Biology, IDP). The instructor is a computational biologist with training in theoretical population genetics, evolutionary theory, and bioinformatics.

### Prerequisites

None specified.

### Catalog Description

**Quantitative models of protein evolution and phylogenetics** – Principles of protein structure and function and their evolution. Mathematical models of protein evolution and their application to the study of phylogenetic relations. Dissection of mathematical and statistical tools implemented in methods of phylogenetic-tree reconstruction. Offered spring semester.

## Course Format and Grading

Quantitative models of protein evolution and phylogenetics will be offered for two credits in the spring semester 2017. Seven and a half hours of lecture and discussion will be conducted weekly. Evaluation will include the following: midterm project: 30%; final exam: 35%; attendance and participation: 35%. Evaluations will follow the letter code grade cutoffs according to the following table:

A	A-	B+	B	B-	C+	C	C-	D+	D	D-	E
95-100	90-94	87-89	83-86	80-82	77-79	73-76	70-72	67-69	63-66	60-62	<60

Information on current UF grading policies for assigning grade points can be found at the following web site: <http://www.registrar.ufl.edu/catalog/policies/regulationgrades.html>.

## Course evaluation

Students are requested to promptly submit evaluation of the course after the end of classes.

## Attendance and make-up exam policies

Attendance to all lessons is expected. Excused absences follow the criteria of the UF Graduate Catalogue (e.g., illness, serious family emergency, military obligations, religious holidays), and should be communicated to the instructor prior to the missed class day when possible. Regardless of attendance, students are responsible for all material presented in class and meeting the scheduled due dates for class assignments. Personal issues with respect to class attendance or fulfillment of course requirements will be handled on an individual basis.

*Due to the participative format of the lectures, recording is not allowed.*

## Topical Outline

Protein evolution and evolutionary distance

- The concept of evolutionary distance
- Models of protein sequence differentiation
- Estimating evolutionary distances
- Distance methods in phylogenetic-tree reconstruction

Principles of protein evolution and phylogenetic methods

- Parsimony
- Minimum evolution
- Protein evolution and heterotachy
- Compositional biases and protein evolution

General probabilistic models of protein evolution

- Continuous markov models in protein evolution
- Empirical amino acid substitution matrices
- Codon-substitution models

Position and branch-specific models of protein evolution

- Position dependent substitution models
- Mixture models of protein evolution
- Branch-specific models

#### Maximum likelihood in protein evolution

- Time-reversible substitution matrices
- Computing the likelihood of a tree
- Searching the tree space

#### Bayesian methods in protein evolution

- Markov Chain Monte Carlo (MCMC) methods
- Bayesian MCMC for phylogenies

#### Diversity

- Indexes of diversity
- Indexes of phylogenetic diversity
- Phylogenetic diversity and functional diversity

### **Accommodations for Students with Disabilities**

Students requesting classroom accommodation must first register with the Dean of Students Office. The Dean of Students Office will provide documentation to the student who must then provide this documentation to the Instructor when requesting accommodation.

### **Required and Recommended Textbooks**

No textbook is required. The motivated student may refer to the following informative books on the topics:

*Inferring phylogenies*, 2004, Felsenstein J, Sinauer Associates Inc, Sunderland, MA.

*Molecular evolution and phylogenetics*, 2000, Nei M and Kumar S, Oxford University Press, Oxford, New York.

*Statistical methods in molecular evolution*, 2005, Nielsen R, Springer, New York, NY.

*Molecular systematics*, 1996, Hillis DM, Moritz C and Mable BK (eds.), Sinauer Associates Inc, Sunderland, MA.

*Mathematics of Evolution and Phylogeny*, 2005, Gascuel O, Oxford University Press, Oxford, New York.

*The Phylogenetic Handbook*, 2009, Lemey P, Salemi M and Vandamme A-M (eds.), Cambridge University Press, New York.