COURSE ANNOUNCEMENT: Summer 2018

Course: Spatial Bioinformatics

Instructors: **Mary Blair (AMNH) and Rob Harbert (AMNH); Peter Galante**

Credits: **1**

Dates: **May 14-23, 2018** Daily except Saturday and Sunday (8 class days)

Spatial data and models are ubiquitous in modern comparative biology and ecology due to the vast amount of available data and ever developing modeling methods. This course will focus lectures on a series of "best-practices" in handling and modeling spatial biological data including data-mobilization, bias detection and reduction, geographic projection management, and comparative modeling frameworks. Labs will concentrate on demonstration of best-practices on a range of datasets including student's personal data. The course will culminate in student's working in "hackathon" style working groups to develop a spatial data analysis pipeline to address a question of mutual interest to be posted on appropriate code-sharing repositories.

Daily schedule:

Lecture/Workshop: 10:00-12:00 Working Lab: 1:30-4:30

Day 1 - Lecture: Introduction to programming

Lab: Unix command line, basic data handling in R

Day 2 - Lecture: Geography, coordinate systems, projections, georeferencing, GIS data types (Vector, Raster), sampling bias, spatial autocorrelation, and remote sensing

Lab: What's in a GPS coordinate? GPS vs. cell-phone locality information

Day 3 - Lecture: Occurrence data and distribution modeling

Lab: Comparative niche modeling. Covering - background/absence sampling, environmental layers, model evaluation and parameterization, niche differentiation and overlap.

Day 4 - Lecture: Survey of advanced modeling topics (see lab topics)

Lab: Advanced modeling exercises. (possibly including: past/future range estimation, biogeographic reconstruction, landscape genetics, landscape analysis, modeling biotic interactions, mechanistic models – may be tailored to class interests)

Day 5 - Lecture: Literature Discussion of spatial data in Comparative Biology - Survey of 2-4 papers.

Lab: Advanced topics in R programming and an introduction to Git for code sharing and management. Will cover scripting, functions, and documentation. A primer in preparation for workshops and hackathon

days.

Day 7-8 - BYOD (Bring Your Own Data) Workshops 1-4 pm
Students will bring or find data relevant to their own research (from i.e., Landsat, GBIF, iNaturalist) and will work through practical issues with their data analysis plans. These sessions will also serve as a brainstorming and guidance period for Hackathon project development.

Day 9 - Project/Hackathon day 9AM - 5PM

Students work in small groups a project that will develop a data analysis pipeline to address a question of interest to them. The goal of a hackathon is to create usable software with a specific focus by collaboration between interested parties. Students will develop an R script, an example dataset, and a demonstration. Students will supplement this usable code with a write-up of the supporting literature and target audience for the code. Final code will be shared on the course GitHub to facilitate reuse and further development