Modeling and analysis issues for understanding species' colonization

Activity report from NIMBioS short visit February 15-19, 2016

Rapporteur: Sean Hoban

Participants in the short visit: Sean Hoban (The Morton Arboretum, previously a NIMBioS postdoc), Allan Strand (College of Charleston, previously a NIMBioS sabbatical visitor).

Since the advent of the field of phylogeography in the late 1980s, scientists have used genetic data to understand the geographic distribution of species (range), and how ranges shift through time. By combining spatial analyses and genetic algorithms to analyze the DNA of living organisms, scientists can identify where species used to live, how abundant they were, how they moved in response to climatic changes, and if similar species had similar histories. *Of particular interest is how species colonized open land after the last Ice Age, about 20,000 years ago; this provides insight into how modern species will move under current environmental change, with practical implications for forestry, carbon sinks, and conservation.* While geneticists have discovered much about European species' responses (e.g. Petit et al 2003), little is known about post-glacial colonization in North America and Asia. In our visit, we worked on several issues that have plagued phylogeographers in North America including complex regional geography, sampling and analysis issues, and interpretation of disparate datasets.

In order to test the phylogeographic hypotheses of interest such as colonization, appropriate study design, analyses and genetic tools must be used to gather data. Our visit was explicitly motivated by the absence of current guidelines for the best sampling sizes, spatial sample distribution, study species, molecular markers, and genetic statistics for documenting these patterns. Our two goals were: (1) lay ground for better future studies (e.g. *better sampling and analysis guidelines*) and (2) better utilize existing data (e.g. *enable re-analysis with new modeling approaches*). We performed several tasks during our visit to facilitate these goals:

- Discussion and evaluation of a literature search that we performed prior to our visit,
- Enumerating potentially useful statistics for describing spatial patterns of genetic variation, to be tested using simulation,
- Writing code for simulation so that we can easily model postglacial range dynamics,
- Writing code to analyse these simulated datasets,
- Running trial simulations and recording simple statistics, from which we already observe distinct patterns depending on the parameters of the range expansion,
- Outlining a review on postglacial dynamics in which we will highlight the issues mentioned above, and describe our simulation results.

As of the end of our visit, we have coded the architecture of a model that will allow us to test different historical biogeographies including multiple refuges and long distance dispersal. We will perform these simulations over the next 3-6 months. We will also expand out literature search and review outline, with a publication planned at the end of 2016.