Distinguishability in Genealogical Phylogenetic Networks
13-17 August, 2018 @ Snellius

Description and aims

The traditional model for representing evolution is the phylogenetic tree. However, trees are not suited to capturing "reticulate" evolutionary phenomena, such as hybridization, recombination or lateral gene transfer, i.e. events in which lineages combine, rather than diversify. Genealogical phylogenetic networks explicitly describe such reticulate events. The workshop focused on indistinguishability issues (both patterns and processes) that significantly complicate the reconstruction, analysis and interpretation of genealogical phylogenetic networks.

Tangible outcomes

Because genealogical phylogenetic networks are still extremely under-represented in the biological and linguistic literature, we have planned the following projects:

1. An "advertisement paper" aiming at a broad biological audience, explaining the importance of genealogical phylogenetic networks and giving an update w.r.t. our *Trends in Genetics* paper, which we published after our 2012 workshop. This update is very needed because, in the last few years, several biologically-appealing methods have been designed; and we all agree that they deserve to be advertised more.
2. A phylogenetic network "cookbook", which explains (in detail) to biologists and linguists how to prepare data, apply various phylogenetic network methods, and how to interpret the results, for various types of data.
3. Enhancements to Philippe Gambette's "Who is Who in Phylogenetic Networks" website, to make it a more general-purpose repository of information about phylogenetic networks.

Scientific breakthroughs

Since the workshop did not aim to solve scientific problems, there were no scientific breakthroughs during the workshop. However, various subgroups started working on interesting computational and practical questions that could eventually lead to scientific breakthroughs. For example, a characterization of networks that can be distinguished based on statistical models, and applications of networks in linguistics.

"Aha" moments

There were many such moments during the workshop, for example:

- Linguists and biologists discovering the similarities as well as the differences between their data, networks and interpretations.
- Mathematicians realizing the similarities and differences between the various biological processes that can cause incongruence, which processes need and which need not be distinguished, and how.
- Computer scientists realizing the importance of developing methods that can deal with large data sets and produce a "consensus network" that summarizes a collection of (near-)optimal solutions.
- Biologists realizing that branch lengths in phylogenetic trees may often be underestimated due to underlying (but ignored) reticulate signals, as well as mis-estimation of topologies.

Format of the workshop
The combination of lectures, plenary discussions and discussions in smaller groups worked very well. In particular, the way we formed the groups was successful. Each participant wrote a question / idea / topic on a post-it, which we clustered to form the groups. An idea for next time would be to schedule more, shorter lectures on the first day, to then have more time to discuss these topics during the rest of the week. A workshop with only mathematicians would also be very useful to work on the concrete computational problems.

Organizers
Leo van Iersel   (Delft, The Netherlands)
Steven Kelk   (Maastricht, The Netherlands)
David Morrison (Uppsala, Sweden)
Céline Scornavacca (Montepellier, France)