MBG FOCUS TALK

Hosted by Mikkel Schierup

Tuesday 21 May 2024 at 09:15-10:00
1870-816 (Faculty Club)

MBG Focus Talks in Molecular Biology

Dr. Magnus Nordborg

Scientific Director,

Gregor Mendel Institute of Molecular Plant Biology,

Vienna, Austria

Towards an unbiased characterization of genetic diversity

Our view of genetic diversity is shaped by methods that provide an incomplete and highly biased picture, effectively limited to single-nucleotide polymorphism in conserved regions of the genome. Long-read sequencing technologies, which are starting to provide nearly complete genome sequences for population samples, should solve the problem—except that characterizing and making sense of non-SNP variation is difficult even with perfect sequence data.

I will describe our attempts to investigate and address this problem using samples of genomes from *Arabidopsis thaliana* as an example. Our analyses reveal substantial and worrying biases in current data that affect everything from GWAS and functional genomics to population genetics and diversity studies. We also discover exciting new biology, especially when it comes to understanding the evolutionary dynamics of transposable elements. We demonstrate that existing genome annotation tools do not predict mobile elements well even in a model plant and present alternative algorithms.

Twenty-five years ago, technical developments that came out of the Human Genome Project ushered in the SNP era, leading to a revolution of population genetics—as predicted by Aravinda Chakravarti, who noted that we needed models to "make sense out of sequence". I will argue that we are in an analogous position now, with technologies making it easy to generate complete genomes sequences of almost any species at a population scale—an enormous breakthrough for anyone interested in the full diversity of life. However, to make sense of these data, we will need a modeling framework rooted in population genetics, but which also incorporates accurate mechanistic models of the mutational and recombination processes that ultimately generate genetic variation. This framework remains to be developed.

AARHUS UNIVERSITY