

Population Genomics

Objective: To give a presentation of about 40-90 minutes duration at the end of the week covering the key aspects of the population genomics of humans.

This 1000 genomes project is devoted to the exciting project, where the genomes of a 1000 individuals are to be sequenced in the next few years. This project will pose many challenges both in terms of data generation, data storage and analysis. The questions and contents below are meant as motivators and need not be followed.

The big questions are:

- How to choose the 1000 individuals out of a population of 6 billion to get maximal information?
- Which techniques should/will be applied to get the data?
- How well can parameters be recovered?
- How well can the history of the sequences be known?
- What is the marginal value of an added genome for the key questions?

Maximal Contents of Presentation:

Choice of individuals

Data generation and assembly

Data storage

Analysis

Mutational processes

 Single Nucleotide Substitutions

 Small Insertion-Deletions

 Large Scale Events

Recombination rates: Spatial and Temporal Distribution

Geography

Selection

Population Size

Ancestral Analysis

Pedigrees and the Global Genealogy

Recommended literature:

David R. Bentley et al. "Accurate whole human genome sequencing using reversible terminator chemistry" *Nature* 456, 53 - 59

Butler et al. (2008) "ALLPATHS: De novo assembly of whole-genome shotgun microreads." *Genome Res* vol. 18 (5) pp. 810-20

Chaisson et al. De novo fragment assembly with short mate-paired reads: Does the read length matter?. *Genome Res* (2009) pp.

Hernandez et al. (2008) De novo bacterial genome sequencing.. *GenRes* vol. 18 (5) pp. 802-9

Levy S et al.. 2007. The Diploid Genome Sequence of an Individual Human. *PLoS Biology* 5(10):e254.

Mardis ER (2008) "Next-Generation DNA Sequencing Methods" *Annua Rev. Genom. Human Genetics* Vol. 9: 387-402

Novembre et al. (2008) "Genes mirror geography within Europe" *Nature* **456**, 98-101

Wang et al. (2008) [The diploid genome sequence of an Asian individual](#). *Nature* 456(7218):60-5.

Wheeler et al. (2008) The complete genome of an individual by massively parallel DNA sequencing. *Nature* 452(7189):872-6

Zerbino et al. (2008) Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res* vol. 18 (5) pp. 821-9

Links

<http://www.1000genomes.org/page.php>