

Presentation title: *De novo* genome assembly

Abstract:

De novo assembly is the process of reconstructing a genome's DNA sequence using only a set of much shorter error-prone sequences (reads) sampled from the genome. It is the "original" genomics-based bioinformatics problem, because it is all we can do when we don't have any related reference genome sequences, with the exemplar being the original human genome project. This presentation will discuss the principles of and approaches to *de novo* assembly of data, and practical issues like computational and memory requirements, limitations of *de novo* assembly, terminology, file formats, available software, and an example run-through of an assembly using the Velvet software if time permits