A KJELDGAARD LECTURE

Friday 18 September 2015 at 13:15
Auditorium AU Foulum
Same location for the PhD session

Anders krogh
Department of Biology
University of Copenhagen
Denmark

Bioinformatics of High-throughput DNA sequencing: From read mapping to RNA-seq assembly

New technologies for DNA sequencing are revolutionizing research in many areas of the life sciences. Not only is it possible to sequence complete genomes at an unprecedented rate, but also use sequencing for genotyping, gene expression analysis (RNA-seq), mapping of protein binding in DNA (ChIP-seq), study of ecosystems (meta-genomics), and many other applications. Computational methods are needed to analyze the large amounts of data from these experiments. In this talk, I will first discuss mapping of DNA reads to a reference genome, which is one of the most common first task. Special care must be taken, when dealing with short reads and I will present a mapping method that builds on a probabilistic basis rather than ad hoc measures of mapping quality. The mapping/alignment problem can be formulated in terms of "pair HMMs" (or automata), and I will discuss how this can be used for a generalized mapping of reads, where for instance exon splice sites can be included in the model, and how the search can be done efficiently using a Burrows-Wheeler transform. Another common task is to assemble short reads from shotgun sequencing into genomes or transcripts. I will discuss this problem briefly and present a probabilistic method for transcriptome assembly that we have recently developed, and which is better at identifying lowly expressed transcripts than competing methods.

Host: Ole Fredslund Christensen, Center for Quantitative Genetics and Genomics, Department of Molecular Biology and Genetics, Aarhus University

The lecture will be followed a chalk-board session for PhD students

The Kjeldgaard Lecture Series is organised by
www.mbg.au.dk/lectures

DEPT. OF MOLECULAR BIOLOGY AND GENETICS
AARHUS UNIVERSITY