PhD Course description

Name of course:
Next Generation Sequencing Analysi**s**

ECTS credits: 5

Course parameters:
Language: English
Level of course: PhD course
Semester/quarter:Q4 - June 23-27 one week course
Hours per week: 35
Capacity limits: 40

Objectives of the course:
At the end of the course the students should have a detailed knowledge of bioinformatics methods for genome analysis using next-generation sequencing data. This includes knowledge of the existing types of sequencing data, how the different types of data can be displayed and analysed, the current methods for genome assembly and analysis, their accuracy, and how to apply them. The course will train the participants to devise and analyse a real genomics project involving NGS data.

Learning outcomes and competences:
At the end of the course the student **should be able to**:

- **Describe** key challenges in analysis of NGS data.
- **Explain** the theoretical foundation of methods using NGS for assembling and analysing genomes.
- **Discuss** the bioinformatic methods for genome analysis.
- **Discuss** original literature within the subjects.
- **Use** bioinformatics tools within the selected application areas.

Compulsory programme:
Course participation and submission of a data analysis report.

Course contents:
The course will start with an overview of current NGS technologies and directions in which these are currently progressing. Methods for assembling genomes from NGS data using either a reference genome or *de novo* assembly will be presented together with some emphasis on assessing NGS data quality (Quality control, accuracy of base calling). The course will then offer an overview of the range of biological/biomedical questions that these new types of data can help to address. This includes:

- The use of NGS resequencing of closely related genomes for detecting mutations. This includes the characterization of somatic mutations in cancer tissues, the *de novo* mutation spectrum in eukaryotic genomes, and the detection of mutations underlying phenotypic changes in experimental or natural populations.
- Combining NGS with methods aimed at reducing genome complexity such as exome resequencing in complex genomes.
- The use of resequencing of a large number of individuals for Single Nucleotide Polymorphism (SNPs) and Copy number variation detection.
- Application of NGS in Transcriptomics, Epigenomics and Chromatin structure studies.
- Environmental/Meta-genomics studies probing the bacterial/viral diversity of understudied biomes.
- Comparative and evolutionary genomics.

To allow hands-on data analysis training, the course includes the following computer exercises, which can be completed using either your own or example data:

- Introduction to Galaxy, GATK and Samtools
- Grooming of reads
- Mapping of Illumina reads to reference
- SNP/indel calling, filtering and annotation
- Gene models and transcript quantification

Prerequisites:
3 years study of molecular biology, biology, mathematics, engineering or computer science.

Name of lecturer[s]:
Mikkel Schierup and Stig Uggerhøj Andersen

Type of course/teaching methods:
One week intensive course including lectures, colloquia and computer exercises.

Literature:
Review and research articles.

Course homepage:
http://birc.au.dk/Studies/NGS/

Course assessment:
Individual written report

Provider:
Bioinformatics Research Center (BiRC)

Time:
June 23-27 2014.

Place:
Aarhus University, Building 1421

Registration:
Deadline for registration is May 15th, 2014.

For registration: contact the teachers by e-mail - sua@mb.au.dk or mheide@cs.au.dk

If you have any questions, please contact Stig Uggerhøj Andersen (sua@mb.au.dkl) or Mikkel Schierup (mheide@cs.au.dk).

PLEASE NOTE

Deadline for registration is May 15th, 2014.
If you have any questions, please contact Stig Uggerhøj Andersen (sua@mb.au.dkl) or Mikkel Schierup (mheide@cs.au.dk).