36626 & 36826: Next Generation Sequencing Analysis
DTU - June 2018
Simon Rasmussen

Simon Rasmussen
Associate Professor
DTU Bioinformatics
Technical University of Denmark
simon@bioinformatics.dtu.dk
@simonrasmu
Who are we?

Organizers:
- Simon Rasmussen (Me)
- DTU Bioinformatics
  - Simon Rasmussen (Me)
  - Peter Wad Sackett
  - Yan-Jun Chen (Angie)
  - Jakob Nissen
  - Katrine Kristensen
  - Jose Izarzugaza (Txema)
  - Francesca Bertolini
- DTU Food
  - Pimlapas Leekitecharoenphon (Shinny)
- University of Copenhagen
  - Martin Sikora
Who am I?

- Msc in Chemistry/Biotechnology from DTU
- Research scientist SSI - TB vaccine development
- PhD in bioinformatics
- Associate Professor - NGS analysis (8 years)
- Head of the Metagenomics & Microbiome group
- Also organiser of Microbiome course in the fall (36636)
What do we do?

Ancient metagenomics & pathogen evolution

Personalised medicine

Metagenomics

Machine Learning & Deep Learning
Main Teaching Assistants

Katrine Kristensen
PhD student
Metagenomics & Bacterial Evolution

Jakob Nybo Nissen
PhD student
Metagenomics & Machine Learning
Who are you?

- According to Campusnet
- Bsc students: 4
- Msc students: 26
- PhD students: 16
- Open education: 6
Feedback

- Sixth time we are running the course

- We are still improving - quite some changes from last year

- Please give us feedback!

  - Please do the evaluation at DTU Inside
Learning objectives

Cooperate in groups

Analytical & reflective

Formulate/perform a project

Unix command line

Strength / Weaknesses

Applications

Explain key steps

Theoretical principles

NGS
Why shell terminal?

• Almost all tools for NGS analysis are command line only

• Generally more efficient/flexible, you can play around with the tools/data

• They can be pipelined, ie. analyzing 100 files in windowed mode is a pain ...

• Alternative approaches: Galaxy, CLC-workbench, Geneious
Course structure

- 3 weeks, 2 tracks:

  Lectures + Exercises + Pres.

  Project work

Date: 31. 12. 21.

= Poster exam

Course program: http://www.cbs.dtu.dk/courses/27626/
Course breakdown

• **Thursday 31st May**
  - Introduction NGS tech.
  - Tech talk groups
  - Unix and first look at data

• **Friday 1st June**
  - Tech talk presentations
  - Data basics & preprocessing
Course breakdown

• Monday 4th
  • Alignment
  • Alignment processing
  • Variant calling

• Wednesday 6th
  • De novo assembly
  • De novo metagenomics
Course breakdown

- **Thursday 7th**
  - Quantitative Metagenomics
  - Test + recap

- **Friday 8th**
  - RNA seq
  - Cancer seq

- **Monday 9th**
  - Genomic Epidemiology
  - Ancient DNA
Course breakdown

- **Tuesday 12th**
  - Project group formation

- **Wednesday 13th**
  - Short project presentations
  - Project work

- **Thursday 14th - Wednesday 20th**
  - Project work

- **Thursday 21st**
  - Poster Exam
Some points

• Learn principles of the analysis

• The exercises will be useful for your projects and later

• Team up 2 and 2 for exercises (or do them yourself but discuss with neighbor)

• Please just ask questions at any time!
Cloud computing

- The course has moved to the Cloud!
- Danish National Supercomputer for Life Science (Computerome) located at DTU Risø
- 16048 cores, 92 Tb RAM an 3Pb storage
- We have 2 nodes
  - 28 cores, 128 Gb RAM
Projects

• Try analyze a “real” dataset and present results on poster
• 4-5 pr. group

• You can find a dataset on SRA/ENA

• You can use your own data if everyone in the group agrees and it can be presented on a poster

• Dont analyze too large datasets (time, resources)
Teachers and TAs will be available to help with your projects

Office hours during project period: 10-14

Use Piazza as a platform to communicate with your peers, TAs and teachers

- Collective knowledge
- Access later today
Exam

- Each group will create a poster
- You can print posters at the DTU library for 20-30 kr
- Each group will present the poster for the examiners
- Then each individual in the group will one-by-one be asked questions on the learning objectives and your project (5-10 min).
Disclaimer

- Sequencing technology changes very rapidly!

- We will dive into many areas - you will not master everything

  but

- There are many opportunities - hopefully you will learn to see them
Be adventurous!

You do not have the ability to do anything destructive

Unless you physically destroy our computers!

The worst that can happen is that you lose your own data

Source: Angus
Course web-page

- Course program, Slides, Handouts, Exercises

- We want the course page to be a repository for you!
Reading + wifi

- There are no text-book for the course
- There are papers uploaded to campusnet that you can read for more information
- Wireless networks
  - Use “dtu” and your dtu/campusnet login to get access to wireless
  - Eduroam
- Alternative wifi: “You can haz wifi”
Pre-test

- Test your knowledge before we start
- Not used for grading or exam
- Used to understand where you are