Overview

MEDT32/33

How to get your grade

1. **Personal project presentation - due Friday PM 29 January - 50%**

Prepare a pilot grant proposal.

Proposal should adhere to guidelines in proposal outline document.

Presentation should discuss (1 slide per section) - 5 slides 10 minute presentation

a) technology approach for sequencing

b) choice of methods for analysis to include variant calling and filtering approach

c) example of a file generated from a chosen dataset and justification of dataset choice for analysis

d) interpretation of results to infer clinical outcome

e) description of adherence to clinical calling standards and best-practice guidelines for reporting the clinical significance of analytical results

Writeup

Personal project writeup/proposal - due Friday 5 pm 5 February - 50%

(3 pages maximum)

What to expect

- This is bioinformatics
  - Its science
- You are in charge
- Genome England is work in progress
- Overload

Day 1

8:30 - 9:00		Welcome and introduction
9:00 - 9:45	L	Introduction to variant analysis using HTS data and its caveats
9:45 - 10:30	L	Introduction to Genomics England data, data access policy, how GE data is generated
10:30 - 10:45		Coffee break
10:45 - 12:00	L	Formats of Fastq and BAM. From raw data to aligned data: present FASTQ format and tools for short read alignment (e.g. BWA)
12:00 - 13:00		Lunch
13:00 - 14:00	L	Introduction to Galaxy
14:00 - 15:00	L/P	Inspect a FASTQ file, manipulate alignments, assess their quality and focus on the importance of quality control
15:00 - 15:30		Coffee break
15:30 - 17:30	P	first steps in Galaxy.
17:30 - 18:00		Q&A

Day 2

9:00 - 10:00	L	Overview of tools for variant calling (e.g. GATK,....)
10:00 - 10:45	L	Variant calling: analytical sensitivity and specificity of genomic tests
10:45 - 11:00		Coffee break
11:00 - 12:30	L/P	Design of studies and clinical trials and how this affects the outcomes - Class project overview
12:30 - 13:30		Lunch
13:30 - 15:30	P	Variant calling practical: generate and work with VCF files
15:30 - 15:45		Coffee break
15:45 - 16:45	L	Filtering strategies for variants, in the context of clinical data, and using control data sets
16:45 - 17:30	P	Variant calling practical: generate and work with VCF files (cont.)
17:30 - 18:00		Q&A

Day 3

9:00 - 9:45	L	Variant annotation with established databases: How do we use of multiple database, in silico tools and literature sources for intepreting the meaning of a variant and evaluate its pathogenicity
9:45 - 10:45	P	Variant annotation/interpretation practical
10:45 - 11:00		Coffee break
11:00 - 12:00	P	Variant annotation/interpretation practical (continued)
12:00 - 13:00		Lunch
13:00 - 13:45	L	Integration with clinical information (e.g. electronic health records)
13:45 - 14:30	L	Principles of biomedical ontologies (e.g. HPO, SNOMED, ICD) and their use for the annotation of GE clinical phenotypes, importance of semantic interoperability for data integration
14:30 - 14:45		Coffee break
14:45 - 16:30	P	More use cases for variant interpretation, based on integration with clinical information. Explore GE clinical phenotypes
16:30 - 17:30	L	Best-practice guidelines for reporting the clinical significance of analytical results
17:30 - 18:00		UV process

Day 4

9:00 - 9:45	L	Increasing the depth of prognistic and diagnistic information: correlation with molecular signatures data and GWAS data
9:45 - 10:45	P	Practical on GWAS catalogue (and perhaps other resources, perhaps not feasible)
10:45 - 11:00		Coffee break
11:00 - 12:00	L	Principles of downstream functional analysis: Phenotyping in mouse (IMPC) or equivalent and cellular models
12:00 - 13:00		Lunch
13:00 - 14:00	P	Working with IMPC data
		Clinical decision support technology and clinical interpretation services
		Practice in examples of analysis of genomic data in the Training Embassy within the Genomics England Data Centre

D Day!

am	Preparation of projects
pm	Presentation of projects