PRAGUE SUMMER SCHOOL
NEXT-GEN SEQ DATA ANALYSIS

Quality Control, Read Mapping, Visualization and Downstream Analyses

7. - 11. 9. 2015

Conference Centre, Institute of Molecular Genetics, Prague, Czech Republic

Who Should Attend

Molecular biologists or data analysts with an interest in Next-Generation Sequencing (NGS) data analysis. You should be computer literate and have a basic understanding of molecular biology (DNA, RNA, gene expression, PCR).

5 Days of Hands-On Training

Scope and Topics

The goal of this workshop is to give you a deeper understanding of the Next-Generation Sequencing technology with a special focus on bioinformatics issues. All workshop participants will perform important steps of NGS data analysis tasks themselves.

Day 1 - Linux for bioinformatics

Go beyond the graphical user interface!

- Introduction to essential tools and file formats required for NGS data analysis assuring that all participants are able to follow the practical parts
- Command line and important commands
- Combining commands by piping and redirection
- Bioinformatics file formats (e.g. FASTA, BED, VCF, WIG) and databases (e.g. UCSC, ENSEMBL)
- Important bioinformatics toolkits (BEDtools, UCSCtools)

Day 2 to 3 - NGS data analysis

Get a full picture of NGS workflow!

- Advantages and disadvantages of current sequencing technologies and their implications on data analysis, NGS file formats and hands-on analyses
- Introduction to sequencing technologies from a data analysts view
- Raw sequence files (FASTQ format)
- Preprocessing of raw reads: quality control (FastQC), adapter clipping, quality trimming
- Introduction to read mapping (Alignment methods, Mapping heuristics)
- Read mapping (BWA, Bowtie2, STAR, segemehl)
- Mapping output (SAM/BAM format)
- Usage of important NGS toolkits (samtools, BEDtools)
- Mapping statistics
- Visualization of mapped reads (IGV, UCSC)

Day 4 to 5 – Getting started with RNA-seq Data Analyses

Explore the open source world of RNA-seq!

- Understanding split-read mapping
- Running different split-read mappers (tophat, segemehl, STAR)
- Understanding the Tuxedo Suite (cufflinks, cuffcompare, cuffmerge, cuffdiff, etc.)
- Prediction of new transcripts/isoforms using cufflinks/cuffmerge
- Quantifying exons/genes/transcripts
- Prediction of
  - Differential exon usage using DEXseq
  - Differential gene expression using DEseq
  - Differential isoform expression using cuffdiff
- Prediction of non-standard transcripts (circularized RNAs and/or fusion transcripts)

Every day – A brief presentation on real-life examples of data analysis given by bench scientists!

Seats and Fees

25 Seats available, First-come – First-served
Registration: www.seqme.eu/courses
Registration fee: 1.450 Euro plus VAT

Registration fee includes: Workshop materials, Lunches/Coffee breaks, Workshop dinner. Please kindly notice that lodging, travel and other incidental expenses are the responsibility of the attendee.

Organizers

ecSeq Bioinformatics solution provider with solid expertise in the analysis of high-throughput sequencing data, Leipzig, Germany.

SEQme Service provider in the field of DNA sequencing and Real-Time PCR, Dobris, Czech Republic.

The course will be done using a bootable USB stick containing a Linux environment with all needed NGS tools already installed! After the course you will get the stick with all exercises and results you created during the week. Thus, you can immediately apply what you have learned with your own data by just booting your machine from the stick! Please kindly notice no labwork is to be performed during the course.