Course Description

Introduction to RNA-Seq Data Analysis

Objective

The analysis of RNA-Seq data has become one of the dominant approaches in Biomedical research especially in gene expression studies due to its speed, low cost and ability to yield useful insights. It is also a highly data-oriented methodology which requires usage of a pipeline of software tools and this CAPOD funded course aims to teach participants how to process such data in an efficient manner.

Location and Date (NEW):
Swallowgate Computer Classroom, University of St Andrews, Fri ~~28th April~~ 12th May 2017.

Intended Audience
Postgraduates and staff starting to work, or currently working on, RNA-seq data and gene expression research projects.

Prerequisites:
Experience with command-line Unix/Linux is required. Good experience with R may also qualify. A preparatory course on the Unix command-line is available 8 days earlier. The link to this course is https://synergy.st-andrews.ac.uk/bioinformatics/course-registration/.

Content
This is a highly practical course which consists in taking some example RNA-Seq data through a pipeline of software tools leading to analysis of its differential gene expression.

Format of course
This is one day course, made up primarily of practical sessions. Laptops are not required as each participant will have a terminal at their disposal. There will be one lunch break and two coffee breaks.

Instructors
Joseph Ward and Ramon Fallon of the St Andrews Bioinformatics Unit. There may also be some assistance for the practicals from colleagues of the Scottish Oceans Institute.

Detailed Schema:

Time	Type	Contents
08:50 – 09:00	Registration	
09:00 – 09:30	Connection and access	* Nature of the data * Connect and setup workspace * Access to example data
09:30-10:30	Data quality	* Quality control with FastQC * Data preprocessing with FastqMcf
10:30-11:00	Break	
11:00-11:45	Mapping	* Mapping to a reference genome with Bowtie2 and Tophat * Visualisation of mapped reads with IGV
11:45-13:00	Mapping quality	* Assessment of mapping quality with samtools, picard-tools

		and RSeQC * Effects of different quality control approaches.
13:00-13:45	Lunch Break	
13:45-14:30	Gene quantification	* Estimating gene count with HTSeq * count file manipulation with awk
14:30-15:30	Practical	* Key aspects of using R * Differential gene expression with edgeR
15:30 – 16:00	Coffee Break	
16:00 – 17:00	Functional view	Functional Analysis using GSEA