Foundation Masterclasses 2020-2021

Primer for Advanced 'Omics' Approaches

optional for first- and second-year EASTBIO PhD students

Workshop leader: **Dr Simon Tomlinson** (University of Edinburgh, Centre for Regenerative Medicine)

Date: 3 May 2021, 10:00-15:00 (approximate times; with lunch break)

The course will be delivered online (via Blackboard Collaborate). To access it, please click on this link: https://eu.bbcollab.com/guest/44652be1429841a0ba747c548123c3d7

Workshop description:

The workshop aims to teach students how to carry out the analysis and interpretation of RNA-seq data. The course focused on RNA-seq data from mixed cell populations rather than single cell data, although the analysis concepts taught are common to both types of data. The course is taught from first principles and so is suitable for someone who is new to RNA-seq analysis. Basic knowledge of applied statistics and molecular biology is assumed although extra reading is provided for students who need to refresh their knowledge of these subjects. In the course, we make extensive use of R and Unix but no prior knowledge is required although basic computer literacy is assumed. We make use of a generic server on the course so that the skills taught can be applied easily outside the local environment.

The course will cover the skills necessary to analyse and interpret RNA-seq data. The typical student on the course may be interested in using skills gained on the course for analysis of data available online or analysing and interpreting their own lab data. The aim is to get students over the most difficult part of the learning curve and then provide a platform for further exploration of these techniques. We make use of several data sets on the course downloaded from public databases. The example data used will be mouse embryonic stem cell data but the skills taught are directly applicable to other biological domains.

Learning outcomes:

At the end of the course students will:

- be familiar with Unix and R and how to access resources hosted on remote servers.
- be able to perform an RNA-seq analysis and interpret the analysis outcomes.
- learn how to use the most commonly used analysis methods; they will gain an understanding of the underlying concepts so they will be ready to adapt their knowledge to related technologies not explicitly covered by the course.

Training Schedule (materials circulated in advance of the masterclass):

- EB_Lecture12_intro_setup- Welcome and getting started for the course
- EB_Tut1_Introduction_to_R_Unix.pdf- Practical introduction to R and Unix
- EB_Tut2_ExploitingConfiguration & RStudio.pdf - Introduction to exploiting R for analysis and R Studio
- EB_Lecture34_RNASeq- RNA-seq lecture covering the background needed
- EB_Tut3_Advanced_Analysis_RNASeq_Using R- Complex practical RNA-seq tutorial

Accounts for accessing the servers will be issued on the day. A detailed guide for setup is also provided as reference (EB_detailed_server_connection_guide.pdf).

This training works like a 'flipped' online class, where most materials can be completed before the day of the training. When the class runs, there will be real-time sessions for the analysis sessions (3 hours), plus the end tutorial (1 hour). Students can complete the analysis sessions before the day of the class but must attend the online tutorial on the day.

There will be several online sessions using the analysis servers, aimed at introducing R and Unix and completing an advanced RNA-seq tutorial. Pre-recorded lectures and suggested reading will be provided to provide the background information. The day will end with a tutorial that everyone should attend online in real time, aimed at discussing and summarising the outcomes from the day. The only part of the class that must take place in real time will be the tutorial; the rest of the course does not have to be completed in real time.

In sum, students registered for this training should commit a day of work time overall, as below:

- 2-3 hours online lectures (completed in advance of the day of the course)
- 3 hours of online analysis tutorials (completed in advance or in real-time online on the day of the course)
- 1-hour online tutorial (in real-time on the day of the course)

Training materials will all be recorded and available online. Tutorials will not be recorded, although a written summary of the class outcomes will be made available online.

Participation requirements:

The course will consist of pre-recorded lectures, tutorials online and interactive tutorial sessions. You will be provided with a login account for the analysis server. We will use SSH software such as MobaXterm to provide an SSH terminal for command line interaction with the server. We will also use RStudio web software for more advanced interaction with the server. Tutorials will use web conferencing software (either Teams or Zoom depending upon the class size). There is a small amount of setup required so that machines can communicate with the server- information on this will be provided before the class and during the first session. The course makes extensive use of open source software or otherwise software that can be used without the purchase of additional licences.

Background reading

A good textbook for Data Analysis using R (free to download &use).
Bioinformatics and Computational Biology Solutions Using R and Bioconductor
Editors: Robert Gentleman, Vincent J. Carey, Wolfgang Huber, Rafael A. Irizarry, Sandrine Dudoit
ISBN: 978-0-387-25146-2 (Print) 978-0-387-29362-2 (Online)
(https://link-springer-com.ezproxy.is.ed.ac.uk/book/10.1007/0-387-29362-0)

A good general article covering many of the R technologies discussed on the course...
http://manuals.bioinformatics.ucr.edu/home/R_BioCondManual

This is quite an accessible review for basic molecular biology
https://www.sciencedirect.com/science/article/pii/B9780080918655000035?via%3Dihub

This is a good RNA-seq review
https://www.annualreviews.org/doi/abs/10.1146/annurev-biodatasci-072018-021255

Useful extra materials for learning R
https://rstudio.com/resources/cheatsheets/

Training web-page: http://www.eastscotbiodtp.ac.uk/foundation-masterclasses

For **further info**, please email enquiries@eastscotbiodtp.ac.uk