UK-India Cancer Bioinformatics Workshop

28th – 31st October 2018

This 3-day workshop is for research scientists, PhD students and postdoctoral researchers who want to strengthen their skills in cancer bioinformatics. The selected 20 participants will learn genomic and transcriptomic techniques to analyse next-generation sequencing (NGS) data in a Unix environment. Participants are expected to have basic familiarity with Unix/Linux and R to participate in this course. The course will provide real-life example data; however, participants are encouraged to bring and discuss their own NGS data.

If selected, participants are expected to bring their own laptop. The minimum recommended specifications are as follows: 1024x768 screen resolution, 1.5GHz CPU, 2GB RAM, 10GB free disk space, and a recent version of Windows, Mac OS X or Linux operating systems. Participants are expected to bring an A1 portrait poster showcasing their current ongoing research. The posters are expected to act as a talking point between participants and the trainers on the course.

The workshop sessions are complemented by 5 faculty presentations, including national and international speakers and a keynote presentation.

Local travel (by train) will be covered. Accommodation for all participants will be provided at the ACTREC guest-house on a twin sharing basis.

Venue: ACTREC, Tata Memorial Center, Navi Mumbai

Day 0: Pre-meeting reception on Sunday 28th October 2018, 5pm

Venue: ACTREC, Tata Memorial Center, Navi Mumbai

Informal reception with faculty, teaching assistants and participants.

Introduction of participants.

Day 1 – Module 1: Whole-exome sequencing

Coordinator: **Dr Amit Dutt, Principal Investigator, ACTREC (India)**
Lead Persons: Dr Pratik Chandrani, Sanket Desai, Hitesh Kore
Date: 29th October 2018 (Monday)

Agenda: Participants will be introduced to whole-exome sequencing (WXS) data, common file formats used to store the raw and processed data, and tools used in variant analysis. The session is designed to give participants hands-on experience with the processing of WXS data, including quality control, alignment, variant calling, functional prediction and variant prioritisation.

08:15 – 09:00	**Faculty presentation 1:** - Speaker to be announced
09:00 – 09:15	**Computer setup** - Ensure all participants have access to the appropriate data and tools
09:15 – 10:00	**Workshop session 1: Introduction to NGS data, formats and tools** - Reference genome versions/ builds - FASTQ, SAM/BAM, VCF, MAF file formats - Overview of the computational infrastructure, NGS data and tool locations on server - Exome analysis pipeline overview
10:00 – 11:00	**Workshop session 2: Quality control and alignment** - Indexing reference genome - Pre-alignment quality control - Alignment using BWA - Post-alignment quality control
11:00 – 11:30	**Coffee break**
11:30 – 13:15	**Workshop session 3: Post-alignment processing and variant calling** - SAM/BAM processing, including conversion, sorting and indexing - Indel realignment and base quality recalibration - Variant calling using GATK
13:15 – 14:15	**Lunch**
14:15 – 15:45	**Workshop session 4: Variant filtering and annotation** - Filtering of germline variants using paired normal and SNP databases, including dbSNP, ExAC and TMC-SNPdb - Variant annotation using Oncotator
15:45 – 16:15	**Coffee break**
16:15 – 17:00	**Faculty presentation 2:** - Speaker to be announced

17:00 – 18:00	**Workshop Session 5: Functional prediction, variant evaluation and visualisation** - Functional scoring and prioritisation of variants using dbNSFP - Statistical evaluation of somatic variants - Visualisation using IGV
18:00 – 18:30	**Session summary and discussion**
18:30 – 20:00	**Poster session and opportunity to discuss participants own data**
20:00 Onwards	***Dinner at ACTREC lawns for all participants, speakers, tutors and guests***

Day 2 – Module 2: RNA sequencing

Coordinator: **Dr Anita Grigoriadis, Cancer Bioinformatics (King's College London, UK)**
Lead Persons: Jelmar Quist, Tom Hardiman
Date: 30th October 2018 (Tuesday)

Agenda: Participants will be introduced to RNA sequencing analysis. The sessions during this module will introduce different tools and methods available for quality control, alignment and differential gene expression analysis.

08:15 – 09:00	**Faculty presentation 3:** - Speaker to be announced
09:00 – 09:15	**Computer setup** - Ensure all participants have access to the appropriate data and tools
09:15 – 11:15	**Workshop session 6: Introduction to RNA sequencing quality control and alignment** - Overview of methods - Introduction of conventional- and pseudo-alignment - Pre-alignment quality control - Alignment, including conventional- and pseudo-alignment methods
11:15 – 11:45	**Coffee break**
11:45 – 13:15	**Workshop session 7: Post-alignment quality control and quantification** - Post-alignment quality control - Quantification
13:15 – 14:15	**Lunch**
14:15 – 16:00	**Workshop session 8: Introduction to differential gene expression and pathway enrichment analysis** - Overview of methods used in translational research - Perform differential expression analysis - Interpretation of differential gene expression results - Pathway enrichment analysis
16:00 – 16:30	**Coffee break**
16:30 – 17:15	**Faculty presentation 4:** - Speaker to be announced
17:15 – 17:45	**Session summary and discussion**
17:45 – 20:00	**Poster session and opportunity to discuss participants own data**
20:00 Onwards	***Dinner at ACTREC lawns for all participants, speakers, tutors and guests***

Day 3 – Module 3: Copy number from DNA sequencing

Coordinator: **Dr Anita Grigoriadis, Cancer Bioinformatics (King's College London, UK)**
Lead Persons: Jelmar Quist, Tom Hardiman
Date: 31st October 2018 (Wednesday)

Agenda: Participants will be introduced to the integration of various data types, both cross-platform and within-platform integration

08:15 – 09:00	**Keynote Speaker presentation:** - Speaker to be announced
09:00 – 09:15	**Computer setup** - Ensure all participants have access to the appropriate data and tools
09:15 – 11:15	**Workshop session 9: Introduction to copy number from DNA sequencing** - Introduction of methods for copy number calling
11:15 – 11:45	**Coffee break**
11:45 – 13:15	**Workshop session 10: Integration of gene expression and copy number** - Basic method of integrating gene expression and copy number data, interpretation and visualisation
13:15 – 14:15	**Lunch**

Day 3 – Module 4: Clonality

Coordinator: **Dr Anita Grigoriadis, Cancer Bioinformatics (King's College London, UK)**
Lead Persons: Jelmar Quist, Tom Hardiman
Date: 31st October 2018 (Wednesday)

Agenda: Participants will be introduced to the concept of clonality using NGS data.

14:15 – 16:00	**Workshop session 11: Introduction to clonality** - Concept of clonality - Inspecting allele frequency and combining with copy number data
16:00 – 16:30	**Coffee break**
16:30 – 17:30	**Workshop session 12: Integration methods for clonal evolution** - Integration of single nucleotide variants with copy number data
17:30 – 17:45	**Session summary and discussion**
17:45 – 18:45	**Faculty presentation 5**
18:45 – 19:15	**Closing remarks and award presentation**
20:00 Onwards	***Gala dinner***