Command-line omics, Linux and Python

Description:

It aims to equip participants with the theory and technical skills to handle the essential data processing stage of omic experiments. This stage, though typically more repetitive, easier to outsource and less time consuming to perform than the R based data analysis, it is still an important final step on the way to becoming a junior bioinformatician.

It covers genomics, epigenomics, and transcriptomics, the Linux operating system (Ubuntu), shell scripting, types of omic files (e.g. fastq, bed, sam, bam, vcf), omic tools (e.g. sam tools, bed tools, bowtie, star, macs, deseq2, sva) and how to write pipelines.

It is highly recommended to those planning on doing genomics or epigenetics, and those that wish to understand the more gritty technical aspects of bioinformatic projects.

Learning objectives:

- Experience of the Linux environment.
- Knowledge and practical experience in command line coding, including using tools such as cat, awk, paste and grep.
- Knowledge and practical experience of writing executable pipelines, including custom input options.
- To gain extensive knowledge and experience of common command line bioinformatic tools
- Knowledge of omic file types, especially precursors.
- Practical experience of writing bespoke pipelines for genomics, transcriptomics and epigenetics.

Course outline:

Day 1: Introduction to Linux and Ubuntu
Day 2: Command line coding
Day 3: Genomes and Coordinates
Day 4: Alignment
Day 5: SNP calling
Day 6: Phylogenetics
Day 7: Splice aware alignment
Day 8: Deseq2 and correction
Day 9: Chip & ATAC
Day 10: Python