A logo with a letter m

Description automatically generated

**Introduction to RNA-seq**

Date and Time: June 27 9H00-13H00

Location: TBD

**Workshop Lead:** Adrien Osakwe **Facilitator:** Jeffrey Yu **Registration link:** [**https://involvement.mcgill.ca/event/264011**](https://involvement.mcgill.ca/event/264011)

**Approximate duration: 4 hours**

**Prerequisites:**

1. **Experience with R**
2. Basic understanding of UNIX and the command line

**Summary: (2-3 sentences summarizing the workshop)**

In this 4-hour workshop, participants will be introduced to the basics of RNA sequencing. The workshop will first cover the biological basis of RNA-seq and the relevant data generating platforms. We will review the basic preprocessing steps to go from raw reads to count data. Key analysis processes including quality control, clustering, differential expression, and gene set enrichment analysis will then be covered. Finally, participants will be introduced to the benefits of single cell technology and how they enable researchers to overcome obstacles found in bulk RNA-seq analyses.

**Learning Outcomes: (List 3-5 learning outcomes participants will learn upon completion of this workshop)**

1. Understand the data generation steps for RNA-seq data.
2. Run basic RNA-seq analyses.
3. Understand the outputs of different RNA-seq analysis tools.

**Content**

1. **Introduction (5 mins)**
   1. What is RNA-seq
   2. Summary of protocol
2. **Protocol & Experimental Design (20 mins)** 
   1. Illumina-based sequencing Protocol (standard components)
   2. Library Size + Number of replicates (power analysis slide)
   3. Read length
   4. Extraction Protocols
3. **Quality Control (25 min) – Galaxy for hands-on**
   1. Read Quality & Alignment Quality
   2. Quantification & Reproducibility
   3. Transcript Identification + Alignment
   4. Transcript Discovery + De novo alignment
   5. Transcript Quantification
4. **Break - mins**
5. **Analysis (1h50) – Google Colab**
   1. EDA vs. Actual Analysis
   2. Highly variable genes
   3. Sample clustering (covariates)
   4. **Differential Expression**
      1. **Standard Results (lfc,p-value,mean-abundance plots)**
      2. **Covariates**
   5. **Gene-Set Enrichment**
6. **Break – 10 mins**
7. **Single cell Data (30 mins) – Google Colab**
   1. Bulk vs. Single Cell
   2. 10X vs smart-seq
   3. **Introduction to Seurat / SingleCellExperiment**
8. **Closing Remarks (10 mins)**
   1. Combining with other genomic assays
   2. Conclusions