RNA-sequencing for "bulk" transcriptomics

Session Intro

The session looks at RNA-sequencing (aka RNA-Seq) and its application to "bulk" gene expression analysis. Different from previous sessions where a specific paper was chosen for in-depth discussion, we switch gears: This session is an introduction to the topic, to prepare for the "project" and for a later, planned, session on single-cell RNA-seq transcriptomics.

Session Plan

As by now you (the students in this class) should have already developed some familiarity with selecting relevant information, and reading and presenting papers, let me get you to teach each other about RNA-Seq. I have selected some specific pertinent topics and assign teams to present them, as below.

I leave each presenting team to decide on what they want to talk about (i.e., it is perfectly ok to leave out some topics/points/details and/or include other topics/points/details.) Also, the presenting team need not just make presentations; they are encouraged to figure out how to engender more class interactions and lead discussions.

Part I, What RNA-Seq is:

This part deals with background knowledge of RNA-Seq. Some suggestions on things to present:

- Overview of how RNA-Seq works and its applications

- Details of some RNA-Seq technologies such as illumina short-read sequencing, PacBio long-read sequencing, and Oxford Nanopore direct RNA sequencing.

Presentation team #9: TANG KAIWEN, NGUYEN HA LINH

Total time limit: 20 minutes (presentation) + 5 minutes (audience questions.) Total slide count: 12 slides max.

Part II, Algorithms/tools for read mapping, transcript assembly and quantification for RNA-Seq.

After reads are produced, they must be mapped to a reference genome. And after mapping, the transcripts must be assembled and quantified to determine gene expression levels. This part presents some methods/tools for read mapping and some ways for transcript quantification. Some suggestions on things to present:

- Describe mapping & alignment methods like BowTie, TopHat, Picard.

- Describe quantification measures such as FPKM, RPKM, TPM, and normalized counts like TMM, and tools such as Cufflinks. It may be worth summarizing these two papers intuitively, https://doi.org/10.1186/s12967-021-02936-w and https://doi.org/10.1186%2Fgb-2010-11-3-r25

Presentation team #7: FU GUOJI, LUO YANG, PRABOWO DJONATAN

Total time limit: 35 minutes (presentation) + 5 minutes (audience questions.) Total slide count: 12 slides max.

Part III, Differential expression analysis

This part discusses differential analysis of gene expression using RNA-Seq data.

- Overview of differential gene expression analysis

- Negative bionomial distribution

- DESeq/DESeq2, https://doi.org/10.1186/gb-2010-11-10-r106

- edgeR, https://doi.org/10.1093%2Fbioinformatics%2Fbtp616

Presentation team #4: DUAN KEYU, LU XINYANG, STEFAN PUTRA LIONAR

Total time limit: 35 minutes (presentation) + 5 minutes (audience questions.) Total slide count: 12 slides max.