Cancer Center
NCI-DESIGNATED COMPREHENSIVE CANCER CENTER

Bulk RNA-seq Planning and Best Practices

General considerations to increase the likelihood of high quality data

Consult Analyst
Engage with an analyst before beginning the experiment

Discuss goals available samples, batches, etc.

Define Experimental Variables
Define experimental question and design experiments to properly test while avoiding confounders

Confirm your experimental groups (e.g. positive controls, genetic change)

Prevent Batch Effects
Do not introduce technical variables that cannot be controlled
(e.g. do not process your control and experimental groups on separate days or slides)

Randomize Sample Order
Minimize batch biases in large experiments by alternating processing of experimental and control group samples

Conduct Pilot Studies
Perform preliminary sequencing to confirm quality and positive/ negative controls

Pilot studies can save money!

Consult Analyst
Review raw data and results

Discuss data management plan

Consult Data Generating Facility
Ensure compliance with sample input policies and understand the deliverables for the data generating facility (e.g. buffers, concentrations)
Use Genomics.Core@CUAnschutz.edu for the Genomics Shared Resource

Ensure Sufficient Sample Size
Verify adequate sample numbers for analysis

Avoid Sample Pooling
Pooling can introduce unwanted variability and mask variability across replicates

Collect Metadata
Gather necessary sample information (e.g. sex, time point, batch)

Perform Quality Control
Conduct quality checks throughout data generation

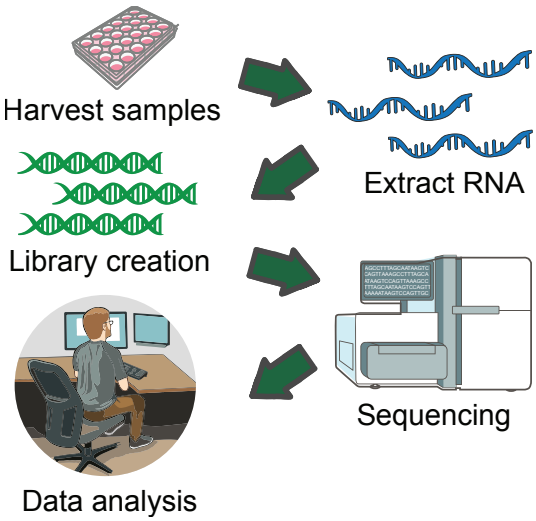
Go/no-go decisions

RNA-seq will capture and sequence the transcriptome

RNA and library kit selection
Selection of poly-adenylated RNA is simple and reliable, but excludes non-poly-A RNA molecules

Ribosomal RNA depletion is more prone to incomplete rRNA removal but enables detection of both poly-A and non-poly-A RNA

A low-RNA input kit might be necessary but they require more PCR cycles and typically result in higher duplication rates, more rRNA, and potentially other quality issues



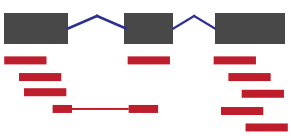
Analysis
Many modalities can be measured such as gene expression, isoforms, splice variants

mRNA quantification does not always correlate with protein

Common analyses include differential gene expression and pathway analysis across groups

The BBSR has user-friendy analysis apps
Contact the BBSR to get started

How to design a RNA-seq experiment?



Review the general considerations (above)

Replicates
≥3 biological reps (more for *in vivo*)

Technical replicates generally aren't necessary - consult an analyst

Cell line replicates should be as independent as possible (e.g. different dishes, days, thaws, etc.)

Coverage depth
Typically 20-40M clusters/sample

More reads may be needed to detect low-abundance transcripts or isoforms

Quality control

Material evaluation
Assess concentration and RIN

Library evaluation
Check size and quality (e.g. bioanalyzer, tapestation)

Data evaluation
Check QC metrics (e.g. depth, contamination, sample variability)

Decide on sequencing
Evaluate material and library QC and seek expert consultation
Go/no-go decision

Contact the Biostatistics & Bioinformatics Shared Resource (BBSR) for cancer-related analysis

James.Costello@CUAnschutz.edu, Bioinformatics@CUAnschutz.edu

Contact CUBE for all non-cancer-related projects