

# Introduction to RNA-sequencing Experimental Design and Analysis (CB321)

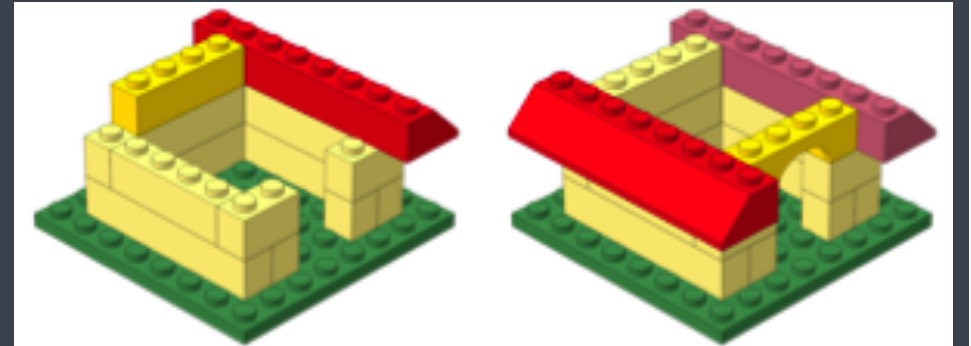
Sarah Boswell, PhD  
Radhika Khetani, PhD

<https://hbctraining.github.io/rnaseq-cb321/>

# Transcriptomics (RNA-seq)

- RNA-seq is the process of sequencing the “transcriptome”
- Uses include:
  - Differential Gene Expression
  - Transcriptional profiling with single cells
  - Transcriptome assembly
  - Metatranscriptomics or community transcriptome analysis
  - Mechanisms of transcription looking at nascent transcripts
- Specific use of RNA-seq will determine experimental protocols, sequencing strategies, and analysis methods

# Learning Objectives



- ✓ Describe sequencing by synthesis, including cDNA library preparation.
- ✓ Describe the design considerations to perform an **RNA sequencing experiment**.
- ✓ Describe the workflow for a standard bulk or single-cell RNA-seq data analysis.
- ✓ List computational skills necessary to implement the various steps in the analysis workflow.

# Base components



## ✓ Unix/Linux shell

- Dealing with large data files (e.g. FASTQ, BAM)
- Using bioinformatics tools (e.g. Salmon, Trinity)
- Accessing and using high-performance compute clusters (O2)

## ✓ R

- Parsing and working with smaller results files
- Statistical analysis, e.g. differential expression analysis (DESeq2) or single-cell analysis (Seurat)
- Generating publication-quality figures from complex data (ggplot2)

# Workshop webpage

<https://hbctraining.github.io/rnaseq-cb321>

*These materials have been developed by Sarah Boswell and members of the teaching team at the Harvard Chan Bioinformatics Core (HBC). These are open access materials distributed under the terms of the Creative Commons Attribution license (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.*

