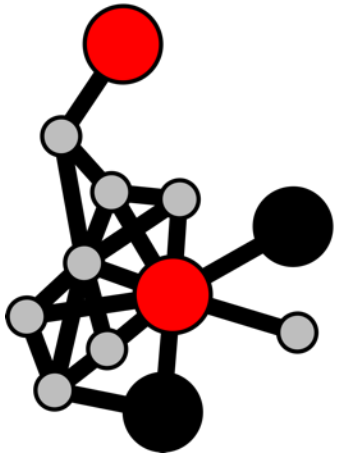
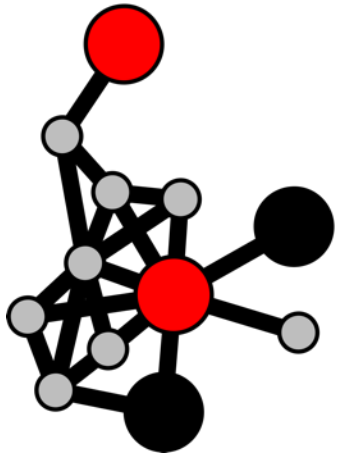
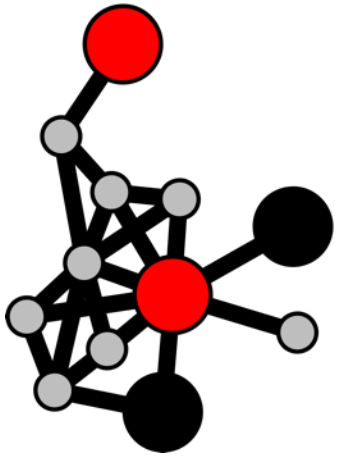
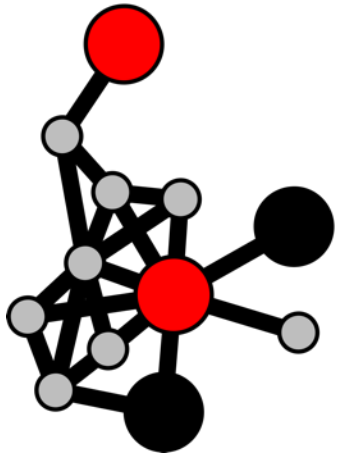




Transcriptomics lesson

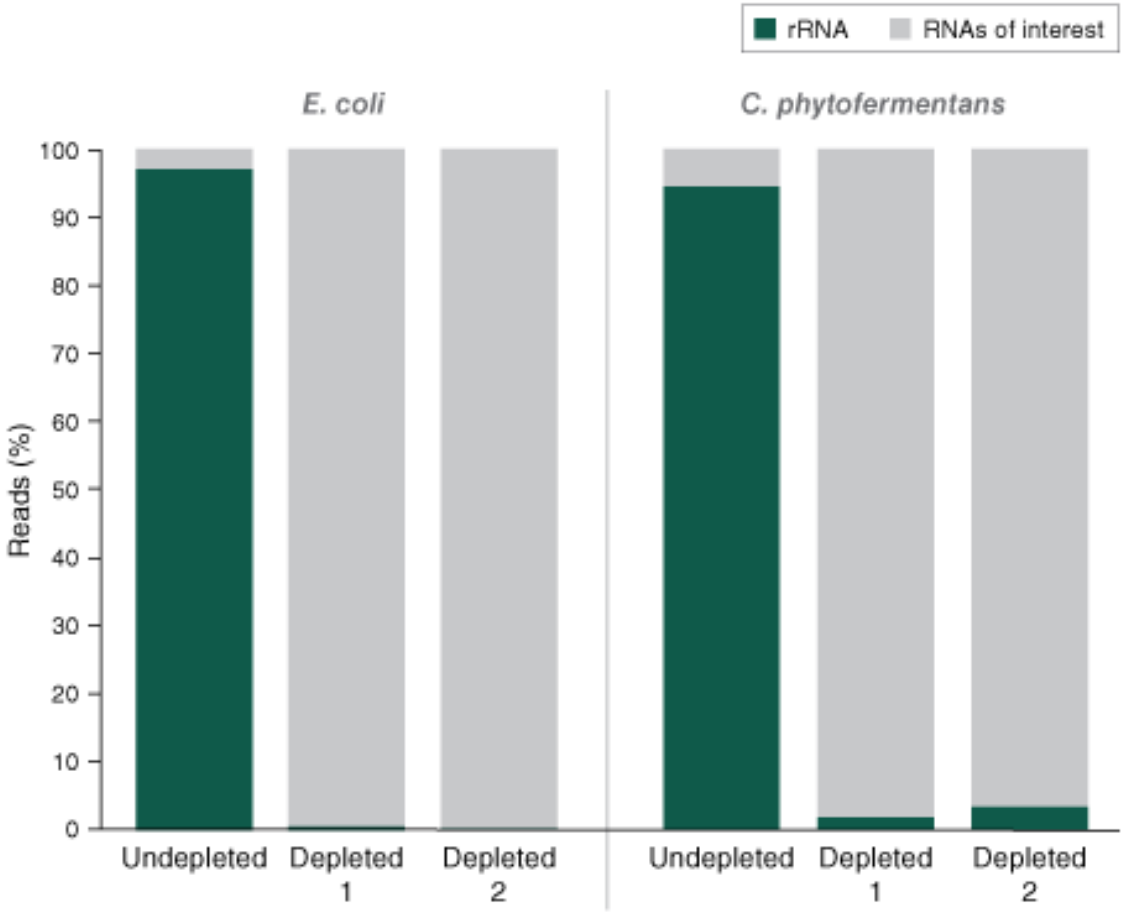




rRNA depletion:

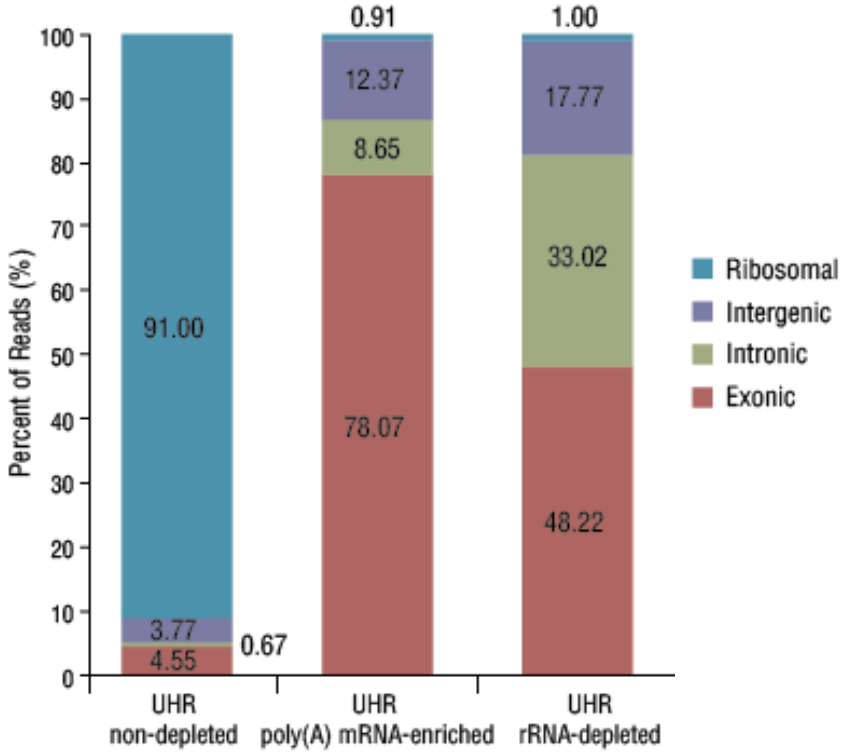
Wet-lab and in silico approaches

Why remove rRNA from you RNA samples?



New England Biolabs

Human RNA

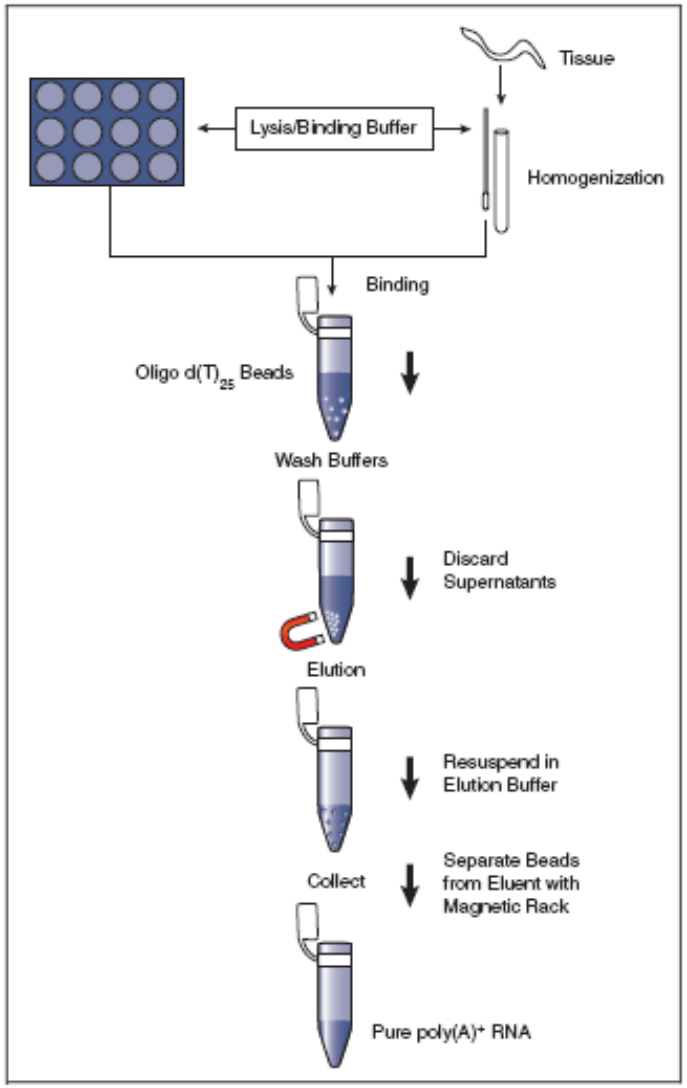


New England Biolabs

Why waste money sequencing rRNA, which makes up >90% of all RNA

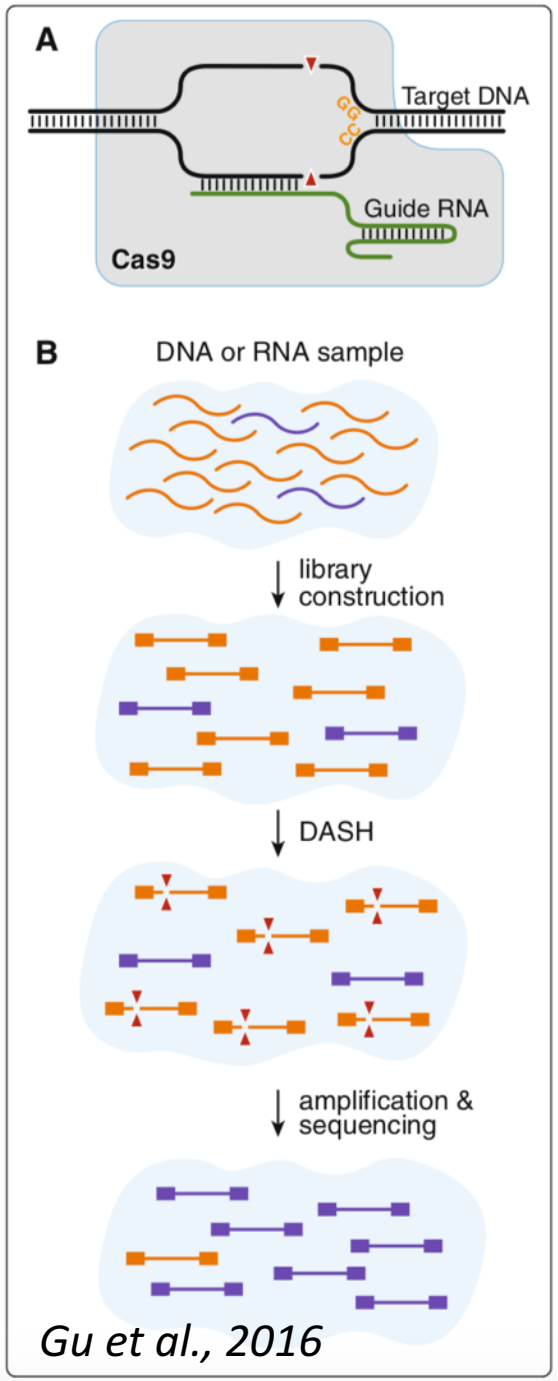
NEBNext rRNA depletion using RNase H

NEBNext mRNA enrichment using magnetic beads (for eukaryotic RNA only)



Depletion of Abundant Sequences by Hybridization (DASH): using Cas9 to remove unwanted high-abundance species in sequencing libraries and molecular counting applications

W. Gu[1†], E. D. Crawford[2,3†], B. D. O'Donovan[4], M. R. Wilson[5], E. D. Chow[6], H. Retallack[7] and J. L. DeRisi[2,3*]



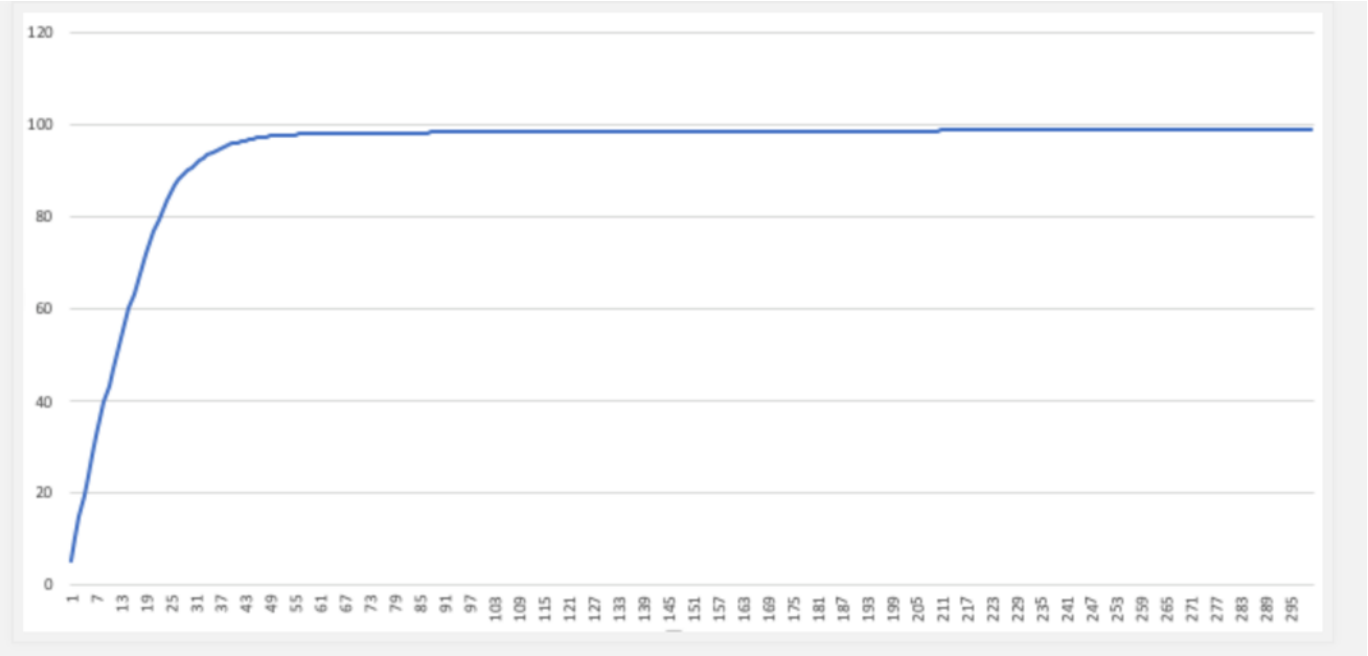
Gu et al., 2016

How does DASH work?

Shallow sequencing to identify abundant rRNA reads, and mRNA reads to avoid

"DASHit" software to identify "off-target" mRNA and "on-target" rRNA sites

Synthesize "on-target" guide RNAs to remove rRNA (or other highly abundant seqs) → **Works on cDNA!!!**



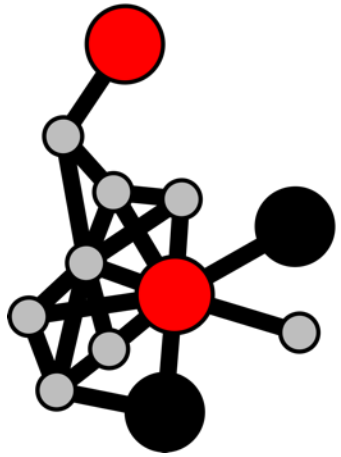
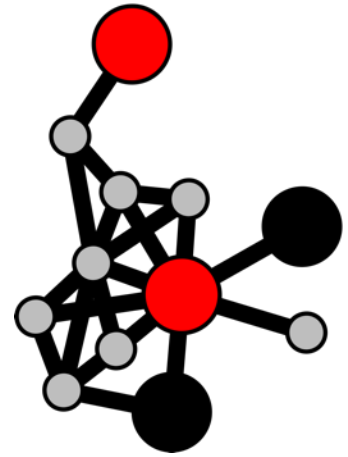
# of on-target sites

So you've completed RNA-seq, now what?

1. Your rRNA-depletion worked, but you still likely have some degree of rRNA in your reads. To avoid potential biases in mapping estimates, you should still remove the remaining rRNA reads.
2. Your rRNA-depletion did not work well for some reason, so you would like to try to salvage your experiment by removing all the rRNA reads from the data.

In silico rRNA removal using SortMeRNA

- Provide single-end or paired reads
- Recommended pipeline:
  1. Quality-trimming (e.g. Trimmomatic)
  2. Combine forward and reverse (e.g. Flash)
  3. **rRNA-depletion SortMeRNA**
- SortMeRNA uses a set of rRNA reference files
  1. Archaeal 16S
  2. Archaea 23S
  3. Bacterial 16S
  4. Bacterial 23S
  5. Eukaryotic 18S
  6. Eukaryotic 28S
  7. Universal 5S

Onto the Jupyter Binder tutorial!