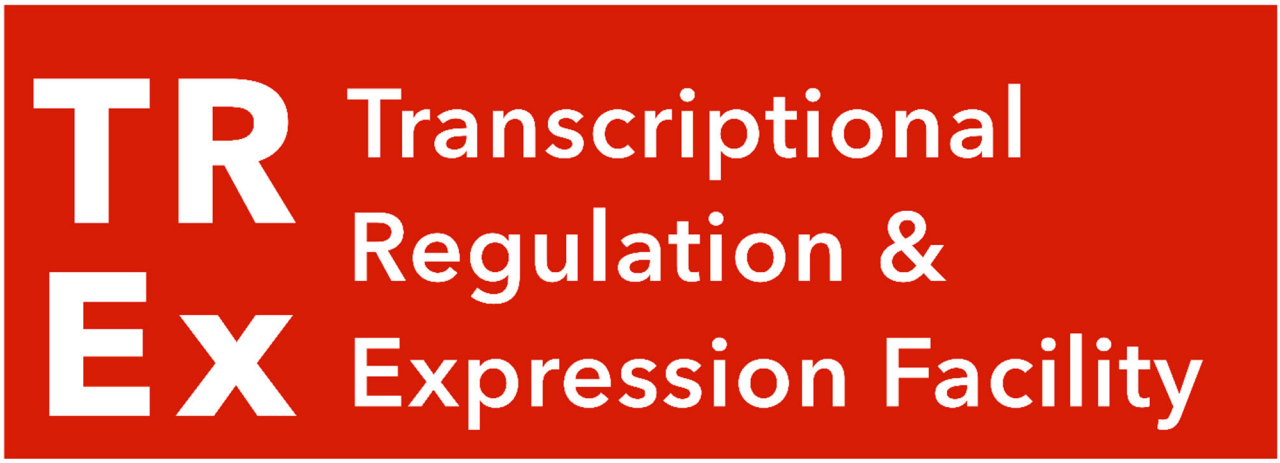
TREx Guide: rRNA Depletion Methods



Poly A Selection vs. Ribosomal Depletion:

How to choose the best method for enriching your RNA for RNAseq

Ribosomal RNA (rRNA) constitutes around 80% of the total RNA in mammalian cells, while protein coding RNA (mRNA) comprise less than 5%. Because mRNA is the typical focus for gene expression analysis, it is critical to pre-enrich RNA samples prior to sequencing. There are two main ways of doing this enrichment: **Poly A selection (polyA+)** and **Ribosomal depletion (rRNA-)**, and it is important to understand how these two methods work in order to choose which is right for your project.

Quick Guide

A quick reference for choosing which depletion method to use

We recommend **poly-A selection** (if compatible with your project) as we find it does a better job of enriching for protein-coding genes. While rRNA depletion is the best choice for some projects (e.g. degraded RNA, bacterial samples, pathogen detection), the resulting RNAseq dataset will have a lower percentage of reads that map to annotated coding genes due to the presence of both residual rRNA and other 'housekeeping' non-coding RNAs. Ribosomal depletion also is usually more expensive than poly-A selection.

Poly A Selection

- Your samples **all have an RQN greater than 7,** meaning your RNA is intact

  AND
- You are interested specifically in coding genes (mRNA)

  AND
- You are working in a species with polyadenylated mRNA (ex: plant, animal)

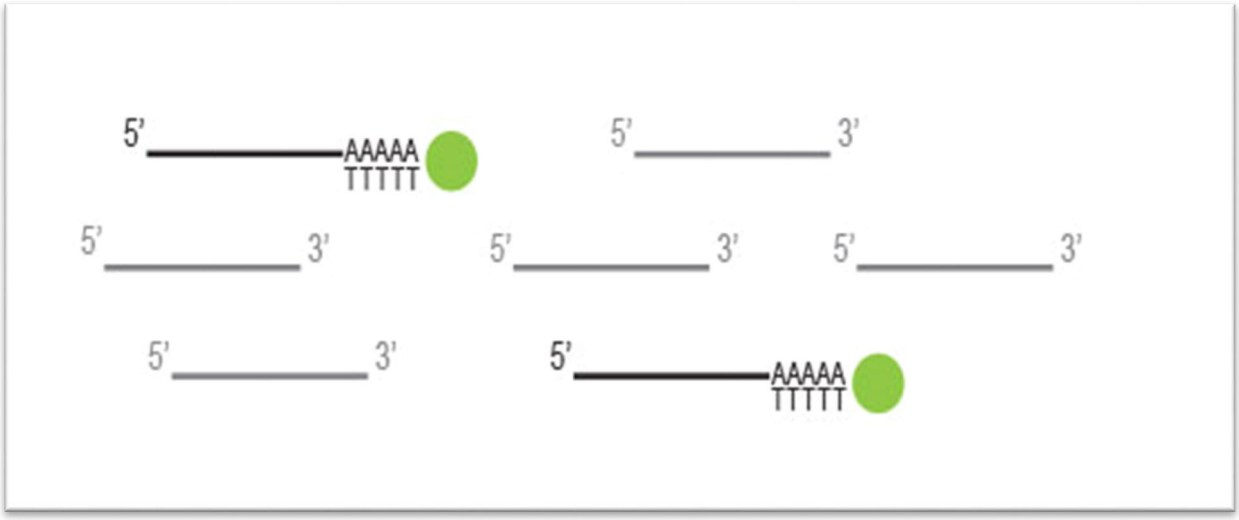
Ribosomal Depletion

- One or more of your samples have an **RQN less than 7**, meaning your RNA is degraded

  OR
- You are interested in non-polyadenylated RNA transcripts

  OR
- You are working in a species that doesn't polyadenylate their mRNA (ex: bacteria, virus)

Poly A Selection

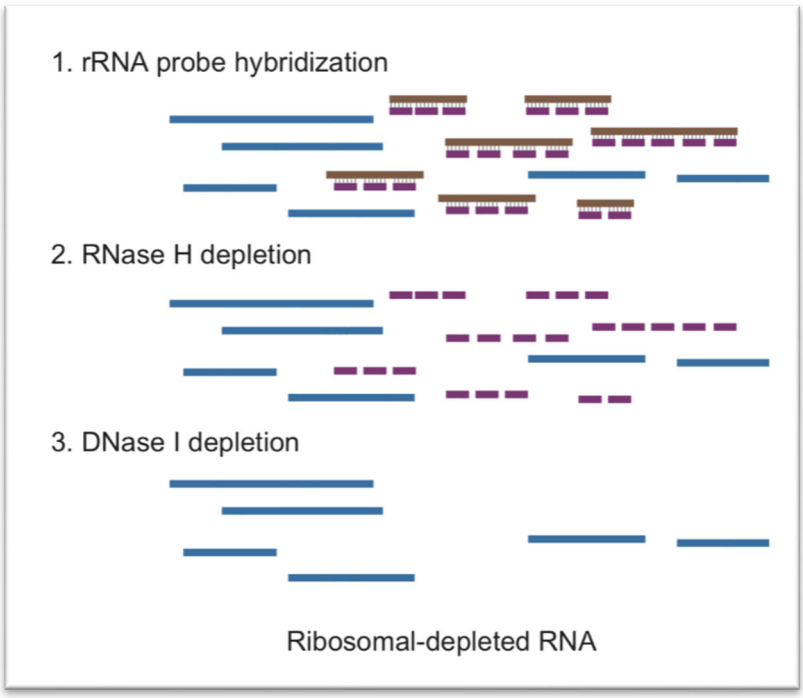


This method relies on the fact that eukaryotic messenger RNA is polyadenylated. Magnetic beads are used to capture mRNA containing a poly-A tail. Any RNA without a poly-A stretch is washed away. This method captures mRNA as well as other polyadenlyated non-coding RNAs (such as many lncRNAs).

In order to capture complete mRNA transcripts, an RNA sample must be intact. Any degradation or fragmentation of the RNA prior to the poly A selection will introduce a 3' bias into the RNAseq dataset. For this reason **TREx recommends using poly A selection only for projects where all RNA samples have an RQN higher than 7 on a Fragment Analyzer** (e.g. BRC Genomics Facility RNA QC results). Because other instruments such as the Bioanalyzer and Tapestation tend to give higher RIN values than the Fragment Analyzer RQN, we recommend a RIN cutoff of 8 for poly A selection.

Ribosomal Depletion (rRNA-)

There are many commercial kits that subtract ribosomal RNAs, however the principles behind them are often the same. First a panel of DNA probes bind to the ribosomal RNA (rRNA). Next the bound rRNA is degraded in order to remove it. Finally the DNA probes are depleted, with the goal to leave behind only RNA transcripts that are not ribosomal.



In order for the rRNA to be efficiently depleted, the rRNA probe set must be complementary to the rRNA from your organism of interest. This also means that if you are working in an organism distantly related to the species for which the rRNA depletion kit is designed, the rRNA depletion may be less efficient. Even when there is a good match between your species and the kit design, up to 10% of the reads in an RNAseq dataset can match rRNA; this percentage will be higher for non-cognate species.

Furthermore, rRNA-depleted samples still retain other 'housekeeping' transcripts that may comprise a significant amount of the final RNAseq dataset (e.g. snRNAs, snoRNAs, scaRNAs, Vault, RNAse P, and more).