J. Jedediah Smith

Week 9 Paper Questions

BIFX-504

10/20/2021

1. How can your RNA extraction technique affect the downstream results of your experiment? Be specific and consider what types of RNAs one might be studying. (2 points)

Silica gel extraction could effect the base-pair length of your RNAs. If too much ethanol is used, only the RNA < 200 bp will be retained. This could have downstream effects if you plan to focus on a certain length of RNA.

Phenol-chloroform extraction has an alcohol precipitation step that requires a salt. The type of salt used has effects on which types of RNA are retained. For example, use of lithium chloride is associated with a loss in tRNA, 5S rRNA, snRNA, and RNA < 250 – 300 bp. Salts should be chosen carefully based on which type of RNA is being studied.

1. RNA that is extracted from amniotic fluid is extracellular RNA and does not contain ribosomal RNA.
   1. Can it ever have a high RIN number? (1 point)

No, the RIN number calculation is determined based on the 28S and 18S bands of ribosomal RNA. With no ribosomal RNA, there would be no bands or RIN number.

* 1. How can one assess its quality? (1 point)

Quality of exRNA can still be assessed based on the number of reads overlap with other RNA of the genome and what proportion of the reads align with the genome.

1. Why is rRNA depleted prior to RNAseq library preparation? (1 point)

Typically, rRNA is 80 – 85% of the total extracted RNA pool. This is a problem if the thing you wish to focus on is protein-coding mRNA. To increase the quantification of mRNA, it is therefore necessary to reduce the amount of excess rRNA.

1. Does rRNA need to be depleted if only polyadenylated transcripts (mostly mRNA) will be sequenced? (1 point)

Not necessarily. One could use Poly-A-Containing mRNA Enrichment instead of rRNA Depletion. But this only works if you are interested in the polyadenylated mRNA, as the non-polyadenylated mRNA will not be captured with this method.

1. If you extract total RNA and then perform an rRNA depletion followed by a standard RNA-seq library preparation technique (as on page 7), will you be able to detect the following kinds of RNA? (Hint: remember the sizes and characteristics of the RNAs from week 6. Look back at the “What is functional and what is junk” review.) (6 points)
   1. Mature mRNA

Around ~1700 bp. More than 150 bp, should be detected.

* 1. hnRNA that includes introns

Around ~10,000 bp. More than 150 bp, should be detected.

* 1. tRNA

Around ~100 bp. Less than 150 bp, shouldn’t be detected.

* 1. circular RNA

Around ~500 bp. More than 150 bp, should be detected.

* 1. miRNA

Around ~20 bp. Less than 150 bp, shouldn’t be detected.

* 1. lncRNA

Around ~1000bp. More than 150 bp, should be detected.

1. If you extract total RNA and then select polyadenylated sequences before using the standard RNA-seq library preparation technique (as on page 7), will you be able to detect the following kinds of RNA? (Hint: remember the sizes of the RNAs from last week) (3 points)
   1. Mature mRNA

Around ~1700 bp. More than 150 bp and polyadenylated, should be detected.

* 1. hnRNA that includes introns

Around ~10,000 bp. More than 150 bp, but not polyadenylated, shouldn’t be detected.

* 1. tRNA

Around ~100 bp. Less than 150 bp and not polyadenylated, shouldn’t be detected.

* 1. circular RNA

Around ~500 bp. More than 150 bp, but not polyadenylated, shouldn’t be detected.

* 1. miRNA

Around ~20 bp. Less than 150 bp and not polyadenylated, shouldn’t be detected.

* 1. lncRNA

Around ~1000bp. More than 150 bp and polyadenylated, should be detected.

1. The paper talks about strand-specific sequencing for when you are studying prokaryotic genomes or interested in antisense transcripts, but why is it always a good idea to do strand-specific sequencing even if you are measure gene expression in a eukaryote? (2 points)

It helps you determine the exact location a gene signal is coming from. This is particularly helpful in determining whether the gene signal is coming from the same loci on both strands, or only one of the strands.

1. According to the paper, what four parameters should be taken into consideration when estimating the necessary sequencing read depth for an RNA-seq experiment? (2 points)

The four parameters for estimating sequencing read depth presented by the paper are: guidelines from the literature or references, the type of experiment and biological question, transcriptome size and complexity, and error rate of sequencing platform.

1. Is it more important to have a lot of technical replicates or biological replicates in an RNA-seq experiment? Why? (2 points)

Biological replicates are typically preferred over technical replicates because they allow one to have a better handle on true mean and variance of expression for the biological population of interest.

1. What are artificial RNA spike-ins used for? (1 point)

They are used to help accurately quantify absolute transcript concentrations. This is done by using the spike-ins to calibrate RNA concentrations in each sample to assess sensitivity, coverage, and linearity.

1. According to the paper, what four steps should you take in planning your experiment in order to avoid bias? (4 points)

The four important steps I would take to help avoid bias when planning my experiment are: identify the question of interest, attempt to identify possible sources of variability, plan the experiment in a way the reduces the effect of expected nuisance factors, and protect against unknown sources of variation.