RNA Sequencing Core

Center for Reproductive Genomics
Department of Biomedical Sciences
College of Veterinary Medicine
Cornell University

rnaseqcore.vet.cornell.edu

Jen Grenier
jgrenier@cornell.edu
RNA Sequencing Core
rnaseqcore.vet.cornell.edu

Next Generation Sequencing

Data Analysis

Testable Hypothesis

Experimental Design

Sample Preparation

Replicates
Input amount
Quality

NGS Sequencing

Plan ahead...

Instrument
Read length
Single/Paired

NGS Library Preparation

Library type
Barcoding

Compute power
Software/parameters
Reference genome
Data management

Research:Technology Partnership
RNA Sequencing Core
rnaseqcore.vet.cornell.edu

Services

• Transcriptome sequencing (RNAseq)
  total RNA $\rightarrow$ polyA$^+$ mRNA
  total RNA $\rightarrow$ rRNA-subtracted RNA
  viral RNA
  custom input

• Small RNA sequencing
  total RNA $\rightarrow$ microRNA, siRNA, piRNA

• All-inclusive Packages
  One price includes library prep, sequencing, and analysis
  Single point of contact
  Project management focus
RNA Sequencing Core
rnaseqcore.vet.cornell.edu

All-inclusive Project Services
Single Point of Contact

- Experimental design
- RNA sample preparation/QC
- Library preparation
  - barcoding
  - QC
  - pooling
- Illumina sequencing
- Standard analysis
  - data QC, preprocessing
  - genome/transcriptome mapping
  - normalization
  - differential expression
- Custom analysis
- Validation, follow-up

2017 Pricing
Cornell  CVM labs

**$370**  **$265**
NextSeq (75nt reads)
20M reads/library
Add’l 20M reads  **$115**  **$115**

**$100**  **$100**
rRNA removal:
Ribo-Zero HMR

**$340**  **$232**
Small RNA seq
10M reads/library
Add’l 10M reads  **$80**  **$80**
RNA Sequencing Core
rnaseqcore.vet.cornell.edu

RNAseq Decision Tree

RNA Quality
RQN ≥ 7

- yes
- no

polyA⁺ OK

- yes
- no

≥ 100 ng
total RNA

- yes
- no

Directional
(stranded)
RNAseq

≥ 100 ng

Standard
(nonstranded)
RNAseq

10 - 100 ng

HMR rRNA⁻
(stranded)
RNAseq

eukaryote

Other rRNA⁻
(stranded)
RNAseq

bacteria/mix

PolyA⁺ RNAseq
RQN ≥ 7 and polyA⁺ OK

rRNA⁻ RNAseq
RQN < 7 or polyA⁺ not OK

Illumina Instrument: NextSeq500 lowest cost and fastest turn-around time
• Experimental Design
  *plan ahead!*
  minimize batch effects, other variables
  test protocols, then be consistent

• Quality Control
  sample quality  *purity, integrity*
  library quality  *yield, insert size range*
  data quality    *read counts/quality, mapping rates, biases*
  biological signal  *sample clustering, DE genes*
Isolation Method

- **Trizol** - highest yield, requires phase extraction/precipitation
- **Silica spin column** - watch out for small RNA recovery

Concentration/Yield

- **Nanodrop (A260)** - less sensitive, but also gives curve/purity ratios
- **Qubit (fluorescence)** - more sensitive, use for samples <20ng/ul

Chemical Purity

- Nanodrop – *shape of curve (ratios) indicate salts/organics/protein*
  *most common impurity: carryover of lysis buffer components*

Biological Integrity

- ‘Bioanalyzer QC’ - *RIN/RQN*
  *Lysis step is critical*
# RNA Sequencing Core
rnaseqcore.vet.cornell.edu

## RNA Samples: Isolation and Storage

<table>
<thead>
<tr>
<th>Sample Collection</th>
<th>cryofreeze (dry)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>homogenize in lysis buffer/freeze or prep</td>
</tr>
<tr>
<td></td>
<td>RNAlater (fridge O/N)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lysis Conditions</th>
<th>rapid, at cellular level</th>
</tr>
</thead>
<tbody>
<tr>
<td>fresh:</td>
<td>homogenize in lysis buffer</td>
</tr>
<tr>
<td>cryofrozen:</td>
<td>grind on dry ice or homogenize in lysis buffer</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>sample handling</th>
<th>keep RNA samples on ice at all times</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>don’t vortex purified RNA (avoid shearing)</td>
</tr>
</tbody>
</table>

| storage                    | store at -80, minimize freeze-thaw cycles            |
RNA Sequencing Core  
rnaseqcore.vet.cornell.edu

Chemical Purity: Nanodrop Absorbance

Concentration: $\alpha$ A260 ($RNA\ extinction\ coeff = 40$)

Purity:
- A260/A230 1.5-2 low: chaotropic salt contamination
- A260/A280 1.8-2 low: protein (A280) or phenol (A270) contamination
RNA Sequencing Core
rnaseqcore.vet.cornell.edu

Biological Integrity: ‘Bioanalyzer’ QC

RQN ≥ 7 : polyA⁺ RNAseq

RQN 9.7

RQN 6.8

RQN < 7 : riboRNA⁻ RNAseq

RQN 4.9

RQN 1.2
Genomic DNA can be co-purified with RNA identified by RNA QC (high MW) or RNA-specific quantification (Qubit) vs nanodrop resolved by (RNAse-free) DNAses Tx.
RNA Sequencing Core  
[rnaseqcore.vet.cornell.edu](rnaseqcore.vet.cornell.edu)

## RNA Samples: Solving QC Problems

<table>
<thead>
<tr>
<th>Category</th>
<th>Solution Details</th>
</tr>
</thead>
</table>
| **Yield**                 | improve lysis efficiency  
try Trizol instead of column                                                        |
| **Chemical purity**       | clean up existing RNA sample  
repeat RNA isolation                                                                |
| **Biological integrity**  | change sample collection procedure  
improve lysis efficiency/speed  
subtract rRNA instead of enriching mRNA                                            |
RNA Sequencing Core
rnaseqcore.vet.cornell.edu

Data Analysis Pipeline

Raw data QC
  FastQC: base quality scores, overrepresented sequences

Preprocessing
  Trim adaptor sequences
  Filter low-quality, short, and contamination sequences

Genome mapping

Gene expression quantification
  Annotated genes (or microRNAs)

Additional analyses
  Hierarchical clustering
  Principal component analysis
  Gene set analysis  
    GO enrichment, pathways, GSEA
  novel transcript and miRNA prediction
  mRNA – microRNA interactions