RNA-Seq Workflow Steps and Examples

Adapted from Illumina Brochure
"Buyer's Guide: Simple RNA Sequencing Workflows"

Step 1. What scientific questions or applications am I interested in?

The first step in RNA sequencing begins with identifying the experimental question or application. What are you interested in studying?

Gene Expression Profiling – Understand and quantify the coding transcriptome.

Whole-Transcriptome Sequencing – Analyze both coding and noncoding transcripts.

Transcriptome Discovery – Identify novel features such as gene fusions, SNVs, splice junctions, and transcript isoforms.

Small RNA Sequencing – Study small RNA species such as miRNAs and other miRNAs with a 5'-phosphate and a 3'-hydroxyl group.

Step 2. What are my study design requirements?

The next step in the process involves understanding your study design needs. The list below describes common study design requirements that will impact your choice of workflow, from library prep to data analysis.

Study Information Required – Determine the kind of figures and tables you want to include in your publication or grant application.

Sample Type and Quality – Whether your samples are human, nonhuman, plant-based, or microbial. Are they low quality or degraded, such as from FFPE preservation?

Sample Abundance – Will your RNA be derived from cell culture, single cells, or another source? Different sample abundance levels require different library prep protocols (GMAK provides library prep options to all abundance levels).

Step 3. What factors will impact my RNA-Seq study cost?

A key consideration when designing an RNA-Seq study is the cost. The factors listed below can be used to calculate RNA-Seq study costs.

Study Size, Sample Throughput, and Replicates – Determine the total number of samples in your study and assess whether or not you will need to prepare and run replicates.

Read Depth – How many reads per sample will you need? The number of reads required depends on the goals of the study. For example, gene count measurements typically require lower read depth compared to measuring low-expressing genes or identification of novel features.

Read Length – For gene expression profiling, 50 bp reads are usually sufficient. For detecting currently unknown transcripts, novel splicing isoforms, gene fusion, etc., longer (150 bp) reads offer an advantage.

Single- or Paired-End Reads – While single-end reads are enough for regular gene expression profiling, paired-end reads enable discovery applications such as detecting or characterizing novel alternative splicing isoforms or gene fusion events.

Examples of Typical RNA-Seq Workflows

Workflow Example #1

• I want to focus on the coding transcriptome and I want to quantify gene expression at the gene level, with one abundance value generated per gene.

Method: mRNA-seq

GMAK recommends: ≥ 15-20 Million reads per sample, paired-end 50 bp (2 × 50 bp) run format.

Workflow Example #2

• I want to focus on the coding transcriptome and I want to quantify gene expression by analyzing abundance values for every transcript isoform from each gene (multiple abundance values per gene).

• I also want to identify novel transcript isoforms, SNVs, gene fusions, and/or identify allelespecific expression.

Method: mRNA-seq

GMAK recommends: ≥ 40 Million reads per sample, paired-end 150 bp (2 × 150 bp) run format.

Workflow Example #3

• I want to focus on both coding and multiple forms of noncoding RNA.

• I want to analyze abundance values for every transcript isoform from each gene (multiple abundance values per gene).

• I also want to identify novel transcript isoforms, SNVs, gene fusions, and/or identify allelespecific expression.

Method: Total RNA-seq

GMAK recommends: ≥ 60 Million reads per sample, 2 × 150 bp run format.

Please contact GMAK (ngs-service@helmholtz-hzi.de; +49-531-6181-5308/5318) to discuss your project needs, schedule a free consultation, or if you have further questions or concerns.