Buyer's Guide:

Simple, Customized RNA-Sequencing Workflows

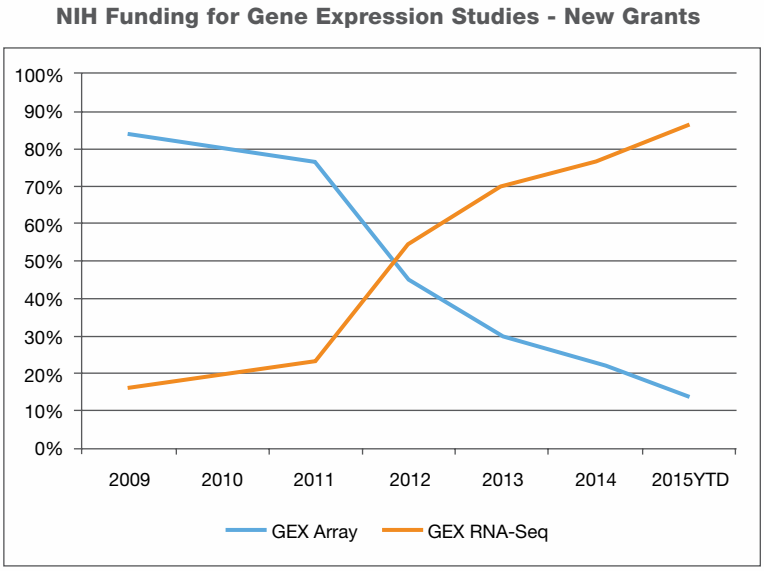
Evaluating Options for Next-Generation RNA Sequencing.

Overcome Limitations with RNA-Seq

Next-generation sequencing (NGS) has revolutionized the study of the transcriptome. Sequence data that once took years to generate can now be produced in a matter of days or hours. Advances in RNA enable simpler, easier workflows from library preparation to data analysis with highly accurate results.

RNA-Seq offers many advantages over previous methods such as qPCR and gene expression (GEX) arrays. RNA-Seq provides a wider dynamic range than GEX arrays, resulting in greater sensitivity and accuracy. And unlike both qPCR and GEX arrays, RNA-Seq can detect both known and novel features, enabling analysis of the transcriptome without the limitation of prior knowledge. Furthermore, RNA-Seq is a popular choice for nonmodel species where genetic tools and resources are limited.

The advantages of RNA-Seq and its impact on research can be seen through the rapid increase in RNA-Seq publications and the increasing success of grant proposals for gene expression studies including RNA-Seq.[1,2]



Comparison of grant awards using gene expression arrays versus RNA sequencing from 2009-2015 in the NIH Research Portfolio. Data derived from an NIH Project Reporter database search (http://projectreporter.nih.gov) in October 2015.

With Illumina® RNA sequencing workflows, RNA-Seq is more accessible than ever before. The most commonly used RNA analysis pipelines are now available through a simple, click-and-go user interface – *which means bioinformatic expertise is no longer required*. Illumina library preparation is available for a broad range of sample types including low-quality samples, such as paraffin-embedded (FFPE) tissue, and for a wide range of input amounts from normal tissue down to the single cell. If you have legacy gene expression data generated on other platforms, such as GEX arrays or qPCR, software solutions are available to integrate these datasets with RNA-Seq data for experimental continuity.

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 isofoms.

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 – Whether your samples are human, nonhuman, plant-based, or microbial, choose library preparation kits designed for the specific sample type.

Sample Quality – For low quality or degraded samples, such as FFPE preserved samples, select library preparation solutions optimized for low quality samples.

Sample Abundance – Will your RNA be derived from cell culture, single cells, or another source? Many library preparation kits are optimized for low-quantity or single-cell samples.

Customized or 'Out of the Box' – Do current out-of-the-box solutions meet your needs? Evaluate whether you have the time and resources for creating new methods or customizing existing assays.

Labor Requirements – Hands-on time (the time technicians must spend managing and monitoring the workflow), walk-away time (the ability to leave the workflow without having to intervene), total assay time, and automation options are important metrics for determining labor requirements.

Technical Expertise – Before implementing a study, determine what skill sets will be required from library preparation through data analysis and ensure that users are appropriately trained.

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. (For examples of common study objective and associated Illumina workflows, see the workflow section of this brochure.)

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

Sample Throughput – How many samples do you need to run per day/week/month? These requirements will help determine the library kit configuration and sequencing instrument that will best meet your throughput needs.

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. While read depth requirements are ultimately based on the preferences of the investigator, guidance provided by published research, resources such as the ENCODE Project RNA-Seq Guidelines v1.0,[3] as well as Illumina Technical Support provide an excellent starting point for first-time users.

Tips for Reducing RNA Sequencing Costs

- Maximize multiplexing when possible.
- Reduce the overall number of sequencing runs by using an Illumina sequencing instrument with the appropriate throughput capacity.
- Perform all pre-library and post-library preparation QC steps.
- Consider automation options vs labor costs.
- Perform a pilot study to test library preparation and analysis pipeline.

Illumina RNA-Seq Workflow Examples

After gathering information about your study goals, design requirements, and budget, you can begin to build a workflow tailored to your specific research needs. The examples below illustrate 3 of the most common RNA sequencing 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: Gene Expression Profiling (≥ 10 Million reads per sample, 1 × 50 bp run format)

Library Prep Kit	Sequencing Systems	Data Analysis Software
NeoPrep™ Library Prep System Automated library preparation, quantification, and normalization for DNA and RNA libraries. With 30 minutes of hands on time, prepare 16 libraries in 1 run. **TruSeq Stranded mRNA Library Prep Kit for NeoPrep** Compatible with the NeoPrep System, delivering high-quality, reproducible libraries. **TruSeq® Stranded mRNA Library Prep Kit** Cost-efficient, scalable library preparation for mRNA-Seq, with precise measurement of strand orientation. For standard RNA samples.*	**MiSeq System** 1 flow cell = 2 sample **NextSeq® 500 System** MO mode = 13 samples HO mode = 40 samples **HiSeq 2500 System** RR mode, 1 flow cell = 30 samples RR mode, 2 flow cell = 60 samples HO mode, 1 flow cell = 200 samples HO mode, 2 flow cell = 400 samples **HiSeq 3000 System** 1 flow cell = 250 samples **HiSeq 4000 System** 1 flow cell = 250 samples 2 flow cell = 500 samples	**RNA Express App in BaseSpace®** Align RNA-Seq reads with the STAR aligner and assign reads to genes. Perform differential gene expression with DESeq2 **NextBio® Annotates RNA-Seq App in BaseSpace** Correlate differential gene expression data from your RNA-Seq studies with disease associations and other findings from curated public data repositories. Integrate and compare legacy qPCR and GEX array data to build upon previous research.

*Standard samples include fresh or frozen tissue (not low quality or FFPE samples).
MO = Medium Output Mode
HO = High Output Mode
RR = Rapid Run Mode

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 allele-specific expression.

Method: mRNA-Seq (≥ 25 Million reads per sample, 2 × 75 bp run format)

Library Prep Kit	Sequencing Systems	Data Analysis Software
NeoPrep™ Library Prep System Automated library preparation, quantification, and normalization for DNA and RNA libraries. With 30 minutes of hands on time, prepare 16 libraries in 1 run. **TruSeq Stranded mRNA Library Prep Kit for NeoPrep** Compatible with the NeoPrep System, delivering high-quality, reproducible libraries. **TruSeq® Stranded mRNA Library Prep Kit** Cost-efficient, scalable library preparation for mRNA-Seq, with precise measurement of strand orientation. For standard RNA samples.* **TruSeq® RNA Access Library Prep Kit** Streamlined solution for mRNA-seq of low quality/FFPE samples	**MiSeq System** 1 flow cell = 1 sample **NextSeq® 500 System** MO mode = 5 samples HO mode = 16 samples **HiSeq® 2500 System** RR mode, 1 flow cell = 12 samples RR mode, 2 flow cell = 24 samples HO mode, 1 flow cell = 80 sample HO mode, 2 flow cell = 160 samples **HiSeq® 3000 System** 1 flow cell = 100 samples **HiSeq® 4000 System** 1 flow cell = 100 samples 2 flow cell = 200 samples	**BaseSpace® Core Apps for RNA** Align RNA-Seq reads, call variants, and detect gene fusions. Identify splice junctions, cSNPs, and novel transcript isoforms **NextBio® Research** Compare expression profiles from RNA-Seq, qPCR, and GEX arrays. Perform integrated analysis between DNA, RNA, and methylation studies. Compare molecular profiles from your own studies with results from curated public data repositories. **NextBio® Annotates RNA-Seq App in BaseSpace** Correlate differential gene expression data from RNA-Seq studies with disease associations and other findings from curated public data repositories.

*Standard samples include fresh or frozen tissue (not low quality or FFPE samples).
MO = Medium Output Mode
HO = High Output Mode
RR = Rapid Run Mode

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 allele-specific expression.

Method: Total RNA Sequencing (normal samples, ≥ 50 Million reads per sample, 2 × 75 bp run format) (low-quality/FFPE samples, ≥ 100 Million reads per sample, 2 × 75 bp run format)

Library Prep Kit	Sequencing Systems	Data Analysis Software
TruSeq® Stranded Total RNA with Ribo-Zero™ Human/Mouse/Rat Standard* or FFPE samples. Provides efficient rRNA removal across human, mouse, and rat as well as additional eukaryotic species.[b] **TruSeq® Stranded Total RNA with Ribo-Zero™ Gold** Standard* or FFPE samples. Removes both cytoplasmic and mitochondrial rRNA. **TruSeq® Stranded Total RNA with Ribo-Zero™ Plant** Standard* or low quality samples. Provides efficient whole-transcriptome analysis of a variety of plant species. **TruSeq® Stranded Total RNA with Ribo-Zero™ Globin** Enables efficient interrogation of valuable human, mouse, and rat blood samples, removing both rRNA and globin mRNA in a single, rapid step.	**NextSeq® 500 System** MO mode = 2 samples HO mode = 8 samples **HiSeq® 2500 System** RR mode, 1 flow cell = 6 samples RR mode, 2 flow cell = 12 samples HO mode, 1 flow cell = 40 sample HO mode, 2 flow cell = 80 samples **HiSeq® 3000 System** 1 flow cell = 50 samples **HiSeq® 4000 System** 1 flow cell = 50 samples 2 flow cell = 100 samples	**BaseSpace® Core Apps for RNA** Align RNA-Seq reads, call variants, and detect gene fusions. Identify splice junctions, cSNPs, and novel transcript isoforms **NextBio® Research** Compare expression profiles from RNA-Seq, qPCR, and GEX arrays. Perform integrated analysis between DNA, RNA, and methylation studies. Compare molecular profiles from your own studies with results from curated public data repositories. **NextBio® Annotates RNA-Seq App in BaseSpace** Correlate differential gene expression data from RNA-Seq studies with disease associations and other findings from curated public data repositories.

*Standard samples include fresh or frozen tissue (not low quality or FFPE samples).

[a] Sample number calculations are for standard samples.

[b] For information on a particular species of interest, please contact Illumina Technical Support.

MO = Medium Output Mode
HO = High Output Mode
RR = Rapid Run Mode

Maximize Performance and Productivity with Illumina Services, Training, and Consulting

Whether immediate help is needed during an instrument run, or in-depth consulting is required for workflows, Illumina can help. Illumina customer service and support teams provide guidance from initial trainings, to instrument support, personalized consultations, and ongoing NGS education. Our support offerings include:

Illumina Technical Support

- Global, 24/5 phone and email support in the Americas, Europe, and Asia-Pacific.
- Illumina TS specialists can perform desktop sharing with GoToAssist—a powerful tool for quick identification and diagnosis of issues over the phone.
- For faster case handling, enter your case number at the main phone menu to be routed directly to the TS specialist handling your case.

Illumina University Training

- Instructor-Led Training at Your Chosen Facility.
- Instructor-Led Training at an Illumina Training Center.
- Online Courses.
- Webinars.

Illumina Consulting

- Proof-of-Concept Services for instrument and library preparation testing.
- Concierge Services for design assistance and product optimization.

Product Care Services

- Tiered Instrument Service Plans + Add-On Services.
- Instrument Compliance Services.
- Instrument On-Demand Services.

For more on Illumina support offerings, visit: www.illumina.com/services/instrument-services-training.html

References

1. Comparison of publications in PubMed from 2010-2015 using the search term “RNA sequencing.” (www.ncbi.nlm.nih.gov/pubmed).
2. Comparison of grant awards using gene expression arrays versus RNA sequencing from 2009-2015 in the NIH Research Portfolio (report.nih.gov/).
3. ENCODE Project RNA-Seq Guidelines v1.0 (genome.ucsc.edu/ENCODE/protocols/dataStandards/ENCODE_RNAseq_Standards_V1.0.pdf).