illumına®

RNA-Seq Solution for Gene Expression Studies

An integrated workflow solution that fully characterizes transcriptome changes for a deep understanding of biology.

Introduction

Because gene expression patterns drive the molecular mechanisms governing biology, scientists are finding a growing body of evidence indicating that the transcriptome is critical to understanding diseases and complex traits.[1–2] Traditional technologies used to analyze gene expression, such as quantitative real-time polymerase chain reaction (qPCR) and microarrays, have been powerful tools for the research community. However, qPCR cannot be scaled to transcriptome-level complexity and gene expression arrays have a limited dynamic range, compromising sensitivity and the accuracy of measured fold change in expression. Furthermore, because they rely on prior knowledge, both approaches lack the ability to discover novel information on a large scale.

In contrast, gene expression analysis with RNA sequencing (RNA-Seq) can address these barriers and provide new insights for a broad range of biological questions. It offers a unique combination of whole-transcriptome coverage, sensitivity, and discovery potential. RNA-Seq can determine which functional pathways might be compromised or induced, what novel transcript isoforms might be involved, and how a therapeutic agent might affect gene expression in health and disease.[3,4]

Using industry-leading next-generation sequencing (NGS) technology, the Illumina workflow solution for RNA-Seq delivers a comprehensive view of the transcriptome. It combines proven kits, systems, and analysis tools with a validated, fully supported workflow to provide exceptional ease of use.

Advantages of RNA-Seq

RNA-Seq provides a detailed snapshot of the transcriptome at a given point in time. It offers numerous advantages over gene expression arrays and other technologies.

A Comprehensive View of the Transcriptome

RNA-Seq enables researchers to measure gene expression changes among both known and novel transcript isoforms. The expression of novel transcript isoforms and the abundance ratio among multiple isoforms have been shown to play a role in many diseases, including rheumatoid arthritis and neurological disorders.[5–6]

Advantages of RNA-Seq Over Gene Expression Arrays

- Requires no *a priori* knowledge of the transcriptome
- Provides qualitative and quantitative transcriptome analysis:
  — Sequence and variant information
  — More accurate measurement of gene expression[7]
  — Higher sensistivty[7–9]
- Can be applied to any species, even if a reference sequence is not available
- Delivers transcriptome information at a lower cost per sample*

Sensitive and Accurate Results

The high sensitivity of RNA-Seq provides more reliable and complete results compared to array methods for gene expression analysis.[10] NGS can quantify RNA activity at much higher resolution than array-based approaches, which is important for capturing subtle gene expression changes associated with biological processes. Microarrays measure continuous signal intensities, with a detection range limited by noise at the low end and signal saturation at the high end. In contrast, NGS quantifies discrete, digital sequencing read counts. By increasing or decreasing the number of sequencing reads, researchers can tune the sensitivity of an experiment to accommodate different study objectives.

Ratio compression is an established technical limitation of gene expression arrays that reduces sensitivity.[7–9] This characteristic ratio compression can mask real transcriptional changes, resulting in missed or misclassified biomarkers.[7] Ratio compression compromises the accuracy of gene expression changes measured by arrays, but it does not affect RNA-Seq results. RNA-Seq has been shown to detect a higher percentage of differentially expressed genes than arrays, especially lowly expressed genes.[11]

Integrated, Supported Workflow

The Illumina RNA-Seq workflow solution provides integrated support that guides researchers from library preparation to results (Figure 1). Experienced Illumina scientists provide installation and training for every system, giving researchers the confidence to start their projects immediately. A single source of technical and field specialists for library preparation, sequencing, and data analysis, combined with the Illumina support team, ensure rapid resolution and minimal potential laboratory downtime.

---

*Lower cost of RNA-Seq compared to gene expression arrays assumes a large number of samples sequenced per run.

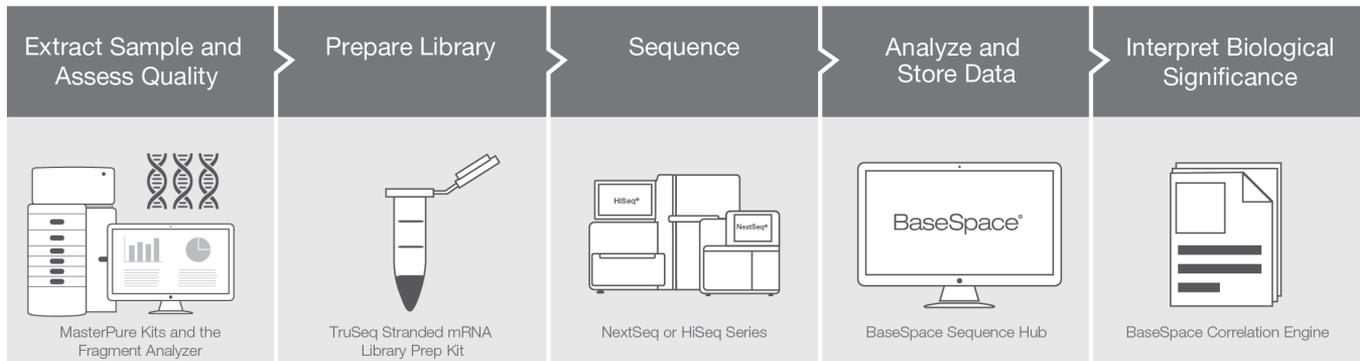**For Research Use Only. Not for use in diagnostic procedures.**

Figure 1: RNA-Seq Workflow—Illumina offers an integrated, fully supported workflow solution for RNA-Seq that provides a comprehensive view of gene expression dynamics while guiding researchers from samples to meaningful results.

Universal Sample Preparation Compatibility

The Illumina RNA-Seq workflow solution is compatible with virtually all commonly used sample preparation methods, including column-based, bead-based, and guanidinium thiocyanate-phenol-chloroform extraction–based purification methods. Epicentre® (an Illumina company) offers MasterPure RNA Purification Kits that yield total cellular RNA from many different sample types. Before preparing libraries, the Illumina workflow supports DNA quantification using the Fragment Analyzer (Advanced Analytical Technologies, Inc.) instrument.

Broad Suite of RNA Library Preparation Kits

Illumina offers an extensive portfolio of library preparation kits to address a wide range of transcriptome studies. Researchers can choose the kit that best fits their experimental goals while addressing common challenges, such as poor RNA quality or limited available sample input. All Illumina library preparation kits are fully supported, analytically validated on Illumina sequencing systems, and optimized as part of a cohesive workflow.

The TruSeq® Stranded RNA suite of library preparation kits delivers precise measurement of strand orientation and high coverage uniformity across transcripts, ensuring efficient use of read output. TruSeq Stranded mRNA Kits provide a cost-effective option for coding RNA analysis. For whole-transcriptome analysis, TruSeq Stranded Total RNA Kits capture coding and multiple forms of noncoding RNA to obtain a more complete picture of biology. TruSeq RNA Access Kits provide a low-cost RNA solution that delivers high-quality data even from degraded samples, including formalin-fixed, paraffin-embedded (FFPE) tissues. In addition, the enhanced TruSeq Stranded mRNA v2 and Total RNA v2 with Ribo-Zero® Kits provide great value and a more complete solution with reduced total assay times and hands-on times.

High NGS Data Quality

All Illumina systems use sequencing by synthesis (SBS) chemistry. Approximately 90% of the world's sequencing data are generated using Illumina SBS chemistry.[†] The flexible NextSeq® Series of Systems combines high throughput with the simplicity of a desktop sequencer for various applications, from gene expression profiling to whole-transcriptome analyses. The HiSeq® 2500, HiSeq 3000, and HiSeq 4000 Systems provide high throughput for production-scale labs to sequence large numbers of samples at one time (Table 1). Sequencing data are transferred directly from Illumina instruments to the BaseSpace® Sequence Hub for storage and analysis.

Table 1: mRNA Samples Sequenced Per Run

Sequencing System	Samples Per Run
NextSeq 500/550 Systems	
Mid Output	5
High Output	16
HiSeq 2500 System	
Rapid Run Mode	24
High Output Mode	128
HiSeq 3000 System	80
HiSeq 4000 System	160

Simplified Bioinformatics with BaseSpace Sequence Hub

RNA-Seq data can be instantly and securely transferred, stored, and analyzed in BaseSpace Sequence Hub, the Illumina genomics computing platform (Cloud or Onsite). Push-button BaseSpace Core Apps for RNA-Seq provide data analysis tools packaged in an intuitive user interface designed for biologists. Based on the most commonly cited RNA-Seq analysis pipelines, these apps deliver preconfigured workflows that support a range of common transcriptome data analysis needs (Figure 2). Researchers can investigate gene fusion events, variants, and the relative abundance of gene and transcript isoforms, as well as perform multiple case-control studies to determine differential expression patterns. Results are visualized in intuitive and interactive reports delivered in publication-ready formats.

[†]Data calculations on file. Illumina, Inc., 2015.

For Research Use Only. Not for use in diagnostic procedures.

BaseSpace Sequence Hub offers the flexibility to modify and rerun analyses on the same data set, compare runs to reference data, and receive feedback from collaborators in the same environment. Example data sets are available in BaseSpace Sequence Hub, allowing researchers to evaluate data generated by Illumina kits and systems before choosing a solution.
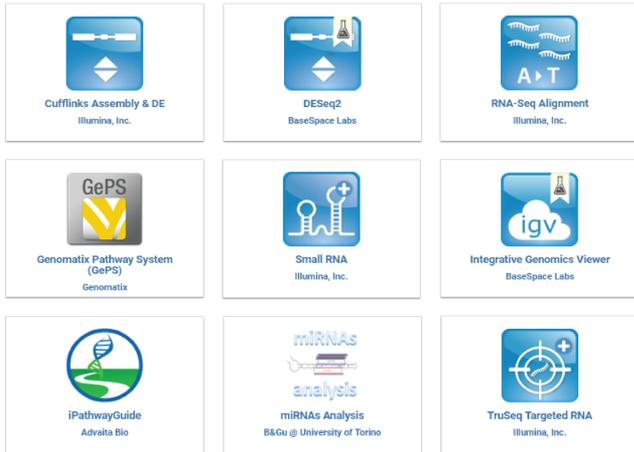

Figure 2: BaseSpace Sequence Hub RNA-Seq Apps—BaseSpace Apps for RNA offer access to expert-preferred data analysis tools via a simple, intuitive interface, so biologists can explore their RNA-Seq data without relying on bioinformatics staff.

For biological interpretation, BaseSpace Correlation Engine contains data sets from over 18,000 public studies, enriching sequencing results with functional information. BaseSpace Correlation Engine provides an interactive gene browser to visualize correlated genes and microRNA (miRNA) targets simultaneously, as well as apps for comparing relevant sets of genes across distinct projects. With BaseSpace Sequence Hub and Correlation Engine, data analysis and biological interpretation are conveniently integrated into a single platform. BaseSpace Sequence Hub and its growing library of apps enable biologists to keep pace with research and manage data easily, all without previous bioinformatics experience.

Figure 3: From Sequencing Data to Biological Effects—After analysis using BaseSpace Sequence Hub RNA-Seq Apps, RNA-Seq data can be imported into the BaseSpace Correlation Engine for functional annotation, to understand the biological effects of gene expression changes.

Summary

RNA-Seq provides a unique combination of transcriptome-wide coverage, sensitivity, and accuracy for a comprehensive view of gene expression changes. Illumina offers a comprehensive and accessible RNA-Seq workflow solution, with expert support at every step. Combining a broad library preparation portfolio, exceptional data quality, and user-friendly analysis apps, the Illumina RNA-Seq workflow enables researchers to investigate the molecular mechanisms of disease.

Learn More

To learn more about RNA-Seq data analysis tools visit: www.illumina.com/RNA-Seq-tools

For answers on common RNA-Seq data analysis questions visit: www.illumina.com/RNA-Seq-FAQs

To read more about transitioning to RNA-Seq, read the RNA-Seq Buyer's Guide: www.illumina.com/RNA-Seq-Buyers-Guide

References

1. Costa V, Aprile M, Esposito R, Ciccodicola A. RNA-Seq and human complex diseases: recent accomplishments and future perspectives. *Eur J Hum Genet*. 2013;21:134–142.

2. Wang X, Cairns MJ. Understanding complex transcriptome dynamics in schizophrenia and other neurological diseases using RNA sequencing. *Int Rev Neurobiol*. 2014;116:127–152.

3. Wang Z, Gerstein M, Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genomics*. 2009;10:57–63.

4. Wilhelm BT, Landry JR. RNA-Seq-quantitative measurement of expression through massively parallel RNA sequencing. *Methods*. 2009;48:249–257.

5. Heruth DP, Gibson M, Grigoryev DN, Zhang LQ, Ye SQ. RNA-seq analysis of synovial fibroblasts brings new insights into rheumatoid arthritis. *Cell Biosci*. 2012;2:43.

6. Tao R, Cousijn H, Jaffe AE, et al. Expression of ZNF804A in human brain and alterations in schizophrenia, bipolar disorder, and major depressive disorder: a novel transcript fetally regulated by the psychosis risk variant rs1344706. *JAMA Psychiatry*. 2014;71:1112–1120.

7. Naef F, Socci ND, Magnasco M. A study of accuracy and precision in oligonucleotide arrays: extracting more signal at large concentrations. *Bioinformatics*. 2003;19:178–184.

8. Shi L, Tong W, Su Z, et al. Microarray scanner calibration curves: characteristics and implications. *BMC Bioinformatics*. 2005;6 Suppl 2:S11.

9. Yuen T, Wurmbach E, Pfeffer RL, Ebersole BJ, Sealfon SC. Accuracy and calibration of commercial oligonucleotide and custom cDNA microarrays. *Nucleic Acids Res*. 2002;30:e48.

10. Liu Y, Morley M, Brandimarto J, et al. RNA-Seq identifies novel myocardial gene expression signatures of heart failure. *Genomics*. 2015;105:83–89.

11. Wang C, Gong B, Bushel PR, et al. The concordance between RNA-seq and microarray data depends on chemical treatment and transcript abundance. *Nat Biotechnol*. 2014;32:926–932.

For Research Use Only. Not for use in diagnostic procedures.

Ordering Information

Sample and Library Preparation Kits	
Kit	**Catalog No.**
MasterPure Complete DNA and RNA Purification Kit (100 RNA purifications)	MC85200
MasterPure Complete DNA and RNA Purification Kit (5 RNA purifications)	MC89010
TruSeq Stranded mRNA LT Set A	RS-122-2101
TruSeq Stranded mRNA LT Set B	RS-122-2102
TruSeq Stranded mRNA HT	RS-122-2103
TruSeq Stranded mRNA LT (with Ribo-Zero Human/Mouse/Rat) Set A	RS-122-2201
TruSeq Stranded mRNA LT (with Ribo-Zero Human/Mouse/Rat) Set B	RS-122-2202
TruSeq Stranded mRNA HT (with Ribo-Zero Human/Mouse/Rat)	RS-122-2203
TruSeq Stranded Total RNA LT (with Ribo-Zero Gold) Set A	RS-122-2301
TruSeq Stranded Total RNA LT (with Ribo-Zero Gold) Set B	RS-122-2302
TruSeq Stranded Total RNA HT (with Ribo-Zero Gold)	RS-122-2303
TruSeq Stranded Total RNA LT (with Ribo-Zero Globin) Set A	RS-122-2501
TruSeq Stranded Total RNA LT (with Ribo-Zero Globin) Set B	RS-122-2502
TruSeq Stranded Total RNA HT (with Ribo-Zero Globin)	RS-122-2503
TruSeq RNA Library Prep Kit v2 Set A (48 reactions)	RS-122-2001
TruSeq RNA Library Prep Kit v2 Set B (48 reactions)	RS-122-2002
TruSeq Small RNA Library Prep Kit Set A	RS-200-0012
TruSeq Small RNA Library Prep Kit Set B	RS-200-0024
TruSeq Small RNA Library Prep Kit Set C	RS-200-0036
TruSeq Small RNA Library Prep Kit Set D	RS-200-0048
TruSeq RNA Access Library Prep Kit Set A	RS-301-2001
TruSeq RNA Access Library Prep Kit Set B	RS-301-2002

Sequencing Systems and Kits	
NextSeq Series	**Catalog No.**
NextSeq 500 System	SY-415-1001
NextSeq 500 Mid Output Kit v2 (150 cycles)	FC-404-2001
NextSeq 500 Mid Output Kit v2 (300 cycles)	FC-404-2003
NextSeq 500 High Output Kit v2 (75 cycles)	FC-404-2005
NextSeq 500 High Output Kit v2 (150 cycles)	FC-404-2002
NextSeq 500 High Output Kit v2 (300 cycles)	FC-404-2004
HiSeq Series	
HiSeq 2500 System	SY-401-2501
HiSeq SBS Kit v4 (50 cycles)	FC-401-4002
HiSeq SBS Kit v4 (250 cycles)	FC-401-4003
HiSeq 3000 System	SY-401-3001
HiSeq 4000 System	SY-401-4001
HiSeq 3000/4000 SBS Kit v4 (50 cycles)	FC-410-1001
HiSeq 3000/4000 SBS Kit v4 (150 cycles)	FC-410-1002
HiSeq 3000/4000 SBS Kit v4 (300 cycles)	FC-410-1003
HiSeq 3000/4000 PE Cluster Kit	PE-410-1001

BaseSpace Onsite Sequence Hub Configurations	
Product	**Catalog No.**
BaseSpace Onsite Hub LT System	20001331
BaseSpace Onsite Hub HT System	20001332
BaseSpace Onsite Hub HT Storage 2X	20001333
BaseSpace Onsite Hub HT Storage 3X + Rack	20001334
BaseSpace Onsite Hub HT Storage 4X + Rack	20001335

Illumina, Inc. • 1.800.809.4566 toll-free (US) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com
For Research Use Only. Not for use in diagnostic procedures.

illumina®