illumina®

RNA Analysis Tools for Cancer Research

From targeted analysis to complete transcriptome profiling.

Introduction

The analysis of RNA is an invaluable tool for the investigation of cancer samples. Understanding changes in gene expression and gene regulation, as well as identifying gene fusions can help researchers and clinicians to:

- Identify mutational profiles and signatures associated with individual tumors and tumor types
- Find novel small RNAs that regulate gene expression
- Focus on somatic mutations that have a clear functional effect
- Detect key driver mutations that are present within each tumor
- Identify gene fusions arising from chromosomal translocations

Illumina RNA Solutions for Cancer Researchers

Illumina now offers a complete suite of end-to-end RNA solutions for cancer researchers. Our comprehensive workflows and analysis solutions take cancer samples from mRNA or total RNA through sequencing and analysis for both discovery and validation applications (Table 1).

Transcriptome Sequencing

Historically, arrays and small-scale RT-PCR have been the only cost-effective solutions to understand cancer gene expression changes. The use of expression arrays in particular has proved to be extremely informative both in the molecular profiling of tumor types and in understanding their therapeutic response.[1,2] However, arrays have

Table 1: Recommended Solutions for Cancer Transcriptome Analysis

Library Preparation Kit	Best Application	Key Features	Instrument	Analysis Solution
TruSeq Stranded Total RNA	• Analysis of non-coding and coding RNA for expression and splicing changes in high-quality or degraded/FFPE samples • Fusion-gene detection in high-quality or degraded/FFPE samples • cSNP analysis	• Compatible with challenging, degraded, or FFPE samples • Ribo-Zero rRNA reduction chemistry enables detection of various forms of non-coding RNA (lncRNA, snRNA, snoRNA) in addition to coding transcripts, even in FFPE samples	NextSeq, HiSeq	BaseSpace Core Apps for RNA
TruSeq Stranded mRNA	• Discovery of expression and splicing changes in fresh samples • Fusion-gene detection in fresh samples • cSNP analysis	• Low input, simple workflow • Precise measurement of mRNA strand orientation and high-coverage uniformity enables detection of antisense transcription, enhanced transcript annotation, and increased alignment efficiency	NextSeq, HiSeq	BaseSpace Core Apps for RNA
TruSeq Targeted RNA Expression	• Profiling or validation of expression and splicing changes in high-quality or degraded/FFPE samples • Fusion-gene analysis in high-quality or degraded/FFPE samples. • Tumor typing • cSNP analysis	• Enable expression profiling of tumors, validation of gene fusions, and tumor typing even from degraded or FFPE samples • Assess the expression levels of up to 2000 targets simultaneously in a single MiSeq run • Select from over 500,000 prede-signed assays to create custom panels or choose pre-validated fixed panels that can also be supplement-ed with custom content	MiSeq	MiSeq Reporter
TruSeq Small RNA	• Analysis of expression changes in miRNA • miRNA variant analysis	Enables detection of miRNA variants such as isomiRs	MiSeq, NextSeq	MiSeq Reporter

several limitations that make them less than ideal for studying cancer. First, gene expression arrays have limited dynamic range, which can lead to less accurate measurement of expression changes. Second, arrays lack the ability to discover novel features such as splice isoforms or fusion transcripts. Third, expression arrays are unable to detect variants that weren't explicitly designed, such as coding SNPs. These weaknesses have led researchers to seek out more enabling technologies for gene expression analysis. As the cost of sequencing decreases, RNA sequencing (RNA-Seq) has emerged as the leading approach.

RNA sequencing has the advantage of providing a complete snapshot of the changes within a cancer, from gene expression profiles and gene regulation information to the identification of structural variants, such as gene fusions. RNA-Seq is particularly suited to gene fusion analysis because it detects only the expressed chromosomal rearrangements, which are the translocations most likely to be functionally relevant in the cancer being studied. Furthermore, because RNA-Seq detects only expressed transcripts, it identifies events that are more likely to be causal in cancer.[3–6] The detailed information that can be obtained from RNA sequencing, often performed in combination with DNA sequencing of the tumor and surrounding "normal" tissue or blood, has proven to be invaluable for the classification, understanding, and treatment choice of a number of cancers.[7,8]

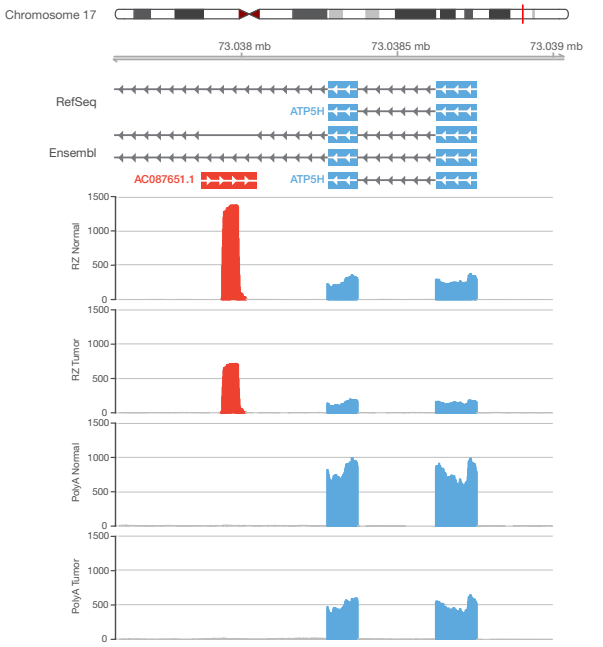
For coding transcriptome or whole transcriptome analysis, Illumina offers TruSeq RNA v2, TruSeq Stranded mRNA, or TruSeq Stranded Total RNA library preparation kits. The TruSeq RNA library preparation kits all feature fragmentation and random priming of RNA for the most uniform sequencing coverage and a streamlined workflow for higher throughput studies.

TruSeq Stranded mRNA kits add the benefit of strand information, knowing from which of the two DNA strands the transcript was derived. This allows the detection of antisense expression and improves transcript annotation, while also increasing the number of uniquely alignable reads.

TruSeq Stranded Total RNA kits with Ribo-Zero enable the detection and quantification of both coding and non-coding transcripts and are compatible with FFPE samples. Ribo-Zero reduction chemistry is included in these kits to efficiently remove highly abundant RNA species (such as ribosomal and globin RNA). By minimizing contamination from these over-represented transcripts, TruSeq Stranded Total RNA kits increase the percentage of reads covering transcripts of interest. These qualities make TruSeq Stranded Total RNA kits ideal for cancer researchers seeking the most complete view of the transcriptome (Figure 1).

Libraries made with any of these kits can be run on any Illumina sequencer, but are ideally suited for higher-throughput instruments, such as the NextSeq™ 500 or HiSeq® 2500. In addition, run data from these instruments can be streamed directly into BaseSpace for analysis with the BaseSpace Core Apps for RNA, providing a complete sample-to-answer solution.

Figure 1: Differential Expression of ncRNA Transcripts



ATP5H expression from chromosome 17 is differentially expressed in breast tumor vs. normal tissue. Using two different library preparation methods (RZ; Ribo-Zero for total RNA or PolyA-based mRNA) shows differential expression in tumor vs. normal tissues in both preps (Blue). However, only Total RNA with Ribo-Zero reveals differential expression at the locus of a pseudogene (Red, AC087651.1), for which reads are detected in the opposite orientation, as expected. This stranded information would have been lost in a standard mRNA prep.

Targeted Transcriptome Analysis

While RNA-Seq is a powerful technology for analyzing the whole transcriptome or coding transcriptome, there are times when a more targeted approach is preferable. For example, screening a large number of samples for a known expression profile or validating differential expression hits from an RNA-Seq experiment would see cost and analysis benefits using a targeted solution. Although RT-PCR is an effective tool, it is limited in the number of genes researchers can practically assess at one time and becomes costly as the gene number increases.

Illumina has developed the TruSeq Targeted RNA Expression kit to address these limitations. It provides a cost-effective alternative for mid plexity expression studies such as these. These kits enable researchers to design panels of up to 1000 targets by choosing from a database of over 500,000 pre-designed assays or from a number of pre-validated fixed panels covering cellular pathways important in cancer. Content can be adjusted at any time using the add-on capability of these products. Up to 1000 targets can be added to any previously designed panel for a total of 2000 targets in a single MiSeq run. These features make Truseq Targeted RNA Expression kits ideal for expression profiling of tumors, validation of gene fusions, and tumor typing, even from degraded or FFPE samples (Figure 2).

These kits are optimized for the MiSeq and MiSeq Reporter.

Gene Regulation

Small RNAs have proven to play key roles in cancer. Because they regulate gene expression, their own deregulation can have striking effects on the expression patterns within cancer cells.[9] Furthermore, small RNAs may be key biomarkers for cancers.[10] More recently, non-coding RNAs have also been identified as playing key roles in cancer,[11,12] driving the need for more comprehensive RNA analysis tools. With Illumina, researchers can choose to study small RNA with the TruSeq Small RNA Library Preparation Kit or mRNA and non-coding RNAs with the TruSeq Stranded Total RNA Library Preparation Kit.

The TruSeq Small RNA Library Preparation Kit is ideally suited to the MiSeq system, which includes a built-in analysis solution with the Small RNA workflow.

Instrumentation and Analysis

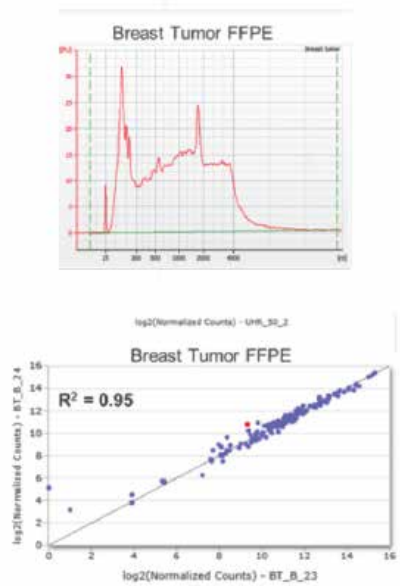
In addition to library preparation solutions, Illumina offers sequencers for any budget or throughput need. Industry-leading Illumina sequencers provide the most accurate data with streamlined workflows, as well as the ability to perform paired-end sequencing, which is key in the detection of gene fusions. Although Illumina sequencing libraries can be run on any of our next-generation sequencing platforms, the ideal instrument for each application depends on both the required read depth and customer throughput Table 1).

MiSeq Reporter provides a built-in analysis solution for samples run on the MiSeq. Two BaseSpace Core Apps, based on the industry-standard Tophat/Cufflinks tools, are available for transcriptome analysis, and can be used for differential expression and the detection of novel isoforms, gene fusions, and cSNPs (Figure 3).

Summary

Transcriptome analysis has been shown to provide critical information to cancer researchers and clinicians regarding gene regulation/dysregulation profiles, key driver mutations, and fusion genes. Although still in the early stages, it has also been shown to provide prognostic and diagnostic information, and can even be used to guide the choice of therapy. With a complete suite of tools from library preparation to analysis, Illumina offers industry leading, sample-to-answer solutions for a complete range of RNA applications.

Figure 2: TruSeq Targeted RNA Expression for Expression Profiling of Tumor Tissues



Panel A: Bioanalyzer trace showing degraded RNA derived from formalin-fixed paraffin embedded (FFPE) breast cancer tissue.
Panel B: Analysis of a TruSeq Targeted RNA Expression library made from RNA in panel A shows very high replicate correlation, indicating assay robustness with FFPE-derived samples.

Figure 3: Using the Differential Expression Gene Browser in the Cufflinks Assembly and Differential Expression App

Interactive assessment of significantly expressed genes can be performed live using the browser.

References

1. Sadi AM, Wang DY, Youngson BJ, Miller N, Boerner S, et al. Clinical relevance of DNA microarray analyses using archival formalin-fixed paraffin-embedded breast cancer specimens. BMC Cancer. 2011 Jun 16;11:253:1-13. doi: 10.1186/1471-2407-11-25.
2. Wesolowski R, Ramaswamy B. Gene expression profiling: changing face of breast cancer classification and management. Gene Expr. 2011;15(3):105-15.
3. Maher CA, Kumar-Sinha C, Cao X, Kalyana-Sundaram S, Han B, et al. Transcriptome sequencing to detect gene fusions in cancer. Nature 2009, 458:97-101.
4. Maher CA, Palanisamy N, Brenner JC, Cao X, Kalyana-Sundaram S, et al. Chimeric transcript discovery by paired-end transcriptome sequencing. Proc Natl Acad Sci USA 2009, 106:12353-12358.
5. Berger MF, Levin JZ, Vijayendran K, Sivachenko A, Adiconis X, et al. Integrative analysis of the melanoma transcriptome. Genome Res 2010, 20:413-427.
6. Edgren H, Murumagi A, Kangaspeska S, Nicorici D, Hongisto V, et al. Identification of fusion genes in breast cancer by paired-end RNA-sequencing. Genome Biol 2011, 12:R6.
7. Sato Y, Yoshizato T, Shiraishi Y, Maekawa S, Okuno Y, et al. Integrated molecular analysis of clear-cell renal cell carcinoma. Nat Genet. 2013 Aug;45(8):860-7.
8. Weiss GJ, Liang WS, Demeure MJ, Kiefer JA, Hostetter G, et al. A pilot study using next-generation sequencing in advanced cancers: feasibility and challenges. PLoS One. 2013 Oct 30;8(10):e76438.
9. Ventura A, Jacks T. MicroRNAs and cancer: short RNAs go a long way. Cell. 2009 Feb 20;136(4):586-91.
10. S P Nana-Sinkam 1,2,3 and C M Croce 2,3,4. Clinical Applications for microRNAs in Cancer. Clinical Pharmacology & Therapeutics (2013); 93 1, 98–104. doi:10.1038/clpt.2012.192.
11. Spizzo R, Almeida MI, Colombatti A, Calin GA. Long non-coding RNAs and cancer: a new frontier of translational research? Oncogene. 2012 Oct 25;31(43):4577-87. doi: 10.1038/onc.2011.621. Epub 2012 Jan 23.
12. Deng G, Sui G. Noncoding RNA in oncogenesis: a new era of identifying key players Int J Mol Sci. 2013 Sep 5;14(9):18319-49. doi: 10.3390/ijms140918319.
13. Genome Biology 2013, 14:R36; Nature Biotechnology 28, 511–515 (2010).

Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY

Pub. No. 470-2013-003 Current as of 11 November 2014

illumina®

AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
TCAACGTACCGTAACGAACGTATCATTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTT
CGACGAAAAGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACC
AACGTACCATTAAGAGCTACCGTGCAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGACGAAAAGAATGATAA
AGAATGATAACAGTAACACACTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAAC
GATTACTTGATCCACTGATTCAACGTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGCTTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAGCAACGAGC
CGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGTCTGTTAACCTTAAGATTACTTGATCCACTGATTCAACGTACCGTAACGAACGTATCAATTGAGACTAAATATTAACGTACCATTAAGAGCTACCGTGCAACGAAAAGAATGATAACAGTA