# Simple, scalable transcriptome-level gene expression analysis with next-generation sequencing

## Overview

RNA sequencing analyzes the continuously changing cellular transcriptome, providing valuable information about the internal state of cells and transcriptional networks. Differential gene expression analysis, one of the most common applications of RNA sequencing, has played a significant role in the identification of pathway interactions or disease-state biomarkers.

Two common types of transcriptome analysis were compared on Ion 550™ Chips using the Ion GeneStudio™ S5 Prime System—whole-transcriptome sequencing, and gene expression analysis using Ion AmpliSeq™ Transcriptome kits. Our results show high correlation in gene expression analysis using these two methods, as well as concordance with Applied Biosystems™ TaqMan® OpenArray™ results.

## Whole-transcriptome sequencing (RNA-Seq) for both discovery and gene expression analysis

This technique enables the sequencing of all RNA types in a high-throughput manner, offering several advantages for detecting gene expression, including high sensitivity, strand-specific detection, and wide coverage of transcripts at the transcriptome level.

Combined with the Ion GeneStudio S5 Prime System, the Ion Total RNA-Seq Kit offers:

- **A consolidated solution**—whole-transcriptome, mRNA, or small RNA sequencing solution in one kit
- **Preservation of strand information**—all mapped reads are aligned in the direction of transcription relative to the chromosomal strand
- **Flexible workflow options**—ability to make library from human, mouse, or any other species

However, whole-transcriptome sequencing is generally expensive compared to targeted sequencing approaches. It requires a significant amount of starting material, and therefore this method is not suitable when starting material is limited or degraded. For large-scale differential gene expression analysis that involves hundreds of samples, the bioinformatics analysis and data storage requirements of RNA-Seq become the limiting factor. In such situations, a targeted quantitative method with faster turnaround time, higher accuracy and sensitivity, and lower cost per sample can offer a better alternative [1].

## Ion AmpliSeq Transcriptome solution for gene-level expression analysis

The Ion AmpliSeq™ Transcriptome Gene Expression Kit used with the Ion GeneStudio S5 Prime System help to enable fast and affordable gene expression analysis from samples limited in quantity and quality. With simultaneous amplification of over 20,000 RefSeq genes in a single tube starting from as little as 10 ng of total RNA, these kits serve as an important tool in oncology and translational research. The Ion AmpliSeq Transcriptome Gene Expression Kits are available for both human and mouse analysis. These kits offer:

- **Low sample input**—as little as 10 ng of total RNA (compatible with formalin-fixed, paraffin-embedded (FFPE) samples)
- **Fast gene expression profiling with simple workflow**—go from RNA to data in <2 days and only ~45 minutes of total hands-on time with the automated workflow of the Ion S5™ and Ion Chef™ Systems (Figure 1)
- **Streamlined data analysis**—export data with Torrent Suite™ Software into existing analysis software

## Materials and methods

### RNA samples

Clontech Mouse Total RNA Master Panel (MMP) (Cat. No. 636644), Biochain Universal RNA-Mouse (MUR) (Cat. No. R4334566-1), and Agilent Universal Mouse Reference (UMR) (Cat. No. 740100) RNAs were used to assess the performance of the Ion AmpliSeq™ Transcriptome Mouse Gene Expression Panel, Chef-Ready Kit (Cat. No. A36412) and the Ion Total RNA-Seq Kit for AB Library Builder™ System (Cat. No. 4482416).

### Ion AmpliSeq panel design

The RefSeq GRCm38.p5 assembly was used as the design reference. Unassigned contigs, and records with no annotated exomes, no transcript ID, or no gene ID were filtered so that only well-annotated RefSeq genes were included in the panel. Amplicon designs targeted exon–exon junctions common to all isoforms when available. A total of 23,930 genes were designed and tested to give good performance in one pool panel.

| Ion AmpliSeq Trancriptome Human or Mouse Gene Expression Panels in Chef-Ready Kits | |
|---|---|
| Fully automated production of Ion AmpliSeq transcriptome libraries with less than 1 hour total of hands-on time from sample to data | Construct libraries during the day; prepare templates and load chips overnight |
| Same chemistry, and performance equivalent to that of a manual workflow | 32 samples per kit; optimal configuration is 8 samples per Ion 540™ Chip |

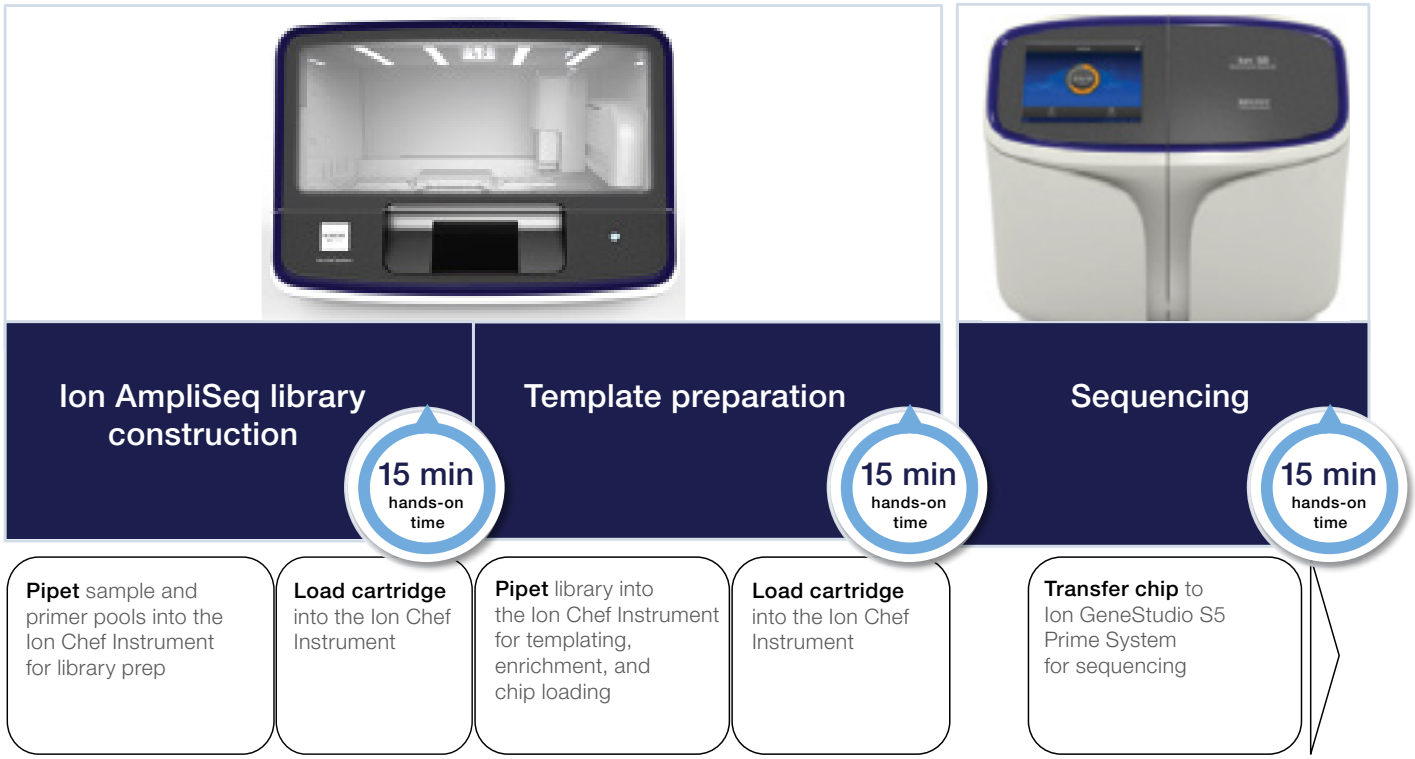



**Figure 1. Streamlined workflow for gene expression analysis.** Fully automated workflow from cDNA to gene expression data requires only 45 minutes of hands-on time using the Ion Chef System and Ion GeneStudio S5 Prime System.

### Library preparation and sequencing

A total of 10 ng RNA was reverse-transcribed using the Invitrogen™ SuperScript™ VILO™ cDNA Synthesis Kit (Cat. No. 11754050). Transcriptome libraries were generated using automated library preparation with the Ion AmpliSeq Transcriptome Mouse Gene Expression Panel, Chef-Ready Kit (Cat. No. A36412) for MMP and MUR samples. The kit enables users to generate 8 equalized Ion AmpliSeq libraries utilizing the robust liquid handling capabilities of the Ion Chef System. The transcriptome libraries were barcoded, templated, and sequenced on the Ion Chef and Ion GeneStudio S5 Prime Systems using the Ion 550™ Kit-Chef (Cat. No. A34541) and Ion 550 Chip Kit (Cat. No. A34537), as one 16-plex library pool. The Ion GeneStudio S5 Prime System combined with the automated library and template preparation on the Ion Chef System enables a simple solution for gene expression analysis, requiring only 45 minutes of hands-on time from 10 ng of RNA (Figure 2).

To prepare ribosomal RNA (rRNA)-depleted total RNA for both MMP and MUR, the Invitrogen™ RiboMinus™ Eukaryote System v2 (Cat. A15026) was used for 2 µg of MMP and MUR. The rRNA-depleted total RNA was fragmented, then used to construct the libraries using the Ion Total RNA-Seq Kit for the AB Library Builder System. The cDNA was amplified, barcoded, templated, and sequenced on the Ion Chef and Ion GeneStudio S5 Prime System using the Ion 550 Kit-Chef (Cat. No. A34541) and Ion 550 Chip Kit (Cat. No. A34537).

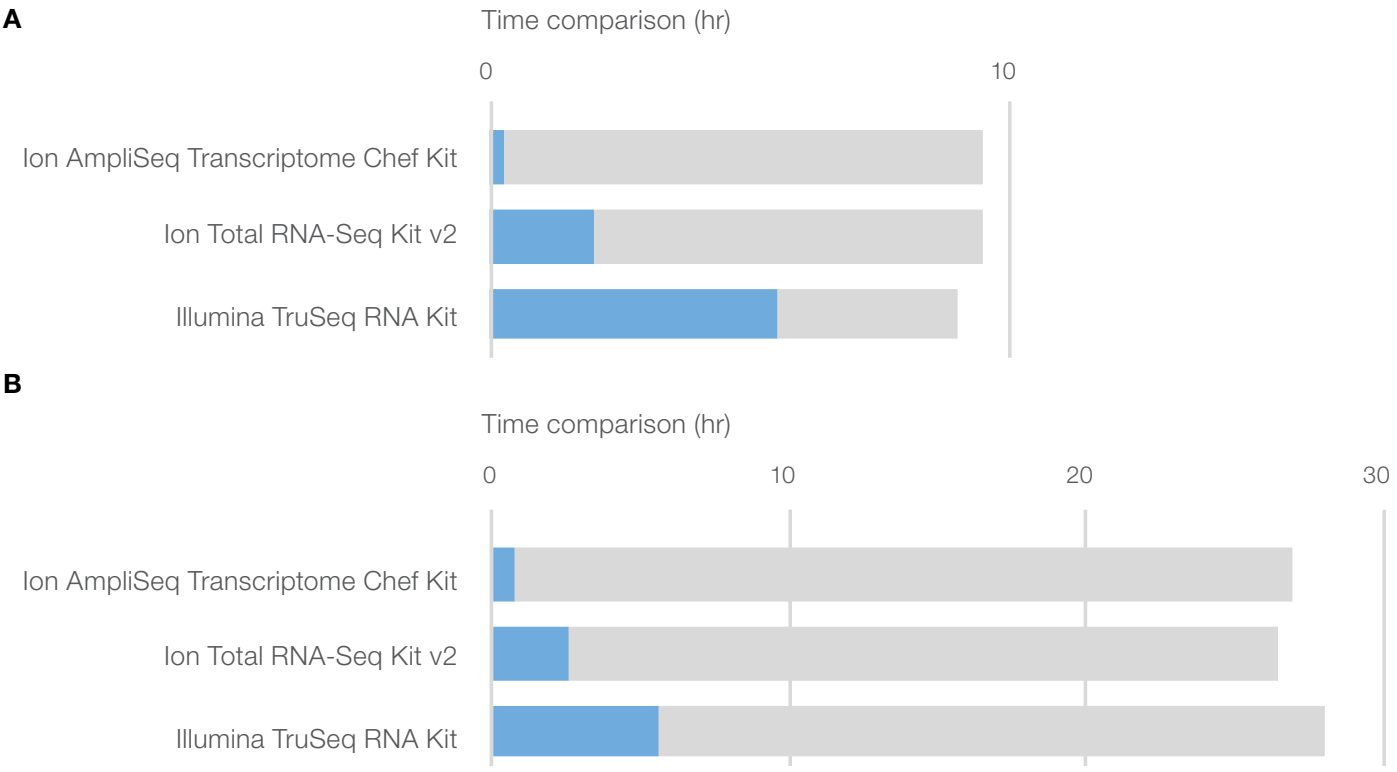


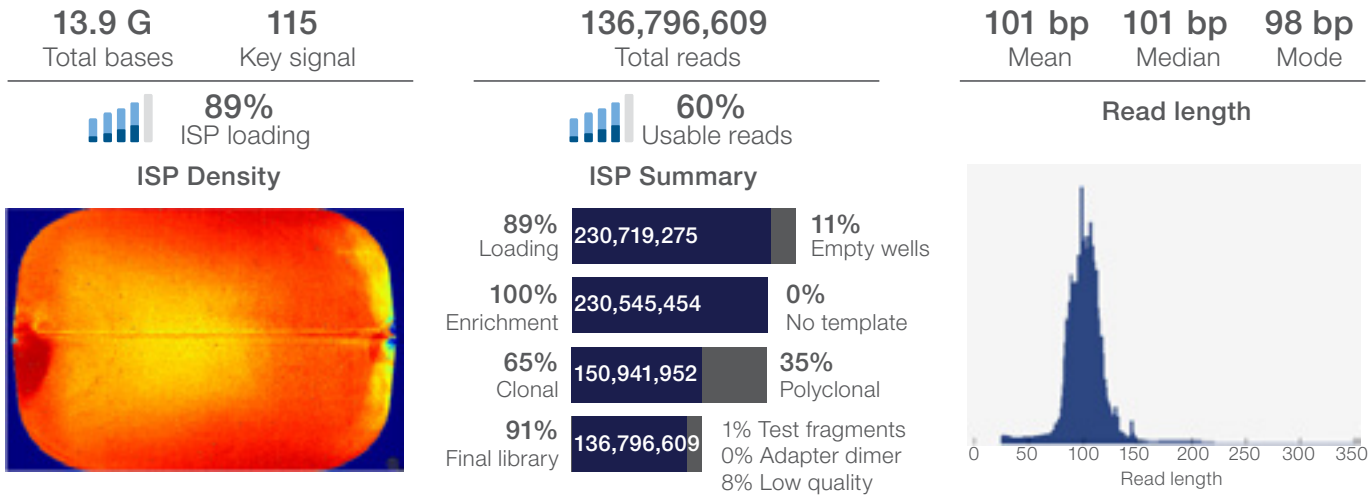

**Figure 2. Comparisons of hands-on and total turnaround time (in hours) between sequencing RNA-Seq kits. (A)** Library preparation. **(B)** Entire workflow, from sample to data files.

## Results

The Ion AmpliSeq Transcriptome Mouse Gene Expression Kit can be used to determine the relative expression levels of >20,000 known and predicted RNA transcripts. To demonstrate the performance of the Chef-Ready version of the kit, 16 libraries utilizing 4 different commercially available RNAs were generated on the automated Ion Chef platform and then sequenced simultaneously on a single Ion 550 Chip using the Ion GeneStudio S5 Prime System.

The run produced over 135 million reads with high rates of mapping to RNA targets (Figure 3). Depending on the RNA source, 52–71% of known RNAs were detected with Pearson's correlation values (*r*) of >0.99. As a QC reference for system performance, *in vitro*–transcribed standards developed by the External RNA Controls Consortium (ERCC) were spiked into each sample; the data showed that the read counts of the spiked standards were highly correlated to their known concentrations (Figure 4).

**A**




**B**

| Sample | Valid reads (%) | Targets detected (%) |
|---|---|---|
| UMR rep 1 | 99.39% | 52.63% |
| UMR rep 2 | 99.40% | 52.99% |
| MUR rep 1 | 99.56% | 65.48% |
| MUR rep 2 | 99.61% | 65.22% |
| MMP rep 1 | 99.49% | 70.98% |
| MMP rep 2 | 99.47% | 70.81% |
| e17 rep 1 | 99.31% | 61.12% |
| e17 rep 2 | 99.20% | 61.49% |
| UMR rep 3 | 99.51% | 53.70% |
| UMR rep 4 | 99.56% | 53.27% |
| MUR rep 3 | 99.59% | 65.93% |
| MUR rep 4 | 99.66% | 65.32% |
| MMP rep 3 | 99.55% | 71.14% |
| MMP rep 4 | 99.59% | 69.97% |
| e17 rep 3 | 99.38% | 58.27% |
| e17 rep 4 | 99.35% | 58.32% |

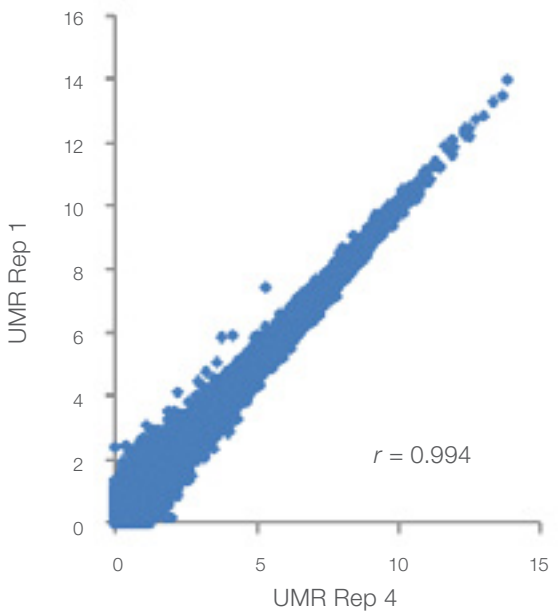
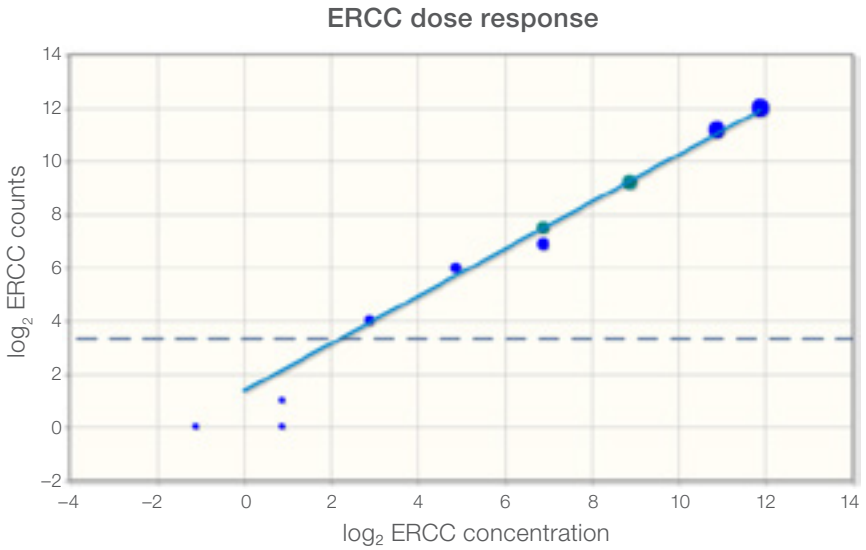



**Figure 3. An example run of 16 libraries made with the Ion AmpliSeq Mouse Transcriptome panel was analyzed using Torrent Suite software and the ampliSeqRNA plug-in. (A)** RNA-Seq using the Ion GeneStudio S5 Prime System produced >135 million reads with high rates of mapping to RNA targets. **(B)** A typical correlation plot of two replicates of the UMR sample from Agilent is presented, with $\log_2$ [(reads per million) + 1] for each gene in the panel compared. Correlation between replicates of any two samples was similar to or higher than the presented value of Pearson's $r = 0.994$. See "Materials and methods" for abbreviations of RNA names; e17 = embryonic day 17, one of the fifteen tissues comprising the MMP RNA sample.




| ERCC plug-in results | | |
|---|---|---|
| **% of total reads** | | |
| ERCC (%) | = | 0.08 |
| Other (%) | = | 99.92 |
| Raw ERCC counts | = | 7,390 |
| **Regression analysis** | | |
| $R^2$ | = | 0.99 |
| Slope | = | 0.89 |
| Y-intercept | = | 1.35 |
| N | = | 7 |
| **Map quality legend** | | |
| Mean MAPQ | = | Dot color |
| More than 90 | = | |
| Less than 90 | = | |
| Less than 80 | = | |
| Less than 70 | = | |
| Less than 60 | = | |

**Figure 4. Correlation data between ERCC transcripts and RNA samples.** Primers for 10 ERCC transcripts were spiked into the Ion AmpliSeq Mouse Transcriptome panel, and those same transcripts were spiked into the 16 RNA samples during the library prep. A typical result of the ERCC plug-in analysis demonstrates high correlation between the observed read counts and the spiked-in ERCC transcript concentrations. Note: Transcripts with fewer than 10 counts (below the dashed line) are not used in calculating $R^2$.

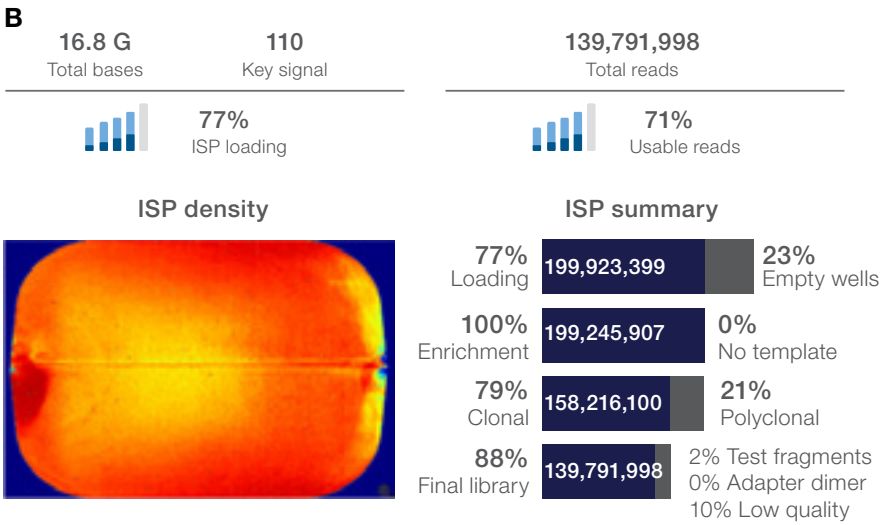
The Ion Total RNA-Seq Kit can discover novel isoforms and transcripts as well as quantitate differential expression. A total of 3 replicate libraries were generated by the automated AB Library Builder system from the MMP sample and sequenced on an Ion 550 Chip (Figure 5).

Again, >135 million total reads were produced with high mapping rates. An average of almost 2 different isoforms were detected for each transcript. Reproducibility among the 3 library replicates exceeded the Pearson's correlation value (*r*) of 0.99.

**A**

| Sample | Total reads | Aligned reads | Percent aligned (%) | Genes detected | Isoforms detected |
|---|---|---|---|---|---|
| MMP_IC04 | 47,581,042 | 46,993,742 | 98.77 | 22,281 | 43,421 |
| MMP_IC05 | 43,153,682 | 42,548,057 | 98.60 | 22,313 | 43,499 |
| MMP_IC06 | 46,332,124 | 45,830,715 | 98.92 | 21,804 | 41,499 |

**B**




**C**

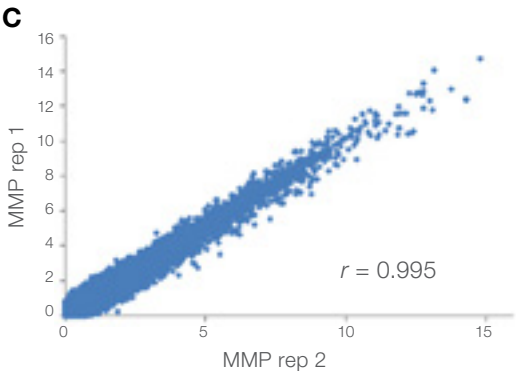



**Figure 5. Wild-type RNA-Seq data for MMP samples. (A)** Whole-transcriptome RNA-Seq libraries were made for three replicates of the BioChain MMP Universal Mouse RNA sample. **(B)** Sequencing data were analyzed using Torrent Suite software, and **(C)** 2 typical replicates of this sample were compared for reproducibility by plotting the $\log_2$ [(reads per million) + 1] and calculating the Pearson's *r*.

To compare the two different RNA sequencing methods described here, differential expression was determined for the replicate MUR and UMR samples and plotted as $\log_2$ of the fold change (Figure 6). Although there are fundamental differences between the two approaches, the Pearson's correlation was found to be higher than 0.9. The performance of the Ion AmpliSeq panel was also compared to a nonsequencing technology by generating data from a TaqMan OpenArray panel. Highly similar differential expression (Pearson's correlation 0.847) was detected by both platforms.

Differential expression files generated in Torrent Suite software can be used in the **Applied Biosystems™ Transcriptome Analysis Console (TAC) software v4** for pathway analysis (Figure 7). These visualizations make it easier to understand how changes in gene expression may affect important processes in the cells or tissue under study.

## Conclusion

The two RNA-Seq tools discussed here, the Ion AmpliSeq Transcriptome panel and Ion Total RNA-Seq Kit, offer valuable benefits in the entire sequencing workflow—from sample prep to analysis. Both solutions provide a wide dynamic range to detect the full breadth of biologically relevant gene expression changes, and automated library preparation solutions. NGS read counts provide a representation of absolute expression, enabling you to identify and characterize low-abundance transcripts.

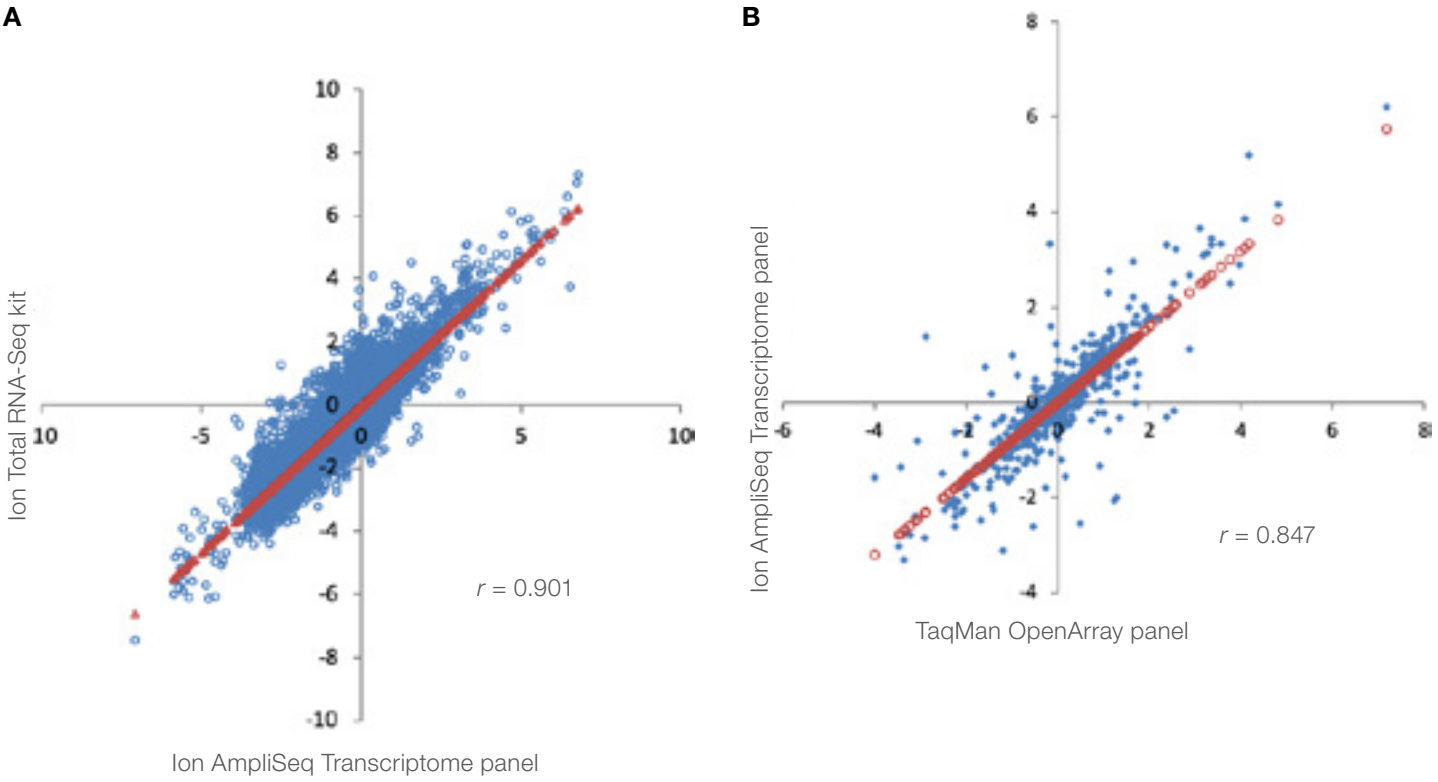



**Figure 6. Fold-change correlation between the Ion Total RNA-Seq Kit, Ion AmpliSeq Transcriptome panel, and TaqMan OpenArray panel.** **(A)** Correlation between the Ion Total RNA-Seq Kit and Ion AmpliSeq Transcriptome panel. The fold changes in gene-level RNA read counts from two different samples (MUR and MMP) were calculated from 3 RNA-Seq libraries and 4 Ion AmpliSeq RNA libraries from a single typical run for each sample. The $\log_2$ of the average fold-change differences for each gene with read counts greater than 10 were plotted above, and the Pearson's r correlation coefficient was calculated using regression analysis. **(B)** Correlation between Ion AmpliSeq panel and TaqMan OpenArray panel. The fold changes in $\Delta C_t$ for gene-level read counts using the Ion AmpliSeq panel and transcript-level read counts using the TaqMan OpenArray panel were calculated on MUR and MMP to assess the concordance between these two platforms. The $\Delta C_t$ values from a total of 12 barcodes of each Ion AmpliSeq panel and one OpenArray panel were adjusted by quantile normalization, and the $\log_2$ of the fold changes were plotted via regression analysis. The Pearson's *r* was calculated both before and after removing 20 outlier data points out of a total of 545 pairwise gene expression comparisons.

The Ion AmpliSeq Transcriptome gene expression kits and panels for mouse and human enable a total of 16 samples to be sequenced simultaneously on an Ion 550 Chip. The Ion Total RNA-Seq Kit can be used for novel transcript/isoform discovery in addition to expression analysis with RNA from any species. For mouse and human samples, up to 3 libraries can be sequenced on an Ion 550 Chip.

Combined with the simple, integrated Torrent Suite software, Ion Torrent™ transcriptome sequencing solutions offer fast identification and reporting of differentially expressed genes. Torrent Suite software takes you from raw sequencing data to informative results, including optimized signal processing, base calling, sequence alignment, and transcriptome analysis. Optional workflows in TAC software can be used to perform multisample differential gene expression analysis and pathway analysis, enabling interactive data visualization and providing more data insight.

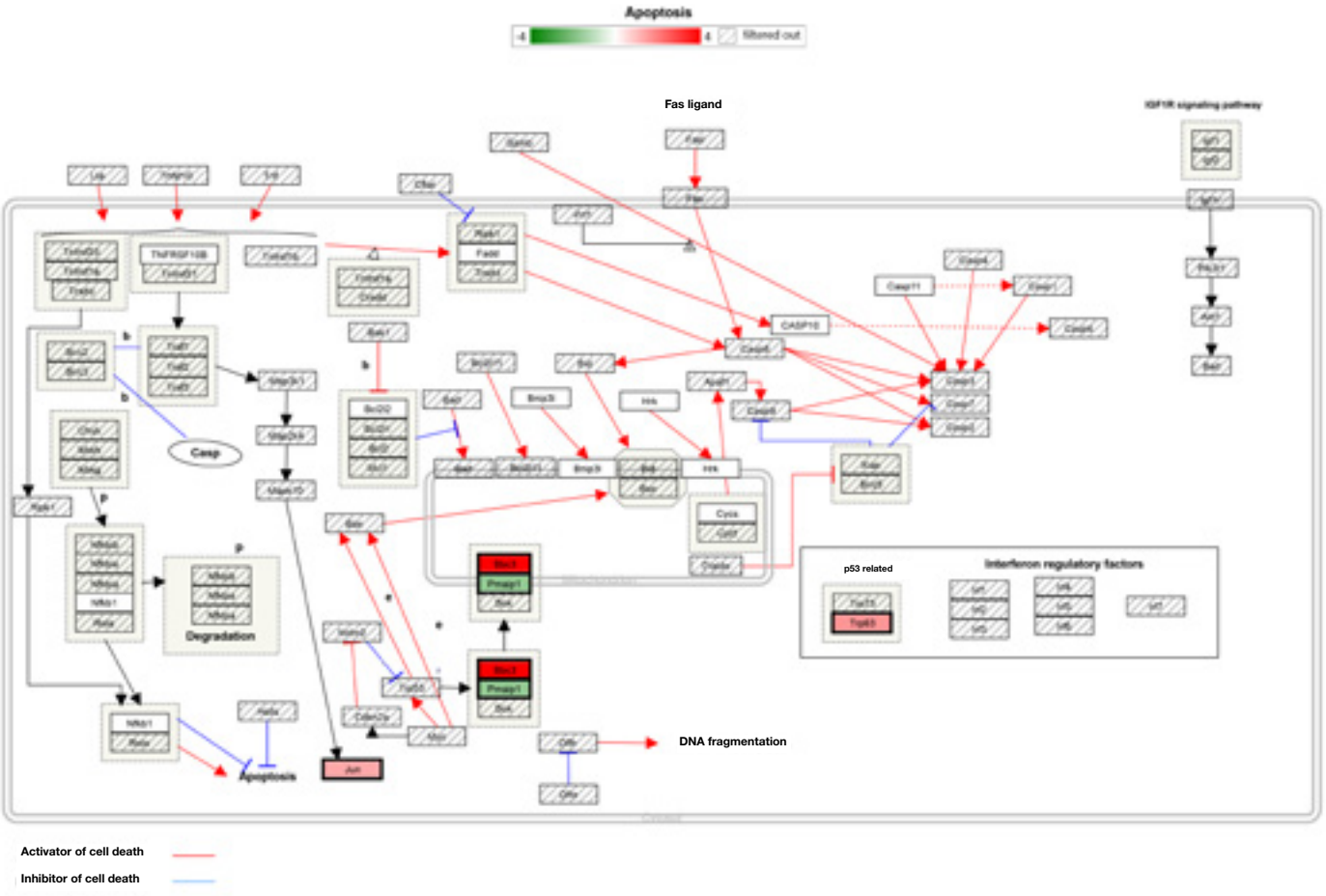



**Figure 7. Pathway analysis using TAC software.** The ampliSeqRNA plug-in for Torrent Suite software produces files in .chp format representing the $\log_2$ of reads per million for each barcode. These files can be used in the TAC v4 software available from Thermo Fisher Scientific. The differential expression analysis by TAC software on two different samples (illustrated above with data generated from UMR and MUR) detects genes whose expression ratios are statistically different from 1. Wikipathway analysis (**wikipathways.org/index.php/wikipathways**) of a typical experiment is displayed here for the apoptosis pathway. Red indicates genes with reduced expression, while green indicates increased expression.

ion torrent

## Ordering information

| Product | Quantity | Cat. No. |
|---|---|---|
| Ion AmpliSeq Transcriptome Mouse Gene Expression Kit* | 24 reactions | A36553 |
| | 96 reactions | A36554 |
| | 384 reactions | A36555 |
| Ion Xpress Barcode Adapters 1-16 Kit | 1 kit | 4471250 |
| Invitrogen SuperScript VILO cDNA Synthesis Kit | 50 reactions | 11754050 |
| Ion AmpliSeq Transcriptome Mouse Gene Expression Panel, Chef-Ready Kit** | 32 reactions | A36412 |
| Ion AmpliSeq RNA ERCC Panel | 24 reactions | A36552 |
| Ion AmpliSeq Transcriptome Human Gene Expression Panel, Chef-Ready Kit | 32 reactions | A31446 |
| Ion AmpliSeq Transcriptome Human Gene Expression Kit | 24 reactions | A26325 |
| | 96 reactions | A26326 |
| | 384 reactions | A26327 |
| Ion Total RNA-Seq Kit v2 | 12 reactions | 4475936 |
| | 48 reactions | 4479789 |
| Ion Total RNA-Seq Kit for AB Library Builder System | 13 reactions | 4482416 |
| Ion 550 Kit-Chef | 8 reactions | A34541 |
| Ion 550 Chip Kit | 4 chips | A34537 |
| | 8 chips | A34538 |

* Ion AmpliSeq Transcriptome Mouse Gene Expression Core Panel and Ion AmpliSeq Library Plus Kit.

** Ion AmpliSeq Transcriptome Mouse Gene Expression Chef-Ready Core Panel and Ion AmpliSeq Kit for Chef DL8.

### Reference


1. Li W et al. (2015) Comprehensive evaluation of AmpliSeq transcriptome, a novel targeted whole transcriptome RNA sequencing methodology for global gene expression analysis. *BMC Genomics* 16:1069.


Find out more at
**thermofisher.com/ampliseqtranscriptome**

ThermoFisher SCIENTIFIC