Concordance of transcriptome sequencing, microarrays, and qPCR using Ion AmpliSeq Transcriptome kits, Clariom D Assays, and TaqMan Assays

Introduction

RNA sequencing (RNA-Seq), gene expression microarrays, and qPCR provide comprehensive solutions for gene expression analysis and biomarker discovery, screening/profiling, and verification. RNA-Seq is an elegant solution for discovery, particularly *de novo* discovery, of differentially expressed genes (DEGs). Another industry standard, gene expression microarrays, can also produce well-annotated results for DEGs. Data from these large-scale transcriptomics technologies are routinely confirmed by qPCR.

The different RNA analysis platforms measure absolute quantities of transcripts by using very different methods and algorithms, making comparisons of absolute expression levels complicated. However, comparisons of transcript-level differences should correlate when the different platforms are compared. Since qPCR produces relative gene expression measurements, comparing gene expression differences between samples with qPCR is the most relevant approach for benchmarking RNA-Seq and gene expression microarrays.

In this technical note, we show high correlation of data for DEGs between Applied Biosystems™ TaqMan® Gene Expression Assays, Applied Biosystems™ Clariom™ D Assays, and the Ion AmpliSeq™ Transcriptome Human Gene Expression Kit. Thus, results across these technologies are comparable. Furthermore, we show that "best coverage" TaqMan Gene Expression Assays can be reliably used to confirm expression data from Ion AmpliSeq Transcriptome kits and Clariom D Assays.

Key observations

- Fold-change measurements using TaqMan Gene Expression Assays have high correlation with data from the Ion AmpliSeq Transcriptome Human Gene Expression Kit and Clariom D Assays.
- Results can be compared across platforms with high correlation and without having to repeat experiments, thus freeing up limitations on the usage of technologies, which can be especially useful in biopharmaceutical settings.

Methods

To compare transcript levels, we used a strategy based on the MicroArray Quality Control (MAQC) study [1,2]. Well-characterized reference RNA transcripts from Stratagene™ Universal Human Reference RNA (UHRR, Cat. No. 740000) and Applied Biosystems™ FirstChoice™ Human Brain Reference RNA (HBRR) were used to compare performance of the different platforms. Fold changes of DEGs in these samples were determined using the Clariom D Assay or Ion AmpliSeq Transcriptome Human Gene Expression Kit, which were then verified by TaqMan Gene Expression Assays.

Results

TaqMan Gene Expression Assays to verify Ion AmpliSeq Transcriptome data

The Ion AmpliSeq Transcriptome Human Gene Expression Kit enables simultaneous measurement of the expression levels of over 20,000 human RefSeq genes in a single assay. We chose a set of 70 TaqMan Gene Expression Assays, including genes spanning a wide range of expression levels, that overlapped with the Ion AmpliSeq profiling experiment. A total of 100 ng of each reference RNA was reverse-transcribed using the Invitrogen™ SuperScript™ VILO™ cDNA Synthesis Kit. The resulting cDNA was analyzed with individual TaqMan Gene Expression Assays using TaqMan Fast Advanced Master Mix, which has been designed for performance superior to standard master mixes and allows for shorter run times of <40 minutes. qPCR data were imported to Thermo Fisher™ Connect and analyzed using the Relative Quantitation (RQ) app. A C_t of 35 was used as a cutoff for qPCR data. Ion AmpliSeq transcriptome libraries were generated for UHRR and HBRR using the Ion AmpliSeq Transcriptome Human Gene Expression Kit, and then were templated for sequencing on the Ion 540™ Chip and Ion S5™ XL System according to supplied protocols [3]. Fold changes in expression between UHRR and HBRR for 70 targets were compared between TaqMan Assays, and the RPKM was generated using the RNA Analysis plugin for Torrent Suite™ Software. The correlation between fold-change measurements made with TaqMan Assays and the Ion AmpliSeq Transcriptome Human Gene Expression Kit is high (Pearson correlation (r^2) of 0.87, Figure 1), despite the different methods used to generate the transcript-level measurements.

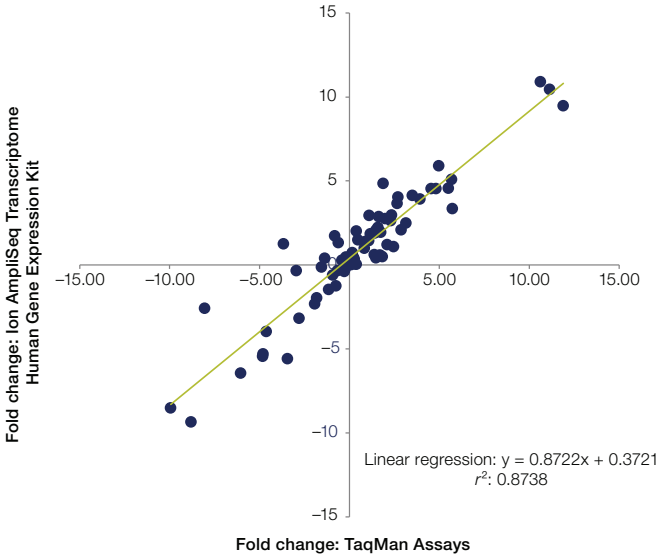

Figure 1. Fold-change correlation between the Ion AmpliSeq Transcriptome Human Gene Expression Kit and TaqMan Gene Expression Assays. The scatter plot shows that the data sets are highly correlated. Ion AmpliSeq data were analyzed on Torrent Suite Software, and qPCR data were analyzed on Thermo Fisher Connect using the RQ app.

TaqMan Gene Expression Assays to verify Clariom D Assay data

The Clariom D Assay enables deep and broad transcriptome analysis and biomarker discovery of both well-annotated and speculative transcripts. When combined with Applied Biosystems™ Transcriptome Analysis Console (TAC) Software 4.0, these microarrays empower scientists to gain information not only about RNA expression at the gene level, but also about RNA alternative splicing at the exon level in different biological samples. Similar to what was done with the Ion AmpliSeq comparison, we used 70 TaqMan Gene Expression Assays that overlapped with Clariom D profiling experiments. Targets from UHRR and HBRR for Clariom D Assays were prepared using the Applied Biosystems™ GeneChip™ WT PLUS Reagent Kit and then hybridized to the Clariom D Array [4]. Figure 2 shows a scatter plot that demonstrates correlation of fold changes between Clariom D and TaqMan Assays. The Pearson correlation (r^2) of 0.89 is indicative of high correlation. We also offer Clariom S Assays that provide a gene-level view of the transcriptome. Since probes in Clariom S Assays are a subset of those in Clariom D Assays, we expect that the gene-level comparison between Clariom S and TaqMan Assays will be similar to the comparison between Clariom D and TaqMan Assays.

Discussion

TaqMan Gene Expression Assays have long been considered the gold standard for studying gene expression, providing a wide dynamic range of greater than 6 logarithmic units of expression levels, high sensitivity, and high specificity. Here we show concordance of qPCR data to transcriptome sequencing and gene expression microarray results for DEGs in two reference RNA samples. The same set of 70 genes was investigated across all three platforms. We found a high degree of concordance between the two comparisons based on differential expression status (i.e., differentially expressed or not differentially expressed). Although the values of the fold-change measurements were not identical in the two comparisons, in general the low fold changes were

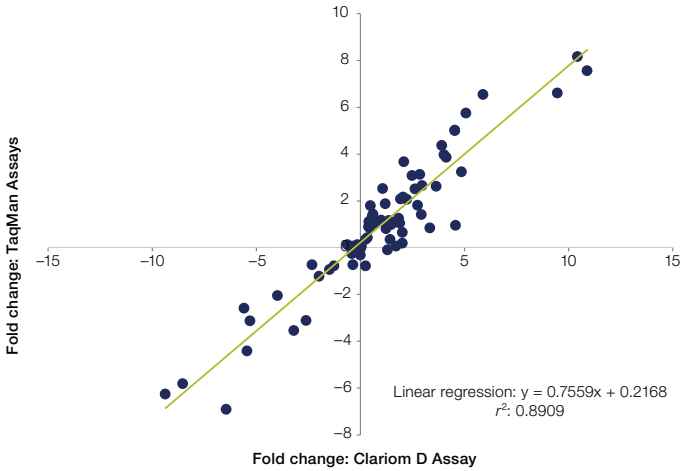

Figure 2. Fold-change correlation between the Clariom D Assay and TaqMan Gene Expression Assays. The scatter plot shows that the data sets are highly correlated. Clariom D Assay data were analyzed on TAC 4.0, and qPCR data were analyzed on Thermo Fisher Connect using the RQ app.

confirmed to be low, and high fold changes were confirmed to be high. This observation confirmed a previous report from Morey et al. that the degree of concordance between microarrays and qPCR increases as fold change increases [5]. Conversely, concordance between microarrays and qPCR dropped among genes with low fold changes (<1.4), due to increased noise [5]. Everaert et al. reported lower concordance between Illumina total RNA-Seq and qPCR data for DEGs when fold change was less than 2 [6]. Nevertheless, we found that genes with low fold changes were consistently low in concordance, even if the absolute value of the magnitude of change varied. Other features such as transcript length, number of exons, splice variants, and read quality may also have decreased the concordance of individual genes in this study, but were not investigated further.

Data presented in this technical note show that TaqMan Gene Expression Assays can be reliably used to verify DEG patterns from transcriptome sequencing and microarrays—our complete offering of gene expression solutions. In general, both transcriptome sequencing and microarray workflows showed high concordance with qPCR data. Thus, it is possible to compare results across these three platforms without having to repeat experiments, which helps to free up limitations on your use of technologies regardless of the number and breadth of expression targets being analyzed. The streamlined workflow of TAC 4.0 facilitates associations between Ion AmpliSeq, microarray, and qPCR data. Gene symbols associated with DEGs from microarray analysis on TAC 4.0 can be readily used to identify "best coverage" TaqMan Gene Expression Assays using the TaqMan online search tool (**thermofisher.com/taqman**). Although not used in this study, TAC 4.0 also enables analysis of Ion AmpliSeq transcriptome panels (**thermofisher.com/ampliseqtranscriptome**) using the plugin available at **apps.thermofisher.com/apps/publiclib/#/plugins** [7]. The ability to analyze data from different platforms makes TAC 4.0 a valuable tool for researchers.

References

1. MAQC Consortium (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* 24:1151–1161.
2. SEQC/MAQC-III Consortium (2014) A comprehensive assessment of RNA-Seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat Biotechnol* 32:903–914.
3. Thermo Fisher Scientific (2016) Application note: Simplified global gene expression profiling. Pub. No. COL21530.
4. GeneChip WT PLUS Reagent Kit. Pub. No. MAN0017939, Rev A.0.
5. Morey JS, Ryan JC, Van Dolah FM (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proceed Online* 8:175–193.
6. Everaert C, Luypaert M, Maag JLV et al. (2017) Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Sci Rep* 7:1559.
7. Ion AmpliSeq Transcriptome Mouse Gene Expression Kit User Guide. Pub. No. MAN0017343, Rev. A.0.

applied
biosystems

Ordering information

Product	Cat. No.
Ion AmpliSeq Transcriptome Human Gene Expression Kit	A26325
Clariom D Assay, human (includes whole-transcriptome Clariom D Array and GeneChip WT PLUS Reagent Kit)	902922
TaqMan Gene Expression Assay, single tube*	4351372
TaqMan Fast Advanced Master Mix	4444963
SuperScript VILO cDNA Synthesis Kit	11754250
FirstChoice Human Brain Reference RNA	AM6050

* For additional TaqMan Assay formats, including Applied Biosystems™ TaqMan® Array Cards and OpenArray™ Plates for screening and verification studies, go to **thermofisher.com/taqman**.

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

ThermoFisher SCIENTIFIC