Abstract No. 1011
Poster No. 2016

Comparison of HTG-Edge Targeted RNA Sequencing Platform with Whole Transcriptome RNA Sequencing for Clinical Biomarker Studies

Dennis O'Rourke, BS [1,3#], Jorge F. Sanchez-Garcia, PhD [1,3#], P. Alexander Rolfe, PhD [2,3], Alice Huang, PhD [1,3], Danyi Wang, PhD [1,3], Juergen Scheuenpflug, PhD [1,4], and Zheng Feng, PhD, MD [1,3*]

1. Global Clinical Biomarkers and Companion Diagnostics, Translational Medicine, Global Development 2. Immunology and Immuno-Oncology Bioinformatics 3. EMD Serono Research and Development Institute, Billerica, MA 4. Merck KGaA, Darmstadt, Germany
Co-first author / Poster presenter * Corresponding author

INTRODUCTION

HTG EdgeSeq Next Generation Sequencing (NGS) technology is a targeted RNA Sequencing (RNA-Seq) platform based on quantitative nuclease protection chemistry, which requires substantially less sample input than other RNA-Seq protocols(one formalin-fixed paraffin-embedded (FFPE) slide versus 4-6 slides). The automated library processor allows fast downstream NGS analysis, making it a promising technology for clinical application and companion diagnostics, especially in samples of long-term storage or with limited quantity of material.

Comparing HTG Edge Technology with traditional RNA-Seq reveals many differences including extraction free chemistry in the HTG workflow, targeted panels for HTG, and differences in downstream analysis of sequencing data such as normalization schemes and QC. Given these differences from standard RNA-Seq protocols and EdgeSeq's clinical implementation potential, it is critical to investigate the performance of HTG EdgeSeq compared to whole transcriptome RNA-Seq technology.

We investigated the concordance & agreement between HTG EdgeSeq and RNA-Seq on a set of 57 FFPE samples. qPCR was used to explore differences which were observed between the two technologies

METHODS

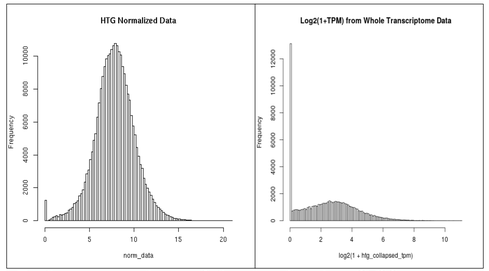
To evaluate the concordance & agreement between the two platforms, 57 FFPE samples from CRC, Gastric, Liver and NSCLC tumors with varying tumor content (16-100%) and sample age (2-24 years) were analyzed. One slide was analyzed per sample with the HTG EdgeSeq Oncology Biomarker Panel covering 2558 genes, while four to six slides were analyzed with whole transcriptome analysis using the Illumina TruSeq Total RNA library kit. A subset of samples and genes were selected for analysis by RT-qPCR using a customized ThermoFisher Taqman 15 gene array. Data from the three platforms was analyzed and compared for concordance & agreement.

RESULTS

Following initial testing 53/57 samples passed defined QC using the HTG Edge platform. In contrast, only 29/57 samples passed QC for traditional RNA-Seq. Although direct comparison of sample failure rate between two platforms could be unfair as there are differences in QC protocols, the difference in the QC-pass rate supports clinical applications of the HTG platform.

Analysis revealed a normal distribution of expression of all genes in the HTG dataset whereas the distribution of expression in the RNA-Seq dataset had a large number of genes with zero or very little expression (Figure 1).

Figure 1: Distribution of gene expression from HTG OBP dataset (Left) and RNA-Seq dataset (Right).



Further comparison of the overlapping 2506 genes in RNA-Seq and the HTG OBP panel showed that 845 of the targets showed low correlation (Spearman correlation less than 0.25), 599 showed moderate correlation (>0.25 but <0.5), 675 showed good correlation (>0.5 but <0.75), and 383 targets showed strong correlation (>0.75) (Figure 2). Genes in the low and moderate correlation grouping tended to have lower dynamic range of expression among the samples resulting in lower concordance (Figure 3).

Figure 2: Distribution of correlation of genes between HTG Edge OGP panel data and RNA-Seq data.

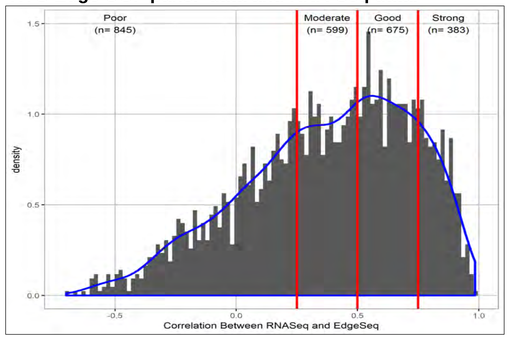
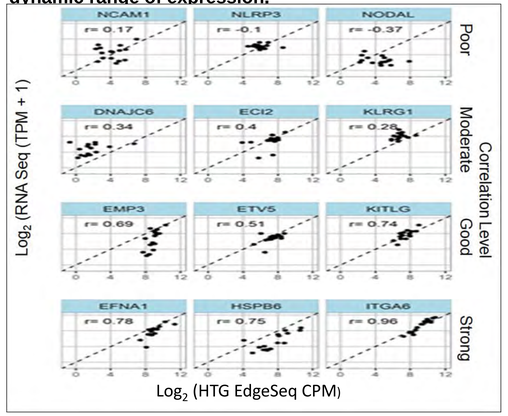
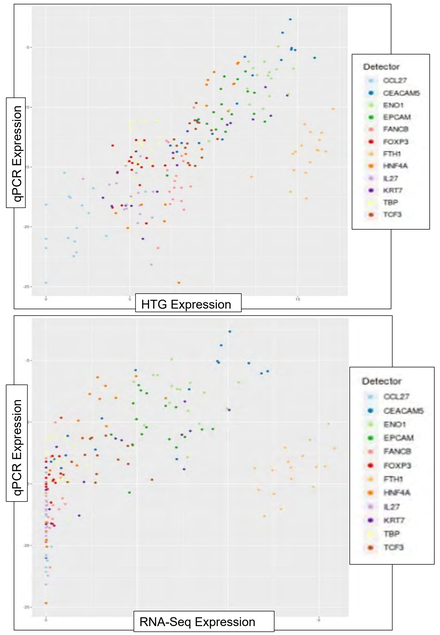


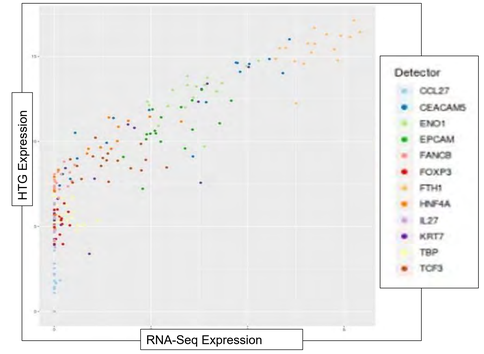
Figure 3: Representation of genes in each correlation category from Figure 2. Low and moderate correlation more often showed low dynamic range whereas those genes in the good and strong category had greater dynamic range of expression.



Following RNA-Seq and HTG EdgeSeq analysis, we further investigated our results using RT-qPCR. A 15-gene ThermoFisher Taqman array (12 target genes and 3 reference genes, see Figure 4 for target gene list) was created to investigate genes that both showed agreement between platforms and gave different results on the platforms. Specifically we aimed to address whether HTG had a superior low-end sensitivity or higher background noise.

Figure 4: Correlation between the three platforms using 12 target genes across 17 samples (Gastric, NSCLC, CRC). Top: qPCR versus HTG, Bottom: qPCR versus RNA-Seq, Top (next column): HTG versus RNA-Seq.





CONCLUSIONS

- HTG EdgeSeq produced fewer QC failures compared to whole transcriptome RNA-Seq while using significantly less tissue input.
- The sound correlation between RNA-Seq, HTG EdgeSeq, and RT-qPCR platforms, **especially for genes with moderate to high expression,** were reported and confirmed.
- Further, these results were consistent across various cancer types.
- The characteristics demonstrated in this study support the potential use of HTG EdgeSeq for NGS based clinical biomarker and companion diagnostics applications.

REFERENCES

Byron SA et al 2016. Translating RNA sequencing into clinical diagnostics: opportunities and challenges. Nat Rev Genet 17(5):257

Costa V et al 2013. RNA-Seq and human complex diseases: recent accomplishments and future perspectives. Eur J Hum Genet 21(2):134

Marioni JC et al 2008. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. Genome Research 18(9):1509

ACKNOWLEDGMENTS AND DISCLOSURES

This study was sponsored by EMD Serono R&D, Inc, a business of Merck KGaA, Darmstadt, Germany

The authors would like to acknowledge HTG Molecular Inc. for their help in generating images and data analysis.