Technical Note

HTG Transcriptome Panel Gene Expression Profiling Data Demonstrate High Correlation Using Illumina and Thermo Fisher Sequencing Platforms

Introduction

HTG Transcriptome Panel (HTP) is the most comprehensive gene expression profiling panel offered by HTG Molecular to date, enabling gene expression analysis of over 19,600 genes. This technical note provides an overview of the HTP process and findings from a comparative experiment using two different sequencing platforms.

The HTP is designed to be run on HTG's EdgeSeq™ platform, a leading extraction-free technology that enables transcriptomic studies, assessing mRNA and miRNA expression from a diverse range of samples such as formalin-fixed, paraffin-embedded (FFPE) tissue (Powles *et al*), core needle biopsies (Girard *et al*), plasma (Wu *et al*), cell lines (Bradbury *et al),* exosomes (García-Flores *et al*), pancreatic cyst fluid (Kane *et al*), cerebrospinal fluid (Reed *et al*) and flow cytometry sorted cells (Szelinski *et al*). HTG's unique chemistry, quantitative nuclease protection assay (qNPA), utilizes a targeted hybridization approach where nuclease protection probes (NPP) are added to the sample lysate allowing for a stoichiometric conversion of the steady-state levels of the target RNAs in the sample to their corresponding NPPs whose levels are then measured using Next Generation Sequencing (NGS).

The EdgeSeq workflow (Figure 1) starts with sample lysate preparation requiring no RNA extraction. The sample lysates then undergo the qNPA reaction on the HTG EdgeSeq Processor, followed by applying molecular barcodes and NGS adaptors via PCR. The libraries are cleaned up, quantified, normalized, and pooled for sequencing. The sequencing counts of the NPP for a gene are a direct representation of its RNA steady-state levels in that sample. A comprehensive suite of software tools, such as the HTG EdgeSeq Parser and Reveal, allow users to align and quantitate the NGS reads for each gene in the panel, as well as carry out a full analysis of differentially expressed genes.

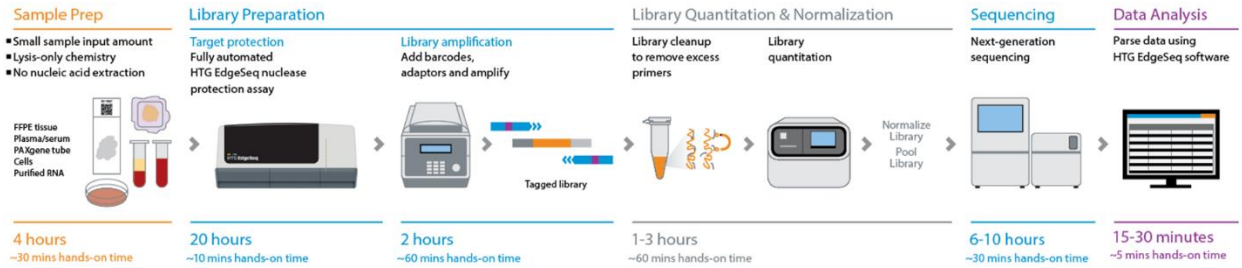

Figure 1: HTG EdgeSeq workflow illustration

The HTP was developed and validated using the Illumina NextSeq 500/550, NextSeq 1000/2000 and NovaSeq 6000 NGS sequencing platforms. Another sequencing platform, the Thermo Fisher Ion Torrent S5, is a cost effective and popular NGS platform that previously had not been run with the HTP. Here we compare the same HTP libraries sequenced with an Illumina NextSeq 550 and a Thermo Fisher Ion Torrent S5.

Methods and Statistical Analyses

Lysates for 89 B-cell lymphoma FFPE samples were prepared as outlined in the HTP User Manual. Following the qNPA run on the EdgeSeq Processor, two different sequencing libraries were prepared from each sample as per the workflow (Figure 1). One library received the sequencing adaptors for the Illumina NextSeq 550, the other for the Thermo Fisher Ion Torrent S5. The libraries were pooled and sequenced on their respective NGS systems at a rate of 12-15 samples per one S5 chip for the Ion Torrent at a loading concentration of 200 pM. For the NextSeq 550, 24 samples per v2 Hight Output flow cell were loaded at 3 pM.

The reads from both the Ion Torrent and NextSeq 550 runs were aligned and quantitated using the HTG Parser software. The data were standardized to $\log_2$-transformed Counts Per Million ($\log_2$CPM) using the HTG Reveal software. All non-control genes were used in the comparison, there was no removal of low expressing genes. Further analysis of the correlations was carried out in R using ggplot2 and LSD libraries.

Results

A total of 89 libraries were run on both the NextSeq 550 and the Ion Torrent S5. The data obtained were then subjected to a series of 4 quality control (QC) checks. Of these 89 libraries, 87 passed all QC criteria which were used in downstream analysis.

The $\log_2$CPM values from the remaining 87 samples were analyzed by a pairwise comparison and the Pearson correlation coefficients calculated for each sample, comparing the Thermo Fisher Ion Torrent S5 results to the results generated on the Illumina NextSeq 550. A high degree of correlation was obtained. Two examples of scatterplots of the respective gene expression profiles are shown in Figure 2.

Overall, the range of Pearson Correlation coefficients was 0.85 to 0.98, with the mean Pearson correlation coefficient across all 87 comparisons being 0.95. These values suggest an exceptionally high level of correlation between the two sequencing platforms.

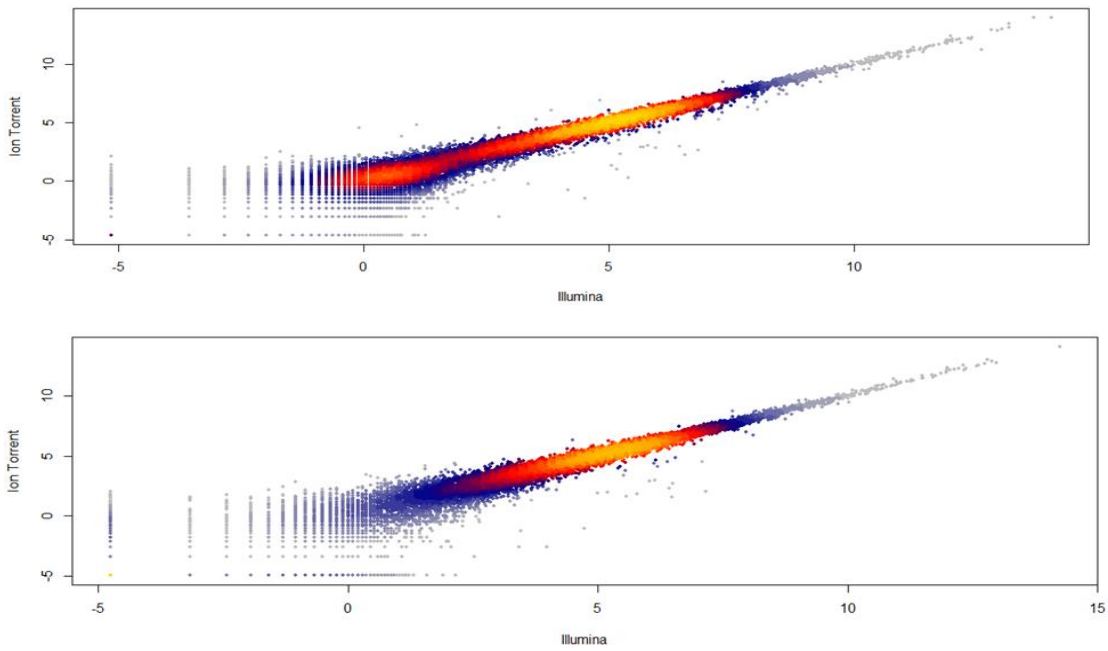

Figure 2: Scatterplots for the log$_2$CPM-transformed gene expression data generated on the Thermo Fisher Ion Torrent S5 sequencer and the Illumina NextSeq 550 sequencer. The color indicates the density of the data in an area with the orange and red colors representing a high density and blue a lower density. The two charts above were selected randomly for illustration purposes.

Conclusions

A consistently high correlation was observed between gene expression profiles generated on the Thermo Fisher Ion Torrent S5 and Illumina NextSeq 550 sequencers. These data demonstrate that HTP libraries generated by the HTG EdgeSeq system can be analyzed with both sequencers while generating remarkably similar results. This enables Ion Torrent users to obtain large scale transcriptomic information across a range of potentially challenging sample types including FFPE.

Acknowledgement

The Early Adopter Program (EAP) was introduced as part of the initial launch of the HTP. The EAP allowed a select group of customers access to the HTP for use in their laboratories or through services provided by HTG prior to the commercial launch of the panel. The data highlighted in this technical note were obtained from the laboratory of Professor Falko Fend at the Institute of Pathology and Neuropathology, University Hospital Tübingen, Germany. We would like to thank Esther Kohler, Dr. Irina Bonzheim, Prof. Falko Fend and everyone involved in producing these results.

References

Bradbury A, O'Donnell R, Drew Y, Curtin NJ, Sharma Saha S. Characterization of Ovarian Cancer Cell Line NIH-OVCAR3 and Implications of Genomic, Transcriptomic, Proteomic and Functional DNA Damage Response Biomarkers for Therapeutic Targeting. *Cancers (Basel)*. 2020;12(7):1939. Published 2020 Jul 17. doi :10.3390/cancers12071939

García-Flores M, Sánchez-López CM, Ramírez-Calvo M, Fernández-Serra A, Marcilla A, López-Guerrero JA. Isolation and characterization of urine microvesicles from prostate cancer patients: different approaches, different visions. *BMC Urol*. 2021;21(1):137. Published 2021 Sep 27. doi:10.1186/s12894-021-00902-8

Girard L, Rodriguez-Canales J, Behrens C, et al. An Expression Signature as an Aid to the Histologic Classification of Non-Small Cell Lung Cancer. *Clin Cancer Res*. 2016;22(19):4880-4889. doi:10.1158/1078-0432.CCR-15-2900

Kane LE, Mellotte GS, Conlon KC, Ryan BM, Maher SG. Multi-Omic Biomarkers as Potential Tools for the Characterization of Pancreatic Cystic Lesions and Cancer: Innovative Patient Data Integration. *Cancers (Basel)*. 2021;13(4):769. Published 2021 Feb 12. doi:10.3390/cancers13040769

Powles T, Yuen KC, Gillessen S, et al. Atezolizumab with enzalutamide versus enzalutamide alone in metastatic castration-resistant prostate cancer: a randomized phase 3 trial. *Nat Med*. 2022;28(1):144-153. doi:10.1038/s41591-021-01600-6

Reed ER, Latourelle JC, Bockholt JH, et al. MicroRNAs in CSF as prodromal biomarkers for Huntington disease in the PREDICT-HD study. *Neurology*. 2018;90(4):e264-e272. doi:10.1212/WNL.0000000000004844

Szelinski F, Stefanski AL, Schrezenmeier E, et al. Antigen-experienced CXCR5^{-} CD19^{low} B cells are plasmablast precursors expanded in SLE [published online ahead of print, 2022 May 4]. *Arthritis Rheumatol*. 2022;10.1002/art.42157.

Wu JW, Yaqub A, Ma Y, et al. Biological age in healthy elderly predicts aging-related diseases including dementia. *Sci Rep*. 2021;11(1):15929. Published 2021 Aug 5. doi:10.1038/s41598-021-95425-5

C-130-06.22.R0