Empowering rare disease cohort biomarker discovery via comparative assessments of gene expression analysis platforms for FFPE pediatric brain tumor specimens

Mateusz Koptyra[1], Namrata Choudhari[1], Zhe Zhang[2], Yuankun Zhu[1], Jennifer Mason[1], Mariarita Santi[2], Phillip B Storm[1,3], Angela J Waanders[1,2], Adam C Resnick[1]

[1]Center for Data Driven Discoveries in Biomedicine, Department of Neurosurgery, Children's Hospital of Philadelphia, Philadelphia, PA, [2]Division of Oncology and Center for Childhood Cancer Research, Children's Hospital of Philadelphia, Philadelphia, PA· [3]Division of Neurosurgery, Children's Hospital of Philadelphia, Philadelphia, PA, [4] Department of Anatomic Pathology, Children's Hospital of Philadelphia, PA

Abstract

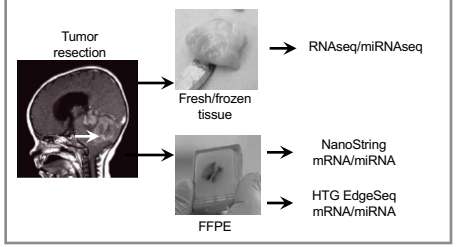
The gene expression landscape for biomarker discovery is still limited for many pediatric brain tumors due to insufficient numbers of biorepository collected fresh frozen specimens. The majority of available tissue are formalin-fixed paraffin embedded (FFPE) pathology diagnostic specimens. These specimens are often of limited quantity and contain compromised RNA material. A number of emerging commercial platforms are described as supporting quantitative expression analysis for low quantity and poor quality materials. Utilizing available platforms, we designed a study to evaluate mRNA and miRNA levels in pediatric brain tumor FFPE specimens with limited/compromised access.

Experiments were performed with specimens and/or data obtained from the Children's Brain Tumor Tissue Consortium (CBTTC) at The Children's Hospital of Philadelphia (CHOP). mRNA gene expression and miRNA target analysis were performed with FFPE material utilizing two commercial platforms: HTG EdgeSeq and Nanostring. The analysis included 5 specimens of high grade glioma or primitive neuroectodermal tumors for mRNA gene expression and 4 specimens of medulloblastoma for miRNA target analysis. FFPE specimens were processed according to manufacturer's protocols. The results were compared with RNAseq or miRNA sequencing data derived from the corresponding tumor's flash frozen specimen.

We evaluated requirements of each platform for data generation and established between platform analysis correlations. The HTG platform was found to a require lower amount of specimen material than NanoString or RNAseq/miRNAseq platforms. For the majority of analyzed genes (>700 genes), the gene expression profile was relatively similar between all three approaches, however each of the platform presented distinctive distribution profile for normalized data. The overall correlation between mean read counts was higher for RNAseq compared to NanoString (0.81) or HTG (0.7) platforms. The RNAseq presented significantly higher variance distribution than other platforms. The miRNA target analysis (>600 genes) distribution of normalized data revealed significantly lower dynamics for NanoString when compared with miRNAseq or HTG panel data. The overall correlation between mean read counts was higher for miRNAseq when compared to the HTG (0.77) or NanoString (0.22) platforms, while miRNASeq presented the highest variance distribution.

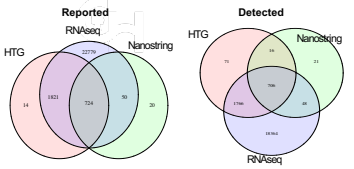
In summary, we found a significant level of agreement between all three platforms tested for gene expression data generation. As for miRNA target analysis, the HTG platform presented significantly higher agreement with miRNAseq data. We conclude that these tested technologies can support data generation from archived FFPE specimens, however, data quality analysis should be performed prior to final platform selection. Sample size requirements, panel gene selection cohort, and pricing evaluation should also be taken into consideration.

Experimental approach

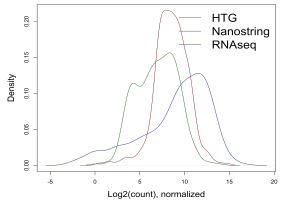


mRNA gene expression analysis

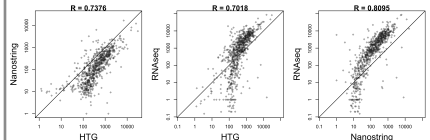
High-grade glioma (HGG) and primitive neuro-ectodermal tumors (PNET) fresh tissue and FFPE specimens were used in the experiment. RNAseq from the fresh tumor RNA was performed by BGI or Nanthealth (data obtained from CBTTC). HTG EdgeSeq FFPE analysis was run using mRNA oncology biomarker panel (OBP panel, 2560 genes). Nanostring FFPE mRNA analysis was performed using NanoString nCounter® PanCancer Pathways panel (770 genes).



Number of mRNA transcripts measured by each platform. The left panel shows the overlap among mRNA transcripts reported by the three platforms. The right panel shows the overlapping of mRNA transcripts with detected expression in five brain tumors measured by three platforms.



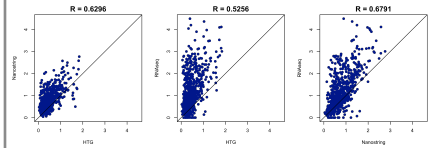
Range and distribution of normalized data from five tumors measured by three platforms.



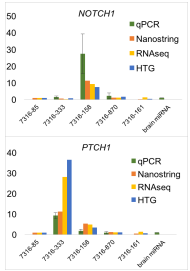
Comparison of mean read counts for 5 tumors between each platform.

HTG	Nanostring	RNAseq
GNAS	*SPP1*	*FN1*
UBB	*UBB*	*GNAS*
FN1	*H3F3C*	*SPP1*
COL1A1	*H3F3A*	*COL1A1*
H3F3A	*FN1*	*FOS*
SPP1	*GNAS*	*UBB*
HSPA1A	*CCND2*	*FLNA*
HSP90B1	*MCM7*	*NOTCH1*
TNC	*AKT3*	*TNC*
RAC1	*HIST1H3H*	*JUN*

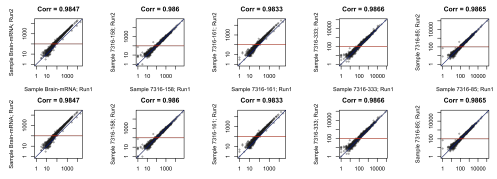
Top 10 genes with the highest total read count for each platform tested.



Comparison of standard deviation for 5 tumors between each platform.



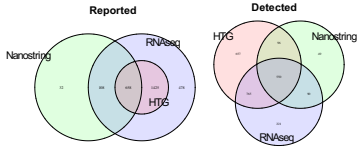
Cross validation of selected genes expression in FFPE mRNA with qPCR. Graphs present relative expression levels (qPCR) and normalized read counts (RNAseq, HTG, Nanostring).



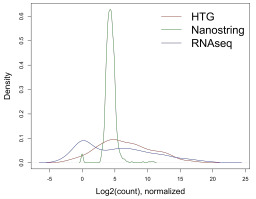
Comparison of read counts for technical replicates between 2 runs for all FFPE tumor specimens tested. Upper row: HTG EdgeSeq; Lower row: Nanostring.

miRNA expression analysis

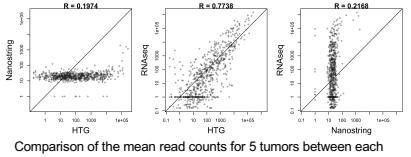
Medulloblastoma fresh frozen tissue and FFPE specimens were used in the experiment. miRNAseq from the fresh tumor was performed by BGI (data obtained from the CBTTC). HTG EdgeSeq FFPE analysis was run using miRNA expression panel including (2083 targets). Nanostring FFPE mRNA analysis was performed using NanoString nCounter® miRNA panel (800 targets).



Number of miRNA targets measured by each platform. The left panel shows the overlapping of all miRNA targets reported by the platforms; the right panel shows the overlapping of miRNA targets with detected expression (read count > 0) in all 4 brain tumors analyzed by the platforms.



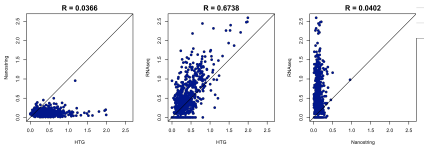
Range and distribution of normalized data from four tumors measured by three platforms.



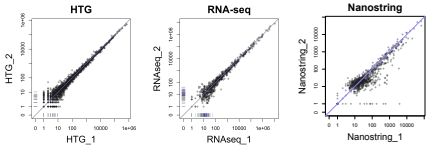
Comparison of the mean read counts for 5 tumors between each platform.

HTG	Nanostring	RNAseq
hsa-miR-125b-5p	hsa-let-7a-5p	hsa-miR-181a-5p
hsa-let-7a-5p	hsa-miR-9-5p	hsa-let-7a-5p
hsa-miR-124-3p	hsa-miR-125b-5p	hsa-let-7f-5p
hsa-let-7c-5p	hsa-let-7b-5p	hsa-miR-9-5p
hsa-miR-26a-5p	hsa-miR-29b-3p	hsa-miR-125b-5p
hsa-miR-99a-5p	hsa-miR-451a	hsa-miR-143-3p
hsa-miR-16-5p	hsa-miR-181a-5p	hsa-miR-26a-5p
hsa-miR-9-5p	hsa-let-7c-5p	hsa-miR-125a-5p
hsa-let-7b-5p	hsa-let-7g-5p	hsa-miR-99b-5p
hsa-miR-204-5p	hsa-miR-26a-5p	hsa-let-7i-5p

Top 10 miRNA with the highest total read count for each platform tested.



Comparison of standard deviation for 5 tumors between each platform.



Comparison of read counts of all miRNAs measured by different runs. Graphs present the scatter plot of read counts for 1 sample.

Ongoing and future plans

- Data validation by qPCR with fresh and FFPR extracted RNA/miRNA
- Pilot study for rare tumor cohort data generation and biomarker discovery (ongoing) with free access to data through online portal (cavatica.org)
- Bouin-fixed specimen evaluation for gene expression analysis to support data generation from archived samples

CAVATICA · CBTTC · D3b Center for Data-Driven Discovery in Biomedicine