Ultra-High-Plex Spatial Proteogenomics of FFPE Tissue Sections

Revealing the tissue architecture of glioblastoma samples with the NanoString GeoMx® Digital Spatial Profiler (DSP) and NGS readout on an Illumina sequencer:

• The GeoMx DSP digitally quantifies hundreds of validated protein targets within selected tissue structures and cell types

• GeoMx spatial proteogenomics enables simultaneous detection and quantification of protein targets and the whole transcriptome

• Illumina sequencing systems radically increase the throughput for spatial profiling

• Integrated software tools like the GeoMx NGS Pipeline with DRAGEN™ Bio-IT platform accessible on BaseSpace™ Sequence Hub provides streamlined data analysis

Introduction

Traditional sequencing methods with dissociated tissue or single cells ignore key spatial information on how cells are organized *in vivo* into tissue and organs. To understand localized expression patterns while maintaining information on tissue architecture, high-plex transcriptomic and proteomic technologies are needed that allow for mapping the spatial location of RNA and protein molecules within tissue sections. Proteomics, the analysis of protein expression patterns in response to specific stimuli or following genomic or transcriptomic changes, is particularly important as translational regulation can lead to a discordance between gene expression and protein expression data. Furthermore, most therapies for disease target a protein rather than the corresponding mRNA transcript. NGS-based proteomic methods are now becoming available and are rapidly complementing transcriptomic-based methods.

The NanoString GeoMx DSP is the first spatial biology platform that measures differential expression changes at the RNA and protein level in formalin-fixed, paraffin-embedded (FFPE), fresh frozen, or fixed frozen tissue sections within selected regions of interest (ROIs) defined by tissue morphology (Figure 1). When coupled with readout from an Illumina next-generation sequencing (NGS) sequencer, tens of thousands of targets can be profiled simultaneously, dramatically increasing the throughput of spatial analysis and discovery power. When assaying spatial proteomics, NGS-based readout on GeoMx DSP allows for profiling hundreds of protein targets alongside the whole transcriptome.

Our previous joint application note described spatial transcriptomics on the GeoMx DSP using Illumina NGS readout. Now, in this follow-up application note, we provide an overview of a comprehensive workflow for spatially resolved proteomics AND simultaneous spatial analysis of RNA and Protein (termed spatial proteogenomics) with the GeoMx DSP, Illumina sequencing and joint data analysis (Figure 2). Simultaneously analyzing gene and protein expression ultimately shows which genes are actively transcribed and translated into proteins, giving a full picture of biological activity in any given sample. The exceptional performance of this spatial method is demonstrated by analyzing human astrocytoma and glioblastoma samples.

Methods
Prepare samples
In this study, 5 µm FFPE core samples from 24 human astrocytoma graded 2-3 and glioblastoma grade 4 were analyzed (BioMax BS17016c). Staining with morphology markers can be performed manually or automated with the BOND RX fully
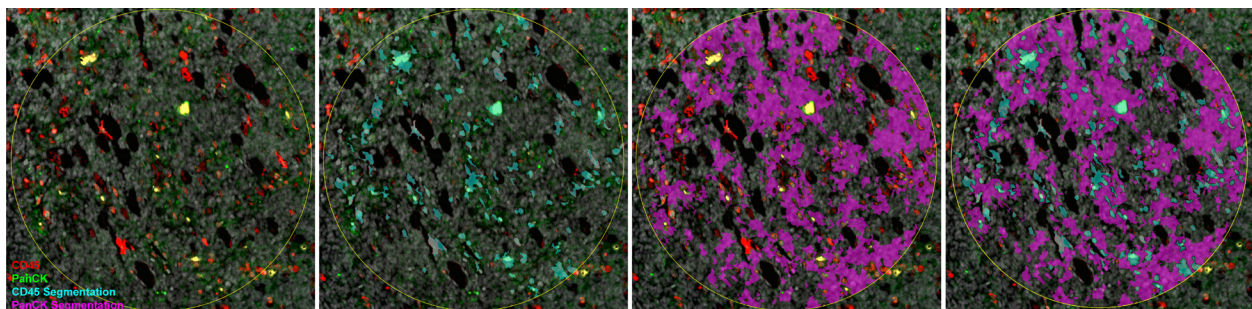

FIGURE 1: Tissue compartmentalization with GeoMx DSP Human Grade 2-3 astrocytoma (n=12, n=20) and Grade 4 glioblastoma (n=16) FFPE cores were stained with fluorescently labeled CD45 (red) and PanCK (green) antibodies. 600 µm regions of interest (ROI) were segmented into immune (CD45+, cyan) and tumor (PanCK+, magenta) areas of interest (AOI).
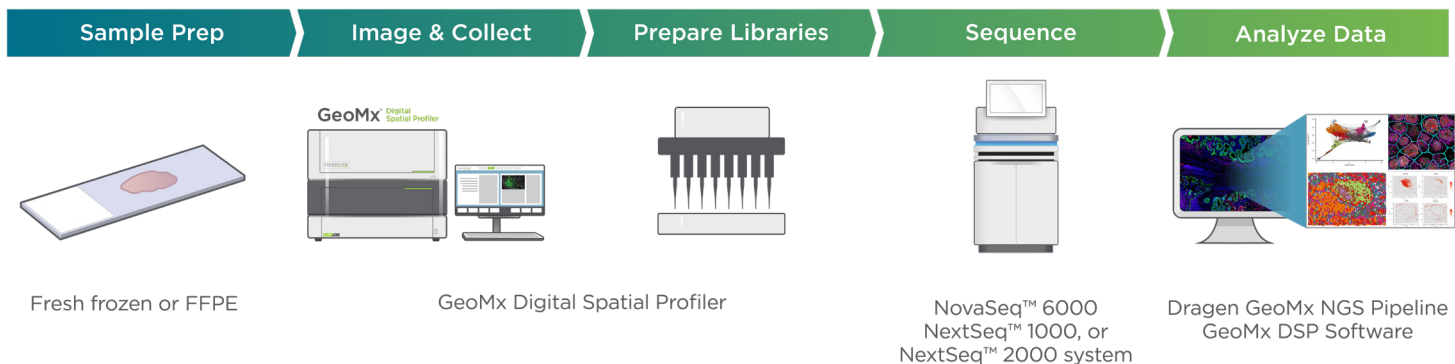
FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.

| Sample Prep | Image & Collect | Prepare Libraries | Sequence | Analyze Data |

Fresh frozen or FFPE | GeoMx Digital Spatial Profiler | | NovaSeq™ 6000 NextSeq™ 1000, or NextSeq™ 2000 system | Dragen GeoMx NGS Pipeline GeoMx DSP Software

FIGURE 2: Spatially resolved transcriptomics workflow The NanoString GeoMx DSP is part of an integrated, sample-to-data workflow for spatially resolved proteomics (Protein) or proteogenomics (Protein + RNA) that includes proven Illumina sequencing on a NovaSeq 6000 or NextSeq 2000 with the DRAGEN GeoMx NGS Pipeline and GeoMx DSP software.

automated research stainer from Leica Biosystems and is adapted from standard histology workflows. Up to four morphology markers can be used for any GeoMx assay and are used to visualize the tissue, to select appropriate regions of interest (ROIs) and segment based on tissue morphology. For this study, fluorescently labeled anti-PanCK (AF532) was used to visualize epithelial cells, anti-CD45 (AF594) was used to visualize immune cells, and anti-Syto13 was used for the nucleus and all three were applied to the tissue array. See guidelines from NanoString for more details on morphology markers and sample preparation.

Image and collect

The GeoMx Whole Transcriptome Atlas (WTA) allows for spatial profiling of every protein-coding transcript in human or mouse samples and GeoMx Protein Assays allow for simultaneous profiling of hundreds of proteins. Both assays are analyzed using an Illumina NGS system for readout. The modular design of GeoMx Protein Assays allows for plex of up to ~150 proteins with the ability to mix and match modules of ~8-10 proteins each to customize coverage for each experiment. Additional custom protein targets can be added to any GeoMx Protein Assay with either NanoString's GeoMx Protein Barcoding Service or Protein Barcoding Kit. All GeoMx Protein Assays undergo extensive validation to ensure high quality data and include content relevant to oncology, immunology, and neuroscience.[1] During the validation process, each antibody is assayed in control tissue(s) pre- and post-conjugation with unique DSP barcodes, tested for sensitivity/specificity and assayed alone or in a high-plex cocktail of multiple antibodies to assess performance.

The GeoMx WTA can be run alone or simultaneously with GeoMx Protein Assays for enabling spatial proteogenomics. In addition, custom GeoMx RNA probes can be made through an RNA Barcoding Service, allowing for custom analysis of up to 200 RNA targets in combination with the GeoMx WTA or as a standalone assay. As a standalone assay, up to two 200 plex custom RNA assays can be combined for spatial transcriptomics of up to 400 RNA targets from any species.

Running a standard single analyte proteomic experiment with GeoMx Protein Assays includes incubating overnight tissue sections mounted on standard microscope slides with a cocktail of highly validated DNA-barcoded antibodies. Running the spatial proteogenomics workflow with the GeoMx WTA and GeoMx Protein Assays adds a day to sample preparation, wherein tissue sections mounted on standard microscope slides are first incubated overnight with a cocktail of DNA-barcoded RNA in-situ hybridization (ISH) probes and then, on the second day, are incubated overnight with DNA-barcoded antibodies (Figure 3). Slides are imaged using the GeoMx DSP and ROIs are selected. ROIs are sequentially exposed to UV light that releases the DNA barcodes and collects them for library preparation and sequencing. ROIs can be further segmented and profiled based on the visualization of morphology markers into multiple Areas of Interest (AOIs), enabling profiling of individual tissue compartments such as the tumor and tumor microenvironment (Figure 1).

Prepare libraries

Libraries are prepared using pre-designed, ready-to-use GeoMx
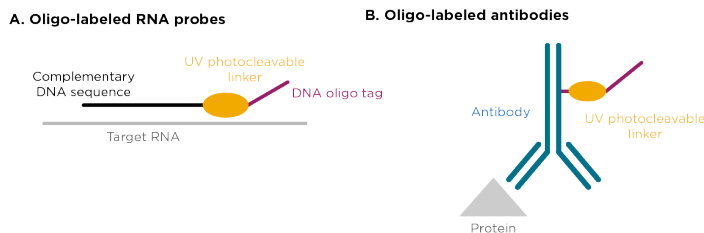

FIGURE 3: GeoMx Assay Profiling Probes GeoMx Assay profiling probes are a cocktail of UV photocleavable oligonucleotides conjugated to (A) in situ hybridization probes for RNA or (B) antibodies for proteins.

Seq Code PCR primers. During library prep, each ROI collected from the tissue section is indexed with a different pair of i7 and i5 unique dual indices. All indexed ROIs are then pooled and purified. The resulting library fragments contain a unique molecular identifier (UMI), target analyte identifier (RTS ID), and the necessary regions for sequencing on Illumina platforms (Figure 4). More details about the libraries can be found here.
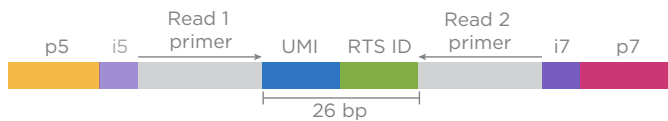

FIGURE 4: GeoMx library Prepared library fragments include a 14-bp UMI for molecular counting, a 12-bp RTS ID for target identification, and i5/i7 unique dual indexes and adaptors compatible with all Illumina sequencing systems.

Sequencing considerations

GeoMx NGS libraries are compatible with any Illumina sequencing system (Figure 2). Higher throughput instruments such as the NextSeq™ 2000 and NovaSeq™ 6000 Systems are recommended to generate the sequencing output required. GeoMx NGS libraries are sequenced with 2 x 27 bp paired-end reads using the

unique dual index workflow with 8 bp for Index 1 (i7) and Index 2 (i5). Library loading concentration by instruments can be found here. A 5% PhiX spike in (v/v) is recommended.

The number of samples (slides) that can be included per sequencing run depends on the sequencing system and ROI/AOI area and number.

As an example, Table 1 shows how to calculate the recommended number of read-pairs for each GeoMx assay based on the total area of illumination and the number of protein targets. Table 2 shows the maximum number of samples that can be run assuming 12 AOIs per slide using the current plex capability of the GeoMx Human Protein Assays (147 plex) for each type of flow cell/sequencing kit combination, depending on the type of Illumina platform used. For example, a common workflow of 12 AOIs per sample with each square AOI measuring 200 µm requires 141M paired-end reads per sample.

Due to the short insert size, P3 50 cycles and NextSeq 2000 and NovaSeq6000 cycles are the most cost-effective options for sequencing; however, GeoMx protein libraires are compatible with all Illumina sequencers and flowcells (see post here)

TABLE 1:

Analyte	Assay	Sequencing Depth Factor	Calculation
RNA	Whole Transcriptome Atlas (WTA)	100	Number of sequencing clusters = total AOI area (µm²) * Sequencing Depth Factor
Protein	Protein (number of antibodies = n)	(2*n)	Number of sequencing clusters = total AOI area (µm²) * Sequencing Depth Factor
RNA + Protein	WTA + Protein (number of antibodies = n)	100 + (3.5*n)	Number of sequencing clusters = total AOI area (µm²) * Sequencing Depth Factor

*Sequencing depth can vary based on inherent experimental variables, including sample quality, tissue type, and expression levels. If the initial sequencing depth is not sufficient, the same library can be re-sequenced and the reads combined to increase sequencing depth.

TABLE 2:

Illumina Platform	No. of NanoString GeoMx Protein Assay Samples (141 Million* Paired-End Reads)	Flow Cell	Sequencing Kit	Catalog No.	No. Of Reads per Flow Cell
NextSeq 2000 System	8	P3	NextSeq 2000 P3 Reagents (50 cycles)	20046810	1200 M
NovaSeq 6000 System	70	S4	NovaSeq 6000 S4 Reagent Kit v1.5 (35 cycles)	20044417	8000-10000 M

*Calculated reads are for the GeoMx Protein Assay single analyte workflow. For information related to the GeoMx Spatial Proteogenomics Workflow, please reach out to NanoString

FOR RESEARCH USE ONLY. Not for use in diagnostic procedures.

nanoString | illumina®

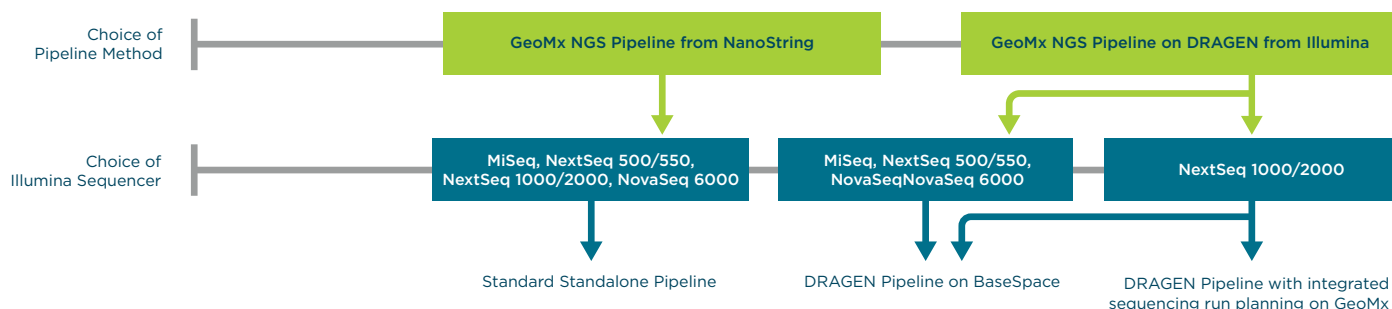## Plan your Sequencing and Pipeline Processing

FIGURE 5: Outline of the NGS analysis pipeline for GeoMx data readout on an Illumina sequencer. Data can be processed via DRAGEN on BaseSpace or with integrated run planning.

Sequencing and analysis setup

Illumina and NanoString have worked together to ensure a seamless end-to-end workflow for processing NGS data from GeoMx DSP experiments (Figure 5). NextSeq 2000 users can now setup a sequencing run directly on the GeoMx instrument (see our blogpost). Analysis can be setup locally on the NextSeq 2000 or on the cloud. This integrated workflow reduces touchpoints and automates secondary analysis. Other users can analyze their data through the DRAGEN GeoMx pipeline available on BaseSpace Sequence Hub or via local analysis on the NextSeq 2000 or NovaSeq 6000 system.

Analyze data

After sequencing, data processing and analysis begins with the GeoMx NGS Pipeline with Illumina DRAGEN, available on BaseSpace Sequence Hub or locally on the NextSeq 2000. The software delivers a reliable and efficient method for converting sequencing reads to counts for spatially resolved expression analysis and visualization using GeoMx software. Digital Count Conversion (DCC) Files generated on the DRAGEN pipeline need to be downloaded and are then processed further on the GeoMx instrument with the DSP Data Analysis software.**

The GeoMx NGS Pipeline with DRAGEN requires two input files: the configuration file from the GeoMx DSP and the FASTQ sequencing file from the Illumina sequencing system. Please see Figure 10 in the GeoMx DSP NGS Readout User Manual for the appropriate GeoMx NGS Pipeline methodology pertinent to the sequencing system used and the Pipeline methodology available. The final output DCC files can be uploaded into the GeoMx DSP Data Analysis Suite for interactive visualization** on the GeoMx instrument. DCC files are also compatible with open-source analysis scripts on GeoScript Hub. For customers who need additional data analysis support, NanoString offers a white-glove Spatial Data Analysis Service to expedite analysis to insight.

**Analysis for the Spatial Proteogenomics Workflow is expected in late 2022

Results

Spatial analysis of tissue structure

Two fluorescently labeled antibodies targeting epithelial cells (PanCK) and immune cells (CD45), and a nucleic acid stain (SYTO-13) were used to identify and discriminate tissue substructures through segmentation. Barcodes from each distinct segmented tissue compartment (CD45+ or PanCK+) were collected independently, yielding two unique proteomic profiles per region, even if the two compartments were adjoining or interdigitated (Figure 1).

Astrocytomas and glioblastomas are diffuse tumors that are difficult to distinguish from neighboring normal tissue. Grade 4 glioblastomas show the most extensive tissue changes from normal brain. Immune segments (CD45+) of Grade 2 and Grade 4 patients show distinct proteomic signatures. Grade 4 glioblastoma immune segments demonstrate increased tumor associated macrophage (TAM) markers CD68, CD14, and anti-inflammatory CD163, and decreased pro-inflammatory TAM CD80 compared to Grade 2 astrocytoma immune segments (Figure 5).[2]

Immune and tumor segments of Grade 3 astrocytomas show distanced proteomic signatures as expected, but also demonstrate the tumor-to-tumor heterogeneity of astrocytomas (Figure 6). For example, EGFR is a prominent treatment target but the heterogeneity of EGFR expression from tumor to tumor has complicated the efficacy of different small molecule inhibitors.[3,4] Ki-67 expression levels may serve as a prognostic marker for survival. Across the tumor samples, Ki-67 expression varies widely, further illustrating the heterogeneity between Grade 3 astrocytomas.[5]

The NEW Spatial Proteogenomic Solution

Spatial proteogenomics offers the advantage of adding proteomic context with the GeoMx WTA with no significant performance loss of the GeoMx protein assay.[6] High concordance (average Pearson correlation coefficient > .90) is observed between matched ROIs assessed with either the GeoMx Protein Assays alone (15 modules, 147 proteins) or the spatial

nanoString | illumina®

FIGURE 6: Proteomics of Grade 3 astrocytoma immune vs tumor
Unbiased hierarchical clustering of detectable protein targets (SNR ≥ 3) of Grade 3 immune and tumor segments and sample identity. Grade 3 astrocytoma cores were stained with 15 stacked GeoMx Human Protein Modules for NGS readout (147 plex). Immune cells (CD45+) and tumor cells (PanCK+) were segmented from 600 μm ROIs within Grade 3 astrocytoma core.

FIGURE 7: Matched ROI to ROI concordance between spatially resolved GeoMx proteomics workflow and GeoMx proteogenomics workflow
Serial Grade 2-3 astrocytoma and Grade 4 glioblastoma cores were stained with 15 stacked GeoMx Human Protein Modules for NGS readout (147 plex). Matched ROIs were assessed for each pair of serial core sections. Pearson R of each matched ROI to ROI detectable proteomic data (SNR ≥ 3) was calculated.
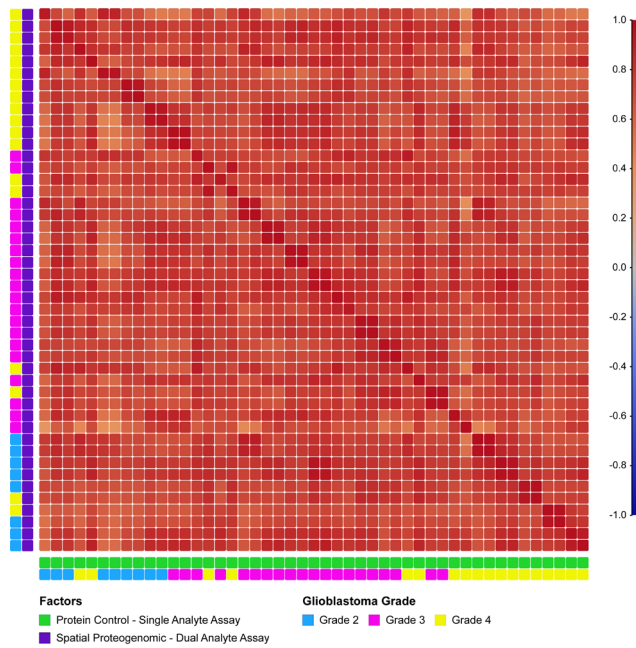

FIGURE 8: Combined volcano plot of Protein and RNA expression Grade 2-3 astrocytoma and Grade 4 glioblastoma cores were co-stained with GeoMx Hu WTA and 15 stacked GeoMx Human Protein Modules for NGS readout (147 plex) using the GeoMx Spatial Proteogenomic workflow. All immune segments were compared to all tumor segments for detectable Protein (SNR ≥ 3) and RNA (SNR ≥ 4) targets. A subset of differentially expressed RNA transcripts (blue) and Protein (magenta) targets are labeled for comparison.

nanoString | illumina®
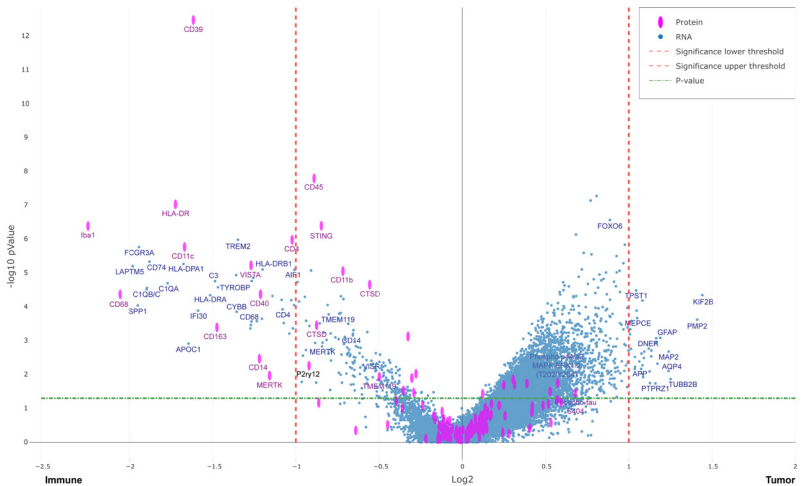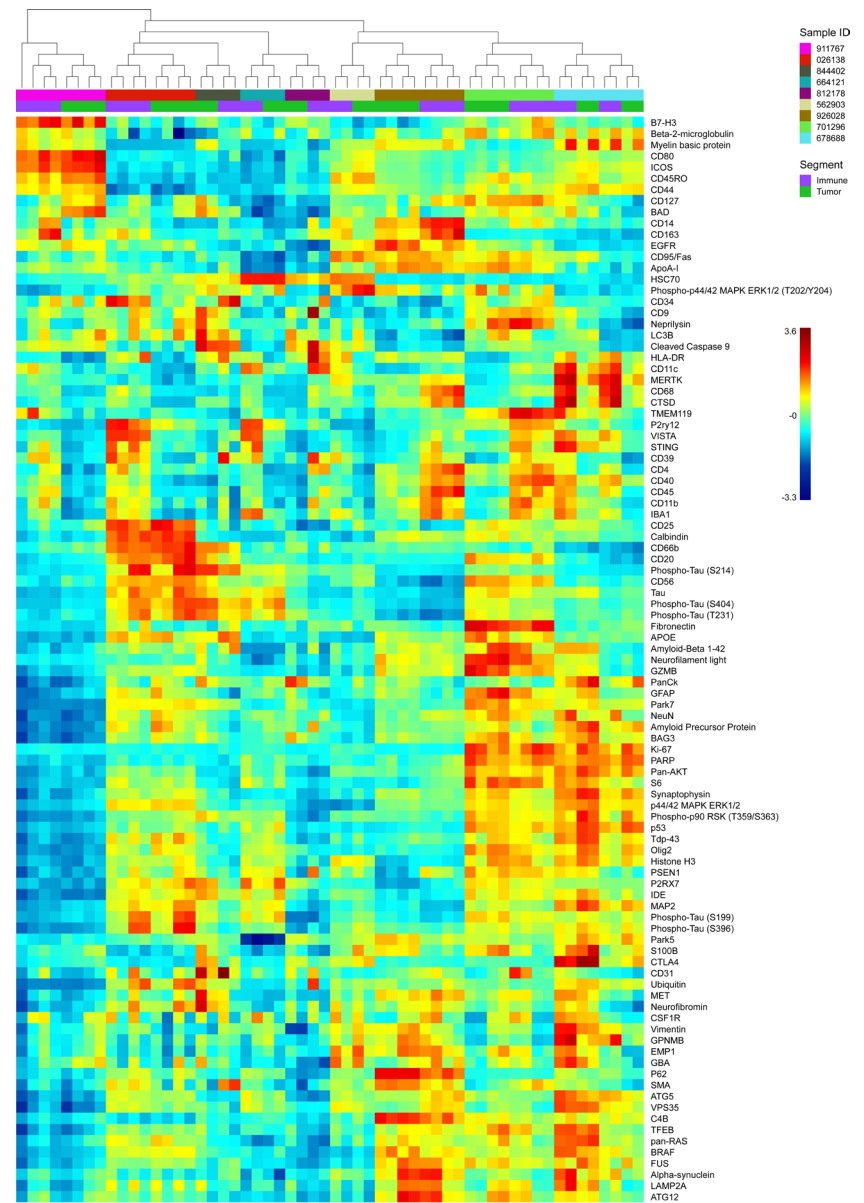
proteogenomic workflow with both the GeoMx Human WTA and Protein Assays run across all tumor grades, patient samples, and tumor/immune segments (Figure 7).

Differential expression of both RNA and protein targets was assessed (Figure 8). Expected expression of RNA and proteins representing immune and tumor targets were seen in their respective segments but differences in the magnitude of expression for either RNA or protein was also observed. Given the fact that many cellular processes regulate translation and proteins often undergo degradation, this lack of concordance is not surprising and is instead biologically informative of cellular processes. Spatial proteogenomics, the ability to simultaneously profile RNA and protein from the same sample, therefore, becomes a powerful tool to give spatial and biological context to healthy and diseased tissue.

Summary

Technological advances have enabled new NGS methods that continue to provide insights into the biology behind health and disease. However, analyses performed on dissociated tissues ignore key spatial context present *in vivo.* The NanoString GeoMx DSP combined with proven Illumina sequencing readout provides high-plex spatial multiomic information on tissue architecture and unique cell populations. This application note outlines new integrated workflows for spatially resolved proteomics and proteogenomics, demonstrating the exceptional performance of these solutions for profiling the heterogeneous pathology of astrocytoma and glioblastoma samples.

Getting help for your experiment

Beyond simply enabling higher throughput by using NGS as a readout, teams from Illumina and NanoString are working together to minimize data analysis time and streamline workflows to improve the customer experience.

For assistance with GeoMx assays or data analysis with the GeoMx NGS pipeline with DRAGEN App on BaseSpace Sequence Hub, contact geomxsupport@nanostring.com For answers to sequencing questions, contact techsupport@illumina.com.

References

1. Rosenbloom, A., Demirkan, G., et al. (2020) Multi-step Antibody Validation for the GeoMx Digital Spatial Profiler. https://nanostring.com/resources/geomx-antibody-validation-whitepaper/

2. Landry, A.P., Balas, M., et al. (2020) Distinct regional ontogeny and activation of tumor associated macrophages in human glioblastoma. Sci Rep 10: 19542.

3. Eskilsson E., Røsland G.V., et al. (2018) EGFR heterogeneity and implications for therapeutic intervention in glioblastoma. Neuro Oncology 20(6): 743-752.

4. Walentynowicz, K.A., Engelhardt, D., et al. (2021) Single-cell genetic heterogeneity linked to immune infiltration in glioblastoma. Preprint at bioRxiv 2021.09.20.461080.

5. Lu, Wc., Xie, H., et al. (2020) Identification of potential biomarkers and candidate small molecule drugs in glioblastoma. Cancer Cell Int 20: 419.

6. Bonnett, S.A., Ong, G., et al. (2020) The Development and Performance of a GeoMx® Based Proteogenomic Workflow for the Detection of Both RNA and Protein on a Single FFPE Slide. https://nanostring.com/resources/the-development-and-performance-of-a-geomx-based-proteogenomic-workflow-for-the-detection-of-both-rna-and-protein-on-a-single-ffpe-slide-white-paper/

Learn more

NovaSeq 6000 System, www.illumina.com/systems/sequencing-platforms/novaseq.html

NextSeq 1000 and NextSeq 2000 Systems, www.illumina.com/systems/sequencing-platforms/nextseq-1000-2000.html

NanoString GeoMx DSP, https://nanostring.com/products/geomx-digital-spatial-profiler/geomx-dsp-overview/

NanoString GeoMx Protein Assays, https://nanostring.com/products/geomx-digital-spatial-profiler/geomx-protein-assays/

NanoString GeoMx Spatial Proteogenomics Workflow, https://nanostring.com/research-focus/spatial-proteogenomics/

For more information, please visit nanostring.com