

High-plex spatial proteogenomics of FFPE tissue sections

Revealing the tissue architecture of astrocytoma and glioblastoma

- The NanoString GeoMx[®] Digital Spatial Profiler (DSP) digitally quantifies protein targets and transcripts *in situ* simultaneously
- Illumina sequencing systems radically increase the throughput for spatial methods
- Integrated software tools available on DRAGEN[™] Bio-IT Platform and in BaseSpace[™] Sequence Hub provide streamlined data analysis

Spatial proteogenomics powered by

nanoString[®]

illumina[®]

Introduction

Traditional sequencing methods using dissociated samples can lose key spatial information present *in vivo*. To understand localized expression patterns while maintaining information on tissue architecture, high-plex spatial technologies are needed for gene and protein profiling. Next-generation sequencing (NGS)-based proteomics methods can complement transcriptome profiling, as translational regulation can lead to differences between gene and protein expression.

The NanoString GeoMx DSP measures differential expression changes at the RNA and protein level within selected regions of interest (ROIs) defined by tissue morphology. When coupled with readout from an Illumina NGS system, 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.

A [previous application note described spatial transcriptomics](#) on the GeoMx DSP using Illumina NGS readout. This application note provides an overview of a comprehensive workflow for spatial proteogenomics, which consists of simultaneous spatial analysis of RNA and protein with the Nanostring GeoMx DSP, Illumina sequencing, and data analysis ([Figure 1](#)). This workflow links insights into gene transcription with translation into proteins, giving a full picture of biological activity in a given sample. The exceptional performance of this spatial method is demonstrated by analyzing human astrocytoma and glioblastoma samples.

Methods

Prepare samples

The NanoString GeoMx DSP is compatible with fresh frozen, fixed frozen, and formalin-fixed, paraffin-embedded (FFPE) samples. In this study, 5 µm FFPE core samples from 24 human astrocytoma (graded 2-3) and glioblastoma (grade 4) were obtained for analysis (US BioMax, Inc. Catalog No. BS17016c). Staining with morphology markers can be performed manually or with the BOND RX Fully Automated Research Stainer (Leica Biosystems) and is adapted from standard histology protocols.

Image and collect

[GeoMx Whole Transcriptome Atlas \(WTA\)](#) and [GeoMx Protein Assays](#) provide a cocktail of GeoMx RNA *in situ* hybridization (ISH) probes or protein antibodies, respectively. Both kits include the ability to add custom targets to tailor the assays to experimental needs. The spatial proteogenomics workflow begins with overnight incubation of tissue sections mounted on standard microscope slides with a cocktail of DNA-barcoded RNA ISH probes. Slides incubated with barcoded ISH probes need to be processed the next day. They are incubated overnight with DNA-barcoded antibodies ([Figure 2](#)).¹

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

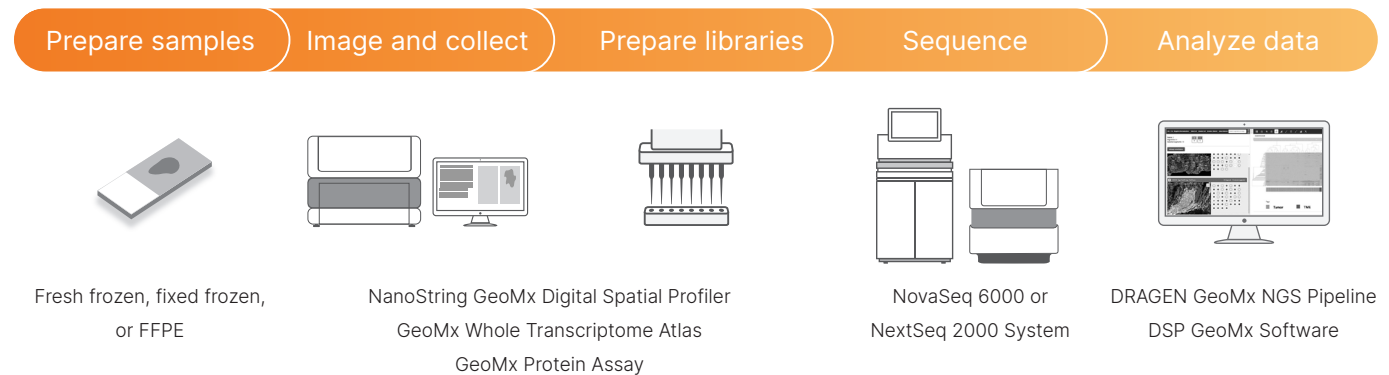
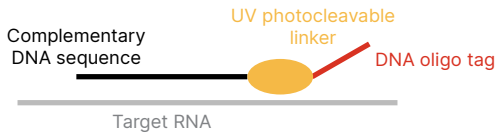


Figure 1: Spatial proteogenomics workflow—The NanoString GeoMx DSP is part of an integrated, sample-to-data workflow for spatially resolved proteomics and transcriptomics that includes proven Illumina sequencing on a NovaSeq 6000 or NextSeq 2000 System and data analysis with the DRAGEN GeoMx NGS Pipeline and GeoMx DSP software.

A. Oligo-labeled RNA probes



B. Oligo-labeled antibodies

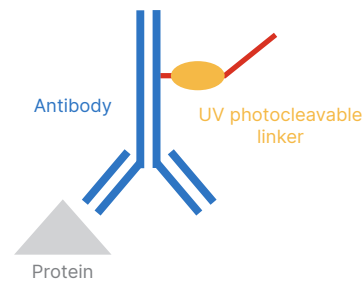


Figure 2: GeoMx imaging probes—GeoMx imaging probes are a cocktail of UV photocleavable oligonucleotides conjugated to (A) ISH probes for RNA or (B) antibodies for proteins.

as the tumor and tumor microenvironment. Up to four morphology markers can be used for any GeoMx assay and are used to visualize the tissue, select appropriate ROIs, and segment based on tissue morphology. Tissue sections were stained with a cocktail of fluorescently labeled antibodies, including anti-PanCK (AF532), anti-CD45 (AF594), and SYTO-13 to visualize epithelial cells, immune cells, and nuclei, respectively.

Prepare libraries

Libraries are prepared using GeoMx Seq Code PCR primers. During the library preparation step, each ROI collected from the tissue section is indexed with a different pair of i7 and i5 unique dual indexes. 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 3).

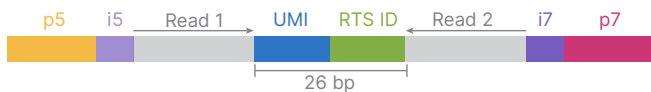


Figure 3: GeoMx library—Prepared library fragments include a 14-bp UMI for molecular counting, a 12-bp RTS ID for target identification, Read 1 and Read 2 primer binding sites, and i5/i7 unique dual indexes and adapters compatible with all Illumina sequencing systems.

Sequence

GeoMx NGS libraries are compatible with any Illumina sequencing system. Higher throughput instruments such as the NovaSeq™ 6000 and NextSeq™ 2000 Systems are recommended to generate the sequencing output required.* An integrated, end-to-end workflow is available enabling users to set up a sequencing run on the GeoMx DSP for the NextSeq 2000 System. This workflow reduces touch points and automates secondary analysis.

GeoMx NGS libraries are sequenced with 2 × 27 bp paired-end reads using the unique dual index workflow with 8 bp for Index 1 (i7) and Index 2 (i5). A 5% PhiX spike in (v/v) is recommended. The recommended number of read pairs for each GeoMx assay can be calculated based on the total area of illumination and the number of protein targets (n) (Table 1).

Table 1: Guidelines for read pairs for GeoMx assays

| Analyte | Assay | Sequencing depth factor | Calculation |
|---------------|---------------|-------------------------|---|
| RNA | WTA | 100 | Number of sequencing clusters = total AOI area (µm ²) * sequencing depth factor |
| Protein | Protein | (2 × n) | |
| RNA + protein | WTA + Protein | 100 + (3.5 × n) | |

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 resequenced and the reads combined to increase sequencing depth.

The maximum number of samples (slides) that can be included per sequencing run, assuming 12 AOIs per slide and a plexity of 147 protein targets, varies depending on the sequencing system (Table 2). For example, a common workflow of 12 AOIs per sample with each square AOI measuring 200 µm × 200 µm requires 282M paired-end reads per sample.†

* Due to the short insert size, sequencing with the NextSeq 2000 P3 (50 cycles) and NovaSeq S4 Reagent Kit v1.5 (35 cycles) are the most cost-effective options.

† The paired-end read strategy recommended by NanoString results in two paired-end reads per cluster: forward and reverse; this can also be described as one read pair.

Table 2: Example sample throughput for protein assay single analyte GeoMx libraries on Illumina sequencing systems

| Illumina platform | No. of NanoString GeoMx Protein Assay samples ^a (141M read pairs ^b) | Flow cell | Sequencing kit | Catalog no. | No. of transcripts per flow cell |
|---------------------|--|-----------|--|-------------|----------------------------------|
| NextSeq 2000 System | 8 | P3 | NextSeq 2000 P3 Reagents (50 cycles) | 20046810 | 1200M |
| NovaSeq 6000 System | 70 | S4 | NovaSeq 6000 S4 Reagent Kit v1.5 (35 cycles) | 20044417 | 8-10B |

a. Calculated reads are for the GeoMx Protein Assay single analyte workflow; reach out to NanoString for information related to the GeoMx Spatial Proteogenomics Workflow.
 b. For the NanoString GeoMx Protein Assay, 141 read pairs equates to 141M reads for sequencing read 1 and 141M reads for sequencing read 2 (282M paired-end reads).

Analyze data

After sequencing, data processing and analysis begins with the GeoMx NGS Pipeline, available in different formats depending on lab infrastructure. Options include a stand-alone pipeline from NanoString or on the DRAGEN Platform from Illumina, available on BaseSpace Sequence Hub or locally on the NextSeq 2000 System (Figure 4). DRAGEN software delivers a reliable and efficient method for converting sequencing reads to counts for spatially resolved expression analysis and visualization using GeoMx software. Two input files are required: the configuration file from the NanoString GeoMx DSP and the FASTQ sequencing file from the Illumina sequencing system. Digital Count Conversion (DCC) Files generated on the DRAGEN pipeline are processed further on the GeoMx instrument with DSP Data Analysis software.‡ The final output DCC files can be

‡ Analysis for the GeoMx Spatial Proteogenomics workflow will be available late 2022.

uploaded into the GeoMx DSP Data Analysis Suite for interactive visualization.‡ DCC files are also compatible with open-source analysis scripts on GeoScript Hub.

Results

Spatial proteomics 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 5).

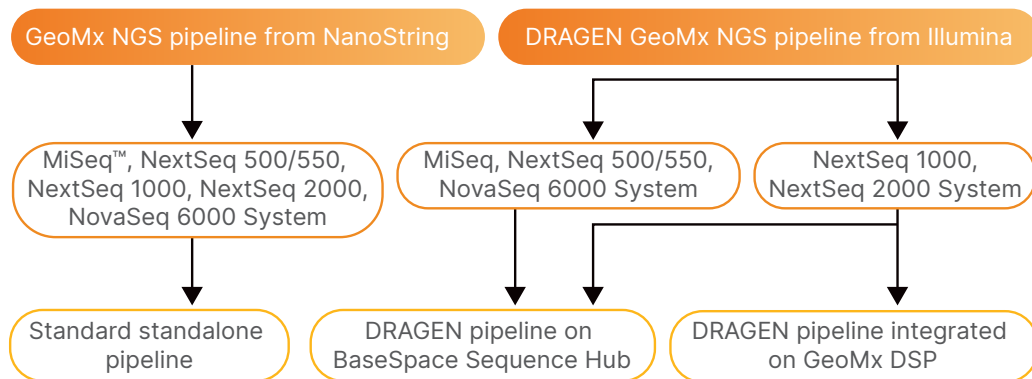


Figure 4: GeoMx NGS Pipeline processing—Customers can choose how to host the GeoMx NGS pipeline. The standard standalone pipeline from NanoString can be hosted on local infrastructure. The GeoMx NGS Pipeline on the DRAGEN Platform from Illumina is available in the cloud on BaseSpace Sequence Hub or locally on the NextSeq 1000 and NextSeq 2000 System.

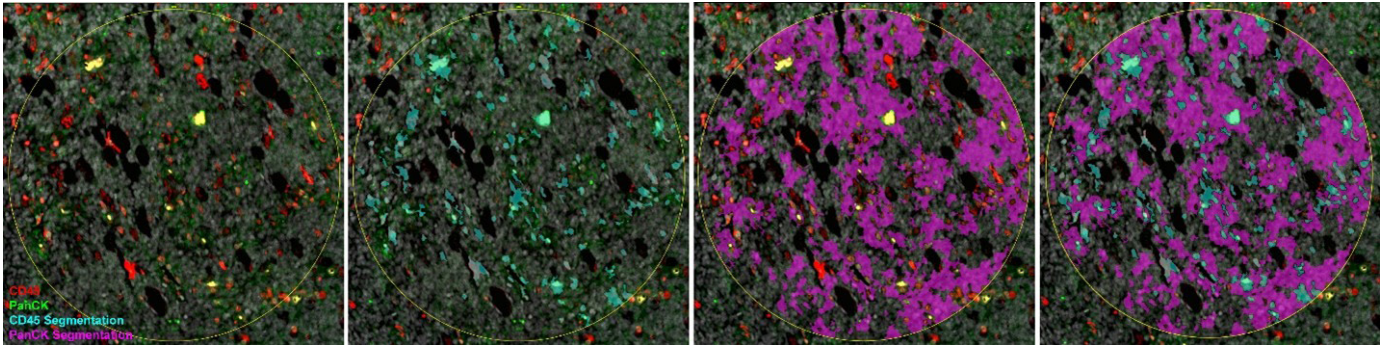


Figure 5: Tissue compartmentalization with the GeoMx DSP—Human grade 2-3 astrocytoma (n=12, n=20) and grade 4 glioblastoma (n=16) FFPE cores were stained with fluorescently labeled CD45 (red) and PanCK (green) antibodies. 600 μ m ROI were segmented into immune (CD45+, cyan) and tumor (PanCK+, magenta) AOI.

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 tumor samples show distinct proteomic signatures. Grade 4 glioblastoma immune segments demonstrate increased tumor-associated macrophage (TAM) markers CD68, CD14, and antiinflammatory CD163, and decreased proinflammatory TAM CD80 compared to grade 2 astrocytoma immune segments (Figure 6).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 7). 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

Spatial proteogenomics reveals tissue complexity

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 > 0.90) is observed between matched ROIs assessed with either the GeoMx Protein Assays alone (15 modules, 147 proteins) or the spatial proteogenomic workflow with both the GeoMx Human WTA and Protein Assays run across all tumor grades, patient samples, and tumor/immune segments (Figure 8).

Differential expression of both RNA and protein targets was assessed (Figure 9). 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, with 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 tissues.

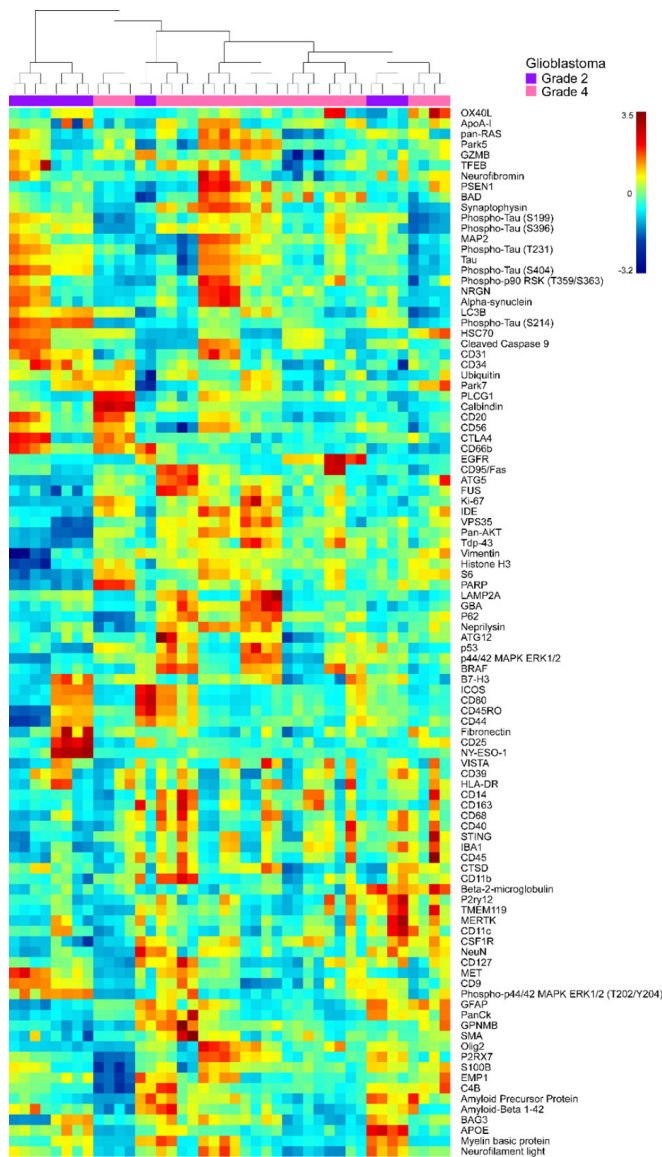


Figure 6: Proteomics of Grade 2 astrocytoma and Grade 4 glioblastoma—Unbiased hierarchical clustering of detectable protein targets (signal-to-noise ratio (SNR) ≥ 3) from Grade 2 and Grade 4 immune segments. Grade 2 astrocytoma and Grade 4 glioblastoma cores were stained with 15 stacked GeoMx Human Protein Modules for NGS readout (147 plex). Immune cells (CD45+) were segmented from 600 μm^2 ROIs within Grade 2 astrocytoma and Grade 4 glioblastoma cores.



Figure 7: Proteomics of Grade 3 astrocytoma immune vs tumor—Unbiased hierarchical clustering 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^2 ROIs within Grade 3 astrocytoma core.

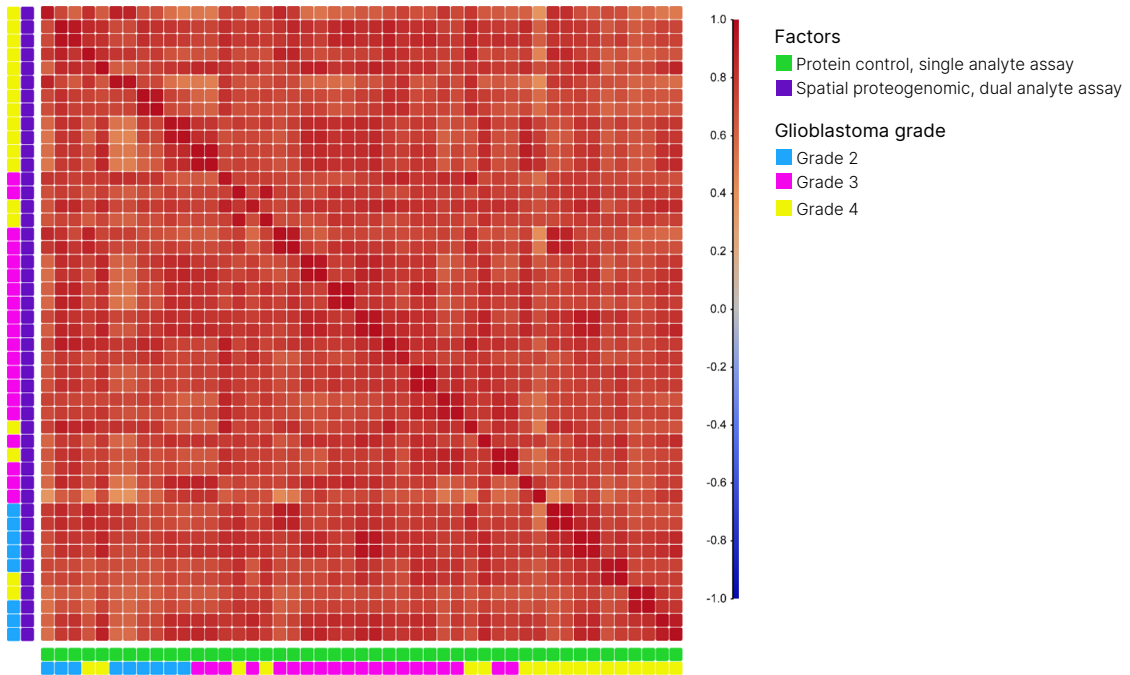


Figure 8: Matched ROI to ROI concordance between spatially resolved GeoMx proteomics workflow and GeoMx proteogenomics workflow—Serial Grade 2-3 astrocytoma and Grade 4 glioblastoma cores were stained with 15 stacked GeoMx Human Protein Modules for NGS readout (147 plex). Matched ROIs were assessed for each pair of serial core sections. Pearson R of each matched ROI to ROI detectable proteomic data (SNR ≥ 3) was calculated.

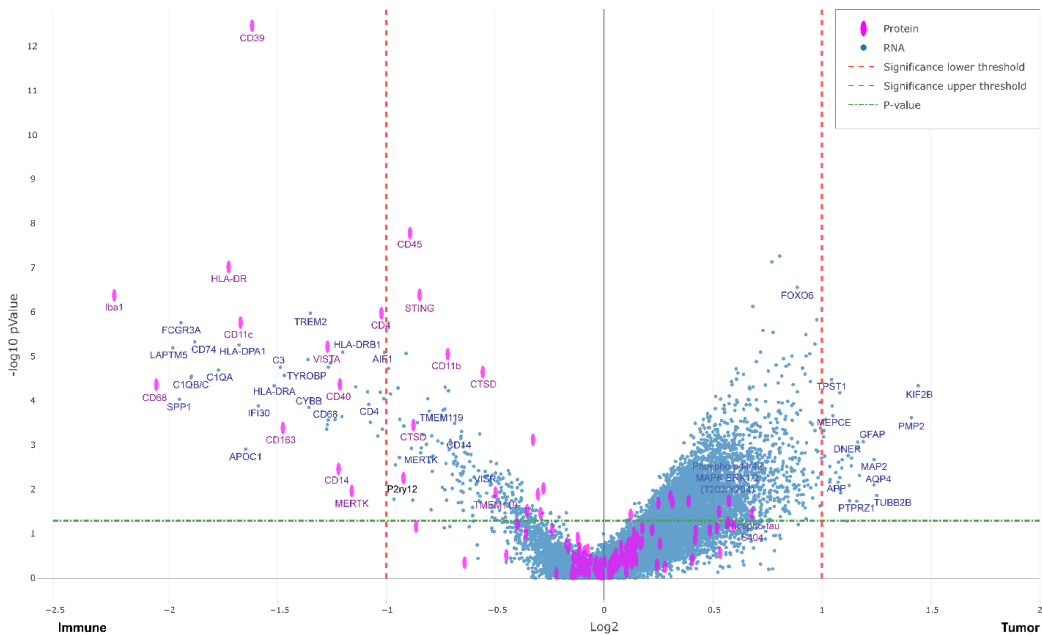


Figure 9: Combined volcano plot of protein and RNA expression—Grade 2-3 astrocytoma and grade 4 glioblastoma cores were costained with GeoMx Human (Hu) WTA and 15 stacked GeoMx Human Protein Modules for NGS readout (147 plex) using the GeoMx Spatial Proteogenomic workflow. All immune segments were compared 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.

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 an integrated workflow 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 have been working together to minimize data analysis times and streamline the workflow to improve the customer experience.

For assistance with the GeoMx assay or data analysis with the GeoMx NGS Pipeline BaseSpace App, contact geomxsupport@nanosttring.com

Send questions regarding sequencing or data analysis with DRAGEN software to: techsupport@illumina.com



1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

© 2022 Illumina, Inc. All rights reserved. All trademarks are the property of Illumina, Inc. or their respective owners. For specific trademark information, see www.illumina.com/company/legal.html.
 M-GL-00968 v1.0

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, www.nanosttring.com/products/geomx-digital-spatial-profiler/geomx-dsp-overview/

NanoString GeoMx WTA, www.nanosttring.com/products/geomx-digital-spatial-profiler/geomx-rna-assays/geomx-whole-transcriptome-atlas/

NanoString GeoMx DSP user guide, nanosttring.com/wp-content/uploads/2022/06/MAN-10153-01-GeoMx-DSP-NGS-Readout-User-Manual.pdf

References

1. Rosenbloom A, Demirkan G, et al. **Multi-step Antibody Validation for the GeoMx Digital Spatial Profiler**. Nanostring. 2020.
2. Landry AP, Balas M, Alli S, Spears J, Zador Z. **Distinct regional ontogeny and activation of tumor-associated macrophages in human glioblastoma**. *Sci Rep*. 2020;10(1):19542. doi: 10.1038/s41598-020-76657-3.
3. Eskilsson E, Røsland GV, Solecki G, et al. **EGFR heterogeneity and implications for therapeutic intervention in glioblastoma**. *Neuro Oncol*. 2018;20(6):743-752. doi: 10.1093/neuonc/nox191.
4. Walentynowicz KA, Engelhardt D, Yadav S, et al. **Single-cell genetic heterogeneity linked to immune infiltration in glioblastoma**. bioRxiv. 2021. doi:10.1101/2021.09.20.461080.
5. Lu WC, Xie H, Yuan C, et al. **Identification of potential biomarkers and candidate small molecule drugs in glioblastoma**. *Cancer Cell Int*. 2020;20:419. doi: 10.1186/s12935-020-01515-1.
6. Bonnett SA, Ong G, Lyssand J, et al. **The Development and Performance of a GeoMx Based Proteogenomic Workflow for the Detection of Both RNA and Protein on a Single FFPE Slide**. Nanostring. 2020.