WHITEPAPER

The Development and Performance of a GeoMx® Based Proteogenomic Workflow for the Detection of Both RNA and Protein on a Single FFPE Slide

Shilah A. Bonnett
Giang Ong
John Lyssand
Gary Geiss
Alyssa Rosenbloom
Joseph Beechem

NanoString Technologies, Inc., Seattle, WA 98109
MK4240 | Jan 2022

nanoString®

GeoMx® Based Proteogenomic Workflow

Abstract

The GeoMx® Digital Spatial Profiler (DSP) enables high-plex, high-throughput spatial profiling and quantification from a single slide for either protein or RNA. To fully understand the interplay between RNA and protein in tissue, we have developed a novel GeoMx Spatial Proteogenomic workflow for NGS readout that allows for the profiling of both analytes from individual area of interest (AOI) on a single slide. Here we describe the development and performance of the proteogenomic workflow on cell pellet array (CPA) and various tissues including non-small cell lung cancer (NSCLC) and colorectal cancer (CRC) using a high plex GeoMx Protein Assay and the GeoMx Human Whole Transcriptome Atlas (GeoMx Hu WTA). We profiled cell lines with known expression profiles and confirmed that the sensitivity and specificity for both analytes under proteogenomic workflow conditions were comparable to the single analyte conditions. Furthermore, the workflow performance was maintained in tissue samples. The high correlation between the performance metrics of the proteogenomic workflow and the single analyte demonstrates successful detection of RNA transcripts and protein from a single FFPE tissue sample. As a result, the Spatial Proteogenomic workflow enables deeper characterization of precious biological samples that are available in limited quantities.

Introduction

The advancement of spatially resolved, multiplex technologies has revolutionized and redefined the approaches to complex biological questions pertaining to tissue heterogeneity, tumor microenvironments, cellular interactions, cellular diversity, and therapeutic response (1). These technologies, including the GeoMx® Digital Spatial Profiler (DSP), have yielded spatially resolved proteomic and transcriptomic datasets from formalin-fixed paraffin-embedded (FFPE) or fresh frozen (FF) samples. Many spatial technologies, though, are specific towards generating either proteomic or transcriptomic datasets. It has been shown, that the correlation between RNA and protein can be poor and often changing depending on the gene and tissue being analyzed (2). To fully capture the biological processes that control transcription, translation, protein turnover and action requires a workflow to accurately measure RNA and protein simultaneously. (Figure 1A) To fully understand the proteogenomic relationship between these

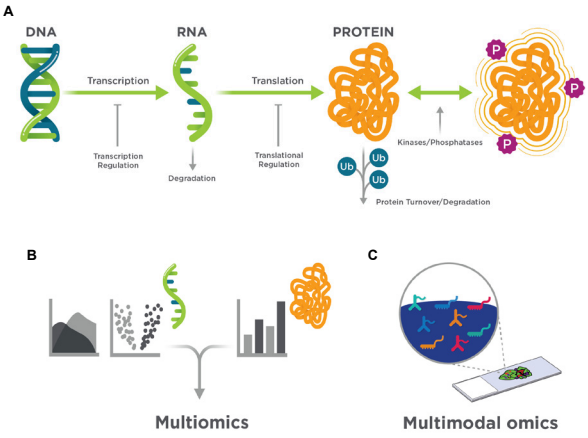

FIGURE 1: (A) Importance of proteogenomic in understanding the relationship between the proteome and either the genome or transcriptome. Current proteogenomic approaches are **(B)** multiomic which entails the integration of individual -omic datasets and **(C)** multimodal which involves the simultaneous, co-detection of multiple 'omes' in a single sample.

various 'omes', the individual datasets need to be integrated from various techniques in a multiomic approach (3-6). (Figure 1B) While this approach provides a deeper understanding of the system understudy, variations stemming from different platforms for RNA and protein, or even on a single platform, section-to-section variability and precisely matching regions of interest (ROIs) across multiple slides needs to be taken into consideration when analyzing and interpreting the data. To gain deeper insight and control for these variables, multimodal omics has been used as an alternative approach, which pertains to the simultaneous, co-detection of multiple 'omes' in a single sample (7, 8). In spatial biology, there has been a growing trend towards the development of multimodal omic workflows with the combination of immunohistochemistry (IHC) or immunofluorescence (IF) methodologies with *in situ* hybridization (ISH) (8-16). (Figure 1C) But to date, multimodal omic workflows with ultrahigh plex analyte co-detection has been lacking.

The GeoMx DSP enables spatially resolved, high-plex digital quantitation of proteins (≥ 100-plex) and RNA (up to 21,000-plex) from FFPE or FF samples (17-20). This technology utilizes unique affinity reagents antibodies for protein or ISH probes for RNA) coupled to UV photocleavable oligonucleotide barcodes. Tissue samples are co-incubated with these affinity reagents

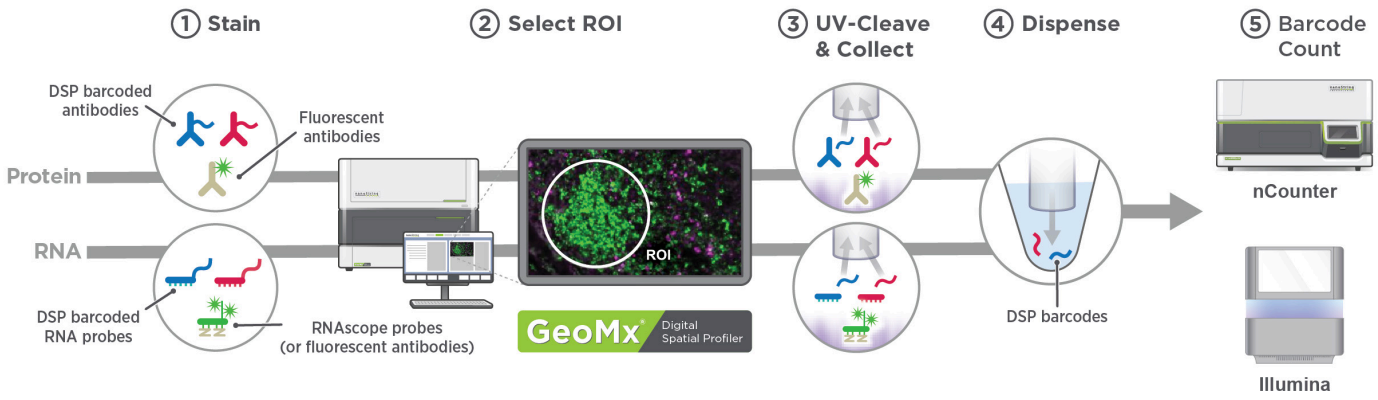

FIGURE 2: Commercially available GeoMx Assays currently enable high-plex, spatially resolved protein and RNA targets on individual tissue sections with nCounter or NGS quantitative readout.

and fluorescent markers, then subsequently imaged using fluorescence microscopy. Oligonucleotide barcodes are then precisely liberated from any area of interest (AOI) with UV-light, collected and quantified with either an nCounter® Platform or Next-Generation Sequencer (NGS). (Figure 2)

Despite the expansive multiplex capabilities of GeoMx DSP, the GeoMx Protein and RNA Assays have been traditionally validated for single analyte detection. Spatial proteogenomic profiling requires two separate FFPE or fresh frozen samples, one to be assayed for each analyte. Integration of a multimodal omic approach into the GeoMx workflow would enable a deeper characterization of precious biological samples that are available in limited quantities. Furthermore, the simultaneous assessment of RNA and protein from a single AOI would reduce technical variation associated with two separate, single analyte workflows. To expand upon GeoMx DSP capabilities, we have developed a novel co-detection workflow for NGS readout that allows for the profiling of both RNA and protein from the area of interest (AOI) on a FFPE tissue section. Here we describe the technical development and performance of the spatial proteogenomic workflow on cell pellet array (CPA) and various tissues using a high plex GeoMx Protein Assay and either the GeoMx Cancer Transcriptome Atlas (GeoMx CTA) or GeoMx Human Whole Transcriptome Atlas (GeoMx Hu WTA).

Experimental Design

FFPE Samples

Sections of formalin-fixed, paraffin-embedded (FFPE) cell pellet arrays (CPA) and tissues, 5 µm in thickness, were used in these studies.

GeoMx Protein Assay

For protein only control, slides were manually processed according to the Protein FFPE Manual Slide Preparation Protocol in the GeoMx NGS Slide Preparation User Manual (MAN-10115-05) and associated published material (17).

GeoMx RNA Assay

RNA only control slides were processed according to the RNA FFPE BOND RX Slide Preparation Protocol in the GeoMx NGS Slide Preparation User Manual for FFPE (MAN-10115-05) and associated published materials (17, 18).

GeoMx Spatial Proteogenomic Workflow Sample Prep

Slides were processed in a similar workflow to those for the RNA control using the Leica Bond-RX system (Leica Biosystems, Melbourne, Australia). Briefly, slides containing FFPE sections were baked, deparaffinized rehydrated in ethanol and washed in Leica BOND Wash Solution. Epitope retrieval was carried out under basic conditions (Tris-EDTA, pH 9.0) for 10 min at 85˚C (cell pellet) or 20 min at 100˚C (tissues). Samples were washed in Bond Wash Solution, digested with 0.1 µg/mL proteinase K (ProK) for 5 min (cell pellet) or 15 min (tissues) at 37˚C. The slides were removed from the Leica and washed one time in 1X PBS for 5 min. Samples were covered with either the GeoMx Cancer Transcriptome Atlas (GeoMx CTA; NanoString) or GeoMx Whole Transcriptome Atlas (GeoMx WTA; NanoString) oligo-conjugated RNA probe set diluted in Buffer R (NanoString), covered with HybriSlip Hybridization Cover (Grace BioLabs) and incubated overnight at 37˚C. The following day, slides were washed two times under stringent conditions to remove unbound probe with 50% formamide in 2X SSC at 37˚C for 25 min each. Slides were

washed two times in 2X SSC at RT for 5 min each and one time in 1X-TBS-T for 5 min. Samples were blocked with Buffer W (NanoString) for 60 min at RT and subsequently incubated with stacked high-plex human GeoMx Protein Assays for NGS readout (oligo-conjugated antibody mix, NanoString) overnight in a closed humidity chamber at 4˚C. For tissue sections, fluorescent markers were added with the antibody cocktail mix for overnight incubation. The following day, slides were washed three times in 1X TBS-T for 10 min each and incubated with 4% PFA for 30 min at RT in a closed humidity chamber. After washing 2 times in 1X TBS-T for 5 min each, sections were counterstained with SYTO 13 for 15 min at RT and then loaded onto the GeoMx DSP instrument platform as described below.

GeoMx DSP Experiments ROI Selection and Collection

GeoMx Digital Spatial Profiling of tissue was carried out according to GeoMx NGS DSP Instrument manual (MAN-10116-05) and as described by Merritt, *et al* (17).

For cell pellet arrays (CPAs), two geometric regions of interest (ROI) of 200 µm in diameter were profiled per cell line. Tissues sections were stained with fluorescent morphology markers to aid in the selection of areas of interest (AOIs), a subsampling of ROI. Human colorectal cancer (CRC) and non-small cell lung cancer (NSCLC) were stained with anti-CD45 (immune) and anti-PanCK (tumor). For NSCLC, circular geometric ROIs of 100 µm in diameter were collected for each marker specific areas of interest (AOI). ROIs were matched across all test slides under study. Advanced ROI selection strategy (segmentation) was implemented on CRC. For segmentation experiments, circular ROIs of 300 µm in diameter for CRC, were segmented into marker specific areas of interest using the DSP auto-segmentation tool.

Next-Generation Sequencing and Data Analysis

Library preparations was carried out according to the GeoMx NGS Readout Library Prep Manual (MAN-10117-05; NanoString) with slight modification. Libraries were sequenced on an Illumina NextSeq2000 or NovaSeq6000 according to the manufacturer's instructions.

The resulting FASTQ files were processed along with a modified GeoMx NGS Pipeline config file using the NanoString GeoMx NGS Pipeline v2.0 or v2.3 according to the GeoMx DSP NGS Pipeline User Manual (MAN-10118-04; NanoString). Analysis was carried out using in-house data processing scripts.

A gene was called detected if the signal to noise (SNR) was ≥ 4, where the limit of quantitation (LoQ) is defined as GeoMean(NegProbes)*GeoSD(NegProbes)^2. Gene counts were normalized by Q3 normalization after the removal of genes below the detection threshold. From the Cancer Cell Line Encyclopedia RNAseq dataset (CCLE; Broad Institute), we identified a true set of expressed genes (TPM > 1) that was used to calculate the sensitivity and specificity of WTA in cell pellets. For tissue, genes were filtered to those above the detection threshold in >15% of AOIs. For protein, the signal-to-noise ratio was calculated by dividing the signal by the geomean of the three IgG negative controls (Mouse IgG1 and Rabbit IgG isotype controls). A protein target with an SNR ≥ 3 was considered detected.

Cluster heat maps were generated with the pheatmap R package. Clustering was carried out on log transformed scaled counts using the pheatmap "correlation" method. For tissue, differential gene analysis between different population of cells was performed using two-sided, unpaired t-test from rstatix R package. The threshold for significance was set at p-value < 0.05 and adjusted for multiple comparison (or multiple hypothesis testing) using the Benjamini-Hochberg method (21).

Results and Discussion

Spatial Proteogenomic Workflow Development

The existing protocol(s) for spatially resolved multiomic profiling on the GeoMx DSP requires two separate serial FFPE tissue sections, one assayed for each analyte. (Figure 2) However, serial sections do not include identical cell populations nor give a clear and direct picture of the distinct regulation of gene expression and protein levels. The collection of high plex transcriptomic and proteomic data from identical cell populations within a single tissue section is the evolution of multiomic analysis into multimodal omic readout and analysis. Therefore, we set out to develop a novel spatial multimodal omic workflow for NGS readout that allows for the simultaneous profiling of high plex transcriptomics and proteomics from a single cell population within an area of interest (AOI) on a single FFPE tissue section. We term this spatially resolved workflow "proteogenomic".

The individual GeoMx Protein and GeoMx RNA Assays for NGS readout make use of existing immunohistochemistry (IHC) or *in situ* hybridization (ISH) methodologies, respectively. The GeoMx Protein Assay uses a single antigen retrieval process of acidic heat-induced epitope retrieval (HIER) buffer (pH 6.0) under high pressure and temperature. The GeoMx RNA Assay uses a two-step, tissue dependent epitope retrieval process with basic HIER buffer (pH 9.0), followed by a proteolytic-induced epitope retrieval (PIER) step. Given two analytes that required distinct and disparate antigen retrieval conditions, we first identified slide treatment conditions compatible to both analytes.

Staining Strategy

ISH methodology involves harsh conditions, such as high salt concentrations, prolonged exposure to elevated temperatures and formamide, all of which may reduce protein antigen detection in FFPE tissue samples. To investigate the impact of ISH conditions on protein antigen detection, we screened different staining strategies.

Two sequential staining strategies were evaluated; ISH followed by IHC (ISH > IHC) and the reverse (IHC > ISH). We found that when IHC staining was performed first, followed by ISH, a decrease correlation (R = 0.86) and a 36% decrease in detected targets, or sensitivity, was observed. In contrast, carrying out ISH first followed by IHC staining, had only a minor impact on protein correlation (R = 0.95) and sensitivity (5% decrease).

We also evaluated a simultaneous strategy, concurrently staining with both antibody (IHC) and RNA (ISH). High concentrations of formamide used in the ISH protocol is known to have a negative impact on protein target detection. While formamide allows for the hybridization to occur at lower temperatures and reduce non-specific binding of RNA probes, it can disrupt antibody-antigen interactions therefore the quality of antibody staining (8, 22). With the simultaneous staining strategy, we observed a 45% decrease in protein target detection, indicating disruption of antibody-antigen binding. Therefore, we determined the strategy for optimal detection of RNA and protein targets was sequential staining with ISH followed by IHC. (Data not shown)

Impact of Epitope Retrieval Conditions

Next, we sought to identify the optimal epitope retrieval conditions for the spatial proteogenomic workflow. The standard GeoMx RNA and Protein Assays are optimized for opposing HIER (basic v acidic) conditions, whereas a single proteogenomic workflow calls for a single epitope retrieval condition. FFPE cell pellet array sections were stained in a sequential fashion with human GeoMx Cancer Transcriptome Assay (GeoMx CTA) followed by a 59-plex protein panel comprised of 6 stacked GeoMx Protein modules (Table 1). For each cell line, the signal was averaged across replicate AOIs and the signal to noise (SNR) was calculated for protein and RNA, respectively. The performance of the spatial proteogenomic workflow was compared to the single analyte workflow control slides.

Comparing the single analyte RNA (GeoMx CTA) workflow control with the spatial proteogenomic workflow on FFPE cell lines, a strong correlation (R > 0.95), regardless of acidic or basic HIER pretreatment conditions, was observed. Additionally, acidic or basic HIER pretreatment conditions had little impact on the correlation between the spatial proteogenomic workflow and the RNAseq CCLE database (23). The spatial proteogenomic workflow FPR under basic HIER conditions was < 10%. In contrast, for the workflow under acidic HIER conditions, a FPR of 30% was observed. The high FPR is consistent with previous observations, where an increase in non-specific hybridization was observed when epitope retrieval was performed under acidic conditions (20).

A

ProK (µg/mL)	Protein	RNA	
		WTA Control	RNAseq CCLE[1]
0	0.85	0.86	0.76
0.1	0.86	0.89	0.8
0.5	0.85	0.91	0.82
1	0.83	0.91	0.83

[1]Pearson coefficient correlation between the CTA control and the CCLE data base is 0.86

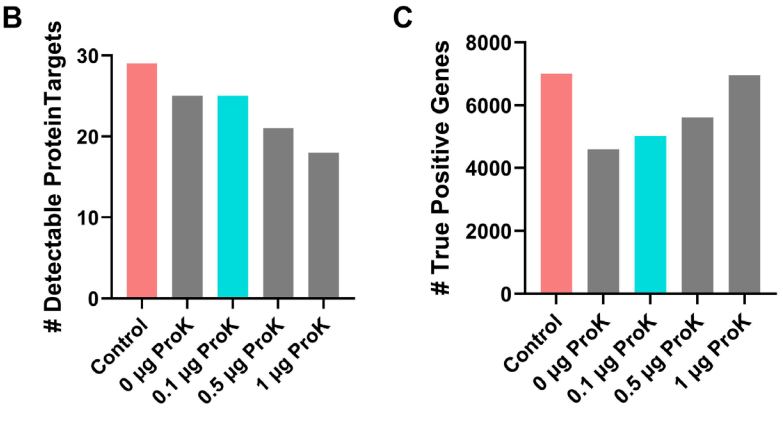

FIGURE 3: Assessment of varying proteinase K on the performance of the spatial proteogenomic workflow. A cell pellet array (CPA) treated with varying concentration of proteinase K (ProK) during epitope retrieval was stained with 6 stacked GeoMx NGS Protein Modules (59-plex) and GeoMx Hu WTA under proteogenomic workflow conditions. **(A)** Pearson correlation on the $\log_2$ transformed SNR data between the proteogenomic workflow and the single analyte controls along with the CCLE RNAseq database. Plots represents the number of **(B)** detectable protein targets and **(C)** true positive detectable RNA targets.

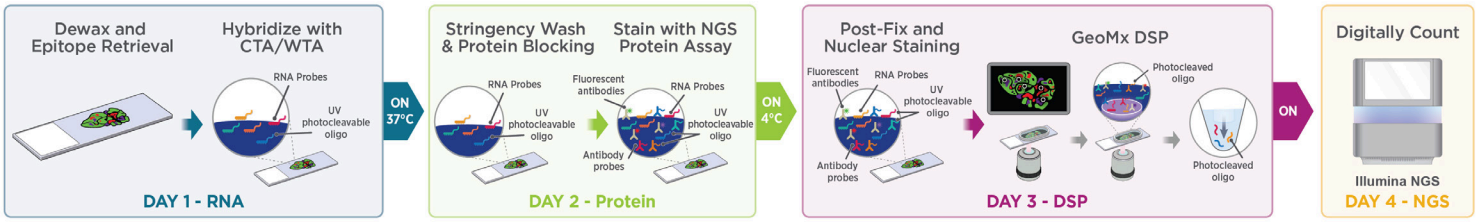

FIGURE 4: GeoMx Spatial Proteogenomic workflow enables multimodal omic profiling on a single slide.

Comparing the single analyte protein workflow control with the proteogenomic workflow, FFPE cell lines pre-treated under acidic HIER gave a higher correlation to the control (R = 0.86) when compared to basic HIER treated (R = 0.77) (Data not shown).

Evaluation of acidic vs basic HIER conditions demonstrated optimal protein detection under acidic HIER and optimal RNA detection under basic conditions, consistent with the standard GeoMx single analyte workflows. We also noted the relatively high concentration of Proteinase K (ProK) (1 µg/mL) used in the PIER epitope retrieval step drove protein target sensitivity loss. We therefore, next assessed the effects of ProK concentration, under basic HIER, on the spatial proteogenomic workflow for protein and RNA target detection.

Impact of varying Proteinase K concentrations

To evaluate the effects of varying concentration of ProK, FFPE cell pellet array sections were stained with the GeoMx Human Whole Transcriptome Atlas (GeoMx Hu WTA) probe set and a 59-plex protein panel comprised of 6 stacked GeoMx Protein Modules. FFPE cell lines were assessed under basic HIER conditions followed by proteolytic treatment (PIER) with varying concentrations of ProK. When comparing the single analyte protein control with the spatial proteogenomic workflow, correlations remained relatively strong and FPR remained < 10% regardless of ProK concentration. (Figure 3A) However, protein target detection (SNR ≥ 3) significantly decreased > 37% at ProK concentrations > 1 µg/mL. At 0.1 µg/mL, the sensitivity of the spatial proteogenomic workflow was comparable to the control. (Figure 3B)

In contrast, the correlation between the RNA single analyte control and the proteogenomic workflow increased with increasing ProK concentration. A similar trend was observed when we compared the RNA targets detected with the proteogenomic workflow to the Cancer Cell Line Encyclopedia (CCLE) RNAseq database. (Figure 3A) Furthermore, the number of true positives increased with increasing concentration of ProK. (Figure 3C) These results demonstrate the critical balance to strike with the importance of ProK proteolytic digestion for optimal RNA detection whereas even the lowest concentrations of ProK had some detrimental effect on protein target detection. Detection of low abundance protein and RNA targets were the most affected by the Proteinase K concentrations.

GeoMx Spatial Proteogenomic Workflow

In conceptualizing the optimal GeoMx Spatial Proteogenomic workflow, we considered the demonstrated effects of staining strategy, the acidic vs basic HIER, and ProK (PIER) concentration on both protein and RNA target detection sensitivity and specificity. We determined the optimal GeoMx Spatial Proteogenomic workflow to consist of a sequential staining strategy of ISH followed by IHC under basic (pH 9.0) HIER conditions and incorporating a low concentration (0.1 µg/mL) ProK digestion (PIER) step. The GeoMx Spatial Proteogenomic workflow takes 4 days, from slide prep to data analysis. Because the GeoMx Spatial Proteogenomic workflow incorporates the GeoMx Hu WTA or GeoMx CTA as its RNA probe sets, the overall workflow is only compatible with NGS readout. (Figure 4)

The overall workflow for FFPE samples is as follows:

Day 1: A two-step epitope retrieval process involving heat-induced epitope retrieval (HIER) under basic conditions followed by a proteolytic-induced epitope retrieval (PIER) step (0.1 µg/mL ProK). Samples are then stained with GeoMx Hu WTA or GeoMx CTA RNA probe cocktails, followed by an overnight hybridization at 37°C

Day 2: Sample are washed under stringent conditions in the presence of formamide and subsequently treated with blocking solution to prevent nonspecific antibody binding. Samples are then stained with GeoMx Protein Assays overnight at 4˚C. Fluorescence conjugated primary antibodies may be added at

this step for tissue morphology visualization.

Day 3: After post-fixing with 4% PFA and staining with a nuclear marker (Syto13), the samples are processed on the GeoMx Digital Spatial Profiler (Manual Reference) and sequenced on the Illumina NextSeq2000 or Illumina NovaSeq6000 as noted in the Experimental Design.

Day 4: Process data with the GeoMx NGS Pipeline as described in Experimental Design.

Profiling Cell Lines using the Optimized GeoMx Spatial Proteogenomic Workflow

Using the optimized GeoMx Spatial Proteogenomic workflow, we profiled FFPE cell pellet array (CPA) sections stained with the GeoMx Hu WTA and the 59-plex protein panel comprised of 6 stacked GeoMx Protein Modules. (Table 1) CPAs were stained with either the 59-plex protein panel or with GeoMx Hu WTA, and were used as the protein control and RNA control respectively.

We first examined the quality of the protein detected with the proteogenomic workflow compared to the protein control. A pairwise correlation analysis was performed between all cell lines and all detectable targets. For the cell line to cell line comparisons, we observed high correlation among the tested cell lines consistently between the same cell line. (Figure 5A) We further observed high correlation between identical protein targets detected with the proteogenomic or the protein control workflows. (Figure 5B)

We then examined then examined the quality of the RNA detected with the proteogenomic workflow and RNA control compared to the CCLE RNAseq dataset. For all overlapping targets between GeoMx Hu WTA and CCLE RNAseq database, each cell line in the proteogenomic workflow and RNA control data were correlated to every cell line in the CCLE dataset (1012 cell lines). A dot plot was generated showing the Pearson R distribution for each CPA cell line that overlapped with the CCLE dataset. (Figure 5C) Cell line labels represent the CCLE cell line with the highest correlation in each experiment and cell line comparison. In both the RNA control and proteogenomic workflow, we observed the highest correlation to the CCLE RNAseq dataset when comparing the same cell line to each other. We further observed high correlation between identical RNA targets regardless of workflow type. (Figure 5D)

An advantage of a proteogenomic workflow over multiomic analysis of two separate sections is the ability to have full protein and RNA data from an identical cell population. Even within homogenous cell line populations, we observe similar or improved correlations with the proteogenomic workflow as compared to a multiomic analysis of the single analyte workflow control slides. In addition, it is well understood that there is not a 1:1 relationship between RNA and protein. Rather, there are critical cellular processes, including RNA stability, translational regulation, and protein degradation, that determine RNA and protein levels at any given timepoint. We observe the positive (PTPRC/CD45) or negative (FN1/Fibronectin 1) correlation of RNA targets and their respective protein targets, consistent with previously described translational and protein degradation regulation (24-27). (Figure 5E)

GeoMx Spatial Proteogenomics in Tissue

FFPE cell lines are a useful homogenous sample type for the development a novel spatial proteogenomic workflow, reducing variability often observed in serial sections of tissue. However, complex and spatially contextual biological questions can only be answered in tissue. We therefore evaluated the spatial proteogenomic workflow on various tissues in comparison to the single analyte workflows for GeoMx Hu WTA or GeoMx Protein Assays. We compared matched immune (CD45 enriched) and tumor (PanCK enriched) regions in serial FFPE NSCLC tissue sections stained with either GeoMx Hu WTA (RNA control), a 59-plex protein panel comprised of 6 stacked GeoMx Protein Modules (protein control), or with both simultaneously (Proteogenomic). (Table 1) Circular 100 µm diameter ROIs were collected for specific CD45+ or PanCK+ areas of interest (AOI). (Figure 6A) We observed a ~25% decrease in sensitivity with respect to the protein analyte. (Figure 6B) When compared to the RNA control, a 20% and 12% decrease in the number of genes with a SNR ≥ 4 in the CD45 and PanCK enriched areas, respectively, was detected. (Figure 6C) The decrease in sensitivity is consistent with previous

6 Stacked GeoMx Human Protein Modules for NGS (59-plex)
GeoMx Human NGS Core
GeoMx Human NGS Immune Cell Typing Panel
GeoMx Human NGS IO Drug Target Panel
GeoMx Human NGS Immune Activation Status Panel
GeoMx Human NGS Pan-Tumor Panel
GeoMx Human NGS Myeloid Panel

TABLE 1: Six stacked GeoMx Human Protein Modules contain 59 antibody-antigen targets.

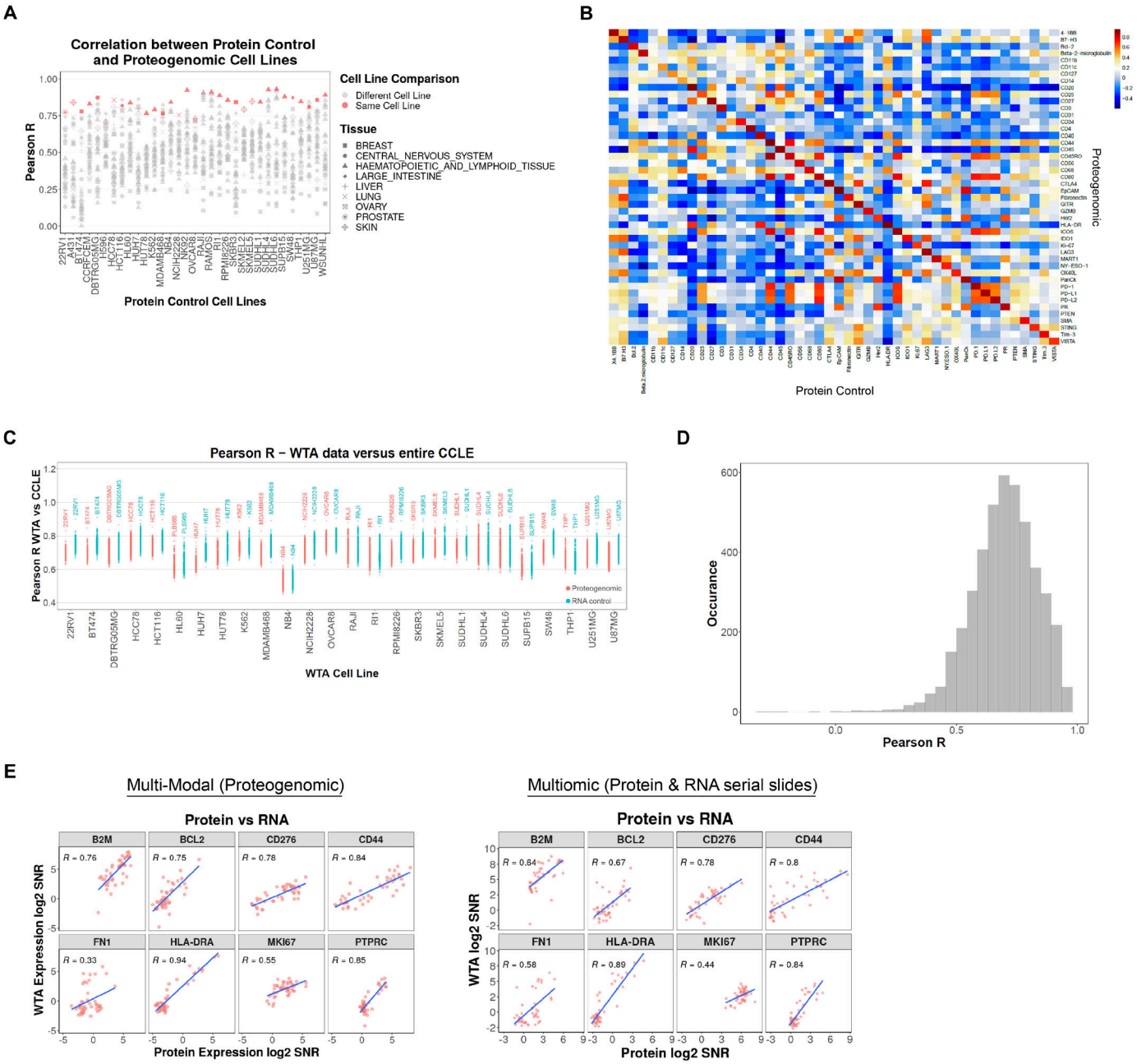

FIGURE 5: Assessment of spatial proteogenomic data quality versus the respective RNA and Protein control data on cell lines. A cell pellet array (CPA) was stained with 6 stacked GeoMx NGS Protein Modules (59-plex) and GeoMx CTA under proteogenomic and standard workflow conditions. **(A)** Cell line to cell line comparison of Protein Control to proteogenomic protein data. For protein targets with SNR ≥ 3, the Pearson R was calculated between each cell line from the Protein Control slide against all the cell lines in the spatial proteogenomic slide. Cell lines from different tissues are denoted by a unique shape and cell lines from spatial proteogenomic workflow matching the Protein Control cell line is denoted in red. **(B)** Target to target comparison of Protein Control to proteogenomic protein data. For protein targets with SNR ≥ 3, the Pearson R between each protein target from the Protein Control slide were calculated against all targets in the spatial proteogenomic slide. Heatmap of R values are displayed. **(C)** Cell line to cell line comparison of GeoMx Hu WTA data from RNA Control and proteogenomic workflow to entire CCLE RNAseq dataset. For all overlapping targets between the CCLE and GeoMx Hu WTA data, the Pearson R in the Protein Control and spatial proteogenomic GeoMx Hu WTA data were calculated against all cell lines in the CCLE RNAseq. Cell line labels in the plot correspond to CCLE cell lines with the highest R correlation to the GeoMx Hu WTA data. **(D)** Target to target comparison of GeoMx Hu WTA control to proteogenomic GeoMx Hu WTA data. For each RNA target with SNR ≥ 4 in 15% of samples, the Pearson R was calculated between GeoMx Hu WTA control $\log_2$ SNR transformed data and the respective proteogenomic GeoMx Hu WTA $\log_2$ SNR transformed data. Histogram shows the distribution of Pearson R. **(E)** Comparison of protein targets to respective GeoMx Hu WTA RNA target. For protein targets with SNR ≥ 3 and GeoMx Hu WTA targets with SNR ≥ 4, the respective protein and RNA analytes were compared using the proteogenomic workflow (left) and multiomic (right).

observations and predominantly effects low abundance targets where signal falls on or near the detection threshold.

We performed unsupervised hierarchical clustering on the matched AOIs between the GeoMx Spatial Proteogenomic workflow and single analyte workflow for GeoMx Protein Assays. We observed high concordance between both matched AOI from the proteogenomic and single analyte protein control as well as high correlation between CD45 enriched AOIs and between PanCK enriched AOIs. (Figure 6D) We then performed unsupervised hierarchical clustering on the matched AOIs between the GeoMx Spatial Proteogenomic workflow and the single analyte workflow for GeoMx Hu WTA. As with the two workflows for protein detection, we observed high concordance between matched AOI from the proteogenomic and single analyte workflows, as well as high correlation between CD45 enriched AOIs and between PanCK enriched AOIs. (Figure 6E)

In NSCLC serial sections, we demonstrated comparable target detection with the GeoMx Spatial Proteogenomic workflow to the standard GeoMx single analyte workflows. We noted a slight decrease in sensitivity, predominantly affecting low abundance or low expression targets. We also demonstrated high concordance between matched AOIs across serial sections as well as high correlation within immune or tumor enriched AOIs. We then explored the greater capabilities of the proteogenomic workflow with the GeoMx segmentation function (optical dissection) and a greater than double the protein plexity.

Segmentation with Spatial Proteogenomic Workflow

The ability to optically dissect a tissue combines the advantages of distinct spatial context and profiling specific cell subpopulations. Using the tissue segmentation capabilities of the GeoMx DSP platform, we next evaluated the spatial proteogenomic workflow on human colorectal cancer. Using the established spatial proteogenomic workflow, tissue sections were stained with GeoMx Hu WTA (18,000 RNA targets) and a 147-plex protein panel comprised of 15 stacked GeoMx Protein Modules along with fluorescence conjugated primary antibodies for CD45 and PanCK to identify the immune and tumor cell subpopulations. (Table 2) 300 µm ROIs were segmented into CD45 enriched immune and PanCK enriched tumor subpopulations. ROI were selected across the various tumor regions within the tissue section including tumor regions proximal to immune rich or immune poor regions. (Figure 7A) We performed unsupervised hierarchical clustering of detected RNA (SNR ≥ 4) and protein targets (SNR ≥ 3). As expected, we observed distinct clustering within immune

15 Stacked GeoMx Human Protein Modules for NGS (147-plex)

GeoMx Human NGS Core	GeoMx Human NGS Myeloid Panel	GeoMx Human NGS Cell Death Panel
GeoMx Human NGS Immune Cell Typing Panel	GeoMx Human NGS MAPK Signaling Panel	GeoMx Human NGS Parkinson's Pathology Panel
GeoMx Human NGS IO Drug Target Panel	GeoMx Human NGS PI3K/AKT Signaling Panel	GeoMx Human NGS Alzheimer's Extended Pathology Panel
GeoMx Human NGS Immune Activation Status Panel	GeoMx Human NGS Neural Cell Typing Panel	GeoMx Human NGS Autophagy Panel
GeoMx Human NGS Pan-Tumor Panel	GeoMx Human NGS Alzheimer's Pathology Panel	GeoMx Human NGS Glial Cell Subtyping Panel

TABLE 2: Fifteen stacked GeoMx Human Protein Modules contain 147 antibody-antigen targets.

segments and tumor segments. (Data not shown) Differential expression analysis of both protein and RNA targets between tumor and immune segments illustrate the robust co-detection and specificity of both analytes within each segment type. (Figure 7B) Examination of expression levels key RNA/Protein target pairs associated with either immune or tumor segments illustrate the variability of certain targets in distinct immune or tumor AOI. (Figure 7C) When considering the correlations of all detectable protein targets and RNA targets in either immune or tumor segmented AOI, we noted distinct patterns of correlation (red) and anti-correlation (blue). For example, IGHG1-4 RNA targets correlate strongly (red arrow) with immune protein targets such as CD44, IDO1, and CD8 but anti-correlated (blue arrow) with cell-adhesion protein targets such as Epcam, CD56, and tumor target B7-H3. (Figure 7D) In the tumor segments, we see observe anti-correlation (blue arrow) of the RNA target for MUC5AC, associated with mucus production in goblet cells, and the protein target PanCK (tumor cell marker) (28). Conversely, we observe positive correlation (red arrow) MUC5AC RNA target and autophagy related protein targets ATG5, ATG12, Lamp2a, and Bag3. Normal regulation of mucus production commonly involves autophagy for regulation and secretion of mucins (29). Additionally, abnormal expression of MUC5AC is commonly associated with malignant colorectal cancerous cells (28). (Figure 7E) Overall, we demonstrate simultaneous high-plex detection of distinct tumor or immune RNA and protein targets from individual spatially resolved CRC cell sub populations using the GeoMx DSP and application of the novel GeoMx Spatial Proteogenomic workflow.

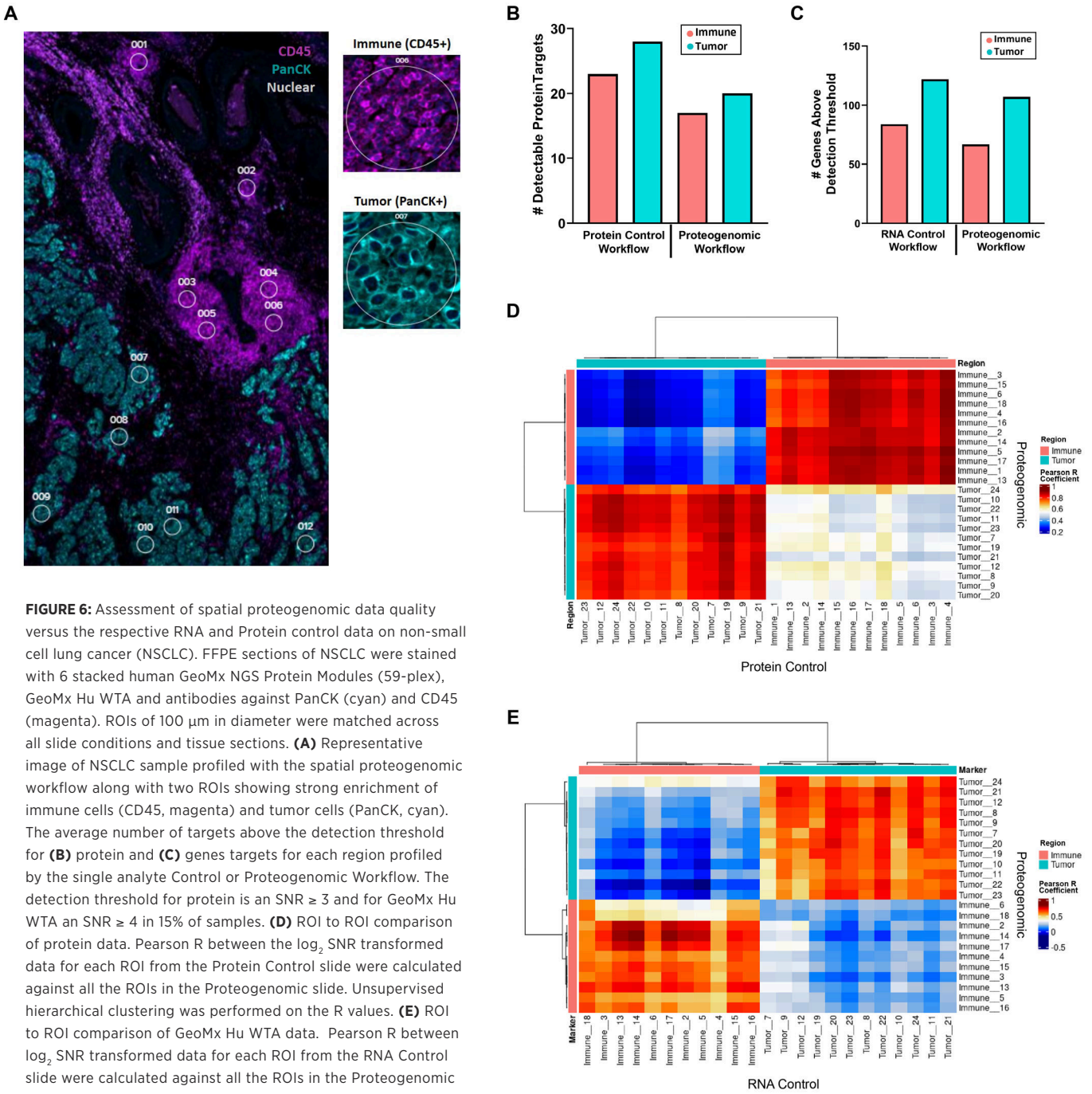

FIGURE 6: Assessment of spatial proteogenomic data quality versus the respective RNA and Protein control data on non-small cell lung cancer (NSCLC). FFPE sections of NSCLC were stained with 6 stacked human GeoMx NGS Protein Modules (59-plex), GeoMx Hu WTA and antibodies against PanCK (cyan) and CD45 (magenta). ROIs of 100 µm in diameter were matched across all slide conditions and tissue sections. **(A)** Representative image of NSCLC sample profiled with the spatial proteogenomic workflow along with two ROIs showing strong enrichment of immune cells (CD45, magenta) and tumor cells (PanCK, cyan). The average number of targets above the detection threshold for **(B)** protein and **(C)** genes targets for each region profiled by the single analyte Control or Proteogenomic Workflow. The detection threshold for protein is an SNR ≥ 3 and for GeoMx Hu WTA an SNR ≥ 4 in 15% of samples. **(D)** ROI to ROI comparison of protein data. Pearson R between the $\log_2$ SNR transformed data for each ROI from the Protein Control slide were calculated against all the ROIs in the Proteogenomic slide. Unsupervised hierarchical clustering was performed on the R values. **(E)** ROI to ROI comparison of GeoMx Hu WTA data. Pearson R between $\log_2$ SNR transformed data for each ROI from the RNA Control slide were calculated against all the ROIs in the Proteogenomic slide. Unsupervised hierarchical clustering was performed on the R values.

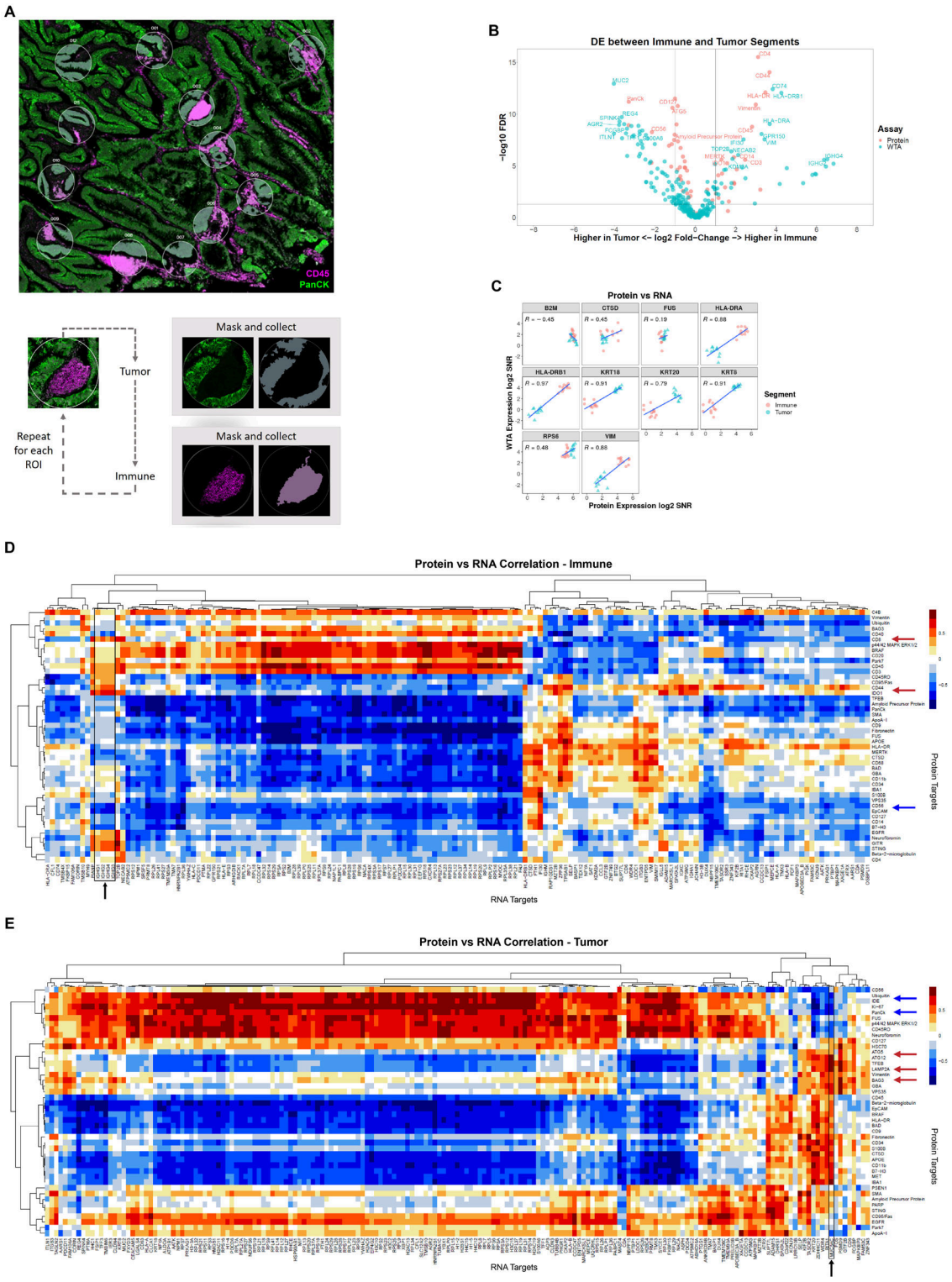

FIGURE 7: Advance ROI selection using segmentation. Multiplexed protein and RNA characterization of CRC sample with 300 µm circular ROIs segmented into immune and tumor regions. Protein and RNA counts were SNR transformed and protein targets with SNR ≥ 3 and WTA RNA targets with SNR ≥ 4 were used in the analysis. **(A)** Colorectal cancer sample profiled with the proteogenomic assay. FFPE sections were stained with 15 stacked GeoMx Protein Modules (147-plex), GeoMx Hu WTA, and antibodies against PanCK (tumor, green) and CD45 (immune, magenta). Tumor and immune segments were generated with PanCK and CD45 immunofluorescence, respectively. **(B)** Combined volcano plot of Protein and RNA expression. All immune segments were compared to all tumor segments for Protein and RNA targets above background. A subset of differentially expressed genes are labeled with colors matching their analyte. **(C)** Concordance between matching protein and RNA targets. For protein targets with SNR ≥ 3 and the respective RNA target with SNR ≥ 4, a pairwise scatterplot was generated to visualize the concordance between respective analytes. Pearson R calculations are shown in each plot. Concordance between Proteogenomic RNA and Protein targets above background in **(D)** immune and **(E)** tumor. For both immune and tumor segments, the Pearson R was calculated between each detected protein target (SNR ≥ 3) and all detected RNA targets (SNR ≥ 4).

Conclusion

In this white paper, we have reviewed the development of a high-plex, spatial proteogenomic workflow to simultaneously detect RNA and Protein from a single tissue section slide. Additionally, we have highlighted several uses cases in cell lines and tissue to demonstrate how this workflow can be leveraged to accurately measure how RNA and protein change from within the same sample. This workflow expands upon the capabilities of GeoMx DSP and enable researchers to preserve precious samples when studying multiple analytes. Future work will continue to expand the proteogenomic capabilities of GeoMx DSP by evaluating the performance of this workflow on fresh frozen (FF) samples, manual pretreatments and with mouse specific reagents. **For more information** on using GeoMx DSP for same-slide, spatial proteogenomics, please visit www.nanostring.com/spatial-proteogenomics.

Acknowledgements

We thank Bridget Kulasekara, Karen Nguyen, Daniel R. Zollinger, Michelle Kriner, Michael Rhodes, and Margaret Hoang from NanoString Technologies Inc. for their insight and guidance. Aric Rininger for their help in editing. We also want to thank Corey Williams from NanoString for their help with NGS.

References

1 J. W. Hickey *et al.,* Spatial mapping of protein composition and tissue organization: a primer for multiplexed antibody-based imaging. *Nature Methods,* (2021).

2 D. P. Nusinow *et al.,* Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **180**, 387-402 (2020).

3 D.-K. Li, W. Wang, Characteristics and clinical trial results of agonistic anti-CD40 antibodies in the treatment of malignancies. *Oncology Letters* **20**, 176 (2020).

4 L. Ramiro et al., Integrative Multi-omics Analysis to Characterize Human Brain Ischemia. *Mol Neurobiol.* **58**, 4107-4121 (2021).

5 Y. Hasin, M. Seldin, A. Lusis, Multi-omics approaches to disease. *Genome Biol.* **18**, (2017).

6 I. Subramanian, S. Verma, S. Kumar, A. Jere, K. Anamika, Multi-omics Data Integration, Interpretation, and Its Application. *Bioinformatics and biology insights* **14**, (2020).

7 Method of the Year 2019: Single-cell multimodal omics. *Nature Methods* **17**, 1-1 (2020).

8 J. Kochan, M. Wawro, A. Kasza, Simultaneous detection of mRNA and protein in single cells using immunofluorescence-combined single-molecule RNA FISH. *Biotechniques* **59**, 209-212 (2015).

9 D. Schulz *et al.*, Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Syst* **6**, 531 (2018).

10 J. Kochan, M. Wawro, A. Kasza, Simultaneous detection of mRNA and protein in single cells using immunofluorescence-combined single-molecule RNA FISH. *Biotechniques* **59**, 209-212, 214, 216 passim (2015).

11 A. Dikshit, H. Zong, C. Anderson, B. Zhang, X.-J. Ma, in *In Situ Hybridization Protocols,* B. S. Nielsen, J. Jones, Eds. (Springer US, New York, NY, 2020), pp. 301-312.

12 T. M. Grabinski, A. Kneynsberg, F. P. Manfredsson, N. M. Kanaan, A method for combining RNAscope in situ hybridization with immunohistochemistry in thick free-floating brain sections and primary neuronal cultures. *PLoS One* **10**, (2015).

13 J. N. Ko *et al.*, Multistaining Optimization for Epstein-Barr Virus-Encoded RNA In Situ Hybridization and Immunohistochemistry of Formalin-Fixed Paraffin-Embedded Tissues Using an Automated Immunostainer. *J Pathol Transl Med.* **53**, 317-326 (2019).

14 S. Ikeda, Novel and simple method of double-detection using fluorescence in situ hybridization and fluorescence immunostaining of formalin-fixed paraffin-embedded tissue sections. *Oncol Lett* **15**, 1084-1088 (2018).

15 S. Chan *et al.*, A method for manual and automated multiplex RNAscope in situ hybridization and immunocytochemistry on cytospin samples. *PLOS ONE* **13**, e0207619 (2018).

16 L. K. Officer, K. E. Andreou, A. V. Teodósio, Z. He, J. P. Le Quesne, in *In Situ Hybridization Protocols,* B. S. Nielsen, J. Jones, Eds. (Springer US, New York, NY, 2020), pp. 245-256.

17 M. Millar, in *In Situ Hybridization Protocols,* B. S. Nielsen, J. Jones, Eds. (Springer US, New York, NY, 2020), pp. 277-298.

18 C. R. Merritt et al., Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nature Biotechnology* **38**, 586-599 (2020).

19 D. R. Zollinger, S. E. Lingle, K. Sorg, J. M. Beechem, C. R. Merritt, GeoMx™ RNA Assay: High Multiplex, Digital, Spatial Analysis of RNA in FFPE Tissue. *Methods Mol Biol* **214**8, 331-345 (2020).

20 S. Gupta, J. Zugazagoitia, S. Martinez-Morilla, K. Fuhrman, D. L. Rimm, Digital quantitative assessment of PD-L1 using digital spatial profiling. *Lab Investigation* **100**, 1331-1317 (2020).

21 S. M. Zimmerman et al., Spatially resolved whole transcriptome profiling in human and mouse tissue using Digital Spatial Profiling. *bioRxiv,* 2021.2009.2029.462442 (2021).

22 Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)* **57**, 289-300 (1995).

23 L. Caruccio, K. Byrne, J. Procter, D. Stroncek, A novel method using formamide for the elution of antibodies from erythrocytes. *Vox Sanguinis* **83**, 63-69 (2002).

24 M. Ghandi et al., Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503-508 (2019).

25 E. Mourgeon, J. Xu, A. K. Tanswell, M. Liu, M. Post, Mechanical strain-induced posttranscriptional regulation of fibronectin production in fetal lung cells. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **277**, L142-L149 (1999).

26 G. Lee, R. Hynes, M. Kirschner, Temporal and spatial regulation of fibronectin in early Xenopus development. *Cell* **36**, 729-740 (1984).

27 N. Wang et al., Tumor Microenvironment Profiles Reveal Distinct Therapy-Oriented Proteogenomic Characteristics in Colorectal Cancer. *Frontiers in Bioengineering and Biotechnology* **9**, (2021).

28 F. Mair et al., A Targeted Multi-omic Analysis Approach Measures Protein Expression and Low-Abundance Transcripts on the Single-Cell Level. *Cell Reports* **31**, 107499 (2020).

29 G.-L. Gan et al., The Diverse Roles of the Mucin Gene Cluster Located on Chromosome 11p15.5 in Colorectal Cancer. *Frontiers in Cell and Developmental Biology* **8**, (2020).

30 K. K. Patel et al., Autophagy proteins control goblet cell function by potentiating reactive oxygen species production. *The EMBO journal* **32**, 3130-3144 (2013).

For more information, please visit nanostring.com

NanoString Technologies, Inc.
530 Fairview Avenue North
Seattle, Washington 98109
T (888) 358-6266
F (206) 378-6288
nanostring.com
info@nanostring.com

Sales Contacts
United States **us.sales@nanostring.com**
EMEA: **europe.sales@nanostring.com**
Asia Pacific & Japan **apac.sales@nanostring.com**
Other Regions **info@nanostring.com**

JAN 2022 MK4240