



# The Cancer Transcriptome Atlas:

Deep Spatial Characterization of the Colorectal  
Cancer Tumor Microenvironment

MK2414 | May 2020

NanoString Technologies, Inc.

Authors: Jason Reeves, Kit Fuhrman, Dan Zollinger, Jamie Kuhar, Sarah Church,  
Jenny Cronin, Michelle Kriner, Margaret Hoang, Sarah Warren

# The Cancer Transcriptome Atlas: Deep Spatial Characterization of the Colorectal Cancer Tumor Microenvironment

## Introduction

Colorectal carcinomas (CRC) are often subdivided based on microsatellite status into microsatellite stable (MSS) or microsatellite instable (MSI) tumors. In general, patients that have MSI tumors are observed to have a hypermutator phenotype leading to an immunogenic microenvironment. Furthermore, inflamed (hot) tumors have both a better prognosis and response to immunotherapeutic intervention. In contrast, MSS tumors are often considered immune cold and unresponsive to immunotherapy.

However, there exists a subset of patients with MSS colorectal cancer that respond to immunotherapy suggestive of additional mechanisms regulating the ability to enable a productive immune response within these tumors<sup>1,2,3</sup>. Similarly, not all MSI patients respond to immunotherapeutic intervention, even while

## CTA Highlights

- CTA covers 1,800+ RNA targets for the discovery of novel spatial biology within cancer samples
- The GeoMx DSP workflow gives scientists ultimate freedom to explore morphological structures
  - Segmentation based on PanCK paired with the CTA illuminates the impact of infiltrating immune cells
  - Contour analysis unveils active antigen sampling and presentation at the invasive margin
  - Rare cell analysis of Intra-tumoral vs Stromal T-cells shows immune checkpoint activation

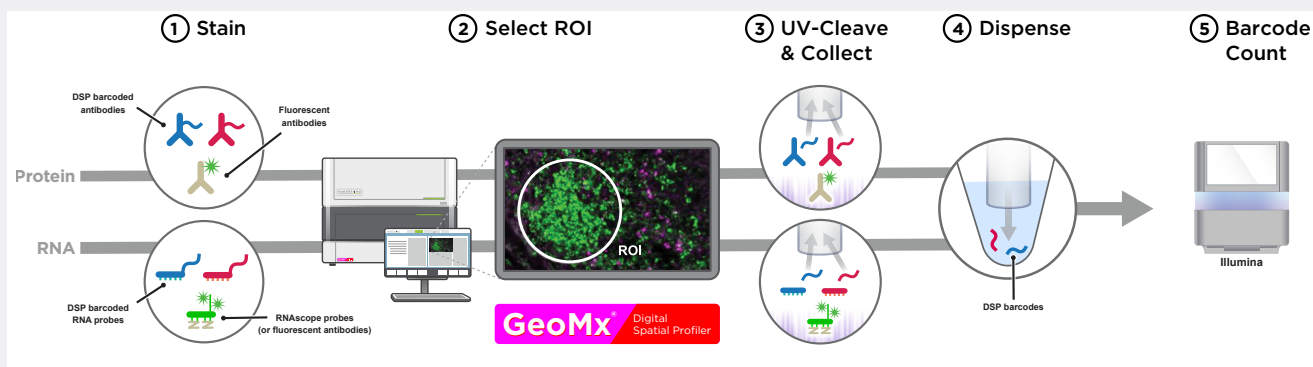


FIGURE 1 GeoMx DSP Workflow with NGS readout.

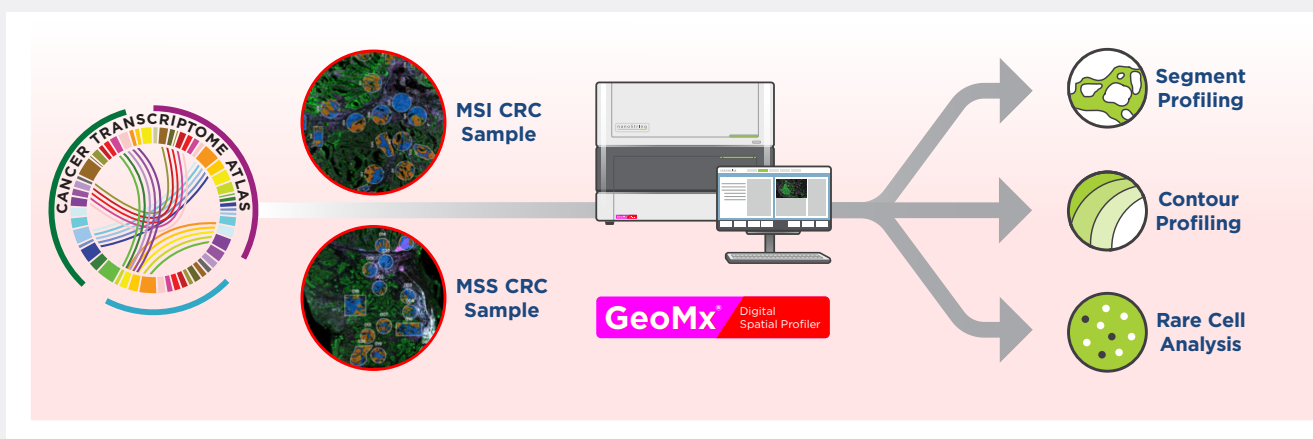


FIGURE 2: Profiling Colorectal Cancer subtypes with CTA on GeoMx DSP

having a generally favorable prognosis. The GeoMx® Digital Spatial Profiler (DSP) allows for high-plex spatial analysis of the molecular mechanisms at work throughout tissues. This approach can identify unique spatial biology and immune characteristics not captured by microsatellite status. The spatial information captured by GeoMx DSP provides more granular and contextual information than traditional gene signatures alone and has the potential to better predict response to therapeutic intervention.

### Exploring the Tumor Microenvironment of Colorectal Cancers with GeoMx DSP

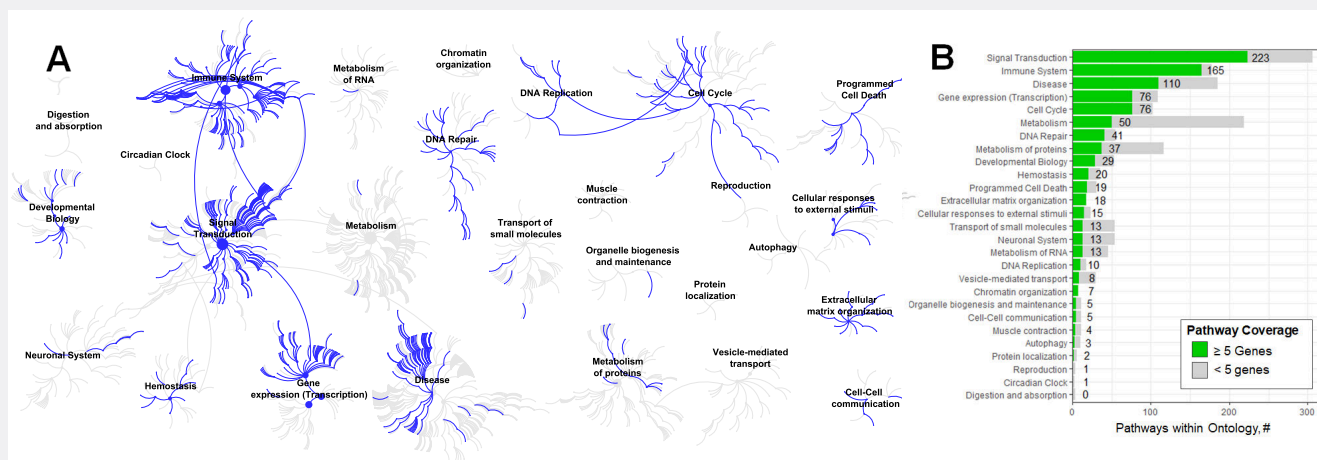
While many technologies focus on either exploring the spatial context of a handful of targets, or expression patterns of hundreds or thousands of targets in bulk samples or dissociated single-cells, the GeoMx Cancer Transcriptome Atlas (CTA) for use with the GeoMx DSP blends high-plex and spatial analysis in a single experiment (**FIGURE 1**). The GeoMx DSP enables spatially resolved, digital characterization of proteins or mRNA in a highly multiplexed (>1,000-plex) assay. The assay relies upon antibody or RNA probes coupled to photocleavable oligonucleotide tags called DSP barcodes. After hybridization of probes to slide-mounted tissue sections, the DSP barcodes are released from discrete regions of the tissue via UV exposure. Released tags are quantitated via sequencing and counts are mapped back to tissue location, yielding a spatially-resolved digital profile of analyte abundance. The CTA uses these RNA probes and was designed with the spatial context of cancer biology in mind to

enable studying key facets of disease without compromising sensitivity or specificity in Formalin Fixed Paraffin Embedded (FFPE) or fresh frozen tissues.

We undertook a series of experiments to demonstrate the ability of the GeoMx DSP and CTA to characterize the dynamics occurring within the tumor microenvironment of MSI and MSS tumors. These experiments focused on three distinct objectives: dissecting the activity within tumor and immune compartments across the tissue, specifically characterizing the invasive margin of highly infiltrated tumors, and studying the changing expression within T-cells as they invade into the tumor microenvironment in hot tumors. The workflow used to accomplish these objectives is outlined in **FIGURE 2**. Each of these questions relies on the ability to sample many genes, while retaining the spatial relationships of cells in the tissue in flexible ways specific to each question and unique tissue biology. The GeoMx DSP combined with the comprehensive coverage provided by the CTA enables each of these questions to be explored within a single slide and across cohorts of banked tissues. In this white paper we will describe the design and use of the CTA in these studies.

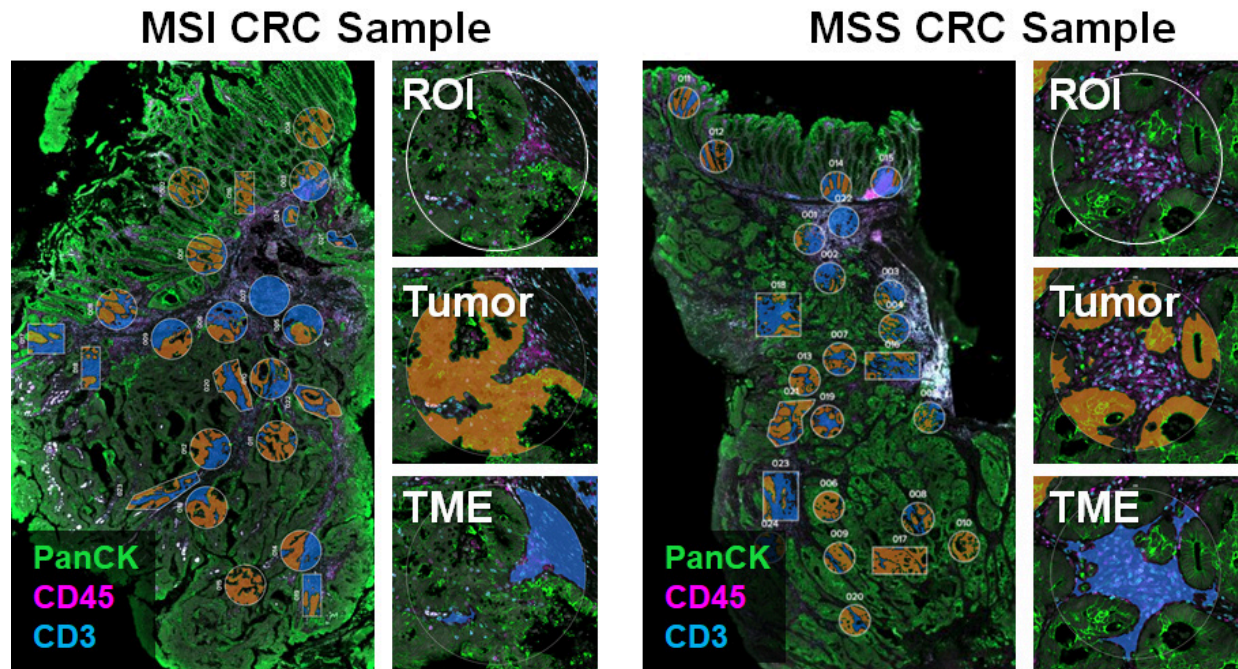
### Comprehensive Cancer Exploration with the CTA

The CTA is designed to deeply profile the interaction between tumor cells and the adjacent surrounding tissues. In developing/designing the CTA we leveraged content from NanoString's 700+ gene expression panels for the nCounter® platform, which

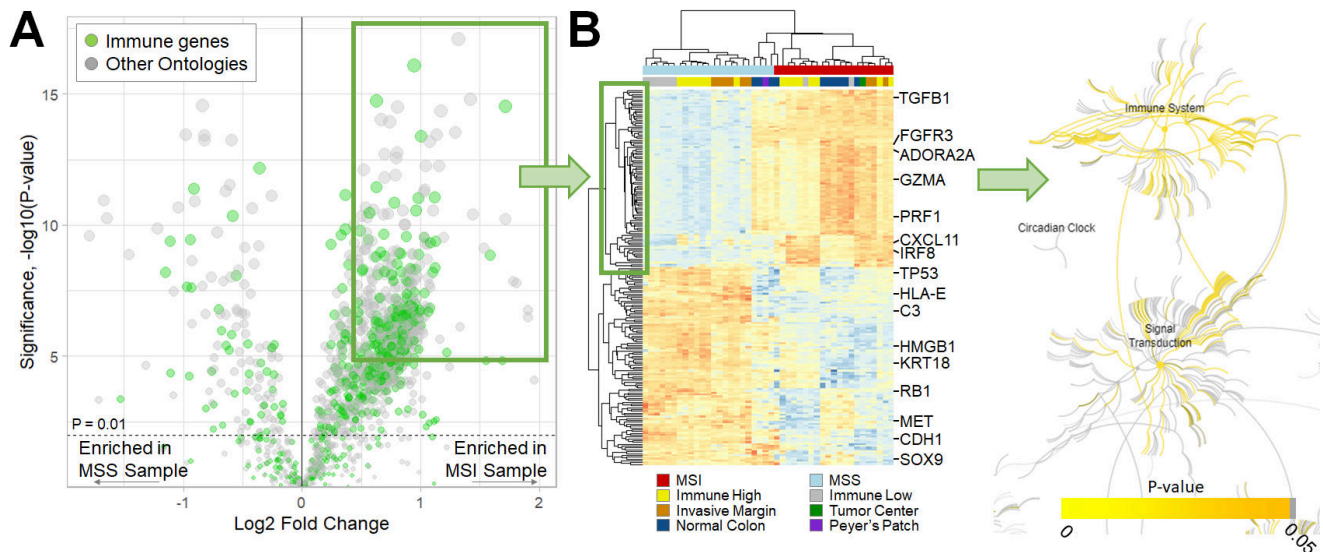


**FIGURE 3:** Overview of the Cancer Transcriptome Atlas.





**FIGURE 4:** Segmentation Strategy to profile MSI and MSS patient samples. Regions of interest shown on whole slide with example regions selected as well as UV illumination path for tumor (orange) or TME (blue) segments.



**FIGURE 5:** Pathway analysis of Tumor Segments within MSI or MSS samples. A) Volcano plot showing association of gene expression with MSI status B) Pathway analysis exploration enabled by Reactome database, showing enrichment of pathways associated with genes highly expressed in within the tumor segments of the MSI patient

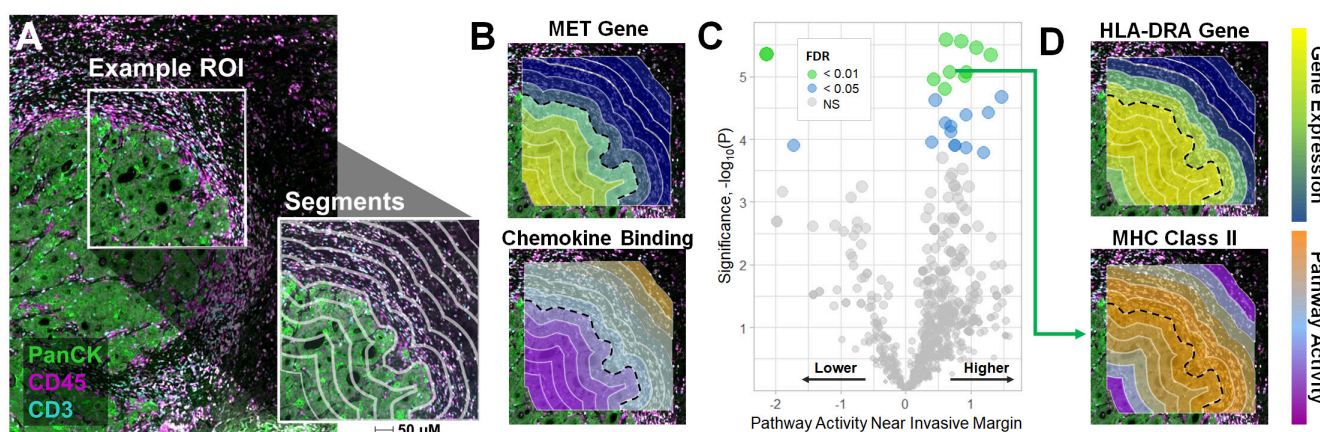
focus on various facets of tumor and immune biology curated through extensive literature review and thought leader input. These bulk expression panels have led to over 3,300 publications and key discoveries within the immuno-oncology field such as the definition of the tumor inflammation signature (TIS) that predicts response to pembrolizumab<sup>4,5</sup>. Additional content included from recent developments in the field of immuno-oncology expands the CTA coverage to new biological areas such as metabolic pathways and the immune microenvironment (**FIGURE 3**). This comprehensive and data driven approach to content development provides complete, spatial analysis of tumor samples. Measuring 1,834 genes simultaneously, this panel covers all immune cell types and checkpoint molecules, 13 canonical cancer pathways, and 55 common functional ontologies related to the microenvironment and tumor biology.

CTA probes are designed to hybridize to a unique sequence within the targeted transcript. Multiple probes are designed for each gene across different locations within the transcript allowing quantification of each mRNA, even if degradation of a molecule at an individual probe binding site occurs and the full-length transcript is no longer intact. Targeting regions throughout the transcript offers distinct advantages over Poly-A based capture methods, including the ability to target partially degraded RNA, non-coding RNA molecules, splice isoforms, and non-human sequences. The DSP barcodes attached to CTA probes are designed to be compatible with standard next-generation sequencing (NGS) platforms and standard library

preparation kits for seamless integration into current laboratory workflows. The resulting FASTQ reads are then processed in a NanoString-provided bioinformatics pipeline and the results are imported into GeoMx DSP Data Analysis software. This software allows for high-plex expression analysis alongside the tissue section images and ROI selections.

Additionally, CTA can be customized with additional genes of interest as a spike-in to the original content. As many tumors have complex and specific etiologies, this provides the flexibility to further understand questions such as tumor subtyping or viral infection that are specific to a given study.

The flexibility of regions of interest (ROI) selection with GeoMx DSP allows important biological questions to be asked in the most appropriate manner. The region illuminated by the UV light can be finely tuned to measure CTA genes expressed within specific ROIs and further segment these into biological compartments. Geometric profiling can be used to assess tissue heterogeneity by profiling with standardized geometric shapes or user-defined polygons across distinct tissue regions. Segmentation uses immunofluorescent morphology markers to identify and profile distinct biological compartments within an ROI. Contour profiling evaluates how proximity affects biological response and the local microenvironment around a central structures or tissue boundaries. Finally, rare cell analysis reveals the function of specific cell types within throughout a tissues architecture by focusing on cell type specific morphology



**FIGURE 6:** Deep Exploration of Pathway Signaling Across the Tumor Invasive Margin. A) Invasive margin profiling segmentation strategy and illuminated regions collected B) Examples of genes and pathways associated with location across the invasive margin MET gene and Chemokine Binding pathway C) Volcano plot showing pathways associated with expression at the invasive margin (\* FDR < 0.05, \*\* < 0.01) D) Examples of genes and pathways with highest expression at the invasive margin including HLA-DRA and MHC Class II Presentation



markers. For the set of experiment explored here, we applied three distinct profiling methods (**FIGURE 2**) that unlock unique aspects of the biology of colorectal carcinomas.

### Biology Driven Segmentation Differentiates Infiltrated from Excluded Tumors

To demonstrate the ability of the CTA to sensitively and specifically investigate the TME of CRC samples, we first compared the PanCK+ tumor compartment within a highly infiltrated MSI sample to the PanCK+ compartment in a non-infiltrated MSS sample (**FIGURE 4**). These samples were chosen because they contained a large T-cell presence within their stroma, but were differentiated based on the ability to recruit T-cells into the TME and their MSI status. In this experiment we used PanCK immunofluorescence (IF) to segment ROIs placed throughout the tissue into Tumor and TME segments. The TME segments were defined by collecting regions that were both PanCK- and were more than 3 microns away from the boundary of the PanCK+ regions<sup>6, 7, 8</sup>.

After collecting and quantifying the gene expression of these discreet regions we compared the tumor regions between the two samples to determine if we were able to confirm the immune infiltration observed during IF staining. Under this analysis we identified key immune-cell lineage specific genes enriched within the tumor compartment of the MSI sample (**FIGURE 4**). Standard hierarchical clustering of the top differentially expressed genes from the MSI and MSS samples identified several unique features about these samples, such as the expression of distinct oncogenes between the two tumors, as well as an enrichment of interferon responsive genes in the MSI sample (**FIGURE 5A-B**). Moreover, we were able to demonstrate that such enrichment was specific to immune pathways within the PanCK+ regions

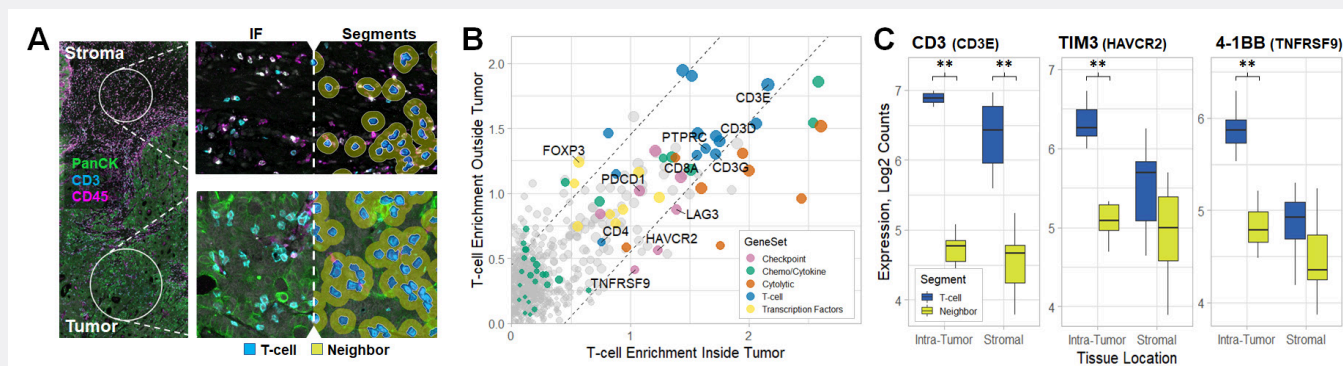
of the infiltrated MSI tumors (**FIGURE 5C**)<sup>9</sup>. To perform this analysis, we extracted the list of differentially expressed targets up-regulated within the MSI sample and queried this set of genes against the Reactome database using over-representation analysis (ORA)<sup>10</sup>, robustly demonstrating the enrichment of immune signatures within the MSI samples.

### The GeoMx CTA Unlocks Transcriptional Program of Tumor Infiltrating Cells

After demonstrating the ease by which the CTA can enable complex interrogation of tissue heterogeneity we wanted to explore additional key structures in these samples. As tumors are heterogenous, there are often structures of interest embedded within the tissue that can only be explored by maintaining the architecture of the tissue. The tumor invasive margin presents a key opportunity for understanding how CRC tumors are interfacing with their environment, and what signaling may be responsible for how the immune system and the tumor interact.

To better understand how the tumor architecture changes as we move towards and across the invasive margin, we defined the boundary between the tumor and microenvironment via IF and created 50  $\mu\text{m}$  wide segments that spanned outward or inward from that boundary in two CRC samples that were highly infiltrated. An example of this ROI selection is shown in **FIGURE 6**. To deeply interrogate the invasive margin, we applied single-sample Gene Set Enrichment Analysis (ssGSEA)<sup>11</sup> to calculate pathway scores for each segment across all ROIs.

To look for pathways that varied relative to distance from the invasive margin, we utilized a linear mixed effect model<sup>12, 13</sup> to test whether any pathways were strongly associated with distance,



**FIGURE 7:** Expression of T-cell specific activation markers within the Tumor A) ROI selection strategy to study the impact of T-cell invasion into a CRC sample B) Expression of specific genes are shared amongst T-cells (within dashed lines) or enriched within specific compartments of the tumor C) Examples of marker expression for CD3E or HAVCR2 (TIM3).

across ROIs and slides tested. We observed expression of key tumor-intrinsic markers increasing inside the tumor, such as the oncogene MET (**FIGURE 6B**), as well as an overall enrichment of immune cell related signaling and pathways outside of the tumor, for example the chemokine binding pathway. To better understand more complex spatial variation, we also examined if any pathways were specifically increased or decreased at the invasive margin (**FIGURE 6B**). Under this more nuanced analysis, we observed that the several pathways are enriched around the tumor invasive margin, such as activation of MHC class II presentation pathway and HLA-DRA expression specifically (**FIGURE 6C-D**), suggesting that immune cells are specifically sampling antigens present at the invasive margin, rather than from deeper within the tumor core. This has direct implications on understanding how antigen immuno-editing happens within a tumor, and how T-cells are being programmed to respond to antigens present within the tissue.

### Polarization of CD3+ Lymphocytes in Highly Infiltrated Tumors

Another key question that exists within the immuno-oncology field is what happens to specific cell types within different compartments of the tumor. Here we studied T-cells dispersed throughout multiple regions of the tissue. We combined segmentation and contouring to understand the distribution of gene expression of cells embedded within tissue, as well as the neighboring cells' expression (**FIGURE 7A**). This provides not only an understanding of T-cells as they invade into a tumor, but also the response of cells around them to their presence within a tissue.

In this experimental design, because the cells are tightly interspersed between stromal or tumor cells, by sampling both T-cells and the surrounding tissue, we can directly examine which cells are expressing which targets. This is done by calculating the fold change between the T-cell and the adjacent contour segment. After calculating this fold change, we can then

determine which markers are specifically expressed by different subsets of T-cells, those within the tumor and those outside of it.

This methodology can specifically identify T-cell and neighboring environmental-specific expression (**FIGURE 7B**). We observed not only canonical cell surface markers such as CD3E (**FIGURE 7C**), but also key markers involved in T-cell activation and polarization, and subsets of T-cells within this data including lowly expressed transcripts such as FOXP3. Furthermore, upon invasion within the tumor, we also identified several checkpoint molecules being expressed specifically by T-cells, such as LAG3, HAVCR2 (TIM3) (**FIGURE 7C**) and TNFRSF9 (4-1BB) (**FIGURE 6C**), as well as enrichment of cytolytic activity genes.

### Summary

In this white paper we reviewed the development of the Cancer Transcriptome Atlas, and several use cases in CRC that demonstrate how this panel can be leveraged to profile the interaction between the tumor and the microenvironment in broader disease areas. Segmentation profiling finely dissected key molecular differences within the tumor microenvironment. Contour profiling followed the unique expression changes across the tumor invasive margin that associate with immune infiltration and activity across and at this boundary. Rare cell profiling illuminated the specific transcriptional profile of T-cells infiltrating the tumor. These workflows relied on data analysis tools readily available as part of the GeoMx DSP platform, as well as common analytical methods developed by the bioinformatics community. As the field of immuno-oncology and spatial biology expands, the GeoMx DSP and CTA will be key tools to interrogate hypotheses about tumor heterogeneity, tissue structure, and response to therapeutic intervention.

To get access to a sample dataset from the first experimental design described in this document, please reach out to [GeoMxSupport@nanosttring.com](mailto:GeoMxSupport@nanosttring.com).

## References

- 1 The Emerging Role of Checkpoint Inhibition in Microsatellite Stable Colorectal Cancer. Hermel DJ, Sigal D. 2019, Journal of Personalized Medicine.
- 2 Immunotherapy with immune checkpoint inhibitors in colorectal cancer: what is the future beyond deficient mismatch-repair tumours? Huyghe N, Baldin P, Van den Eynde M. 2020, Gastroenterology Report.
- 3 Mismatch repair deficiency/microsatellite instability-high as a predictor for anti-PD-1/PD-L1 immunotherapy efficacy. Zhao P, Li L, Jiang X, Li Q. 2019, Journal of Hematology & Oncology.
- 4 IFN- $\gamma$ -related mRNA profile predicts clinical response to PD-1 blockade. Ayers M, Lunceford J, Nebozhyn M, Murphy E, Loboda A, Kaufman DR, Albright A, Cheng JD, Kang SP, Shankaran V, Piha-Paul SA, Yearley J, Seiwert TY, Ribas A, McClanahan TK. 2017, Journal of Clinical Investigation.
- 5 Pan-cancer adaptive immune resistance as defined by the Tumor Inflammation Signature (TIS): results from The Cancer Genome Atlas (TCGA). Danaher P, Warren S, Lu R, Samayoa J, Sullivan A, Pekker I, Wallden B, Marincola FM, Cesano A. 2018, Journal for Immunotherapy of Cancer.
- 6 Digital quantitative assessment of PD-L1 using digital spatial profiling. Gupta S, Zugazagoitia J, Martinez-Morilla S, Fuhrman K, Rimm DL. 2020, Laboratory Investigation.
- 7 Tertiary lymphoid structures improve immunotherapy and survival in melanoma. Cabrita, R, et al. 2020, Nature.
- 8 Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Merritt, Christopher T, Ong, G T and Church, Sarah E. 2020, Nat Biotechnol, pp. 8(5):586-599.
- 9 The reactome pathway knowledgebase. Jassal, B, et al. 2020, Nucleic Acids Research.
- 10 GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes. Boyle EI, Weng S, Gollub J, Jin H, Botstein D, Cherry JM, Sherlock G. 2004, Bioinformatics.
- 11 GSVA: gene set variation analysis for microarray and RNA-Seq data. Hanzelmann S, Castelo R, Guinney J. 2013, BMC Bioinformatics.
- 12 Fitting Linear Mixed-Effects Models Using lme4. Bates D, Machler M, Bolker B, Walker S. 2015, Journal of Statistical Software.
- 13 On fitting generalized linear mixed-effects models for binary responses using different statistical packages. Zhang H, Lu N, Feng C, Thurston SW, Xia Y, Zhu L, Tu XM. 2011, Statistics in Medicine.
- 14 Comparative analysis of differential gene expression tools for RNA sequencing time course data. Spies D, Renz PF, Beyer TA, Ciaudo C. 2017, Briefings in Bioinformatics.
- 15 SpatialDE: identification of spatially variable genes. Svensson V, Teichmann SA, Stegle O. 2018, Nature Methods.

**For more information, please visit [nanosttring.com](https://nanosttring.com)**

### NanoString Technologies, Inc.

530 Fairview Avenue North  
Seattle, Washington 98109

T (888) 358-6266  
F (206) 378-6288

[nanosttring.com](https://nanosttring.com)  
[info@nanosttring.com](mailto:info@nanosttring.com)

### Sales Contacts

United States [us.sales@nanosttring.com](mailto:us.sales@nanosttring.com)  
EMEA: [europe.sales@nanosttring.com](mailto:europe.sales@nanosttring.com)

Asia Pacific & Japan [apac.sales@nanosttring.com](mailto:apac.sales@nanosttring.com)  
Other Regions [info@nanosttring.com](mailto:info@nanosttring.com)

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

© 2020 NanoString Technologies, Inc. All rights reserved. NanoString, NanoString Technologies, GeoMx, the NanoString logo and nCounter are trademarks or registered trademarks of NanoString Technologies, Inc., in the United States and/or other countries. All other trademarks and/or service marks not owned by NanoString that appear in this document are the property of their respective owners.