LETTERS
https://doi.org/10.1038/s41587-020-0472-9

nature biotechnology

Check for updates

Multiplex digital spatial profiling of proteins and RNA in fixed tissue

Christopher R. Merritt[1]✉, Giang T. Ong[1], Sarah E. Church[1], Kristi Barker[1], Patrick Danaher[1], Gary Geiss[1], Margaret Hoang[1], Jaemyeong Jung[1], Yan Liang[1], Jill McKay-Fleisch[1], Karen Nguyen[1], Zach Norgaard[1], Kristina Sorg[1], Isaac Sprague[1], Charles Warren[1], Sarah Warren[1], Philippa J. Webster[1], Zoey Zhou[1], Daniel R. Zollinger[1], Dwayne L. Dunaway[1], Gordon B. Mills[2] and Joseph M. Beechem[1]✉

Digital Spatial Profiling (DSP) is a method for highly multiplex spatial profiling of proteins or RNAs suitable for use on formalin-fixed, paraffin-embedded (FFPE) samples. The approach relies on (1) multiplexed readout of proteins or RNAs using oligonucleotide tags; (2) oligonucleotide tags attached to affinity reagents (antibodies or RNA probes) through a photocleavable (PC) linker; and (3) photocleaving light projected onto the tissue sample to release PC oligonucleotides in any spatial pattern across a region of interest (ROI) covering 1 to ~5,000 cells. DSP is capable of single-cell sensitivity within an ROI using the antibody readout, with RNA detection feasible down to ~600 individual mRNA transcripts. We show spatial profiling of up to 44 proteins and 96 genes (928 RNA probes) in lymphoid, colorectal tumor and autoimmune tissues by using the nCounter system and 1,412 genes (4,998 RNA probes) by using next-generation sequencing (NGS). DSP may be used to profile not only proteins and RNAs in biobanked samples but also immune markers in patient samples, with potential prognostic and predictive potential for clinical decision-making.

Understanding the spatial distribution of proteins and RNAs has led to important discoveries for defining tissue pathology and elucidating biomarkers that predict patient response to therapy[1–6]. However, there remains much to learn about diagnostics, targets for drugs, predictors of response to therapy and biological mechanisms of action through protein and RNA localization within tissues. Immunohistochemistry (IHC) and RNA in situ hybridization (ISH) have traditionally been used to assess the spatial distribution of proteins and RNA in fixed tissue samples. However, these techniques are limited in their multiplexing and quantification of markers owing to the optical nature of chromogenic and fluorescent detection chemistries. The importance of investigating the expression of high numbers of markers in a spatially resolved manner has motivated the development of new methodologies to increase multiplexing capabilities. These approaches tend to fall into two broad classes: (1) sequential analyses, which examine tissues over multiple reagent processing cycles across a relatively small number of fluorescence channels, and (2) simultaneous analyses, which label targets of interest in one cycle with higher-plex readouts that do not involve light detection and are thus not constrained by the number of fluorescent probes that can be used in parallel.

Standard fluorescence microscopy methods often rely on antibody-based protein detection[7] and oligonucleotide-based RNA detection (ACD)[8], with tyramide signal amplification (TSA) to increase sensitivity. Because these approaches are limited by their ability to resolve independent fluorophores, their maximum multiplexing capability is typically 4–8 targets per tissue section, depending on the imaging platform used. More recent sequential spatial analysis methods greatly increase target multiplexing for protein and RNA. For protein detection, CODEX[9] and InSituPlex[10] methodologies increase target plex by labeling antibodies with unique oligonucleotide tags that are identified with fluorescence detection during amplification of the probe sequence or through cyclical hybridization of fluorescently labeled reverse-complement sequences. MERFISH[11] and FISSEQ[12] methodologies for RNA detection use on-tissue sequencing approaches to determine RNA expression. A general limitation of these sequential spatial analysis approaches is the experimental time it takes to collect data from multiple interrogations of a single section, which both decreases sample throughput and lengthens the time to generate results.

Simultaneous analysis methods increase throughput using various approaches. For protein detection, these developments have been driven by coupling the readout with mass spectrometry (MS). Spatial techniques such as matrix-assisted laser desorption/ionization imaging mass spectrometry (MALDI–IMS)[13] do not use affinity reagents but instead directly profile very large numbers of proteins simultaneously[14]. Spatial technologies using affinity reagents coupled with MS readout, such as imaging mass cytometry (IMC)[15] and multiplex ion beam imaging (MIBI)[16], rely on metal-isotope-labeled antibodies and are limited primarily by the number of metal isotopes that can be conjugated to the affinity reagents (approximately 100). All MS-based approaches are also limited by the relatively high cost of MS instrumentation. Additional methods using antibody–oligonucleotide conjugates also have the potential for high-plex analyses (for example, CITE-seq[17] and Abseq[18]) but sacrifice spatial information. RNA analysis using simultaneous spatial methods has primarily been driven by the ability to release and capture mRNA molecules from tissue sections. The Spatial Transcriptomics approach captures mRNA from non-fixed tissue on a predefined grid of spatially barcoded poly(T) tails, which are identified and deconvolved using a sequencing approach[19]. The recently demonstrated SlideSeq method uses a similar capture-and-sequence approach for a tissue section and has also been shown to provide single-cell or subcellular resolution for mRNAs captured with barcoded beads[20,21].

The benefit of most of these approaches (except Spatial Transcriptomics, CITE-seq and Abseq) is that they output

[1]NanoString Technologies, Inc., Seattle, WA, USA. [2]Precision Oncology, Oregon Health & Science University Knight Cancer Institute, Portland, OR, USA.
✉e-mail: cmerritt@nanostring.com; jbeechem@nanostring.com

10x Genomics EX1013

high-resolution images, which are very similar to those produced with traditional imaging systems and enable sophisticated analyses with digital image processing software[22,23]. In general, most of the methods are limited by the ability to examine only one analyte type (protein or RNA); most were also developed to profile fresh-frozen tissues rather than the standard FFPE samples banked by pathologists, although recent studies describe the use of some of these approaches for protein profiling in FFPE tissues[22,24,25]. In addition, most of these methods are not available commercially as an integrated system.

Here we describe the capabilities and applications of GeoMx Digital Spatial Profiling (DSP) in addressing these unmet needs (https://www.nanostring.com). DSP is an integrated commercial system comprising hardware, software and nCounter chemistry that enables simultaneous, highly multiplex spatial profiling of proteins or RNA in FFPE tissues.

DSP platform overview

The DSP platform quantifies the abundance of protein or RNA by counting unique indexing oligonucleotides assigned to each target of interest (Fig. 1a). Indexing oligonucleotides are covalently attached to primary antibodies (illustrated in Fig. 1a) or to mRNA hybridization probes (illustrated in Fig. 2a) with a UV-photocleavable (PC) linker. The profiling process begins on a slide-mounted FFPE tissue section that undergoes antigen retrieval and incubation with a cocktail of PC-oligo-labeled primary antibodies or mild proteinase K digestion and hybridization of RNA probes. The same section is also stained with up to four fluorescently labeled imaging reagents to identify tissue features and markers of interest. After incubation or hybridization steps are complete, slides are loaded into the DSP instrument and each sample is scanned to produce a digital image of the tissue morphology based on the fluorescent markers. This scanned image guides the selection of ROIs for profiling the PC-oligo-conjugated antibodies or RNA probes. ROIs of any shape, including noncontiguous areas, can be profiled with a programmable digital micromirror device (DMD) that directs UV light to precisely illuminate the ROI and release PC-oligos in a region-specific manner. The DSP prototype shows high correlation in counts in comparison to the UV area illuminated, down to areas as small as a single cell (described below). Cross-talk between neighboring compartments was characterized and can be minimized by using various approaches (Supplementary Fig. 1). The released indexing oligonucleotides are collected via microcapillary aspiration, dispensed into a microtiter plate and digitally counted using the nCounter system or an NGS readout. On the prototype DSP instrument, the entire workflow (Fig. 1a), from tissue preparation to data output, can be completed in 1.5 to 2.5 d. Most steps can be combined to process multiple samples in parallel, allowing the collection of ~384 ROIs from one or multiple specimens within 48 h.

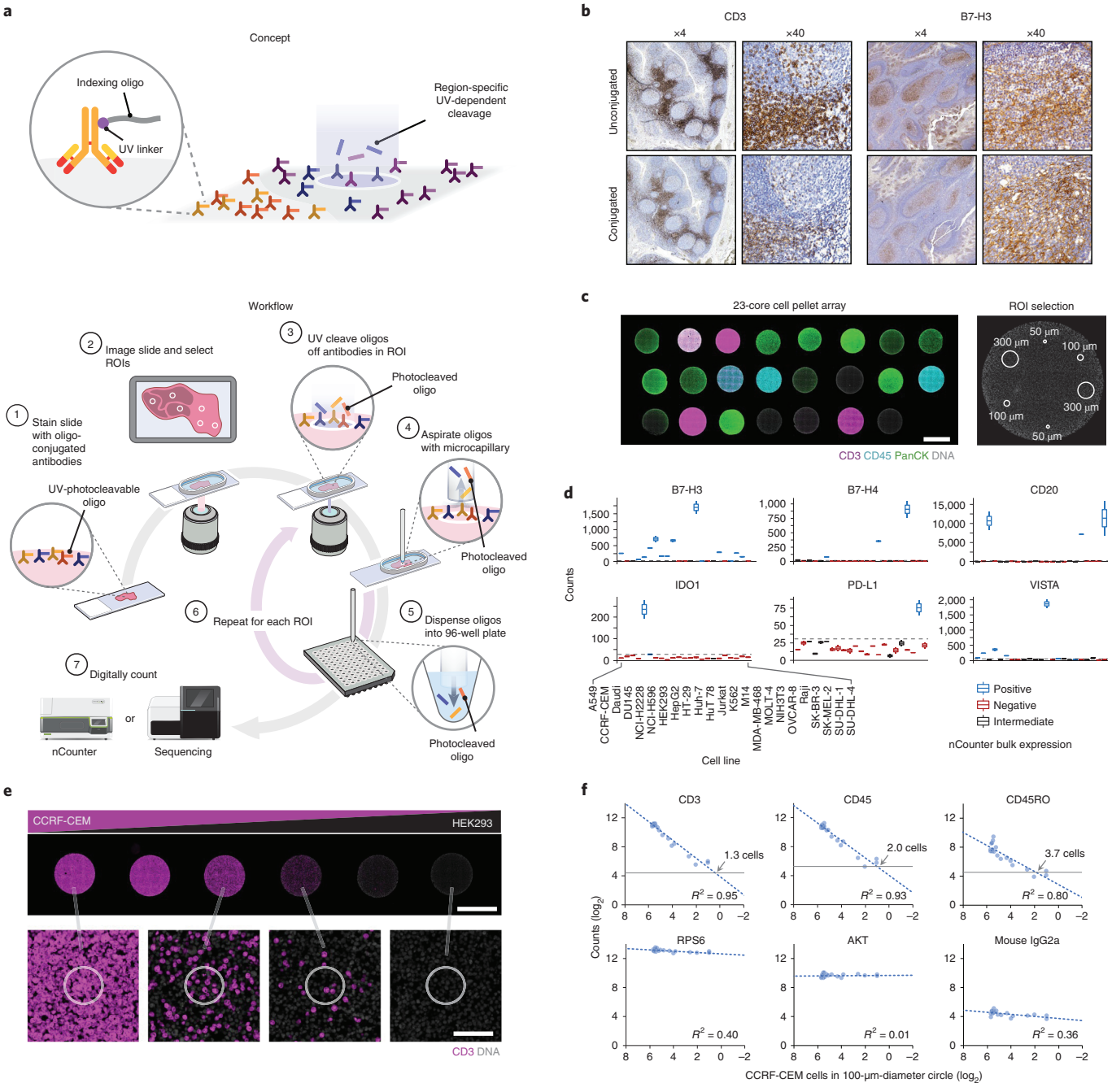
Antibody specificity after oligonucleotide conjugation

FFPE-qualified antibodies were first conjugated with PC-oligo sequences compatible with readout on the nCounter system. Antibodies were tested to validate the retention of staining specificity after conjugation (Fig. 1b and Supplementary Table 1). Standard IHC was performed on positive-control FFPE tissue samples or cell pellets on serial sections with conjugated and unconjugated antibodies. Oligonucleotide conjugation typically maintained specificity but led to lower IHC staining intensity.

Fig. 1 | Overview of the DSP workflow and validation of DSP with oligonucleotide-conjugated antibodies. a, DSP concept. Affinity reagents (for example, antibodies or RNA probes) are covalently linked to a DNA indexing oligonucleotide with a UV-cleavable linker (PC-oligo). These affinity reagents are used to stain tissue sections, and focused UV light liberates indexing oligonucleotides from any ROI. These oligonucleotides are then collected and digitally counted. Overview of the localized protein profiling workflow on the DSP instrument. (1) Process tissue section: FFPE slide-mounted tissue is subjected to a standard antigen retrieval protocol, followed by incubation with fluorescent imaging reagents (specific antibodies or biological stains and dyes) and a cocktail of primary antibodies conjugated to PC-oligos. (2) View and select ROIs: DSP scans a ×20 image of up to four-channel fluorescence. Fluorescent imaging reagents establish the overall architecture of the tissue (for example, fluorescently labeled antibodies to PanCK and CD45 identify epithelium and immune cells, respectively, and a nucleic acid stain identifies nuclei). ROIs are drawn or automatically placed on the basis of fluorescence for multiplex profiling. (3) Localized PC-oligo release: oligonucleotides from the selected ROI are released upon localized exposure to UV light. (4,5) Oligonucleotide collection: photocleaved oligonucleotides that have been released into the aqueous liquid above the slide are collected via a microcapillary tube and deposited in a microtiter plate for subsequent quantitation. (6) Repeat: localized UV exposure and PC-oligo collection (steps 3–5) are repeated for each ROI, with extensive washing between each cycle. (7) Digital or NGS counting: spatially resolved pools of photocleaved oligonucleotides are either hybridized to NanoString fluorescent barcodes, enabling the digital counting of up to ~1 million binding events per ROI using the standard NanoString nCounter analysis system, or quantified by NGS, where the entire plate can be pooled into a single tube, purified and sequenced. Reads are processed into digital counts and mapped back to each ROI, generating a map of protein and transcript activity within the tissue architecture. **b**, Comparison of the staining specificity for unconjugated and oligonucleotide-conjugated antibodies to CD3 and B7-H3 on tonsil with DAB IHC. IHC was performed at 0.25 μg ml^{-1} and 0.25 μg ml^{-1} concentrations for unconjugated antibodies and at 1 μg ml^{-1} and 4 μg ml^{-1} concentrations for conjugated antibodies, respectively. Antibodies to CD3 and B7-H3 showed similar staining patterns before and after conjugation. These experiments were performed independently two times with similar results. **c**, A 23-core FFPE cell pellet array was stained with fluorescently conjugated antibodies to CD3, CD45 and PanCK and with SYTO13 nuclear dye, along with a mix of 41 oligonucleotide-conjugated antibodies. $n = 2$ replicate ROIs across $n = 3$ ROI areas (circles of 50 μm, 100 μm and 300 μm in diameter) were profiled ($n = 6$ total technical replicates). Scale bar, 2 mm. **d**, Box plots of replicate 100-μm ROIs across all cell lines for six targets (additional targets shown in Supplementary Fig. 3). nCounter counts from 100 ng of purified RNA were used to estimate which cell lines were expected to be positive (>100 counts), negative (<20 counts) or have intermediate expression (20–100 counts). The LOD of each target was calculated as the mean + 3 s.d. for all negative cell lines. Box plots, used to clearly show the replicate data points, indicate the median and upper and lower quartiles with whiskers extending to the minimum and maximum values. **e**, Cell pellet TMA with variable percentages of two cell lines with differential expression of CD3 (magenta). CCRF-CEM cells were titrated into a background of HEK293 cells with approximate proportions of CCRF-CEM cells of 100%, 95%, 70%, 30%, 5% and 0%. Circles of 100 μm in diameter were selected across cell pellet titrations. Representative ROIs are shown for the 100%, 70%, 30% and 0% proportions. This experiment was performed independently two times with similar results. A second experiment was performed with a different set of cell lines (Supplementary Fig. 5). Top scale bar, 2 mm; bottom scale bar, 100 μm. **f**, nCounter counts plotted against the number of CCRF-CEM cells (CD3-positive cells) in each ROI for three targets expressed exclusively in CCRF-CEM cells (CD3, CD45 and CD45RO), two targets expressed in both cell lines (RPS6 and AKT) and one negative-control antibody (mouse IgG2a). ROIs were profiled that contained a variable number of CCRF-CEM cells ($n = 24$). R^2 (from Pearson's R) values are shown. The LOD of each target was calculated as the mean + 3 s.d. for all ROIs containing no CCRF-CEM cells. Arrows indicate the intersection of the LOD with the best fit regression line. The cell count on each graph enumerated from this intersection is the estimated number of cells needed for detection of the antibody.

Validation of oligonucleotide-conjugated antibodies on the DSP system

Validation of the DSP methodology was first performed on well-characterized cell lines with stable expression of most protein targets of interest. Antibodies for 44 immuno-oncology proteins and isotype controls were conjugated with PC-oligos (Supplementary Table 1). Positive- and negative-control cell lines were identified for targets of interest with bulk gene expression analysis on purified RNA samples. Selected cell lines were included in FFPE cell pellet tissue microarrays (TMAs) to simultaneously quantify conjugated antibody specificity and sensitivity on the DSP system (Supplementary Fig. 2). Fluorophore-labeled antibodies to CD3, CD45 and pan-cytokeratin (PanCK) were included with a mix of 44 oligonucleotide-conjugated antibodies, and samples were stained with a modified IHC protocol followed by nuclear fluorescence staining and then profiled on a prototype DSP instrument. Fluorescence images from the DSP instrument showing the location of the 23 TMA cores were used to guide ROI selection (Fig. 1c and Supplementary Fig. 2). Digital counts obtained from the nCounter system were compared to the fluorescence images obtained on the same section for the antibodies to CD3, CD45 and PanCK (Supplementary Fig. 2). Average counts from DSP profiling for the replicate ROIs showed expression patterns that reflected the intensities seen with fluorescence images for the corresponding target. To analyze the other 41 oligonucleotide-conjugated antibodies in the cocktail, replicate circular ROIs of 100 µm in diameter were plotted across cell lines with expected expression annotated as positive, negative or intermediate on the basis of counts obtained with purified RNA using the nCounter system or publicly available RNA-seq datasets (Fig. 1d and Supplementary Figs. 3 and 4). In general, DSP counts had a wide dynamic range, with negative and positive cell lines showing various levels of differential expression



(Fig. 1d and Supplementary Fig. 3). CD20 and PD-L1 are examples of the extremes in performance. These differences likely reflect differential abundance of the targets in the control cell lines and the intrinsic sensitivity and specificity of each antibody. Unsupervised clustering of data across the entire dataset was also performed to assess reproducibility across replicate ROIs from the cell pellet TMA (Supplementary Fig. 3). The six ROIs selected for each cell type clustered together across the entire dataset, verifying the reproducibility of different ROI sizes and replicates for the slide. Expected clustering was observed for cell lines with similar tissues of origin, including leukemia and lymphoma or breast cancers. Targets also tended to cluster by protein target type (housekeeping proteins, T cell markers or isotype controls).

Correlation of DSP counts with titrated cells

We next generated a mixed-proportion cell pellet TMA containing CCRF-CEM cells (CD3-positive, acute lymphoblastic leukemia-derived cells) titrated into a background of HEK293 cells (CD3-negative, embryonic kidney-derived cells) to test the ability of DSP to quantify variable cell composition within an ROI using oligonucleotide-conjugated antibodies. We did not focus on the precise quantification of proteins because antigen levels were unknown across these targets and cell lines and the affinity of each antibody was unknown. For this, CCRF-CEM-specific targets were quantified within ROIs of 100 µm in diameter (Fig. 1e) and counts were plotted against the number of CD3-positive CCRF-CEM cells identified using fluorescence staining for each marker within the ROI (Fig. 1f). Targets expressed exclusively in CCRF-CEM cells showed the expected linear decrease in counts with a decrease in the proportion of CCRF-CEM cells present in each ROI. Targets expressed in both cell lines, ribosomal protein S6 (RPS6) and AKT, did not show a decrease or increase in counts. The negative control mouse IgG2a showed low counts across all ROIs. These data demonstrate the capability of the DSP platform to quantify relative levels of protein, as measured by known numbers of cells, across ROIs.

Because this mixed-proportion cell pellet assay revealed strong correlations between nCounter counts and positive cell numbers in an ROI, we were able to calculate the limit of detection (LOD) measured in number of cells for targets expressed exclusively in CCRF-CEM cells. Note that such an LOD is impacted by both the level of protein expression of the target in a given cell type and the affinity of the antibody reagent. Figure 1f shows the LOD and the LOD intercept with a linear regression trendline to estimate the numbers of CCRF-CEM cells necessary to detect each target. CD3, CD45 and CD45RO had LODs approaching single-cell detection at 1.3, 2.0 and 3.7 cells, respectively. Additional targets were analyzed with this mixed-proportion TMA and a second cell line mixture (Supplementary Fig. 5). In all cases, counts decreased proportionally to the abundance of target-expressing cells. Across the 12 targets analyzed with cell line-specific expression, LODs ranged from <1 cell (CD20 and CD45) to 74 cells (PD-L1). PD-L1 was a clear outlier in these experiments, likely owing to the low specificity of the antibody, with the second highest target LOD measured as 5.5 cells (VISTA). Additional experiments were performed on FFPE cell pellets, tonsil sections and primary cells to further characterize the LOD of the DSP system (Supplementary Fig. 6). We found that CD3 and CD20 could also be detected down to a single-cell level in these alternative model systems. In all, these results demonstrate the high specificity and sensitivity of the DSP system.

In situ detection of RNA with DSP

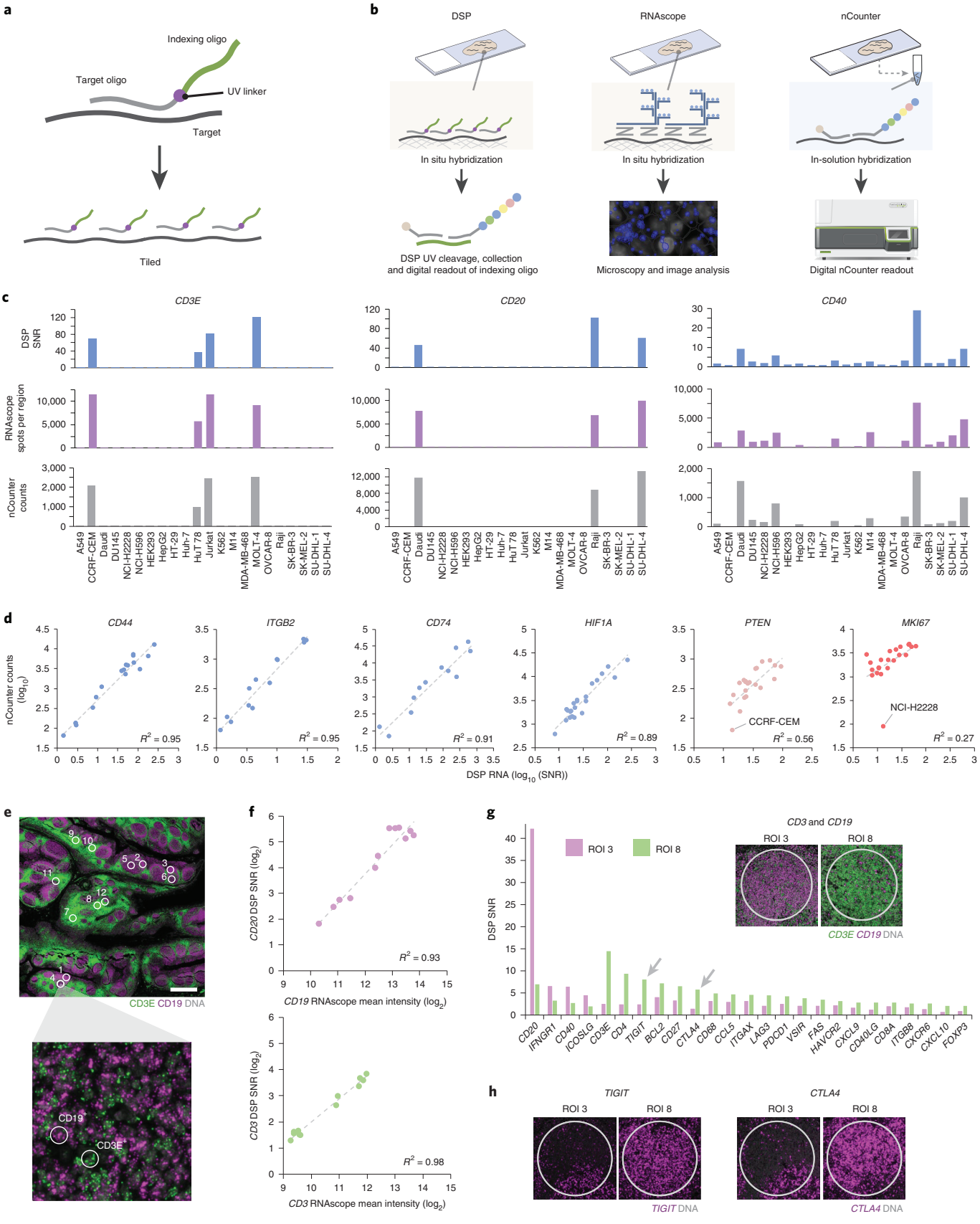
We also developed oligonucleotide probes and protocols to enable quantitative in situ detection of RNA targets with DSP. The DSP RNA detection assay uses oligonucleotide probes containing sequences complementary to targets of interest and an indexing PC-oligo sequence (Fig. 2a). Multiple probes with unique target oligonucleotides but identical indexing PC-oligo sequences ('tiles') were spaced across a transcript to increase the sensitivity and isoform coverage of the assay.

Cross-platform validation of DSP RNA detection

We confirmed the specificity of the DSP RNA assay by comparing it to two orthogonal approaches: fluorescence in situ hybridization (FISH) and bulk RNA expression analysis (Fig. 2b). FISH, direct fluorescence detection of RNA transcripts on mounted FFPE sections with microscopy, was performed using commercially available RNAscope assays (ACD) and subsequent quantification of fluorescent spots (each representing one transcript) with image analysis software (Supplementary Fig. 7). Bulk RNA expression analysis for additional comparison was performed with direct detection and digital counting of purified RNA transcripts. A 96-plex DSP RNA panel was built comprising probes covering 88 immuno-oncology-focused targets plus 8 nonspecific control RNA sequences, with 8–10 tiles per target (Supplementary Table 2).

Fig. 2 | Validation of oligonucleotide-conjugated RNA probes for DSP. a, Oligonucleotide probes were designed with a target-specific complementary sequence joined to an indexing sequence by a UV-photocleavable linker. Multiple RNA probes were designed to target various regions of each transcript (tiled). **b**, Overview of the orthogonal approaches used for validation of RNA DSP. Left, DSP: RNA targets (dark gray) are bound to tissue sections via formalin cross-linking. DSP probes bind directly to cross-linked targets (ISH), and the indexing portion of the probe is digitally counted after UV cleavage. Middle, RNAscope: ISH is performed with a series of amplifying oligonucleotides and a final layer allowing visualization with fluorescent probes (either directly conjugated to oligonucleotides or through a TSA reaction). Fluorescent spots, each representing individual transcripts, are quantified with image analysis software. Right, nCounter: RNA transcripts are purified, hybridized in solution with nCounter barcodes and counted on the nCounter platform. Note that nCounter barcodes hybridize directly to RNA molecules, whereas, for DSP, nCounter barcodes hybridize to the indexing PC-oligos. **c**, Comparison of *CD3E*, *CD20* and *CD40* transcripts across 22 cell pellet types with all three platforms. Top row, DSP signal-to-noise ratio (SNR; target counts/average negative-control counts) for an average of $n = 2$ replicate circular ROIs of 400 µm in diameter (0.1320 mm^2 total area). Middle row, RNAscope spot counts for $n = 2$ replicate ROIs of 300 µm × 400 µm (0.24 mm^2 total area). Bottom row, nCounter PanCancer IO360 counts for 100 ng of purified RNA from FFPE sections. Similar experiments were performed independently two times (Supplementary Fig. 8 and 9). **d**, Correlation of log-transformed cell pellet data from DSP RNA analysis and nCounter data from purified RNA. Representative targets are shown with at least $n = 10$ data points significantly above background from bulk RNA profiling and covering at least one order of magnitude of expression. R^2 (from Pearson's R) values are shown. Samples in blue have $R^2 > 0.75$, and samples in red have $R^2 \leq 0.75$. **e**, RNAscope versus DSP RNA profiling in tonsil. A tonsil section was imaged using three-color fluorescence of CD3E (green), CD19 (magenta) and DNA (gray) to establish overall tissue morphology. After RNAscope fluorescent labeling, DSP RNA probes were hybridized to the tissue. Circular ROIs of 500 µm in diameter ($n = 12$) were selected for detailed molecular profiling with a 96-plex RNA probe cocktail. Scale bar, 2 mm. **f**, Correlation of log-transformed data from quantification of RNAscope intensities and DSP RNA SNRs for each ROI. **g**, Targets with a >2-fold difference for two representative ROIs. ROI 3 and ROI 8 showed strong enrichment of B cells (*CD19*) and T cells (*CD3E*), respectively, with RNAscope (images shown) and strong enrichment of B cell markers and T cell markers, respectively, with DSP. **h**, *TIGIT* and *CTLA4* RNAscope visualization on serial sections to confirm signal enrichment quantified with DSP profiling (indicated with arrows in the bar graph in **g**). This experiment was performed once for both the *TIGIT* and *CTLA4* targets on sections flanking the section profiled with DSP.

A cell pellet TMA was profiled with ROIs of 500 µm in diameter using this panel. PC-oligos were quantified using the nCounter system. We observed highly concordant expression patterns with all three methods across three targets, thus validating the specificity of the DSP RNA assay (Fig. 2c). Additional analyses performed across a total of 11 targets with smaller ROIs of 200 µm in diameter



showed similar concordance across these three approaches (Supplementary Fig. 8). These data also allowed us to estimate the LOD of this assay, at the transcript level, by comparing targets just above background in the DSP assay to the corresponding transcript counts obtained with RNAscope (Supplementary Figs. 8 and 9). For example, *CD40* was found to be just above background when ~318 NIH-H2228 cells were profiled with DSP (Supplementary Fig. 8). Because the average *CD40* transcript abundance in NIH-H2228 cells was ~1.8 transcripts per cell with RNAscope, we could estimate the LOD of *CD40* to be ~580 transcripts (318 cells × 1.8 transcripts per cell; Supplementary Fig. 9). Including data for additional target and cell combinations, we estimate the limit of RNA detection at approximately 2 copies per cell across 300 cells, or 600 total individual mRNA transcripts (Supplementary Fig. 9). By extension, we expected to be able to detect higher expressors in fewer cells. To examine this, ROIs of 50 µm in diameter containing ~20–40 cells were profiled (Supplementary Fig. 9). As expected, we could readily detect six more highly expressed targets in ROIs containing a minimum of 21 cells, thus confirming the LOD estimates of this system.

Further validation focused on the comparison of DSP RNA expression analysis versus bulk gene expression analysis using the nCounter system. Correlation of the two measurements for transcripts that were well above background in the nCounter system in ten or more cell lines, with a dynamic range of expression of >10-fold, was analyzed (Fig. 2d). Most targets (24 of 27) had R^2 values above 0.75, with a range of correlation values from 0.76 to 0.97 (Supplementary Fig. 7). *CD44*, *CD74*, *ITGB2* and *HIF1A* showed the highest correlation values, from 0.89 to 0.95 (Fig. 2d). *PTEN* and *MKI67* (Ki-67) were the greatest outliers, with each target showing aberrantly low nCounter counts for a single cell line. This discrepancy might be explained by alternative transcript detection by probes targeting different regions of the transcript in the two assays.

We next characterized the concordance between DSP and RNAscope FISH data in FFPE tissues. RNAscope probes targeting the B cell marker *CD19* and the T cell marker *CD3E* were used to identify ROIs (Fig. 2e). After RNAscope visualization, slides were incubated with the DSP RNA 96-probe cocktail and 12 ROIs were profiled (Fig. 2e). DSP counts were quantified with the nCounter system, and RNAscope signal intensities were quantified using image analysis software. Concordance was tested for B cell transcripts whose spatial expression profiles are known to be similar but have no probe sequence overlap (*CD19* for RNAscope and *MS4A1* (also known as *CD20*) for DSP) and also for a T cell transcript for which probe sequences targeted overlapping regions (*CD3E* for both RNAscope and DSP). We found that both probe configurations displayed high concordance in both tissue types (Fig. 2f). In all, these data show that RNAscope and DSP signals are highly correlated when analyzed in the same tissue section.

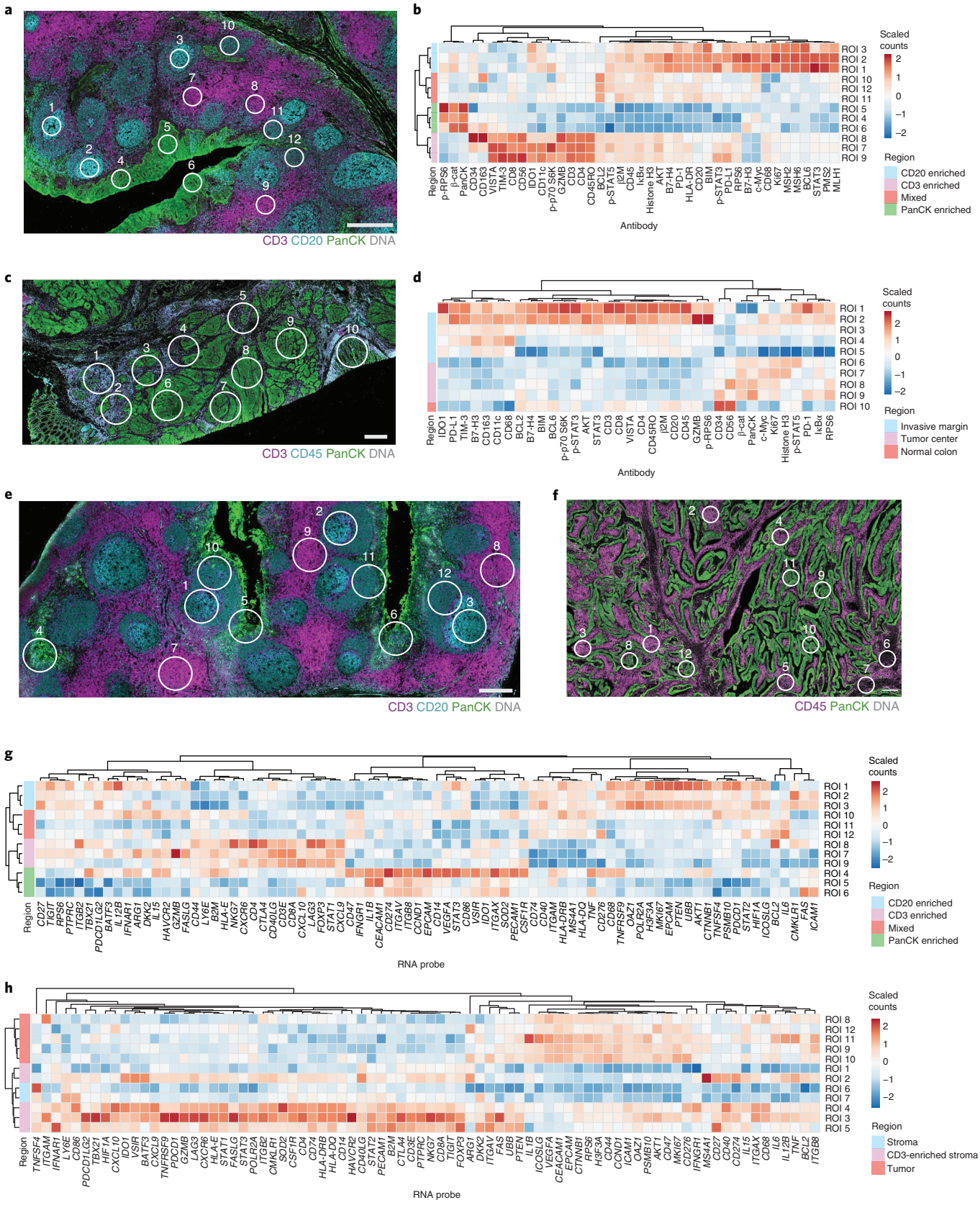
RNAscope assays were also used for validation of DSP results in serial sections of the same FFPE tissue samples. Two representative ROIs for each sample type, with one ROI enriched for B cells and another ROI enriched for T cells, are shown in Fig. 2g. As expected, higher expression of *CD20* was seen for the B cell-enriched ROI, along with higher counts for other markers preferentially expressed in B cells (Fig. 2g). Similarly, for the T cell-enriched ROI, higher expression of *CD3E* was seen, along with higher counts for other transcripts preferentially expressed in T cells. We validated the preferential expression seen in T cell-enriched ROIs by performing RNAscope on two selected targets, on serial sections (Fig. 2h). RNAscope visualization of *TIGIT* and *CTLA4* transcripts showed enrichment of these targets in the *CD3E*-enriched ROIs, thus confirming results obtained with DSP.

Profiling of protein and RNA with circular ROIs

Our first thorough examination of this platform on FFPE tissue sections was performed with simple circular ROIs. We profiled lymphatic tissues, which have well-documented expression of many of the selected targets, as well as a sampling of tumor tissues with unknown expression profiles. Twelve ROIs encompassing a range of morphological tissue features were selected for 44-plex antibody profiling on the basis of enrichment for CD3, CD20 or PanCK expression (Fig. 3a and Supplementary Table 1). DSP counts for the three markers used to visualize ROIs strongly correlated with each morphologically enriched region (Fig. 3b). Unsupervised clustering of regions with similar visual profiles (enriched for each morphology marker) demonstrated the correspondence between CD3, CD20 or PanCK morphology marker identity and the high-plex molecular profiles based on T cells, B cells and epithelium, respectively. We also found high section-to-section reproducibility across two serial FFPE tonsil sections profiled for this experiment (Supplementary Fig. 10). A similar approach was used to profile colorectal cancer (CRC) tissue (Fig. 3c). Unsupervised clustering of counts from these ROIs (Fig. 3d) showed that ROIs clustered according to the spatial immune biology of the tumor at the invading margin, the center of the tumor and normal tissue. As would be expected, the invasive margin expressed the highest levels of immune cell-related markers, including those associated with an adaptive (T cells and B cells) and innate (myeloid) immune response, while the tumor center was associated with tumor markers such as PanCK, β2-microglobulin (β2M) and Ki-67 (a proliferation marker). We also performed similar geometric profiling of in situ RNA expression with similar results (Fig. 3e–h).

Fig. 3 | Spatially resolved, multiplexed protein characterization of geometric regions on FFPE tissue. a, Protein profiling of tonsil. A tonsil section was imaged using four-color fluorescence of CD3 (magenta), CD20 (cyan), PanCK (green) and DNA (gray) to establish overall tissue morphology. Twelve circular ROIs of 200 µm in diameter were selected for detailed molecular profiling with a 44-plex oligonucleotide–antibody cocktail. Scale bar, 500 µm. **b**, Unsupervised hierarchically clustered heat map of region-specific nCounter counts across all protein targets on tonsil. Regions with similar visual profiles cluster together, demonstrating the correspondence between local biological identity and molecular profiles. Region annotations were manually curated on the basis of immunofluorescence staining patterns. **c**, Protein profiling of CRC. A CRC section was imaged using four-color fluorescence of CD3 (magenta), CD45 (cyan), PanCK (green) and DNA (gray) to establish overall tissue morphology. Ten circular ROIs of 665 µm in diameter were then selected for detailed molecular profiling with a 44-plex oligonucleotide–antibody cocktail. Scale bar, 500 µm. **d**, Unsupervised hierarchically clustered heat map of ROI-specific nCounter digital counts across all protein targets on CRC tissue. Note that ROIs were not clustered for this heat map to allow direct comparison to the heat map shown in Fig. 5c. **e**, RNA profiling of tonsil. A tonsil section was imaged using four-color fluorescence of CD3 (magenta), CD20 (cyan), PanCK (green) and DNA (gray) to establish overall tissue morphology. Twelve circular ROIs of 300 µm in diameter were selected for detailed molecular profiling with a 96-plex RNA probe cocktail. Scale bar, 500 µm. **f**, RNA profiling of CRC. A CRC section was imaged using three-color fluorescence of CD45 (magenta), PanCK (green) and DNA (gray) to establish overall tissue morphology. Twelve circular ROIs of 300 µm in diameter were then selected for detailed molecular profiling with a 96-plex RNA probe cocktail. Scale bar, 500 µm. **g**, Unsupervised hierarchically clustered heat map of region-specific nCounter counts across all RNA targets on tonsil that had at least one ROI above background. Regional annotations based on fluorescence are labeled as CD3 enriched, CD20 enriched, Mixed and PanCK enriched. **h**, Unsupervised hierarchically clustered heat map of region-specific nCounter digital data across all RNA targets on CRC that had at least one ROI above background. Regional annotations based on fluorescence are labeled as Stroma, CD3-enirched stroma and Tumor (PanCK enriched).

With these promising results in tissues, we were interested in testing unique features of this system: the ability to reuse sections after DSP and the ability to perform high-throughput ROI profiling of high-density TMAs. While tissue sections were still intact after RNA profiling on DSP, it was unknown whether this protocol would allow tissue section reuse for follow-on studies. Re-profiling of RNA in tissues was of primary interest, as the use of multiple cycles with antibodies has already been established[7]. We performed various reuse



experiments after DSP (Supplementary Fig. 11) and found that tissue sections that were stripped and then re-profiled with DSP gave highly concordant results with the initial profiling and that RNA purified before and after DSP gave highly concordant gene expression results using the nCounter system. In addition, hematoxylin and eosin (H&E) staining was very similar before and after DSP. In all, these data suggest that the DSP RNA assay is only mildly destructive, thus allowing reuse of tissue sections for further experimental analyses. We also demonstrated the high-throughput capability of DSP by profiling 44 protein targets across a 384-core colorectal tumor TMA (Supplementary Fig. 12). As expected, we found that targets profiling similar cell types had the best correlation across all cores. Interestingly, cores from only one patient (of 192) were identified as having very high levels of PD-L1, thus highlighting the capability of DSP to detect rare phenotypes across a large set of samples.

Gridded profiling of protein and RNA

We next used the DSP instrument to profile tissue in a gridded manner, with square ROIs arrayed adjacent to each other for deep expression profiling of one concentrated area (Fig. 4a). A total of 196 gridded ROIs of 100 µm × 100 µm in size (~170 cells per ROI) were profiled with a PC-oligo-conjugated antibody panel (Fig. 4a). Remarkably, the distribution of each morphology marker could be directly recapitulated by spatially distributing nCounter counts for CD3, CD20 and PanCK (Fig. 4b). The 41 additional PC-oligo-conjugated antibodies also showed distinct spatial patterns of protein expression with different morphology visualization markers (Fig. 4c and Supplementary Fig. 13). Morphology-related patterns were observed for CD4, CD8, CD3, B7-H3, PD-1 and Ki-67 counts, while HLA-DR had similar spatial profiles to CD20. β-catenin had similar spatial profiles to PanCK, as both are related to epithelium, and VISTA and BCL2 each had a unique spatial profile distinct from those of the three visualization markers. Nearly all of the 44 PC-oligo-conjugated antibodies showed distinct spatial profiles, while the isotype negative controls (mouse and rabbit IgG) did not (Supplementary Fig. 13). For an unbiased analysis of these gridded data, hierarchically clustered heat map analysis was performed across all ROIs and targets (Fig. 4d). Specific clusters, or nodes, were selected and highlighted as unique expression signatures. ROIs within each signature were highlighted in a grid format matching the grid used for DSP profiling (Fig. 4e) and combined (Fig. 4f), recreating structures similar to those seen in the original four-channel immunofluorescence tissue image. Gridded profiling was also performed on tonsil tissue using the RNA panel. Similar results were seen, with RNA expression counts mimicking morphology patterns based on distinct immune biology within the tonsil (Fig. 4g–i).

Specific profiling of tumor regions versus surrounding microenvironment

To demonstrate the capability of DSP to capture biology on the basis of the organic tissue structure, we further characterized the colorectal tumor that had been profiled with circular ROIs in Fig. 3c. Starting with the same circular regions shown in Fig. 3c, we segmented the profiling to directly assess protein expression in two separate compartments of the tissue: (1) tumor cells and (2) the surrounding tumor microenvironment (TME), such as stroma and immune cells (Fig. 5a). The PanCK-positive tumor areas within the ROIs were exposed to UV and collected, followed by UV illumination of the PanCK-negative areas within the same ROIs (Fig. 5b). A heat map of tumor (PanCK-positive) and surrounding tissue (PanCK-negative) segments (Fig. 5c) revealed two distinct expression profiles, based on each segment, with strong enrichment of immune markers in the surrounding microenvironment or stromal compartment. The tumor compartment of ROI 1, which is located at a highly immune-infiltrated region of the invasive margin, showed the greatest enrichment of immune markers as compared to other tumor compartments. This increased expression of immune markers is likely due to high prevalence of immune infiltrates within this region of the tumor. ROI-overlayed heat maps of surrounding tissue segments showed that high, but varied, expression of CD3 and CD11c was observed in the microenvironment at the invasive margin, but CD34 was downregulated in these same regions (Fig. 5d).

Rare cell profiling and contour profiling of inflammatory bowel disease

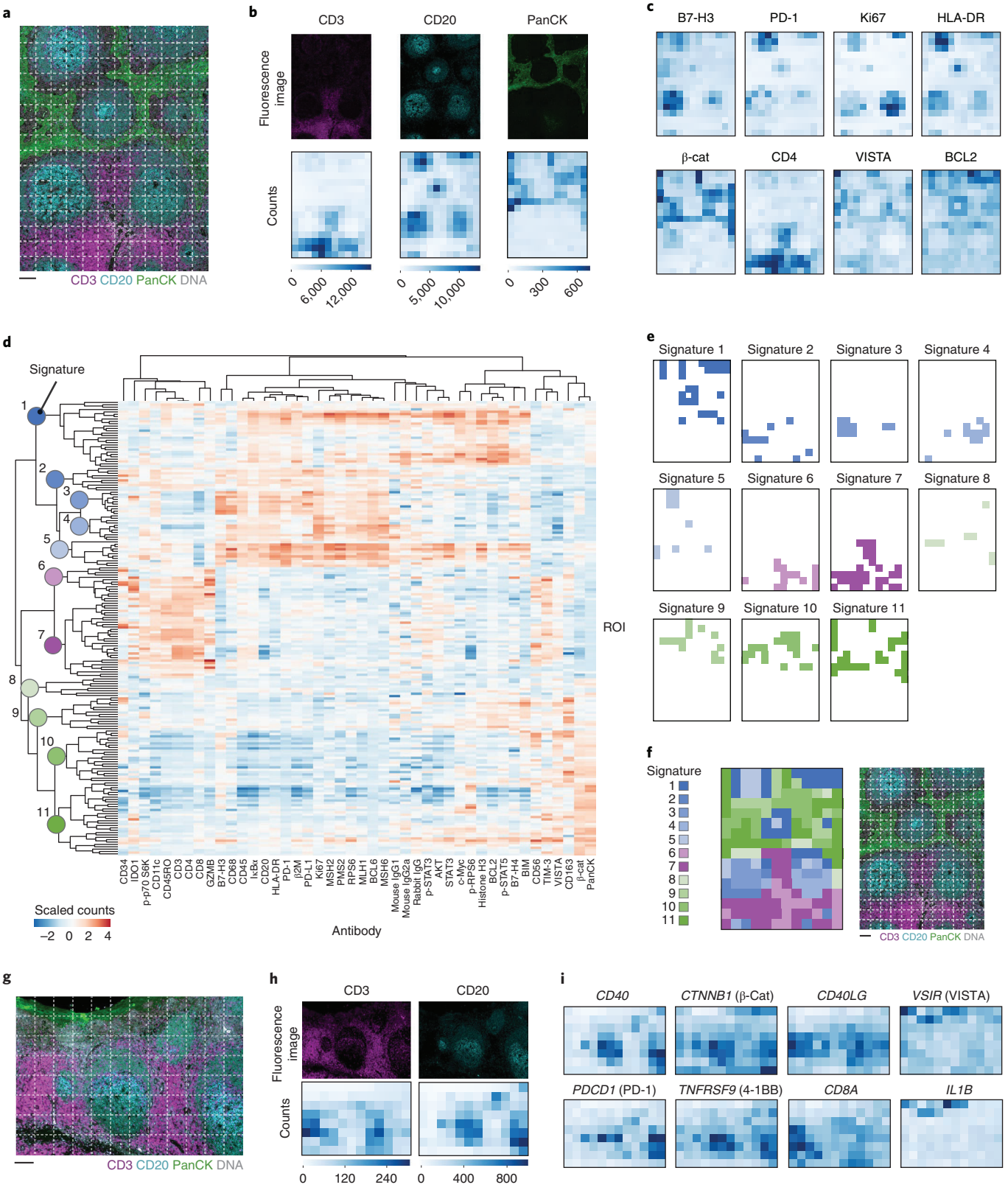
Because autoimmune diseases that manifest in the gastrointestinal system, such as inflammatory bowel disease (IBD), have diverse proinflammatory immune cells, we used DSP to characterize rare cells and areas adjacent to immune cells in the IBD tissue microenvironment. We stained samples for macrophages (CD68), T cells (CD3), epithelium (PanCK) and nuclei, along with the cocktail of 44 PC-oligo-conjugated antibodies, to interrogate IBD regions containing concentrated macrophage populations (Fig. 5e–i). We used two approaches for assessing macrophage biology in IBD samples: (1) unique cell type profiling of CD68-positive cells in various inflamed regions of the tissue (Fig. 5e,f) and (2) contour profiling of markers at varying distances from macrophage-enriched regions (Fig. 5g–i).

For rare cell profiling, macrophages identified using fluorescently labeled antibody to CD68 were profiled in six different IBD samples (Fig. 5e). This rare cell profiling of macrophages in IBD showed that some samples had higher wound healing proinflammatory macrophage markers (CD163 and B7-H3), while other samples had antigen-presenting profiles (CD11c) and other markers had no enrichment (PD-L1) (Fig. 5f), thus showing that different samples and regions had unique macrophage biology that was only captured by multiplexing immune markers.

One IBD sample used for rare cell profiling displayed an interesting pattern of macrophages (CD68) that followed the tissue architecture of the colorectal epithelium (Fig. 5g). To further assess this spatial feature, we performed contour profiling to measure protein markers that were expressed near the macrophage-enriched areas.

Fig. 4 | Protein and RNA profiling with gridded ROIs. a, Protein profiling of tonsil. A different section of the same tonsil from Fig. 3 was imaged using fluorescence of CD3 (magenta), CD20 (cyan), PanCK (green) and DNA (gray). Gridded profiling was performed by selecting 192 adjacent 100 µm × 100 µm square ROIs for detailed molecular profiling with a 44-plex oligonucleotide–antibody cocktail. Scale bar, 100 µm. **b**, Heat maps of nCounter counts obtained for CD3 (magenta), CD20 (cyan) and PanCK (green) for all ROIs compared to single-channel visualization images for the same targets. **c**, Heat maps of protein counts for other targets showing distinct patterns of expression across the grid, which are similar to the patterns for the fluorescent visualization markers. The entire dataset for this one ROI containing spatial information for 44 different antibodies, all with unique expression profiles, is in Supplementary Fig. 13. **d**, Unsupervised hierarchically clustered heat map of all ROIs and targets. Highlighted nodes represent unique signatures that were further analyzed. **e**, ROIs within each signature highlighted in a grid format matching the grid used for DSP profiling. **f**, All signatures combined, next to the original fluorescence image for comparison. Scale bar, 100 µm. **g**, RNA profiling of tonsil. A different section of the same tonsil from Fig. 3 was imaged using fluorescence of CD3 (magenta), CD20 (cyan), PanCK (green) and DNA (gray). Gridded profiling was performed by selecting 96 adjacent 100 µm × 100 µm square ROIs for detailed molecular profiling with a 96-plex RNA probe cocktail. Scale bar, 100 µm. **h**, Heat maps of digital counts obtained for *CD20* and *CD3* across ROIs compared to single-channel images of visual staining markers for the same targets. **i**, Heat maps of RNA counts for other targets showing distinct patterns of expression across the grid, which are similar to the patterns for the fluorescent visualization markers.

ROIs were selected that contained ribbons of macrophages that were aligned with the colon epithelium (Fig. 5g) and assessed for biology related to macrophages using contour profiling of markers at varying distances from the macrophage-dense region (Fig. 5h). Hierarchical clustering of digital counts normalized to ROI area for each marker clearly distinguished three clusters of immune biology on the basis of proximity to the macrophage-enriched region (Fig. 5i). For all samples, (1) moving toward the lumen (left of the CD68-positive region), there were higher amounts of T cell (CD8, CD3, CD4 and CD45RO) and natural killer (NK; CD56) effector cells; (2) as expected, the regions closest to the CD68-dense segments expressed a number of macrophage (CD163 and CD11c)



and myeloid cell markers and activation/inhibitory myeloid markers (HLA-DR, B7-H3, PD-L1 and IDO1); and (3) moving closer to the epithelium (right of the CD68-positive region), there was high expression of activation/effector molecules (PD-1, VISTA and GZMB), proliferation marker (Ki-67) and STAT signaling components (phospho-STAT3 and phospho-STAT5). Using this approach, we were able to identify three distinct clusters of immune biology on the basis of proximity to a CD68 (macrophage)-dense area, with clear effector and cytokine-related signaling in the epithelium of the IBD samples that may be part of the dysregulated inflammatory response.

High multiplexing capabilities of RNA using next-generation sequencing

We next sought to increase DSP multiplexing capabilities for RNA by using NGS as an alternative readout to the nCounter system. The NGS indexing oligonucleotide was designed to contain three features: (1) PCR primer binding sites for the addition of Illumina adaptors and dual-indexing sequences; (2) a unique molecular identifier (UMI) for the post-analysis removal of PCR duplicates; and (3) a barcoded sequence for identification of the RNA target. We generated a 96-gene NGS readout panel (928 independent RNA probes) with targeting sequences identical to the previously described nCounter RNA panel (Supplementary Table 2). In contrast to the nCounter system indexing oligonucleotide, the NGS design allows for the ability to resolve each tile along the RNA transcript (Fig. 6a,b). DSP analysis was performed on a cell pellet TMA using the nCounter and NGS panels to compare the two readout methods. We found high cross-platform concordance ($R^2 > 0.8$) among targets that showed positive expression with the nCounter system (Supplementary Fig. 14), thus validating the NGS readout workflow.

Notably, this concordance demonstrated the accuracy of the UMI-based digital counting strategy. NGS readout negative controls (glass surface only and mouse NIH3T3 cell pellet) showed significantly lower counts than the 22 human cell lines (Supplementary Fig. 15). Furthermore, the 928 probes showed differential expression patterns between the cell pellets and across all tiles. For example, the T cell marker *CD3E* showed significantly higher expression in T cell-derived cell lines (CCRF-CEM, HuT 78, Jurkat and MOLT-4) than in the other 19 cell lines (Fig. 6b). Interestingly, we observed cases where probes targeting specific exons were differentially counted between cell pellets, suggesting alternative splicing events in some cell lines (see, for example, *PTEN* in CCRF-CEM cells; Fig. 6b).

We next investigated the feasibility of expanding our analyses beyond 96 genes and 928 probes. Additional RNA probes were created to cover 1,412 genes with a total of 4,998 RNA probes. We found that the data obtained with this much larger panel demonstrated reliable scalability (Supplementary Fig. 15). Furthermore, we found high reproducibility across serial sections analyzed with this higher-plex panel (Supplementary Fig. 15).

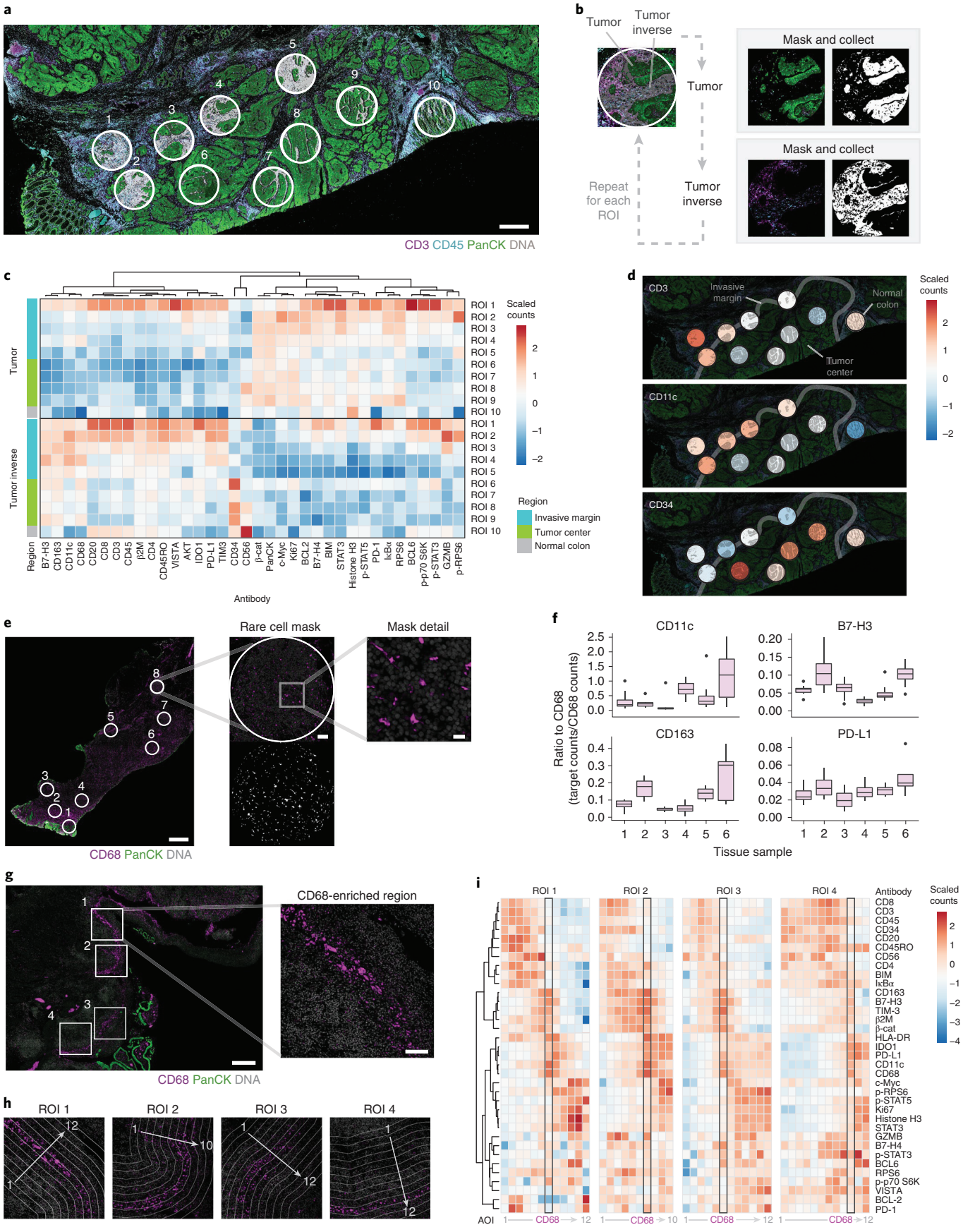
To study the performance of the 1,412-gene, 4,998-probe plex on FFPE tissues, we selected seven FFPE CRC samples with microsatellite stability (MSS) or microsatellite instability (MSI) phenotypes as assessed by IHC and evidence of immune infiltration as assessed by the tumor inflammation signature (TIS)[26] (Fig. 6c). To reveal coexpression patterns beyond the signal driven by relative abundance of immune and tumor cells, we analyzed the residual expression of genes from estimated tumor and immune cell abundances (Fig. 6d). Residuals were calculated by first estimating tumor and immune cell abundances and then calculating the $\log_2$ (fold change) between the observed and expected expression for genes. Many targets displayed differential expression across ROIs. For example, stroma-associated collagen genes were enriched across invasive margin ROIs, while remaining low in tumor ROIs. Interferon-γ-associated chemokines showed high spatial variability in three tumors, where they were consistently low across most ROIs but were elevated two- to threefold in a few other ROIs. B cell markers were generally stable and low in abundance but exhibited sporadic spikes in various ROIs. Within the tumor ROIs of one tissue, a cluster of extracellular matrix-associated genes was enriched ~10-fold, while a cluster of antigen-presentation genes was variable within all spatial contexts.

Next, gene expression signatures originally developed for bulk gene expression profiling were calculated from the residual expression data (Fig. 6f). Many interesting patterns could also be seen with this analysis. For example, the hypoxia signature was elevated sporadically within tumor interior ROIs but was uniformly low in invasive margin ROIs; the signature for antigen processing by the immunoproteasome and major histocompatibility complex (MHC) class I were elevated in similar ROIs; the mast cell signature was enriched within the immune-rich ROIs of a single tissue; and signatures of lymphoid compartment cell types (T cells, $CD8^+$ T cells, NK cells and T helper type 1 cells) were enriched within the same ROIs.

Finally, we examined chemokines in this dataset, as these high-interest targets have traditionally been difficult to quantify in FFPE tissues (Fig. 6e). Many chemokines showed high expression for at least one ROI across these samples. Notably, the observed

Fig. 5 | Segmented, rare cell and contour profiling of different compartments of colorectal tumor and IBD tissues. a, Segmented profiling on the same ROIs from CRC shown in Fig. 3c. Two segments were profiled within each circular ROI of 665 μm in diameter: the PanCK-positive area (shown in green; tumor or normal epithelium) and the remaining PanCK-negative area of the ROI (inverse; shown in white). Scale bar, 500 μm. **b**, Workflow for collection of multiple segments within the same ROI. Segments are first identified (tumor and tumor inverse). The PanCK fluorescence image is used to create a mask for only PanCK-positive cells, UV illumination is performed only within this mask, and the eluate is collected and dispensed in a microtiter plate. The slide is washed, to remove any residual PC-oligos, and the procedure is repeated for the inverse of the PanCK mask. This workflow is then repeated for each ROI across the sample. **c**, Hierarchical clustering of ROIs arranged by location (invasive margin, tumor center or normal colon) and by segment type (tumor or tumor inverse). Note that ROI 10 contained only normal colon tissue, so the 'tumor' mask for this ROI is simply used as a negative control. **d**, Heat map of tumor inverse segments for selected targets. Landmarks are annotated to highlight the relationship between target counts and the location of each ROI. **e**, Rare cell profiling of IBD tissue. Example are shown of eight ROIs and the masking scheme used to profile CD68-positive (macrophage) segments. Scale bars (left to right), 1 mm, 100 μm, 25 μm. **f**, Six IBD samples underwent rare cell profiling with $n = 8$ ROIs each. nCounter counts for all targets were normalized to CD68 counts. Differential expression of macrophage biological markers (CD11c, B7-H3, CD163 and PD-L1) varies between different IBD samples. Box plots indicate the median and upper and lower quartiles with whiskers extending to the minimum and maximum values within 1.5 times the distance between the first and third quartiles. Data beyond the whiskers are plotted as individual data points. **g**, An IBD sample was used to select four 665 μm × 665 μm square ROIs that contained a CD68-enriched ribbon of cells. Within each ROI, contour selection based on the CD68-enriched ribbon of cells was used to generate a central 50-μm-wide segment. Scale bars (left to right), 500 μm, 100 μm. **h**, Within each ROI, contour selection based on the CD68-enriched trail of cells was segmented by a 50-μm region. From each CD68-enriched area, additional segments with widths of 50 μm were selected from each side toward the lumen or epithelium. Each segment was collected and analyzed for 44-plex protein markers. **i**, Unsupervised hierarchical clustering of markers in each ROI. Black boxes highlight the CD68-enriched region.

enrichment of *CXCL14* in the MSS TIS-high sample and enrichment of *CXCL16* in the MSI TIS-high sample parallel results previously published by others[27]. In all, this 1,412-gene-plex pilot study demonstrates the feasibility of complex pathway-level spatial gene expression analysis on FFPE samples along with targeted analysis of key signaling molecules, such as chemokines.



Discussion

Spatial profiling holds great potential for revealing new aspects of biology that can give insight into disease etiology, progression, treatment and resolution. However, the complexity of these studies means they must be undertaken with consideration for extracting the most information from the most relevant samples. Multiplexed spatial analyses of FFPE samples have lagged behind modern genomic high-plex techniques for molecular profiling because they have been limited in the number of markers that can be simultaneously analyzed (typically under 10-plex). In this study, we describe DSP, a methodology that combines modern genomic detection technologies (NGS and digital optical barcodes), photochemistry (photocleavable oligonucleotides), customizable digital light projection technology (DMDs) and microfluidic sampling to advance classic IHC and ISH techniques to an integrated system capable of multiplex, spatially resolved genomic and proteomic profiling. We show that DSP can accurately detect over 40 proteins or 5,000 mRNA probes simultaneously. The DSP technology allows up to 800-plex profiling of either mRNA or protein using an optical barcode readout[28,29] (Figs. 1–5) and has the potential for even greater multiplexing using NGS readout (Fig. 6). While the plex of DSP mRNA profiling has thus far been tested for approximately 1,400 genes and 5,000 probes, there is no indication that these numbers cannot be increased, possibly to a whole-transcriptome panel.

Full results from 10 to 20 tissue sections (depending on the number of ROIs) can be generated within 1.5 to 2.5 d. The system is highly sensitive and produces direct digital counts of antibody or RNA probe abundance for quantitative downstream analysis (Figs. 1 and 2). We show that DSP has single-cell sensitivity for protein detection within an ROI (Fig. 1e,f and Supplementary Figs. 5 and 6) and RNA detection down to approximately 600 individual mRNA transcripts, or ~20 to 300 cells (Supplementary Figs. 8 and 9). Notably, this technology is only minimally destructive, allowing additional rounds of profiling on the same tissue section (Supplementary Fig. 11).

A key component of the DSP instrument is the use of a programmable DMD to perform the spatially resolved molecular profiling. This technology allows regions of different shapes or sizes, including noncontiguous areas, to be analyzed as small groups of single cells (Fig. 5). The DSP platform also allows automated profiling of tumor and stroma compartments in the TME (Fig. 5a–d), as well as rare localized cell types within different locations in a tissue, on the basis of fluorescent markers (Fig. 5e,f). The ability to assess discrete tissue compartments and cell types will aid the understanding of heterogeneity within tumor and stromal regions, which is increasingly necessary for tumor stratification and drug discovery[30,31] but remains a key challenge in the development of targeted cancer therapeutics[32].

Another unique feature of DSP is the ability to profile both protein and RNA. Protein profiling allows rapid validation of IHC results, but protein profiling technologies depend upon FFPE-compatible antibodies, which can be difficult to generate for some targets. Spatial mRNA profiling provides an alternative for targets where existing antibodies are not available. High-plex RNA and protein detection on serial sections can also identify targets where divergent mRNA and protein expression patterns reveal post-transcriptional regulation or post-translational modifications that alter protein activity or stability. On a practical level, transcript-specific synthetic oligonucleotide probes can be made rapidly, in a high-throughput manner, allowing completely custom high-plex panels to be easily developed. DSP RNA readout chemistry using NGS allows each tile to be quantified independently (Fig. 6a,b), enabling the spatial investigation of both the expression level and alternatively spliced transcripts of a given gene.

An advantage of imaging-based methods over DSP is the ability to obtain multiplexed information on each cell in the tissue slice, whereas DSP provides information within a selected ROI. However, the extremely high plex, the high dynamic range and the quantitative readout of target abundance, in conjunction with the four-color fluorescence images that are generated from each slide, allow for a nuanced analysis of tissues with DSP. We view DSP and imaging-based approaches as being complementary, and, in fact, cycling-based imaging approaches can be used in advance of a DSP workflow for those cases where multiplex guided ROI selection for DSP is desired.

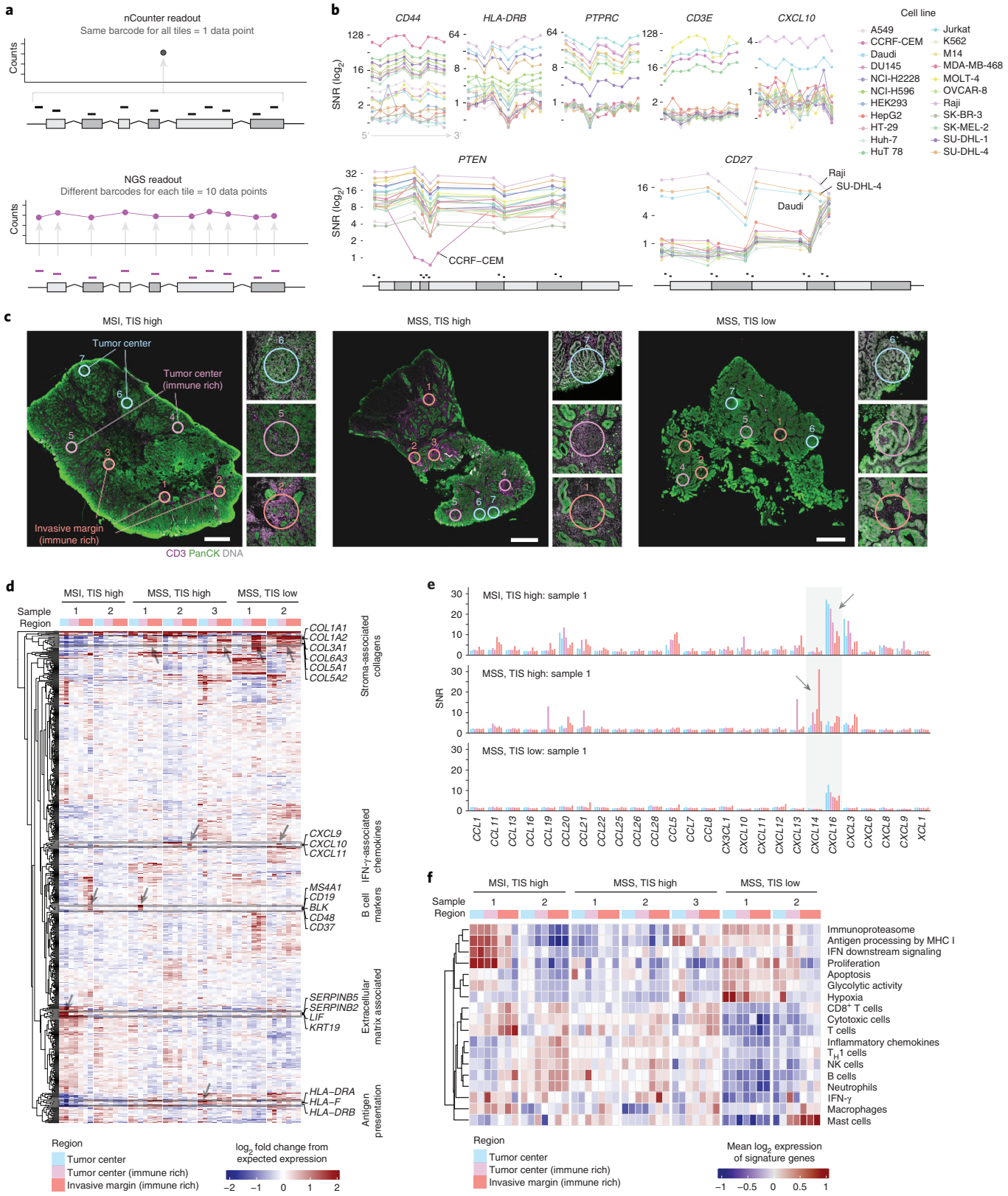
DSP does not require any complex detection instrumentation (such as MS), is highly automated and is routinely operated by nonspecialized personnel. To further simplify DSP experimentation, panels of prevalidated antibodies and RNA probes can be used, in addition to custom target detection. Additionally, clinical specimens can be stored following DSP profiling to be reprobed at a later time or used for further downstream analysis (Supplementary Fig. 11). The simplicity of using the DSP platform makes it easily optimized for future clinical applications. Early DSP prototype analysis of clinical trial samples has already provided key insights into the mechanism of action of combination immunotherapy for treatment of melanoma[33,34] and identification of biomarkers

Fig. 6 | Next-generation sequencing readout allows high-plex analyses of each transcript. a, Comparison of nCounter and NGS readouts. For the nCounter readout, all tiles of a transcript are read out with the same barcode, resulting in one data point per target. For the NGS readout, each tile has its own unique barcode, resulting in a data point for each tile. **b**, Expression of representative transcripts across all cell lines using a 928-plex NGS readout panel identical to the nCounter RNA panel except for the readout portion of the oligonucleotide. SNR (*y* axis) was calculated for each tile by dividing the geomean of the deduplicated NGS counts by the geomean of the negative probe counts in each sample. For each target, data from ten tiles are shown 5′ to 3′, from left to right. Each point corresponds to the measurement for an independent tile (*x* axis) for the transcript. For the purpose of visualization, lines have been drawn between independent points to distinguish between cell pellets. Bottom, expression of two targets with differential expression across transcripts shown with tile positions roughly indicated along the exons. Tile counts with corresponding exon position for *PTEN* (left) and *CD27* (right) are shown. Line graphs show SNR with each point aligned above the tile position in the gene structure below the graph. For the purpose of visualization, lines have been drawn to connect data from unique tiles from the same cell pellet. **c**, Gene expression was profiled spatially across seven colorectal tumors characterized for microsatellite stability status (MSS or MSI) and tumor inflammation signature (TIS). Three selected samples are shown. Within each tumor, expression data were collected from seven circular ROIs of 400 µm in diameter. Fluorescently conjugated antibodies targeting PanCK (tumor marker) and CD3 (T cell marker) were used to guide selection of ROIs composed of tumor, tumor center (immune rich) or invasive margin (immune rich) regions. Full FFPE sections with representative ROIs are shown at ×20 magnification. This experiment was performed once. Similar experiments have been performed on different samples with identical high-plex readout. Scale bars, 1 mm. **d**, Heat map of gene expression values after adjusting for tumor and immune cell content. The most variable 50% of genes are shown. Selected groups of targets are highlighted: stroma-associated collagen genes, interferon-γ-associated chemokines, B cell markers, extracellular matrix-associated genes and antigen-presentation genes (*HLA-DRA*, *HLA-DRB* and *HLA-F*). Arrows indicate areas with high levels of expression across each set of targets. **e**, Heat map of gene expression signatures of biological processes. Signatures are interpreted on a $\log_2$ scale, such that a unit difference corresponds to a doubling of the named biological process. **f**, SNR for chemokines in the panel are shown for the three samples above. Seven ROIs were profiled for each sample. Similar patterns for *CXCL14* and *CXCL16* (arrows) have previously been published by others[27].

associated with response to immune therapy in melanoma[35], as well as mechanisms of immunotherapy toxicity[36]. Thus, the DSP platform has the potential to become a widespread tool for high-resolution, multiplexed, spatially resolved protein and RNA characterization in both the research and clinical environments.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and



code availability are available at https://doi.org/10.1038/s41587-020-0472-9.

Received: 5 April 2019; Accepted: 21 December 2019;
Published online: 11 May 2020

References

1. Garon, E. B. et al. Pembrolizumab for the treatment of non-small-cell lung cancer. *N. Engl. J. Med.* **372**, 2018–2028 (2015).
2. Yu, H. et al. PD-L1 expression by two complementary diagnostic assays and mRNA in situ hybridization in small cell lung cancer. *J. Thorac. Oncol.* **12**, 110–120 (2017).
3. Yu, H., Boyle, T. A., Zhou, C., Rimm, D. L. & Hirsch, F. R. PD-L1 expression in lung cancer. *J. Thorac. Oncol.* **11**, 964–975 (2016).
4. Ting, D. T. et al. Aberrant overexpression of satellite repeats in pancreatic and other epithelial cancers. *Science* **331**, 593–596 (2011).
5. Garber, K. Oncologists await historic first: a pan-tumor predictive marker, for immunotherapy. *Nat. Biotechnol.* **35**, 297–298 (2017).
6. Sokolenko, A. P. & Imyanitov, E. N. Molecular tests for the choice of cancer therapy. *Curr. Pharm. Des.* **23**, 4794–4806 (2017).
7. Stack, E. C., Wang, C., Roman, K. A. & Hoyt, C. C. Multiplexed immunohistochemistry, imaging, and quantitation: a review, with an assessment of tyramide signal amplification, multispectral imaging and multiplex analysis. *Methods* **70**, 46–58 (2014).
8. Wang, F. et al. RNAscope: a novel in situ RNA analysis platform for formalin-fixed, paraffin-embedded tissues. *J. Mol. Diagn.* **14**, 22–29 (2012).
9. Goltsev, Y. et al. Deep profiling of mouse splenic architecture with CODEX multiplexed imaging. *Cell* **174**, 968–981 (2018).
10. Jungmann, R. et al. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat. Methods* **11**, 313–318 (2014).
11. Chen, K. H., Boettiger, A. N., Moffitt, J. R., Wang, S. & Zhuang, X. Spatially resolved, highly multiplexed RNA profiling in single cells. *Science* **348**, aaa6090 (2015).
12. Lee, J. H. et al. Highly multiplexed subcellular RNA sequencing in situ. *Science* **343**, 1360–1363 (2014).
13. Stoeckli, M., Chaurand, P., Hallahan, D. E. & Caprioli, R. M. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* **7**, 493–496 (2001).
14. Cornett, D. S., Reyzer, M. L., Chaurand, P. & Caprioli, R. M. MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat. Methods* **4**, 828–833 (2007).
15. Giesen, C. et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–422 (2014).
16. Angelo, M. et al. Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442 (2014).
17. Stoeckius, M. et al. Simultaneous epitope and transcriptome measurement in single cells. *Nat. Methods* **14**, 865–868 (2017).
18. Shahi, P. et al. Abseq: ultrahigh-throughput single cell protein profiling with droplet microfluidic barcoding. *Sci. Rep.* **7**, 44447 (2017).
19. Stahl, P. L. et al. Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science* **353**, 78–82 (2016).
20. Rodriques, S. G. et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. *Science* **363**, 1463–1467 (2019).
21. Vickovic, S. et al. High-density spatial transcriptomics arrays for in situ tissue profiling. *Nat. Methods* **16**, 987–990 (2019).
22. Keren, L. et al. A structured tumor-immune microenvironment in triple negative breast cancer revealed by multiplexed ion beam imaging. *Cell* **174**, 1373–1387 (2018).
23. Feng, Z. et al. Multispectral imaging of formalin-fixed tissue predicts ability to generate tumor-infiltrating lymphocytes from melanoma. *J. Immunother. Cancer* **3**, 47 (2015).
24. Tsujikawa, T. et al. Quantitative multiplex immunohistochemistry reveals myeloid-inflamed tumor-immune complexity associated with poor prognosis. *Cell Rep.* **19**, 203–217 (2017).
25. Remark, R. et al. In-depth tissue profiling using multiplexed immunohistochemical consecutive staining on single slide. *Science Immunol.* **1**, aaf6925 (2016).
26. Ayers, M. et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017).
27. Mlecnik, B. et al. Integrative analyses of colorectal cancer show Immunoscore is a stronger predictor of patient survival than microsatellite instability. *Immunity* **44**, 698–711 (2016).
28. Geiss, G. K. et al. Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26**, 317–325 (2008).
29. Lee, J. et al. Implementation of a multiplex and quantitative proteomics platform for assessing protein lysates using DNA-barcoded antibodies. *Mol. Cell Proteomics* **17**, 1245–1258 (2018).
30. Falcon, B. L. et al. High-content multiplexed tissue imaging and quantification for cancer drug discovery. *Drug Discov. Today* **18**, 510–522 (2013).
31. Pages, F. et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet* **391**, 2128–2139 (2018).
32. Bodenmiller, B. Multiplexed epitope-based tissue imaging for discovery and healthcare applications. *Cell Syst.* **2**, 225–238 (2016).
33. Amaria, R. N. et al. Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nat. Med.* **24**, 1649–1654 (2018).
34. Blank, C. U. et al. Neoadjuvant versus adjuvant ipilimumab plus nivolumab in macroscopic stage III melanoma. *Nat. Med.* **24**, 1655–1661 (2018).
35. Toki, M. et al. High-plex predictive marker discovery for melanoma immunotherapy treated patients using Digital Spatial Profiling. *Clin. Cancer Res.* **25**, 5503–5512 (2019).
36. Johnson, D. B. et al. A case report of clonal EBV-like memory CD4$^+$ T cell activation in fatal checkpoint inhibitor-induced encephalitis. *Nat. Med.* **25**, 1243–1250 (2019).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

Methods

Microscope and fluidics system overview. For most of the data presented, an automated imaging and sample collection instrument was developed by modifying a standard microscope and controlling it with Metamorph (Molecular Devices) and custom software. For protein or RNA detection, a multiplexed cocktail of primary antibodies or RNA-binding probes, each with UV-photocleavable indexing oligonucleotides, and/or 1–4 fluorescent markers (antibodies and/or DNA dyes) was applied to a slide-mounted FFPE tissue section (staining protocols described in the following sections). Antibodies and RNA probes used in these studies are listed in Supplementary Tables 1 and 2. The slide-mounted tissue section was placed on the stage of an inverted microscope (Nikon, Ti-E), and a custom gasket was then clamped onto the slide, allowing the tissue to be submerged in 1.5 ml of buffer solution (TBS-T, protein assay; 1× PBS-T or 2× SSC + 0.1% Tween-20, RNA assay). The gasket clamp design allowed the buffer to be accessed from above by a microcapillary tip (inner diameter of 100 µm). The microcapillary tip was connected to a syringe pump (Cavro XCalibur) primed with buffer solution, allowing for accurate aspiration of small volumes (<2.5 µl). Additionally, the tip was mounted to a separate vertically aligned *z* stage (ASI LS-50), which provided submicron tip position accuracy over the tissue. Under the microscope, wide-field fluorescence imaging was performed with epi-illumination from a visible LED light engine (Lumencor, SOLA). The tissue area of interest was then located using fluorescence imaging with a ×4 objective (Nikon). This was followed by ×20 (Nikon, ELWD) fluorescence scanning. Each ×20 image corresponds to 665 µm × 665 µm of tissue area with a CMOS camera (Hamamatsu, Flash 4.0). The ×20 images were assembled to yield a high-resolution image of the tissue area of interest. The specific ROIs for molecular profiling were then selected on the basis of the fluorescence information and sequentially processed by the microscope automation. For the data presented in Figs. 1e,f and 2e–h, and Supplementary Figs. 1, 5, 7–9, and 11, the same overall workflow (as above) is accomplished with a custom-built high-speed automated system (GeoMx Digital Spatial Profiler, NanoString) and an integrated instrument software. Also, the single-DMD configuration (custom design by NanoString) was enhanced to use a dual-DMD configuration, with two semiconductor chips operated in series, to provide enhanced background signal reduction.

The steps performed for each ROI by the microscope automation were as follows. First, the microcapillary tip was washed by dispensing clean buffer out of the capillary and into a wash station. Next, the tissue slide was washed by exchanging the buffer solution on the slide via the inlet and outlet wash ports on the gasket clamp. The microcapillary tip was then moved into position 50 µm above the ROI. The local area of tissue around the ROI was washed by dispensing 100 µl of buffer solution from the microcapillary. Then, the ROI was selectively illuminated with UV light to release the indexing oligonucleotides by coupling UV LED light with a DMD module (Andor, Mosaics3). The dual-DMD configuration used is a custom optical module built directly from a Texas Instruments DMD chip. UV LED light was collimated to be reflected from the DMD surface into the microscope objective and focused at the sample tissue (365 nm, ~125 mW cm^{-2} or 385 nm, ~800 mW cm^{-2}; 1- to 10-s exposure). Each micromirror unit in the DMD corresponds to an area of ~1 $µm^2$ of sample and reflects the UV light in a controlled pattern on the basis of the ROI selection in the image. After each UV illumination cycle, the eluent was collected from the local region via microcapillary aspiration and transferred to an individual well of a microtiter plate. As a note, in general, diffusion of oligonucleotides is not of concern for this system. To estimate potential diffusion of oligonucleotides after UV cleavage, we used a diffusion constant (D) = 12×10^{-7} cm^2 s^{-1}. This D was measured for 20-nucleotide (nt) DNA oligonucleotides in Stellwagen et al.[37]. Note that the UV-cleaved oligonucleotides for DSP are 60–66 nt. For a conservative estimate of the time of diffusion, we used 10 s, to cover the total amount of UV exposure and aspiration time. Using the equation $2Dt = \langle X^2 \rangle$ and these estimates, we calculated the maximum diffusion distance for a 20-nt oligonucleotide to be approximately 49 µm. Volume-wise, this is below nanoliter total displacement, which is a much smaller volume that the total collection volume of the ~10 µl used for these studies. Because of this, we do not expect signal loss due to diffusion. Several washing steps between each ROI collection are performed to reduce cross-contamination. Once all ROIs were processed, pools of released indexing oligonucleotides were hybridized to NanoString optical barcodes for digital counting and subsequently analyzed with an nCounter analysis system or NGS readout using the protocols below.

Custom UV illumination mask creation. Custom masks were created to define custom ROIs for UV illumination for experiments described in Fig. 5. These custom masks were used by the DMD device (described above) to determine which mirrors would be used to direct UV light to these custom regions. For each ROI, ×20 stacked tiff images of four fluorescence images were processed with ImageJ[38] to create these custom masks. For tumor and tumor inverse segments (Fig. 5a–d), each stacked tiff was split into four separate images for custom thresholding. The image for the tumor marker PanCK was thresholded manually to match the PanCK staining pattern and converted to a binary mask to generate the original tumor mask. To fill holes, a 'fill holes' secondary mask was generated using Analyze Particles with settings size = 0–700 pixels, circularity = 0.35–1.00 and show = masks. A selection was generated on the 'fill holes' secondary mask and copied onto the tumor mask, and the holes were filled to generate a tumor_filled-holes mask. To remove small particles, a 'remove particles' secondary mask was generated by inverting the tumor_filled-holes mask and using 'analyze particles' with settings size = 0–60 circularity = 0–1.00 and show = masks. A selection was generated on the 'remove particles' secondary mask and copied onto the tumor_filled-holes mask, holes were filled, the selection was removed and the mask was inverted to generate a tumor_filled-hole_removed-particles mask that will be referred to as the Original_Tumor_Mask. The Original_Tumor_Mask was inverted and eroded three times, small particles were removed, a centered circular selection of 665 µm in diameter was created and clearing was performed outside to generate the final tumor segment. The tumor inverse mask was generated by inverting the Original_Tumor_Mask and dilated four times, a centered circular selection of 665 µm in diameter was created and clearing was performed outside to generate the final tumor inverse segment.

For rare cell profiling of macrophages (Fig. 5e,f), each stacked tiff was split into four separate images. The image for the macrophage marker CD68 was thresholded manually to match the CD68 staining pattern, converted to a binary mask and dilated two times to generate the original CD68 mask. To fill holes, a 'fill holes' secondary mask was generated using 'analyze particles' with settings size = 0–1,000 pixels, circularity = 0.30–1.00 and show = masks. A selection was generated on the 'fill holes' secondary mask and copied onto the CD68 mask, and the holes were filled to generate a CD68_filled-holes mask. To remove small particles, a 'remove particles' secondary mask was generated by inverting the CD68_filled-holes mask, eroding two times and using 'analyze particles' with settings size = 0–60, circularity = 0–1.00 and show = masks. A selection was generated on the 'remove particles' secondary mask and copied onto the CD68_filled-holes mask, holes were filled, the selection was removed and the mask was inverted to generate the final CD68 segment.

For the contour profiling of macrophage-enriched regions (Fig. 5g–I), each stacked tiff was split into four images and the image for the macrophage marker CD68 was used for contour profiling. The polygon selection tool was used to trace along the center of the macrophage region, and the selection was enlarged on each side by 25 µm to create the macrophage-enriched segment of 50 µm in width. The macrophage-enriched segment was then serially enlarged by 50 µm to create radiating 50-µm-thick segments away from the macrophage-enriched segment. The wand tool was then used to move each 50-µm segment to its respective 665 µm × 665 µm mask.

Sample preparation for protein profiling and IHC. Each selected primary antibody was coupled to a unique 60-nt indexing oligonucleotide targeting either lysine or cysteine residues (NanoString, custom conjugation service). After conjugation, antibody–oligonucleotide conjugates were HPLC purified to separate the conjugated products from any remaining unconjugated antibodies or free oligonucleotide tag. All assays were performed on 5-µm FFPE sections mounted onto charged slides. Deparaffinization and rehydration of tissue was performed by incubating slides in three washes of CitriSolv (Decon Labs, 1601) for 5 min each, two washes of 100% ethanol for 10 min each, two washes of 95% ethanol for 10 min each and two washes of distilled water for 5 min each. For antigen retrieval, slides were then placed in a plastic Coplin jar containing 1× Citrate Buffer pH 6.0 (Sigma, C9999) and covered with a lid. The Coplin jar was placed into a pressure cooker (BioSB, BSB7008) that was run at high pressure and temperature for 15 min. The Coplin jar was removed from the pressure cooker and cooled at room temperature for 25 min. Slides were washed with five changes of 1× TBS-T (Cell Signaling Technology, 9997) for 2 min each. Excess TBS-T was removed from the slide, and a hydrophobic barrier was drawn around each tissue section with a hydrophobic pen (Vector Laboratories, H-4000). Slides were then incubated with blocking buffer (1× TBS-T, 5% goat serum (Sigma-Aldrich, G9023-5ML), 0.1 mg ml^{-1} salmon sperm DNA (Sigma-Aldrich, D7656) and 10 mg ml^{-1} dextran sulfate (Sigma-Aldrich, 67578-5G)) for 1 h. Slides were washed with three changes of 1× TBS-T for 2 min each. Primary antibodies were diluted in antibody diluent (Signal Stain Antibody Diluent (Cell Signaling Technology, 8112), 0.1 mg ml^{-1} salmon sperm DNA and 10 mg ml^{-1} dextran sulfate). Tissue sections were then covered with diluted primary antibody solution. Slides were incubated at 4 °C in a humidity chamber overnight. Primary antibody was aspirated from slides, and slides were washed with three changes of 1× TBS-T for 10 min each. For some experiments, antibodies were postfixed with 4% paraformaldehyde for 30 min at room temperature and then wasted twice in TBS-T. DNA was counterstained with 100 nM SYTO83 (Thermo Fisher, S11364) or 500 nM SYTO13 (Thermo Fisher, S7575) in 1× TBS-T or TBS for 15 min. Excess DNA counterstain was removed with five changes of TBS-T, and slides were processed in an automated fashion on the instrument described above.

RNA profiling probe design. DNA oligonucleotide probes were designed to bind mRNA targets. From 5′ to 3′, they each comprised a 35- to 50-nt target-complementary sequence, an iSpPC UV-photocleavable linker (Integrated DNA Technologies) and an indexing oligonucleotide sequence. Up to 13 oligonucleotide RNA detection probes were designed per target mRNA.

Sample preparation for RNA profiling. To perform ISH, 5-µm FFPE tissue sections mounted on positively charged histology slides were processed on a Leica

Bond Rx system or manually. Sections were baked, deparaffinized, rehydrated in ethanol and washed in PBS or BOND wash solution (Leica Biosystems, AR9590). Targets were retrieved for 10 to 20 min in 1× Tris-EDTA pH 9.0 buffer (Sigma-Aldrich, SRE0063; manual assay) or BOND Epitope Retrieval Solution 2 (Leica Biosystems, AR9640) at 85 °C (tonsil and cell pellet arrays) or 100 °C (colorectal samples). Next, tissues were washed in PBS, incubated in 1 μg ml^{-1} proteinase K (Thermo Fisher Scientific, AM2546) in PBS for 5 to 25 min at 37 °C and washed again in PBS. Tissues were removed from the Leica Bond, if applicable, and incubated overnight at 37 °C with hybridization solution (4–75 nM RNA detection probe per target, 100 μg ml^{-1} denatured, sonicated salmon sperm DNA (Sigma-Aldrich, D7656), 2.5% dextran sulfate (Sigma-Aldrich, 67578-5G), 0.2% BSA (Sigma, A1933), 40% deionized formamide (Thermo Fisher Scientific, AM9344) and 2× SSC (Sigma, S6639)). During incubation, slides were covered with HybriSlip Hybridization Covers (Grace BioLabs, 714022). After incubation, HybriSlip covers were gently removed by soaking in 2× SSC + 0.1% Tween-20. Two 25-min stringent washes were performed in 50% formamide in 2× SSC at 37 °C. Tissues were washed for 5 min in 2× SSC and then incubated in blocking buffer (composition above) for 30 min at room temperature in a humidity chamber. SYTO13 (100 μM) and fluorescently conjugated antibodies targeting PanCK and CD3 in blocking buffer were applied to each section for 1 h at room temperature, and sections were washed for 5 min in fresh 2× SSC. Indexing oligonucleotides were then cleaved and quantified as previously described.

nCounter hybridization assay for photocleaved oligonucleotide counting. Hybridization of cleaved indexing oligonucleotides to fluorescent barcodes was performed using nCounter Protein TagSet reagents or GeoMx Hyb code (NanoString Technologies). For some experiments where the indexing oligonucleotide was double stranded, aspirates were denatured at 95 °C for 3 to 5 min and placed on ice for 2 min. A master mix was created by adding hybridization buffer (NanoString Technologies) to the TagSet or Hyb Code. To read out the RNA detection probes, 9.6 nM of in situ capture probe was added. A 7-μl aliquot of master mix was added to each well. Depending on the experiment, 2 to 10 μl of aspirate was added to each tube, and each hybridization was brought to a final volume of 15 μl with DEPC-treated water. Hybridizations were performed at 65 °C overnight in a thermocycler. After hybridization, samples were processed using the nCounter Prep Station and Digital Analyzer according to the manufacturer's instructions.

RNAscope in situ hybridization. RNAscope ISH was performed on 5-μm FFPE tissue sections using the RNAscope LS Multiplex Fluorescent Reagent kit (ACD, 322800) according to the manufacturer's instructions. Epitope retrieval was performed for 15 min at 88 °C, and proteinase K digestions was performed at 40 °C for 15 min. Probes used included probes to *CD3* (ACD, 553978), *CD20* (ACD, 426778) and *CD40* (ACD, 578478) and were visualized using either the TSA Plus Cyanine 3 (PerkinElmer, NEL744001KT) or TSA Plus Cyanine 5 (PerkinElmer, NEL745001KT) system.

All cell pellet slides were fluorescently imaged using the Nikon Eclipse TE2000-E microscope with a ×40 objective. Images were captured with Nikon Elements commercial software. For imaging, *z* stacks (comprising ~10 focal planes) were taken from the top to bottom focal planes of each cell pellet. Maximum *z*-projection images were created with Nikon Elements software across all channels. QuPath software (https://qupath.github.io/) was used to quantify the number of RNAscope spots imaged per field of view. For this, scripts were run for each field of view that calculated the numbers of cells, through nucleus counting in the DAPI channel, and the number of fluorescent spots from the RNAscope assay. Nucleus counting did not appear accurate for some cell types that had nuclei that could not be readily distinguished; therefore, total spots per field of view were reported. Watershed detection of nuclei ('qupath.imagej.detect.nuclei.WatershedCellDetection' plugin) was performed with the following settings: 'requested Pixel Size Microns': 0.5; 'background Radius Microns': 8.0; 'median Radius Microns': 0.0; 'sigma Microns': 1.5; 'min Area Microns': 10.0; 'max Area Microns': 400.0; 'threshold': 100.0; 'watershed Post Process': true; 'cell Expansion Microns': 20.0; 'include Nuclei': true; 'smooth Boundaries': true; 'make Measurements': true. RNAscope spot detection ('qupath.imagej.detect.cells.SubcellularDetection' plugin) was performed with the following settings: 'detection[Channel]': 300.0; 'do Smoothing': true; 'split By Intensity': true; 'split By Shape': true; 'spot Size Microns': 4.0; 'min Spot Size Microns': 0.2; 'max Spot Size Microns': 25.0; 'include Clusters': true (where the detection channel was the corresponding channel used for RNAscope visualization).

Cell pellet purified RNA quantification on nCounter. RNA expression for FFPE cell pellets was analyzed to select positive and negative controls for IHC (RNA expression data not shown). One mouse cell line, NIH3T3, was included as a potential negative control for human antibodies, but gave positive signal for several antibodies in the cocktail and was therefore not rigorously analyzed with the human cell lines. This positive signal is likely due to high sequence similarity between species because many clones are developed to be cross-compatible across a wide range of species. For this, deparaffinization and rehydration of FFPE cell pellets was performed by incubating slides in two washes of CitriSolv (Decon Labs, 1601) for 2 min each and in 100% ethanol for 2 min. The slides were air dried and then dipped in 3% glycerol (MP Bio, 3055-034). Excess 3% glycerol was removed from slides, and the cell pellet was scraped in a single direction on the slide with a clean razor and transferred to a microcentrifuge tube. Lysis buffer mix consisting of 100 μl lysis buffer (Vial 1, Roche FFPET Isolation Kit, 06650775001), 100 μl of distilled water and 200 μl of 20 mg ml^{-1} proteinase K (Roche FFPET Isolation Kit, 06650775001) was added to each tube. Tubes were mixed, and samples spun down and incubated at 55 °C for 30 min while shaking at 600 r.p.m. RNA hybridization was performed using the nCounter PanCancer IO360 Gene Expression Panel. Target lists for this panel can be found on NanoString's website. For the assay, a master mix was created by adding 70 μl of hybridization buffer to the Reporter Probe tube. Aliquots (8 μl) of the master mix were added to each of the 12 hybridization tubes, and 5 μl of the RNA samples (100 ng total RNA) was added to each tube followed by 2 μl of the Capture Probe set. Tubes were mixed, and samples were spun down. Hybridizations were performed at 65 °C overnight in a thermocycler with a 70 °C heated lid. After hybridization, samples were processed using the nCounter Prep Station and Digital Analyzer according to the manufacturer's instructions. Digital counts from barcodes corresponding to protein probes were normalized to ERCC counts and housekeeping gene counts.

Tissue immunohistochemistry with DAB-based secondary detection. IHC of FFPE tissues (Fig. 1b and Supplementary Fig. 11) was performed using Leica Biosystems Bond Rx. Pretreatment of FFPE slides consisted of deparaffinization for 30 min at 72 °C using Bond Dewax solution followed by antigen retrieval for 20 min using Bond ER 1 solution. Slides were blocked for 1 h with blocking buffer (1× TBS-T, 5% goat serum (Sigma-Aldrich, G9023-5ML), 0.1 mg ml^{-1} salmon sperm DNA (Sigma-Aldrich, D7656) and 10 mg ml^{-1} dextran sulfate (Sigma-Aldrich, 67578-5G)) and then incubated with the primary antibody for 60 min. The primary antibody was diluted in antibody diluent (Signal Stain Antibody Diluent (Cell Signaling Technology, 8112), 0.1 mg ml^{-1} salmon sperm DNA and 10 mg ml^{-1} dextran sulfate). To obtain similar staining intensity for comparison, conjugated antibodies were stained with higher concentrations of antibody relative to their unconjugated counterparts. Secondary detection was performed with the Bond Polymer Refine Detection kit (Leica, DS9800). Tissue sections were imaged using a Nikon Eclipse Ci light microscope with a ×4 or ×20 objective and Nikon Elements software.

Next-generation sequencing readout for photocleaved oligonucleotide counting. Each ROI was uniquely indexed using Illumina's i5×i7 dual-indexing system. Four microliters of the aspirate containing the photocleaved oligonucleotides was used in a PCR reaction with 1 μM i5 primer, 1 μM i7 primer and 1× NSTG PCR Master Mix. Thermocycler conditions were 37 °C for 30 min, 50 °C for 10 min, 95 °C for 3 min, 18 cycles of 95 °C for 15 s, 65 °C for 60 s and 68 °C for 30 s, and a final extension of 68 °C for 5 min. PCR reactions were purified with two rounds of AMPure XP beads (Beckman Coulter) at a 1.2× bead-to-sample ratio. Libraries were paired-end sequenced on a MiSeq (2 × 75) or NextSeq (2 × 37) instrument. Each library (or ROI) was demultiplexed on instrument to produce one pair of FASTQ files per ROI. Raw sequencing reads were processed for high quality using TrimGalore and Paired-End Read Merger. Reads were then aligned to analyte barcode with Bowtie. For NGS digital counts, PCR duplicates were removed using UMI-tools with the Hamming distance set at three. The geometric mean of tiles was calculated when multiple tiles were collapsed to a single count value. In Fig. 6d–f, poorly performing probes were removed from analysis if they were outliers (Grubbs test, 3–9 tile probes; Rosner test, 10–13 tile probes) or gave low counts relative to other probes targeting the same gene.

Data analysis and visualization. For nCounter data analysis, digital counts from barcodes corresponding to protein or RNA probes were first normalized to internal spike-in controls (ERCC) to account for system variation. Counts were also compared to negative-control barcodes, which were present in the hybridization but contained no target. For these assays, the nCounter platform-specific background was lower than the background obtained from nonspecific binding of antibodies or RNA probes. The antibodies or probes with the lowest counts were typically tenfold above the nCounter system background. Box plots (Figs. 1c and 5f, and Supplementary Figs. 1 and 3) were created with the ggplot2 R package. The upper whisker extends from the hinge to the highest value that is within 1.5 times IQR of the hinge, where IQR is the interquartile range, or the distance between the first and third quartiles. The lower whisker extends from the hinge to the lowest value within 1.5 times the IQR of the hinge. Data beyond the ends of the whiskers are outliers and are plotted as points. IO360 nCounter counts from 100 ng of purified RNA were used to classify which cell lines were likely to be positive (>100 counts) or negative (<20 counts) for protein expression. The conservative threshold of 10 counts for a negative cell line was selected because this approach is likely less accurate in estimating true positives as compared to true negatives. The true negative is also likely to be more accurately called because RNA is necessary but not sufficient for protein expression. Because there was high confidence in calling these true negatives with this approach, LODs were estimated with the negative cell lines in each target and ROI condition. For this, the average plus 3 s.d. was used to estimate the LOD of each antibody, for this experiment. Targets with

nCounter counts between 20 and 100, targets expected to be expressed in all cell lines (housekeeping factors) or antibodies with no specificity (IgG controls) were labeled as having 'intermediate' or 'unknown' expression. Dot plots (Figs. 1f and 2d,f, and Supplementary Figs. 5, 6, 10, 11, 12b and 14) were created and R^2 analysis was performed with Excel (Microsoft) or were created with the ggplot2 R package (Fig. 6b and Supplementary Fig. 15). Clustered heat maps were created with the pheatmap R package (Figs. 3b,d,g,h, 4d and 5c,i and Supplementary Figs. 3 and 12) or the ComplexHeatmap package (Fig. 6d,e). Data were $\log_{10}$ transformed (log10 baseR function) and scaled (scale baseR function) for each target before clustering. Clustering was with the pheatmap 'correlation' function. For Supplementary Fig. 3b, each ROI was first normalized to the average of the geomean of all targets, and for Fig. 5c,i each ROI was first normalized by area to account for differences in ROI sizes. Gridded heat maps (Fig. 4b,c,h,i and Supplementary Fig. 13) were created with the heatmap function in the Seaborn Python package. TMA arrayed heat maps (Supplementary Figs. 2 and 12e) were created with the ggplot2 R package. Final figures were created in Adobe Illustrator.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the authors on reasonable request (see the author contributions for details on specific datasets). All unique materials used are readily available from NanoString Technologies.

References

37. Stellwagen, N. C., Magnusdottir, S., Gelfi, C. & Righetti, P. G. Measuring the translational diffusion coefficients of small DNA molecules by capillary electrophoresis. *Biopolymers* **58**, 390–397 (2001).
38. Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. The ImageJ ecosystem: an open platform for biomedical image analysis. *Mol. Reprod. Dev.* **82**, 518–529 (2015).

Acknowledgements

We thank A. VanSchoiack for editing the manuscript and H. Metz for performing IHC.

Author contributions

C.R.M., S.E.C., P.J.W. and J.M.B. wrote the manuscript, and C.M. prepared the figures. J.M.B. and G.M. conceived the project. J.M.B., D.L.D. and C.R.M. supervised the project. C.R.M. supervised experimental studies. G.G. supervised antibody labeling. C.R.M., G.T.O., S.E.C., D.Z., K.S., K.B., M.H. and J.M.B. designed experiments. K.B. performed antibody validation experiments. G.T.O. and K.B. performed protein profiling experiments. Y.L. and S.E.C. performed IHC interpretation. C.M. and P.W. developed the IHC protocol. C.M., P.W. and K.S. developed the RNA profiling protocol. D.Z. and J.M.-F. performed RNA profiling experiments. K.N. and K.S. performed RNAscope experiments. C.R.M., G.T.O., D.Z., K.S., K.B., M.H. and P.D. performed data analysis. Z.N. established the NGS data analysis pipeline. J.J., I.B.S. and D.L.D. developed initial instrumentation setup. J.J. configured the microscope and performed preliminary spatial protein profiling experiments. C.W. and I.B.S. developed the final instrumentation and fluidics setup. I.B.S. developed the instrumentation automation.

Competing interests

Patent applications have been filed related to the subject matter of this publication. G.B.M. is a co-inventor of the technology and receives research support from NanoString. All NanoString Technologies employees (C.R.M., G.T.O., S.E.C., K.B., P.D., G.G., M.H., J.J., Y.L., J.M.-F., K.N., Z.N., K.S., I.S., C.W., S.W., P.J.W., Z.Z., D.R.Z., D.L.D. and J.M.B.) declare that they are employees and shareholders of NanoString Technologies.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41587-020-0472-9.

Correspondence and requests for materials should be addressed to C.R.M. or J.M.B.

Reprints and permissions information is available at www.nature.com/reprints.

nature research

Corresponding author(s): Joseph M. Beechem, Christopher R. Merritt

Last updated by author(s): Nov 21, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Metamorph software was used to control the DSP prototype system.
Data analysis	Data analysis software packages/programs used are outlined in detail throughout the Methods section of the manuscript. These include the R and Python programing languages, Microsoft Excel, and QuPath software (v0.1.2).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the authors on reasonable request, see author contributions for specific data sets.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	The largest practical sample numbers were typically chosen, with the goal of obtaining ~96 samples per study since this is the number of samples that can be collected in a 96-well microtiter plate. We typically attempted to maximize technical replicates for cell pellets system validation studies. RNAscope studies and cell pellet titration studies to characterize the limit-of-detection of the system were limited due to the relatively high amount of effort necessary to obtain this data. In the case of RNAscope versus DSP comparisons, higher-throughput nCounter System studies on bulk purified RNA versus DSP were utilized for additional validation. No calculations to optimize sample sizes were performed.
Data exclusions	Data was excluded from some heatmaps to help highlight targets that were most relevant to the study and because of space constraints. For example, for antibody studies IgG controls would typically be removed because they provide no specific signal. For most RNA studies, targets that only gave signal above background were plotted because of space constraints. Much of the raw imaging data (e.g. IHC validation, RNAscope quantification) was not provided due to manuscript publishing space limitations.
Replication	Replicate ROIs were used for all cell pellet microarray validation studies. In most cases, experiments were repeated at least once. Given the inherent nature of FFPE sections where serial sections are not true replicates, and the inherent nature of the DSP system where small ROIs are examined, multiple sections or ROI studies are not true replicates. When we attempted studies on section-to-sections reproducibility on tissues, replicate ROIs from similar regions showing different morphology (based on fluorescent markers) were excluded.
Randomization	Randomization was not necessary for these studies, but any inadvertent experimental bias was avoided whenever possible. For example, cell pellet microarrays with 23 cell lines were constructed to cover as many targets as possible, cell lines were chosen strictly based on their commercial availability, and microarrays were constructed with cell lines in alphabetical order (they were not arranged based on cell type). For gridding experiments, samples were collected from one corner to the opposite corner by scanning across one row at a time.
Blinding	Blinding was not necessary for these studies. In most cases, these studies did not involve subjective measurements of any kind. The exception to this was the subjective interpretation of IHC for interpretation of antibody validation, which is inherent to these studies. Data collection and data analyses were typically performed by several researchers, allowing cross comparison during data interpretation to avoid any inadvertent experimental misinterpretation.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used

Fluorescently labeled antibodies:
Name, Clone, Fluorophore, Vendor, Catalog Number, Host, Clonality, Validated for IHC, Working Concentration (ug/mL)
CD20, IGEL/773, DL594, Novus Biologicals, NBP2-47840DL594, Mouse, Monoclonal, Yes, 0.5-3.0
CD3, UMAB54, AF647, Origene, UM500048, Mouse, Monoclonal, Yes, 1.0
CD3, UMAB54, DL594, Origene, UM500048, Mouse, Monoclonal, Yes, 1.3
CD3, UMAB54, AF647, Origene, UM500048, Mouse, Monoclonal, Yes, 1.0
CD45, 2B11+PD7/26, DL594, Novus Biologicals, NBP2-34528DL594, Mouse, Monoclonal, Yes, 1.0
CD45, D9M8I, AF594, Cell Signaling Technology, 14579, Rabbit, Monoclonal, Yes, 1.0
CD68, KP1, AF647, Santa Cruz Biotechnology, sc-20060 AF647, Mouse, Monoclonal, Yes, 5.0
PanCK, AE1/AE3, AF594, Biolegend, 914201, Mouse, Monoclonal, Yes, 1.9
PanCK, AE1/AE3, AF532, Biolegend, 914201, Mouse, Monoclonal, Yes, 2.0-2.5
PanCK, AE1/AE3, DL550, Novus Biologicals, NBP2-33200R, Mouse, Monoclonal, Yes, 1.0-3.0

Oligo conjugated antibodies:
NanoString provides a Protein Probe ID for each unique supplier/antibody clone in lieu of specific identifying information about the antibodies as Nanostring considers this confidential information. The custom antibody number is provided below, and in the

supplementary material, and will allow investigators who wish to replicate the work to use the same reagents. Antibodies selected for these studies were all validated for human FFPE samples. We did not validated these antibodies for other species/ applications.

Name, Nanostring Protein Probe ID, Host, Clonality, Validated for IHC, Working Concentration (ug/mL), Oligo-Conjugated Antibody IHC Validation
Akt, DPROT_00318.1, Rabbit, Monoclonal, Yes, 2.5, Yes
B7H3, DPROT_00032.1, Rabbit, Monoclonal, Yes, 2.5, Yes
B7H4, DPROT_00324.1, Rabbit, Monoclonal, Yes, 2.5, Yes
Bcl2, DPROT_00057.1, Mouse, Monoclonal, Yes, 2.5, Yes
BCL6, DPROT_00326.1, Rabbit, Monoclonal, Yes, 2.5, No
B2M, DPROT_00010.1, Rabbit, Monoclonal, Yes, 2.5, Yes
Beta Catenin, DPROT_00128.1, Rabbit, Monoclonal, Yes, 2.5, Yes
BIM, DPROT_00254.1, Rabbit, Monoclonal, Yes, 2.5, No
CD11c , DPROT_00011.1, Rabbit, Monoclonal, Yes, 2.5, Yes
CD163, DPROT_00052.1, Rabbit, Monoclonal, Yes, 2.5, No
CD20 , DPROT_00012.1, Mouse, Monoclonal, Yes, 2.5, Yes
CD3, DPROT_00013.1, Rabbit, Monoclonal, Yes, 2.5, Yes
CD34, DPROT_00048.1, Rabbit, Monoclonal, Yes, 2.5, Yes
CD4, DPROT_00014.1, Rabbit, Monoclonal, Yes, 2.5, Yes
CD45, DPROT_00323.1, Mouse, Monoclonal, Yes, 2.5, Yes
CD45RO, DPROT_00046.1, Mouse, Monoclonal, Yes, 2.5, Yes
CD56, DPROT_00322.1, Rabbit, Monoclonal, Yes, 2.5, Yes
CD68, DPROT_00006.1, Rabbit, Monoclonal, Yes, 2.5, No
CD8, DPROT_00313.1, Mouse, Monoclonal, Yes, 2.5, Yes
cMyc, DPROT_00327.1, Rabbit, Monoclonal, Yes, 2.5, No
GZMB, DPROT_00019.1, Mouse, Monoclonal, Yes, 2.5, Yes
Histone H3, DPROT_00005.1, Rabbit, Monoclonal, Yes, 0.2, Yes
HLADR, DPROT_00007.1, Rabbit, Monoclonal, Yes, 2.5, Yes
IDO1, DPROT_00033.1, Rabbit, Monoclonal, Yes, 2.5, Yes
Rabbit IgG, DPROT_00001.1, Rabbit, Monoclonal, Yes, 2.5, Yes
IkBa, DPROT_00325.1, Mouse, Monoclonal, Yes, 2.5, No
Ki67 , DPROT_00009.1, Mouse, Monoclonal, Yes, 2.5, Yes
MLH1, DPROT_00316.1, Rabbit, Monoclonal, Yes, 2.5, No
Mouse IgG1, DPROT_00002.1, Mouse, Monoclonal, Yes, 2.5, Yes
Mouse IgG2a, DPROT_00003.1, Mouse, Monoclonal, Yes, 2.5, Yes
MSH2, DPROT_00317.1, Rabbit, Monoclonal, Yes, 2.5, No
MSH6, DPROT_00319.1, Mouse, Monoclonal, Yes, 2.5, No
PanCK, DPROT_00022.1, Mouse, Monoclonal, Yes, 2.5, Yes
PD1 , DPROT_00004.1, Mouse, Monoclonal, Yes, 2.5, Yes
PD-L1, DPROT_00021.1, Rabbit, Monoclonal, Yes, 2.5, Yes
Phospho-p70S6K, DPROT_00314.1, Rabbit, Monoclonal, Yes, 2.5, No
Phospho-RPS6, DPROT_00315.1, Rabbit, Monoclonal, Yes, 2.5, Yes
Phospho-STAT5, DPROT_00328.1, Rabbit, Monoclonal, Yes, 2.5, No
PMS2, DPROT_00312.1, Mouse, Monoclonal, Yes, 2.5, No
S6, DPROT_00008.1, Mouse, Monoclonal, Yes, 1.3, Yes
STAT3, DPROT_00321.1, Mouse, Monoclonal, Yes,2 .5, Yes
Phospho-STAT3, DPROT_00320.1, Rabbit, Monoclonal, Yes, 2.5, Yes
Tim3, DPROT_00029.1, Rabbit, Monoclonal, Yes, 2.5, Yes
VISTA, DPROT_00030.1, Rabbit, Monoclonal, Yes, 2.5, Yes

Validation

Antibody validation is described in Figure 1, Supplementary Fig. 2, Supplementary Fig. 3, Supplementary Fig. 3, and Supplementary Table 1

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

All cell lines used in these studies were sourced from, propagated by, and processed by Acepix Biosciences, Inc. (Hayward, CA)

Authentication

None of the cell lines used were authenticated by Nanostring Technologies..

Mycoplasma contamination

Cell lines were not tested for Mycoplasma contamination by Nanostring Technologies.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified lines were used in these studies.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Suppliers and population characteristics for FFPE and primary PBMC samples used in these studies are described in Supplementary Table 3.
Recruitment	N/A - Patients were not recruited for these studies. All samples were obtained from commercial suppliers.
Ethics oversight	FFPE and primary PBMC samples were obtained from six commercial suppliers AMSBIO (Cambridge, MA), Astarte Biologics (Bothell, WA), BioIVT (Hicksville, NY), iSpecimen (Lexington, MA), Proteogenex (Inglewood, CA), and US Biomax Inc. (Rockville, MD). All patient samples submitted for analysis were done so with the full informed consent of the donors as described by the commercial providers: AMSBIO (Cambridge, MA) www.amsbio.com Informed consent documentation is provided with each Certificate of Analysis supplied with each sample. Example: http://www.amsbio.com/datasheets/SP-F-1.pdf Astarte Biologics (Bothell, WA) https://astartebio.com/about/ "Quality Assurance - Our dedication to quality starts with our collection procedures. We have Institutional Review Board (IRB) approved protocols — reviewed on an annual basis — for collection of blood and leukapheresis of healthy and autoimmune donors. Our reagents and kits are performance tested, and we document our manufacturing processes for consistent lot-to-lot quality. We are committed to assuring the safety and privacy of our study participants. All participants have given informed consent and are negative for HIV-1 and 2, Hepatitis B and C, and other blood-borne pathogens." BioIVT (Hicksville, NY) https://bioivt.com/about/quality-assurance (Asterand bio is now bioivt) “Informed Consent - The Department of Health and Human Services regulations for the protection of human subjects (45 CFR §46.116 and §46.117) and Good Clinical Practice (GLP), (ICH E6) require that informed consent will be sought from each prospective subject or the subject’s legally authorized representative and will be appropriately documented in writing. Furthermore, we have various forms of informed consent that are utilized for the collection of various biospecimens in different disease areas. These forms, along with their corresponding protocols, are reviewed and approved by appropriate regulatory and ethics authorities. Special situations may exist where the informed consent of the subject (or the subject’s legal representative) to participate in research is unavailable. For example, BioIVT has worked with hospitals to recover pathology specimens and data no longer needed for diagnostic purposes. In such circumstances, BioIVT obtains a waiver of informed consent from an institutional review board to enable the research.” iSpecimen (Lexington, MA) https://www.ispecimen.com/company/ethics-compliance/ "Ethical Oversight - All specimens from iSpecimen are collected under the oversight of an IRB/IEC or other appropriate ethics committee. In the U.S., this includes conformance to 45 CFR part 46, HIPAA, 21CFR Part 50/56, and the Uniform Anatomical Gift Act. Specifically: Specimens Collected Specifically for Research (“Human Subject Research”) are collected under iSpecimen’s IRB protocol with informed consent or under our partners’ IRB/IEC protocol with informed consent, both of which have been reviewed and approved by iSpecimen’s compliance team. Remnant Clinical Specimens are collected under a Non-Human Subject Research determination without consent (45 CFR 46 exemption 4), with a waiver of consent, or with informed consent, depending on the sourcing institution. Researchers can specify the level of consent desired when procuring specimens and only those meeting proper consent requirements will be delivered. Cadaver Specimens cadavers are collected under guidelines set forth by the Uniform Anatomical Gift Act with consent provided by the donor or next of kin." Proteogenex (Inglewood, CA) https://www.proteogenex.com/about-us/ethics-policy/ “Ethics Policy - ProteoGenex is committed to only providing specimens collected under ethical regulations and in accordance all applicable (local and international) laws. We have defined and adopted a strong bioethics policy including state-of-the-art procedures to achieve the highest ethical standards in all of our operations. ProteoGenex is obligated to ensure: •Every participating medical institution is committed to make certain that all clinical materials collected do not interfere in any way with patient care. •Stringent procedures are followed to protect each and every volunteer/donor. •All donors have voluntarily signed legal Informed Consent documents, which clearly state the intent of use for their donated specimens: medical research purposes with possible benefits to scientific discovery and humanity as a whole. •Clinical specimens can only be obtained with signed consent from the donor him- or herself, or in post-mortem cases, the donor’s next-of-kin. •All donors are anonymized and all identifying information is redacted from all pathology reports and summarizing clinical spreadsheets. •All specimens are de-identified with a Sample ID number that can only be traced back to the donor by ProteoGenex’s network of collection sites. •All clinical materials are obtained following official protocols, with appropriate Institutional Review Board/Independent Ethics Committee (IRB/IEC) approval, which operates in accordance with the current Federal Regulations in addition to ICH, HIPAA, and GCP guidelines pertaining to the protection of human subjects. •ProteoGenex takes pride in practicing Ethical standards set within the guidelines of the World Health Organization.” US Biomax Inc. (Rockville, MD) https://www.biomax.us/FAQs “All tissue is collected under the highest ethical standards with the donor being informed completely and with their consent. We make sure we follow standard medical care and protect the donors' privacy. All human tissues are collected under HIPPA

approved protocols. All animal tissues are collected under IACUC protocol. All samples have been tested negative for HIV and Hepatitis B or their counterparts in animals, and approved for commercial product development."

Note that full information on the approval of the study protocol must also be provided in the manuscript.