USCAP

TECHNICAL REPORT

Reproducible, high-dimensional imaging in archival human tissue by multiplexed ion beam imaging by time-of-flight (MIBI-TOF)

Candace C. Liu[1,5], Marc Bosse[1,5], Alex Kong[1], Adam Kagel[1], Robert Kinders[2], Stephen M. Hewitt[3], Sushama Varma[1], Matt van de Rijn[1], Stanisław H. Nowak[4], Sean C. Bendall[1]✉ and Michael Angelo[1]✉

Multiplexed ion beam imaging by time-of-flight (MIBI-TOF) is a form of mass spectrometry imaging that uses metal labeled antibodies and secondary ion mass spectrometry to image dozens of proteins simultaneously in the same tissue section. Working with the National Cancer Institute's (NCI) Cancer Immune Monitoring and Analysis Centers (CIMAC), we undertook a validation study, assessing concordance across a dozen serial sections of a tissue microarray of 21 samples that were independently processed and imaged by MIBI-TOF or single-plex immunohistochemistry (IHC) over 12 days. Pixel-level features were highly concordant across all 16 targets assessed in both staining intensity ($R^2 = 0.94 \pm 0.04$) and frequency ($R^2 = 0.95 \pm 0.04$). Comparison to digitized, single-plex IHC on adjacent serial sections revealed similar concordance ($R^2 = 0.85 \pm 0.08$) as well. Lastly, automated segmentation and clustering of eight cell populations found that cell frequencies between serial sections yielded an average correlation of $R^2 = 0.94 \pm 0.05$. Taken together, we demonstrate that MIBI-TOF, with well-vetted reagents and automated analysis, can generate consistent and quantitative annotations of clinically relevant cell states in archival human tissue, and more broadly, present a scalable framework for benchmarking multiplexed IHC approaches.

Laboratory Investigation (2022) 102:762–770; https://doi.org/10.1038/s41374-022-00778-8

INTRODUCTION

Immunohistochemistry (IHC) is commonly used in clinical diagnostics and basic research to visualize proteins in intact tissue using chromogenic or fluorescent reporters[1–3]. IHC staining is used routinely to guide diagnoses and therapeutic selection in the vast majority of solid tissue malignancies[4–6]. Although it remains an indispensable tool in anatomic pathology, chromogenic IHC has inherent limitations that hinder quantitative interpretation and prevent routine multiplexed staining[1,7–10]. These drawbacks are particularly limiting in the field of cancer immunotherapy, where accurate evaluation of the tumor immune microenvironment requires the simultaneous mapping of dozens of proteins[11,12]. The emerging field of spatial-omics with a multitude of analytical solutions is working toward replacement of conventional IHC-driven decision making[11,13–16]. While it is clear that new technologies and assays have the potential to generate new types of data, it is unclear whether they can reliably replicate the decision-making information in current 'gold standard' assays.

Our lab has developed multiplexed ion beam imaging by time-of-flight (MIBI-TOF), which avoids the limitations of optical imaging by using secondary mass spectrometry to image dozens of proteins on the same tissue section[17,18]. In the place of chromogenic or fluorescent reporters, MIBI-TOF uses primary antibodies that are labeled with isotopically enriched metal reporters that can be cleanly delineated and quantified using time-of-flight (TOF) mass spectrometry. Tissue sections are treated with all metal labeled primary antibodies simultaneously using a simple protocol that does not include secondary antibodies, enzymatic amplification, or cyclical staining. During MIBI-TOF analysis, the tissue is sputtered by a primary ion beam in a pixel-by-pixel fashion that liberates the metal tags as secondary ions that are subsequently quantified by TOF. Our lab routinely quantifies 40 targets simultaneously and are currently working toward increasing this capability to 60 or more[18–21]. Notably, MIBI-TOF is compatible with formalin-fixed, paraffin-embedded (FFPE) samples and can detect both low and high abundant targets with a dynamic range that spans six orders of magnitude [18].

For technologies like MIBI-TOF to be used in large translational studies and ultimately for routine clinical diagnostics, robustness and reproducibility studies are needed, as are standardized workflows for interpreting these complex data. Furthermore, it is important to show that new technologies are concordant with existing, established clinical assays. In collaboration with the National Cancer Institute (NCI) as part of the Cancer Immune Monitoring and Analysis Centers-Cancer Immunologic Data Commons (CIMAC-CIDC) network[22], we benchmarked the reproducibility of multiplexed antibody staining using MIBI-TOF. To achieve this, we compared MIBI-TOF imaging data across six

[1]Department of Pathology, Stanford University, Stanford, CA 94304, USA. [2]Office of the Director, Division of Cancer Treatment and Diagnosis, National Cancer Institute, Bethesda, MD 20892, USA. [3]Experimental Pathology Laboratory, Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA. [4]Ionpath Inc, Menlo Park, CA 94025, USA. [5]These authors contributed equally: Candace C. Liu, Marc Bosse. ✉email: bendall@stanford.edu; mangelo0@stanford.edu

Received: 17 October 2021 Revised: 23 February 2022 Accepted: 25 February 2022
Published online: 29 March 2022

Table 1. Description of tissue cores and parameter randomization.

Slide number	Serial section 1	Serial section 2	Serial section 3	Serial section 4	Serial section 5	Serial section 6
Stain Order	1	2	6	5	3	4
Day of Analysis	5	2	6	3	4	1
FOV	Order of acquisition					
CONTROL TONSIL	14	4	17	1	19	1
CONTROL TONSIL	18	20	21	7	20	2
CONTROL TONSIL	19	10	6	5	13	3
CONTROL TONSIL	11	1	2	18	5	4
CONTROL LYMPH NODE	9	3	18	2	11	5
CONTROL LYMPH NODE	15	6	11	12	10	6
CONTROL LYMPH NODE	3	7	15	9	4	7
CONTROL THYMUS	7	11	14	20	3	8
CONTROL THYMUS	5	12	5	13	8	9
CONTROL SPLEEN	17	19	16	16	21	10
SALIVARY WARTHIN SALIVARY	2	17	4	4	9	11
HIGH GRADE, CARCINOMA BLADDER	16	15	12	19	18	12
CONTROL PLACENTA	8	9	20	14	7	13
DUCTAL CARCINOMA BREAST	20	18	3	21	1	14
DUCTAL CARCINOMA BREAST	1	2	13	15	12	15
LEIOMYOSARCOMA, MET SOFT TISSUE	12	13	7	8	6	16
HIGH-GRADE MYXOFIBROSARCOMA SOFT TISSUE	6	16	19	10	15	17
FOREIGN BODY GIANT CELLS SOFT TISSUE	21	21	9	6	14	18
ADENOCARCINOMA COLON	10	8	1	3	2	19
PLEOMORPHIC DERMAL SARCOMA SKIN	13	14	10	11	16	20
SQUAMOUS CELL CARCINOMA SKIN	4	5	8	17	17	21

Table 2. MIBI-TOF staining panel.

Target	Clone
dsDNA	35I9 DNA
beta-tubulin	D3U1W
CD4	EPR6855
CD11c	EP1347Y
CD56	MRQ-42
CD31	EP3095
CD68*	D4B9C
CD8*	C8/144B
CD3*	D7A6E
Vimentin	D21H3
CD20	L26
HLA DR	EPR3692
CD45	2B11 & PD7/26
Na-K-ATPase alpha 1	D4Y7E
Pax5*	D7H5X
PanCK*	AE1/AE3

*Target used for IHC concordance analysis.

independent adjacent tissue microarray (TMA) serial sections. Antibody staining and MIBI-TOF imaging of each serial section were carried out independently and randomized with respect to all experimental parameters. In addition, we assessed MIBI-TOF concordance with single-plex IHC chromogenic staining. In keeping with the goal of CIMAC-CIDC to develop comprehensive, standardized immune monitoring analysis for biomarker discovery, we show that MIBI-TOF is a highly reproducible assay and concordant with single-plex IHC.

MATERIALS AND METHODS

Tissue microarray construction and sectioning

A TMA was constructed using human FFPE blocks from Stanford Pathology. The TMA consisted of disease-free controls as well as multiple types of carcinomas, sarcomas, and central nervous system lesions. A table of the cores included in this study can be found in Table 1 (1 mm cores). For each TMA tissue block, 13 consecutive serial sections (4 µm section thickness) were acquired. The IHC and MIBI-TOF recuts were alternated so that for the IHC concordance analysis, MIBI-TOF and IHC could be compared in adjacent sections. The assays performed on each section were as follows:

Section 1: H&E
Sections 2,4,6,8,10: Chromogenic IHC
Sections 3,5,7,9,11,13: MIBI-TOF full panel staining
Section 12: MIBI-TOF unstained control

Antibody preparation

The full MIBI-TOF panel containing 16 antibodies can be found in Table 2. The lyophilized antibody panel was obtained in a preformulated format from a commercial vendor (Ionpath Inc., "Cell Classification" human panel, https://www.ionpath.com/antibody-panels/). The panel of antibodies was provided in a lyophilized format. Pre-mixed lyophilized antibodies are designed to reduce variability and improve stability. The titer for each antibody has been validated before lyophilization by the vendor. Each vial-

aliquot contains all the antibodies pre-mixed, and a vial is sufficient to stain a single slide. For each staining procedure, a new vial of the antibody panel was suspended in antibody diluent buffer (TBS-IHC tween, Donkey serum 3%) and filtered with a 0.1 µm filter (Millipore). Individual metal labeled antibodies for IHC were also obtained from Ionpath: CD3 (D7A6E), CD8 (C8/144B), CD68 (D4B9C), Pax5 (D7H5X), PanCK (AE1/AE3). These individual antibodies were provided in solution format. The vendor recommends a 1:100 dilution factor. We used this dilution factor for all markers except for CD3, for which we determined a 1:60 dilution factor was appropriate.

Tissue staining

Tissue staining for the six MIBI-TOF recuts was carried out over 6 separate days, randomized with respect to serial section number (Table 1). Detailed staining protocols can be found here: dx.doi.org/10.17504/protocols.io.byzrpx56, dx.doi.org/10.17504/protocols.io.bf6ajrae, dx.doi.org/10.17504/protocols.io.bhmej43e. We used the Sequenza manual staining system for both MIBI-TOF and IHC staining (dx.doi.org/10.17504/protocols.io.bmc6k2ze).

Briefly, slides were baked at 70 °C overnight followed by deparaffinization and rehydration with successive washes of 30 s and three dips each in xylene (three washes), 100% ethanol (two washes), 95% ethanol (two washes), 80% ethanol (one wash), 70% ethanol (one wash), and ddH2O (two washes) with a Leica ST4020 Linear Stainer (Leica Biosystems). Tissues next underwent antigen retrieval by submerging slides in 3-in-1 Target Retrieval Solution (pH 9, DAKO Agilent) and incubating at 97 °C for 40 min in a Lab Vision PT Module (Thermo Fisher Scientific). After cooling to room temperature, slides were assembled with a cover plate on a Sequenza rack and washed with PBS wash buffer. The tissue was blocked for 1 h at room temperature with 1× TBS IHC Wash Buffer with Tween 20 with 3% (v/v) normal donkey serum (Sigma-Aldrich), 0.1% (v/v) cold fish skin gelatin (Sigma Aldrich), 0.1% (v/v) Triton X-100, and 0.05% (v/v) sodium azide. The blocking buffer was washed by adding 200 µL of antibody diluent buffer in the Sequenza upper chamber. After the diluent buffer flow through, 120 µL of the suspended panel of antibodies was added to the slides. The Sequenza rack was then placed at 4 °C overnight (16 h).

Following the overnight incubation, slides were washed twice with 1 mL of PBS wash buffer and fixed in a solution of 2% glutaraldehyde (Electron Microscopy Sciences) solution in low-barium PBS for 5 min. Slides were successively washed with 30 s and three dips per wash in PBS (one wash), 0.1 M Tris at pH 8.5 (three washes), ddH2O (two washes), and then dehydrated by washing in 70% ethanol (one wash), 80% ethanol (one wash), 95% ethanol (two washes), and 100% ethanol (two washes). Slides were dried under vacuum prior to imaging.

Single-plex chromogenic IHC was performed for five targets using the same antibody clone as used in the MIBI-TOF staining panel (Table 2). Dewaxing, epitope retrieval, blocking, hybridization, and washing for IHC tissue sections were identical to that of MIBI, with the addition of blocking endogenous peroxidase activity with 3% H_2O_2 (Sigma Aldrich) in ddH2O after epitope retrieval. After overnight primary antibody staining, tissues were washed twice with 1 mL wash buffer and the antigen:antibody complex detected with the ImmPRESS universal (Anti-Mouse/Anti-Rabbit) kit (Vector labs). Chromogenic solution (DAB) was incubated for 40 s and the reaction was stopped with PBS. The slides were counterstained with hematoxylin.

H&E tissue sections were reviewed to identify missing tissue cores, cores with large areas of necrosis, tissue folding, and any other macroscopic defect. All IHC and MIBI-TOF data were manually reviewed for excessive background staining.

MIBI-TOF imaging

The order of image acquisition for each recut was randomized with respect to section number and carried out on 6 separate days (Table 1). Within each run for each recut, the order of acquisition for the TMA cores was randomized as well (Table 1). Area normalized Xe+ primary ion dose of 9 nA*hr*mm^{-2} was used for all MIBI-TOF image acquisitions. At the end of each imaging run, sensitivity and mass resolution were quantified using a molybdenum foil standard (Supplementary Fig. 1).

Image processing

MIBI-TOF data processing was performed using a standardized pipeline from Ionpath. During data extraction, to only capture the monoatomic peaks that correspond to the monoatomic isotopic labels and avoid the polyatomic ions containing hydrogen, we integrated all peaks using a mass window of (X-0.3, X) where X is the nominal mass-to-charge (m/z) for each metal reporter tag.

Background signal arising from bare slide was removed as described previously using predefined thresholds for Ba, Ta, and Au[19,23]. Isobaric interferences arising from diatomic reporter adducts and polyatomic hydrocarbons were removed iteratively using empirically defined compensation coefficients (Supplementary Table 1). Lastly, we used a combination of several density metrics, including reachability density, connectivity, and Voronoi tessellation to estimate and remove noise from each channel (Supplementary Table 2). These density metrics leverage both spatial and intensity information. After background removal and noise subtraction, any pixel with a non-zero value was considered as positive.

Hematoxylin counterstained IHC slides were digitized with the NanoZoomer slide scanner at ×40 magnification (Hamamatsu). The threshold for hematoxylin and each DAB channel was set collectively for all images in the staining batch using a DAB channel extracted from the original counterstained image via spectral unmixing. An initial threshold approximation was set with automatic image histogram thresholding. The thresholds were then confirmed by three trained users on controls and additional images for validation and adjusted until consensus was reached (Supplementary Table 3).

Marker quantitation

In all MIBI-TOF images, for each marker, the frequency and mean intensity of positive pixels in each tissue core was quantified. The resultant value for percent positive pixels (PPP) and mean pixel intensity (MPI) for these images was then normalized with respect to the signal intensity of endogenous carbon-containing mass peaks originating from the tissue itself. Because the amount of carbon per unit area is relatively consistent between tissue samples, we can use this signal as an internal standard to normalize for variations in detector sensitivity through the course of the study. Peaks at 36, 37 and 38 m/z corresponding to 3 C, 3C1H, and 3C2H were integrated for each image (referred to here as C_i) and the average of these values across all images in all six serial sections was calculated (referred here to C_{avg}). To calculate the normalization coefficient (C_N) for a given tissue core, C_i for that tissue core was divided by C_{avg} ($C_N = C_i / C_{avg}$). Median and standard deviation for C_N across all tissue cores was 1.03 ± 0.3. Before normalization, there was a positive correlation between normalization coefficient and PPP. After normalization, this relationship was no longer evident (Supplementary Fig. 2). For IHC images, the frequency of positive pixels was quantified using deconvolved and thresholded DAB images (PPP).

Reproducibility and concordance

To assess serial section reproducibility, least squares linear regression was used to compare PPP for tissue cores in each section to the average PPP across all six serial sections. Similarly, least squares linear regression was used to compare MPI for tissue cores in each section to the average MPI across all six serial sections. For each comparison, we calculated the slope (m) and coefficient of determination (R^2). To assess concordance of MIBI-TOF with single-plex IHC, we used least squares linear regression to compare values for PPP attained by MIBI-TOF and IHC in adjacent serial tissue sections.

Cell segmentation and phenotyping

To delineate the location of single cells in MIBI-TOF images, we performed cell segmentation using the pre-trained Mesmer convolutional neural network architecture[24]. We used dsDNA as the nuclear marker and HLA class I and Na-K-ATPase as the membrane markers as input to the network. The output of Mesmer is the location of each cell in the image.

After cell segmentation, the next step of the analysis pipeline was to determine the phenotype of each individual cell. We used an unsupervised clustering approach based on the cells' mean marker expression. This method has been used in several of our previous publications[19–21,25]. For each cell, the total expression of each marker was determined, then normalized by cell size. A 99.9% normalization was applied for each marker. Cells were clustered into 100 clusters using FlowSOM[26] based on the expression of 12 phenotypic markers: CD3, CD4, CD8, CD11c, CD20, CD31, CD45, CD56, CD68, PanCK, Pax5, and Vimentin. The average expression of each of the 100 clusters was found and the z-score for each marker across the 100 clusters was computed. All z-scores were capped at

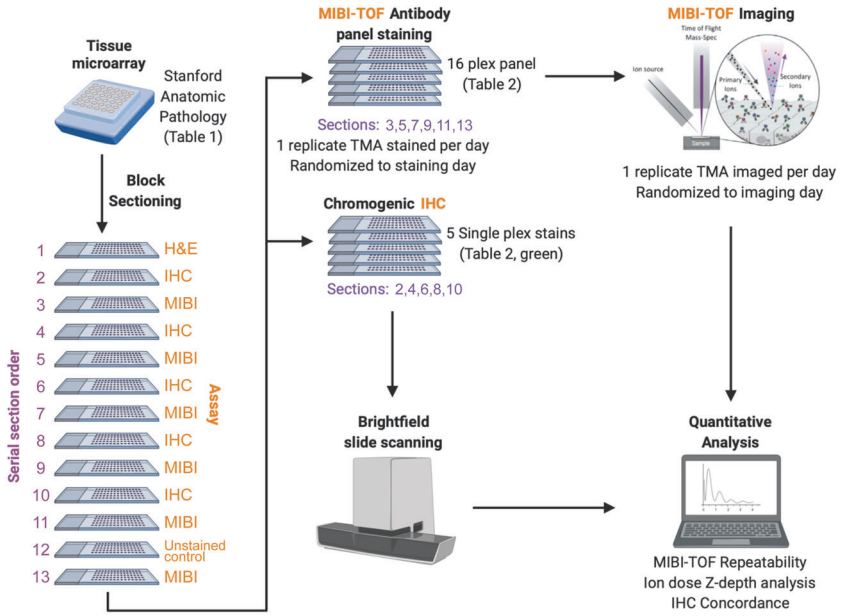

Fig. 1 MIBI-TOF validation overview. A tissue microarray (TMA) was constructed using human FFPE tissue blocks, and the TMA was serially sectioned for analysis by single-plex chromogenic IHC and MIBI-TOF. The order that each serial section was stained and imaged was randomized. We then analyzed the concordance between IHC and MIBI-TOF, as well as assessed reproducibility of MIBI-TOF between serial sections.

3, such that the maximum *z*-score was 3. Using these *z*-scored expression values, the 100 clusters were meta-clustered using hierarchical clustering into 20 meta-clusters. Each of the cell meta-clusters was then manually annotated with its cell phenotype by assessing marker expression, resulting in a total of eight cell types: B cells, T cells, dendritic cells, macrophages, NK cells, fibroblasts, endothelial cells, and epithelial cells. To assess the concordance of cell phenotyping between serial sections of the same tissue core, we quantified the number of each cell type in each image, then calculated the average Spearman correlation of these frequencies between serial section images of each core.

RESULTS

Experimental design for MIBI-TOF serial section and IHC concordance

To assess the reproducibility of MIBI as well as benchmark against single-plex chromogenic IHC, we took 13 serial sections from a TMA constructed for this study (see Methods) that included disease-free controls and multiple types of carcinomas, sarcomas, and central nervous system lesions (Fig. 1). Here, every other slide was stained and processed for either MIBI-TOF analysis or single-plex chromogenic IHC for the indicated target, in addition to slides for H&E and unstained tissue control. To control for confounding factors due to batch effects resulting from day of tissue staining and imaging, the order that each TMA was stained and imaged were randomized with respect to serial sectioning order. In addition, the order that each tissue core was acquired for each TMA was randomized as well (Table 1).

MIBI-TOF serial section concordance

For MIBI-TOF, six serial sections of each of the 21 tissue cores were stained with a 16-plex antibody panel and analyzed for a total of 126 MIBI-TOF images (Table 2). Instrument stability was assessed prior to data acquisition using a molybdenum foil standard, for which we know the expected range of ion counts. Primary ion current and mass resolution varied insignificantly over the course of the study (CV = 1.8% and 7%, respectively, Supplementary Fig. 1A, B). Minor variations in ion detector sensitivity (CV = 22%) were within historical norms and were not found to impact subsequent quantitative comparisons (Supplementary Fig. 1C). Fields in each tissue core were manually co-registered with respect to a slide scan of a reference H&E serial section, which showed high concordance of histological features (Supplementary Fig. 3). High quality MIBI-TOF imaging data was obtained for all targets in the antibody staining panel in all six TMA slides. Manual evaluation of each marker demonstrated robust and specific staining to the target cells of interest in both normal and disease tissue cores (Fig. 2).

Assay reproducibility was assessed by comparing marker intensity and abundance in each TMA serial section. An example of co-registered fields-of-view (FOVs) from six serial sections of a representative tissue core is shown in Fig. 3A, illustrating the high degree of reproducibility between the serial sections. PPP values for CD3, a marker for T cells, and CD20, a marker for B cells, in the fields for this tissue core were 22.6% ± 2.5 and 20.6% ± 8.1 (mean ± SD), respectively. The significantly higher standard deviation observed with CD20 is attributable to section-to-section changes in the area occupied by two germinal centers (marked by expression of the B cell marker CD20) as sections were taken from the tissue core.

Linear regression of normalized values for marker intensity (MPI) and marker frequency (PPP) for each tissue core in each TMA serial section (see Methods) were used to determine the degree of concordance with respect to the average across all six serial sections. For MPI, which quantifies the average marker intensity in positive pixels, the resultant slope (m) and coefficient of determination (R^2) values for all markers in all six serial sections were close to 1, with a m of 1.0 ± 0.13 and a R^2 of 0.94 ± 0.04 (mean ± SD, Fig. 3B–D). MPI captures marker intensity, thus m and R^2 values close to 1 indicate that when a marker is present, the brightness of the marker is highly concordant across serial sections. Similarly, for PPP, which quantifies the total number of pixels that are positive for a given marker, m and R^2 in all serial sections and markers were also close to 1, with a m of 1.0 ± 0.17 and a R^2 of 0.95 ± 0.04 (mean ± SD, Supplementary Fig. 4). PPP m and R^2 values close to 1 indicate that the number of positive pixels across serial sections of the same core is consistent. We also examine inter-core variability in addition to whole TMA concordance (Supplementary Fig. 5), demonstrating that serial

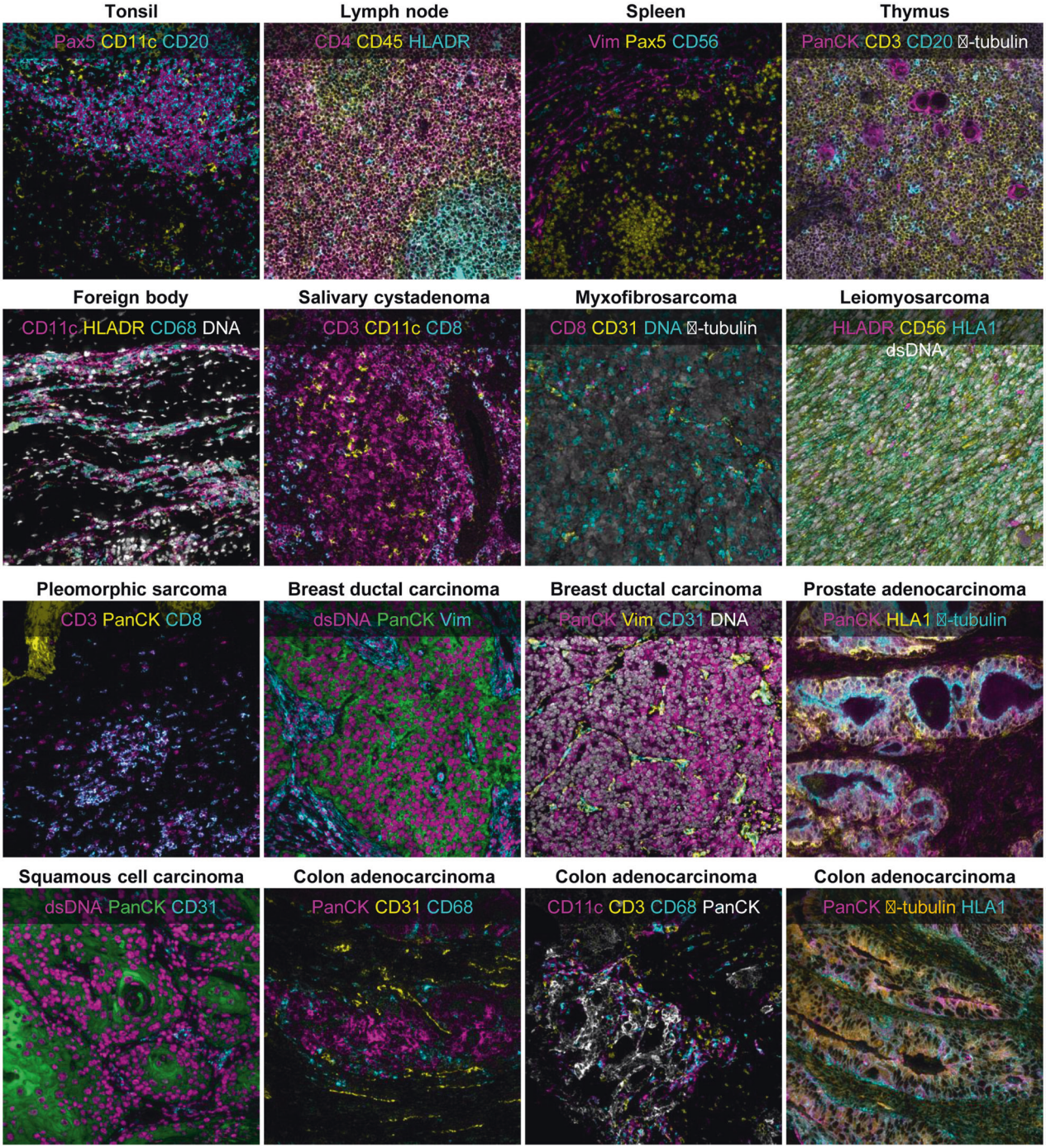

Fig. 2 Representative MIBI-TOF imaging data. High-quality MIBI-TOF imaging data was attained for 16 markers for all serial section TMA slides. Representative images of the various tissue types in the TMA are shown here.

sections within a core were relatively concordant and that variation was consistent across markers. Taken together, MPI and PPP demonstrate that MIBI-TOF images of serial sections of the same tissue core that were randomized for staining and imaging day were highly concordant, lending evidence for the reproducible nature of MIBI-TOF.

Reproducibility of cell immunophenotypic assignments by MIBI-TOF

An important step in the analysis of high-dimensional imaging data is the identification of single-cell lineages by immunophenotype in the images. Thus, we next assessed the reproducibility of single-cell annotations across serial sections of each tissue core. Our lab has recently developed a cell segmentation algorithm that can accurately identify the location of cells in solid tissue without the need for manual fine-tuning or user adjustment[24]. We used this method to delineate the location of single cells in the images (Supplementary Fig. 6), then used an unsupervised clustering approach to classify the single cells into eight cell lineages: B cells, T cells, dendritic cells, macrophages, NK cells, fibroblasts, endothelial cells, and epithelial cells (Fig. 4A, Supplementary Fig. 7). After classifying the cells in each of the images, we quantified the number of each cell type in each image and assessed concordance between serial sections by calculating the average Spearman correlation of

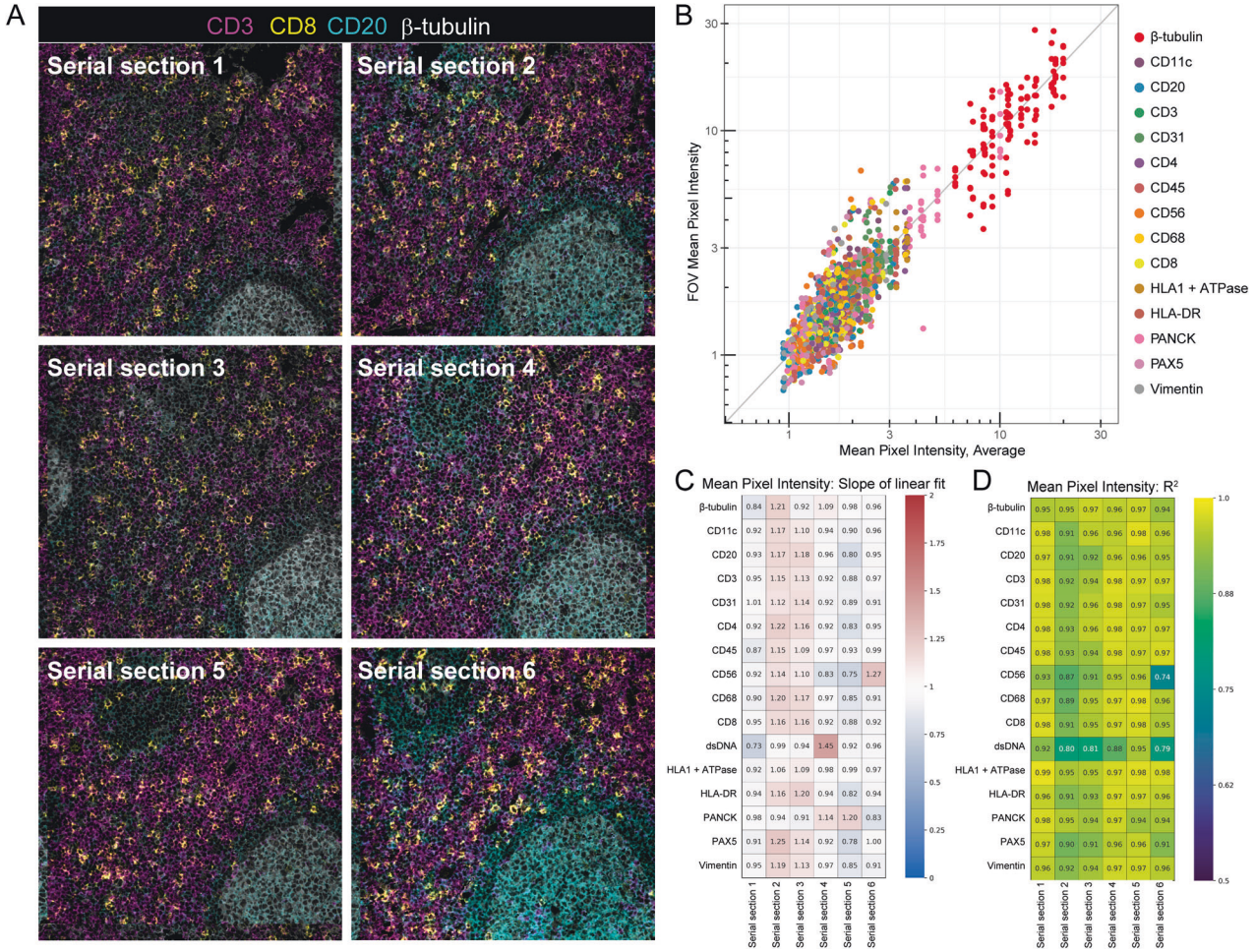

Fig. 3 Concordance of Mean Pixel Intensity by MIBI-TOF. A Representative MIBI-TOF images from serial sections of the same TMA core of lymph node tissue. **B** Plot of the average Mean Pixel Intensity (MPI) of all FOVs of the same TMA core vs. the MPI of each FOV. Each color represents a different marker. The gray line is a reference line with slope 1 and intercept 0. We performed least squares linear regression of the average MPI of the core vs. the MPI of all six serial sections and found the slope *m* (**C**) and coefficient of determination R^2 (**D**) for all 16 markers.

these frequencies between serial sections of each tissue core (Fig. 4B). When comparing serial section MIBI-TOF images, we observed that broad histological features were conserved, such as germinal centers or vessels (Figs. 3A, 4C, Supplementary Fig. 7A, B). The average Spearman correlation was high, 0.94 ± 0.05 (mean ± SD), demonstrating that cell phenotyping, a necessary step in high-dimensional image analysis, was reproducible between MIBI-TOF images of the same tissue core.

We expected variation in these Spearman correlation values, since the serial sections of the same tissue core are sections of 4 µm each (with additional serial sections taken for IHC concordance analysis). In total, 13 serial sections were obtained for each tissue for MIBI-TOF or IHC (see Methods). Therefore, we would expect true biological differences when comparing serial sections up to 48 µm apart. To quantify these differences, we compared serial sections of each tissue core as a function of slide distance (Supplementary Fig. 8). For tissues with multicellular histological structures, there is an evident linear relationship between slide similarity and slide distance (Supplementary Fig. 8A). For example, the size of germinal centers in lymph node varied in a predictable manner consistent with serial sectioning of an ellipsoid structure. However, for other tissues such as carcinomas that lack periodic structures, we see very little difference in cell composition as we cut through the tissue core (Supplementary Fig. 8B). Taken together, the majority of differences were determined to be attributable largely to variations in tissue composition.

Concordance of MIBI-TOF with single-plex chromogenic IHC

Chromogenic IHC is the standard modality for disease prognosis and therapeutic selection in the vast majority of solid tumors. Therefore, we assessed concordance to determine if multiplexed imaging by MIBI-TOF can quantitatively recapitulate single-plex immunoassays. Similar to the analyses described above that were used to compare MIBI-TOF data across serial sections, the frequency of marker positive pixels in MIBI-TOF images co-registered with single-plex chromogenic IHC stains were compared by linear regression across all 21 tissue cores. Because we alternated recuts for MIBI-TOF and IHC, we could compare MIBI-TOF and IHC on adjacent serial sections. Single-plex chromogenic IHC images were acquired for five targets: CD3, CD8, Pax5, PanCK, and CD68. Marker positivity by single-plex IHC for CD3, CD8, Pax5, PanCK, and CD68 in adjacent serial sections were compared with the respective values found by MIBI-TOF using linear regression (Fig. 5, Supplementary Fig. 9). For all five targets, concordance of MIBI-TOF and IHC was high, with an R^2 of 0.85 ± 0.08 (mean ± SD, all $R^2 > 0.7$). Of note, the slope of the linear fits comparing IHC and MIBI-TOF were consistently less than unity (Fig. 5, Supplementary Fig. 9A). The reason for this proportional difference between the two methods can be attributed to differences in signal detection. IHC uses secondary antibody amplification, while MIBI-TOF only directly measures a tagged primary antibody. Consequently, the non-linear amplification used in chromogenic IHC will result in more positive pixels on a per cell basis[27].

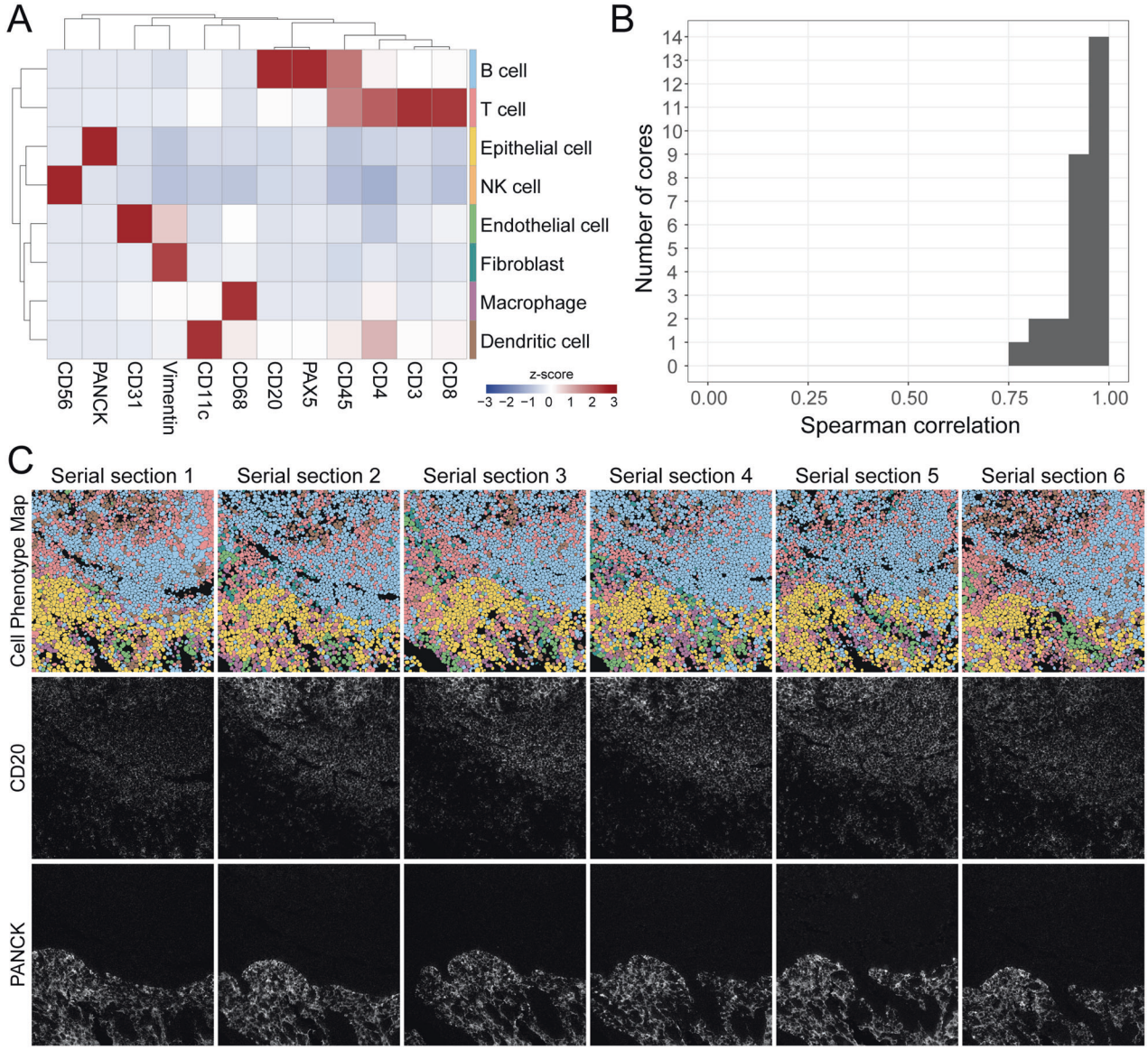

Fig. 4 Reproducibility of cell phenotyping by MIBI-TOF. A Cells were assigned to a phenotype using an unsupervised clustering approach and manually annotated into eight major cell types. Expression values were *z*-scored for each marker. High *z*-scores are an indicator of high marker specificity. **B** The Spearman correlation between all serial sections of each TMA core using the frequency of cell types in each FOV. **C** Representative images of six serial sections of the same TMA core of tonsil tissue. Top row: Cell phenotype map colored according to the eight cell types shown in (**A**). Middle and bottom rows: MIBI-TOF images showing CD20 and PanCK expression.

DISCUSSION

In this study, we assessed the reproducibility of MIBI-TOF by evaluating assay concordance on six serial sections from a TMA, including disease-free tonsil, lymph node, thymus, and spleen, in addition to multiple types of carcinomas, sarcomas, and central nervous system lesions. Importantly, each serial section was stained independently on 6 separate days with a 16-marker staining panel and subsequently analyzed on 6 different days. 21 tissue cores were analyzed in each serial section for a total of 126 MIBI-TOF images. The staining intensity and frequency of positive pixels for each serial section (MPI and PPP, respectively) were compared against average values across all runs using least squares linear regression. In all comparisons, concordance between serial sections of the same tissue core was high, with m of 1.0 ± 0.13 and R^2 of 0.94 ± 0.04 when assessing MPI and m of 1.0 ± 0.17 and R^2 of 0.95 ± 0.04 when assessing PPP (mean ± SD). The staining pattern of each marker was manually inspected to confirm, where relevant, appropriate subcellular localization and histologic distribution. For example, CD20 staining in positive regions of tonsil and lymph node in all images was verified to be membrane-localized and enriched in follicles. Furthermore, we assessed the reproducibility of cell phenotyping across serial sections and found that across serial sections, the location and lineage of cells identified using our data analysis pipeline were highly concordant, with an average Spearman correlation of cell type frequencies between serial sections of 0.94 ± 0.05 (mean ± SD). Importantly, the data analysis pipeline used to go from MIBI-TOF images to cell phenotypes was nearly fully automated. Cell segmentation and clustering were fully automated with no user input and only minimal user intervention was needed when labelling cell clusters with cell annotations, demonstrating that the biological annotation of MIBI-TOF images is robust without the need for significant researcher intervention. Taken together, the combination of consistent staining, imaging, and feature extraction illustrates the highly reproducible nature of MIBI-TOF. In addition, we assessed concordance of PPP by MIBI-

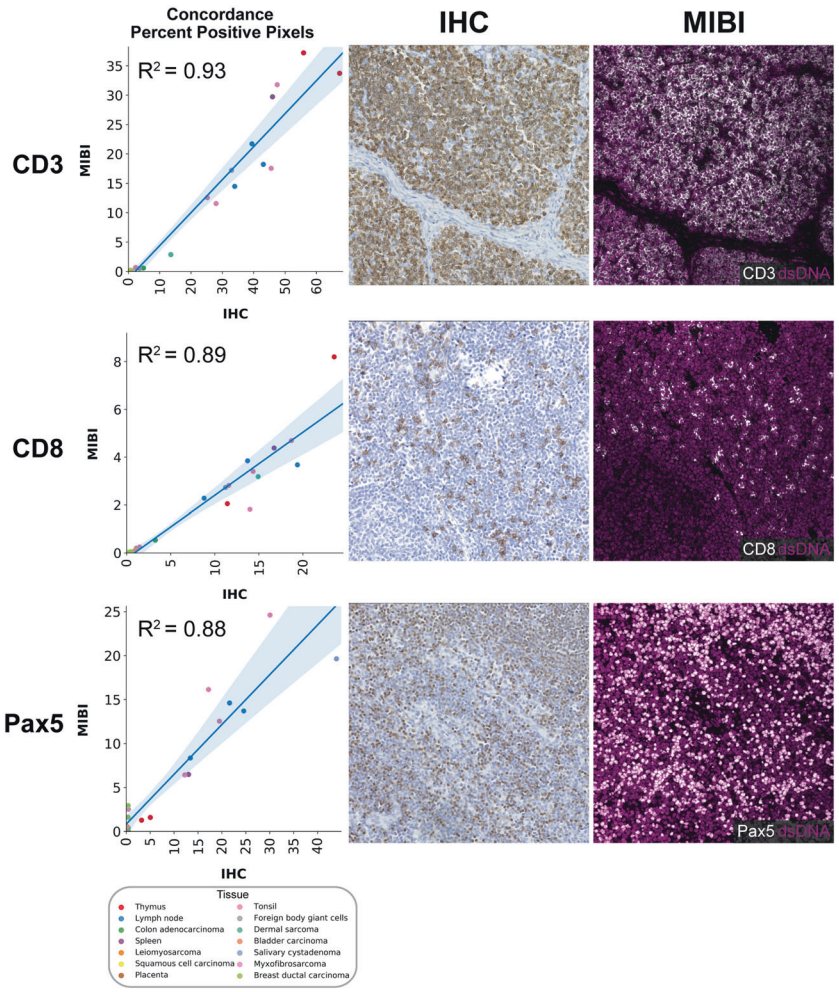

Fig. 5 Concordance of MIBI-TOF with single-plex chromogenic IHC. Representative images (CD3: tonsil, CD8: lymph node, Pax5: lymph node) of the comparison of MIBI-TOF images co-registered with single-plex chromogenic IHC stains. Each data point represents the PPP by MIBI-TOF and IHC for a single tissue core. Color of data points indicates tissue type. Shaded area represents 95% confidence interval. Additional markers are shown in Supplementary Fig. 9.

TOF for five markers with single-plex IHC and found that MIBI-TOF and single-plex IHC were highly concordant with an average R^2 of 0.85.

One limitation of validation studies involving serial tissue sections is that exact biological replicates cannot be obtained since each tissue section is somewhat different and is not a homogenized bulk sample. In this study, since the tissue sections are 4 µm each and a total of 13 serial sections were acquired, the first and last tissue section were separated axially by 48 µm. Also, when balancing assay replication vs. concordance with an orthogonal assay, like single-plex IHC here, one needs to prioritize what physically adjacent slides are used for. In this case, we selected physically adjacent slides for benchmarking single-plex IHC since any one marker's distribution would be heavily dependent on cell composition. Lymphocytes that are approximately ~10 µm would be best preserved in adjacent sections. Although there will be true biological differences in this volume of tissue, this study represents the next best situation, in which we compared adjacent serial sections as a proxy for true replicates. Furthermore, this was a single-site study, in which all staining and imaging was performed in our laboratory. We are currently planning a multi-site study to assess inter-institutional differences in instrument performance and tissue staining. Still, we believe the overall metrics for benchmarking outlined here will be sufficient to assess these future efforts.

This study was performed as a part of the Cancer Immune Monitoring and Analysis Centers Cancer Immunology Data Commons (CIMAC-CIDC) network[22], which is a NCI Cancer Moonshots initiative to provide the technology and expertise for immunotherapy clinical trials. Multiplexed tissue imaging is an integral assay for fully characterizing the tumor immune microenvironment, which requires the simultaneous profiling of multiple tumor and immune cell types. MIBI-TOF can routinely provide quantitative, multiplexed imaging data that is back compatible with archival FFPE tissue and conventional anatomic pathology workflows. Technical innovations in reagents and instrumentation will further increase the throughput and multiplexing ability of MIBI-TOF. In the future, we envision that MIBI-TOF will not only be used for reproducible basic science research, but also enable the adoption of quantitative spatial signatures in the clinic for more accurate diagnosis and therapeutic selection. At the same time, as efforts to develop other IHC-centric spatial imaging technologies evolve, the analytical framework and approach laid out in this study should serve as a guide for their assessment.

DATA AVAILABILITY

The datasets generated and/or analyzed during the current study are available on Zenodo: https://doi.org/10.5281/zenodo.5945388. The analysis code is available on Github: https://github.com/angelolab/publications/tree/master/2022-Liu_Bosse_etal_CIMAC.

REFERENCES

1. Levenson, R. M., Borowsky, A. D. & Angelo, M. Immunohistochemistry and mass spectrometry for highly multiplexed cellular molecular imaging. *Lab. Investig.* **95**, 397–405 (2015).
2. Matos, L. L., de, Trufelli, D. C., de Matos, M. G. L. & da Silva Pinhal, M. A. Immunohistochemistry as an important tool in biomarkers detection and clinical practice. *Biomark. Insights* **5**, 9–20 (2010).
3. Chlipala, E. A., et al. An Image Analysis Solution For Quantification and Determination of Immunohistochemistry Staining Reproducibility. *Appl. Immunohistochem. Mol. Morphol.* **28**, 428–36 (2020).
4. Thunnissen, E., et al. The Use of Immunohistochemistry Improves the Diagnosis of Small Cell Lung Cancer and Its Differential Diagnosis. An International Reproducibility Study in a Demanding Set of Cases. *J. Thoracic Oncol.* **12**, 334–46 (2017).
5. Yatabe, Y., et al. Best Practices Recommendations for Diagnostic Immunohistochemistry in Lung Cancer. *J. Thoracic Oncol.* **14**, 377–407 (2019).
6. Zaha, D. C. Significance of immunohistochemistry in breast cancer. *World J. Clin. Oncol.* **5**, 382–92 (2014).
7. McCabe, A., Dolled-Filhart, M., Camp, R. L. & Rimm, D. L. Automated quantitative analysis (AQUA) of in situ protein expression, antibody concentration, and prognosis. *J. Natl Cancer Inst.* **97**, 1808–15 (2005).
8. Taube, J. M., et al. The Society for Immunotherapy of Cancer statement on best practices for multiplex immunohistochemistry (IHC) and immunofluorescence (IF) staining and validation. *J. Immunother. Cancer* **8**, e000155 (2020).
9. Zimak, J., Schweller, R. M., Duose, D. Y., Hittelman, W. N. & Diehl, M. R. Programming in situ immunofluorescence intensities through interchangeable reactions of dynamic DNA complexes. *Chembiochem* **13**, 2722–8 (2012).
10. Neher, R. & Neher, E. Optimizing imaging parameters for the separation of multiple labels in a fluorescence image. *J. Microsc.* **213**, 46–62 (2004).
11. Patel SS, Rodig SJ. Overview of Tissue Imaging Methods. in *Biomarkers for Immunotherapy of Cancer: Methods and Protocols* (eds. Thurin, M., Cesano, A. & Marincola, F.M.) 455–65 (Springer New York, 2020). https://doi.org/10.1007/978-1-4939-9773-2_21.
12. Binnewies, M., et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat. Med.* **24**, 541–50 (2018).
13. Bodenmiller, B. Multiplexed Epitope-Based Tissue Imaging for Discovery and Healthcare Applications. *Cell Syst.* **2**, 225–38 (2016).
14. Goltsev, Y., et al. Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging. *Cell* **174**, 968–.e15 (2018).
15. Schulz, D., et al. Simultaneous Multiplexed Imaging of mRNA and Proteins with Subcellular Resolution in Breast Cancer Tissue Samples by Mass Cytometry. *Cell Syst.* **6**, 25–36.e5 (2018).
16. Giesen, C., et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods* **11**, 417–22 (2014).
17. Angelo, M., et al. Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–42 (2014).
18. Keren, L., et al. MIBI-TOF: a multiplexed imaging platform relates cellular phenotypes and tissue structure. *Sci. Adv.* **5**, eaax5851 (2019).
19. Keren, L., et al. A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373–.e19 (2018).
20. Risom, T., et al. Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma. *Cell* **185**, 299–310.e18 (2022).
21. McCaffrey, E. F., et al. The immunoregulatory landscape of human tuberculosis granulomas. *Nat. Immunol.* https://doi.org/10.1038/s41590-021-01121-x (2022).
22. Chen, H. X., et al. Network for Biomarker Immunoprofiling for Cancer Immunotherapy: Cancer Immune Monitoring and Analysis Centers and Cancer Immunologic Data Commons (CIMAC-CIDC). *Clin. Cancer Res.* **27**, 5038–48 (2021).
23. Baranski, A., et al. MAUI (MBI Analysis User Interface)—An image processing pipeline for Multiplexed Mass Based Imaging. *PLOS Comput. Biol.* **17**, e1008887 (2021).
24. Greenwald, N. F., et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-021-01094-0 (2021).
25. Liu, C. C., et al. Multiplexed Ion Beam Imaging: insights into Pathobiology. *Annu. Rev. Pathol. Mech. Dis* **17**, 403–23 (2022).
26. Van Gassen, S., et al. FlowSOM: Using self-organizing maps for visualization and interpretation of cytometry data. *Cytometry A* **87**, 636–45 (2015).
27. Rimm, D. L. What brown cannot do for you. *Nat. Biotechnol.* **24**, 914–6 (2006).

AUTHOR CONTRIBUTIONS

S.V. and M.R. created the tissue microarray. M.B. performed experiments and imaging. A.K., A.K., C.C.L., and S.H.N. performed data analysis. C.C.L. and M.A. wrote the paper. S.C.B., R.K., and S.M.H. reviewed the paper. S.C.B. and M.A. supervised the work.

FUNDING

This work was supported by DOD EOH W81XWH2110143, 1-DP5-864 OD019822, 1R01AG056287, 1R01AG057915, R01AG068279, 1UH3CA246633, 1U24CA224309, and the Bill and Melinda Gates Foundation.

COMPETING INTERESTS

M.A. and S.C.B. are inventors on patents related to MIBI technology. M.A. and S.C.B. are consultants, board members and shareholders in Ionpath Inc.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41374-022-00778-8.

Correspondence and requests for materials should be addressed to Sean C. Bendall or Michael Angelo.