Full Title: Micro-region transcriptomics of fixed human tissue using Pick-Seq

Short Title: mrSEQ using Pick-Seq

Zoltan Maliga[1†], Ajit J. Nirmal[1,2†], Nolan G. Ericson[3†], Sarah A. Boswell[1,4], Lance U'Ren[3‡], Rebecca Podyminogin[3§], Jennifer Chow[3], Yu-An Chen[1], Alyce A. Chen[1], David M. Weinstock[2], Christine G. Lian[4], George F. Murphy[4], Eric P. Kaldjian,[3] Sandro Santagata[1,5] and Peter K. Sorger[1,4]*

[1]Laboratory of Systems Pharmacology, Harvard Program in Therapeutic Science, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA.

[2]Department of Medical Oncology, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215, USA.

[3]RareCyte Inc., 2601 4th Avenue. Suite 500, Seattle, WA 98121, USA.

[4]Department of Systems Biology, Harvard Medical School, 200 Longwood Avenue, Boston, MA 02115, USA.

[5]Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 75 Francis Street, Boston, MA 02115, USA.

[†]These authors contributed equally

[‡]Current address: Gilead Sciences Inc., 199 E Blaine Street, Seattle, WA 98102, USA.

[§]Current address: Sana Biotechnologies, 188 East Blaine Street, Seattle, WA 98102, USA.

*Corresponding author P.K.S. (peter_sorger@hms.harvard.edu) cc: alyce_chen@hms.harvard.edu

ABSTRACT

Spatial transcriptomics and multiplexed imaging are complementary methods for studying tissue biology and disease. Recently developed spatial transcriptomic methods use fresh-frozen specimens but most diagnostic specimens, clinical trials, and tissue archives rely on formaldehyde-fixed tissue. Here we describe the Pick-Seq method for deep spatial transcriptional profiling of fixed tissue. Pick-Seq is a form of micro-region sequencing in which small regions of tissue, containing 5-20 cells, are mechanically isolated on a microscope and then sequenced. We demonstrate the use of Pick-Seq with several different fixed and frozen human specimens. Application of Pick-Seq to a human melanoma with complex histology reveals significant differences in transcriptional programs associated with tumor invasion, proliferation, and immuno-editing. Parallel imaging confirms changes in immuno-phenotypes and cancer cell states. This work demonstrates the ability of Pick-Seq to generate deep spatial transcriptomic data from fixed and archival tissue with multiplexed imaging in parallel.

INTRODUCTION

Although tissues have long been imaged using chemical stains, immunohistochemistry (IHC), and *in situ* hybridization (*1*), the introduction of single-cell RNA sequencing (scRNA-seq) has revealed an unexpected diversity of cell types and states (*2*, *3*). Recently announced tissue atlases (*4–6*) combine multiplexed tissue imaging (*7–9*) with transcriptomics (*10*) to enable joint molecular and morphological analysis of human and animal tissues. Because the molecular programs that specify tissue architecture are of inherent interest, and tissues are mixtures of many cell types, there is long-standing interest in subjecting specific regions of a tissue to RNA and DNA sequencing. The earliest approaches involved manual tissue dissection using needles (*11*) but Laser Capture Microdissection (LCM)(*12*) was the first widely used approach. In LCM, an infrared laser melts an ethylene-vinyl acetate layer onto selected regions of tissue, allowing cells in that region to be recovered and processed for sequencing. More

recent spatial transcriptomics approaches (*10*, *13*) have made it possible to interrogate gene expression at near single-cell resolution.

Most methods for spatial transcript profiling require or work far better with, fresh frozen samples (*13*) for the simple reason that fixation damages nucleic acids. Frozen sections (those mounted in optimal cutting temperature medium; OCT) are used for intra-operative patient management, but in both pre-clinical and clinical settings formaldehyde-fixed paraffin-embedded (FFPE) specimens are more common for multiple reasons. FFPE sections are the diagnostic standard in clinical pathology, with pathology services in many teaching hospitals processing $>10^5$ specimens/year; morphology is also better preserved in FFPE than OCT specimens. Archives of FFPE tissue exist for many diseases and multicenter clinical trials prefer FFPE specimens because fixed samples are easily stored and exchanged. Finally, many specimens cannot be allocated for use in research studies until they have undergone microscopic evaluation for diagnostic purposes, for example, to exclude the possibility of invasive disease. Thus, a substantial need exists for transcript profiling methods that are optimized for use with FFPE tissue in conjunction with multiplexed imaging of the same specimen. Remarkably, sequencing of single cells has recently been demonstrated using micro-regions of FFPE subjected to LCM(*14*). This result inspired us to develop approaches to micro-dissection and sequencing that were compatible with multiplexed immunofluorescence-based tissue imaging.

In this paper, we describe a method for isolating and sequencing micro-regions of interest (mROIs) from tissue guided by imaging of the same or serial (adjacent) tissue sections. The Pick-Seq approach evolved from simple and robust methods for mechanical recovery of tissue for sequencing (*15*) with the integration of compact robotic manipulators and automated mechanical recovery of tissue samples. Whereas many multiplexed imaging technologies focus on small fields of view (*8*, *9*), our approach (*7*) involves whole-slide, multiplexed, subcellular resolution imaging of specimens up to

several square centimeters in area; this generates single-cell imaging on a scale sufficient to provide physiological context for mROIs (*16*). Whole slide imaging makes it possible to analyze tissue structures of a wide range of spatial scales and is regarded by the FDA as a diagnostic necessity (*17*).

RESULTS

The Pick-Seq method uses immunofluorescence whole-slide imaging of a standard 5-10 µm thick section followed by aspiration of an mROI into a liquid-filled 40 µm bore needle using a robotic arm ("picker") and subsequent deposition into a PCR tube containing lysis buffer (**Fig. 1)** (*18*). Successful deposition can optionally be confirmed by imaging through a flat-bottom tube. To capture and sequence RNA from the mROIs, lysed cells were de-crosslinked, mRNA purified using Oligo(dT) beads, and sequencing libraries then prepared (Methods). Pick-Seq has been implemented in a commercial instrument (*19*) (RareCyte CyteFinder®) that integrates a high-resolution, slide-scanning fluorescence microscope with a robotic picker but could be performed with other microscopes and robotic manipulators.

To test Pick-Seq on a well-characterized tissue, we stained a section of FFPE human tonsil with B- (CD20) and T- (CD3, CD4, CD8) cell markers and, from an adjacent section, collected five mROIs from two B-cell follicles and five mROIs from an inter-follicular region rich in T cells ("T-cell zone") (**Fig. 2, A and B**); subsequent post-pick imaging of the picked section showed that each picked mROI contained ~5-10 cells. The resulting bar-coded RNA-sequencing (RNA-seq) libraries detected, on average, 2.700 genes per mROI, and 13,033 unique genes across all ten mROIs (**Table S1, Fig. 2C**). Differentially expressed genes (DEGs) between the T-cell zone and B-cell follicles included the T- and B-cell lineage markers *CD3D* and *CD19*, respectively, as expected (FDR < 0.05; **Fig. 2D**). We used CIBERSORT to deconvolve the cell composition of the mROIs from their gene expression profiles (*20*) and identified a high abundance of T or B cells expected within each compartment (**Fig. 2E**), but

transcripts from other cell types were also detected (note that relative gene counts and CIBERSORT fractions are not identical because deconvolution uses gene set signatures, not single genes). Fluorescence microscopy confirmed the presence of some B cells in T-cell zones and T cells in B-cell follicles (**Fig. 2F**), but principal component analysis (PCA) was able to resolve the two B-cell follicles from each other as well as from the T-cell zone (**Fig. 2G)**. When considering only the B-cell follicles, DEGs included *JCHAIN*, *MZB1,* and *CD21* (**Fig. 2H**), and subsequent imaging confirmed spatially restricted expression of CD21 (**Fig. 2B**), consistent with the absence of a follicular dendritic cell network in the sectioned plane of one follicle (*21*). We conclude that Pick-Seq can uncover spatially restricted gene expression patterns from FFPE tissue and these patterns can be confirmed by imaging.

Frozen tissues have less RNA degradation than FFPE specimens, and, if available, allow for deeper RNA sequencing than FFPE. We performed Pick-Seq on a frozen section from an ER/PR double-positive breast cancer biopsy that had been stained to identify tumor (cytokeratin) and T (CD3, CD8) cells. Fourteen mROIs were recovered from three tissue microenvironments: cancer cells alone (4 mROIs), T cells outside the tumor (5 mROIs), and regions containing tumor-infiltrating T lymphocytes (TILs; 5 mROIs). In the case of TIL mROIs, imaging showed that each pick contained one (mROIs TIL1-3) or three (mROI TIL4) T cells along with an estimated four to six cancer cells (**Fig. 3A**). The RNA-seq libraries prepared from these picks detected, on average, 4,640 unique genes per mROI (~16,800 genes in all mROIs; **Table S2**, **Fig. 3B**). We used single-sample gene set enrichment analysis (ssGSEA) to identify classes of genes that were over-represented, and we confirmed enrichment of breast cancer-related and T-cell signatures in tumor and T-cell containing mROIs, respectively (**Fig. 3C**). DEGs included estrogen receptor 1 (*ESR1)* and progesterone receptor (*PGR*; **Fig. 3D**), consistent with an ER^+/PR^+ luminal A tumor subtype). *ANXA1*, a gene that promotes Th1 differentiation of T cells, was found only in regions of tumor-infiltrating T cells (TILs; **Fig. 3D**), suggesting tumor specific

immune-suppression (*22*). Based on these data, we conclude that Pick-Seq can identify a single T cell in a background of transcriptionally unrelated cells, although signal improves when an mROI contains three rather than one T cell (compare TIL4 to TIL1, for example; **Fig. 3C**).

Comparing data from FFPE tonsil with frozen breast cancer tissue (**Fig. 2C and 3B**), we found that transcriptome coverage was 50-70% greater in frozen than fixed tissue; we interpret this as arising from the higher RNA quality of frozen samples. Because polyA tail purification was used to generate sequencing libraries, RNA degradation in mROIs from FFPE specimens resulted in read data that was strongly biased toward the 3' ends of genes as revealed by the cumulative distribution of read positions along the length of each gene (**Fig. S1, A and B**). The fraction of aligned reads was also greater for mROIs from frozen than FFPE specimens (an average of ~ 95% for picks from frozen breast cancer samples vs. 50-80% for tonsil and melanoma; **Fig S2A**). Despite this 3' bias, FFPE sequencing appeared to correctly capture cellular composition, as judged by imaging performed in parallel(*23*).

To study transcriptional changes associated with differences in tumor histology that impact disease outcome, we analyzed an archival FFPE *BRAF*-wild type *NF1*-mutant melanoma. This sample contained five distinct histologic regions: (i) early-stage melanoma *in situ* (MIS) (**Fig. 4, A and B**); (ii) tumor, (iii) invasive tumor margin involving tumor growth into the normal dermis; (iv) tumor-adjacent infiltrating lymphocytes (TILs) representing a brisk TIL response (**Fig. 4B**); and (v) exophytic melanoma projecting up towards the surface of the skin (not shown). Seventy-five mROIs samples were collected from this sample, each containing approximately 5-20 cells, with an average of 2,377 genes detected per mROI (**Table S3, Fig. S2B**), consistent with results from the FFPE tonsil sample (**Fig. 2C**). The PCA landscape across all mROIs (25% of variance explained in PC1 and PC2) revealed four clusters with mROIs from exophytic and invasive tumor near each other and distinct from the MIS and brisk TIL regions (**Fig. 5A**). The mROIs from the invasive margin clustered close to, but still largely

137 separate from, the exophytic and invasive tumor clusters. Pairwise analysis among all five histological
138 sites identified 208 to 947 DEGs; data in **Fig. 5B** shows 705 DEGs for tumor vs. MIS regions. Tumor
139 enriched DEGs included *S100B,* a progression marker (*24*), and *CD63*, a negative regulator of the
140 epithelial-mesenchymal transition in melanoma (*25*) (**Fig. 5C**). Imaging of an adjacent tissue section
141 confirmed the melanocyte-restricted expression of S100B and CD63 in tumor regions (**Fig. 5D**),
142 consistent with the annotation of these proteins as progression markers. When Pick-Seq was repeated on
143 adjacent specimens, batch-independent clustering by region was observed, demonstrating the
144 reproducibility of the method (**Fig. 5, E and F**).

145 Enrichment of immune-related signatures in mROIs from the TIL and MIS regions **(Fig. 6)** was
146 consistent with imaging of an adjacent tissue section for the presence of macrophages and T cells **(Fig.**
147 **4B**). The ratio between signatures for MITF (a transcription factor) and AXL (a receptor tyrosine kinase)
148 has been studied extensively in melanoma (*26*, *27*), and in our data, GSEA demonstrated enrichment of
149 MITF programs and downstream targets in the exophytic melanoma region as compared to MIS (**Fig. 7,**
150 **A and B**); in MIS, a MITF-low, AXL-high transcriptional state was observed (**Fig. 7A**). Fluorescence
151 microscopy confirmed higher MITF protein levels in melanocytes within the tumor region relative to
152 MIS (**Fig. 7, C and D)**, but AXL expression in MIS was restricted to the membranes of epithelial
153 keratinocytes (**Fig. 7E)**. Thus, fluorescence microscopy shows that differential AXL gene expression
154 between MIS and tumor is a consequence of the cellular composition of the mROIs rather than a change
155 in melanocyte biology, demonstrating the value of combining RNA expression data with protein
156 expression data.

157 NanoString GeoMx^{TM} DSP is a leading, commercially available method for high-plex, spatial
158 transcriptomic analysis of FFPE tissues (*28*, *29*). We used it to evaluate the performance of Pick-Seq on
159 serial sections containing regions of exophytic melanoma and MIS. GeoMx measures the abundance of

~1,800 transcripts: 1,571 of these were detected by Pick-Seq (**Table S4, Fig. 8)**. Most of the 94 DEGs (exophytic melanoma vs. MIS) common to GeoMx and Pick-Seq were concordantly up- or down-regulated in both assays (**Fig. 9**). Concordance between Pick-Seq and GeoMx strongly suggests that the two methods correctly capture significant differences between tissue regions. However, Pick-Seq sampled regions approximately 25-fold smaller in area than GeoMx (40 μm vs. ~200 μm diameter) while identifying ~2.2-fold more DEGs (FDR < 0.05; **Fig. 9** inset). The 40 μm size of Pick-Seq mROIs compares favorably to the 100 μm raster used for spatial transcriptomics on frozen specimens (*10*, *30*), and the depth of sequencing and integration with imaging retains the advantages of other emerging methods involving "*high-definition spatial transcriptomics*" (*31*) while extending the application to fixed specimens.

DISCUSSION

The data in this paper establishes that micro-mechanical isolation of tissue mROIs using compact robotics integrated with multiplexed fluorescence imaging is a simple and sensitive means of spatially-resolved transcriptional profiling of fixed and frozen tissue. Pick-Seq is sufficiently reliable that in a typical use case, it proved possible to sequence 75 successive micro-regions spanning multiple tumor stages and histomorphologies in a single complex melanoma. Pairwise comparison of RNA from melanoma *in situ*, invasive tumor, and areas of active immunosurveillance identified ~200 to 950 DEGs, which included genes known to be differentially expressed by tumor stage or extent of lymphocyte infiltration as well as newly identified genes. In our work on tumor atlases, tight integration of imaging and transcript profiling is essential, but it is also convenient for the now-common practice of validating cRNA-Seq or micro-region sequencing (mrSeq) results using immunofluorescence imaging (*32*).

Mechanical isolation avoids complex chemistry and RNA-damaging lasers, and can, in principle, be used with many different sequencing and imaging approaches; the implementation described here has

been commercialized, making it readily accessible. A wide variety of methods are being developed for spatial transcriptomics of frozen specimens (*33*), but FFPE tissue presents additional challenges. These challenges merit overcoming because FFPE tissues are more widely available in research and clinical settings (including in histopathology archives), better preserve morphology and have fewer pre-analytical variables; we therefore expect Pick-Seq to have its greatest impact in the analysis of fixed and archival tissue.

In most settings, Pick-Seq is not a true "single-cell" method, but we have shown that it is possible to pick one cell of interest when it is surrounded by other cell types and deconvolve the data; with improvements in sequencing technology (e.g. use of Smart-3SEQ or Smart-seq3)(*14*, *34*) and computation, this may approach single-cell resolution (*20*). In our specimens, Pick-Seq outperformed GeoMx, the leading micro-region sequencing (mrSeq) platform, with respect to spatial resolution as well as the number of DEGs, and both appear to be more sensitive than laser capture microdissection as usually deployed (*35*). Implementing different sequencing workflows, including methods for single-cell sequencing of T-cell receptors (*35*), represents a direct and potentially impactful extension of the method described here. We also anticipate continued improvement in Pick-Seq through the use of smaller needles, faster picking, and integration with whole-slide imaging to enable a straightforward comparison of molecular and morphological features of normal and diseased tissues. More generally, the integration of transcript profiling with highly-multiplexed tissue imaging on microscope platforms promises to substantially advance the goal of linking our understanding of disease genomics with histopathology for research and diagnostic purposes(*5*).

MATERIALS AND METHODS

Tissue procurement

Frozen breast cancer blocks were obtained from a commercial vendor (Origene, Rockville MD) and sectioned at a thickness of 10 microns by the University of Washington Pathology Core. FFPE tonsil sections were obtained from Zyagen (San Diego, CA). Freshly harvested tissue was fixed by the vendor in 10% neutral buffered formalin and processed for paraffin embedding. Paraffin blocks were sectioned at a thickness of 5 microns and mounted on positively charged slides.

FFPE melanoma sections were obtained under IRB oversight as 5 micron sections from the archives of the Dermatopathology Core at Brigham and Women's Hospital. The sections originated from a biopsy removed from a 61 year-old male, non-smoker, with extensive sun exposure. He presented with a pigmented lesion on his forearm with recurrent intermittent bleeding that became continuous after a trauma. The patient was also diagnosed with basal cell carcinoma on the right side of the chest. The clinical pathology reported a malignant melanoma, depth of 14.0 mm, anatomic level IV, with extensive associated melanoma in situ. Superficial spreading, intraepidermal component, vertical growth phase, and ulceration were also noted as present. Non-brisk tumor-infiltrating lymphocytes were observed. There was no report of perineural or vascular invasion. The sample tested negative for BRAF V600E and PD-L1. This study was approved by the Institutional Review Board of the Harvard Faculty of Medicine (FWA00007071, Protocol IRB18-1363). A waiver of the requirement to obtain consent was deemed appropriate.

Instrumentation

The RareCyte® imaging and picking platform performs four or six-color fluorescence imaging and has a microscope stage that employs a kinematic mount to ensure highly reproducible positioning of the slide;

X–Y displacement upon reloading is approximately 2–3 μm. Scanning of each slide in four channels takes about 12 min inclusive of image plane determination. The fluid-coupled picking system positions a needle above the slide stage for retrieval of individual mROIs. The needle tip mechanically dislodges the tissue region with positive pressure into an imaging tube; visual confirmation of deposition is optional. The system is automated and does not require high technical skills; the rate of successful cell retrieval is ~80–90%. Additional feature of the system, originally developed for isolation of circulating tumor cells, have been described previously (*19*)

OCT tissue staining and imaging

Frozen biopsy samples were sectioned 10 μm thick in optimal cutting temperature (OCT) mounting medium, thawed (25˚C, 5 minutes), dehydrated in acetone (-20°C, 10 min), rehydrated in PBS/0.25% Triton X/0.0025% RNasin Plus (Promega) (twice, 0˚C, 3 min each) then washed PBS/0.0025% RNasin Plus (0˚C). Prior to staining slides were blocked with ice-cold blocking buffer 1 (PBS/6% BSA/0.1% RNasin Plus) for 10 min on ice. Sections were stained with primary antibodies (**Table S5)** and SYTOX Orange (nuclear stain) diluted in a blocking buffer and incubated on ice for 30 min. After staining, slides were washed twice for 3 min in ice-cold 1X PBS/0.25% Triton X/0.0025% RNasin Plus, followed by a wash in ice-cold 1X PBS/0.0025% RNasin Plus. Stained sections were mounted with 1X PBS/0.1% RNasin Plus and scanned on a CyteFinder instrument to identify regions of interest. Coverslips were removed, and sections were dehydrated in a series of ice-cold solutions each containing 0.0025% RNasin Plus: 1X PBS for 1 min, 1X PBS for 1 min, 75% ethanol for 1 min, 95% ethanol for 1 min, 100% ethanol for 1 min. Slides were left in ice-cold 100% ethanol prior to mROI retrieval.

FFPE tissue staining and imaging

FFPE sections were deparaffinized and rehydrated using the Histogene Refill Kit (Arcturus). Slides were immersed in xylene for 10 min, a second jar of xylene for 15 min then incubated in an ethanol series (100% ethanol for 4 min, 95% ethanol for 4 min, 75% ethanol for 4 min) followed by 1X PBS for 4 min. For antigen retrieval, slides were incubated in Leica 1X Tris/EDTA pH 9 retrieval solution at 95°C for 10 min, then washed in 1X PBS/0.25% Triton X three times (3 min each wash). To reduce autofluorescence, slides were submerged in 1X PBS with 4.5% hydrogen peroxide and 24 mM NaOH and photobleached between two light sources for 1 hr. Slides were washed twice in 1X PBS/0.25% Triton X (3 min per wash), followed by a 1X PBS wash, and stored overnight at 4°C. Melanoma sections were dewaxed and subjected to antigen retrieval using a Leica Biosystems BOND RX automated slide stainer, following the manufacture's protocols.

Prior to staining, slides were blocked for 1 hour at room temperature with blocking buffer 2 (1X PBS/10% goat serum/6% BSA/0.1% PEG). Primary antibodies diluted in blocking buffer 2 (**Table S5**) were applied to the section and incubated for 1 hr at room temperature. After the primary antibody incubation, slides were washed two times with 1X PBS/0.25% Triton X (3 min per wash), followed by a wash with 1X PBS. Secondary antibodies diluted in blocking buffer 2, were applied to the sections and incubated for 1 hr at room temperature. Slides were then washed twice with 1X PBS/0.25% Triton X (3 min per wash) followed by a wash with 1X PBS. For the second round of staining on the same section, the above primary antibody incubation step was repeated after a blocking step. To visualize the nuclei, slides were stained with a 1:100,000 dilution of SYTOX Orange for 15 min at room temperature followed by two washes in 1X PBS/0.25% Triton X (3 min per wash). Stained sections were mounted with RareCyte mounting media and coverslip applied. Stained slides were scanned on a CyteFinder instrument.

Pick-Seq tissue section preparation

For each FFPE tissue sample, a serial section adjacent to the immunofluorescence (IF) stained section was prepared for microregion retrieval and downstream RNA-seq analysis. Sections were deparaffinized and rehydrated using the Histogene Refill Kit. Slides were immersed in xylene for 5 min, followed by incubation in a second jar of xylene for 5 min. Slides were then incubated in a series of ice-cold solutions with 0.0025% RNasin Plus (Promega): 100% ethanol for 1 min, 95% ethanol for 1 min, 75% ethanol for 1 min, 1X PBS for 1 min, and another tube of 1X PBS for 1 min. Slides were stained with 50 µM DRAQ5,™ a Far-Red DNA Dye (ThermoFisher) in PBS, with 0.1% RNasin Plus for 2 min on ice. Sections were dehydrated in a series of ice-cold solutions with 0.0025% RNasin Plus: 1X PBS for 1 min, 1X PBS for 1 min, 75% ethanol for 1 min, 95% ethanol for 1 min, 100% ethanol for 1 min. Slides were left in ice-cold 100% ethanol prior to microregion retrieval.

Pick-Seq micro-region retrieval

IF stained sections were evaluated to identify regions of interest for transcriptional analysis (**Table S6**). Tissue architecture of the IF-stained and DRAQ5-stained sections were compared to identify the corresponding regions of interest on the DRAQ5-stained slide. Prior to microregion retrieval, slides were removed from 100% ethanol and allowed to air dry. Slides were loaded into a CyteFinder instrument (RareCyte), and microregions were retrieved using the integrated CytePicker module with 40 µm diameter needles. Tissue microregions were deposited with 2 µl PBS into PCR tubes containing 18 µl of lysis buffer: 1:16 mix of Proteinase K solution (QIAGEN) in PKD buffer (QIAGEN), with 0.1% RNasin Plus. After deposit, tubes were immediately placed in dry ice and stored at -80°C until ready for downstream RNA-seq workflow.

294

295 **Tissue microregion lysis and mRNA enrichment**

296 PCR tubes containing tissue microregions in the lysis buffer were removed from the freezer, allowed to
297 thaw at room temperature for 5 min, and incubated at 56°C for 1 hr. Tubes were briefly vortexed, spun
298 down, and placed on ice. Dynabeads Oligo$(dT)_{25}$ beads (ThermoFisher) were washed three times with
299 ice-cold 1X hybridization buffer (NorthernMax buffer (ThermoFisher) with 0.05% Tween 20 and
300 0.0025% RNasin Plus) and resuspended in original bead volume with ice-cold 2x hybridization buffer
301 (NorthernMax buffer with 0.1% Tween 20 and 0.005% RNasin Plus). A volume of 20 µl of washed
302 beads was added to each lysed sample, mixed by pipette, and incubated at 56°C for 1 min followed by
303 room temperature incubation for 10 min. Samples were placed on a magnet and washed twice with an
304 ice-cold 1X hybridization buffer, then once with ice-cold 1X PBS with 0.0025% RNasin Plus. The
305 supernatant was removed, and the pellet was resuspended in 10.5 µl nuclease-free water. Samples were
306 incubated at 80°C for 2 min and immediately placed on a magnet. The supernatant was transferred to
307 new PCR tubes or plates, and placed on ice for subsequent whole transcriptome amplification or stored
308 at -80°C.

309

310 **Whole transcriptome amplification, RNA-seq library preparation, and sequencing**

311 Reverse transcription and cDNA amplification were performed using the SMART-Seq v4 Ultra Low
312 Input RNA Kit for Sequencing (Takara Bio, Kusatsu, Shiga, Japan). The resulting amplified cDNA
313 libraries were assessed for DNA concentration using the Qubit dsDNA HS Assay Kit (ThermoFisher)
314 and for fragment size distribution using the BioAnalyzer 2100 High Sensitivity DNA Kit (Agilent). The
315 fragment size distribution is used to determine how to make the cDNA into a sequencing library.
316 Tissues, where the majority of cDNA was >500 bp in length, was prepared into a library using Nextera

tagmentation method (Illumina) while if the majority of the cDNA was <500 bp in length the cDNA was cleaned up and adapters were ligated following the ThruPLEX method (Takara Bio).

cDNA from tonsil and breast cancer tissue microregion samples were prepared into sequencing libraries using Nextera XT library preparation (Illumina), while cDNA from melanoma microregion libraries were prepared with ThruPLEX DNA-seq Kit (Takara Bio). The resulting libraries were quantitated using the Qubit dsDNA HS Assay Kit and quality was assessed on a BioAnalyzer 2100 High Sensitivity DNA Kit. Libraries were pooled at equimolar ratios and sequenced in-house using an Illumina MiSeq or on an IlluminaNextSeq at Biopolymers Facility at Harvard Medical School.

RNA-seq data processing

The raw FASTQ files were examined for quality issues using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to ensure library generation and sequencing are suitable for further analysis. The reads were processed using the bcbio pipeline v.1.2.1 software (*36*). Briefly, reads were mapped to the GRCh38 human reference genome using HISAT2 (*37*) and Salmon (*38*). Length scaled transcripts per million (TPM) derived from Salmon was passed through the ARSeq pipeline v.2.2.14 (https://github.com/ajitjohnson/arseq) for downstream analysis. The DESeq2 R package (*39*) was used to generate the normalized read count table based on their estimateSizeFactors() function with default parameters by calculating a pseudo-reference sample of the geometric means for each gene across all samples and then using the "median ratio" of each sample to the pseudo-reference as the sizeFactor for that sample. The sizeFactor is then applied to each gene's raw count to get the normalized count for that gene. DESeq2 was also used for differential gene expression analysis. A corrected P value cut-off of 0.05 was used to assess significant genes that were up-regulated or down-regulated using Benjamini-Hochberg (BH) method. The combat function from the R package SVA was

used to account for batch effects when combining two independent Pick-Seq experiments for assessing technical reproducibility.

Pathway enrichment analyses

A compendium of biological and immunological signatures was identified from publicly available databases or published manuscripts for performing enrichment analysis. To perform gene set enrichment analysis, two previously published methods (Gene Set Enrichment Analysis (GSEA) (*40*) and single-sample GSEA (ssGSEA)) were primarily used. The R package clusterProfiler (*41*) was used to perform GSEA and the R package GSVA (*42*) was used to perform ssGSEA which calculates the degree to which the genes in a particular gene set are coordinately up- or down-regulated within a sample. The breast cancer signatures were curated from MsigDB (*43*), and immune cell-related and melanoma-related (MITF pathway and AXL pathway) signatures were curated from published studies (*32*, *44–46*).

CIBERSORT analysis

The normalized counts' table was inverse log2 transformed and uploaded to the CIBERSORT web app v1.06 (https://cibersort.stanford.edu/) and run with default parameters and quantile normalization disabled. The LM22 signature was used for inferring cellular proportions. For visualization purposes, different cell subtypes are combined: mast cells, eosinophils, and neutrophils are represented as granulocytes; and monocytes and macrophages are represented as mono/macrophages.

GeoMx analysis

NanoString GeoMx gene expression analysis using the cancer transcriptome array probe set was performed by the Technology Access Program at NanoString using methods, as previously described

(*29*). Briefly, a 5 μm section of FFPE melanoma was dewaxed and stained overnight for DNA, melanocytes (PMEL), epithelia (pan-cytokeratin), and immune cells (CD45) to define areas of interest on the NanoString GeoMx instrument for transcriptional analysis using the human cancer transcriptome array probe set. Twenty-nine ROIs representing five morphological sites (melanoma in situ, invasive tumor, invasive tumor margin, brisk tumor-infiltrating lymphocytes, and exophytic melanoma) were selected. All sample processing and sequencing were performed by the Technology Access Program at NanoString. Tissue images, probe measurements (**Table S4**), and quality control data were provided by NanoString, then analyzed using DESeq2 for differential expression analysis.

Comparison of Pick-Seq and GeoMx

Similar, but not identical, regions were assayed using GeoMx to compare with Pick-Seq (differences in the areas sampled precluded a one-to-one comparison). Differential gene expression analysis using DESeq2 compared regions of exophytic melanoma and melanoma in situ. Differentially expressed genes with an FDR < 0.05 were compared between Pick-Seq and GeoMx. As the dynamic range of expression scales between the two methods was largely different, we sought to compare them by directional concordance (i.e. up/downregulation) rather than absolute correlation in fold change. Of the 94 common differentially expressed genes, seventeen genes were not directionally concordant. However, only two of those seventeen showed a log-fold change of >1, suggesting that the 15 others might represent low expression noise. Together, this analysis suggested that differential gene expression analysis by Pick-Seq and GeoMx was largely concordant.

SUPPLEMENTARY MATERIALS

Fig. S1: Positions of sequence reads

Fig. S2: Reads mapped and genes detected for mROIs

Table S1: Normalized gene expression data for tonsil

Table S2: Normalized gene expression data for breast

Table S3: Normalized gene expression data for melanoma

Table S4: Normalized gene expression data for NanoString GeoMx

Table S5: Antibodies for FFPE tissue staining and imaging

Table S6: Regions of interest for transcriptional analysis

REFERENCES

1. G. D. Hammer, S. J. McPhee, *Pathophysiology of disease: An Introduction to Clinical Medicine 7/E* (McGraw-Hill Education, 2014).

2. D. A. Jaitin, E. Kenigsberg, H. Keren-Shaul, N. Elefant, F. Paul, I. Zaretsky, A. Mildner, N. Cohen, S. Jung, A. Tanay, I. Amit, Massively parallel single cell RNA-Seq for marker-free decomposition of tissues into cell types. *Science*. **343**, 776–779 (2014).

3. D. Grün, A. Lyubimova, L. Kester, K. Wiebrands, O. Basak, N. Sasaki, H. Clevers, A. van Oudenaarden, Single-cell messenger RNA sequencing reveals rare intestinal cell types. *Nature*. **525**, 251–255 (2015).

4. HuBMAP Consortium, The human body at cellular resolution: the NIH Human Biomolecular Atlas Program. *Nature*. **574**, 187–192 (2019).

5. O. Rozenblatt-Rosen, A. Regev, P. Oberdoerffer, T. Nawy, A. Hupalowska, J. E. Rood, O. Ashenberg, E. Cerami, R. J. Coffey, E. Demir, L. Ding, E. D. Esplin, J. M. Ford, J. Goecks, S. Ghosh, J. W. Gray, J. Guinney, S. E. Hanlon, S. K. Hughes, E. S. Hwang, C. A. Iacobuzio-Donahue, J. Jané-Valbuena, B. E. Johnson, K. S. Lau, T. Lively, S. A. Mazzilli, D. Pe'er, S. Santagata, A. K. Shalek, D. Schapiro, M. P. Snyder, P. K. Sorger, A. E. Spira, S. Srivastava, K. Tan, R. B. West, E. H. Williams, Human Tumor Atlas Network, The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell*. **181**, 236–249 (2020).

6. N. Rajewsky, G. Almouzni, S. A. Gorski, S. Aerts, I. Amit, M. G. Bertero, C. Bock, A. L. Bredenoord, G. Cavalli, S. Chiocca, H. Clevers, B. De Strooper, A. Eggert, J. Ellenberg, X. M. Fernández, M. Figlerowicz, S. M. Gasser, N. Hubner, J. Kjems, J. A. Knoblich, G. Krabbe, P. Lichter, S. Linnarsson, J.-C. Marine, J. C. Marioni, M. A. Marti-Renom, M. G. Netea, D. Nickel, M. Nollmann, H. R. Novak, H. Parkinson, S. Piccolo, I. Pinheiro, A. Pombo, C. Popp, W. Reik, S. Roman-Roman, P. Rosenstiel, J. L. Schultze, O. Stegle, A. Tanay, G. Testa, D. Thanos, F. J. Theis, M.-E. Torres-Padilla, A. Valencia, C. Vallot, A. van Oudenaarden, M. Vidal, T. Voet, LifeTime and improving European healthcare through cell-based interceptive medicine. *Nature*. **587**, 377–386 (2020).

7. J.-R. Lin, B. Izar, S. Wang, C. Yapp, S. Mei, P. M. Shah, S. Santagata, P. K. Sorger, Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-CyCIF and conventional optical microscopes. *eLife Sciences*. **7**, e31657 (2018).

8. M. Angelo, S. C. Bendall, R. Finck, M. B. Hale, C. Hitzman, A. D. Borowsky, R. M. Levenson, J. B. Lowe, S. D. Liu, S. Zhao, Y. Natkunam, G. P. Nolan, Multiplexed ion beam imaging of human breast tumors. *Nat. Med.* **20**, 436–442 (2014).

9. C. Giesen, H. A. O. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schüffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Günther, B. Bodenmiller, Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. *Nat. Methods*. **11**, 417–422 (2014).

10. P. L. Ståhl, F. Salmén, S. Vickovic, A. Lundmark, J. F. Navarro, J. Magnusson, S. Giacomello, M. Asp, J. O. Westholm, M. Huss, A. Mollbrink, S. Linnarsson, S. Codeluppi, Å. Borg, F. Pontén, P. I. Costea, P. Sahlén, J. Mulder, O. Bergmann, J. Lundeberg, J. Frisén, Visualization and analysis of gene expression in tissue sections by spatial transcriptomics. *Science*. **353**, 78–82 (2016).

11. J. J. Going, R. F. Lamb, Practical histological microdissection for PCR analysis. *J Pathol*. **179**, 121–124 (1996).

12. M. R. Emmert-Buck, R. F. Bonner, P. D. Smith, R. F. Chuaqui, Z. Zhuang, S. R. Goldstein, R. A. Weiss, L. A. Liotta, Laser capture microdissection. *Science*. **274**, 998–1001 (1996).

13. D. J. Burgess, Spatial transcriptomics coming of age. *Nature Reviews Genetics*. **20**, 317–317 (2019).

14. J. W. Foley, C. Zhu, P. Jolivet, S. X. Zhu, P. Lu, M. J. Meaney, R. B. West, *Genome Res.*, in press, doi:10.1101/gr.234807.118.

15. A. N. Snow, A. A. Stence, J. A. Pruessner, A. D. Bossler, D. Ma, A simple and cost-effective method of DNA extraction from small formalin-fixed paraffin-embedded tissue for molecular oncologic testing. *BMC Clin Pathol*. **14**, 30 (2014).

16. F. Aeffner, M. D. Zarella, N. Buchbinder, M. M. Bui, M. R. Goodman, D. J. Hartman, G. M. Lujan, M. A. Molani, A. V. Parwani, K. Lillard, O. C. Turner, V. N. P. Vemuri, A. G. Yuil-Valdes, D. Bowman, Introduction to Digital Image Analysis in Whole-slide Imaging: A White Paper from the Digital Pathology Association. *J Pathol Inform*. **10**, 9 (2019).

17. C. for D. and R. Health, Technical Performance Assessment of Digital Pathology Whole Slide Imaging Devices. *U.S. Food and Drug Administration* (2019), (available at http://www.fda.gov/regulatory-information/search-fda-guidance-documents/technical-performance-assessment-digital-pathology-whole-slide-imaging-devices).

18. D. E. Campton, A. B. Ramirez, J. J. Nordberg, N. Drovetto, A. C. Clein, P. Varshavskaya, B. H. Friemel, S. Quarre, A. Breman, M. Dorschner, S. Blau, C. A. Blau, D. E. Sabath, J. L. Stilwell, E. P. Kaldjian, High-recovery visual identification and single-cell retrieval of circulating tumor cells for genomic analysis using a dual-technology platform integrated with automated immunofluorescence staining. *BMC Cancer*. **15**, 360 (2015).

19. E. P. Kaldjian, A. B. Ramirez, Y. Sun, D. E. Campton, J. L. Werbin, P. Varshavskaya, S. Quarre, T. George, A. Madan, C. A. Blau, R. Seubert, The RareCyte® platform for next-generation analysis of circulating tumor cells. *Cytometry A*. **93**, 1220–1225 (2018).

20. B. Chen, M. S. Khodadoust, C. L. Liu, A. M. Newman, A. A. Alizadeh, Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. **1711**, 243–259 (2018).

21. E. M. Wohlford, P. C. Baresel, J. R. Wilmore, A. J. Mortelliti, C. B. Coleman, R. Rochford, Changes in Tonsil B Cell Phenotypes and EBV Receptor Expression in Children Under 5-Years-Old. *Cytometry B Clin Cytom*. **94**, 291–301 (2018).

22. F. D'Acquisto, A. Merghani, E. Lecona, G. Rosignoli, K. Raza, C. D. Buckley, R. J. Flower, M. Perretti, Annexin-1 modulates T-cell activation and differentiation. *Blood*. **109**, 1095–1102 (2007).

23. N. Bossel Ben-Moshe, S. Gilad, G. Perry, S. Benjamin, N. Balint-Lahat, A. Pavlovsky, S. Halperin, B. Markus, A. Yosepovich, I. Barshack, E. N. Gal-Yam, E. Domany, B. Kaufman, M. Dadiani, mRNA-seq whole transcriptome profiling of fresh frozen versus archived fixed tissues. *BMC Genomics*. **19** (2018), doi:10.1186/s12864-018-4761-3.

24. A. R. Bresnick, D. J. Weber, D. B. Zimmer, S100 proteins in cancer. *Nat Rev Cancer*. **15**, 96–109 (2015).

25. A. Lupia, S. Peppicelli, E. Witort, F. Bianchini, V. Carloni, N. Pimpinelli, C. Urso, L. Borgognoni, S. Capaccioli, L. Calorini, M. Lulli, CD63 Tetraspanin Is a Negative Driver of Epithelial-to-Mesenchymal Transition in Human Melanoma Cells. *Journal of Investigative Dermatology*. **134**, 2947–2956 (2014).

26. J. Müller, O. Krijgsman, J. Tsoi, L. Robert, W. Hugo, C. Song, X. Kong, P. A. Possik, P. D. M. Cornelissen-Steijger, M. H. G. Foppen, K. Kemper, C. R. Goding, U. McDermott, C. Blank, J. Haanen, T. G. Graeber, A. Ribas, R. S. Lo, D. S. Peeper, Low MITF/AXL ratio predicts early resistance to multiple targeted drugs in melanoma. *Nature Communications*. **5**, 5712 (2014).

27. C. Levy, M. Khaled, D. E. Fisher, MITF: master regulator of melanocyte development and melanoma oncogene. *Trends in Molecular Medicine*. **12**, 406–414 (2006).

28. B. A. Helmink, S. M. Reddy, J. Gao, S. Zhang, R. Basar, R. Thakur, K. Yizhak, M. Sade-Feldman, J. Blando, G. Han, V. Gopalakrishnan, Y. Xi, H. Zhao, R. N. Amaria, H. A. Tawbi, A. P. Cogdill, W. Liu, V. S. LeBleu, F. G. Kugeratski, S. Patel, M. A. Davies, P. Hwu, J. E. Lee, J. E. Gershenwald, A. Lucci, R. Arora, S. Woodman, E. Z. Keung, P.-O. Gaudreau, A. Reuben, C. N. Spencer, E. M. Burton, L. E. Haydu, A. J. Lazar, R. Zapassodi, C. W. Hudgens, D. A. Ledesma, S. Ong, M. Bailey, S. Warren, D. Rao, O. Krijgsman, E. A. Rozeman, D. Peeper, C. U. Blank, T. N. Schumacher, L. H. Butterfield, M. A. Zelazowska, K. M. McBride, R. Kalluri, J. Allison, F. Petitprez, W. H. Fridman, C. Sautès-Fridman, N. Hacohen, K. Rezvani, P. Sharma, M. T. Tetzlaff, L. Wang, J. A. Wargo, B cells and tertiary lymphoid structures promote immunotherapy response. *Nature*. **577**, 549–555 (2020).

29. D. R. Zollinger, S. E. Lingle, K. Sorg, J. M. Beechem, C. R. Merritt, GeoMx^{TM} RNA Assay: High Multiplex, Digital, Spatial Analysis of RNA in FFPE Tissue. *Methods Mol. Biol.* **2148**, 331–345 (2020).

30. R. Moncada, D. Barkley, F. Wagner, M. Chiodin, J. C. Devlin, M. Baron, C. H. Hajdu, D. M. Simeone, I. Yanai, Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. *Nat. Biotechnol.* **38**, 333–342 (2020).

31. S. Vickovic, G. Eraslan, F. Salmén, J. Klughammer, L. Stenbeck, D. Schapiro, T. Äijö, R. Bonneau, L. Bergenstråhle, J. F. Navarro, J. Gould, G. K. Griffin, Å. Borg, M. Ronaghi, J. Frisén, J. Lundeberg, A. Regev, P. L. Ståhl, High-definition spatial transcriptomics for in situ tissue profiling. *Nature Methods*. **16**, 987–990 (2019).

32. I. Tirosh, B. Izar, S. M. Prakadan, M. H. Wadsworth 2nd, D. Treacy, J. J. Trombetta, A. Rotem, C. Rodman, C. Lian, G. Murphy, M. Fallahi-Sichani, K. Dutton-Regester, J. R. Lin, O. Cohen, P. Shah, D. Lu, A. S. Genshaft, T. K. Hughes, C. G. Ziegler, S. W. Kazer, A. Gaillard, K. E. Kolb, A. C. Villani, C. M. Johannessen, A. Y. Andreev, E. M. Van Allen, M. Bertagnolli, P. K. Sorger, R. J. Sullivan, K. T. Flaherty, D. T. Frederick, J. Jane-Valbuena, C. H. Yoon, O. Rozenblatt-Rosen, A. K. Shalek, A. Regev, L. A. Garraway, Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. **352**, 189–196 (2016).

33. J. R. Moffitt, D. Bambah-Mukku, S. W. Eichhorn, E. Vaughn, K. Shekhar, J. D. Perez, N. D. Rubinstein, J. Hao, A. Regev, C. Dulac, X. Zhuang, Molecular, Spatial and Functional Single-Cell Profiling of the Hypothalamic Preoptic Region. *Science*. **362** (2018), doi:10.1126/science.aau5324.

34. R. Burgemeister, Laser capture microdissection of FFPE tissue sections bridging the gap between microscopy and molecular analysis. *Methods Mol. Biol.* **724**, 105–115 (2011).

35. M. De Simone, G. Rossetti, M. Pagani, Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges. *Front. Immunol.* **9** (2018), doi:10.3389/fimmu.2018.01638.

36. B. Chapman, R. Kirchner, L. Pantano, S. Naumenko, M. D. Smet, L. Beltrame, T. Khotiainsteva, I. Sytchev, R. V. Guimera, J. Kern, C. Brueffer, V. Savelyev, G. Carrasco, M. Giovacchini, M. Ahdesmaki, P. Tang, Sehrish Kanwal, J. J. Porter, S. Möller, Vang Le, A. Coman, Bogdang989, V. Svensson, B. Pedersen, J. Hammerbacher, M. Edwards, M. Mistry, Apastore, P. Cock, S. Turner, *bcbio/bcbio-nextgen: v1.2.6* (Zenodo, 2021; https://zenodo.org/record/3564938).

37. D. Kim, J. M. Paggi, C. Park, C. Bennett, S. L. Salzberg, Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol*. **37**, 907–915 (2019).

38. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. **14**, 417–419 (2017).

39. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. **15**, 550 (2014).

40. A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, J. P. Mesirov, Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. **102**, 15545–50 (2005).

41. G. Yu, L.-G. Wang, Y. Han, Q.-Y. He, clusterProfiler: an R Package for Comparing Biological Themes Among Gene Clusters. *OMICS: A Journal of Integrative Biology*. **16**, 284–287 (2012).

42. S. Hänzelmann, R. Castelo, J. Guinney, GSVA: gene set variation analysis for microarray and RNA-Seq data. *BMC Bioinformatics*. **14**, 7 (2013).

43. A. Liberzon, A. Subramanian, R. Pinchback, H. Thorvaldsdóttir, P. Tamayo, J. P. Mesirov, Molecular signatures database (MSigDB) 3.0. *Bioinformatics*. **27**, 1739–1740 (2011).

44. A. J. Nirmal, T. Regan, B. B. Shih, D. A. Hume, A. H. Sims, T. C. Freeman, Immune Cell Gene Signatures for Profiling the Microenvironment of Solid Tumors. *Cancer Immunol Res*. **6**, 1388–1400 (2018).

45. A. M. Newman, C. L. Liu, M. R. Green, A. J. Gentles, W. Feng, Y. Xu, C. D. Hoang, M. Diehn, A. A. Alizadeh, Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. **12**, 453–457 (2015).

46. B. B. Shih, A. J. Nirmal, D. J. Headon, A. N. Akbar, N. A. Mabbott, T. C. Freeman, Derivation of marker gene signatures from human skin and their use in the interpretation of the transcriptional changes associated with dermatological disorders. *J. Pathol.* **241**, 600–613 (2017).

ACKNOWLEDGEMENTS

We thank Jerry Lin for scientific advice and technical assistance and NanoString Inc. for the acquisition of GeoMx DSP data via its Technology Access Program.

Funding:

National Institutes of Health, NCI grant U54-CA225088 (PKS)

National Institutes of Health, NCI grant R41-CA224503 (PKS, EPK)

National Institutes of Health, NCI grant R35-CA231958 (DMW)

Ludwig Cancer Research Foundation (PKS, SS)

Author contributions:

Study design: ZM, YAC, AJN

Quantitative data analysis: AJN, YAC

Data collection: ZM, NGE, LU, RP, JC, YAC, SAB

Pathology analysis and image review: CGL, GFM, SS

Supervision: DMW, EPK, SS, PKS

Writing: ZM, AJN, AAC, PKS

Manuscript was read and approved by all authors.

Competing interests: NE, LU, RP, JC, and EJK are or have been employees of RareCyte Inc. PKS is a member of the SAB or BOD of Applied BioMath, RareCyte Inc., and Glencoe Software, which distributes a commercial version of the OMERO database; PKS is also a member of the NanoString SAB. SS is a consultant for RareCyte Inc. All other authors declare they have no competing interests.

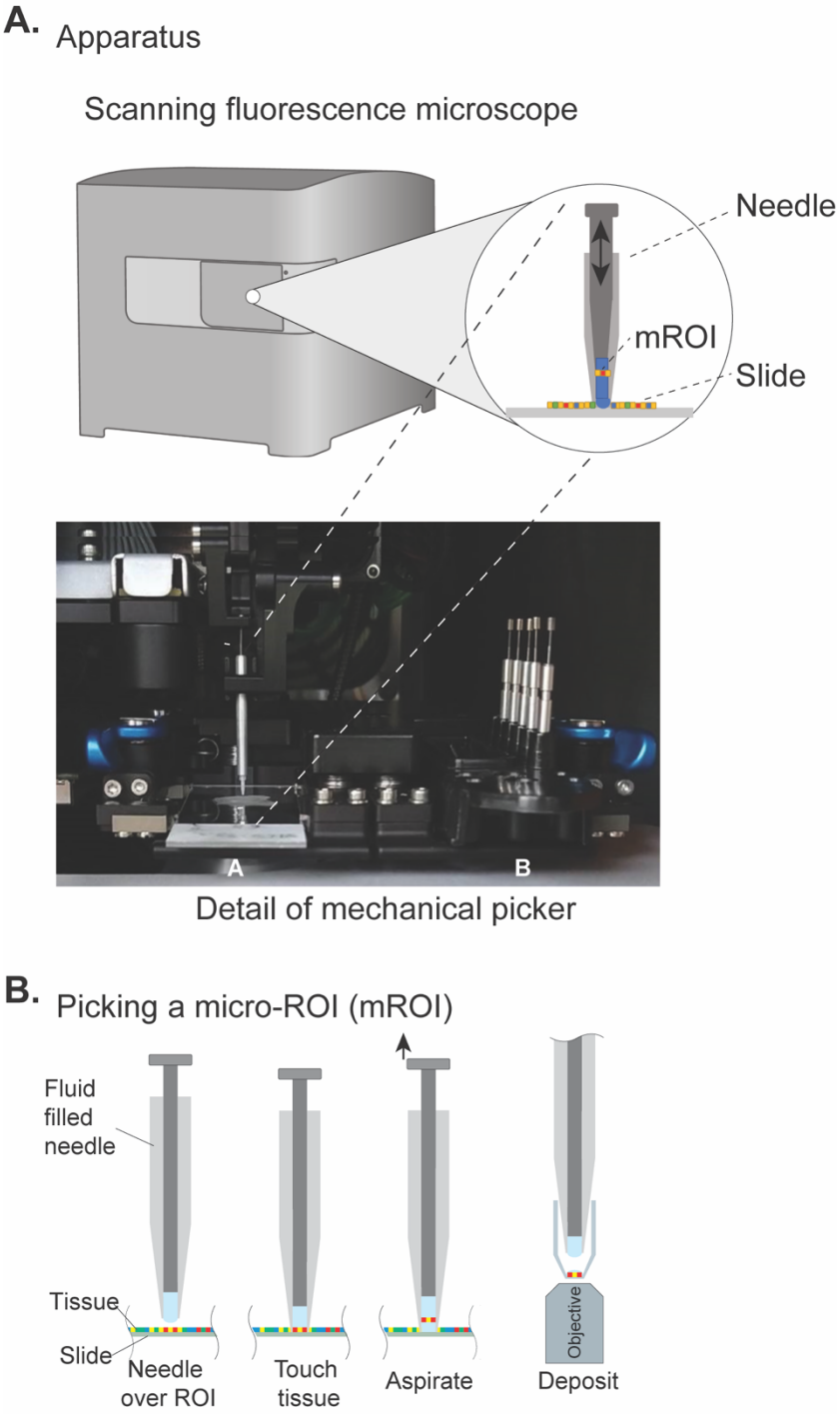
573

574 **Data and materials availability:** Raw sequence data and processed counts table can be accessed from

575 NCBI's GEO repository under accession GSE158564. All code has been deposited on GitHub at

576 https://github.com/sorgerlab.

577 **FIGURES**



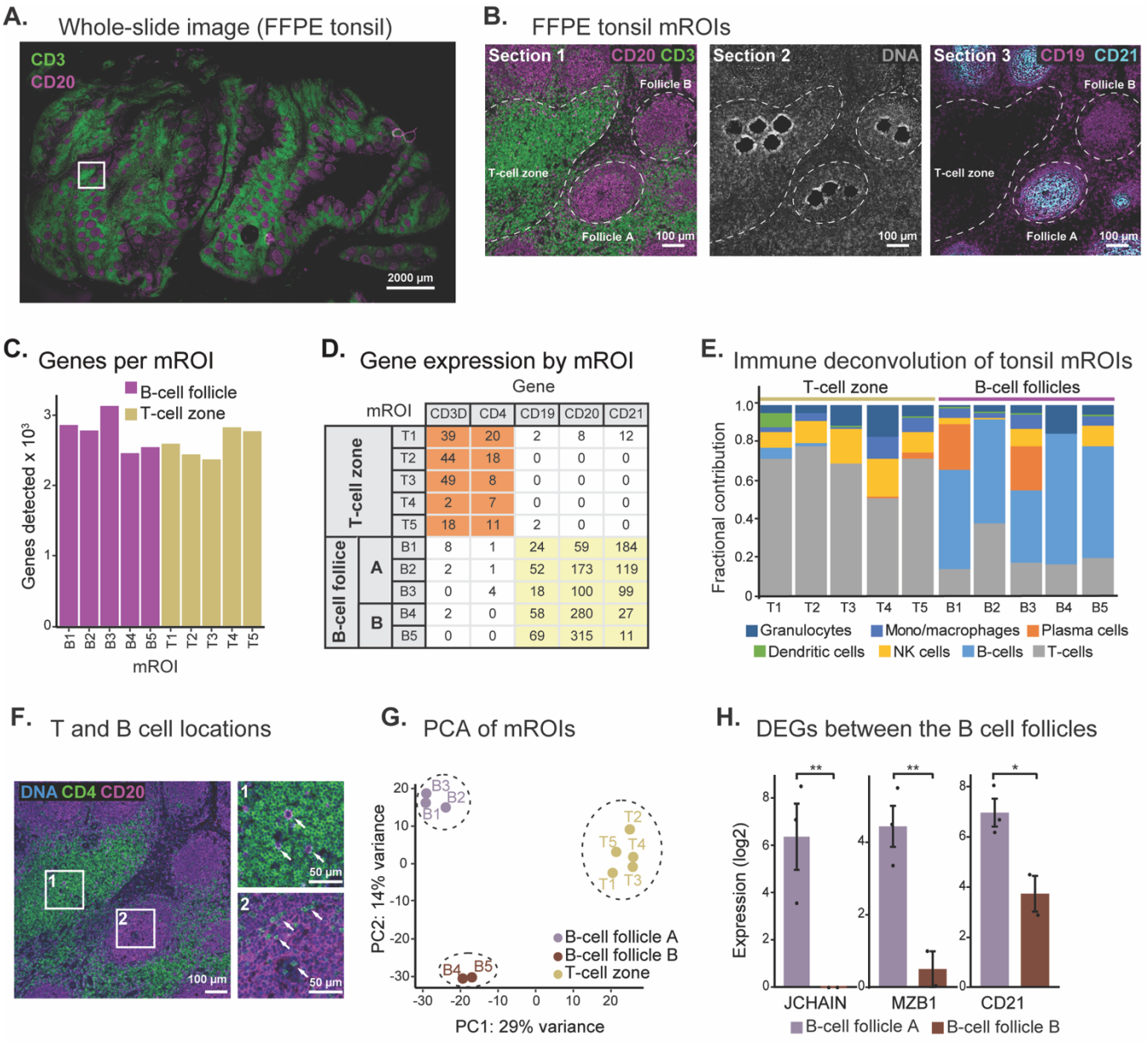
578

579 **Fig. 1: Pick-Seq technology and retrieval of micro-regions from tissue sections**

580 **(A)** External drawing of the RareCyte CyteFinder fluorescence slide scanning microscope (left), a

581 schematic of the picker needle (middle), and a photograph of a needle mounted on a robotic arm above a

582 target slide in stage position A and spare needles and target tubes in position B (right). **(B)** Schematic of
583 mROI retrieval. A fluid-filled needle attached to a robotic arm recovers tissue from a mROI. Sample
584 recovery can be confirmed by imaging the PCR tube if desired.



585

586 **Fig. 2: Pick-Seq analysis of an FFPE tonsil tissue section**

587 **(A)** Whole slide fluorescence microscopy of a 5 µm thick FFPE tonsil section stained for T (CD3:

588 green) and B (CD20: violet) cells. Scale bar, 2000 µm. **(B)** Inset region from Fig. 2A (left);

589 corresponding region of the adjacent 5 µm thick FFPE tissue section after mROI recovery, stained to

590 visualize nuclear DNA (DRAQ5: gray) (center); the next adjacent section was stained for a lineage

591 marker of B cells (CD19: violet) and tonsil follicular DEG (CD21: cyan) (right). Histologic features in

592 tonsil that guide picking are outlined (dotted lines). Scale bars, 100 µm. **(C)** Number of genes detected
593 in each mROI extracted from FFPE tonsil. **(D)** Expression (transcripts per million; TPM) of selected cell
594 lineage genes in each mROI. **(E)** Immune cell deconvolution for each tonsil mROI transcriptome using
595 CIBERSORT LM22 signature. **(F)** Left, fluorescence microscopy of FFPE tonsil stained for DNA
596 (blue), helper T cells (CD4: green), and B cells (CD20: purple). Scale bar, 100 µm. Right, insets 1 and 2.
597 Arrows indicate scattered B-cells in T-cell zone (top) and T cells in B-cell follicle (bottom). Scale bars,
598 50 µm. **(G)** Principal component analysis (PCA) of tonsil mROI transcriptome data colored by
599 histologic feature: B-cell follicle A (violet), B-cell follicle B (brown), and T-cell zone (yellow). **(H)**
600 Expression of selected DEGs (JCHAIN, MZB1, and CD21) in mROIs from B-cell follicles A (n=3;
601 violet) and B (n=2; brown). Data is mean ± SEM. *P<0.01; **P<0.05.
602

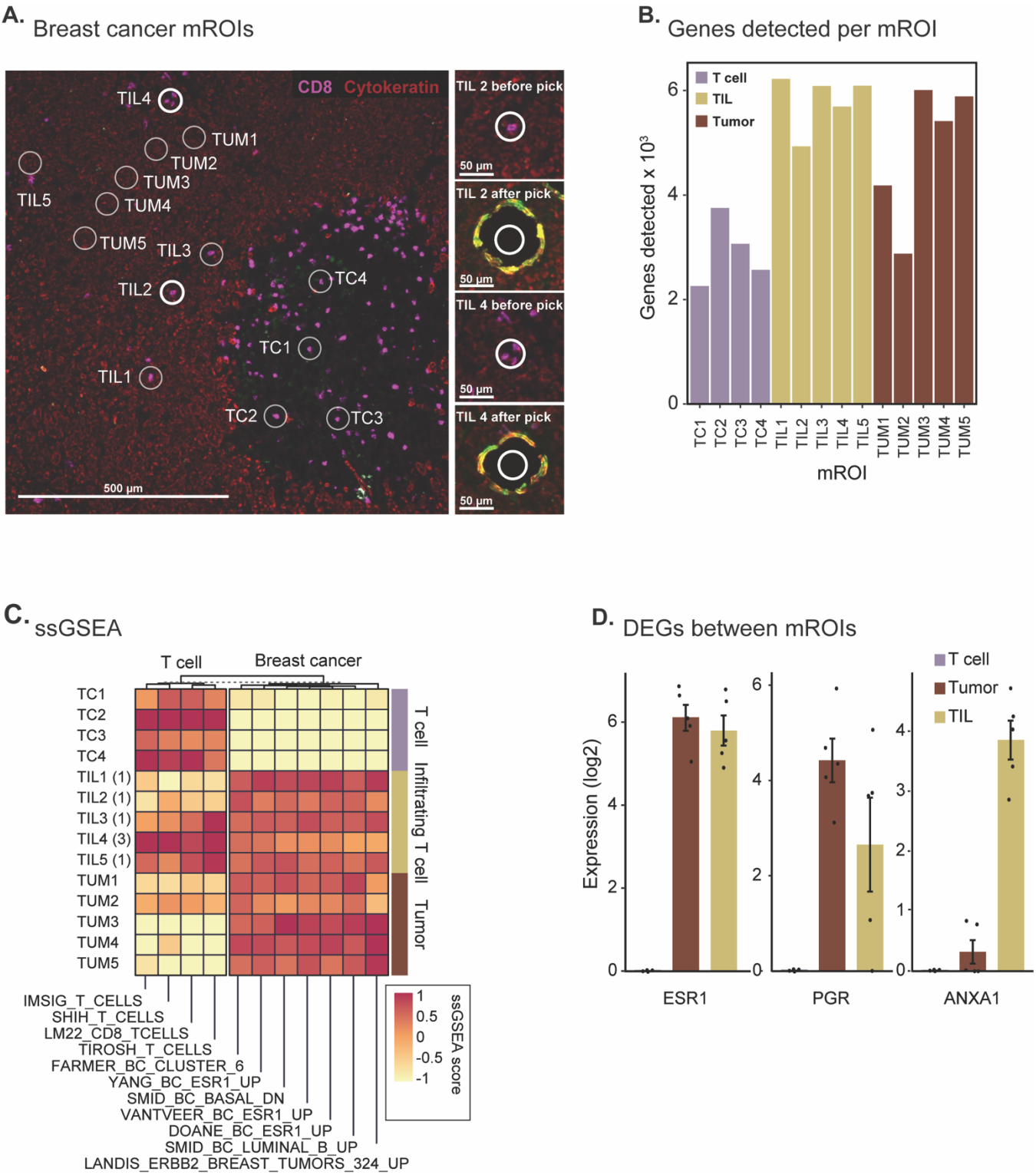


Fig. 3: Pick-Seq of a frozen breast cancer surgical biopsy specimen

(A) Fluorescence microscopy of a 10 µm thick frozen section of breast cancer surgical biopsy (left) stained for tumor (cytokeratin: red) and cytotoxic T cells (CD8: magenta). Numbered mROIs are indicated as tumor (TUM), T cell (TC), and TIL. Scale bar, 500 µm. Magnified mROI containing regions before and after sample recovery of TIL2 and TIL4 samples (right). Yellow ring is due to localized tissue compaction from picking and reflects the size of the picking needle. Scale bars, 50 µm. **(B)** Number of genes detected in each mROI recovered from frozen breast cancer section. Samples are coded by tissue microenvironments: T cell (TC, purple), TIL (yellow), and tumor (TUM, brown). (**C**) Single-sample gene sample enrichment analysis (ssGSEA) of mROI-derived sequence data for breast cancer (BC) and T cell related gene signatures. Number of T cells in each TIL mROI is indicated (parentheses). ssGSEA scores highlight enrichment of breast cancer-related gene signatures in tumor mROIs and T cell related signatures in the T cell rich mROIs. **(D)** Expression of selected cancer-related genes in T-cell (n=4; purple), tumor (n=5, brown), and TIL (n=5; yellow) mROIs recovered from breast tumor sample. Data is mean ± SEM.

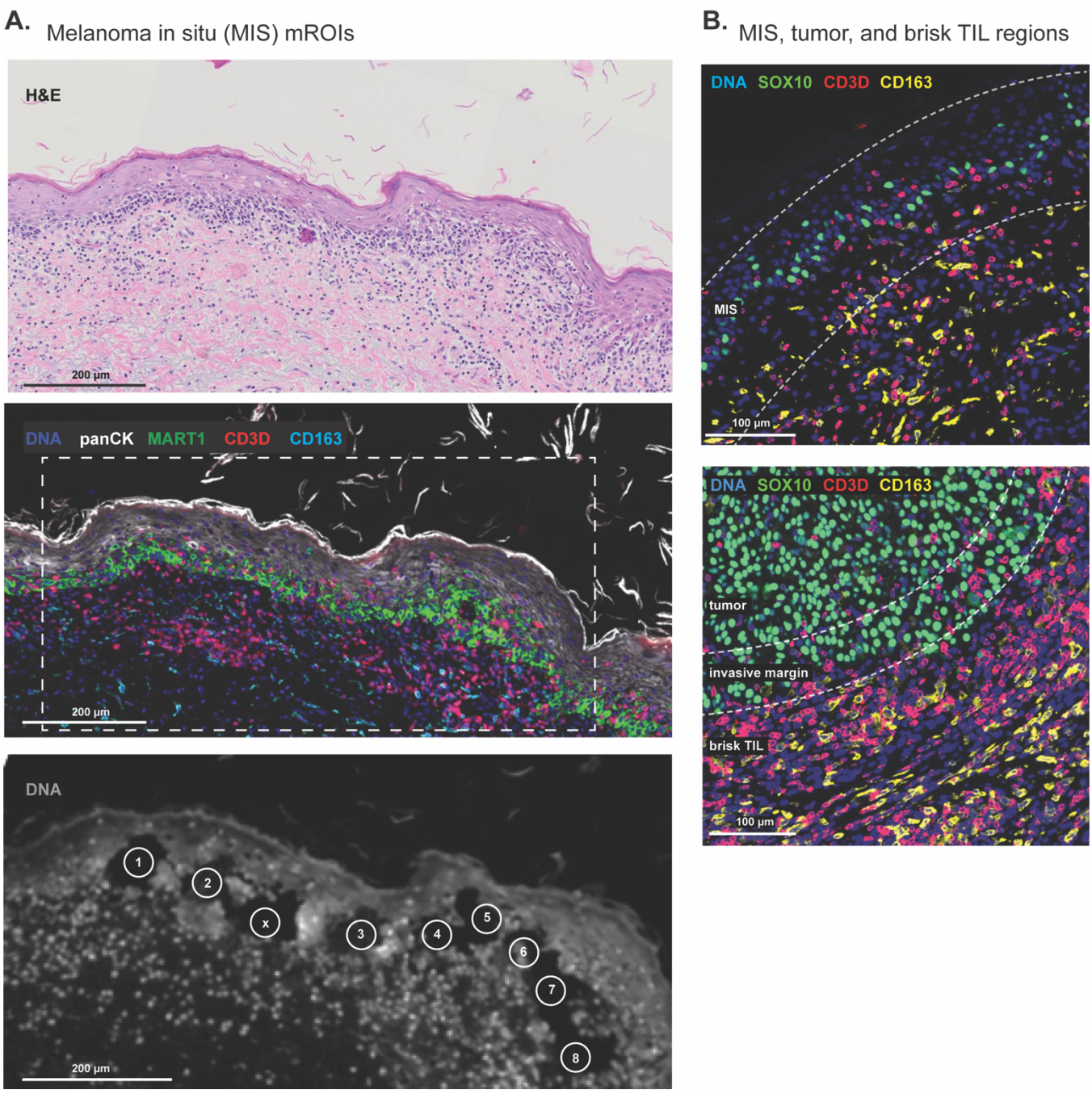
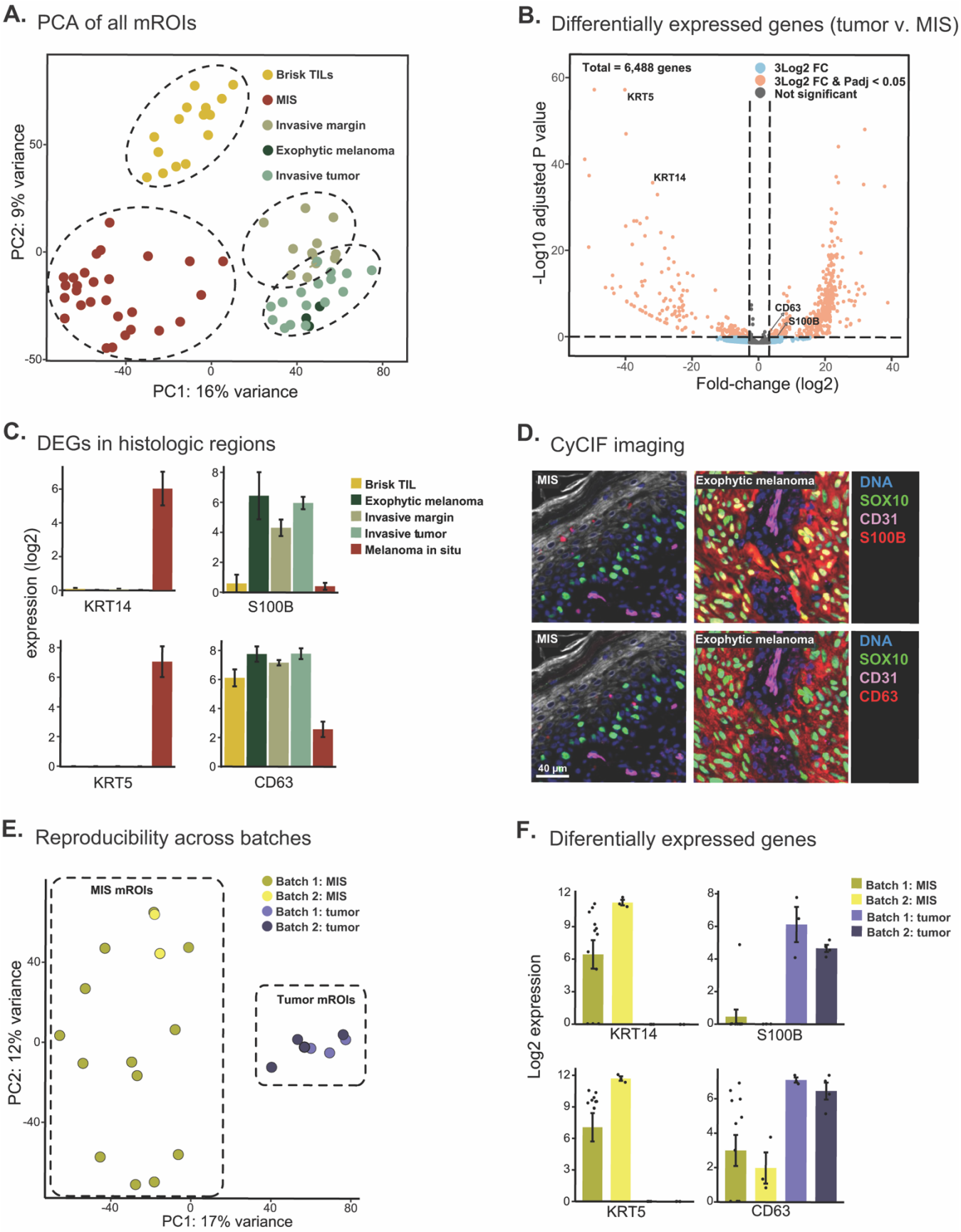


Fig. 4: A complex archival melanoma FFPE specimen

(A) H&E (top panel) and fluorescence microscopy (middle panel) images of a melanoma in situ (MIS) region stained for DNA (blue), keratinocytes (cytokeratin: white), melanocytes (MART1: green), T cells (CD3D: red), and macrophages (CD163: cyan). Magnified corresponding region (bottom panel) in an adjacent tissue section stained for DNA (DRAQ5: gray). Sites of MIS mROI extraction are indicated.

625 Scale bars, 200 µm. **(B)** Fluorescence microscopy images of histologic features in melanoma tissue
626 stained for DNA (blue), melanocytes (SOX10: green), T cells (CD3D: red), and macrophages (CD163:
627 yellow). Top, a region of melanoma in situ. Bottom, a region of invasive melanoma containing tumor
628 center, invasive margin, and adjacent brisk TIL region. Scale bars, 100 µm.



629

Fig. 5: Pick-Seq of archival FFPE melanoma specimen examining transcriptional differences associated with distinct tumor domains

(A) Principal component analysis (PCA) of melanoma mROI transcriptomes. Colors indicate regional histopathology: brisk TIL (BTIL: yellow), MIS (red), invasive margin (IM: olive green), exophytic melanoma (EM: dark green), and center of invasive melanoma tumor (IT: light green). EM and IT are considered tumor samples in this analysis. **(B)** Fold-difference and significance for expression of 6,488 genes between tumor (n=19) and MIS (n=28) samples. DEGs above (blue) and below (orange) a significance threshold (P-adjusted = 0.05) are indicated. **(C)** Expression of selected genes (KRT14, KRT5, S100B, and CD63) in each histologic region of melanoma (same mROIs as **(A)**). Mean ± SEM for mROIs of same histology. **(D)** Fluorescence microscopy image of MIS (left) and exophytic melanoma (right) stained for DNA (blue), melanocytes (SOX10: green), blood vessels (CD31: violet), and S100B (red, top panel) or CD63 (red, bottom panel). Scale bar, 40 μm. **(E)** Principal component analysis (PCA) of mROIs retrieved from exophytic melanoma tumor (purple) or MIS regions (yellow) from different tissue sections of the same patient in separate experiments. Samples are coded by batch and histologic feature. **(F)** Expression of selected genes (KRT14, KRT5, S100B, and CD63) in exophytic melanoma tumor (purple) or MIS (yellow) mROIs. Mean ± SEM for mROIs of same histology.

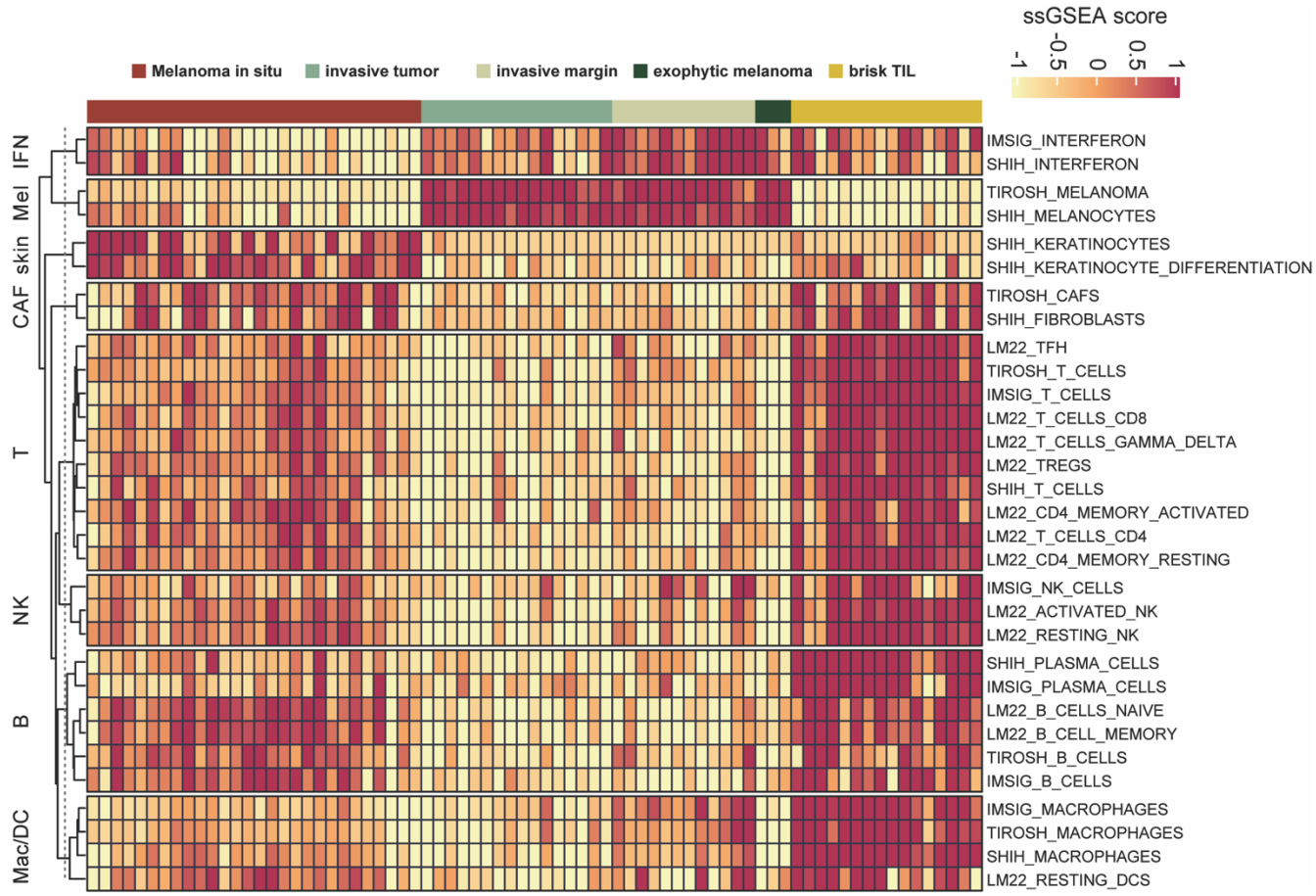
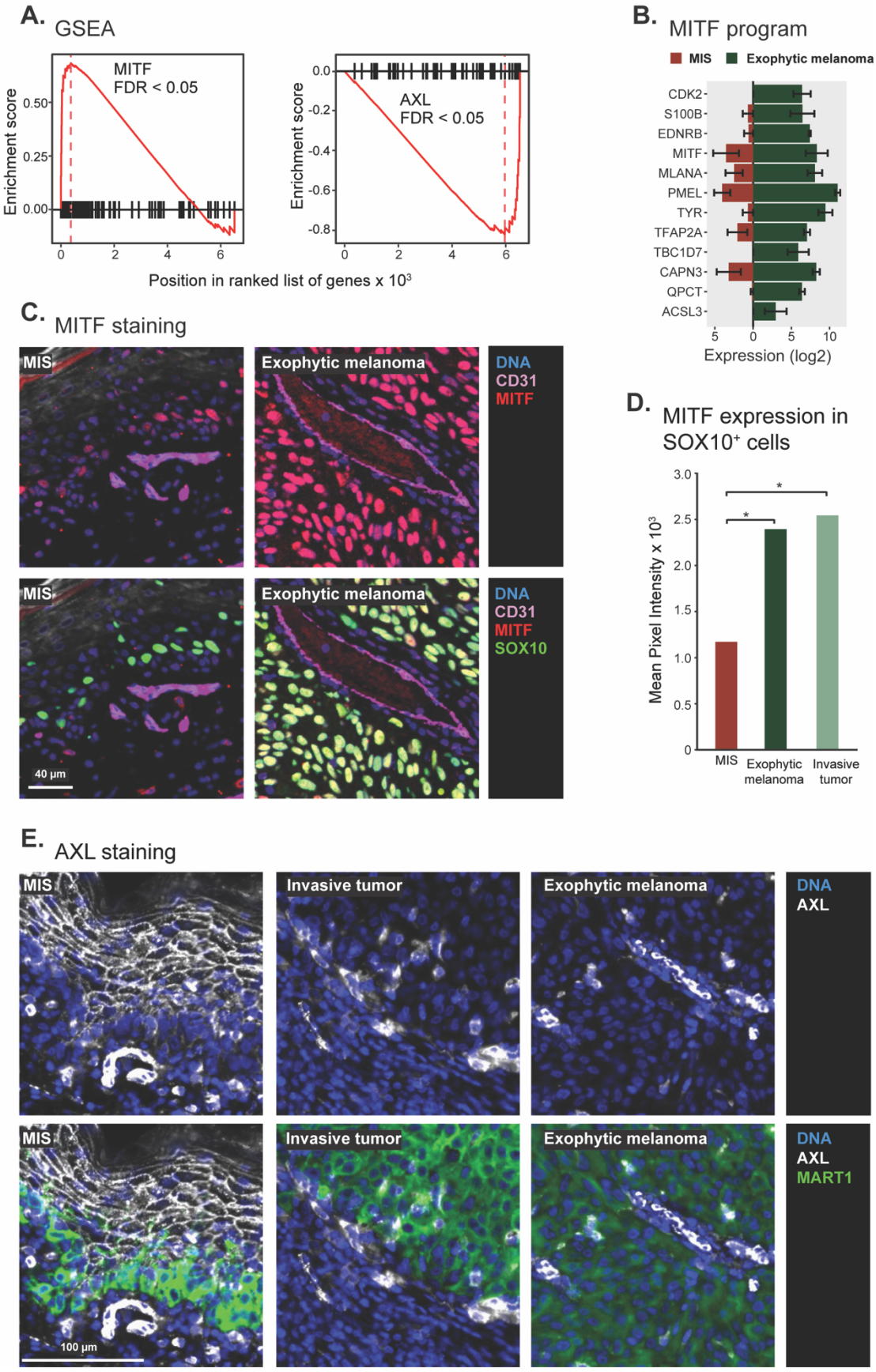


Fig. 6: ssGSEA of immune signatures for melanoma mROIs

Single-sample gene set enrichment analysis (ssGSEA) for immune cell, skin, and melanoma-related gene signatures (rows) in mROIs of similar histology (columns) from FFPE melanoma.



651

Fig. 7: Analysis of transcriptional signatures in melanoma

(A) Gene set enrichment analysis for MITF pathway (left panel) and AXL pathway (right panel) in exophytic melanoma (n=3) and MIS (n=28) samples. FDR < 0.05. **(B)** Expression (log2) of selected genes in the MITF pathway in MIS (n=28; red bars to the left) and exophytic melanoma (n=3; green bars to the right) samples. Data is mean ± SEM. **(C)** Fluorescence microscopy of MIS (left) and exophytic melanoma (right) stained for DNA (blue), blood vessels (CD31: violet), and the melanoma progression marker MITF (red); the bottom panels are also stained for melanocytes (SOX10: green). Scale bar, 40 µm. **(D)** MITF protein detected by fluorescence microscopy in melanocytes (SOX10 positive cells) in MIS, exophytic melanoma, or invasive tumor regions (* P value < 0.05, t-test). **(E)** Fluorescence microscopy images of MIS (left), invasive tumor (center), or exophytic melanoma (right) stained for DNA (blue), AXL (gray), and, in the bottom panels, melanocytes (MART1: green). Scale bar, 100 µm.

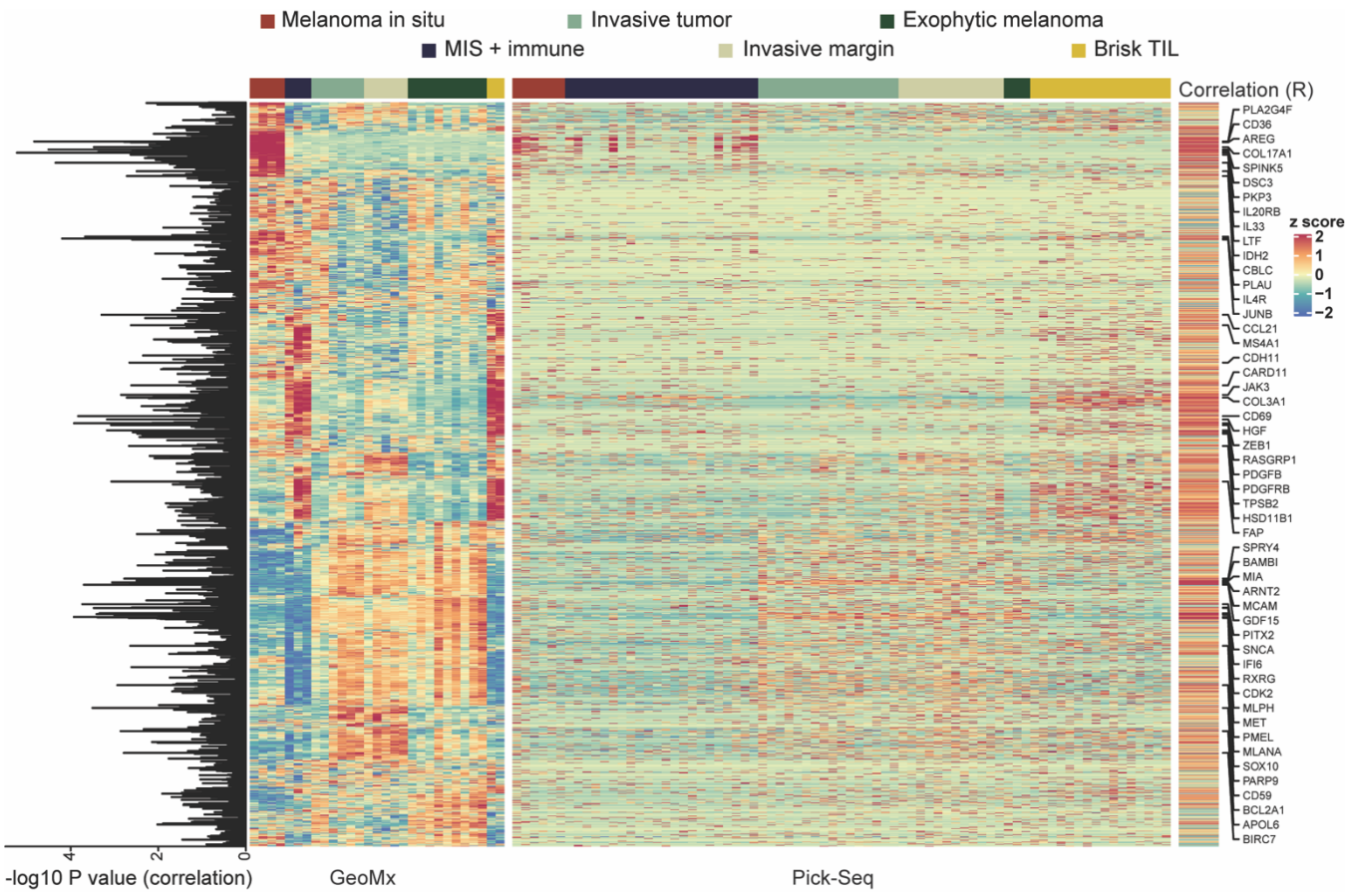


Fig. 8: Gene expression analysis with Pick-Seq and GeoMx

Expression of 1571 genes (rows) detected by both GeoMx and Pick-Seq organized by mROI and tissue histology (columns). Correlation for each gene between the two methods (right) and significance of correlation (P-value) between the two technologies (left) are indicated.

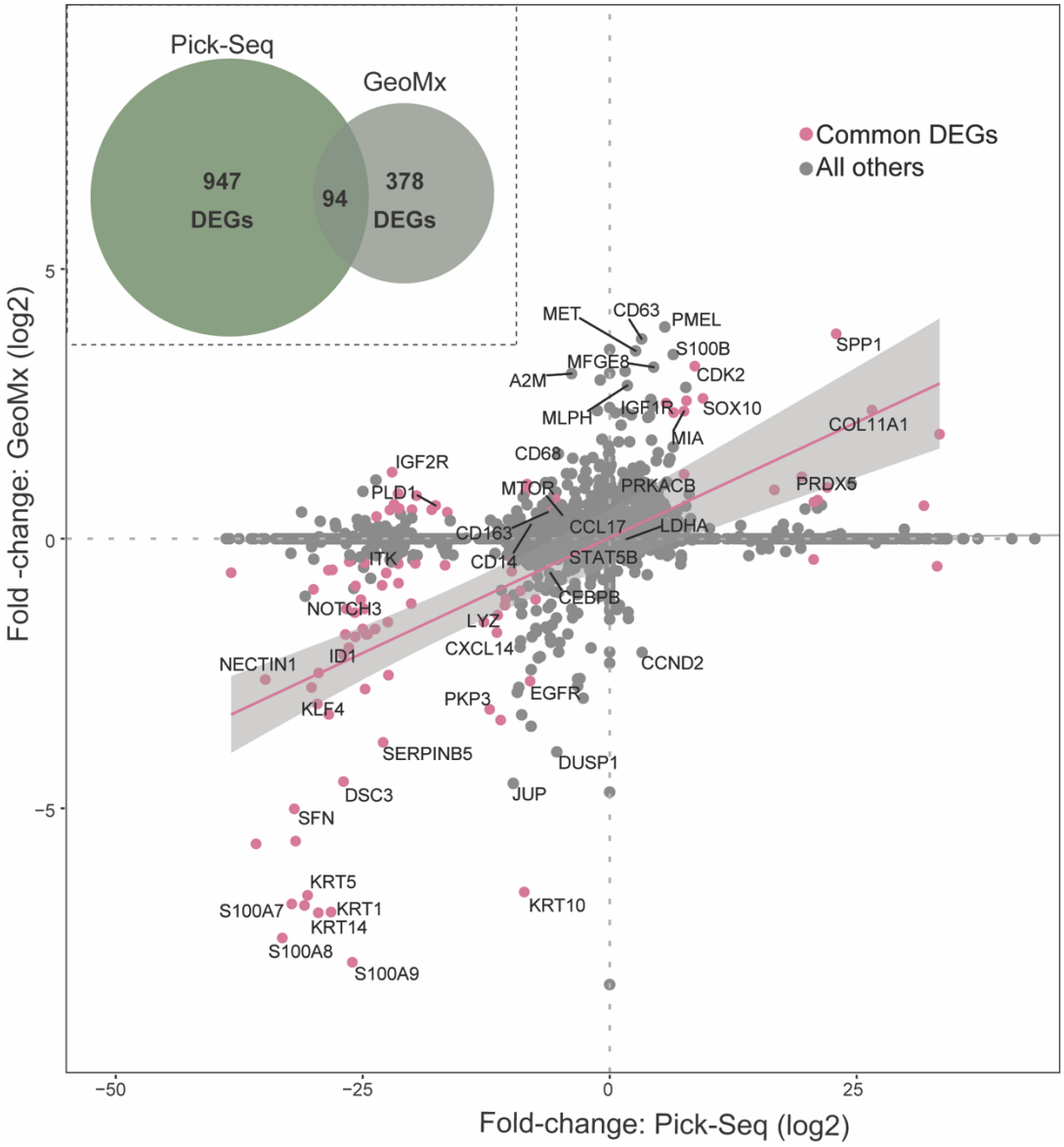
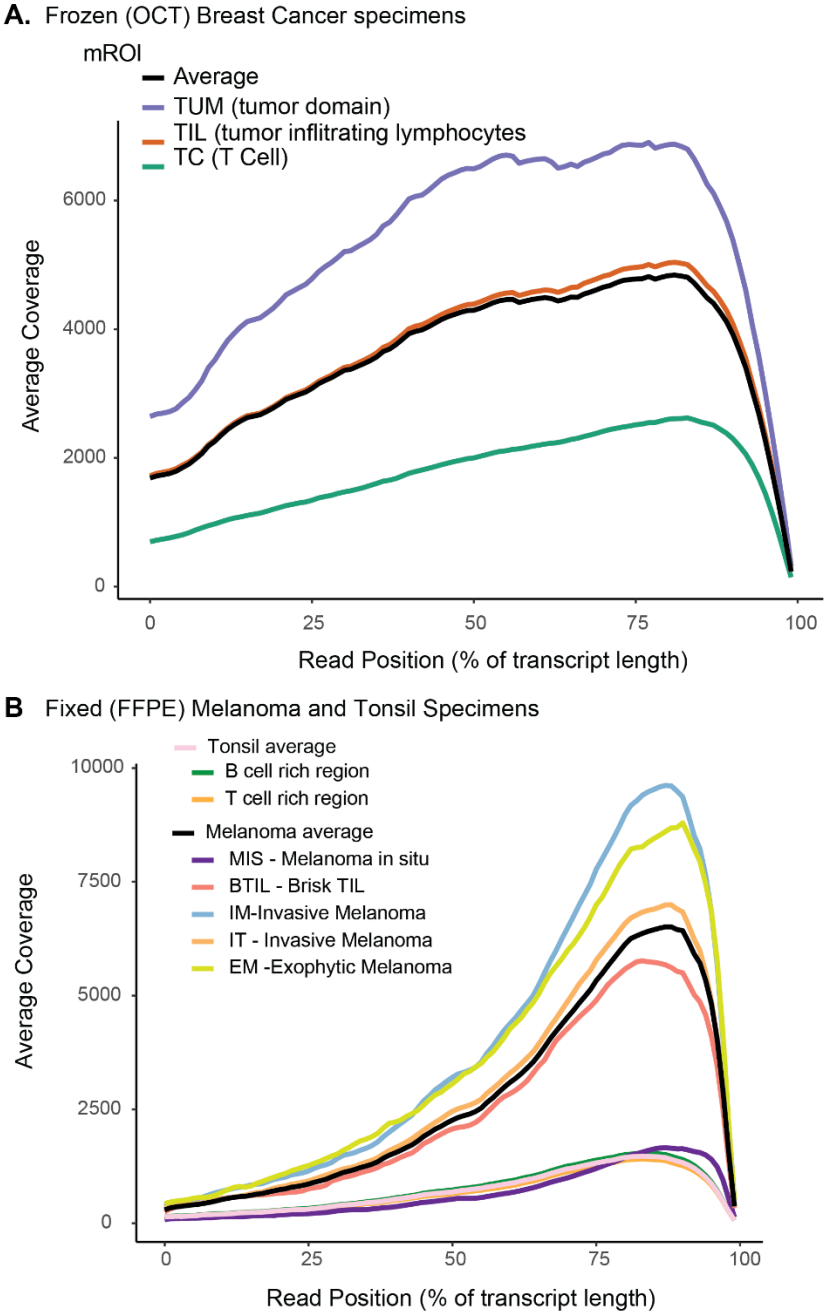


Fig. 9: Comparison of Pick-Seq and GeoMx differentially expressed genes

Differential expression (log2) for 94 DEGs (pink) measured using Pick-Seq and GeoMx. All other genes (gray). Trend-line (pink) and confidence interval (95%, gray) are indicated. For Pick-Seq, n=3 exophytic melanoma and n=28 MIS samples. For GeoMx, n=9 exophytic melanoma and n=7 MIS samples. Inset, top left, Venn diagram indicates the number of DEGs measured using one or both technologies.

675 **SUPPLEMENTARY MATERIALS**



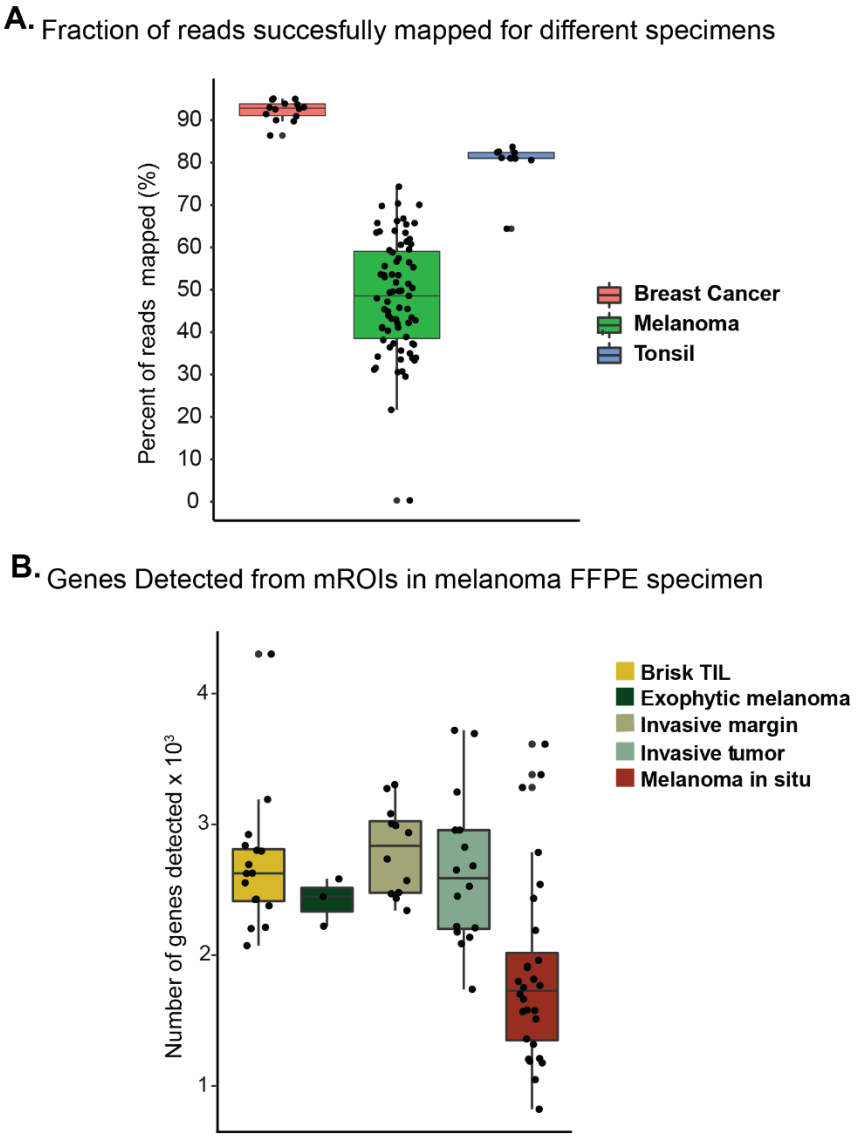
676

677 **Fig. S1: Positions of sequence reads**

678 Plots show mapped-read depth (y-axis) at each relative transcript position (x-axis). The read-depth is

679 averaged across mROI's belonging to the same sample group. The overall average for each tissue is also

680 shown. (**A)** Frozen (OCT) breast cancer mROIs. (**B)** FFPE tonsil and melanoma mROIs.



681

682 **Fig. S2: Reads mapped and genes detected for mROIs**

683 **(A)** Boxplot shows the median percentage of reads that could be mapped to the reference genome; the
684 lower and upper hinges correspond to the first and third quartiles. The points represent values for
685 individual mROI's. **(B)** Number of genes detected in each mROI sorted by histology feature; brisk TIL
686 (n=16), exophytic melanoma (n=3), invasive margin (n=12), invasive tumor (n=16), and MIS (n=28).

687 **Table S1: Normalized gene expression data for tonsil**

688 **Table S2: Normalized gene expression data for breast**

689 **Table S3: Normalized gene expression data for melanoma**

690 **Table S4: Normalized gene expression data for NanoString GeoMx**

691 **Table S5: Antibodies for FFPE tissue staining and imaging**

692 **Table S6: Regions of interest for transcriptional analysis**