Multi-modal digital pathology for colorectal cancer diagnosis by high-plex immunofluorescence imaging and traditional histology of the same tissue section

Jia-Ren Lin[1,2,*], Yu-An Chen[1,2,*], Daniel Campton[3,*], Jeremy Cooper[3], Shannon Coy[1,4], Clarence Yapp[1,2], Juliann B. Tefft[1,2], Erin McCarty[3], Keith L. Ligon[4], Scott J. Rodig[4], Steven Reese[3], Tad George[3], Sandro Santagata[1,2,4,±], Peter K. Sorger[1,2,±]

* These authors contributed equally ± These authors contributed equally

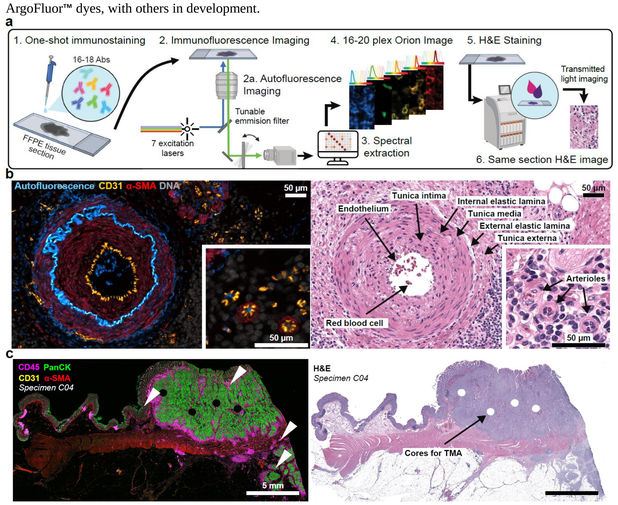
Human Tissue Atlas Center 1.Laboratory of Systems Pharmacology, Harvard Medical School, Boston, MA. 2. Ludwig Center at Harvard, Harvard Medical School, Boston, MA. 3.RareCyte, Inc., 2601 Fourth Ave., Seattle, WA.
4. Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA.

Abstract

Precision medicine is critically dependent on better methods for diagnosing and staging disease and predicting drug response. Histopathology using Hematoxylin and Eosin (H&E) stained tissue - not genomics – remains the primary diagnostic modality in cancer. Moreover, recently developed, highly multiplexed tissue imaging represents a means of enhancing histology workflows with single cell mechanisms. Here we describe an approach for collecting and analyzing H&E and high-plex immunofluorescence (IF) images from the same cells in a whole-slide format suitable for translational and clinical research and eventual deployment in diagnosis. Using data from 40 human colorectal cancer resections (60 million cells) we show that IF and H&E images provide human experts and machine learning algorithms with complementary information. We demonstrate the automated generation and ranking of computational models, based either on immune infiltration or tumor-intrinsic features, that are highly predictive of progression-free survival. When these models are combined, a hazard ratio of ~0.045 is achieved, demonstrating the ability of multi-modal digital pathology to generate high-performance and interpretable biomarkers.

Methods

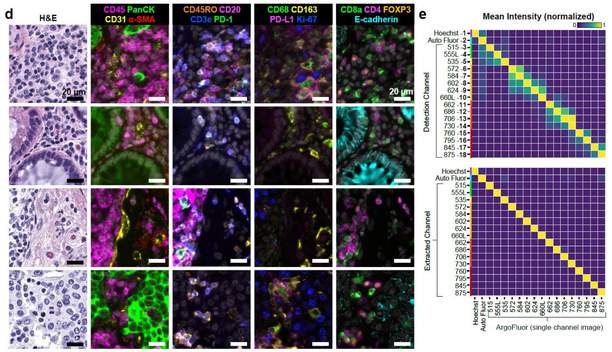
The Orion™ instrument is designed to work with an optimized set of fluorophores from RareCyte, branded as ArgoFluor™ dyes whose emission peaks cover the spectrum from green to far-red (Extended Data Table 2). Although the instrument can also be used with other commercially available dyes, the ArgoFluor™ dyes have been strategically chosen based on a combination of properties that include resistance to photobleaching, narrow excitation and emission spectra, and high quantum efficiency. To date, the company has optimized 18 ArgoFluor™ dyes, with others in development.



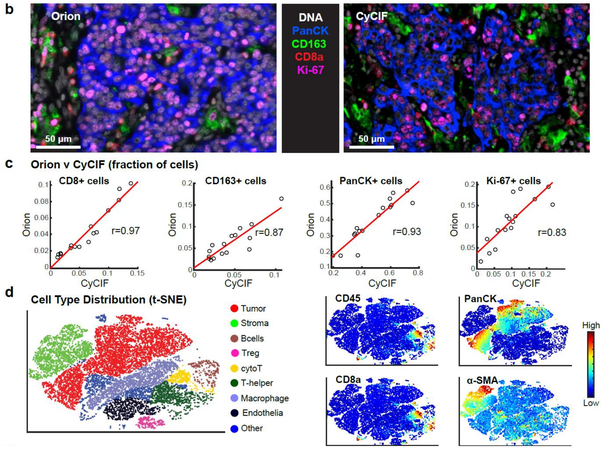
a, Schematic of one-shot 16 to 20-channel multiplexed immunofluorescence imaging with the Orion™ method followed by Hematoxylin and Eosin (H&E) staining of the same section using an automated slide stainer and scanning of the H&E-stained slide in transillumination (brightfield) mode. This method of discriminating the emission spectra of fluorophores is repeated using seven excitation lasers spaced across the spectrum (see Extended Data Fig. 1a and Methods section). Using polychroic mirrors and tunable optical filters, emission spectra are extracted to discriminate up 20 channels including signal from fluorophore-labelled antibodies (15-20 in most experiments), the nuclear stain Hoechst 33342, and tissue intrinsic autofluorescence. **b,** Left panels: Orion multiplexed immunofluorescence image showing CD31, α-SMA, Hoechst (DNA), and signal from the tissue autofluorescence channel (AF) from a colorectal cancer FFPE specimen (C04); this highlights an artery outside of the tumor region with red blood cells in the vessel lumen and elastic fibers in the internal and external elastic lamina of the vessel wall, numerous smaller vessels (arterioles), and stromal collagen fibers (inset displays arterioles). Right panels: images of the H&E staining from the same tissue section (histologic landmarks are indicated). Scalebars 50 μm. . **c,** Orion multiplexed immunofluorescence image (showing CD45, pan-cytokeratin, CD31, and α-SMA) from a whole tissue FFPE section of a colorectal cancer (C04) and matched H&E from the same section. Holes in the images are regions of tissue ('cores') removed in the construction of TMAs. Scalebar 5 mm.

Results

• Constructing and testing the Orion platform

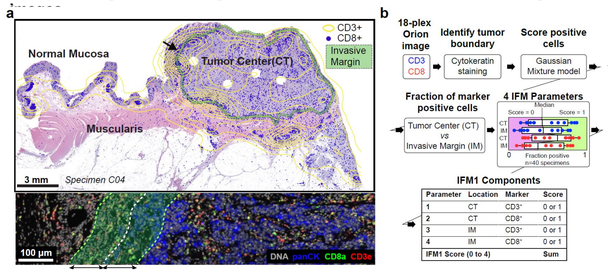


d, Zoom-in views of the regions indicated by arrowheads in panel c; marker combinations indicated. Scalebar 20 μm. **e,** Intensities of fluorochromes (columns in heatmaps) in each Orion channel (rows in heatmaps) prior to (top) and after (bottom) spectral extraction. The extraction matrix was determined from control samples scanned using the same acquisition settings that were used for the full panel. The control samples included: unstained lung tissue (for the autofluorescence channel), tonsil tissue stained with Hoechst, and tonsil tissue stained in single-plex with ArgoFluor-conjugates used in the panel (for the biomarker channels). The values in each column were normalized to the maximum value in the column.

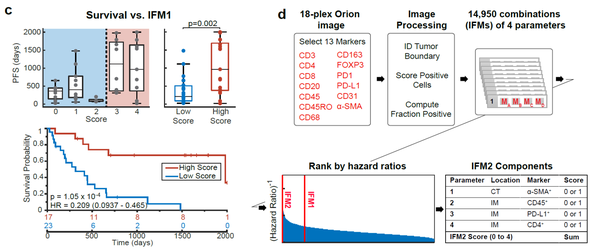


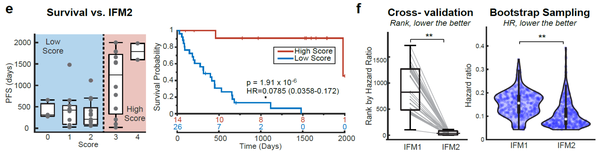
b, Orion IF images and cyclic immunofluorescence (CyCIF) images from neighboring sections of an FFPE colorectal adenocarcinoma; Scalebars 50 μm. The CyCIF images were done with 2x2 binning while Orion images were obtained with no binning. **c,** Plots of the fraction of cells positive for the indicated markers from whole slide Orion IF and CyCIF images acquired from neighboring sections from 16 FFPE colorectal cancer specimens. Pearson correlation coefficients are indicated. **d,** t-distributed stochastic neighbor embedding (t-SNE) plots of cells derived from CyCIF (left panels) of a FFPE colorectal cancer specimen (C01) with the fluorescence intensities of immune (CD45, pan-cytokeratin, CD8a, α-SMA) markers overlaid on the plots as heat maps

Recapitulating and extending the Immunoscore tissue immune test using Orion



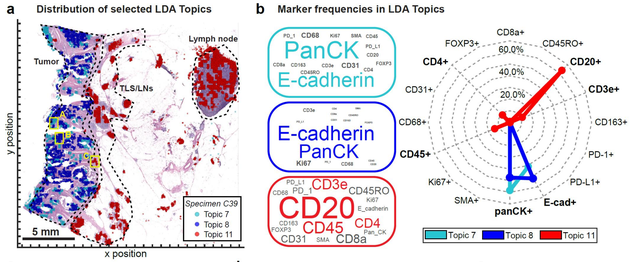
a, Map of tumor center and invasive-margin compartments for specimen C04 overlaid on an H&E image with the density of CD3+ cells shown as a contour map (yellow) and the positions of CD8+ T cells as blue dots. The arrow indicates the zoom-in images shown below. Lower panel shows selected channels from a portion of the Orion image for C04 spanning the invasive boundary (denoted by green shading). **b,** Flow chart for the calculation of Image Feature Model 1 (IFM1) that recapitulates key features of the Immunoscore test. c, Upper panel: Box-and-whisker plots for progression-free survival (PFS) for 40 CRC patients based on actual IFM1 scores (midline = median, box limits = Q1 (25th percentile)/Q3 (75th percentile), whiskers = 1.5 inter-quartile range (IQR), dots = outliers (>1.5IQR) or scores stratified into two classes as follows, low: score ≤ 2, high: score = 3 or 4 (pairwise two-tailed t-test p = 0.002. Lower panel: Kaplan Meier plots computed using IFM1 binary classes (HR, hazards ratio; 95% confidence interval; logrank p-value).

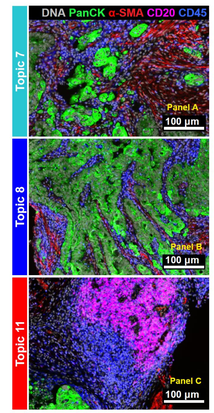


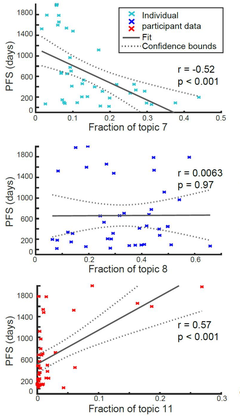


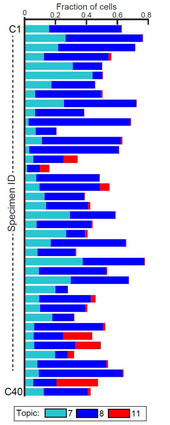
c, Upper panel: Box-and-whisker plots for progression-free survival (PFS) for 40 CRC patients based on actual IFM1 scores (midline = median, box limits = Q1 (25th percentile)/Q3 (75th percentile), whiskers = 1.5 inter-quartile range (IQR), dots = outliers (>1.5IQR) or scores stratified into two classes as follows, low: score ≤ 2, high: score = 3 or 4 (pairwise two-tailed t-test p = 0.002. Lower panel: Kaplan Meier plots computed using IFM1 binary classes (HR, hazards ratio; 95% confidence interval; logrank p-value). **d,** Flow chart for calculation of additional models that use the underlying logic of Immunoscore but considering 13 markers. The image processing steps are the same as in panel a. The rank positions of IFM1 and IFM2 are shown relative to all other 14,950 combinations of parameters that were considered. **e,** (Left) Box-and-whisker plots for PFS for 40 CRC patients based on IFM2 scores. (Right) Kaplan Meier plots computed using IFM2 binary classes stratified into two classes as follows, low: score ≤ 2, high: score = 3 or 4 (HR, hazards ratio; 95% confidence interval; logrank p-value). **f.** Plots of leave-one-out cross-validation of ranks from IFM1 and IFM2 (unadjusted p = 3.35 x 10-14 and adjusted using the Benjamini-Hochberg Procedure; p=5.0 x 10-9) and bootstrapping of hazard ratios from IFM1 and IFM2 (unadjusted p = 1.94 x 10-25 and adjusted p = 2.91 x 10-20). Detailed analysis was described in the methods section and pairwise two-tailed t-test were used unless otherwise mentioned.

• Bottom-up development of a tumor-intrinsic image feature model.









a, Positions in specimen C39 of three selected topics identified using Latent Dirichlet Allocation (LDA). Topic locations are overlaid on an H&E image; Scalebar 5 mm. **b,** Left: Markers making up selected LDA topics as shown with size of the text proportional to the frequency of the marker but with colored text scaled by 50% for clarity; Radar plot indicating the fraction of cells positive for each marker in Topic 7, 8, and 11 (data for all others topics show in Extended Data Fig. 5). **c,** Immunofluorescence images showing expression of pan-cytokeratin, α-SMA, CD45, and CD20 for the indicated LDA topics. The position of each image frame is denoted by the yellow boxes in panel a. Scalebars 100 μm. **d,** Pearson correlation plots of progression-free survival (PFS) and Fraction of Topic 7, 8 and 11 in 40 CRC patients. Topic 11 corresponded to TLS, whose presence is known to correlate with good outcome77. **e,** Fraction of Topics 7, 8, and 11 in colorectal cancer specimens C1-C40.

Acknowledgement

This work was supported by NCI grants U54-CA225088 and U2C-CA233262 (PKS, SS), an NCI SBIR small business grant to RareCyte and PKS (R41-CA224503), and commercial investment from RareCyte; data processing software was developed with support from a Team Science Grant from the Gray Foundation and Ludwig Cancer Research (PKS, SS). SS is supported by the BWH President's Scholars Award. We are grateful to all members of the HMS Laboratory of Systems Pharmacology (LSP) engaged in tissue imaging (see https://www.tissue-atlas.org/), to Joe Victor, and to members of the RareCyte software and hardware development teams.