Supplementary Material for

Analysis of multispectral imaging with the AstroPath platform informs efficacy of PD-1 blockade

Sneha Berry, Nicolas A. Giraldo, Benjamin F. Green, Tricia R. Cottrell, Julie E. Stein, Elizabeth Engle, Haiying Xu, Aleksandra Ogurtsova, Charles Roberts, Daphne Wang, Peter Nguyen, Qingfeng Zhu, Sigfredo Soto-Diaz, Jose Loyola, Inbal B. Sander, Pok Fai Wong, Shlomit Jessel, Joshua Doyle, Danielle Signer, Richard Wilton, Jeffrey S. Roskes, Margaret Eminizer, Seyoun Park, Joel C. Sunshine, Elizabeth M. Jaffee, Alexander Baras, Angelo DeMarzo, Suzanne L. Topalian, Harriet Kluger, Leslie Cope, Evan J. Lipson, Ludmila Danilova, Robert A. Anders, David Rimm, Drew M. Pardoll, Alexander S. Szalay, Janis M. Taube*

*Corresponding author. Email: jtaube1@jhmi.edu

Published 11 June 2021, *Science* **372**, eaba2609 (2021)
DOI: 10.1126/science.aba2609

This PDF file includes:

Other Supplementary Material for this manuscript includes the following:
(available at science.sciencemag.org/content/372/6547/eaba2609/suppl/DC1)

Figure S1

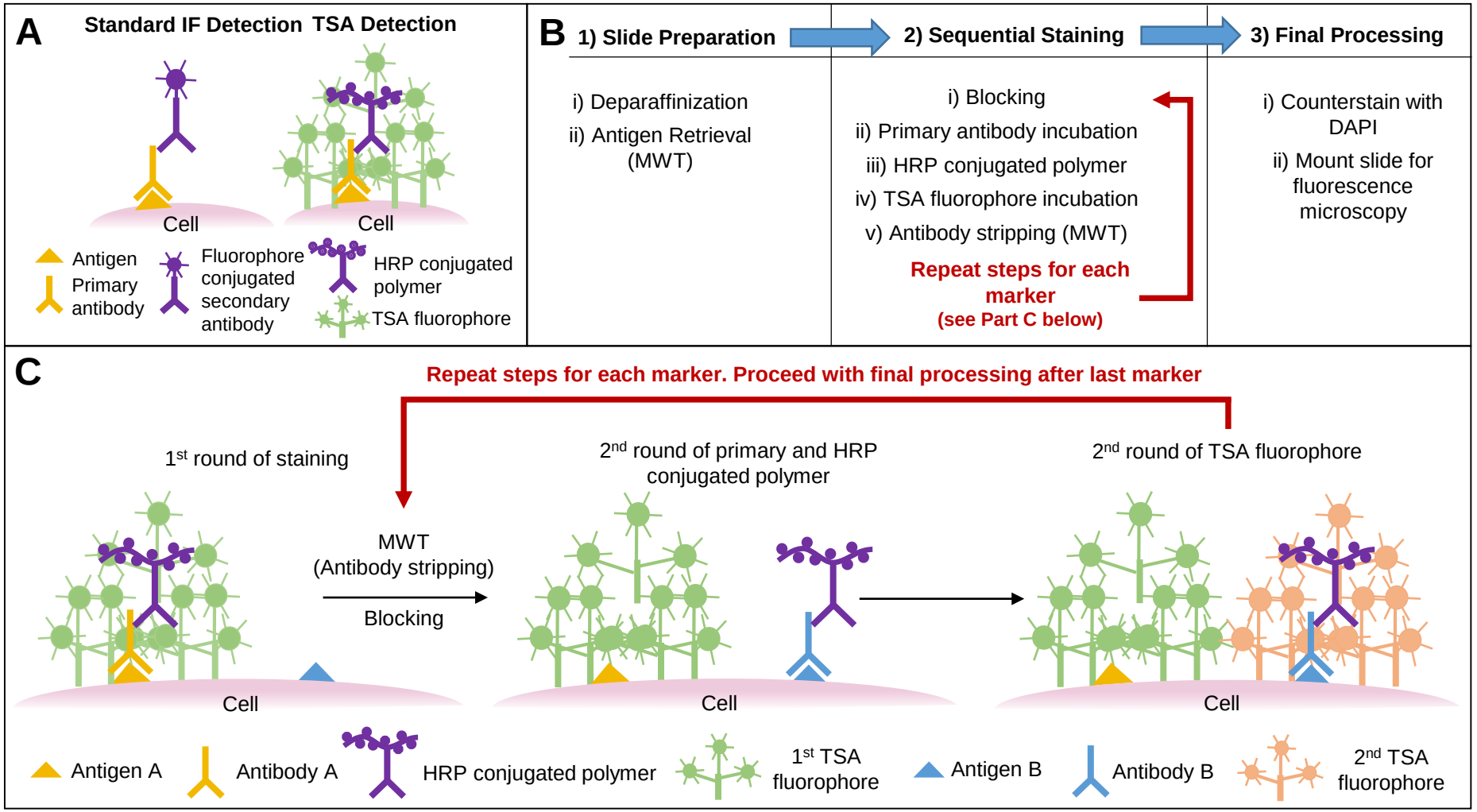


Figure S1. Tyramide signal amplification (TSA) technology can be used to amplify signal and visualize multiple markers on a single slide. **(A)** TSA detection allows for greater amplification (~1000 fold) of signal when compared to staining using a fluorophore tagged secondary antibody. This ability is attributed to the deposition of multiple TSA fluorophore molecules by an enzyme catalyzed reaction. **(B)** The multiplex staining process can be broken down in to three phases: slide preparation, sequential staining and final processing. **(C)** In the sequential staining phase, microwave treatment (MWT) strips off antibodies from prior staining rounds while retaining the deposited TSA fluorophore due its stronger binding to the tissue. The staining process can be repeated for multiple markers without any cross-reactivity.

Figure S2

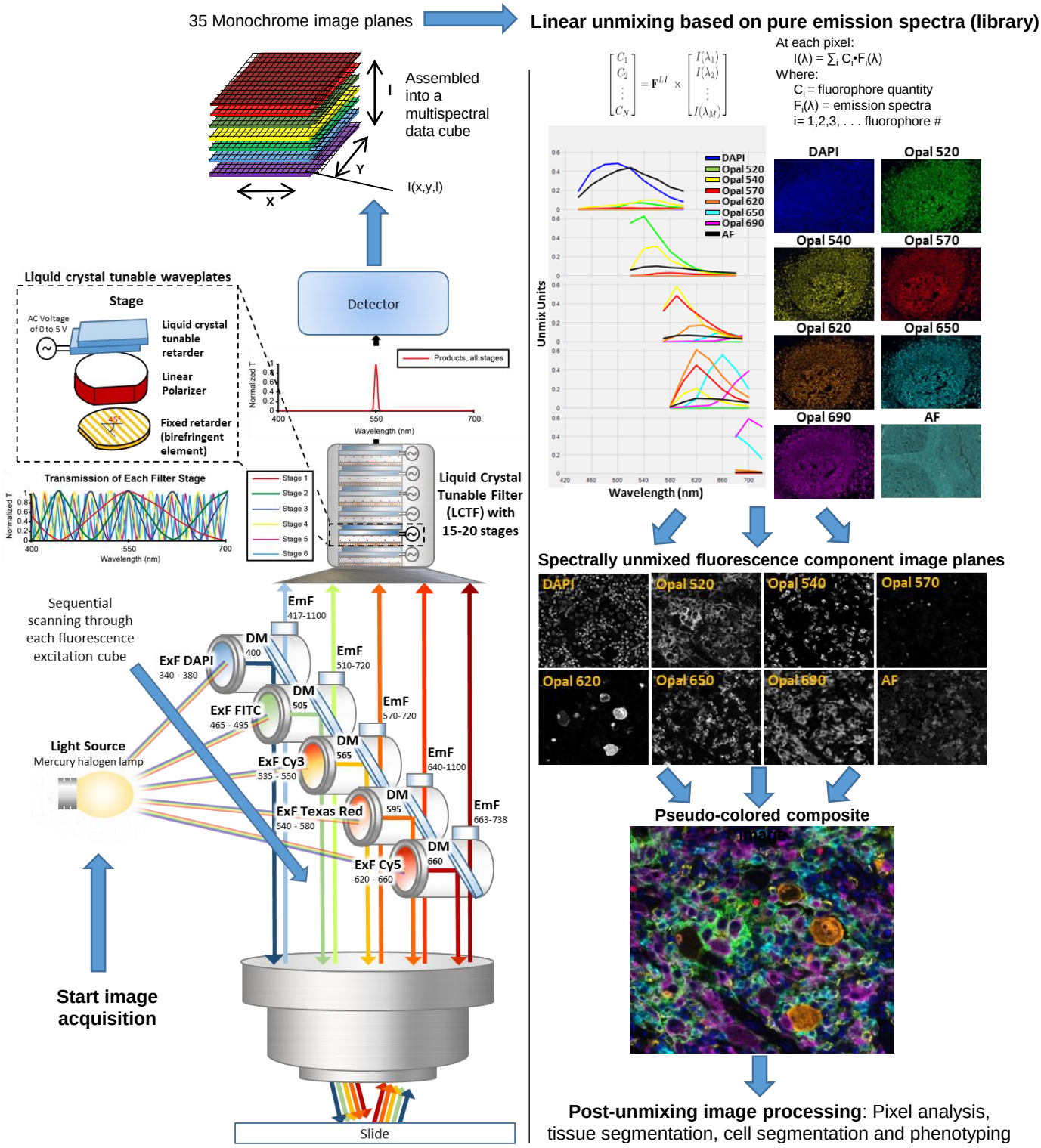
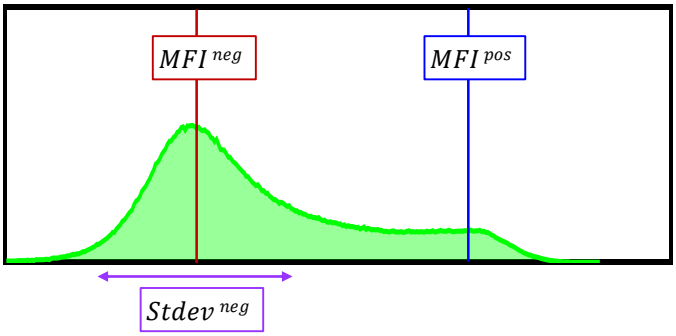


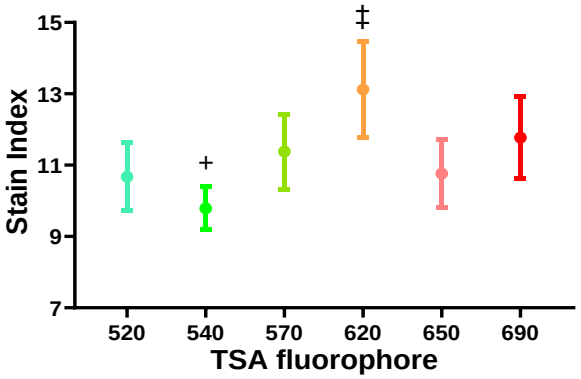
Figure S2. Multispectral image acquisition using the Vectra 3.0 system allows for the simultaneous visualization of six channels of interest plus DAPI. A mercury-halogen lamp emits light that is received by five excitation cubes whose wavelengths span the visible spectrum. Light is next received by the liquid crystal tunable filter which allows specific wavelengths to pass through, each one forming an individual monochromatic image plane. The resultant images are then unmixed using a library of pure spectra for each fluorophore. The individual images for each fluorophore are then pseudo-colored and overlayed to form a composite image. Unmixed images are then further processed using inForm software.

Figure S3

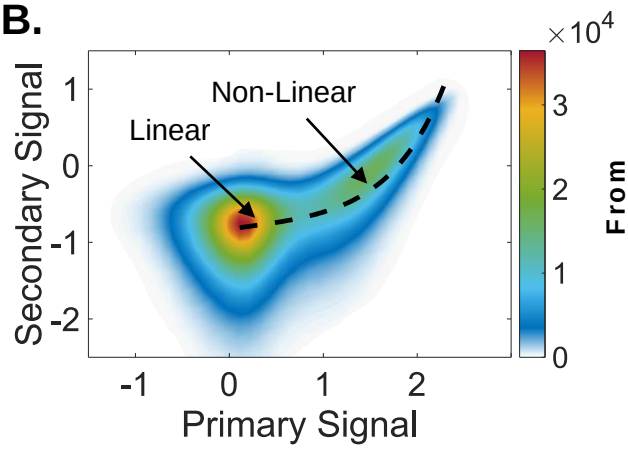
A.

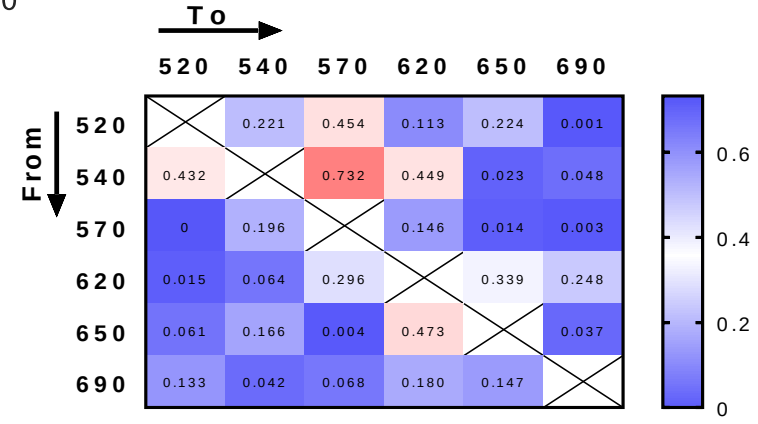


$$Stain\ Index = \frac{MFI^{pos} - MFI^{neg}}{2 * Stdev^{neg}}$$



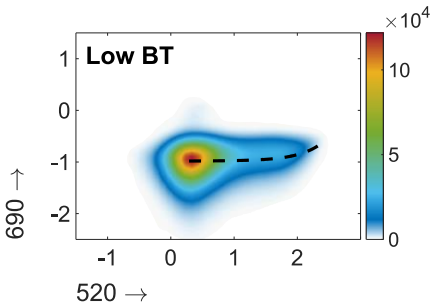
B.

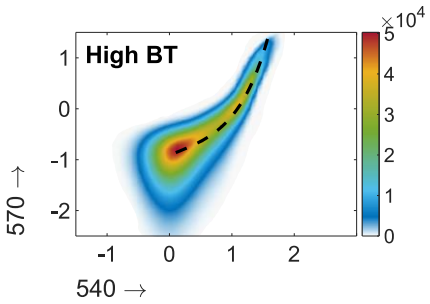




Parameterized hyperbolic sine function:

$$f(x) = \frac{A}{a} * sinh\big(a * (x - h)\big) + k + (A * h)$$





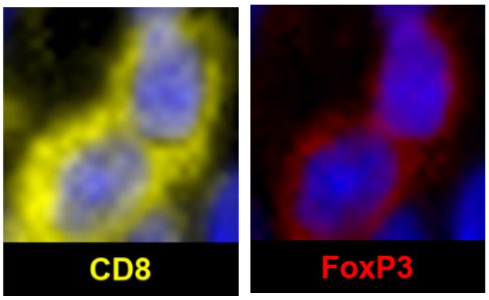


Figure S3. Characterization of TSA fluorophores for stain index (SI) and bleed-through (BT) **(A)** The SI is a signal to background metric useful for quantifying the brightness of immunofluorescent reagents. Fluorophores 540 and 620 had the lowest and the highest SIs respectively. TSA fluorophores with lower SIs were paired with more abundant markers e.g. CD8 (an abundant, strong antigen) is paired with the 540 fluorophore. Data represent mean ± SEM of n=5 tonsil specimens, 10 HPF each, stained with each TSA fluorophore at 1:50, using anti-CD8 at 1:100. + significantly different from 570, 620, 690, $p<0.001$. ‡ significantly different from all other fluorophores, $p<0.001$. **(B)** BT is the detection of false positive signal in a channel due to spillover from a different channel. The propensity for BT of each fluorophore, when used at a dilution of 1:50, was characterized. Top left: The logarithm of the normalized intensity of fluorescence for each possible TSA fluorophore-TSA fluorophore pair was plotted. A parameterized hyperbolic sine curve was fitted as shown on the graph. Top right: A*a in the parameterized hyperbolic sine function shows the propensity for BT from each channel to another. The most substantial BT between fluorophores ranked from high to low are: 540 to 570, 650 to 620, 520 to 570, 540 to 620 and 540 to 520. Bottom left: Examples of low and high BT. Bottom right: Representative 540 to 570 BT is shown in the photomicrographs where membranous signal from CD8 cells is seen in the 570 (FoxP3) channel (real FoxP3 staining is nuclear). The propensity for BT is further reduced by diluting the TSA fluorophores during subsequent steps of panel optimization. For example a 4 fold reduction in BT (percentage FoxP3 positive pixels) was seen when slides were stained for CD8 with a dilution of 1:100 versus 1:50.

Figure S4

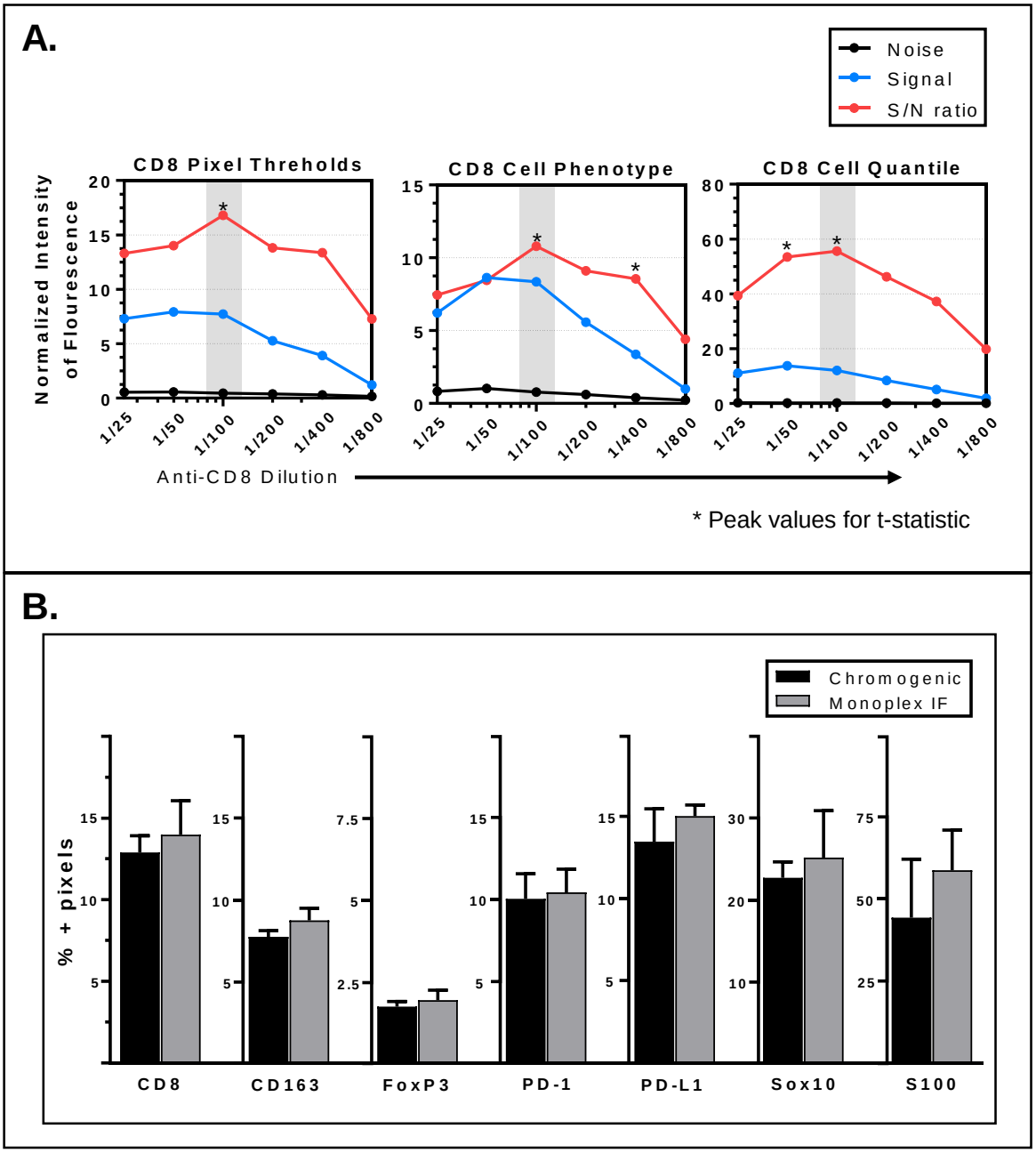


Figure S4. Primary antibody optimization is required to maximize IF staining specificity using chromogenic IHC as the gold standard. After selection of the appropriate HRP-conjugated polymer **(Figure 2B),** primary antibody dilutions were performed to optimize the signal to noise (S/N) ratio. **(A)** Representative figure for CD8 monoplex IF staining indicating that 1:100 is the dilution with the optimal S/N ratio. Concordance was seen between three different approaches for signal quantification (see **Materials and Methods**). **(B)** When the appropriate HRP-conjugated polymer is paired with the optimal primary antibody concentration, monoplex IF yields equivalent signal to chromogenic IHC. Welch's t test, * $p \leq 0.05$

Figure S5

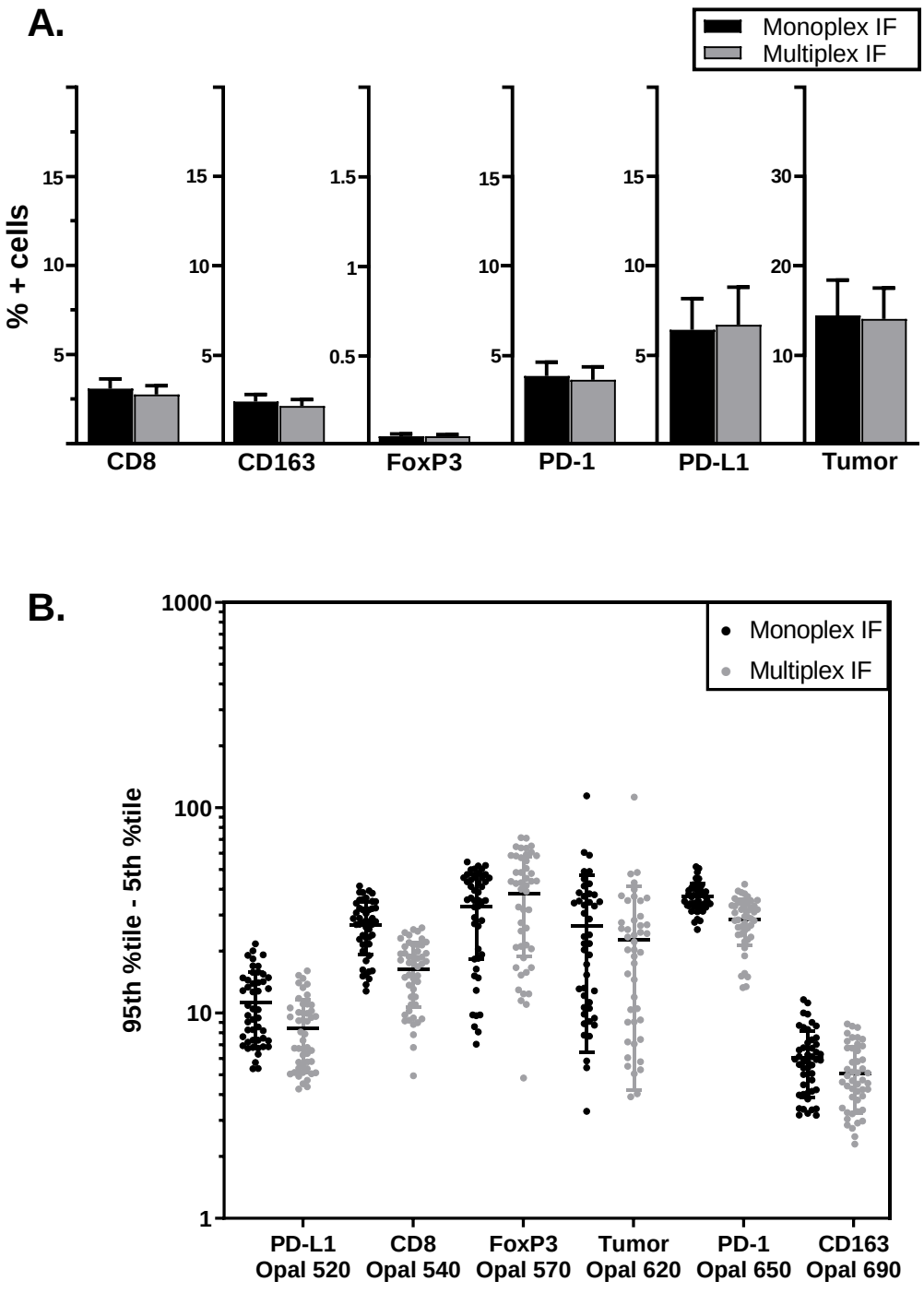
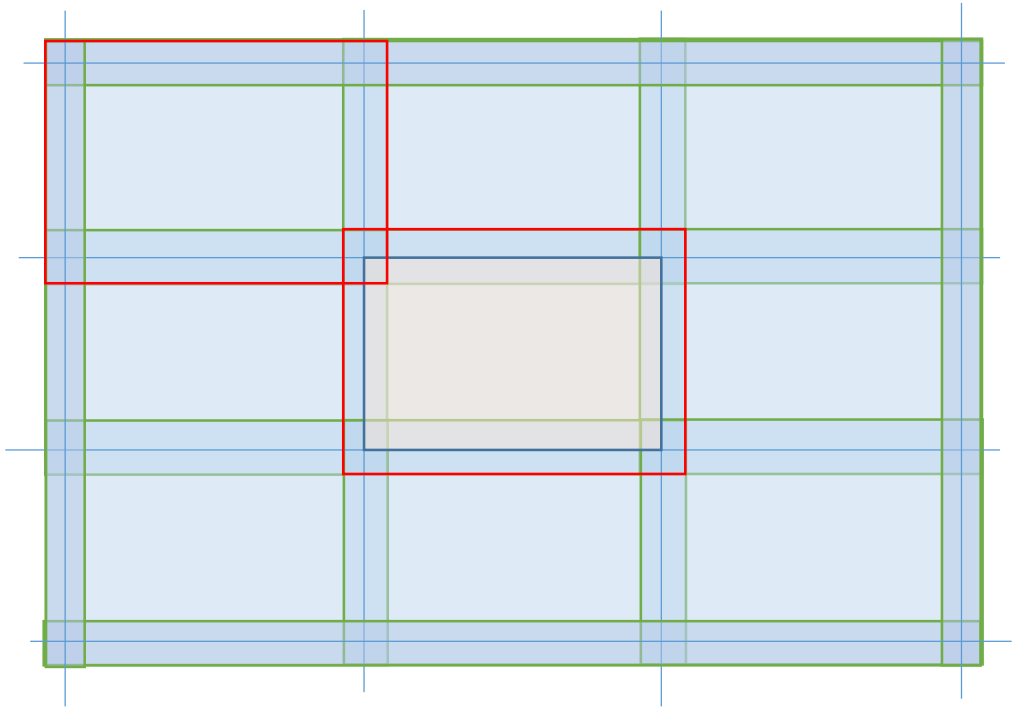


Figure S5. Comparison of monoplex IF and multiplex IF staining (A) When optimized as detailed herein, multiplex IF yields an equivalent percentage of positive cells to monoplex IF. **(B)** The usable dynamic range of the epitope was reduced by 13% on average in multiplex IF format. Each dot represents the difference between the 95th percentile and 5th percentile of the mean normalized fluorescence intensity of positive cells for a single HPF.

Figure S6

A.



B.

$$f(x) = \frac{(T-O)^2}{T^2} = \left(1 - \frac{4x}{1-x}\right)^2 = \frac{(1-5x)^2}{(1-x)^2}$$

C.

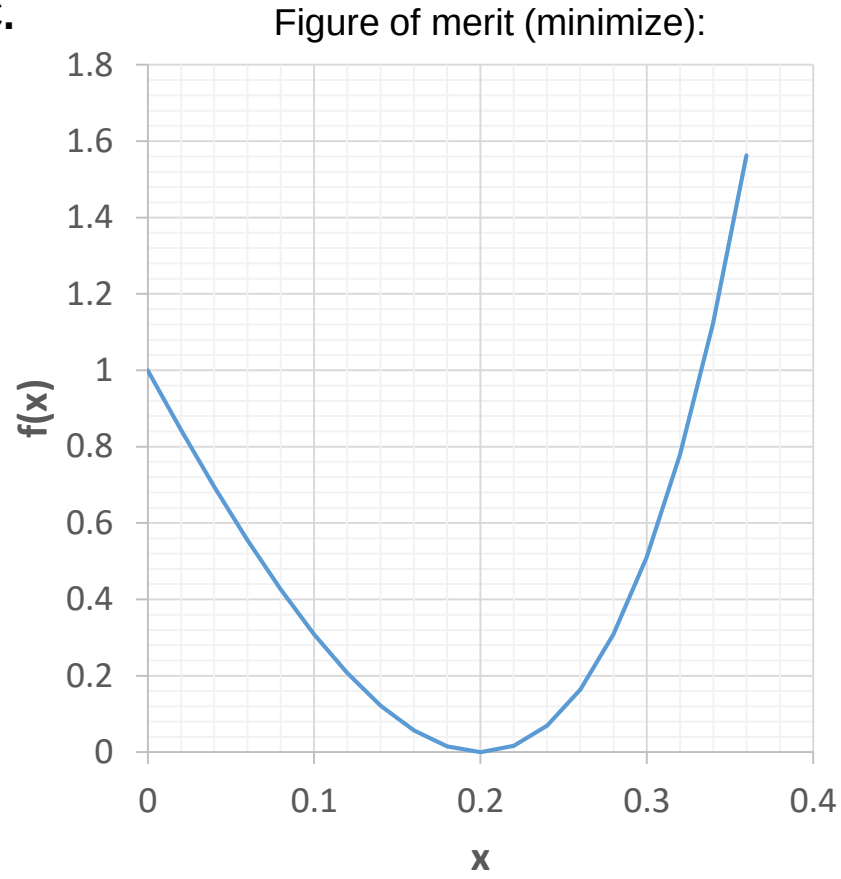


Figure S6. The optimum overlap of neighboring tiles is x=20% of the tile width and height. (A) Overlapping image tiles (examples shown in red) are used to create a seamless coverage of the whole area, built from the central rectangles of each image (blue lines, with peach shading showing one central rectangle). These central rectangles form the statistical sample for our analysis, and fully cover the tissue. The overlaps (areas shaded in darker blue), are observed multiple times and are used for intrinsic error estimates. **(B)** Too much overlap is "costly" in terms of data resources and time, while too little overlap fails to provide enough information to correct for imaging deficiencies. The information content (inverse variance) in estimating the corrections is proportional to the areas T and O, respectively. In the equation above, the useful area is T, representing the tissue area on the slide, and O is the area of the overlaps. **(C)** The optimum solution is x=0.2, i.e. 20%, corresponding to the case when the area that is imaged multiple times (O) is equal to the area of the tissue itself (T).

Figure S7

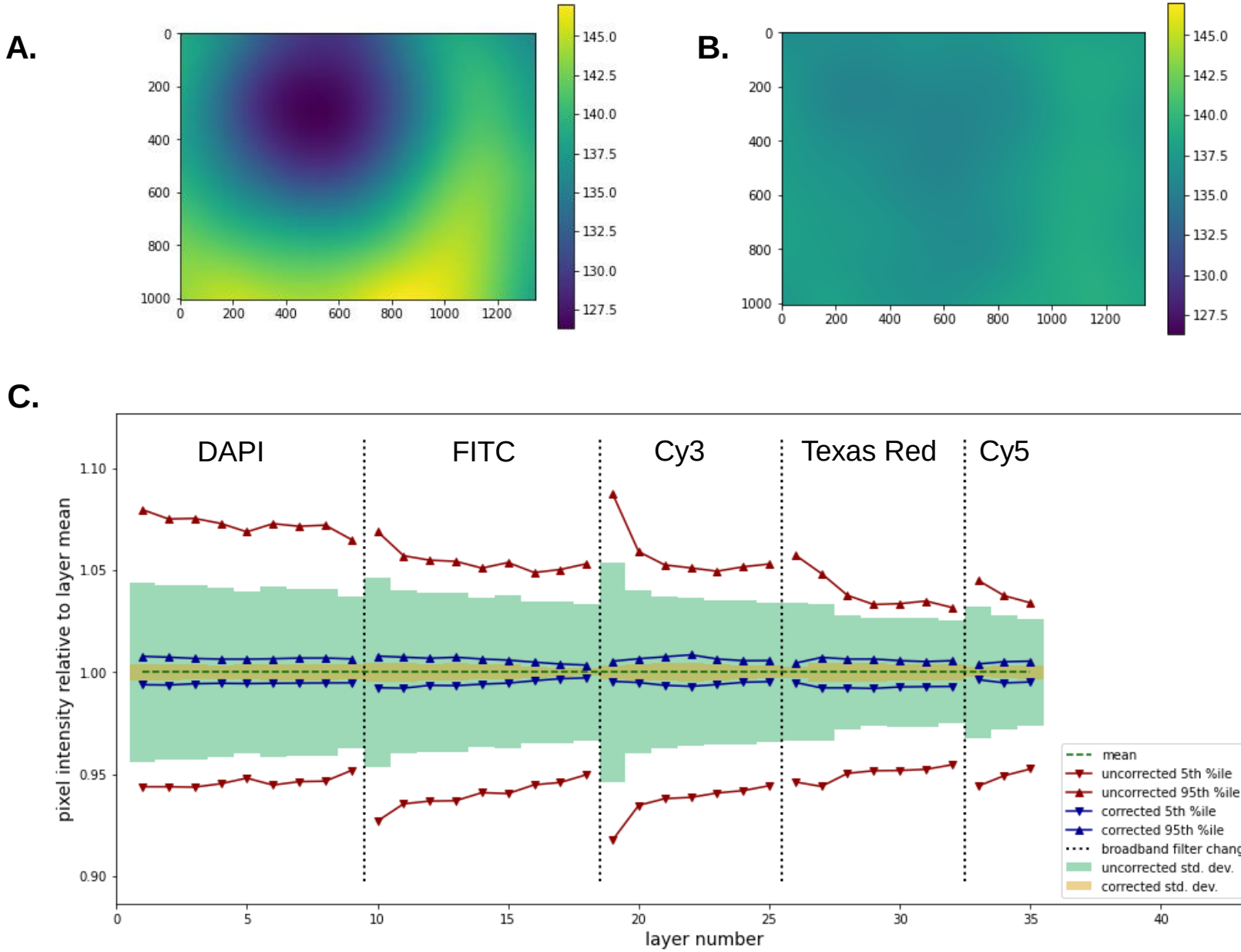
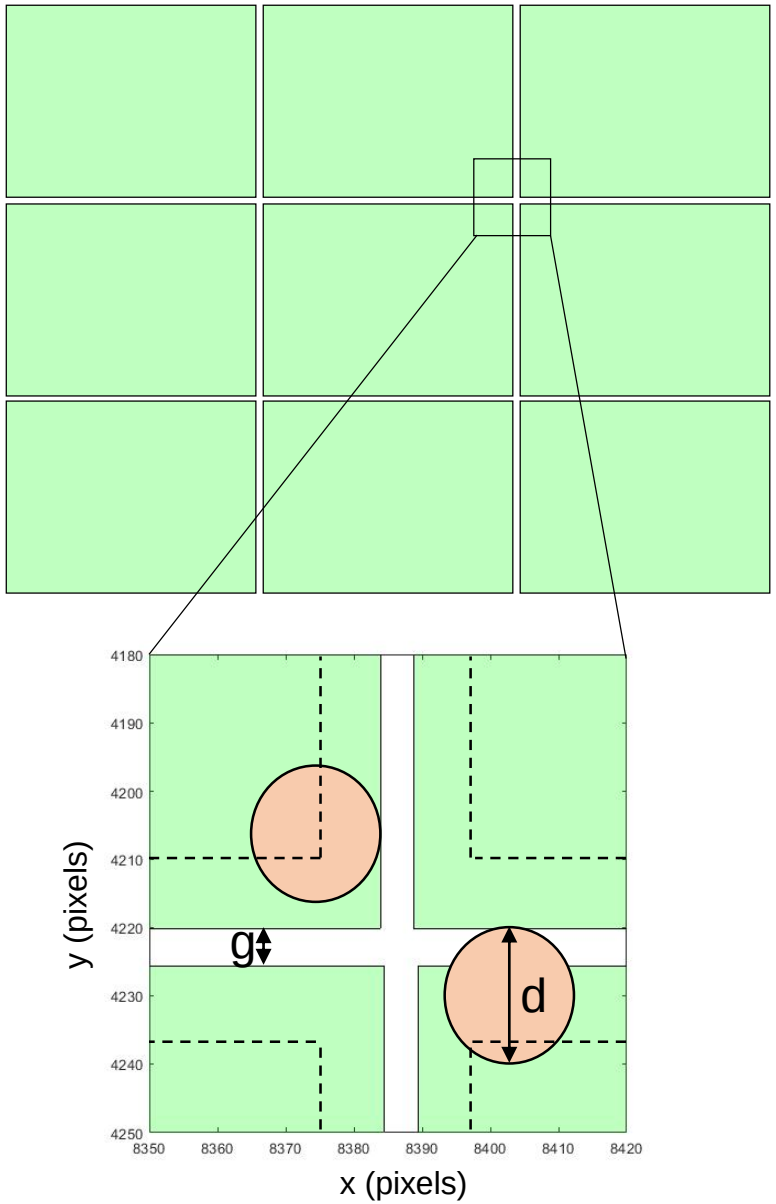


Figure S7. Image processing of individual fields included flatfield corrections for systematic illumination variation. (A) An average of 11,508 images were stacked to define the average illumination variation by image layer across a single HPF. Shown is the uncorrected, smoothed mean image for layer 13 (FITC broad band filter, PD-L1 in this study). **(B)** A flatfield model was developed and applied, and the resultant smoothed, corrected mean image is shown. **(C)** Relative pixel intensities between uncorrected and corrected images showed a consistent 9-fold reduction of illumination variation (11.2% to 1.2% for the 5^{th}-95^{th} percentile and 3.6% to 0.4% standard deviation on average). Pixel intensities relative to mean layer intensities are shown here across all image layers for one representative sample.

Figure S8

A.



$$\%\ \text{cells at edge(s) of HPFs} = \frac{(d+g)(L+W)}{(LW)}$$

L (length of HPF) = 1300 pixels
W (width of HPF) = 1000 pixels
g (pixel gap/random relative position error) = 0-10 pixels
d (cell diameter) = 18-24 pixels

Cell diameter (pixels)	Error before correction*
20 (average cell size in all specimens)	**3.5 - 5.3%**
18 (typical lymphocyte)	**3.2 - 5%**
24 (typical tumor cell)	**4.2 - 6%**

* Cells at HPF edges are prone to erroneous segmentation and flux measurements

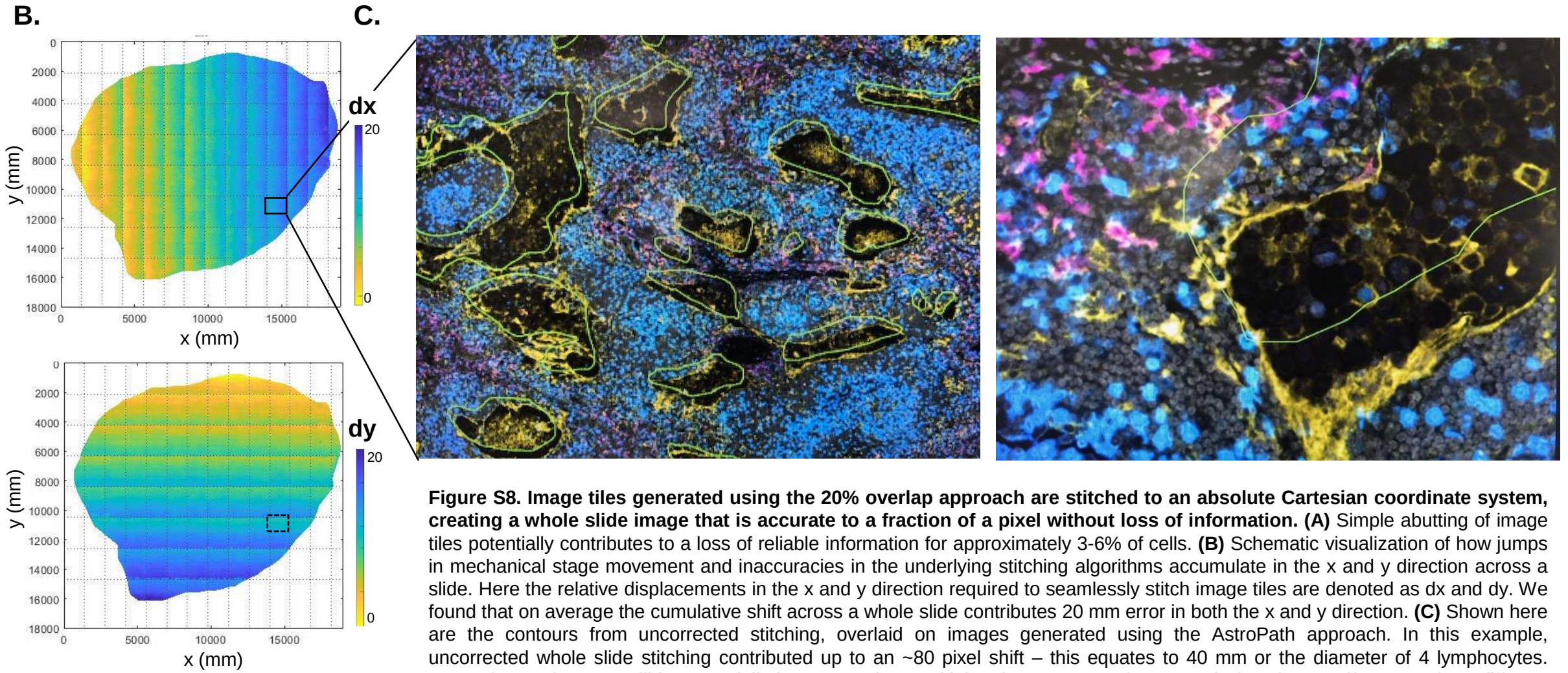
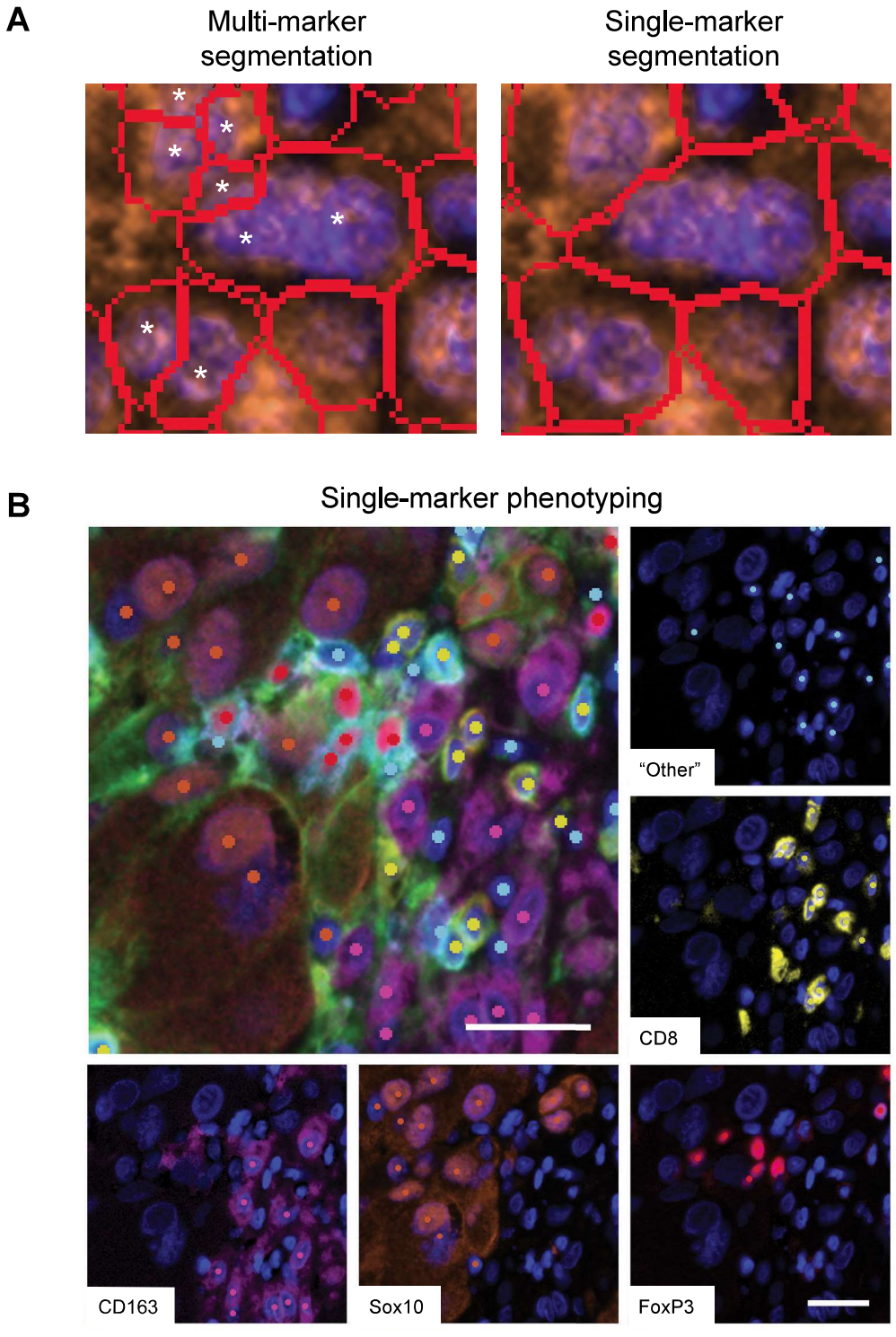
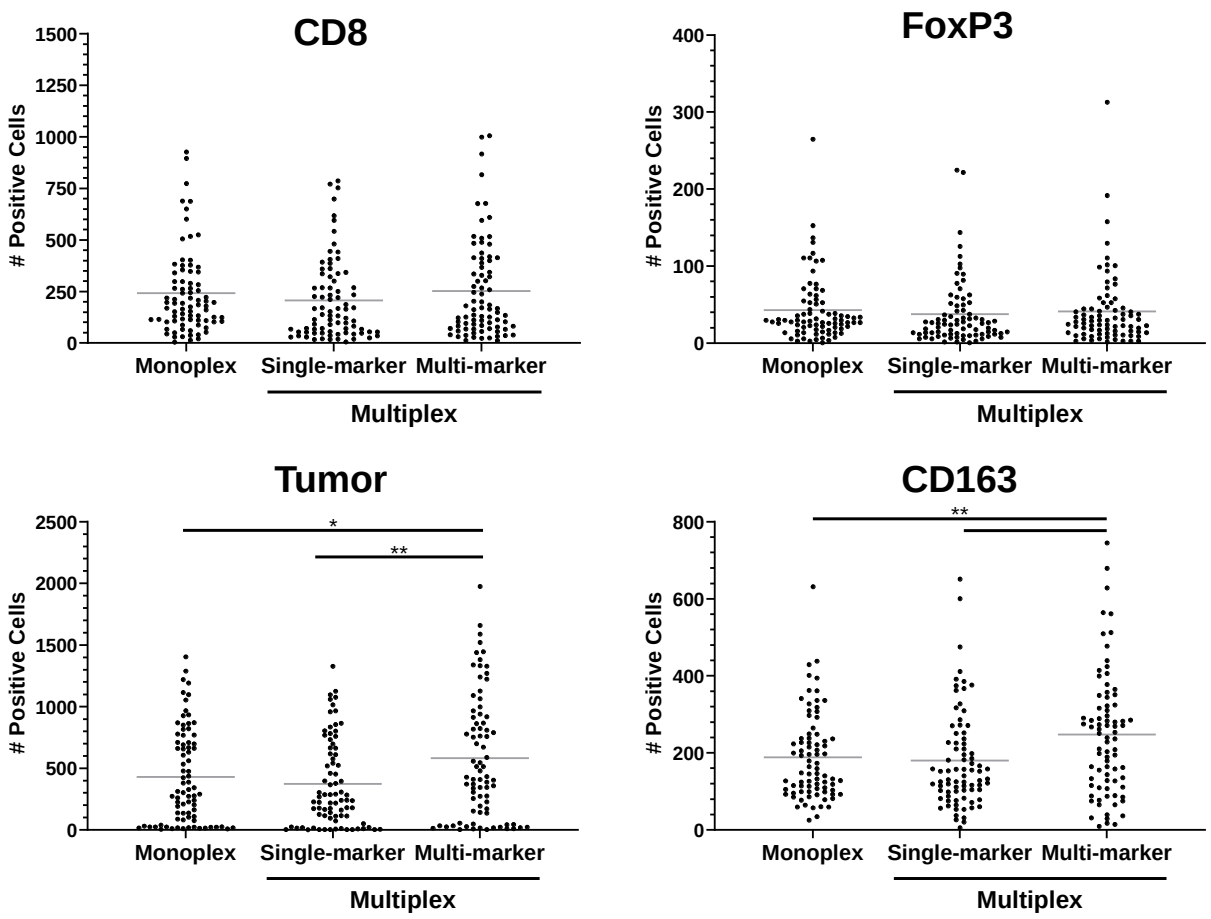


Figure S8. Image tiles generated using the 20% overlap approach are stitched to an absolute Cartesian coordinate system, creating a whole slide image that is accurate to a fraction of a pixel without loss of information. (A) Simple abutting of image tiles potentially contributes to a loss of reliable information for approximately 3-6% of cells. **(B)** Schematic visualization of how jumps in mechanical stage movement and inaccuracies in the underlying stitching algorithms accumulate in the x and y direction across a slide. Here the relative displacements in the x and y direction required to seamlessly stitch image tiles are denoted as dx and dy. We found that on average the cumulative shift across a whole slide contributes 20 mm error in both the x and y direction. **(C)** Shown here are the contours from uncorrected stitching, overlaid on images generated using the AstroPath approach. In this example, uncorrected whole slide stitching contributed up to an ~80 pixel shift – this equates to 40 mm or the diameter of 4 lymphocytes. Correcting such errors will be especially important when multiple microscopes, software analysis suites, and/or scans from different systems are used. Such directional, cumulative shifts could also contribute to inaccuracies in slide registration when Z-stacking images (i.e. overlaying a second slide image on top of the first image from the same specimen).

Figure S9


A
Multi-marker segmentation
Single-marker segmentation
B
Single-marker phenotyping
"Other"
CD8
CD163
Sox10
FoxP3

C



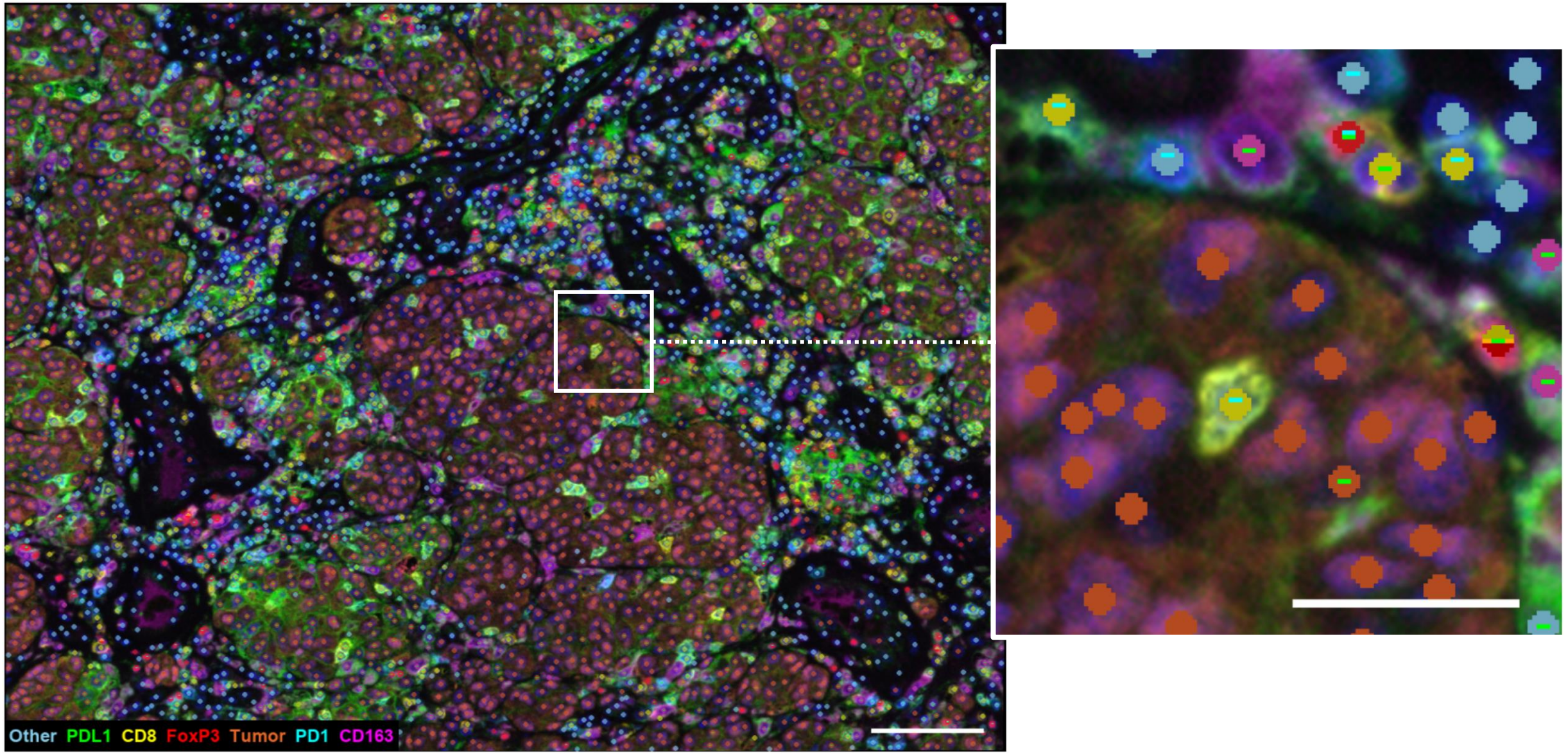
D

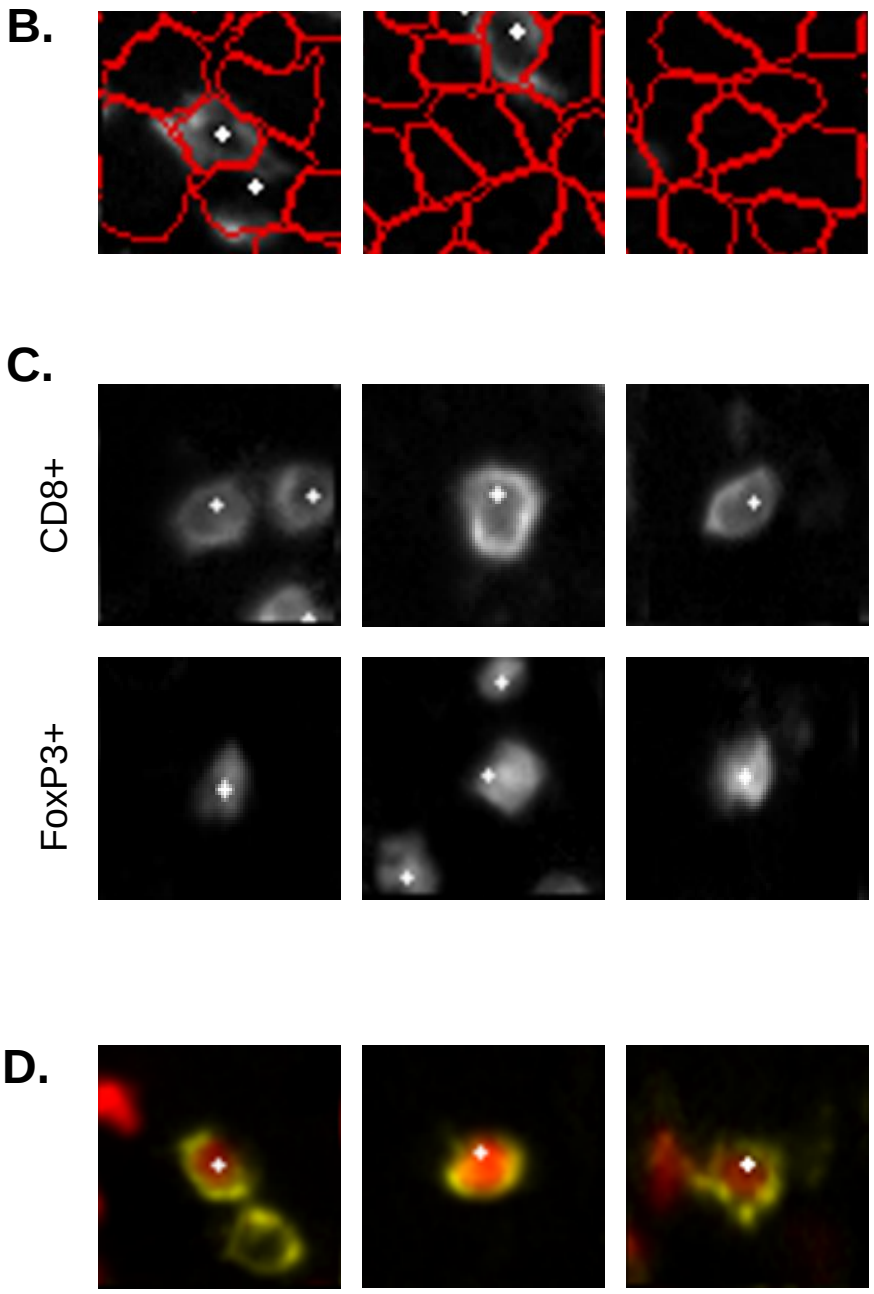
	Cell counts multi-marker approach	Cell counts single-marker approach	% difference
Tumor	920740	695911	-26.35
CD8	78805	81162	2.77
CD163	123609	96642	-36.42
FoxP3	25212	24118	-6.80

Figure S9. Single-marker phenotyping approach minimizes error in dataset due to over-segmentation of large cells (A) Representative images displaying improved cell segmentation using a single-marker approach (red lines = cell boundaries, * = over-split tumor nuclei). **(B) Top left**: Representative image of merged phenotype output following single-marker phenotyping, **Bottom and Right:** Corresponding output of each individual single-marker phenotype algorithm before merging. Scale bars are 25 μm. **(C)** Number of positive cells quantified by the single-marker approach reflects the 'gold standard', while the multi-marker approach overestimates the tumor and CD163+ cells. The 'gold standard' is defined as segmentation/phenotyping performed for each lineage marker on monoplex IF (i.e. individual stain). Data displayed represents analysis of 80 HPFs from 10 different melanoma specimens. * $p \leq 0.05$, ** $p \leq 0.01$ by unpaired t-test. **(D)** This systematic error was further characterized by testing the number of cells counted in CD8 hotspots from 46 specimens by the single-marker and multi-marker approaches. Percent differences between cell counts show the multi-marker approach leads to a 30% over counting of tumor and CD163 cells, compared to the single-marker approach.

Figure S10

A.





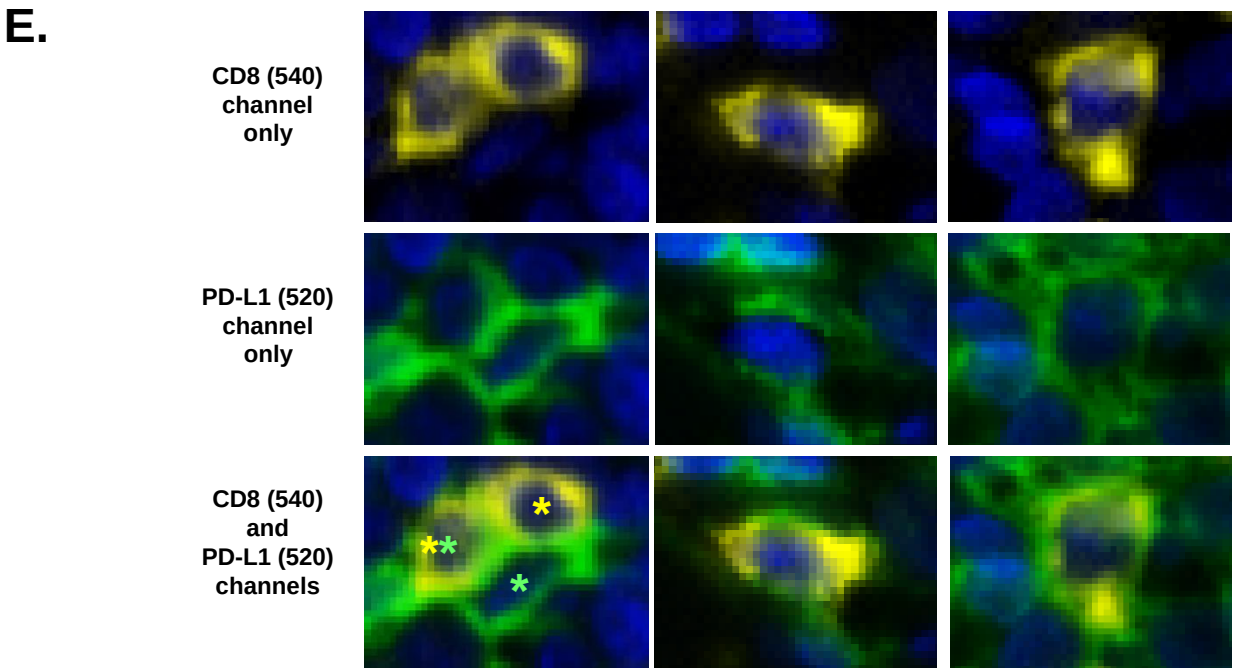


Figure S10. Representative output from algorithms that facilitate visual inspection of segmentation and phenotyping performance. (A) The first quality assurance viewer is a custom display that shows ~1250 cells per view. A colored dot is placed on each cell in the mIF image indicating the lineage. Additionally, a dash is placed over the cell if PD-L1 (green dash) and/or PD-1 (cyan dash) is expressed. 20 views per specimen were visually inspected, except for the rare cases with less tissue availability. Scale bar is 75μm and 25 μm for large and small image respectively. **(B)** The second custom quality assurance viewer facilitates inspection of up to 25 randomly selected positive and negative cells/stamps for each marker across all HPFs in a given specimen. The results of the segmentation algorithm are shown in red, and each cell that is positive for a given marker is labeled with a white "+". A minimum of 2000 cells displaying each marker was visually inspected per specimen using these stamps. **(C)** Additional representative QA/QC stamps shown without the overlying cell segmentation map. The "+" shows each cell that was called positive by the algorithm. **(D)** The QA/QC stamp viewer can also be used to visually inspect co-expression profiles of interest. Representative images of CD8+FoxP3+ cells are shown (FoxP3 in red; CD8 in yellow). An average of 200 CD8+FoxP3+ cells per specimen were visually inspected to verify this finding. **(E)** Shown here are three examples of CD8+ cells (yellow) that are PD-L1+ (green). For this example, these three images were obtained from three different patient specimens to show the generalizability of this finding. The top row shows the CD8 channel only; the middle row shows the PD-L1+ channel only; and the bottom row shows the CD8 and PD-L1 channels together. Specifically, the left panel in the bottom row shows a cell that is CD8+PD-L1- cell (single yellow asterisk), a cell that is CD8-PD-L1+ (single green asterisk), and a cell that is CD8+PD-L1+ (one yellow and one green asterisk). The DAPI is also displayed in this example in blue.

Figure S11

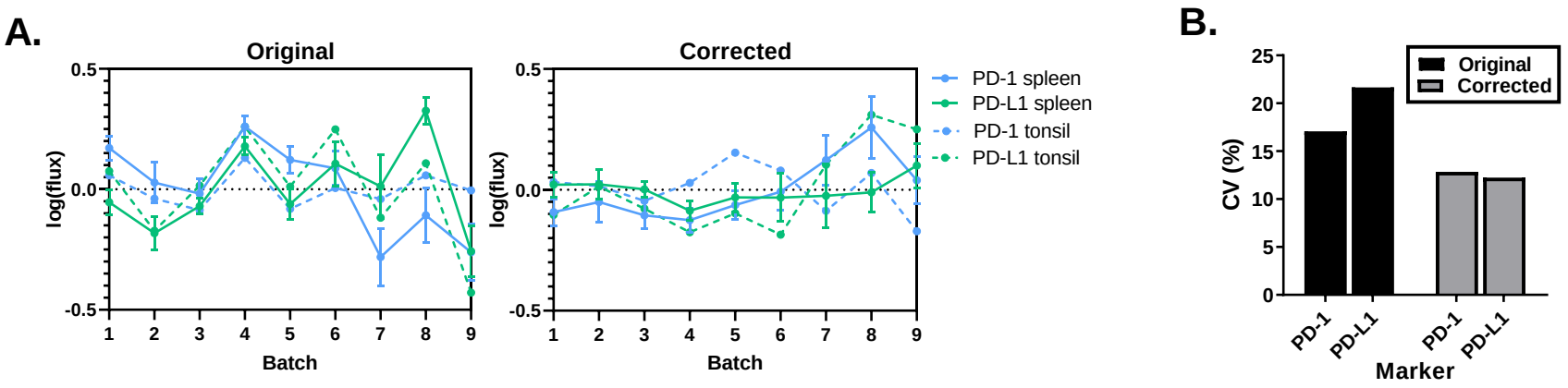


Figure S11. Accurate comparison of specimens stained at different times requires the correction of batch-to-batch variation. (A) Batch-to-batch variation was evident for PD-1 and PD-L1 expression intensities. It was corrected through normalization to a tissue microarray slide containing tonsil and spleen (n=3 each), which is run with each batch. **(B)** The percent coefficient of variation across the 9 batches included in this study was 17% for PD-1 and 22% for PD-L1, and was reduced by ~50% for both markers once normalized.

Figure S12

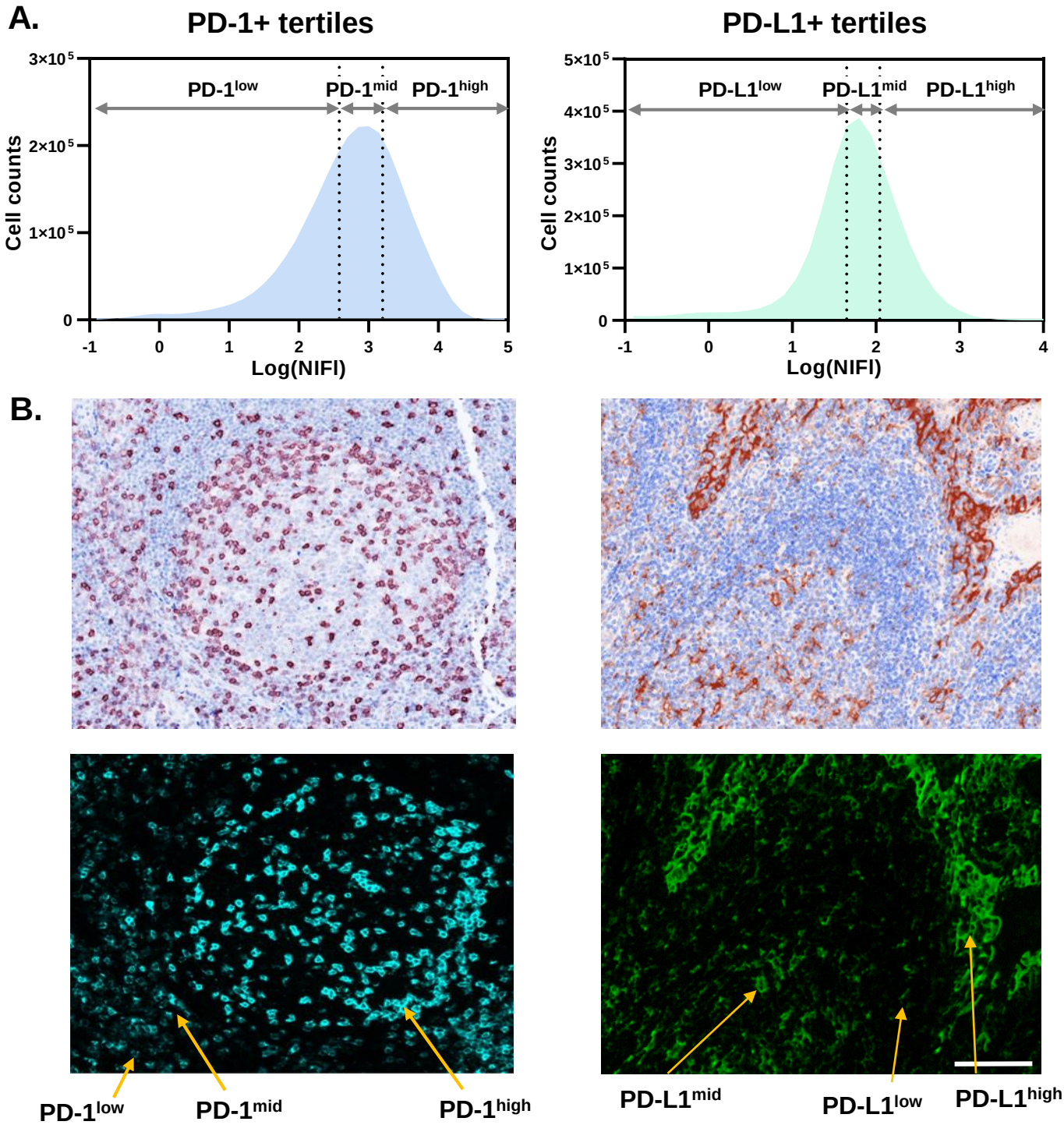


Figure S12. PD-(L)1^{low}, PD-(L)1^{mid} and PD-(L)1^{high} intensity levels. (A) Histogram showing PD-1 (left) and PD-L1 (right) intensity cut-offs that were defined by pooling all PD-1^{pos} or PD-L1^{pos} cells and dividing the population into tertiles (PD-1^{neg} and PD-L1^{neg} cells are not included with PD-1^{low} and PD-L1^{low}, respectively). **(B) Left:** Photomicrographs (brightfield on top and IF on bottom) show PD-1+ populations vary by specific regions in tonsil tissue. PD-1^{high} cells are predominantly located in germinal center T-cells in the light zone, while PD-1^{low} and PD-1^{mid} cells are found in the interfollicular zone. **Right:** Photomicrographs showing the location of PD-L1+ populations also vary by microanatomic location within tonsil. The tonsillar crypts show PD-L1^{high} cells. PD-L1^{mid} and PD-L1^{low} cells are observed in the germinal centers, and scattered PD-L1^{low} perifollicular cells may also be seen. For assay optimization of PD-1 and PD-L1 signal in the mIF assay, anatomic regions of low and mid expression were preferentially selected. Scale bar is 75 μm.

Figure S13

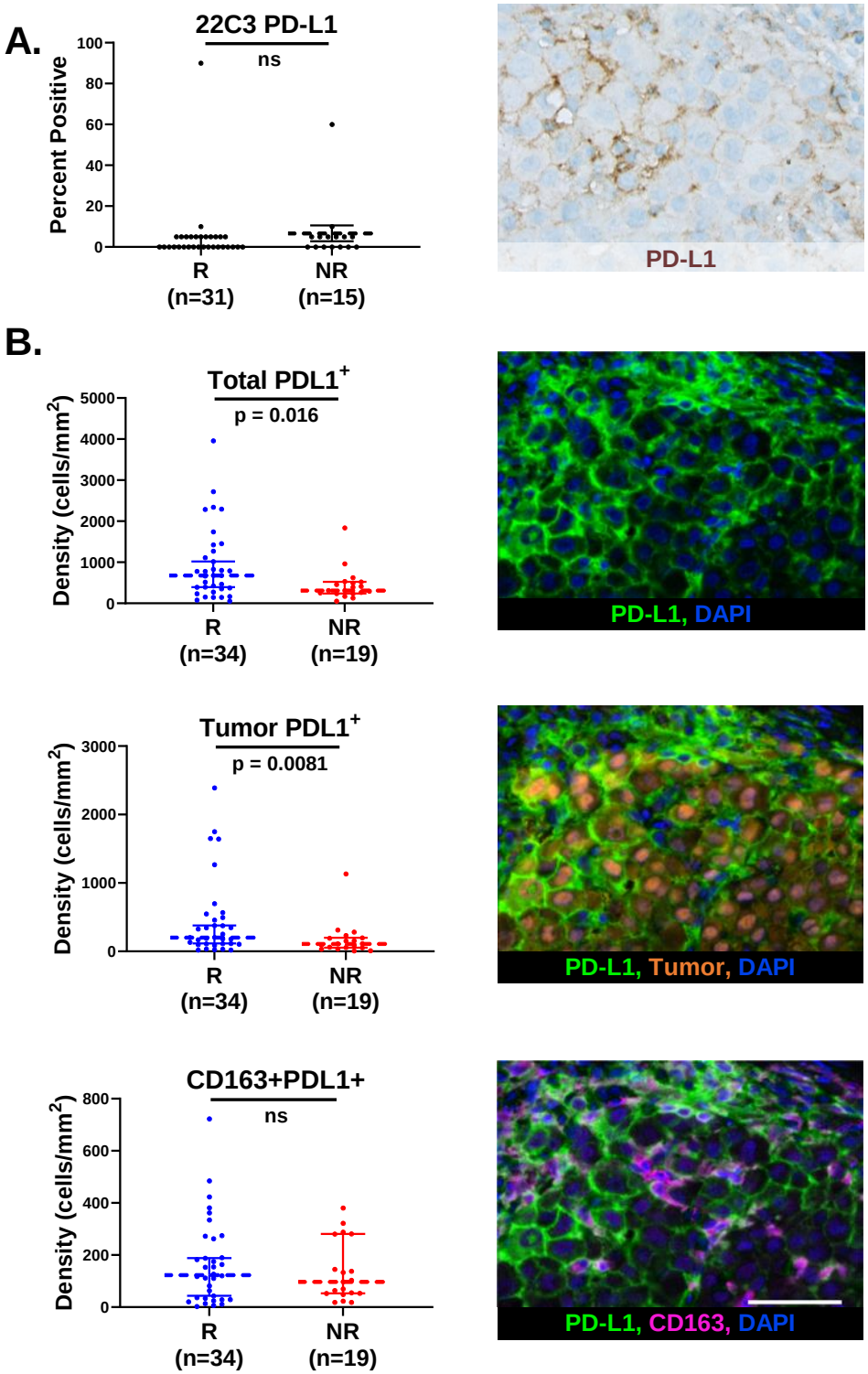


Figure S13. Densities of specific cell populations in responders vs. non responders across the entire TME. The mean tumor area analyzed among the 53 patients was 61 mm^2 (range 5 – 308 mm^2). **(A)** There was no significant difference in densities of PD-L1 positive cells between responders and non-responders to anti-PD-1 when scoring for % tumor cell expression using the commercially available chromogenic 22C3 IHC assay and interpreted by a pathologist using light microscopy. Representative photomicrograph of PD-L1 IHC shown on right. **(B)** Total and tumor cell PD-L1+ cell densities across the entire TME (whole slide analysis using 6-plex mIF assay on the AstroPath platform) were associated with response while no significant associations were seen for CD163+PD-L1+ cell densities. Median +/- 95% CI, one-tailed Mann-Whitney. Representative photomicrographs shown in right column with PD-L1 on any cell type by mIF assay (top row), PD-L1 and tumor (middle row), and PD-L1 and CD163 (bottom row). Scale bar is 50 μm.

Figure S14

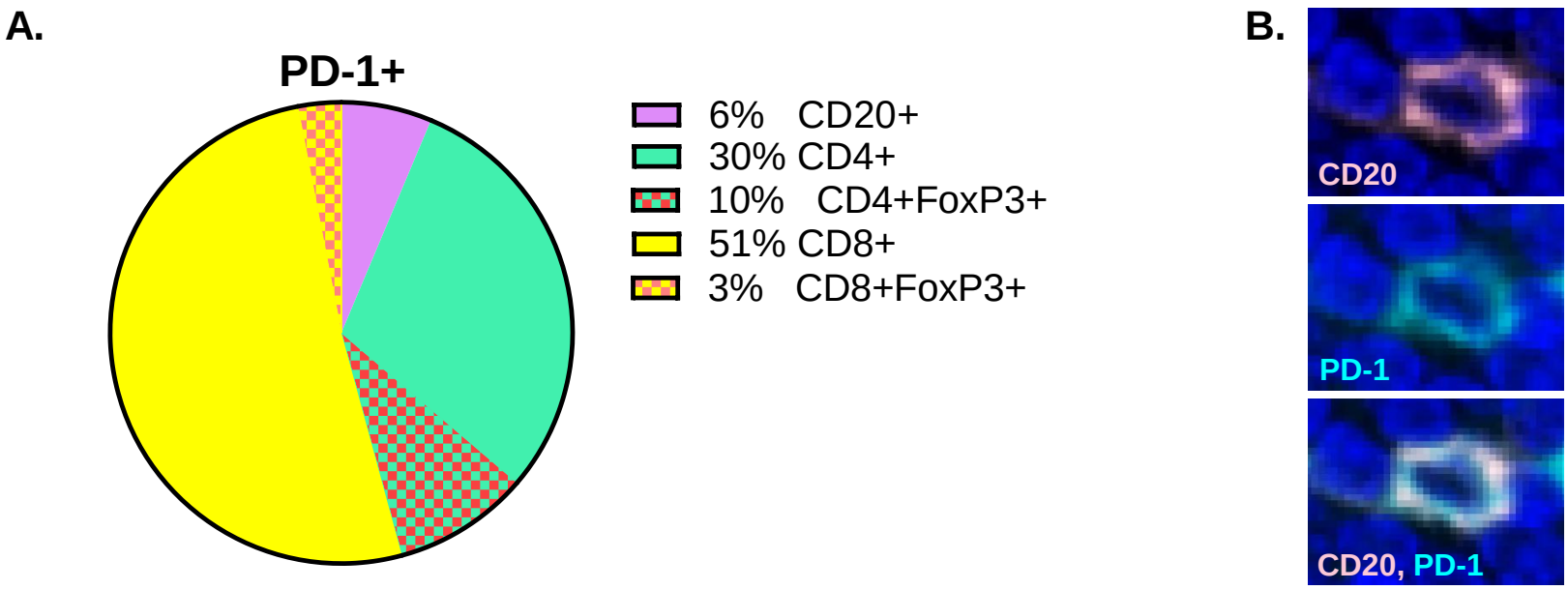


Figure S14. PD-1 is displayed by multiple lymphocyte subsets within the melanoma TME. (A) 94 archival melanoma specimens in TMA format were stained using a mIF assay for PD-1, CD8, CD4, CD20, FoxP3, and tumor (Sox10/S100), see **Materials and Methods** for additional detail. Pie chart shows the PD-1 expression proportion within the melanoma TME by cell type. CD8+ cells contributed the majority of PD-1 to the melanoma TME. Of the non-CD8+ cells contributing PD-1, 86% labeled as CD4+ (65% of conventional CD4+ cells and 21% of CD4+FoxP3+ cells). **(B)** Photomicrograph of a representative CD20+PD-1+ cell.

Figure S15

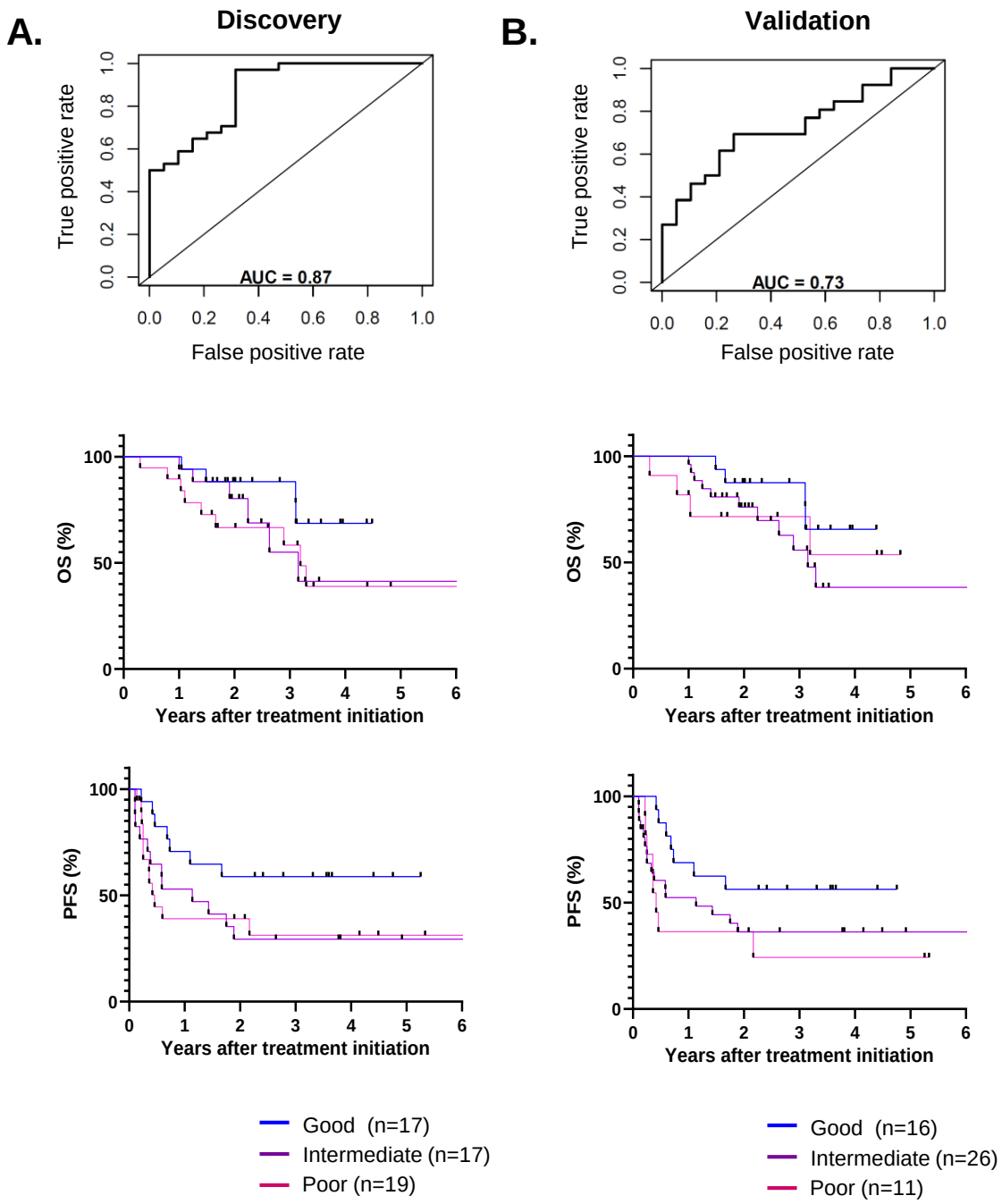


Figure S15. Analysis of the whole TME (100% sampling) was not as effective at stratifying patients as 30% sampling, when similar strategies were applied. At 100% TME sampling, the features with AUCs having p-values <0.05 after correction for multiple tests were identified (positive features: CD8+PDL1^{low}, CD8+FoxP3+, CD8+FoxP3+PD-1^{low}, CD8+FoxP3+PD-1^{mid}, tumor PD-L1^{low}, and negative features: CD163+PD-L1^{neg} and CD163+ cells), (See **Table S4**). These features were used to generate combinatorial ROC curves and Kaplan-Meier curves for the **(A)** Discovery cohort, as well as **(B)** a second, independent Validation cohort.

Figure S16

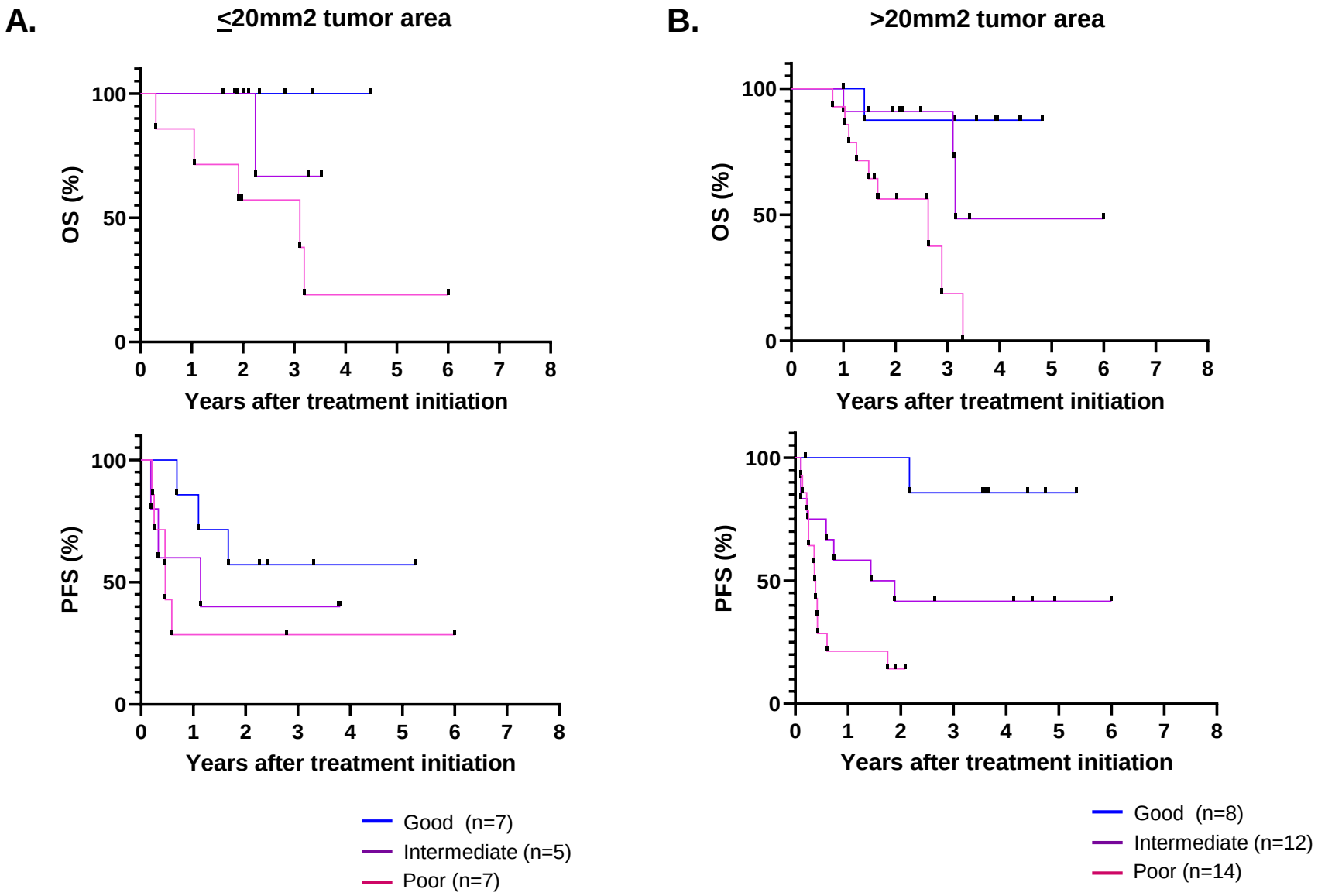


Figure S16. TMEs defined by specific cell types and association with long-term survival by Kaplan-Meier analysis for smaller specimens. The minimum tumor area for inclusion in the study was 5 mm^2. Patients where **(A)** ≤20mm^2 and **(B)** >20mm^2 tumor area was present on the slide are separated into good, intermediate and poor prognosis using scoring rules defined for **Figure S6B**. 20 mm^2 in surface area was chosen because it represents the size of 3 core biopsies (each 1 mm x 15 mm in size) with ~50% tumor in each core. The same trends were observed between the two groups, though larger cohorts will be required to set standards of minimum specimen adequacy.

Figure S17

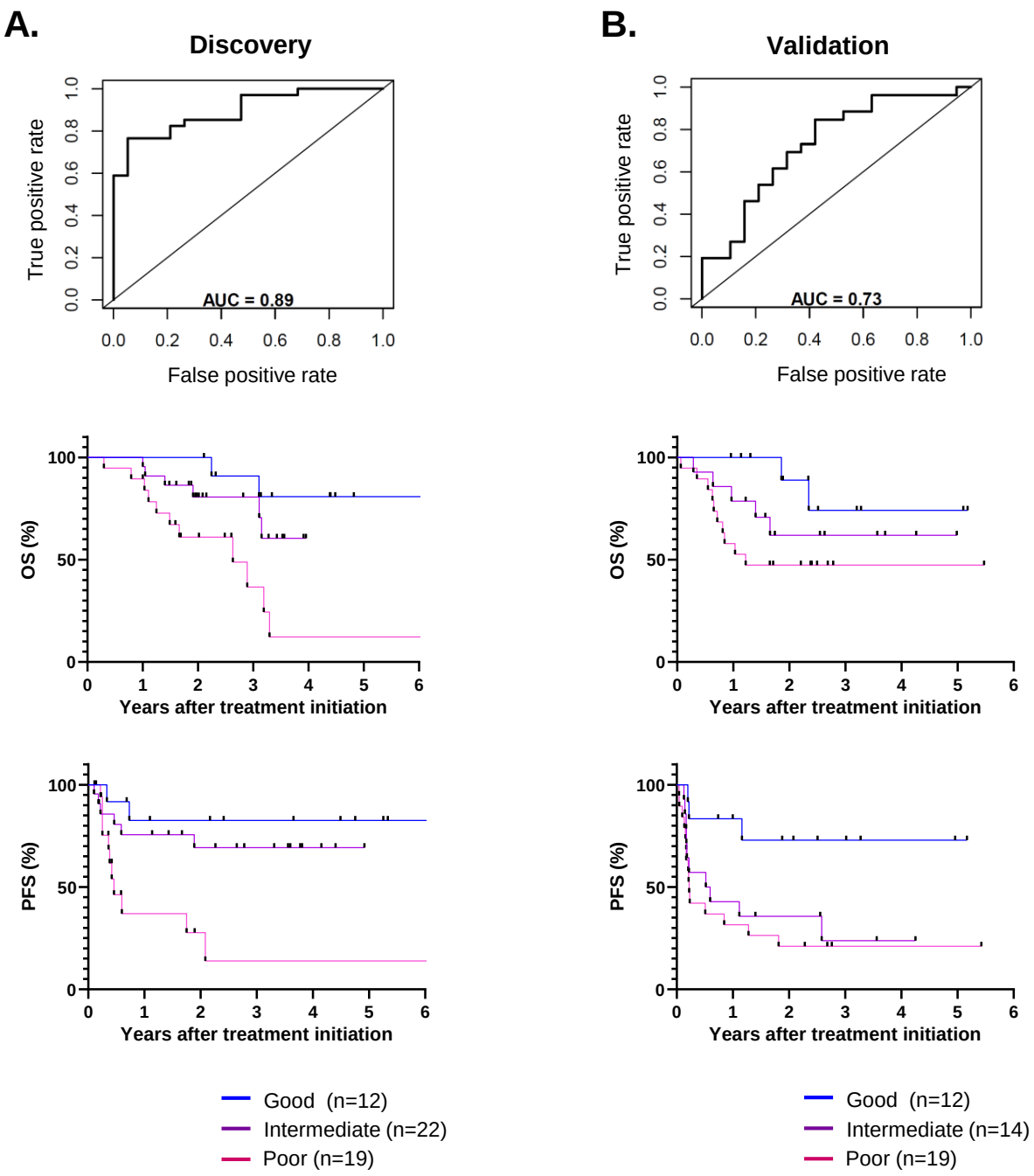


Figure S17. Reduction of mIF assay from 6-plex to 4-plex is not as efficacious for predicting objective response and stratifying overall survival. In the index 6-plex assay (**Figure 6**), CD8+ subsets were used to predict patients with good vs. intermediate long-term outcomes. Here, we tested whether total CD8+ cell densities alone could be used for this distinction, potentially reducing the number of requisite markers from 6 to 4 (CD8, CD163, PD-L1, Sox10/S100). These four features were used to generate combinatorial ROC curves and Kaplan-Meier curves for the **(A)** discovery cohort and **(B)** validation cohort.

Figure S18

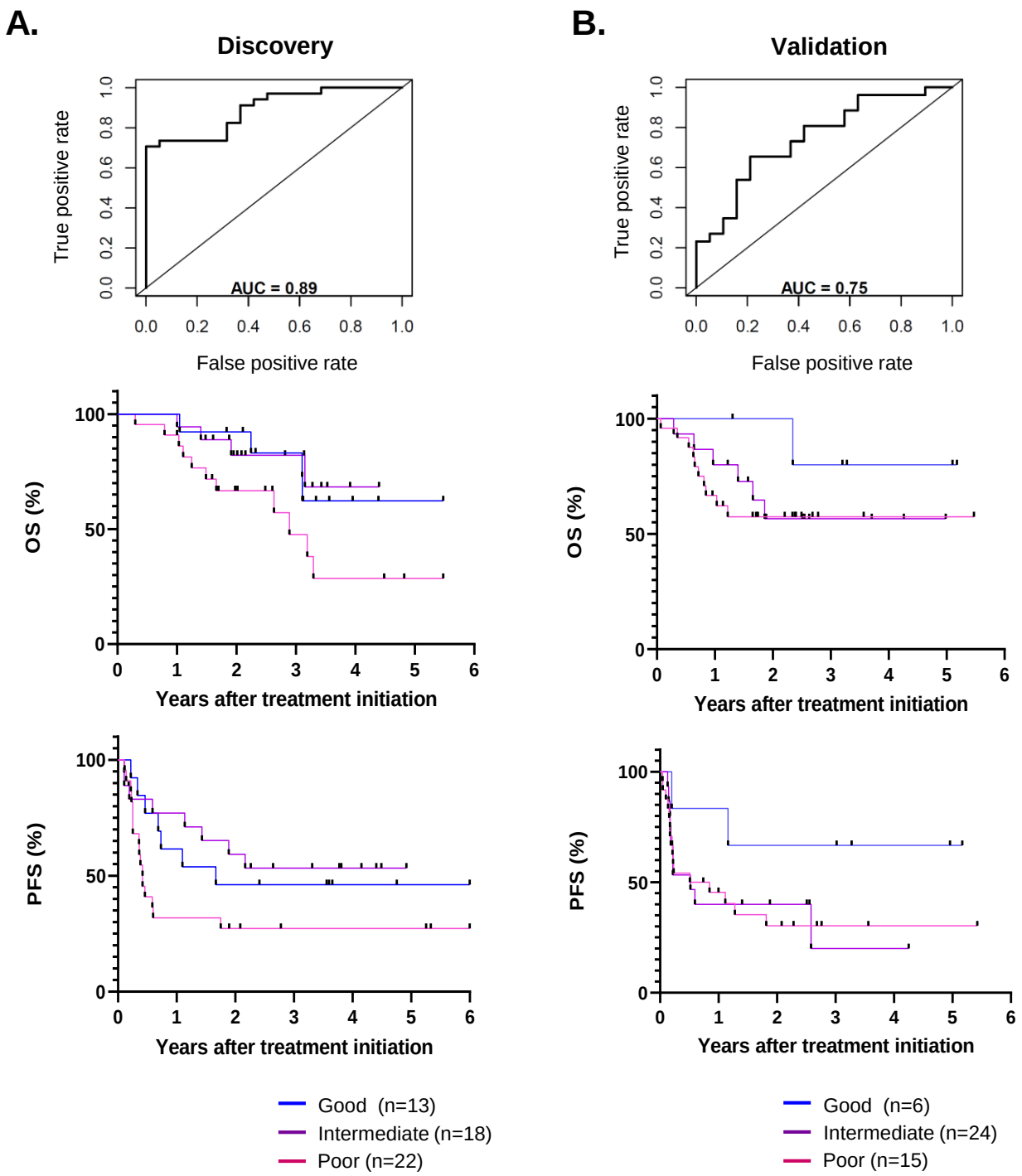


Figure S18. 6-plex mIF assay without assessments of PD-1 and PD-L1 expression intensity for predicting objective response and stratifying overall survival. In the index 6-plex assay (**Figure 6**), PD-1(+) and PD-L1(+) low vs. mid vs. high subsets were used to predict patient outcomes. Here, we tested whether total PD-1(+) and PD-L1(+) cell densities, without grouping by intensity of positive expression, could be used for this distinction. These features were used to generate combinatorial ROC curves and Kaplan-Meier curves for the **(A)** discovery cohort and **(B)** validation cohort, which were less effective than when intensities were assessed.

Table S1. Contributions/improvements compared to commercial platform for mIF panel optimization

Task (in order of operation*)	Commercial kit / manufacturer recommendation	AstroPath recommendation	Example of contribution / improvement beyond commercial platform
Pairing marker to fluorophore	Exposure time 50 – 250 ms. or intensity 5 to 30 after primary antibody optimization.	Consideration given to (1) fluorophore staining index, (2) target protein expression intensity, and (3) subcellular location (nucleus vs. membrane).	Facilitates balancing of signals through pairing stronger fluors with weaker markers and vice versa. Also, facilitates mitigation of any potential residual bleed-through at later stages of panel development by capitalizing on differential subcellular localization, **Figure S3**.
Selection of secondary antibody	Secondary antibody provided with commercial kit.	Select markers require replacement of commercial kit secondary antibody with an alternative to meet 'gold-standard' chromogenic IHC.	Capture populations with lower levels of marker expression (e.g. PD-$1^{low/mid}$), improving sensitivity by 50%, **Figures 2** and **S12**.
Primary antibody optimization	See "Pairing of marker to fluorophore" above.	Titration of primary antibody to optimize the signal to noise ratio (SNR), **Figure S4.**	Improved sensitivity and specificity through optimized SNR e.g. Sox10 had a 3 fold higher SNR when using our approach vs. manufacturer recommendation (optimal concentration resulted in intensity counts of up to 100).
TSA optimization	Recommended dilution of 1:100 (recent update).	Titration of TSA to identify concentration required to reduce potential bleed-trough and steric hindrance (with excess TSA) without signal loss (with insufficient TSA).	Improved specificity through reduced false positive signal (bleed-through) from adjacent channel, e.g. reduced 4-fold (12% to 3%) from 570 (FoxP3) to 540 (CD8) channel. The remaining 3% is further reduced during image analysis by capitalizing on differential subcellular localization (see pairing marker to fluorophore above).
Validation of final mIF panel against chromogenic IHC	None provided.	mIF validated against chromogenic IHC.**	mIF panel sensitivity and specificity is comparable to gold-standard, **Figure 2** and **S5**.

*The defined step-wise order of operations provided here is a contribution beyond manufacturer recommendations.

**This may also be achieved through two-step validation: (1) monoplex IF to IHC and (2) mIF to monoplex IF.

Table S2. **Image acquisition and processing workflow by task, software, and description.**

	Task	Commercial (software) or custom (GitHub code name)	Description	Contributions / improvements beyond commercial platform
Image acquisition and processing of individual HPFs	Modified image acquisition protocol	Custom (Phenochart.config*)	Change settings in the Phenochart software ROI functionality to generate 20% HPF overlap (qpTIFF output).	Facilitates image corrections and whole slide stitching, including accurately mapping the 3-6% of cells found at HPF edges, **Figures S6-S8.**
	Image acquisition	Commercial (Vectra 3.0 software)	Image the slides using the Vectra platform (im3 file output).	N/A
	Spectral unmixing of image	Commercial (inForm)	Deconvolution of spectral signatures for the seven detected colors (6 markers+DAPI) and removal of autofluorescence (component TIFF output).	N/A
	Image correction/ processing	Custom (flatw*)	Correct the images for illumination variation and lens distortion effects.	This step reduces systematic error in the HPFs themselves, *e.g.* illumination variation reduced 9x (11.2% → 1.2%), **Figure S7**.
Lineage assignment/cell classification and normalized PD-1/L1 expression levels per individual cell	Segmentation and phenotyping	Commercial (with modified usage)	Commercial cell segmentation and phenotyping routine is run multiple times (segmentation for larger vs. smaller cells, and phenotyping once for each marker), i.e. a "multipass approach" for each.	When run as a single pass, the cell segmentation/ phenotyping algorithm overestimates the number of large cells (tumor cells and macrophages) by 25%, **Figure S9.**
	Merge multipass data	Custom (MaSS*)	Outputs from the cell segmentation/phenotyping routine for each marker are merged into a single data set.	Multipass phenotyping allows for training for each marker individually. Individual markers are then combined at this stage, simplifying training algorithms and facilitating the identification of rare cell phenotypes.
	QA/QC phenotyping	Custom (Create image QA/QC, segmaps*)	Shows images to visually inspect performance of multipass phenotyping and merging algorithms.	Commercial platform cannot be used to visually inspect results of multipass approach. Also, functionality showing individually, randomly selected positive and negative cells for each marker is provided, **Figure S10.**
	Batch Normalization	Custom (calib*)	Reduce potential batch-to-batch intensity variation by normalizing to control tissues.	Average batch-to-batch variation for PD-1 and PD-L1 expression intensity was ~20%, and this was reduced by half through normalization, **Figure S11.**

Image handling and individual cell mapping in whole-slide format	Image stitching and mapping to absolute coordinate system	Custom (align, shift*)	Seamless stitching of image tiles into a whole slide. Scaling all inputs (multiple scanners, images, and annotations) to an absolute Cartesian coordinate system.**	This step corrects the combination of different errors generated when re-assembling numerous HPFs into a single image. (~5% loss of cells around perimeter of HPFs, and 40 μm cumulative shift with regard to relative cell position from the left to right edge of a whole slide).
	Image annotation	Commercial (Halo)	Pathologist manual annotation of tumor-stromal boundary and removal of tissue artifacts (tears, folds, etc) (qpTIFF output).	N/A
	Image annotation overlay	Custom (annowarp*)	Pathologist manual annotation of tumor-stromal boundary, etc is applied to whole slide, stitched image.	Annotations are stored in the database, lending ease to spatial statistics and visualizations. This makes data consistency easier and is of particular interest for anticipated tumor-immune atlas generation and use.**
	Geometric conversions	Custom (geomcell, geom*)	The different geometric regions are created both in vector formats and some as raster images, and we create a unified geographic information system (GIS) representation that is loadable into the database.	Detailed representations of millions of geometries, enabling large-scale spatial analysis with distinct shapes for each cell.
	Image tiling	Custom/open source (Vips*)	After building the whole slide images for each layer, a set of hierarchical tiles for visualization is created.	Enables standard geospatial visualization packages for quick zoom and pan of the images (openlayers.org).
Loading the Database	Database load	Custom (dbload, zoomload, mergedb*)	A two-phase load of the database is performed. First, a database is made for each slide and the data undergoes an extensive validation. Second, all individual slide databases are merged into a cohort database.	No commercial equivalent.

* The computer code has been archived with Zenodo *(48)*, and is maintained at https://github.com/AstroPathJHU/AstroPathPipeline.

** Additional measures that ensure data consistency across different imaging platforms and image analysis software include the implementation of a universal scaling system (2 pixels = 1 μm).

Table S3. AUC for each feature at 30% hot spot sampling, ranked by association with response/non-response to anti-PD-1.

Feature	AUC	Uncorrected p value	Benjamini-Hochberg corrected p value	Association with response
CD163_PDL1_neg	0.751	0.001	0.036	-
Tumor_PDL1_neg	0.743	0.002	0.036	-
CD8FoxP3_PD1_mid	0.725	0.004	0.036	+
Tumor_PDL1_low	0.721	0.004	0.036	+
CD8FoxP3_PD1_low	0.717	0.005	0.036	+
CD8_PDL1_low	0.711	0.006	0.036	+
CD8FoxP3	0.710	0.006	0.036	+
CD8FoxP3_PDL1_low	0.695	0.009	0.045	+
CD8FoxP3_PDL1_neg	0.693	0.011	0.045	+
Tumor	0.692	0.011	0.045	-
CD8_PD1_neg	0.681	0.015	0.057	+
CD163	0.675	0.018	0.063	-
CD8FoxP3_PD1_neg	0.672	0.020	0.064	+
CD8	0.661	0.027	0.075	+
CD8FoxP3_PDL1_mid	0.661	0.027	0.075	+
CD8_PD1_low	0.656	0.031	0.077	+
CD8_PDL1_mid	0.653	0.034	0.077	+
CD8FoxP3_PD1_high	0.653	0.034	0.077	+
CD8_PDL1_neg	0.644	0.043	0.093	+
CD8_PD1_mid	0.635	0.054	0.106	+
CD8FoxP3_PDL1_high	0.634	0.054	0.106	+
CD8_PDL1_high	0.602	0.112	0.209	+
FoxP3_PDL1_low	0.596	0.127	0.226	+
Other_PD1_high	0.593	0.135	0.230	+
Other_PD1_mid	0.573	0.194	0.314	+
Tumor_PDL1_mid	0.571	0.199	0.314	+
CD8_PD1_high	0.567	0.215	0.315	+
Other_PD1_low	0.567	0.215	0.315	+
Tumor_PDL1_high	0.563	0.226	0.320	+
FoxP3_PD1_neg	0.560	0.238	0.325	+
FoxP3_PDL1_mid	0.554	0.261	0.342	+
FoxP3	0.553	0.267	0.342	+
FoxP3_PD1_low	0.548	0.286	0.355	+
FoxP3_PD1_high	0.533	0.352	0.405	+
CD163_PDL1_low	0.533	0.352	0.405	+
FoxP3_PD1_mid	0.529	0.366	0.405	+
Other_PD1_neg	0.529	0.366	0.405	+
FoxP3_PDL1_neg	0.525	0.387	0.407	+
FoxP3_PDL1_high	0.525	0.387	0.407	+
CD163_PDL1_high	0.508	0.467	0.478	+
CD163_PDL1_mid	0.503	0.489	0.489	+

Table S4. AUC for each feature at whole slide (100%) sampling, ranked by association with response/non-response to anti-PD-1.

Feature	AUC	Uncorrected p value	Benjamini-Hochberg corrected p value	Association with response
CD163_PDL1_neg	0.759	0.001	0.041	-
Tumor_PDL1_neg	0.726	0.003	0.043	-
Tumor_PDL1_low	0.721	0.004	0.043	+
CD163	0.715	0.005	0.043	-
CD8FoxP3_PD1_mid	0.709	0.006	0.043	+
CD8_PDL1_low	0.709	0.006	0.043	+
CD8FoxP3	0.699	0.009	0.049	+
CD8FoxP3_PD1_low	0.697	0.009	0.049	+
CD8FoxP3_PDL1_low	0.688	0.012	0.054	+
Tumor	0.681	0.015	0.061	-
CD8FoxP3_PDL1_neg	0.679	0.016	0.061	+
CD8_PD1_neg	0.663	0.026	0.090	+
CD8FoxP3_PDL1_mid	0.647	0.039	0.108	+
CD8FoxP3_PD1_high	0.646	0.041	0.108	+
CD8_PDL1_mid	0.646	0.041	0.108	+
CD8FoxP3_PD1_neg	0.644	0.043	0.108	+
CD8	0.642	0.045	0.108	+
CD8FoxP3_PDL1_high	0.634	0.054	0.122	+
CD8_PD1_low	0.630	0.061	0.131	+
CD8_PDL1_neg	0.627	0.065	0.134	+
CD8_PD1_mid	0.616	0.084	0.163	+
CD8_PDL1_high	0.587	0.152	0.283	+
Tumor_PDL1_mid	0.582	0.165	0.294	+
Tumor_PDL1_high	0.557	0.249	0.418	+
FoxP3_PDL1_low	0.556	0.255	0.418	+
CD8_PD1_high	0.551	0.273	0.424	+
Other_PD1_high	0.550	0.280	0.424	+
Other_PD1_mid	0.543	0.305	0.435	+
FoxP3_PDL1_mid	0.542	0.312	0.435	+
CD163_PDL1_low	0.540	0.318	0.435	+
Other_PD1_low	0.529	0.366	0.484	+
FoxP3_PD1_neg	0.517	0.423	0.520	+
FoxP3_PD1_low	0.517	0.423	0.520	+
FoxP3_PDL1_high	0.512	0.445	0.520	+
FoxP3	0.511	0.452	0.520	+
FoxP3_PD1_mid	0.509	0.459	0.520	+
CD163_PDL1_high	0.506	0.474	0.520	+
CD163_PDL1_mid	0.505	0.482	0.520	+
FoxP3_PDL1_neg	0.491	0.548	0.576	+
FoxP3_PD1_high	0.485	0.577	0.592	+
Other_PD1_neg	0.472	0.634	0.634	+

Table S5. Clinicopathologic features and outcomes for individual patients in the discovery cohort.

Sample ID	Response	Treatment	Age At Collection	Sex	Anatomic Site	Tumor area on slide (mm^2)	OS (years)	OS Status	PFS (years)	PFS Status	Survival Group	Prior Immunotherapy	BRAF+/-MEK Rx
1	Responder	Pembrolizumab	64	Female	Skin_soft-tissue	5.811	1.660	Dead	0.600	Progressor	Poor	-	-
2	Responder	Pembrolizumab	59	Male	Lymph-node	5.560	3.953	Alive	3.599	Non-Progressor	Good	-	-
3	Responder	Nivolumab	74	Female	Skin_soft-tissue	113.011	4.397	Alive	2.167	Progressor	Good	-	-
4	Responder	Ipilimumab + nivolumab	48	Male	Skin_soft-tissue	25.608	3.110	Dead	0.462	Progressor	Good	-	-
5	Non-Responder	Ipilimumab + nivolumab	78	Male	Skin_soft-tissue	95.243	3.296	Dead	0.250	Progressor	Poor	-	-
6	Responder	Ipilimumab + nivolumab	64	Female	Skin_soft-tissue	15.372	3.529	Alive	3.809	Non-Progressor	Int	-	-
7	Responder	Nivolumab	80	Male	Skin_soft-tissue	15.102	3.912	Alive	4.406	Non-Progressor	Good	-	-
8	Non-Responder	Nivolumab	67	Male	Viscera	57.953	2.019	Alive	0.137	Progressor	Poor	-	-
9	Responder	Pembrolizumab	69	Male	Skin_soft-tissue	121.292	3.562	Alive	3.560	Non-Progressor	Good	-	-
10	Responder	Ipilimumab + nivolumab	68	Male	Viscera	11.493	2.151	Alive	4.917	Non-Progressor	Int	-	-
11	Responder	Ipilimumab + nivolumab	57	Male	Mucosal	55.978	3.137	Alive	4.492	Non-Progressor	Int	-	-
12	Responder	Pembrolizumab	76	Male	Viscera	78.560	2.485	Alive	0.586	Progressor	Poor	-	-
13	Responder	Nivolumab	61	Male	Skin_soft-tissue	38.366	9.411	Alive	9.408	Non-Progressor	Poor	-	-
14	Responder	Nivolumab	49	Male	Viscera	32.767	9.701	Alive	10.199	Non-Progressor	Int	-	-
15	Responder	Ipilimumab + nivolumab	63	Female	Lymph-node	55.080	3.340	Alive	1.099	Progressor	Good	-	-
16	Responder	Pembrolizumab	92	Male	Lymph-node	25.919	3.121	Alive	3.656	Non-Progressor	Good	-	-
17	Responder	Ipilimumab + nivolumab	71	Male	Skin_soft-tissue	206.252	4.389	Alive	4.755	Non-Progressor	Good	-	-
18	Non-Responder	Ipilimumab + nivolumab	57	Female	Skin_soft-tissue	41.004	1.104	Dead	0.250	Progressor	Poor	-	-
19	Responder	Pembrolizumab	67	Male	Brain	309.440	3.279	Alive	3.774	Non-Progressor	Good	-	-
20	Responder	Pembrolizumab	68	Female	Skin_soft-tissue	60.048	2.244	Dead	0.333	Progressor	Int	-	-
21	Non-Responder	Pembrolizumab	60	Female	Skin_soft-tissue	14.464	1.047	Dead	0.218	Progressor	Good	-	-
22	Non-Responder	Ipilimumab + nivolumab	66	Male	Mucosal	197.926	1.030	Dead	0.363	Progressor	Poor	-	-
23	Responder	Pembrolizumab	63	Female	Skin_soft-tissue	13.209	1.877	Alive	1.139	Progressor	Int	-	-
24	Responder	Ipilimumab + nivolumab	59	Female	Skin_soft-tissue	31.925	2.014	Alive	2.264	Non-Progressor	Good	-	-
25	Non-Responder	Pembrolizumab	76	Male	Lymph-node	41.026	1.003	Dead	0.105	Progressor	Int	-	-
26	Responder	Pembrolizumab	82	Male	Skin_soft-tissue	94.481	1.403	Dead	0.183	Non-Progressor	Int	-	-
27	Responder	Nivolumab	74	Female	Skin_soft-tissue	14.075	2.321	Alive	0.685	Progressor	Good	-	-
28	Responder	Pembrolizumab	53	Male	Skin_soft-tissue	35.623	2.110	Alive	2.411	Non-Progressor	Good	-	-
29	Non-Responder	Pembrolizumab	83	Female	Lymph-node	13.081	1.608	Alive	0.191	Progressor	Int	-	-
30	Non-Responder	Pembrolizumab	67	Female	Mucosal	7.997	1.688	Alive	0.110	Progressor	Poor	-	-
31	Responder	Ipilimumab + nivolumab	38	Female	Skin_soft-tissue	14.320	4.485	Alive	5.255	Non-Progressor	Good	-	-
32	Non-Responder	Pembrolizumab	91	Female	Lymph-node	217.077	1.249	Dead	0.383	Progressor	Poor	-	-
33	Responder	Nivolumab	76	Male	Skin_soft-tissue	202.748	2.819	Alive	3.310	Non-Progressor	Good	-	-
34	Responder	Nivolumab	53	Female	Lymph-node	7.961	3.151	Dead	1.887	Progressor	Int	-	-
35	Responder	Pembrolizumab	73	Male	Skin_soft-tissue	15.002	1.978	Alive	2.778	Non-Progressor	Int	-	-
36	Non-Responder	Nivolumab	71	Male	Brain	37.062	3.192	Dead	0.460	Progressor	Poor	IL-2, Ipilimumab	-
37	Responder	Ipilimumab + nivolumab	73	Female	Mucosal	309.708	3.427	Alive	4.151	Non-Progressor	Int	-	-
38	Non-Responder	Pembrolizumab	92	Female	Mucosal	38.304	1.482	Alive	0.106	Progressor	Int	-	-
39	Responder	Ipilimumab + nivolumab	72	Male	Skin_soft-tissue	151.159	2.893	Dead	2.086	Non-Progressor	Poor	-	-
40	Responder	Ipilimumab + nivolumab	64	Male	Skin_soft-tissue	74.266	4.822	Alive	5.336	Non-Progressor	Good	-	-
41	Responder	Ipilimumab + nivolumab	55	Female	Lymph-node	36.876	3.104	Dead	0.733	Progressor	Int	-	-
42	Responder	Ipilimumab + nivolumab	50	Male	Skin_soft-tissue	102.751	1.951	Alive	2.645	Non-Progressor	Int	-	-
43	Non-Responder	Nivolumab	67	Female	Skin_soft-tissue	17.739	2.605	Alive	0.417	Progressor	Poor	-	Yes
44	Non-Responder	Pembrolizumab	33	Male	Skin_soft-tissue	23.484	2.630	Dead	1.753	Progressor	Poor	-	-
45	Non-Responder	Nivolumab	68	Female	Skin_soft-tissue	187.446	1.000	Alive	0.228	Progressor	Poor	-	-
46	Non-Responder	Nivolumab	50	Male	Lymph-node	5.472	1.488	Dead	0.419	Progressor	Poor	-	-
47	Non-Responder	Nivolumab	63	Male	Skin_soft-tissue	39.769	1.699	Alive	0.357	Progressor	Poor	-	-
48	Non-Responder	Ipilimumab + nivolumab	35	Female	Mucosal	6.869	0.792	Dead	0.218	Progressor	Poor	-	-
49	Responder	Ipilimumab + nivolumab	51	Female	Skin_soft-tissue	13.667	2.085	Alive	1.433	Progressor	Int	-	-
50	Non-Responder	Pembrolizumab	71	Male	Skin_soft-tissue	52.684	1.912	Dead	0.589	Progressor	Int	-	-
51	Non-Responder	Ipilimumab + nivolumab	81	Female	Mucosal	43.078	0.299	Dead	0.250	Progressor	Poor	-	-
52	Responder	Pembrolizumab	51	Female	Skin_soft-tissue	8.732	1.836	Alive	1.669	Progressor	Good	-	-
53	Responder	Pembrolizumab	75	Female	Skin_soft-tissue	75.748	1.592	Alive	1.895	Non-Progressor	Poor	-	-

Table S6. Clinicopathologic features and outcomes for individual patients in the validation cohort.

Sample ID	Response	Treatment	Age At Collection	Sex	Tumor area on slide (mm^2)	OS (years)	OS Status	PFS (years)	PFS Status	Survival Group	Prior Immunotherapy	BRAF+/-MEK Rx
1	Non-Responder	Ipilimumab + nivolumab	72	Male	121.292	1.712	Alive	0.134	Progressor	Poor	-	Yes
2	Responder	Pembrolizumab	67	Female	113.011	2.397	Alive	1.814	Progressor	Poor	Ipilimumab	-
3	Non-Responder	Ipilimumab + nivolumab	60	Female	60.048	2.545	Alive	0.173	Progressor	Int	IL-2	-
4	Responder	Pembrolizumab	55	Male	102.751	3.277	Alive	3.274	Non-Progressor	Good	IL-2, Ipilimumab	-
5	Responder	Nivolumab	63	Male	309.440	2.496	Alive	0.844	Progressor	Poor	-	-
6	Responder	Nivolumab	44	Male	41.026	3.710	Alive	2.581	Progressor	Int	Interferon	-
7	Responder	Ipilimumab + nivolumab	76	Male	38.366	2.685	Alive	2.682	Non-Progressor	Poor	-	-
8	Non-Responder	Ipilimumab + nivolumab	74	Male	11.493	1.222	Dead	0.222	Progressor	Poor	-	-
9	Non-Responder	Ipilimumab + nivolumab	28	Female	55.080	0.811	Dead	0.205	Progressor	Poor	-	Yes
10	Responder	Ipilimumab + nivolumab	51	Male	94.481	2.627	Alive	2.553	Non-Progressor	Int	-	-
11	Responder	Nivolumab	60	Male	74.266	2.345	Dead	1.162	Progressor	Good	Ipilimumab	-
12	Non-Responder	Pembrolizumab	75	Female	15.372	0.351	Dead	0.167	Progressor	Poor	Ipilimumab	-
13	Non-Responder	Nivolumab	69	Male	31.925	0.638	Dead	0.145	Progressor	Int	Ipilimumab	-
14	Non-Responder	Ipilimumab + nivolumab	61	Female	32.767	0.715	Dead	0.156	Progressor	Poor	-	-
15	Non-Responder	Ipilimumab + nivolumab	77	Female	38.304	0.288	Dead	0.129	Progressor	Good	-	-
16	Responder	Ipilimumab + nivolumab	38	Female	37.062	1.573	Alive	1.403	Non-Progressor	Good	-	-
17	Responder	Pembrolizumab	73	Male	23.484	1.137	Alive	0.997	Non-Progressor	Good	-	-
18	Responder	Ipilimumab + nivolumab	68	Male	151.159	2.337	Alive	2.079	Non-Progressor	Good	-	-
19	Non-Responder	Pembrolizumab	88	Male	17.739	1.304	Alive	0.197	Progressor	Good	-	-
20	Non-Responder	Pembrolizumab	86	Female	13.209	0.967	Dead	0.170	Progressor	Int	-	-
21	Non-Responder	Nivolumab	50	Male	5.811	0.627	Dead	0.211	Progressor	Poor	Interferon, Ipilimumab	-
22	Responder	Ipilimumab + nivolumab	49	Female	25.608	2.781	Alive	2.762	Non-Progressor	Poor	Interferon	-
23	Non-Responder	Ipilimumab + nivolumab	60	Male	14.075	1.655	Dead	0.181	Progressor	Int	Interferon	-
24	Responder	Ipilimumab + nivolumab	56	Female	55.978	1.652	Alive	0.510	Progressor	Poor	-	Yes
25	Responder	Ipilimumab + nivolumab	62	Female	35.623	4.984	Alive	0.518	Progressor	Int	-	-
26	Responder	Pembrolizumab	64	Female	14.464	3.567	Alive	3.562	Non-Progressor	Int	-	-
27	Non-Responder	Pembrolizumab	74	Male	15.002	1.222	Alive	0.219	Progressor	Good	-	Yes
28	Responder	Pembrolizumab	40	Female	25.919	2.205	Alive	1.279	Progressor	Poor	-	-
29	Non-Responder	Nivolumab	67	Male	57.953	0.066	Dead	0.041	Progressor	Poor	-	-
30	Non-Responder	Nivolumab	16	Male	41.004	0.847	Dead	0.044	Progressor	Poor	Interferon, IL-2	Yes
31	Non-Responder	Pembrolizumab	72	Female	13.081	1.858	Dead	0.219	Progressor	Int	Ipilimumab	-
32	Responder	Pembrolizumab	52	Male	197.926	3.203	Alive	3.019	Non-Progressor	Int	IL-2, Ipilimumab	-
33	Responder	Nivolumab	69	Female	7.997	1.397	Dead	0.595	Progressor	Int	Ipilimumab	-
34	Non-Responder	Ipilimumab + nivolumab	59	Male	78.560	0.652	Dead	0.099	Progressor	Poor	-	Yes
35	Non-Responder	Nivolumablumab	82	Male	15.102	0.548	Dead	0.170	Progressor	Poor	-	-
36	Responder	Ipilimumab + nivolumab	52	Male	206.252	5.471	Alive	5.425	Non-Progressor	Poor	-	-
37	Responder	Pembrolizumab	71	Female	75.748	0.959	Alive	0.740	Non-Progressor	Good	-	-
38	Responder	Pembrolizumab	85	Male	309.708	2.512	Alive	2.510	Non-Progressor	Good	-	-
39	Responder	Ipilimumab + nivolumab	69	Female	14.320	4.260	Alive	4.252	Non-Progressor	Int	-	-
40	Responder	Ipilimumab + nivolumab	64	Male	36.876	5.178	Alive	5.170	Non-Progressor	Good	Ipilimumab	-
41	Responder	Nivolumab	34	Male	202.748	1.742	Alive	1.115	Progressor	Good	TIL therapy, IL-2	Yes
42	Non-Responder	Ipilimumab + nivolumab	56	Female	95.243	1.033	Dead	0.233	Progressor	Poor	-	-
43	Responder	Pembrolizumab	63	Male	217.077	1.885	Alive	1.877	Non-Progressor	Int	Ipilimumab	-
44	Responder	Pembrolizumablizumab	82	Male	5.560	2.378	Alive	2.282	Non-Progressor	Poor	Ipilimumab	-
45	Responder	Ipilimumab + nivolumab	58	Male	7.961	5.101	Alive	4.959	Non-Progressor	Good	-	-

Table S7. Chromogenic immunohistochemistry antibodies and staining conditions.

Marker	Primary Antibody				Secondary Antibody		
	Species	Clone	Source	Final [µg/mL]	Amplification	Source	Final [µg/mL]
FoxP3	Mouse	236A/E7	Affymetrix	5.00	Anti-Mouse Polymer HRP	MP-7402, Vector Labs	RTU
CD8	Mouse	4B11	AbD Serotec	*	Anti-Mouse Polymer HRP	MP-7402, Vector Labs	RTU
Sox10	Mouse	BC34	BioCare Medical	0.24	Anti-Mouse Polymer HRP	MP-7402, Vector Labs	RTU
S100	Mouse	4C4.9	Abnova	0.33	Anti-Mouse Polymer HRP	MP-7402, Vector Labs	RTU
PD-1	Mouse	NAT105	Abcam	1.00	Anti-Mouse Biotinylated	553441, BD Pharmigen	1.00
PD-L1	Rabbit	SP142	Spring Bioscience	0.10	Anti-Rabbit Biotinylated	550338, BD Pharmingen	1.00
CD163	Mouse	10D6	Leica Biosystems	0.49	Anti-Mouse Polymer HRP	MP-7402, Vector Labs	RTU

*RTU anti-CD8 was diluted 1:60. All antibodies were diluted in Antibody Diluent Background Reducing (S3022, Dako).

Table S8. Monoplex IF primary antibody titrations for PD-1/PD-L1 axis mIF assay.

Marker	Primary Antibody					Secondary Antibody			TSA Fluorophore	
	Species	Clone	Source	Final [µg/mL]	Incubation time (min)	Amplification	Source	Dilution (in PBS)	Opal	Dilution
FoxP3	Mouse	236A/E7	Affymetrix	0.31 - 10.00	30	Anti-Mouse PowerVision Poly-HRP	PV6114, Leica Biosystems	1:1	570	1:50
CD8	Mouse	4B11	AbD Serotec	*	30	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	540	1:50
Sox10	Mouse	BC34	BioCare Medical	0.01 - 0.25	60	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	620	1:50
S100	Mouse	4C4.9	Abnova	0.03 - 1.00						
PD-1	Rabbit	EPR4877	Abcam	0.06 - 7.96	30	Anti-Rabbit PowerVision Poly-HRP	PV6119, Leica Biosystems	1:1	650	1:50
PD-L1	Rabbit	SP142	Spring Bioscience	0.01 - 1.54	60	Anti-Rabbit PowerVision Poly-HRP	PV6119, Leica Biosystems	1:1	520	1:50
CD163	Mouse	10D6	Leica Biosystems	0.06 - 1.96	60	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	690	1:50

*RTU anti-CD8 antibody was diluted in the range 1:25-1:800. All primary antibodies were diluted in Antibody Diluent Background Reducing (S3022, Dako).

Table S9. Monoplex IF – TSA fluorophore titrations for PD-1/PD-L1 axis mIF assay.

Marker	HIER Steps		Primary Antibody					Secondary Antibody			TSA Fluorophore	
	Before staining	After Staining	Species	Clone	Source	Final [µg/mL]	Incubation time (min)	Amplification	Source	Dilution (in PBS)	Opal	Dilution
FoxP3	pH9 x 1, pH6 x 1	pH6 x 6	Mouse	236A/E7	Affymetrix	5.00	30	Anti-Mouse PowerVision Poly-HRP	PV6114, Leica Biosystems	1:1	570	1:50 - 1:1600
CD8	pH9 x 1, pH6 x 4	pH6 x 3	Mouse	4B11	AbD Serotec	*	30	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	540	1:50 - 1:1000
Sox10	pH9 x 1, pH6 x 3	pH6 x 4	Mouse	BC34	BioCare Medical	0.13	60	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	620	1:50 - 1:250
S100			Mouse	4C4.9	Abnova	0.06						
PD-1	pH9 x 1, pH6 x 2	pH6 x 5	Rabbit	EPR4877	Abcam	0.50	30	Anti-Rabbit PowerVision Poly-HRP	PV6119, Leica Biosystems	1:1	650	1:50 - 1:800
PD-L1	pH9 x 1, pH6 x 5	pH6 x 2	Rabbit	SP142	Spring Bioscience	0.19	60	Anti-Rabbit PowerVision Poly-HRP	PV6119, Leica Biosystems	1:1	520	1:50 - 1:800
CD163	pH9 x 1, pH6 x 6	pH6 x 1	Mouse	10D6	Leica Biosystems	0.49	120	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	690	1:50 - 1:100

*RTU anti-CD8 antibody was diluted 1:100. All primary antibodies were diluted in Antibody Diluent Background Reducing (S3022, Dako).

Table S10. Antibodies and staining conditions for PD-1/PD-L1 axis mIF assay.

Position	Marker	Primary Antibody					Secondary Antibody			TSA Fluorophore	
		Species	Clone	Source	Final [µg/mL]	Incubation time (min)	Amplification	Source	Dilution (in PBS)	Opal	TSA dilution
1	FoxP3	Mouse	236A/E7	Affymetrix	5.00	30	Anti-Mouse PowerVision Poly-HRP	PV6114, Leica Biosystems	1:1	570	1:200
2	CD8	Mouse	4B11	AbD Serotec	*	30	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	540	1:100
3	Sox10	Mouse	BC34	BioCare Medical	0.13	60	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	620	1:100
	S100	Mouse	4C4.9	Abnova	0.06						
4	PD-1	Rabbit	EPR4877	Abcam	0.50	30	Anti-Rabbit PowerVision Poly-HRP	PV6119, Leica Biosystems	1:1	650	1:100
5	PD-L1	Rabbit	SP142	Spring Bioscience	0.19	60	Anti-Rabbit PowerVision Poly-HRP	PV6119, Leica Biosystems	1:1	520	1:100
6	CD163	Mouse	10D6	Leica Biosystems	0.49	120	Opal Polymer HRP Ms + Rb	NEL811001KT, Perkin Elmer	RTU	690	1:50

*RTU anti-CD8 antibody was diluted 1:100. All primary antibodies were diluted in Antibody Diluent Background Reducing (S3022, Dako).

Table S11. Antibodies and staining conditions for mIF assay to assess PD-1 expression by cell type.

Position	Marker	Primary Antibody					Secondary Antibody			TSA Fluorophore	
		Species	Clone	Source	Final [µg/mL]	Incubation time (min)	Amplification	Source	Dilution (in PBS)	Opal	TSA dilution
1	FoxP3	Mouse	236A/E7	abcam Inc.	6.67	30	Mouse Powervision	Leica Biosystems	1:1	570	1:150
2	CD8	Mouse	4B11	Leica Biosystems	0.285	60	Opal Polymer	Akoya Biosciences	None	620	1:100
3	CD20	Mouse	L26	Leica Biosystems	0.24	60	Opal Polymer	Akoya Biosciences	None	520	1:300
4	PD-1	Rabbit	EPR4877	abcam Inc.	0.96	60	Rabbit Powervision	Leica Biosystems	1:1	650	1:150
5	CD4	Rabbit	EP204	Millipore Sigma	0.18	120	Rabbit Powervision	Leica Biosystems	1:1	540	1:300
6	Sox10	Mouse	BC34	BioCare Medical	0.06	60	Opal Polymer	Akoya Biosciences	None	690	1:150
	S100	Mouse	4C4.9	Abnova Corp	0.06						

Data S1-4. Densities for each of 41 features studied using the different slide sampling strategies.

Data S1. 30% ‘Hot spot’ HPF slide sampling approach for the discovery cohort.

Data S2. 30% ‘Hot spot’ HPF slide sampling approach for the validation cohort.

Data S3. Whole TME (100% sampling) approach for the discovery cohort.

Data S4. Whole TME (100% sampling) approach for the validation cohort.