



# Digital Pathology Analysis Quantifies Spatial Heterogeneity of CD3, CD4, CD8, CD20, and FoxP3 Immune Markers in Triple-Negative Breast Cancer

Haoyang Mi<sup>1\*</sup>, Chang Gong<sup>1</sup>, Jeremias Sulam<sup>1,2</sup>, Elana J. Fertig<sup>1,3</sup>, Alexander S. Szalay<sup>4,5</sup>, Elizabeth M. Jaffee<sup>3,6</sup>, Vered Stearns<sup>3</sup>, Leisha A. Emens<sup>7</sup>, Ashley M. Cimino-Mathews<sup>3,8</sup> and Aleksander S. Popel<sup>1,3</sup>

## OPEN ACCESS

### Edited by:

Russell C. Rockne,  
City of Hope National Medical Center,  
United States

### Reviewed by:

Hermann Frieboes,  
University of Louisville, United States  
Krithika Kodumudi,  
Moffitt Cancer Center, United States

### \*Correspondence:

Haoyang Mi  
hmi1@jhmi.edu

### Specialty section:

This article was submitted to  
Computational Physiology  
and Medicine,  
a section of the journal  
Frontiers in Physiology

Received: 14 July 2020

Accepted: 24 September 2020

Published: 19 October 2020

### Citation:

Mi H, Gong C, Sulam J, Fertig EJ, Szalay AS, Jaffee EM, Stearns V, Emens LA, Cimino-Mathews AM and Popel AS (2020) Digital Pathology Analysis Quantifies Spatial Heterogeneity of CD3, CD4, CD8, CD20, and FoxP3 Immune Markers in Triple-Negative Breast Cancer. *Front. Physiol.* 11:583333. doi: 10.3389/fphys.2020.583333

<sup>1</sup> Department of Biomedical Engineering, Johns Hopkins University School of Medicine, Baltimore, MD, United States,

<sup>2</sup> Johns Hopkins Mathematical Institute for Data Science, Whiting School of Engineering, Johns Hopkins University,

Baltimore, MD, United States, <sup>3</sup> Department of Oncology, Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University, Baltimore, MD, United States, <sup>4</sup> Henry A. Rowland Department of Physics and Astronomy, Krieger School of Arts

and Sciences, Johns Hopkins University, Baltimore, MD, United States, <sup>5</sup> Department of Computer Science, Whiting School of Engineering, Johns Hopkins University, Baltimore, MD, United States, <sup>6</sup> The Bloomberg-Kimmel Institute for Cancer

Immunotherapy, Johns Hopkins School of Medicine, Baltimore, MD, United States, <sup>7</sup> Department

of Medicine/Hematology-Oncology, Hillman Cancer Center, University of Pittsburgh Medical Center, Pittsburgh, PA,

United States, <sup>8</sup> Department of Pathology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Overwhelming evidence has shown the significant role of the tumor microenvironment (TME) in governing the triple-negative breast cancer (TNBC) progression. Digital pathology can provide key information about the spatial heterogeneity within the TME using image analysis and spatial statistics. These analyses have been applied to CD8+ T cells, but quantitative analyses of other important markers and their correlations are limited. In this study, a digital pathology computational workflow is formulated for characterizing the spatial distributions of five immune markers (CD3, CD4, CD8, CD20, and FoxP3) and then the functionality is tested on whole slide images from patients with TNBC. The workflow is initiated by digital image processing to extract and colocalize immune marker-labeled cells and then convert this information to point patterns. Afterward invasive front (IF), central tumor (CT), and normal tissue (N) are characterized. For each region, we examine the intra-tumoral heterogeneity. The workflow is then repeated for all specimens to capture inter-tumoral heterogeneity. In this study, both intra- and inter-tumoral heterogeneities are observed for all five markers across all specimens. Among all regions, IF tends to have higher densities of immune cells and overall larger variations in spatial model fitting parameters and higher density in cell clusters and hotspots compared to CT and N. Results suggest a distinct role of IF in the tumor immuno-architecture. Though the sample size is limited in the study, the computational workflow could be readily reproduced and scaled due to its automatic

nature. Importantly, the value of the workflow also lies in its potential to be linked to treatment outcomes and identification of predictive biomarkers for responders/non-responders, and its application to parameterization and validation of computational immuno-oncology models.

**Keywords:** digital pathology, image informatics, spatial patterns, breast cancer, tumor heterogeneity, immuno-architecture, QuPath

## INTRODUCTION

Triple negative breast carcinoma (TNBC) is an aggressive form of breast cancer that is negative for estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER-2). Treatments for TNBC have historically been confined to surgery, radiation and chemotherapies due to the lack of biologic targets which enable endocrine therapy (ER and PR) and targeted therapy (HER2) in other subgroups of breast cancer. Recent studies deciphering the role of the immune system in cancer revealed a significant effect of the tumor microenvironment (TME) in modulating the tumor progression, especially how the tumor hijacks the anti-inflammatory mechanism of inhibitory immune checkpoint molecules to develop its immune resistance and evasion capability. These studies inspired the emergence of anti-cancer immunotherapy by promoting the host anti-tumor immunity. This idea has led to an array of successful treatments against cancers. Recently, the Impassion130 study demonstrated that first-line treatment with atezolizumab and nab-paclitaxel resulted in an overall survival benefit in patients with advanced programmed death-ligand 1 (PD-L1) positive TNBC, and this is a new standard of care (Schmid et al., 2020). Meanwhile, immune checkpoint blockade also advances the treatment outcomes for other cancer types including melanoma (Long et al., 2016), non-small cell lung cancer (Melosky et al., 2020), and renal-cell carcinoma (Motzer et al., 2019). However, not all patients experience therapeutic benefit from immunotherapy. Heterogeneity within the TME may account for some of the variability in patient response to immunotherapies (Klemm and Joyce, 2015). Therefore, characterization of the tumor immuno-architecture is a critical step toward understanding the complex interplay between pro- and anti-tumor immunity.

Previous research indicates that the existence of tumor-infiltrating lymphocytes (TILs) possesses a unique predictive or prognostic value in different types of cancer (Haanen et al., 2006; Kawai et al., 2008; Suzuki et al., 2010; Ladányi, 2015; Lianyuan et al., 2018), including TNBC (Denkert et al., 2018) as the infiltrating profiles are associated with favorable patient outcomes. Specifically, high overall survival scores are often accompanied by high levels of cytotoxic CD8+ T cells, whereas forkhead box protein 3+ (FoxP3+) regulatory T cells and Type 2/Type 17 CD4+ helper T cells (Th<sub>2</sub> and Th<sub>17</sub> T cells) diminish this effect (Fridman et al., 2011). Therefore, monitoring the distribution of different TILs and their associations should yield insights into how cancer progresses. Hence, further studies are needed to elucidate the underlying mechanisms from the spatio-temporal perspective.

Digital pathology is an emerging discipline that allows quantitative analysis of digital images of histological specimens using computational approaches. Digitized images provide easy access for pathologists to high-resolution histological data with typically gigapixel content (Al-Janabi et al., 2012). Therefore, tissue contexts are well preserved and amenable to computer-assisted techniques for quantitative analysis of the spatial immuno-architecture. We started the development of a workflow for spatial statistical analysis in a previous study with a single immune marker for CD8+ T cells (Gong et al., 2018). The process started with image processing, during which tumor specimen images were segmented to map CD8+ T cells into a Cartesian coordinate system, then the density of point pattern within each subregion (first-order property in spatial statistics terminology) were gathered to reveal spatial variations. For each subregion, cell coordinates were converted into spatial point patterns, then Thomas cluster process was fitted to clustered patterns (assessed by complete spatial randomness test). For the entire point pattern, cell clustering morphometrics were performed. Collectively, fitted clustering parameters and morphometric measurements were harvested to characterize the immuno-architecture (Gong et al., 2018). Other investigators examined the CD8+ T cells infiltration profile by constructing profiles of cell pixel density vs. distance from the tumor boundary and then used a computational model to interpret the data (Li et al., 2019). Alternative approaches introduced the pattern of tumor cells as a reference to measure the infiltration of lymphocytes. CD8+ T cells and the tumor cells were colocalized and then a series of metrics were introduced to measure the spatial interactions such as quantifying the nearest neighbor distribution function for two different cell types using spatial G-function (Barua et al., 2018) and evaluating the spatial clustering using Morisita index and Getis-Ord hotspots analysis (Yuan, 2016). These studies adopt different metrics to interpret spatial distributions of various entities extracted from pathology data; spatial heterogeneity is a universally observed hallmark of cancer whether it is gauged in terms of density, model fitting parameters, clustering size, or infiltration level. Such variations can also be linked to treatment outcomes to better understand the effects of the TME and to assist clinicians in making more accurate diagnoses.

The integration of image processing, statistical analysis, and computational biology has already shown to be powerful in characterizing and interpreting the spatial heterogeneity in multiple tumor types including breast cancer (Brown et al., 2014; Mani et al., 2016; Altan et al., 2018; Du et al., 2019; Wong et al., 2019). Nevertheless, high dimensional quantitative measurements of the interaction between immune markers

and assessment of their spatial correlations have not been conducted. Such metrics will not only provide additional layers of tumor spatial heterogeneity information valuable for patient stratification, but also allow us to better understand the mechanisms behind the formation of the patterns we observe in the TME. In this study, we propose a multi-module workflow to quantify the spatial patterns of five immune markers that control the functional status of T cells, on consecutive pathology sections: CD3, CD4, CD8, CD20, and FoxP3, and from a small patient population ( $n = 5$ ) with TNBC to obtain statistically and pathologically meaningful results. For each patient, our workflow starts with image processing, evaluation of point patterns from three perspectives, and implementation of region-based characterization. To the best of our knowledge, our analysis evaluates the heterogeneity of TNBC in a broad immune context for the first time, therefore paving the way to the identification of reliable predictive biomarkers and the design of innovative therapies when properly correlated with clinical outcomes.

In addition, the cell densities derived from the workflow can be converted to 3D numerical densities to facilitate development and calibration of spatially resolved computational immunology models. For example, 3D densities of different cell types calculated from point patterns can be utilized to populate *in silico* computational agent-based models (ABMs) that have the potential of predicting spatio-temporal TME (Gong et al., 2017; Norton et al., 2017, 2018, 2019). Such three-dimensional ABM could be combined with Quantitative Systems Pharmacology (QSP) models for whole patient to enable mechanistic systems biology modeling of different drugs or combinations (Cheng et al., 2017; Rieger et al., 2018; Bai et al., 2019; Jafarnejad et al., 2019; Milberg et al., 2019; Wang et al., 2019, 2020; Sové et al., 2020).

## MATERIALS AND METHODS

### Pathology Specimen Materials and Methods

This study was approved by the Institutional Review Board of the Johns Hopkins Medical Institutions. Digitally scanned slides from a subset of previously described primary breast tumors were evaluated (Cimino-Mathews et al., 2016). Briefly, formalin fixed, paraffin embedded blocks from surgically resected primary breast carcinomas with no prior neoadjuvant chemotherapy were randomly selected from the pathology archives at Johns Hopkins Medical Institutions (associated response data to treatment not available). TNBC was defined as negative for ER, PR, and the HER-2. Consecutive sections (approximately 5  $\mu\text{m}$  each) from whole tumor were individually stained for CD3 (mouse monoclonal, clone PS1, catalog no. ORG-8982; Leica Microsystems, Bannockburn, IL, United States), CD4 (rabbit monoclonal, clone Sp35, catalog no. 790-4423; Ventana Medical Systems), CD8 (mouse monoclonal, clone C8/C8144B, catalog no. 760-4250; Cell Marque, Rocklin, CA, United States), CD20 (monoclonal, clone MS/L26, catalog no. 760-2531; Ventana Medical Systems, Tucson, AZ, United States), and FoxP3 (mouse monoclonal, clone 236A/E7, catalog no. 14-4777-80,

dilution 1:50; eBioscience; San Diego, CA, United States). Immunohistochemically labeled slides were scanned at 20 $\times$  objective (0.49 microns/pixel) using the Aperio Scanscope AT (Aperio/Leica Biosystems, Vista, CA, United States). Five (5) cases of TNBC were selected for this current study based on intact tissue integrity on the scanned images (i.e., complete cross sections and lack of tissue folds). To simplify the analysis, two tumor islands in Case 1 are split into Case 1A (upper left island) and Case 1B (lower right island). **Supplementary Figure S1** shows a representative biomarker panels across Cases 1–5.

### Computational Methods

The overall workflow includes a central module and four submodules (**Figure 1**). First, the stained (positive) cell nuclei are detected and the tissue annotation module is launched to identify regions of normal tissue (N), central tumor (CT), and invasive front (IF). In this module, image registration is performed for each case on the five slides with different labels, and raw coordinates obtained from image segmentation are mapped to the same reference coordinate system using the transformation matrix. Cell densities are the output of this step, and therefore can be directly visualized using 3D and waterfall graphical representations to visualize intra- and inter-tumor heterogeneity. Also, cell density vs. distance profiles can be constructed. In the spatial point model-fitting module, for each slide, the full point pattern is divided into smaller patches with overlaps, and the Thomas point process model is then fitted based on subregion data if complete spatial randomness (CSR) hypothesis is rejected for this subregion (Baddeley et al., 2015). In the clustering and morphometrics module, for the full point pattern of each slide, the cell clusters are detected using a hierarchical clustering algorithm (Tang et al., 2016). For each detected cluster, morphometrics including convexity, circularity, and eccentricity are calculated and recorded. In the correlation analysis module, a clustering and degree of colocalization-based method is applied to quantify the correlations between immune marker pairs. All spatial statistical measurements used in this study have shown promising application values in the context of digital pathology. Results from these metrics are classified based on tissue type regions and then used for statistical comparison to reveal intratumoral heterogeneity. Such workflow is repeated for all cases, thus capturing inter-tumoral heterogeneity.

### Image Processing

#### *Cell nucleus segmentation and coordinate extraction from IHC slides*

Stained (positive) nucleus segmentation is performed using software platform QuPath (v0.2.0-m10) (Bankhead et al., 2017). QuPath is selected for this study because it is a flexible open-source software with well-managed version control and technical support, and it is capable of a wide range of digital pathology analyses. As the IHC staining may vary both between and within each case, the image processing workflow is initiated by manual correction to the stain estimates for each whole slide image using the 'estimate stain vectors' function. Nucleus detection is then carried out using an unsupervised watershed algorithm with custom parameters tuned on a validation set of WSIs from Case

1, 2, and 3. This built-in algorithm has been implemented by a wide range of peer-reviewed studies (Bankhead et al., 2018; Zhang et al., 2018; Acs et al., 2019a,b; Ferré et al., 2019; Habets et al., 2019; Kather et al., 2019; Santiago et al., 2019; Berben et al., 2020; Blagih et al., 2020; Tsakiroglou et al., 2020). Importantly, the performance was found equivalent to commercial software and pathologists' manual annotations (Bankhead et al., 2018; Acs et al., 2019b; Berben et al., 2020). While the nuclei are identified, centroid coordinates are recorded to represent the cells' location. Afterward, pseudo cell objects are formed by expanding the nuclei boundaries for 7.5  $\mu\text{m}$  and then a list of features is calculated based on intensity and morphometry measurements. For each IHC biomarker, 25 regions of interests (ROIs) are randomly selected for Cases 1–3 (75 in total) and a Random Tree classifier (Breiman, 2001) is trained using the aforementioned features by annotating regions in a subset of subregions. The classification results are updated in the form of color marks when each annotation is drawn. The classifier is then validated on the remaining WSIs to ensure the robustness. A low-resolution image in the testing set is shown in **Figure 2A**. Exemplar segmentation and classification results are shown in **Figures 2B,C**. Segmentation and classification settings are shown in **Supplementary Material** and **Supplementary Tables S1, S2**.

To evaluate the performance of the image segmentation algorithm, 20 subregions are sampled from each slide for each case using the random sampling method. **Figure 2D** shows four exemplar subregions for performance evaluation, where red outlines indicate QuPath segmentation results and green dots indicate manual approach. For each subregion, we also manually detect labeled cells, and then measure the sensitivity (recall) and precision of our algorithm. Results indicate that there is a strong correlation between manual and automatic approaches (Spearman's correlation coefficient  $\rho = 0.978$ ). Details of the evaluation of QuPath can be found in the **Supplementary Material**.

### **Registration and coordinate transformation of IHC slides**

The pathology images available for this study were single label IHC slides produced with consecutive sections from each tumor excision. In this process, z-axis difference for each section, location and rotation when placed onto slides, as well as possible folding of tissue during preparation, all contributed to discrepancies between coordinate systems of each slide from the same patient. These discrepancies are minimized by image registration. As the cutting sequence of these immune marker slides was unknown, all slides are treated equally and the CD4+ slide is selected as the reference for all cases. Global automatic registration by Matlab application "Registration Estimator" is first performed for all pairs (**Figure 3A**). The registration accuracies are manually assessed based on tissue overlap level: among 20 registration pairs, we find that global registration produces high accuracy for five pairs; for the remaining, the local registration is required, which is performed using software Icy (De Chaumont et al., 2012) (**Figure 3B**). Both global and local registrations generate transformation matrices for the corresponding regions. These matrices can be used to estimate registration accuracy. First, global registrations are performed

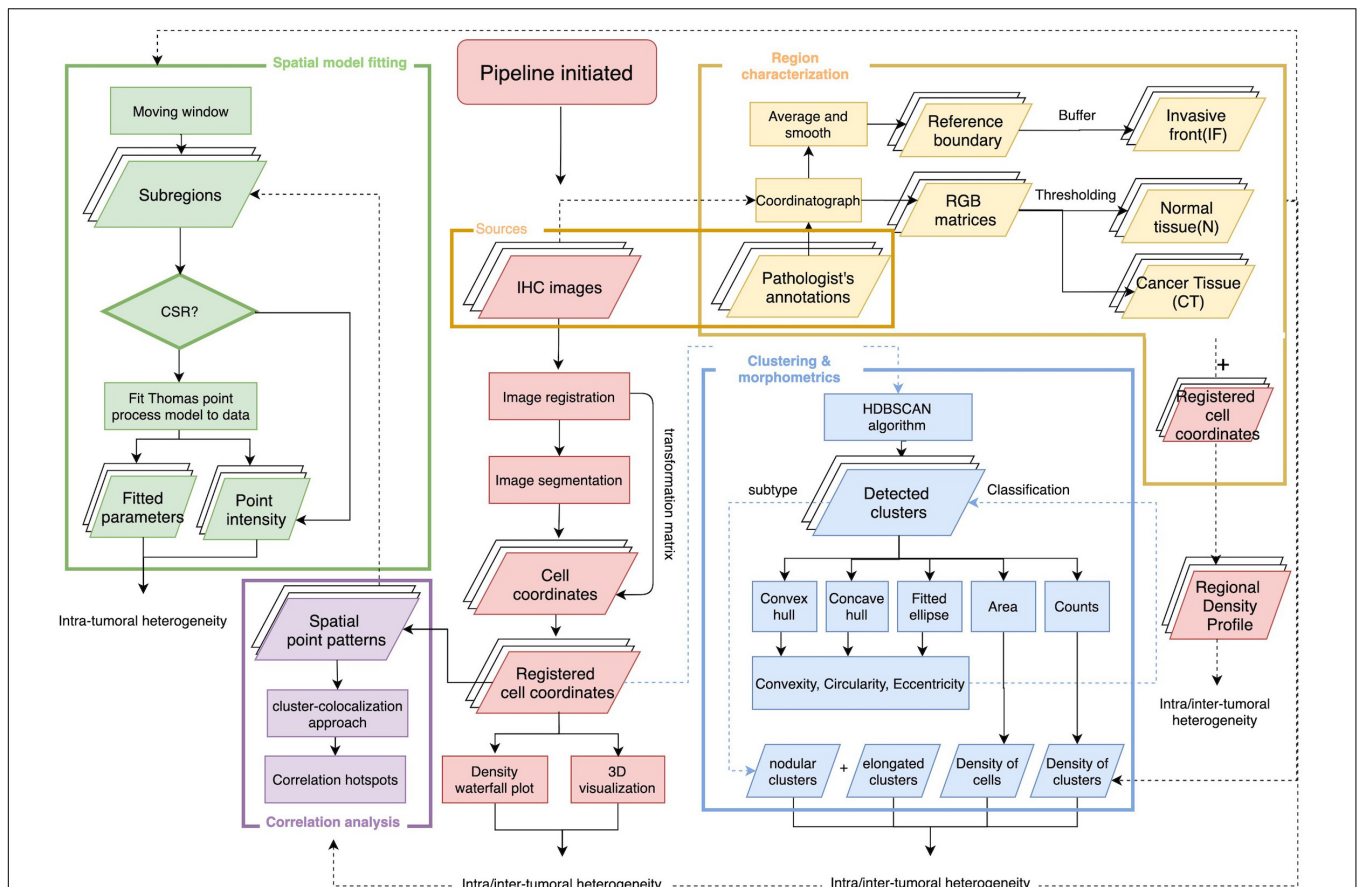
for all pairs as the baseline. Next, transformation matrices generated from global and local registrations are used to register tissue contours, separately. Dice Similarity Coefficients (DSCs) (Guy et al., 2019) are calculated, respectively, and cumulative results are collected (**Figure 3C**). Finally, we compared the registration accuracy by performing the Wilcoxon rank-sum test between global and local DSC groups (**Figure 3D**) and observed a significant improvement ( $p = 4.90e-3$ ). Results also show that the average global registration DSC scores for those five pairs (0.916) are very similar to the average local registration DSC scores (0.917). Technical details on global and local registrations and performance evaluation can be found in the **Supplementary Material**.

### **Measuring Intra- and Inter-tumoral Heterogeneity**

#### ***Region characterization based on pathologist's annotations***

For each case, the breast cancer pathologist (AMC-M) annotated (outlined) the tumor boundary, which was considered the 'ground truth' for the present analysis. Green contours in **Figure 4A** and **Supplementary Figure S2** indicate the annotations for Case 1 and Cases 2–5. Annotations are converted into coordinate sets and then registered to the reference slide (CD4+ slide) using the transformation matrices. Each annotation is a closed curve so that the corresponding coordinate set can form a polygon. Next, we create a score map by overlaying the five coregistered polygons (from five slides, one for each immune marker) and recording the number of polygons each pixel of the WSI resides within. Now that each pixel is assigned a score, we apply a smoothing filter to the score map and threshold all pixels with scores exceeding 2.5 to determine the consensus tumor boundary. Finally, we obtain a dense region, the contour of which functions as the averaged boundary between normal tissue [sometimes referred to as stroma (Tyekucheva et al., 2017)] and CT. Next, we buffered the boundary with 0.5 mm inward and outward to create the IF, which separates the normal tissue (N) and CT with a band of 1 mm (Pages et al., 2009; Halama et al., 2011; Hendry et al., 2017). Note that in some studies the region is considered 0.5 mm wide (Gong et al., 2018; Li et al., 2019); we will show below the quantitative implications of either assumption. Afterward, we remove those pixels that fall within the IF from the tumor mask, thus the contour of remaining points gives the outline of CT. Similarly, to extract the normal tissue, we computed the mean value of the RGB channels for each pixel. To exclude background and noise, we set a customized threshold to rule out high-intensity pixels. The remaining pixels contain complete WSI foreground information, and the normal region can be easily extracted when IF and CT pixels are removed. For all pixels associated with each region, we obtain the outline to form a polygon to represent the region. In this step, specimen images and annotations from pathologist are the only inputs. Pixel coordinate maps are generated using python. All subsequent computations are performed using R: the point-in-polygon test is performed using 'point.in.polygon' function from R package 'sp' (Pebesma and Bivand, 2005); the outlines are generated using 'concaveman' function from R package 'concavemann' (Gombin et al., 2017).





**FIGURE 1 |** Overall workflow of spatial pattern quantification for immune markers. The workflow is initiated by two steps, first the image processing for IHC slides to extract coordinates of immune marker-labeled cells; second, the tissue type regions are characterized based on pathologist’s annotations of the tumor boundary and original IHC images. The results from these two steps construct regional density profiles. The point patterns are fed to all remaining submodules to quantify intra- and inter-tumoral heterogeneity. In the spatial model-fitting submodule, point patterns within subregions are tested for CSR, and the Thomas model is fitted to data if test is rejected. In the clustering and morphometric submodule, point clusters are detected and multiple shape descriptors are calculated for each cluster. In the correlation analysis submodule, a cluster-colocalization based method gauges the spatial distributions near each point to identify correlation hotspots, where highly correlated immune marker pairs are engaged in. For each slide, the collective results from the aforementioned metrics capture intra-tumoral heterogeneity and the analyses repeated across all cases capture inter-tumoral heterogeneity.

### Whole slide image partitioning and extraction of first-order properties

We partition the WSI into subregions for local spatial analysis using a moving rectangular window with edge lengths of  $x_{window}$  and  $y_{window}$ , which traverses the WSI with step size of  $x_{step}$  and  $y_{step}$ . The window size should be large enough to capture local density variations, and sufficiently small to have multiple subregions and stationary underlying point pattern processes. Based on these considerations, we performed fractal analysis (see **Supplementary Material** and **Supplementary Figure S3**) and determined the window lengths and step size as  $x_{window} = y_{window} = 0.4$  mm and  $x_{step} = y_{step} = 0.2$  mm (complete discussion is in **Supplementary Material**). In this study, we define all individual rectangular areas that the moving window has scanned as subregions. As the window is moving through the whole slides, first-order properties such as number and density of points are recorded for subsequent visualization and local statistical analysis.

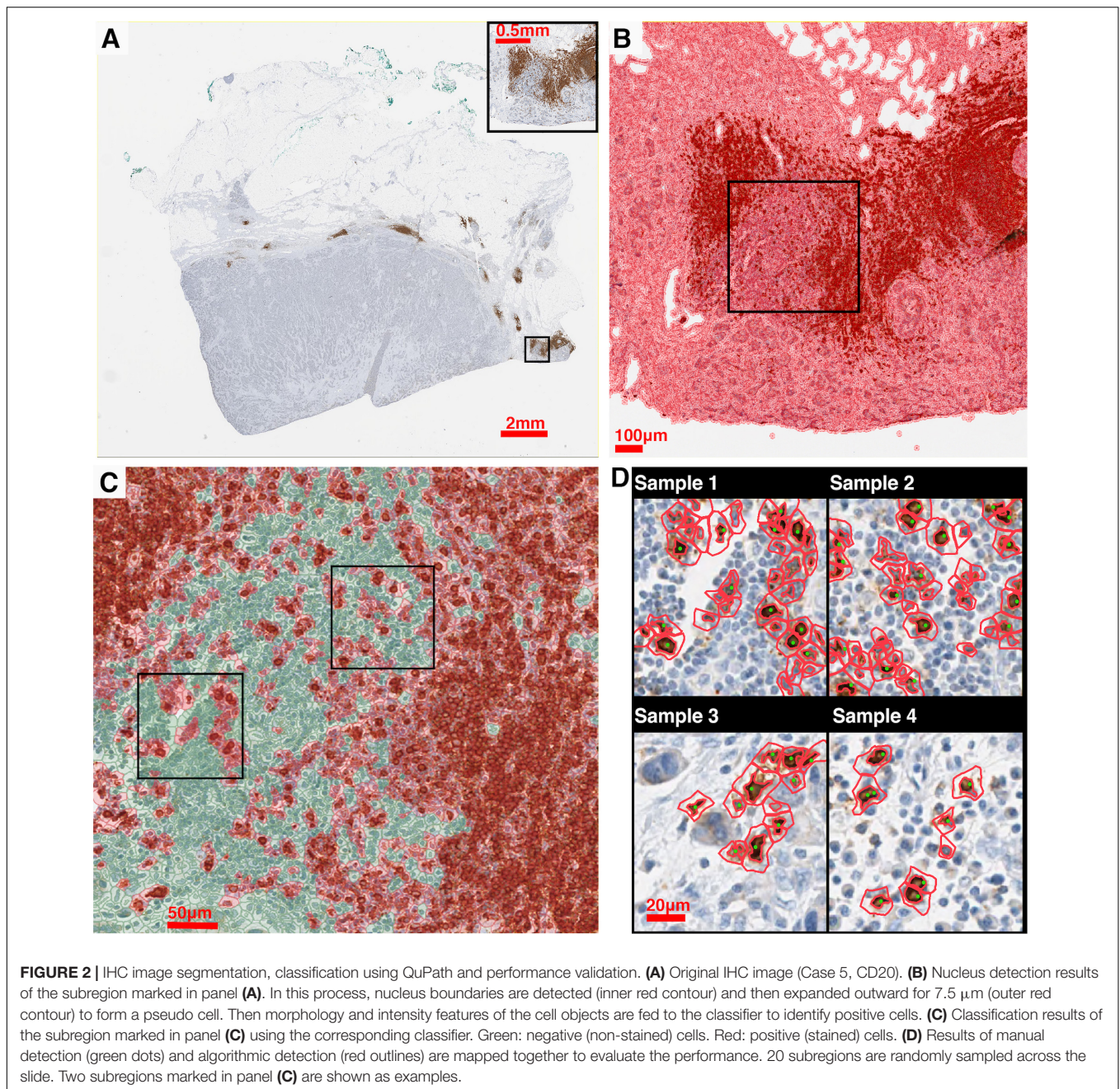
### Measuring the heterogeneity with spatial entropy measurement

A form of Shannon’s entropy (Claramunt, 2005) uses the entropy as a measure of diversity of density of multiple point species in space. This modified version incorporates the factor of distance. Assuming that the increase of distance between same type of points and the decrease of distance between different types of points will result in the increase of entropy, this spatial entropy is defined as:

$$H_{SC} = - \sum_{i=1}^n \frac{d_i^{int}}{d_i^{ext}} p_i \log_2 p_i \tag{1}$$

where  $d_i^{int}$  is the average Euclidean distance between all points of type  $i$ ;  $d_i^{ext}$  is the average Euclidean distance between all points of type  $i$  and the points of other types;  $p_i$  is the percentage of type  $i$  within the subregion.

For each case, we first map registered and reference point patterns into a Cartesian coordinate system. Next, we calculate



the  $H_{sc}$  for the multi-type point pattern within subregions. Collective results are then classified based on tissue type associated with their locations. Then we examine intra- and inter-tumoral heterogeneity of spatial entropy in each tissue type by showing their distributions with a series of probability density functions.

#### *Constructing cell density versus distance profile for whole slide image*

To compute cell density-distance profiles, infiltration intensities are quantified as the immune-marker labeled cell densities at the corresponding distance from consensus tumor boundary

obtained in the previous step. To generate the density-distance profile, we utilized pixel-distance based algorithm to segment the whole tissue into multiple equal-width band sections: all foreground pixels are classified according to their locations, denoted as CT-, IF-, and N-pixels; for both CT- and IF-pixels, the distance toward the IF inner (adjacent to CT) boundary are computed; for each N-pixel, the distance toward the IF outer (adjacent to normal tissue) boundary are computed; next we group the pixels into multiple intervals with the interval length of 150  $\mu\text{m}$  (first interval: 0–150  $\mu\text{m}$ ; second interval: 150–300  $\mu\text{m}$ ; ...) based on their distance value; then we extract the shape outline for each group by computing the



concave hull; consequently each group comprises a polygon with 150  $\mu\text{m}$  width. Minor adjustments are needed to ensure all points within each shape outline are properly arranged so that they can be connected in clockwise or counterclockwise direction, if necessary.

To ensure the density-distance profile proceeds along the direction of immune infiltration, first we assign index 0 to the central band polygon of the IF as the reference polygon. Then we render negative indexes, decreasing from 0 until we reach the most distant band within the normal tissue, and render positive indexes, increasing from 0 until we reach the most distant polygon within the tumor. With all groups are arranged in a consecutive numerical order that expands from the edge of normal tissue toward the CT, the cell density is generated by calculating the area and counting the cells inside. The described methods are summarized in **Figures 4A–C**.

For all aforementioned calculations, the concave hull is computed using a function from R package ‘concaveman’; the pixel distance and polygon areas are computed using the function ‘gDistance’ and ‘gArea’ from R package ‘rgeos’ (Bivand and Rundel, 2017).

#### Constructing cell density versus distance profile within invasive front

The IF is segmented into sections along its own direction: the length of central reference line is first calculated; then multiple equispaced points along the line are sampled based on a given interval. A Voronoi tessellation was created within the IF region based on Euclidian distance to each chosen point. The area of each resulting polygon and the number of cells it encloses are computed to determine cell density, and all polygons are indexed in a numerical order starting from 1 with the left-most point, in a clockwise direction. The methods are summarized in **Figure 4D**. The length of the reference line is calculated using the function ‘lineLength’ from R package ‘SDraw’ (McDonald, 2016); the length of interval is set to be 0.2 mm; the points are sampled using the function ‘spsample’ from R package ‘sp.’

#### Constructing 95% confidence interval (CI) along with the cell density profile

Depending on the shape of whole slide images, polygons at a certain distance may be small in size, which could result in inaccurate estimation of the average cell density of that distance. Therefore, to test the reliability of all average cell density point estimates along the 2D projection, we construct a 95% CI along the profile. For better accuracy, we assume the variance of density between each window within one distance polygon band is equal to the variance of the entire region with the same tissue type as the polygon of interest, which is denoted as  $\sigma$ . Hence the confidence interval can be computed according to the formula (Efron, 1981):

$$D \pm T_c \cdot \frac{\sigma}{\sqrt{n}} \quad (2)$$

where  $D$  is the density of cells within a polygon,  $T_c$  is the critical  $t$ -value. In this study  $T_c$  is defined as 1.96 (95% confidence level);  $n$  is the number of samples, and is calculated according to the equation:

$$n = \frac{A}{s_l \times s_w} \quad (3)$$

where  $A$  is the area of the polygon;  $s_l$  and  $s_w$  are the length and width of the window. In this study, window size is defined as  $s_l = s_w = 0.4$  mm.

Confidence intervals are calculated along with the density profile; however, when visualizing the data, we truncated the portion below zero, and use 80% of the density as the threshold to filter out locations at which the density estimates are not reliable as the mean of the region (labeled as red dots).

#### Constructing the three-dimensional immune landscape

For each slide, the recorded density mapped to original locations to construct the landscape. The entire landscape is then characterized to reflect region-specific information (N, IF, and CT). The landscape data visualization is implemented using software Blender 2.80 (Hess, 2007).

#### Measuring the heterogeneity from spatial point pattern process model fitting results

In our study, the local point pattern is defined as the point pattern of the immune marker within a given subregion. For each captured local point pattern, Complete Spatial Randomness (CSR) is tested using the Clark-Evans test with the null hypothesis being a uniform Poisson process (one-tailed,  $H_A$ : clustered distribution, significance level  $\alpha = 0.05$ ) (Baddeley et al., 2015). If the pattern failed to pass the CSR test, we fit a Thomas point process model to the local point pattern and record fitted parameters. The model assumes cluster patterns are generated in two steps: in the first step, a pattern of parent points within the window is generated according to a homogeneous Poisson process given the intensity  $\kappa$ ; in the second step, a random number of offspring points is generated, so that the number of offspring points that belong to any parent point also follows Poisson distribution with intensity  $\mu$ , at the same time the location follows isotropic Gaussian distribution with standard deviation  $\sigma$ . The theoretical Ripley’s  $K$  function of the Thomas process is:

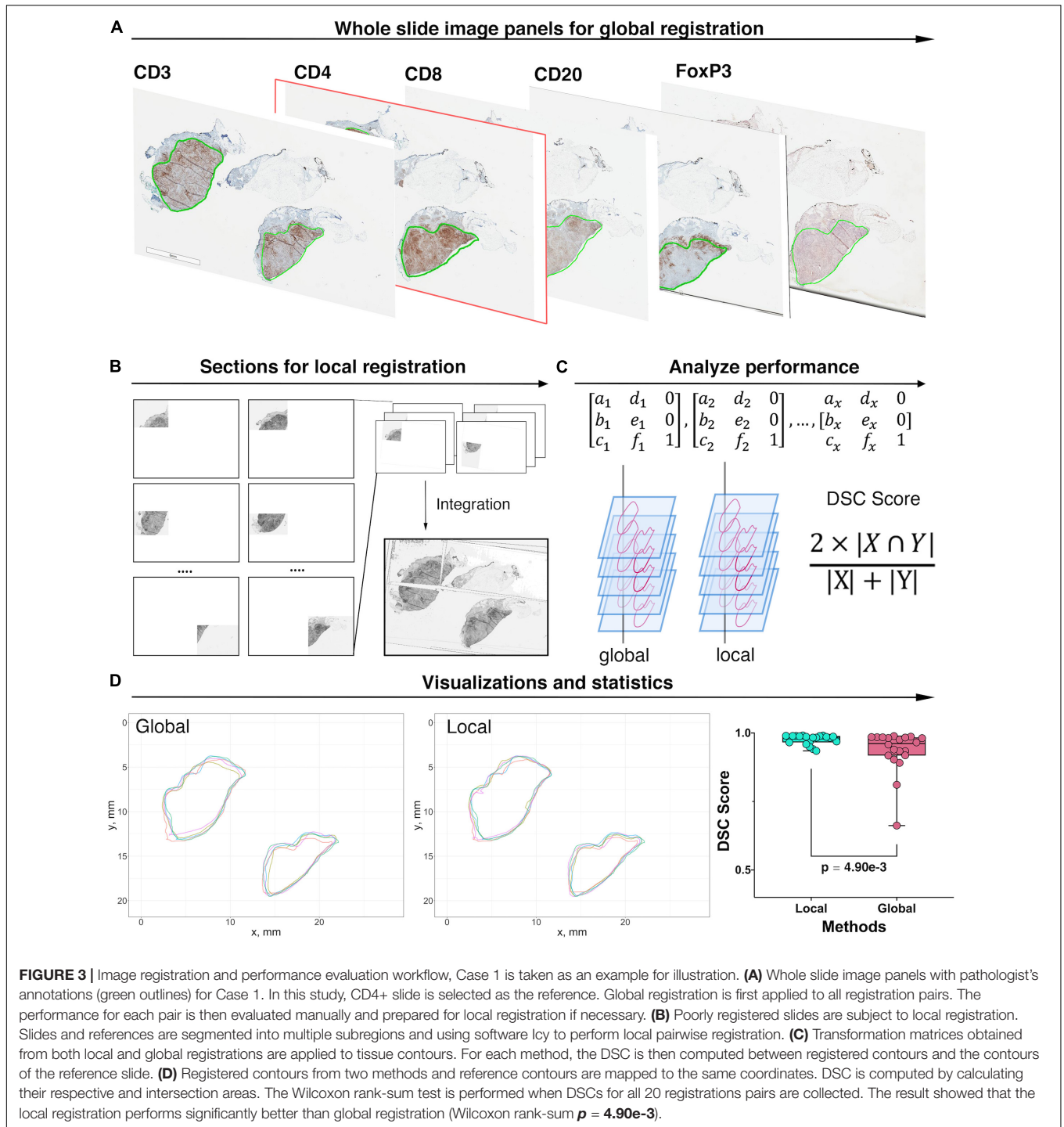
$$K(r) = \pi r^2 + \frac{1}{\kappa} (1 - e^{-\frac{r^2}{4\sigma^2}}) \quad (4)$$

where  $r$  is the distance of a sample random point of the point pattern within which the function is evaluated. For each sub-region, the fitted parameters  $\kappa$ ,  $\mu$ , and  $\sigma$  are biologically interpreted as features of the clustering pattern of immune marker-labeled cells.  $\kappa$  stands for the number of labeled cell clusters per unit area;  $\mu$  is the number of labeled cells per cluster. We further use the distance of the cell toward the cluster center to quantify the internal cell distribution of each cluster (Thomas, 1949; Waagepetersen, 2007; Tanaka et al., 2008). We observe that the components of distance vector are normally distributed and independent since each point in the generated clustered pattern is produced from an isotropic Gaussian process so that the collective distance profile within each cluster follows a Rayleigh distribution with a probability density function:

$$G(r; \sigma) = \frac{r}{\sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad (5)$$

The average moment can be calculated as

$$\mu(r) = \sigma\sqrt{\pi/2} \quad (6)$$



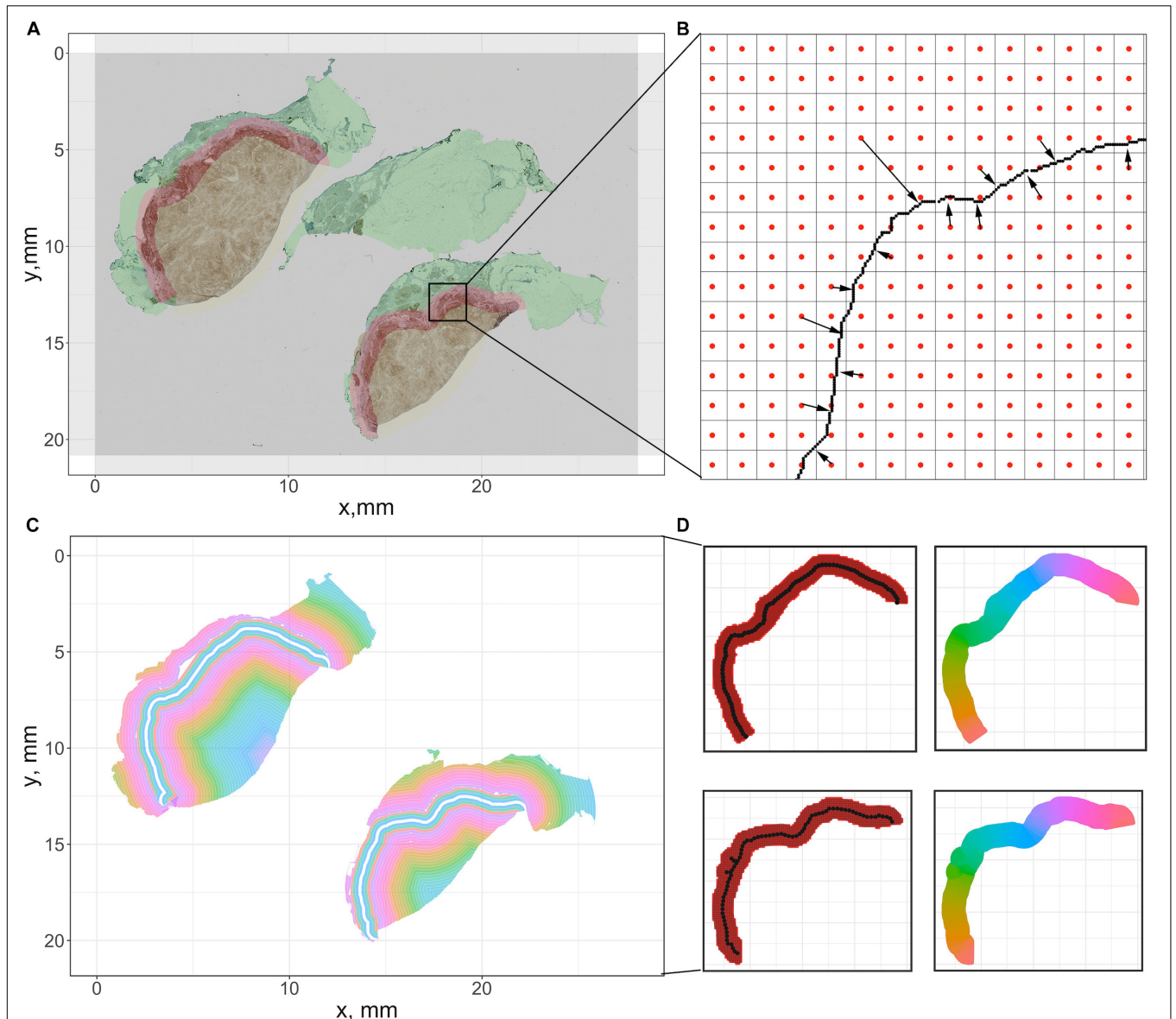
and the radius of the circle where 95% of the cells would fall in is calculated as:

$$Q(F, \sigma) = \sigma \sqrt{-2 \ln(1 - F/100)} \quad (7)$$

where  $F = 95$ . The CSR testing is performed using functions “clarkevans.test” and “kppm” from R package “spatstat,” with parameters “clustered” and “Thomas,” respectively (Baddeley et al., 2015).

For each slide, we measure and compare spatial statistics between different regions (intra-tumoral) and between different cases (inter-tumoral). We use the quartile coefficient of dispersion (QCoD) and the coefficient of variation (CoV) to assess the variability of the metrics (density and spatial model fitting parameters) within one slide. Furthermore, the median value of each metric is used to represent the corresponding case for the case-wise comparison. QCoD





**FIGURE 4 |** Summary of the pixel-distance based method to generate sections for cell density-distance profiles. We use Cases 1A and 1B here for illustration. **(A)** For the downsampled image, we generate a grid whose size is equal to the number of pixels. Then the center coordinate of each square is recorded to represent the pixel. **(B)** For all pixels within tissues, we calculated the distance toward the tumor boundary of the corresponding case. **(C)** All pixels are grouped based on their distances to generate equidistant sections. **(D)** To segment the invasive front, multiple equidistant points are sampled along the boundary, and each pixel within the IF is assigned to their nearest sample point. Then all pixels associated with each sample points form a polygon to calculate density. The color codes in the right panels indicate different sections.

and CoV are computed using the following formulas:

$$QCoD = \frac{Q_3 - Q_1}{Q_3 + Q_1} \tag{8}$$

$$CoV = \frac{\sigma}{\mu} \tag{9}$$

where  $Q_1$ ,  $Q_3$ ,  $\mu$ , and  $\sigma$  are the first quartile, third quartile, standard deviation, and mean of each metric.

**Measuring the heterogeneity from clustering and morphometric analysis**

Immune contexture heterogeneity in the TME holds a significant value to the study of the anti-tumor immune response (Beck et al., 2011; Schwen et al., 2018). Therefore, we performed immune cell cluster analysis to assess the intra- and inter-tumoral heterogeneity. We first identify clusters from the global point patterns using an adjusted version of the clustering algorithm Hierarchical DBSCAN (HDBSCAN) from R package ‘largeVis’ (Tang et al., 2016). This method

generates the cluster hierarchy based on density-adjusted distance connectivity, and parent and child cluster stabilities are compared to extract clusters. The algorithm arguments are defined as minPts (minimum cells per cluster) = 30, and  $K$  (the number of cells in the core neighborhood) = 4. Next, morphological analysis is performed on each previously identified cluster by calculating shape descriptors. To describe the structure of clusters, we introduced  $\alpha$ -shape, which envelops a set of points by point pairwise segments that could be regarded as a chord of a circle with a given radius  $\alpha$ . To identify the  $\alpha$ -shape which exactly harbors all points within the given region, the value of  $\alpha$  is increased from 10  $\mu\text{m}$  until the polygon reaches the ideal size.  $\alpha$ -shapes are computed using the function “ashape” from R package ‘alphahull’ (Rodríguez Casal and Pateiro López, 2010). Then the morphometrics for each cluster are calculated for the following measurements:

**Convexity:** Measures the degree of the object. Convexity is mathematically defined as:

$$f_{convex} = \frac{A_{\alpha}}{A_{convex}} \quad (10)$$

where  $A_{\alpha}$  is the area of the  $\alpha$ -shape and  $A_{convex}$  is the area of the convex hull, generated upon the same dataset.

**Circularity:** Measures the roundness of the object. Circularity is mathematically defined as:

$$f_{circularity} = \frac{4 \cdot \pi \cdot A_{\alpha}}{P_{\alpha}^2} \quad (11)$$

where  $P_{\alpha}$  is the perimeter of the  $\alpha$ -shape.

**Eccentricity:** Measures the degree of deviation of the object from being circular. To generate the ellipse, we first assume the points within each cluster follow chi-squared distribution  $Q \sim \chi^2(k)$ . Then the eigenvectors can be calculated from the covariance matrix, which indicates orientations. Now the semi-major and -minor axis lengths can be computed as:

$$a = \sqrt{\lambda_1 X_2^2(0.95)}, b = \sqrt{\lambda_2 X_2^2(0.95)} \quad (12)$$

where  $\lambda_1$  and  $\lambda_2$  are eigenvalues of the covariance matrix. By this definition, the ellipse is represented as the contour where 95% of points were covered. Thus, the eccentricity is computed as:

$$e = \sqrt{1 - \frac{b^2}{a^2}} = \sqrt{1 - \frac{\lambda_2}{\lambda_1}} \quad (13)$$

## Correlating the Spatial Patterns of Different Immune Markers

The metrics above all referred to spatial distributions of a single marker. However, it is of interest to know if the distributions of cells with different markers are correlated. For example, CD8+ cells generally inhibit tumor growth whereas FoxP3+ cells generally promote tumor growth; if they are colocalized their effects might cancel each other. Such

assumption simplifies the definition of T cell lineages due to the limitations in materials and the scale of biomarker panel. In this study, we implement a variation of the Clus-DoC (clustering-degree of colocalization) approach to analyze spatial correlation between different immune cell labels (Pageon et al., 2016). We focused on the correlation between three pairs of the spatial patterns: CD3+/CD8+, CD4+/FoxP3+, and CD8+/FoxP3+ as representations for anti-tumor immunity regions, immunosuppression regions, and immune-crosstalk regions. With each pair of full point patterns, for both channels, a DoC score is assigned to each point. This step requires the comparison of the spatial distribution of all the neighboring points from both channels for every single point. Centered at each point of type A, circles with increasing radius are formed to compute the associated density gradients of points from both channels. Then for each point of type A, the correlations between the density gradients between both channels are measured by Spearman's rank coefficient  $\rho_{AB}$ . Next, each coefficient  $\rho_{AB}$  is converted to a DoC score by normalization using the equation:

$$DoC_A = \rho_{AB} \cdot e^{-\left(\frac{N_{AB}}{R_{max}}\right)} \quad (14)$$

where  $N_{AB}$  is the distance of the current point of type A to the nearest neighbor of type B,  $R_{max}$  is the maximum search radius. Thus, the DoC score is bounded within  $[-1, 1]$ , where 1 indicates a strong correlation (colocalization) and  $-1$  indicates anti-correlation (segregation). These calculations are performed for both channels and DoC scores are then used to identify correlated points. To select a proper DoC threshold, we create a synthetic point pattern by shifting the full point patterns of a given slide (we use CD8+ slide from Case 2 in our study, but it can be any WSI within the study cohort) to a given direction by a minor distance to simulate a well-localized pattern pair. For simplicity, we unify the shift directions for all points to left and with a distance to 10  $\mu\text{m}$  plus uncertainty caused by the registration error. The averaged DoC score is then selected as the threshold. Spatially, points with high DoC scores (correlated) are close to other points of both channels, whereas points with low DoC scores (non-correlated) are not close to points of at least one channel.

In the second step, the threshold is used to select highly correlated points for each channel. Points from both channels are then mapped to the same coordinate system. Next, we use the density-based clustering algorithm described in the morphometric module to detect clusters that contain points of both types A and B. Such clusters highlight regions with strong mutual interactions of immune markers. The search of neighbors for each point is calculated and accelerated using the C++ implemented k-dimensional tree search algorithm in Python library ‘SciPy’ (Virtanen et al., 2020); the distance of a point to its nearest neighbor is calculated using function ‘nncross’ in R package ‘spatstat’ (Baddeley et al., 2015); the density-based clustering is performed using function HDBSCAN in R package ‘largeVis,’ with arguments  $K = 4$ , and minPts = 30.

## RESULTS

### Assessing Intra- and Inter-tumoral Heterogeneity With Multiple Metrics Immune Cell Density Distribution and Infiltration Profiles

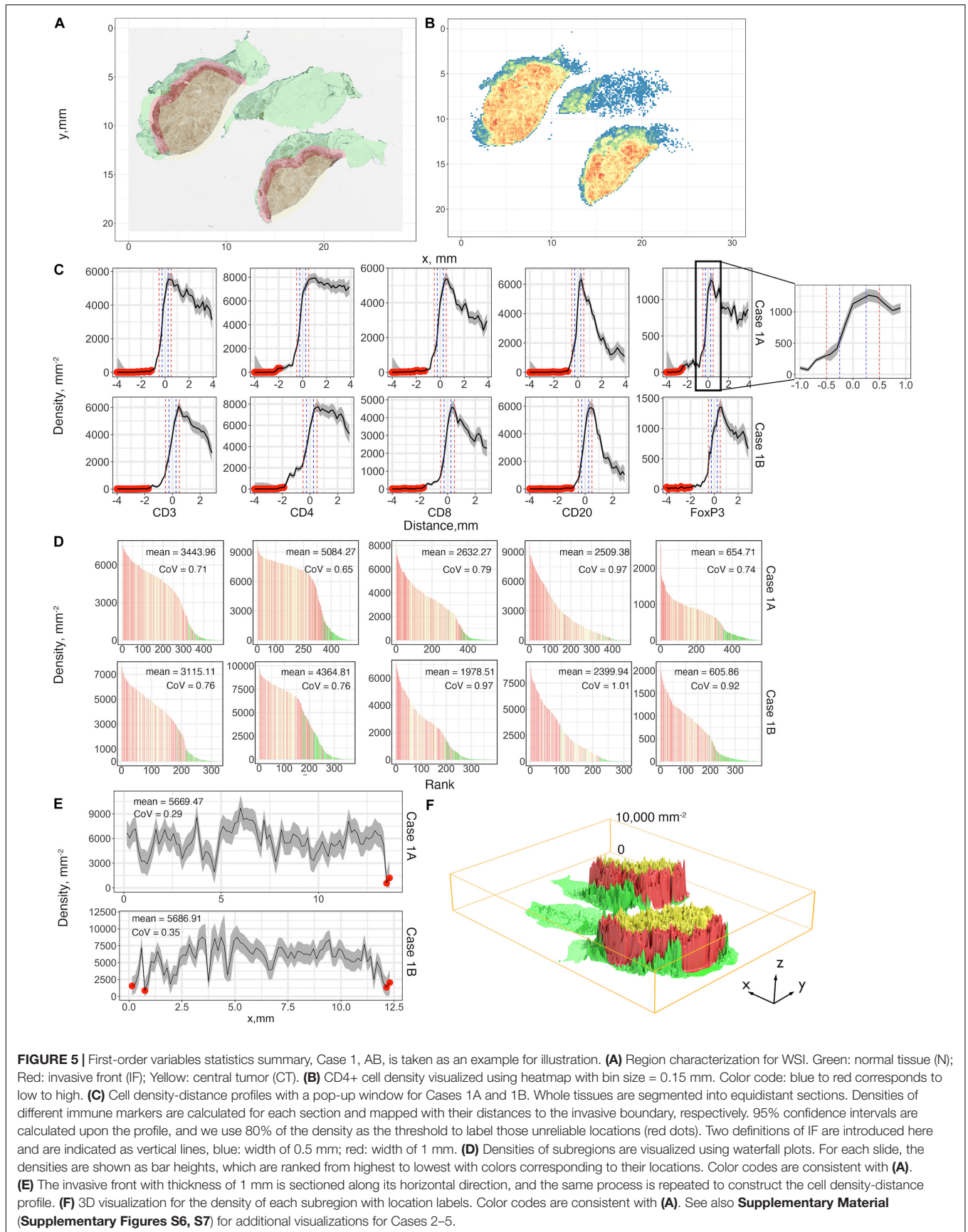
A summary of first-order properties analysis is shown in **Figure 5**. Cases 1A and 1B are used here for illustration; both parts of the tumor are present on the same slide but are separated. For each slide, the region annotations (**Figure 5A**) and coordinates extraction (**Figure 5B**) are performed first to characterize the spatial distribution. **Figure 5C** shows profiles of cell density vs. distance from the boundary in mm for each case for the five labels. These are the immune-infiltration profiles for different cell types. Two definitions of the IF are introduced here, with a width equals 0.5 mm (blue vertical dashed lines) and 1 mm (red vertical dashed lines) in accordance with pathologists' convention. 95% confidence intervals are depicted in gray bands. Most of the sections consist of sufficient  $0.15 \times 0.15$  mm windows to estimate CI; however, near the edges of the slide, in normal tissue (N) and CT, the areas may be small resulting in wide confidence intervals; in these cases we first exclude all the small regions (area  $< 1.1$  mm<sup>2</sup>) from analysis and then we manually set a threshold with  $\pm 80\%$  of real density to label the remaining unreliable sections (red dots). Importantly, we observe that the immune-infiltration profiles are unbiased regardless of how IF is defined. The maxima for the different immune cells are shifted from the boundary toward the CT, but still are within the IF. In other words, the cell densities peak around the 0.2–0.35 mm band and then drop gradually toward the innermost of the CT. To further corroborate that such infiltration profiles are caused by tumor heterogeneity, we compare the actual cell distribution pattern to a binomial distribution pattern by quadrat test. Theoretically, if a point set is randomly generated over a region which consists of multiple sections, then the expected number of points each section harbored can be calculated as the total number of points multiplied by the probability a point happened to be in this section. Therefore, for each WSI, we harvest the actual cell counts (frequencies) and theoretical frequencies in each section and performed chi-square independent test. Results show that the null hypothesis (two types of observation are independent) is rejected by all trials, suggesting that the actual cell distribution pattern is not a realization of randomness, rather it is driven by heterogeneity (**Supplementary Figure S4**).

Waterfall plots are frequently used in other studies to present results of clinical trials, when patients' responses are ranked from best to worse, using tumor size as a metric, each patient is represented as a bar in the plot. In this case we use waterfall plot to rank cell densities from largest to smallest, from  $0.4 \times 0.4$  mm windows throughout the tissue, for each label with colors corresponding to N (green), IF (red) and CT (yellow). The results are shown in **Figure 5D**. Waterfall plots indicate that CT and IF tend to have higher cell densities whereas fewer cells tend to accumulate in N, as the left-hand side of the chart contains more red (IF) bars and right-hand side have more

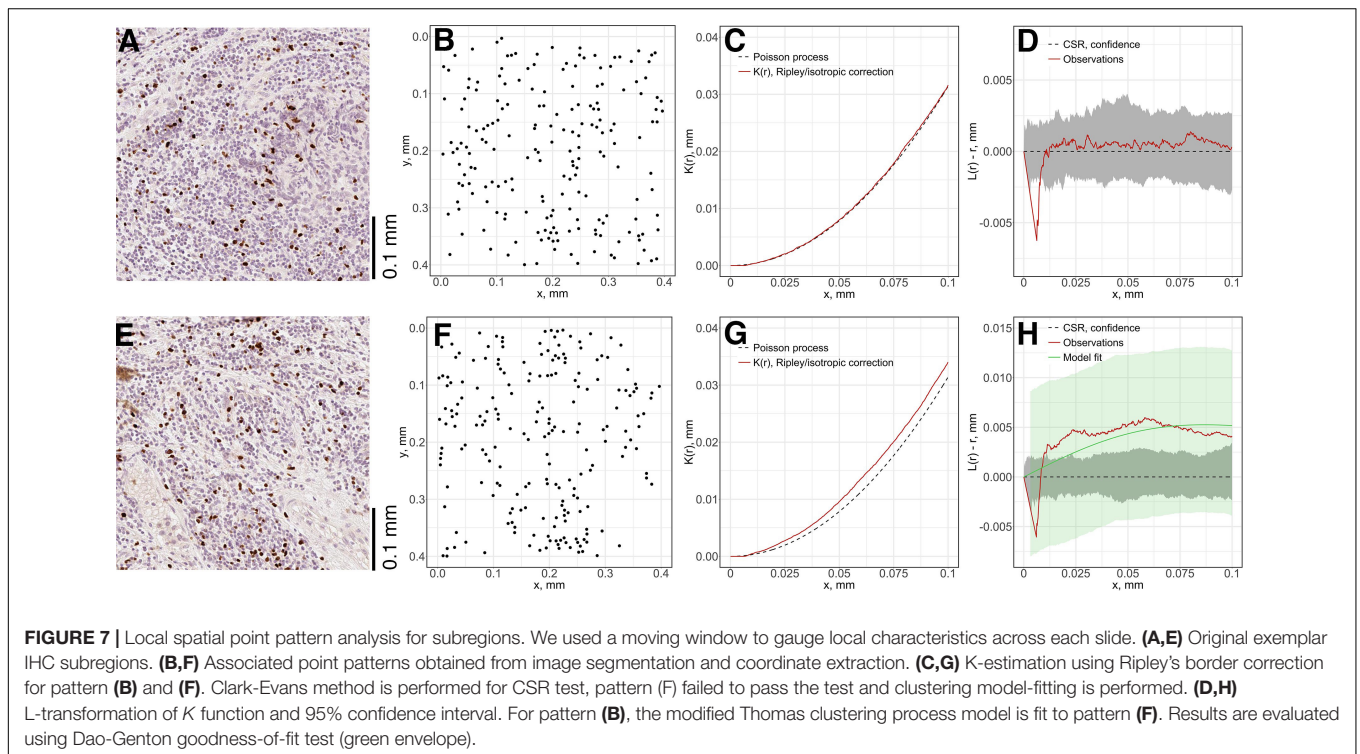
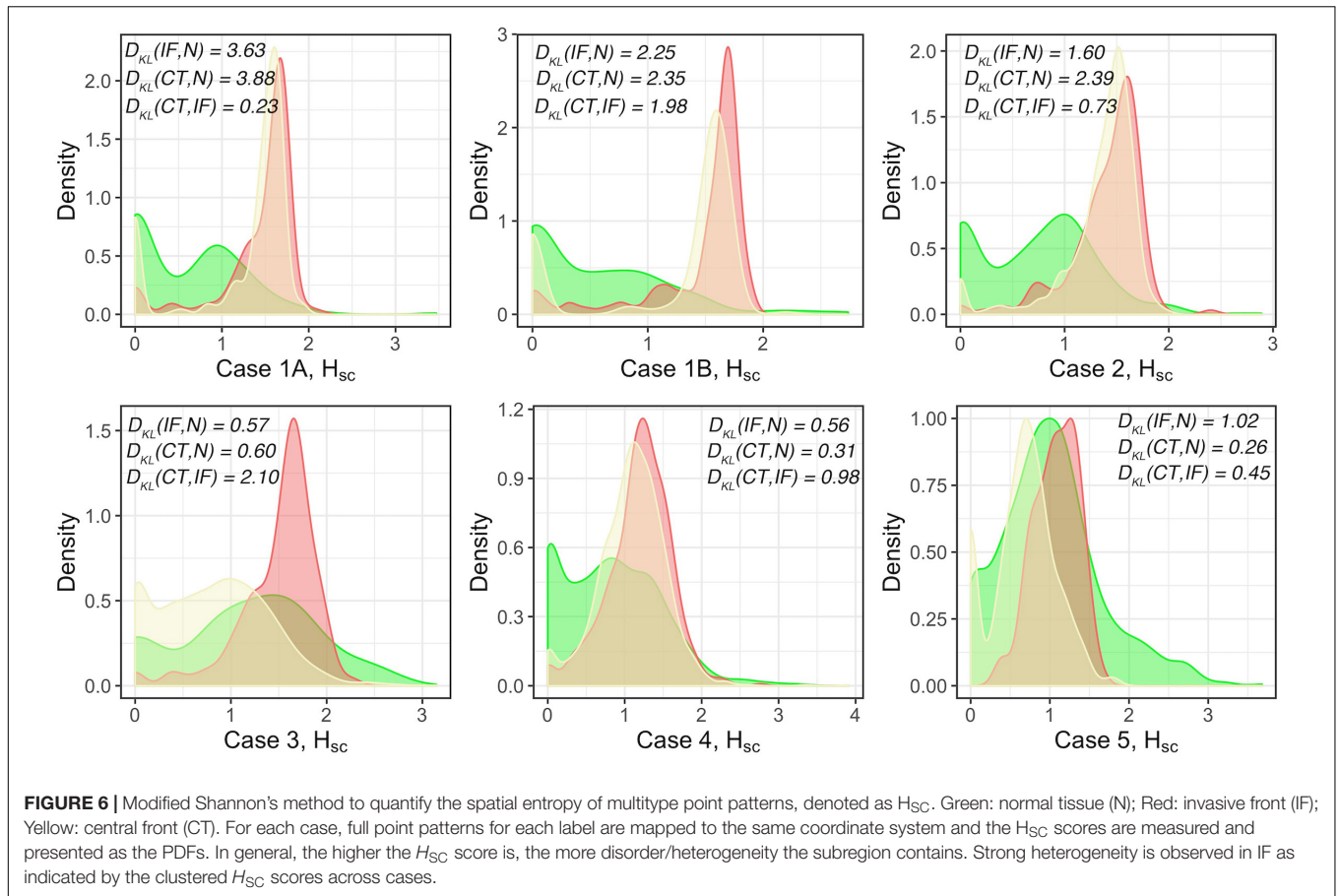
green bars; the plots illustrate a high degree of heterogeneity as bars of different color are interspersed throughout the tissue. As our study focuses on heterogeneity of tumor characteristics, it is important to assess the level of heterogeneity within the IF. **Figure 5E** depicts the cell density distribution of CD4+ T cell plotted as a function of the distance along the middle of the IF; the densities are averaged over the width of the IF of 1 mm. Clearly, the spatial heterogeneity is present not only between different tumor regions, but also within the IF. Again, quadrat test is performed to assess the distribution pattern across all tessellations within each IF. Similarly, tests reject the null hypothesis so that the tumor heterogeneity is also the key factor that dominate the infiltration profiles in IF (**Supplementary Figure S5**). We then depict the immune landscape by visualizing the cell distribution in 3D (**Figure 5F**). For each slide, we map recorded subregion densities to their corresponding locations and generate surface plot with density represented by magnitude and region categories represented by different colors. 3D landscape representation directly depicts the regional density variations. We repeat the analysis for Cases 2–5 for all five labels and the results are presented in **Supplementary Figure S6** (infiltration profiles), **Supplementary Figure S7** (waterfall plots), **Supplementary Figure S8** (infiltration profiles in IF), and **Supplementary Figures S9, S10** (3D plots). Note that the cell density level in Cases 1A and 1B are significantly higher compared to other cases, which may reflect an efficient immune infiltration.

### Spatial Entropy of Multitype Point Patterns

The results above visualize the heterogeneity of cell density distributions within and between the different regions of the specimens. The coefficient of variation CoV is one metric that characterizes the level of heterogeneity. We will also use a spatially adjusted Shannon's entropy as a formal metric of spatial heterogeneity. For each WSI, the point pattern for each subregion is mapped to a Cartesian coordinate system to form a series of multitype point patterns. For each multitype point pattern, the modified Shannon's entropy is calculated, and collective statistics are presented using probability density functions (PDFs, **Figure 6**). The results show clear clustering patterns of the entropy scores around 1.5 over  $H_{SC}$  measurement spaces in IF across all cases. The results indicate that regions with higher entropies are more likely to associate with IF. Biologically, such 'chaos' is possibly driven by the engagement of various components within the TME, namely the spatial intra-tumoral heterogeneity. We also note that while the  $H_{SC}$  scores in CT also appear to cluster in Cases 1A, 1B, 2, 4, and 5, the distribution in Case 3 is comparatively flatten and is similar to N; whereas the  $H_{SC}$  scores in N are normally flatten with lower magnitude, the distribution in Case 5 is intense and sharp. This phenomenon is possibly caused by the different infiltration level of lymphocytes. An efficient immune response can facilitate the recruitment of infiltrating T cells into the battlefield to either fuel the immunosuppression or promote immunoactivation, depending on the recruited T cell subtypes. Once the infiltration barriers (as seen in **Figure 5C** and **Supplementary Figure S6**) are broken, the







immuno-architecture may tend to uniform and mitigate the spatial heterogeneity within TME.

### Spatial Point Pattern Model Fitting

For each subregion (Figures 7A,E), we perform the Complete Spatial Randomness (CSR) to check whether the associated point pattern (Figures 7B,F) follows homogeneous Poisson distribution (Figures 7C,G); if that window failed to pass the CSR test, we further fit the Thomas process model to the point pattern to quantify the clustering (Figures 7D,H). In this study, we either directly use these parameters, e.g.,  $\mu$  as the number of immune marker-labeled cells per cluster, or perform transformation for intuitive interpretation, e.g.,  $\sigma^2$  is used to calculate the average distances of points to cluster areas and center.

As the window is moved through the slide, local features of spatial point patterns are quantified. Collective results are further classified based on their regions; therefore, the variance captures the intra-tumoral heterogeneity. Statistics for each region are also comparable among cases, which can reveal inter-tumoral heterogeneity. In this study, we use the following metrics to characterize each region for each case: mean cell density (counts/mm<sup>2</sup>), average number of cells per cluster, mean distance to cluster core, and mean cluster area. QCoD of the mean values for each parameter across cases are summarized in Table 1. They collectively reflect intra- and inter-tumoral heterogeneity.

### Cell Cluster Distributions and Morphometrics

We also extend our heterogeneity analysis to a global scale by quantifying cell clustering patterns. For each slide, the whole point pattern is clustered using a hierarchical clustering algorithm (Supplementary Figures S11A,B). For each detected cell cluster, we describe and characterize the shape by morphometrics, including convexity, circularity, and eccentricity. We first generate the alpha-hull (Supplementary Figure S11C) and convex hull from the point set that forms the cluster, upon which we derive the convexity and circularity; then we obtain the minor and major axes of the fitted ellipse (Supplementary Figure S11D) that covers 95% of the points of the cluster, from which we derive eccentricity.

Similar to model-fitting analysis, the regional variations of morphometrics capture intra-tumoral heterogeneity, and variations among cases capture inter-tumoral heterogeneity. We use the following metrics to characterize regions, respectively: average density of cells within clusters, average density of clusters, nodular cluster density (Supplementary Figures S11E,G) and elongated cluster density (Supplementary Figures S11F,H). For each label, we use boxplots to show variations between cases, and we perform the Wilcoxon rank-sum test between groups. Figure 8 shows that even though IF is not always distinguished across labels and metrics, the nodular cluster densities and elongated cluster densities in IF are consistently higher than N. We also observe that CD3+ and CD8+ elongated cluster densities in IF are higher than CT. Such unique immuno-architecture may suggest that the elongated cytotoxic T cell cluster density in IF is a potential biomarker for further exploration.

## Correlation Analysis of the Spatial Patterns of Different Immune Markers

The threshold for classification of the degree of colocalization (DoC) scores is determined by comparison of synthesized colocalization pairs. Considering the overall accuracy of our local registration algorithm is 0.976, we shift the point patterns of CD8+ left by 0.01/0.976 mm. This results in the majority ( $\geq 90\%$ ) of DoC for each individual point are larger than 0.84, hereby the threshold is established. In this study, two types of correlations are considered: positive correlation means co-occurrence of points of both channels are likely to be observed in subjects' neighborhood. High DoC scores of points from both channels account for this type; negative correlation means co-occurrence of points of only one channel are likely to be observed in subject's neighborhood. High DoC scores of one channel whereas low DoC scores of the other account for this type. Based on such criteria, we analyze the correlations between CD3+/CD8+, CD4+/FoxP3+, and CD8+/FoxP3+. Results are shown in Figure 9 with Cases 1A, B as an example, and clusters are represented by outlines in different colors.

### Positive Correlations of CD3+ and CD8+ Immune Markers Identify Possible Anti-tumor Hotspots

For each case, the DoC scores, defined by Eq. 14, are assigned to each point of CD3+ and CD8+ markers. Next we use the threshold established above to select points with high DoC scores as candidates for clustering analysis using the same algorithm in cell clustering and morphometrics module (Figures 9A,B), and finally, the detected clusters that contain correlated cells from two channels are defined as hotspots of correlation and the cells within such hotspots are defined as correlated cells (Figure 9C).

We first examined the correlations between CD3+ and CD8+ marker pairs. As Table 2 shows, the numbers of correlated cells differ drastically from case to case, however, the ratios are relatively consistent across cases. The density of hotspots in the IF and CT are significantly higher than normal tissue (N) (Wilcoxon rank-sum  $p = 2.20e-3$  and  $8.70e-3$ ), but there is no difference between IF and CT (Wilcoxon rank-sum  $p = 0.3095$ ). Such distribution pattern of correlation clusters of CD3+ and CD8+ cells indicates possible sites of tumor infiltrate conferring anti-tumor immunity (Figure 9D).

### Positive Correlations of CD4+ and FoxP3+ Immune Markers Identify Possible Immunosuppression Hotspots

The same workflow is repeated for the CD4+ and FoxP3+ pair. A summary of the results is presented in Table 3. Similar patterns are observed when compared to CD3+ and CD8+ statistics. However, the ratios of correlated cells are generally higher than CD3+ and CD8+ pairs. The density of hotspots within IF and CT are significantly higher than N (Wilcoxon rank-sum  $p = 2.20e-3$  and  $8.70e-3$ ), plus the density at IF is also significantly higher than CT (Wilcoxon rank-sum  $p = 4.11e-2$ ).

Collectively, our analysis identifies several hotspots in which CD3+/CD8+ and CD4+/FoxP3+ pairs are

**TABLE 1** | Quartile coefficient of dispersion (QCoD), Eq. 8, for spatial model fitting parameters (range and mean).

Marker	Region	Density	Cell/cluster	Mean distance	Cluster area
CD3	N	0.61–0.82 (0.67)	0.50–0.73 (0.64)	0.12–0.22 (0.16)	0.24–0.41 (0.31)
	IF	0.55–0.71 (0.64)	0.67–0.87 (0.77)	0.15–0.20 (0.17)	0.28–0.38 (0.32)
	CT	0.42–0.68 (0.57)	0.58–0.79 (0.71)	0.16–0.20 (0.18)	0.30–0.39 (0.34)
CD4	N	0.48–0.72 (0.62)	0.58–0.73 (0.64)	0.13–0.19 (0.17)	0.26–0.37 (0.32)
	IF	0.49–0.68 (0.57)	0.64–0.90 (0.76)	0.13–0.18 (0.15)	0.26–0.35 (0.30)
	CT	0.26–0.68 (0.53)	0.46–0.78 (0.65)	0.11–0.21 (0.16)	0.22–0.41 (0.32)
CD8	N	0.51–0.82 (0.61)	0.60–0.77 (0.67)	0.16–0.21 (0.18)	0.31–0.40 (0.35)
	IF	0.48–0.69 (0.59)	0.59–0.83 (0.71)	0.14–0.20 (0.17)	0.28–0.38 (0.33)
	CT	0.53–0.78 (0.64)	0.54–0.75 (0.66)	0.14–0.21 (0.18)	0.28–0.40 (0.35)
CD20	N	0.48–0.68 (0.59)	0.55–0.67 (0.60)	0.14–0.17 (0.15)	0.27–0.32 (0.29)
	IF	0.67–0.76 (0.71)	0.65–0.84 (0.78)	0.14–0.20 (0.17)	0.28–0.38 (0.33)
	CT	0.36–0.70 (0.50)	0.51–0.65 (0.61)	0.13–0.20 (0.16)	0.26–0.38 (0.32)
FoxP3	N	0.42–0.62 (0.48)	0.44–0.88 (0.60)	0.13–0.55 (0.23)	0.25–0.88 (0.42)
	IF	0.52–0.63 (0.58)	0.60–0.85 (0.73)	0.16–0.24 (0.19)	0.31–0.45 (0.36)
	CT	0.36–0.64 (0.50)	0.56–0.76 (0.67)	0.11–0.24 (0.18)	0.21–0.45 (0.35)

co-localized. The majority of identified hotspots are located within the IF. Such hotspots are characterized by strong correlations of CD3+/CD8+ and CD4/FoxP3+. These findings further reveal a strong heterogeneity within IF, as these two T cell subpopulations carry distinct immune characteristics.

### Negative Correlations of CD8+ and FoxP3+ Immune Markers May Identify an Immune Response Landscape

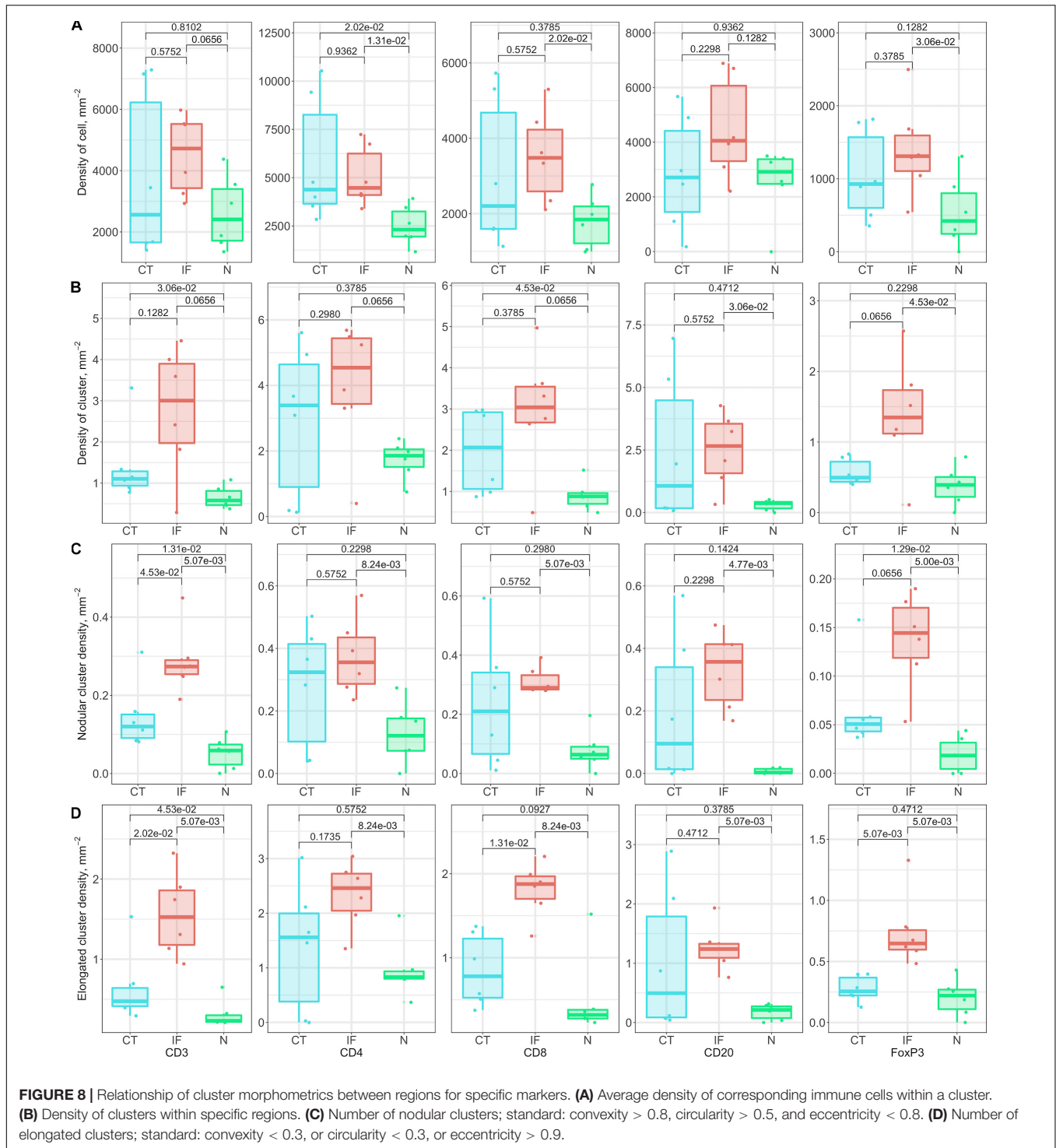
We analyzed the negative correlation of CD8+ and FoxP3+ immune markers as a potential indicator of immune response landscape involving cytotoxic T cells (CTLs) and regulatory T cells (Tregs). In this section, two types of hotspots are defined. First, the hotspot with correlated CD8+ cells and non-correlated FoxP3+ cells; such clusters include FoxP3 cells that are surrounded by CD8+ cells, namely CD8-dominant hotspot. Biologically such hotspots may indicate places where CTLs may efficiently inhibit the strong immunosuppression of Tregs. Second is the opposite type, namely FoxP3-dominant hotspot, where the anti-tumor immunity is possibly impaired by Tregs.

The identification of such landscape can contribute to evaluating the immunotherapy treatment outcomes. Interactions are visualized by identifying negative correlation hotspots of two channels using previous workflow. Unlike the positive hotspots, negative hotspots are widespread throughout the entire tumor tissue for both types. For both types, the density of hotspots at CT and IF (**Supplementary Table S3**) are significantly higher than N (Wilcoxon rank-sum  $p = 1.52e-2$  and  $2.20e-3$ ) but no significant difference is observed between IF and CT (Wilcoxon rank-sum  $p = 0.0649$ ). For FoxP3-dominant hotspots, the CT and IF (**Supplementary Table S4**) are significantly higher than N (Wilcoxon rank-sum  $p = 1.52e-2$  and  $2.20e-3$ ) and the density in IF is also significantly higher than CT (Wilcoxon rank-sum  $p = 2.20e-3$ ). Results are shown in **Supplementary Figure S12** with Cases 1 A, B as an example, and clusters are represented by outlines in different colors.

## DISCUSSION

In this study, we proposed a digital pathology computational workflow to systematically analyze whole slide images (WSI) and quantify the intra- and inter-tumoral heterogeneity through multiple metrics. We analyzed immunohistochemistry (IHC) slides of tumor resections from five patients with TNBC. The sample size is limited to make inferences at the population level, but it is sufficient to look in-depth at each sample with multiple immune markers and to develop a methodology for spatial statistical characterization and build a platform that could be extended to large number of samples, including multiplex IHC and immunofluorescence microscopy (mIF). It should also be noted that each WSI contains enormous amount of information and large numbers of cells of different type to fulfill our need to obtain statistically and biologically meaningful results. In principle, this approach is consistent with personalized medicine where inferences could be made at the level of individual patient. For each patient, five biomarkers: CD3, CD4, CD8, CD20, and FoxP3, were labeled using IHC staining methods. The whole computational workflow starts with image processing: we use cell nucleus segmentation to obtain location information of labeled cells in their original slides and perform image registration using a multimodal protocol to calculate transformation matrices which map all slides to the reference CD4+ slide. We further proposed a pixel-distance based method, and with its application, we can characterize the whole slide into normal tissue (N), IF, and CT. From this point, all subsequent analysis results can be classified into tissue type/region categories to reveal intra-tumoral heterogeneity; for each region, we compare the results between cases for inter-tumoral heterogeneity.

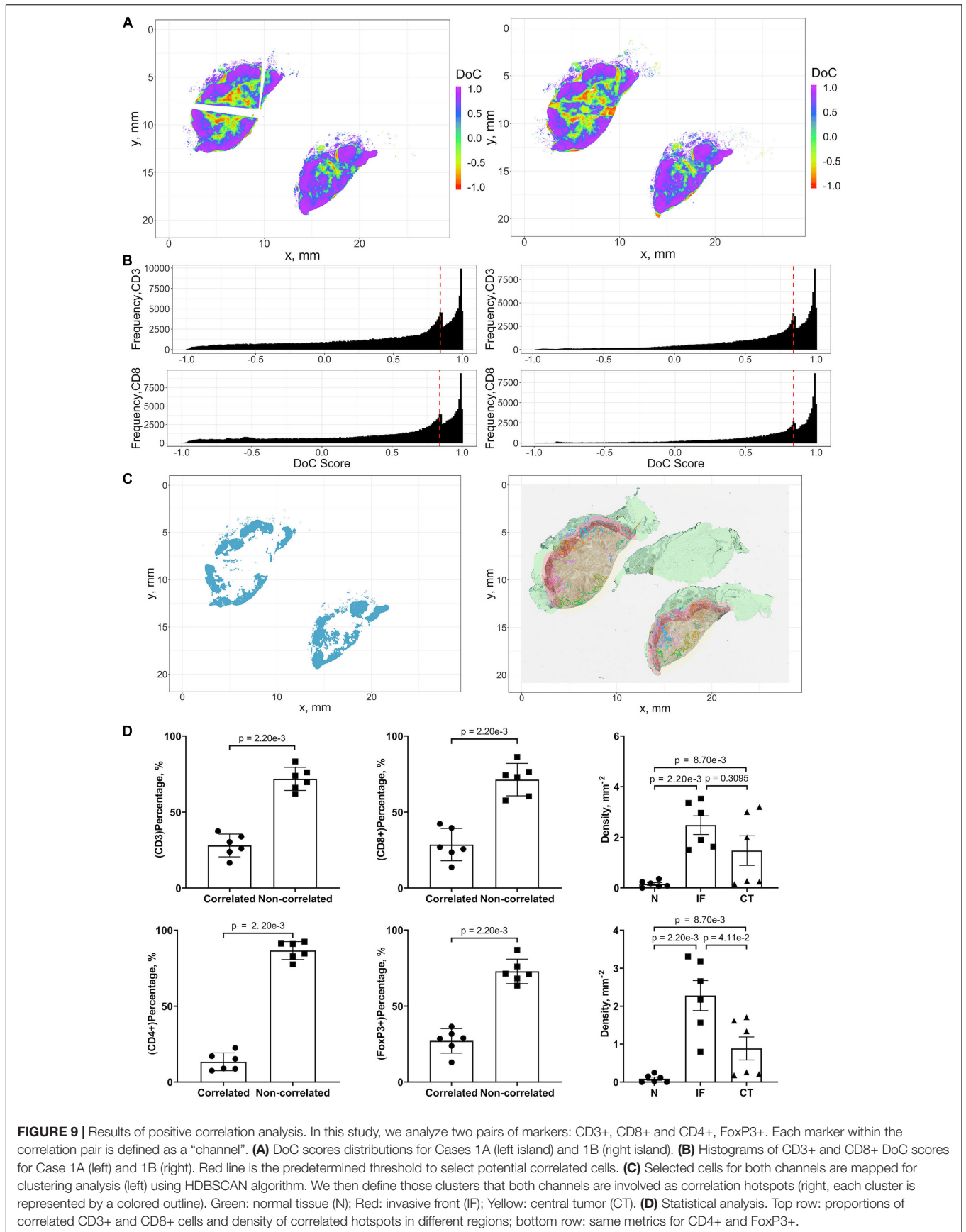
By visualizing the density distributions for each slide and spatial entropy analysis, we identified significant spatial variations of cell densities within and across slides, which qualitatively characterized the intra- and inter-tumoral heterogeneity. In addition, we are particularly interested in the spatial profiles along the direction from N through the IF to the innermost of CT. We observe that for each slide, the cell densities increase sharply within IF and then drop, in some



cases to a plateau and in some precipitously and exhibiting fluctuations; hence corroborating the distinct role of IF in the immuno-architecture. This allows us to propose a hypothesis that the source of the immune cells in the IF is not in the normal tissue into which the tumor grows, but rather the cells extravasate from the tumor vasculature whose density is known to be higher at the rim of the tumor (Stamatelos et al., 2019);

this hypothesis needs to be tested in future studies. We then fit a spatial point process model to data within subregions to capture local variabilities. Statistical results indicate that strong intra- and inter-tumoral heterogeneities co-exist across our study cohort. For each slide, we also evaluated the cell clustering using a hierarchical based algorithm against full point patterns. We then gauged the first-order properties and morphometrics of each





cluster. Results revealed that variations are more likely to occur in CT and IF but less likely in N. We also identified a unique distribution pattern of nodular cytotoxic T cell clusters. As our recent study has shown (Gong et al., 2018), the distribution and shape of clusters have certain relations to the treatment outcomes, thus our findings may lead to predictive biomarkers that could eventually be used clinically when tested on a large number of specimens. Finally, we performed correlation analysis and discovered that the IF is multifaceted and may bear pro- and anti-tumor functions simultaneously, e.g., with higher expressions of CD8+ and FoxP3+ cells.

In addition to characterizing intra- and inter-tumoral heterogeneity, the characteristics obtained from tissue samples, such as spatial cell density profiles of different immune cells, the magnitude of spatial cell density fluctuations, and the spatial correlations between the densities of different immune cell types, can facilitate development and parameterization of spatially resolved computational immuno-oncology models. Recently, QSP models have been applied to immuno-oncology research as a platform for conducting virtual clinical trials (Cheng et al., 2017; Bai et al., 2019; Jafarnejad et al., 2019; Milberg et al., 2019; Ma et al., 2020). These models capture system scale behavior in cancer patients and are capable of population level predictions of disease trajectories in response to intervention. On tissue-cellular scale, ABMs have been employed and used for spatially explicit simulations to investigate emergent behavior arising from interactions between cancer and immune cells, such as spatial and spatio-temporal variations in tumor morphology

and immuno-architecture (Kim et al., 2009; Shi et al., 2014; Wells et al., 2015; Gong et al., 2017; Norton et al., 2017, 2019; Pourhasanzade et al., 2017; Hoehme et al., 2018; Ji et al., 2019). When combining QSP models with ABM, cancer models can be further enhanced by taking advantage of both model types: while the QSP module captures whole-body temporal dynamics including lymph nodes, blood, peripheral compartment, and tumor, ABM simulation accounts for crucial aspects of high-granularity features such as cancer cell clonal evolution and TME heterogeneity. The resulting hybrid model will be able to closely track and predict the course of cancer development, both primary tumors and metastases, and potentially during treatment in individual patients by incorporating patient-specific TME characteristics, which can be quantified using our digital pathology platform. Such synergy would enable a better understanding of impact of spatial heterogeneities in the CT and IF on the pathophysiological parameters and variables. Strictly speaking, cell densities calculated directly from digitally segmented pathology images as described in this study represent the number of cell profiles per unit area in the tissue slide typically with a 4–5 micron thickness, which is a common metric in pathology, where a cell signature is a section of cell with an area larger than the detectable threshold set in our segmentation algorithm. However, in computational models the cell concentrations are usually represented as the number of cells per unit volume rather than unit area. Using methods from the field of stereology (Weibel et al., 1966), 3D numerical densities ( $N_V$ ) can be estimated from 2D density ( $N_A$ ) using

**TABLE 2** | Statistical summary for CD3+ and CD8+ immune markers correlation analysis.

Case	QDoC (DoC Score)		Correlated cell counts		Percentage, %		Cluster density, mm <sup>-2</sup>		
	CD3	CD8	CD3	CD8	CD3	CD8	N	IF	CT
1A	0.95	0.94	71,630	66,097	23.9	25.7	0.36	3.36	2.00
1B	0.35	0.29	61,342	53,715	33.9	39.6	0.24	2.97	3.20
2	0.15	0.15	91,147	70,859	37.5	42.3	0.08	3.53	3.00
3	0.29	0.39	25,773	17,804	30.3	26.9	0	1.90	0.25
4	0.24	0.30	26,340	22,180	26.1	23.5	0.07	1.63	0.27
5	0.61	0.91	11,672	7,214	16.7	13.7	0.18	1.51	0.15

*N*, normal tissue; *IF*, invasive front; *CT*, central tumor.

**TABLE 3** | Statistical summary for CD4+ and FoxP3+ immune markers correlation analysis.

Case	QDoC (DoC Score)		Correlated cell counts		Percentage, %		Cluster density, mm <sup>-2</sup>		
	CD4	FoxP3	CD4	FoxP3	CD4	FoxP3	N	IF	CT
1A	0.68	0.60	83,363	17,902	17.1	29.0	0.25	3.18	1.71
1B	0.40	0.36	65,267	14,719	22.5	36.5	0.12	3.31	1.62
2	0.52	0.53	59,281	15,630	15.3	23.9	0.02	2.17	1.34
3	1.30	0.39	25,172	12,085	9.0	31.8	0	2.66	0.26
4	0.85	0.90	19,385	6,547	8.8	13.0	0	0.80	0.21
5	0.87	0.49	18,596	7,586	7.5	28.6	0.15	1.57	0.18

*N*, normal tissue; *IF*, invasive front; *CT*, central tumor.

the following equation:  $N_V = N_A/(t + D - 2h)$ , where  $t$  is the thickness of the section,  $D$  is the diameter of stained cells (which are lymphocytes in the scope of this study), and  $h$  is the minimum height of detectable spherical cap (which can be derived from cellular segmentation algorithm parameters) (Royet, 1991). In this equation,  $2h$  in the denominator accounts for loss of undetected parts of the cell.  $N_V$  indicates number of cells per unit volume and can directly be used to inform 3D spatial models of tumor-immune interactions. Using this equation, one could convert the 2D densities (in  $\text{mm}^{-2}$ ) to 3D densities (in  $\text{mm}^{-3}$ ); in the conversion the slide thickness is typically  $t = 4.5\text{--}5\ \mu\text{m}$ ,  $h = D/2 - \sqrt{(D/2)^2 - A_{crit}/\pi}$ , where  $A_{crit}$  is the minimum area detectable during the segmentation, typically  $\sim 10\ \mu\text{m}^2$ ; diameter values reported for T cell ( $5\text{--}7.1\ \mu\text{m}$ ) and for B cell ( $5.5\text{--}9\ \mu\text{m}$ ) are also necessary for the conversion (Chapman et al., 1981; Turgeon, 2005; Tsourkas et al., 2007; Strokovov et al., 2009; El Hentati et al., 2010; Mrozek-Gorska et al., 2019; Renner et al., 2020).

Depending on the purpose of each computer simulation, one can either derive overall 3D density and use it to populate the *in silico* TME; or if spatial heterogeneity is of interest, the variability of cell density can be taken into account by sampling multiple  $N_A$  from different regions of the digital pathology analysis output to initiate the simulated TME with a range of  $N_V$  values in space. After simulation, the same methods employed in this study to analyze spatial correlations between different cell types can be applied to virtual sections of model-generated three-dimensional tumor, which would enable quantitative comparisons between model-generated spatial patterns of cancer and immune cells and patient pathology images. QSP and ABM have been used to model the tumor growth and invasion of several cancer types, such as melanoma (Wang et al., 2013; Milberg et al., 2019), breast cancer (Bates et al., 2006; Bianca and Pennisi, 2012), colorectal (Kather et al., 2017), and non-small cell lung cancer (Jafarnejad et al., 2019).

Future work should focus on increasing the scale of the current workflow. In this study, cells expressing CD8/FoxP3 are considered as cytotoxic/regulatory T cells. Such loose criterion serves the need to test the functionality of the workflow using preliminary computational results from pathology images. However, a comprehensive biomarker panel is required to account for the complexity in cell lineage definition and further to characterize the components in TME. Such materials will be obtained in subsequent studies applied to multiplex labeled specimens. Improving the performance of image processing is another critical issue. We recognize the power of artificial intelligence in digital pathology and such techniques could be incorporated in an extension of the workflow. For example, traditional segmentation algorithms may not adequately distinguish cell boundaries due to staining issues. A possibility is to introduce convolutional neural network (CNN) trained on well-defined ground-truth (Khosravi et al., 2018). In this study, immune markers are stained on consecutive slices of tumor resections, therefore artifacts may be introduced such as distortions. Registration reduces uncertainty introduced by these artifacts, but cannot fully compensate for the location mismatch, and errors may be introduced in derived point

patterns and subsequent analysis. Such problems can be alleviated by harvesting data from multiplex images in the first place by labeling different cells and molecules on the same slide; in this case artificial location shifts, sample folds and z-axis differences are essentially eliminated. In the tissue type characterization step, we identified the IF by averaging annotations provided by expert pathologist. To pinpoint IF, deep learning methods can be applied for automated tissue segmentation. In point pattern analysis stage, we extract copious intra- and inter-tumoral heterogeneity information from collective slides; when correlated with treatment outcomes, these results can provide more useful information for pathologists and immuno-oncologists.

## DATA AVAILABILITY STATEMENT

The datasets generated or analyzed during this work are available from the corresponding author on reasonable request. The codes for computational methods are made available at <https://github.com/popellab/SpatHeterogeneity-TNBC>. The open source software QuPath may be downloaded at <https://qupath.github.io/>. The open source software Blender may be downloaded at <https://www.blender.org/download/>. The open source software Icy may be downloaded at <http://icy.bioimageanalysis.org/download/>.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Institutional Review Board of the Johns Hopkins Medical Institutions. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

## AUTHOR CONTRIBUTIONS

HM, CG, and AP designed the workflow. HM processed the images, implemented the statistical tests, and produced the results. AC-M provided the specimens, assisted in the interpretation of results, and edited the manuscript. HM, CG, JS, EF, AS, EJ, VS, LE, AC-M, and AP critically edited the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

The authors gratefully acknowledge financial support from National Institutes of Health (Grant Nos. R01CA138264 and U01CA212007).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2020.583333/full#supplementary-material>

## REFERENCES

- Acs, B., Ahmed, F. S., Gupta, S., Wong, P. F., Gartrell, R. D., Pradhan, J. S., et al. (2019a). An open source automated tumor infiltrating lymphocyte algorithm for prognosis in melanoma. *Nat. Commun.* 10, 1–7. doi: 10.1038/s41467-019-13043-2
- Acs, B., Pelekanou, V., Bai, Y., Martinez-Morilla, S., Toki, M., Leung, S. C., et al. (2019b). Ki67 reproducibility using digital image analysis: an inter-platform and inter-operator study. *Lab. Invest.* 99, 107–117. doi: 10.1038/s41374-018-0123-7
- Al-Janabi, S., Huisman, A., and Van Diest, P. J. (2012). Digital pathology: current status and future perspectives. *Histopathology* 61, 1–9. doi: 10.1111/j.1365-2559.2011.03814.x
- Altan, M., Kidwell, K. M., Pelekanou, V., Carvajal-Hausdorf, D. E., Schalper, K. A., Toki, M. I., et al. (2018). Association of B7-H4, PD-L1, and tumor infiltrating lymphocytes with outcomes in breast cancer. *NPJ Breast Cancer* 4:40. doi: 10.1038/s41523-018-0095-1
- Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R. R Package Version 1.64-1*. Available online at: <https://cran.r-project.org/web/packages/spatstat/index.html> (accessed September 30, 2020).
- Bai, J. P., Earp, J. C., and Pillai, V. C. (2019). Translational quantitative systems pharmacology in drug development: from current landscape to good practices. *AAPS J.* 21:72. doi: 10.1208/s12248-019-0339-5
- Bankhead, P., Fernández, J. A., Mcart, D. G., Boyle, D. P., Li, G., Loughrey, M. B., et al. (2018). Integrated tumor identification and automated scoring minimizes pathologist involvement and provides new insights to key biomarkers in breast cancer. *Lab. Invest.* 98, 15–26. doi: 10.1038/labinvest.2017.131
- Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., Mcart, D. G., Dunne, P. D., et al. (2017). QuPath: open source software for digital pathology image analysis. *Sci. Rep.* 7, 1–7. doi: 10.1038/s41598-017-17204-5
- Barua, S., Solis, L., Parra, E. R., Uraoka, N., Jiang, M., Wang, H., et al. (2018). A functional spatial analysis platform for discovery of immunological interactions predictive of low-grade to high-grade transition of pancreatic intraductal papillary mucinous neoplasms. *Cancer Inform.* 17:1176935118782880. doi: 10.1177/1176935118782880
- Bates, G. J., Fox, S. B., Han, C., Leek, R. D., Garcia, J. F., Harris, A. L., et al. (2006). Quantification of regulatory T cells enables the identification of high-risk breast cancer patients and those at risk of late relapse. *J. Clin. Oncol.* 24, 5373–5380. doi: 10.1200/JCO.2006.05.9584
- Beck, A. H., Sangoi, A. R., Leung, S., Marinelli, R. J., Nielsen, T. O., Van De Vijver, M. J., et al. (2011). Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci. Transl. Med.* 3:108ra113. doi: 10.1126/scitranslmed.3002564
- Berben, L., Wildiers, H., Marcelis, L., Martinez, A. A., Bosisio, F., Hatse, S., et al. (2020). Computerized scoring protocol for identification and quantification of different immune cell populations in breast tumor regions using QuPath software. *Histopathology* 77, 79–91. doi: 10.1111/his.14108
- Bianca, C., and Pennisi, M. (2012). The triplex vaccine effects in mammary carcinoma: a nonlinear model in tune with SimTriplex. *Nonl. Anal.* 13, 1913–1940. doi: 10.1016/j.nonrwa.2011.12.019
- Bivand, R., and Rundel, C. (2017). *rgeos: Interface to Geometry Engine-Open Source (GEOS). R Package Version 0.3-23*. Available online at: <https://cran.r-project.org/web/packages/rgeos/index.html> (accessed September 30, 2020).
- Blagih, J., Zani, F., Chakravarty, P., Hennequart, M., Pilley, S., Hobor, S., et al. (2020). Cancer-specific loss of p53 leads to a modulation of myeloid and T cell responses. *Cell Rep.* 30, 481–496. doi: 10.1016/j.celrep.2019.12.028
- Breiman, L. (2001). Random forests. *Machine Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Brown, J. R., Wimberly, H., Lannin, D. R., Nixon, C., Rimm, D. L., and Bossuyt, V. (2014). Multiplexed quantitative analysis of CD3, CD8, and CD20 predicts response to neoadjuvant chemotherapy in breast cancer. *Clin. Cancer Res.* 20, 5995–6005. doi: 10.1158/1078-0432.CCR-14-1622
- Chapman, E. H., Kurec, A. S., and Davey, F. (1981). Cell volumes of normal and malignant mononuclear cells. *J. Clin. Pathol.* 34, 1083–1090. doi: 10.1136/jcp.34.10.1083
- Cheng, Y., Thalhauser, C. J., Smithline, S., Pagidala, J., Miladinov, M., Vezina, H. E., et al. (2017). QSP toolbox: computational implementation of integrated workflow components for deploying multi-scale mechanistic models. *AAPS J.* 19, 1002–1016. doi: 10.1208/s12248-017-0100-x
- Cimino-Mathews, A., Thompson, E., Taube, J. M., Ye, X., Lu, Y., Meeker, A., et al. (2016). PD-L1 (B7-H1) expression and the immune tumor microenvironment in primary and metastatic breast carcinomas. *Hum. Pathol.* 47, 52–63. doi: 10.1016/j.humpath.2015.09.003
- Claramunt, C. (2005). “A spatial form of diversity,” in *Lecture Notes in Computer Science*, eds A. G. Cohn, and D. M. Mark, (Berlin: Springer), 218–231. doi: 10.1007/11556114\_14
- De Chaumont, F., Dallongeville, S., Chenouard, N., Hervé, N., Pop, S., Provoost, T., et al. (2012). Icy: an open bioimage informatics platform for extended reproducible research. *Nat. methods* 9:690. doi: 10.1038/nmeth.2075
- Denkert, C., Von Minckwitz, G., Darb-Esfahani, S., Lederer, B., Heppner, B. I., Weber, K. E., et al. (2018). Tumour-infiltrating lymphocytes and prognosis in different subtypes of breast cancer: a pooled analysis of 3771 patients treated with neoadjuvant therapy. *Lancet Oncol.* 19, 40–50. doi: 10.1016/S1470-2045(17)30904-X
- Du, Z., Lin, J.-R., Rashid, R., Maliga, Z., Wang, S., Aster, J. C., et al. (2019). Qualifying antibodies for image-based immune profiling and multiplexed tissue imaging. *Nat. Protoc.* 14, 2900–2930. doi: 10.1038/s41596-019-0206-y
- Efron, B. (1981). Nonparametric standard errors and confidence intervals. *Can. J. Stat.* 9, 139–158. doi: 10.2307/3314608
- El Hentati, F. Z., Gruy, F., Iobagiu, C., and Lambert, C. (2010). Variability of CD3 membrane expression and T cell activation capacity. *Cytom. Part B Clin. Cytom.* 78, 105–114. doi: 10.1002/cyto.b.20496
- Ferré, E. M., Break, T. J., Burbelo, P. D., Allgauer, M., Kleiner, D. E., Jin, D., et al. (2019). Lymphocyte-driven regional immunopathology in pneumonitis caused by impaired central immune tolerance. *Sci. Transl. Med.* 11:eav5597. doi: 10.1126/scitranslmed.aav5597
- Fridman, W. H., Galon, J., Pagès, F., Tartour, E., Sautès-Fridman, C., and Kroemer, G. (2011). Prognostic and predictive impact of intra- and peritumoral immune infiltrates. *Cancer Res.* 71, 5601–5605. doi: 10.1158/0008-5472.CAN-11-1316
- Gombin, J., Vaidyanathan, R., and Agafonik, V. (2017). *Concaveman: A Very Fast 2D Concave Hull Algorithm. R Package Version 1.1.0*. Available online at: <https://cran.r-project.org/web/packages/concaveman/index.html> (accessed September 30, 2020).
- Gong, C., Anders, R. A., Zhu, Q., Taube, J. M., Green, B., Cheng, W., et al. (2018). Quantitative characterization of CD8+ T cell clustering and spatial heterogeneity in solid tumors. *Front. Oncol.* 8:649. doi: 10.3389/fonc.2018.00649
- Gong, C., Milberg, O., Wang, B., Vicini, P., Narwal, R., Roskos, L., et al. (2017). A computational multiscale agent-based model for simulating spatio-temporal tumour immune response to PD1 and PDL1 inhibition. *J. R. Soc. Interface* 14, 20170320. doi: 10.1098/rsif.2017.0320
- Guy, C. L., Weiss, E., Che, S., Jan, N., Zhao, S., and Rosu-Bubulac, M. (2019). Evaluation of image registration accuracy for tumor and organs at risk in the thorax for compliance with TG 132 recommendations. *Adv. Radiat. Oncol.* 4, 177–185. doi: 10.1016/j.adro.2018.08.023
- Haanen, J., Baars, A., Gomez, R., Weder, P., Smits, M., De Gruijl, T., et al. (2006). Melanoma-specific tumor-infiltrating lymphocytes but not circulating melanoma-specific T cells may predict survival in resected advanced-stage melanoma patients. *Cancer Immunol. Immunother.* 55, 451–458. doi: 10.1007/s00262-005-0018-5
- Habets, R. A., De Bock, C. E., Serneels, L., Lodewijckx, I., Verbeke, D., Nittner, D., et al. (2019). Safe targeting of T cell acute lymphoblastic leukemia by pathology-specific NOTCH inhibition. *Sci. Transl. Med.* 11:eau6246. doi: 10.1126/scitranslmed.aau6246
- Halama, N., Michel, S., Kloor, M., Zoernig, I., Benner, A., Spille, A., et al. (2011). Localization and density of immune cells in the invasive margin of human colorectal cancer liver metastases are prognostic for response to chemotherapy. *Cancer Res.* 71, 5670–5677. doi: 10.1158/0008-5472.CAN-11-0268
- Hendry, S., Salgado, R., Gevaert, T., Russell, P. A., John, T., Thapa, B., et al. (2017). Assessing tumor infiltrating lymphocytes in solid tumors: a practical review for pathologists and proposal for a standardized method from the International Immuno-Oncology Biomarkers Working Group: Part 2: TILs in melanoma, gastrointestinal tract carcinomas, non-small cell lung carcinoma and mesothelioma, endometrial and ovarian carcinomas, squamous cell



- carcinoma of the head and neck, genitourinary carcinomas, and primary brain tumors. *Adv. Anat. Pathol.* 24:311. doi: 10.1097/PAP.0000000000000161
- Hess, R. (2007). *The Essential Blender: Guide to 3D Creation with the Open Source Suite Blender*. San Francisco, CA: No Starch Press.
- Hoehme, S., Bertaux, F., Weens, W., Grasl-Kraupp, B., Hengstler, J. G., and Drasdo, D. (2018). Model prediction and validation of an order mechanism controlling the spatiotemporal phenotype of early hepatocellular carcinoma. *Bull. Math. Biol.* 80, 1134–1171. doi: 10.1007/s11538-017-0375-1
- Jafarnejad, M., Gong, C., Gabrielson, E., Bartelink, I. H., Vicini, P., Wang, B., et al. (2019). A computational model of neoadjuvant PD-1 inhibition in non-small cell lung cancer. *AAPS J.* 21:79. doi: 10.1208/s12248-019-0350-x
- Ji, Z., Zhao, W., Lin, H.-K., and Zhou, X. (2019). Systematically understanding the immunity leading to CRCP progression. *PLoS Comput. Biol.* 15:e1007344. doi: 10.1371/journal.pcbi.1007344
- Kather, J. N., Hörner, C., Weis, C.-A., Aung, T., Vokuhl, C., Weiss, C., et al. (2019). CD163+ immune cell infiltrates and presence of CD54+ microvessels are prognostic markers for patients with embryonal rhabdomyosarcoma. *Sci. Rep.* 9, 1–12. doi: 10.1038/s41598-019-45551-y
- Kather, J. N., Poleszczuk, J., Suarez-Carmona, M., Krisam, J., Charoentong, P., Valous, N. A., et al. (2017). In silico modeling of immunotherapy and stroma-targeting therapies in human colorectal cancer. *Cancer Res.* 77, 6442–6452. doi: 10.1158/0008-5472.CAN-17-2006
- Kawai, O., Ishii, G., Kubota, K., Murata, Y., Naito, Y., Mizuno, T., et al. (2008). Predominant infiltration of macrophages and CD8+ T cells in cancer nests is a significant predictor of survival in stage IV nonsmall cell lung cancer. *Cancer* 113, 1387–1395. doi: 10.1002/cncr.23712
- Khosravi, P., Kazemi, E., Imielinski, M., Elemento, O., and Hajirasouliha, I. (2018). Deep convolutional neural networks enable discrimination of heterogeneous digital pathology images. *EBioMedicine* 27, 317–328. doi: 10.1016/j.ebiom.2017.12.026
- Kim, P. S., Levy, D., and Lee, P. P. (2009). Modeling and simulation of the immune system as a self-regulating network. *Methods Enzymol.* 467, 79–109. doi: 10.1016/S0076-6879(09)67004-X
- Klemm, F., and Joyce, J. A. (2015). Microenvironmental regulation of therapeutic response in cancer. *Trends Cell Biol.* 25, 198–213. doi: 10.1016/j.tcb.2014.11.006
- Ladányi, A. (2015). Prognostic and predictive significance of immune cells infiltrating cutaneous melanoma. *Pigment Cell Melanoma Res.* 28, 490–500. doi: 10.1111/pcmr.12371
- Li, X., Gruosso, T., Zuo, D., Omeroglu, A., Meterissian, S., Guiot, M.-C., et al. (2019). Infiltration of CD8+ T cells into tumor cell clusters in triple-negative breast cancer. *Proc. Natl. Acad. Sci. U.S.A.* 116, 3678–3687. doi: 10.1073/pnas.1817652116
- Lianyan, T., Dianrong, X., Chunhui, Y., Zhaolai, M., and Bin, J. (2018). The predictive value and role of stromal tumor-infiltrating lymphocytes in pancreatic ductal adenocarcinoma (PDAC). *Cancer Biol. Ther.* 19, 296–305. doi: 10.1080/15384047.2017.1416932
- Long, G. V., Weber, J. S., Infante, J. R., Kim, K. B., Daud, A., Gonzalez, R., et al. (2016). Overall survival and durable responses in patients with BRAF V600-mutant metastatic melanoma receiving dabrafenib combined with trametinib. *J. Clin. Oncol.* 34, 871–878. doi: 10.1200/JCO.2015.62.9345
- Ma, H., Wang, H., Sove, R. J., Jafarnejad, M., Tsai, C.-H., Wang, J., et al. (2020). A quantitative systems pharmacology model of T cell engager applied to solid tumor. *AAPS J.* 22:85. doi: 10.1208/s12248-020-00450-3
- Mani, N. L., Schalper, K. A., Hatzis, C., Saglam, O., Tavassoli, F., Butler, M., et al. (2016). Quantitative assessment of the spatial heterogeneity of tumor-infiltrating lymphocytes in breast cancer. *Breast Cancer Res.* 18:78. doi: 10.1186/s13058-016-0737-x
- McDonald, T. (2016). *SDraw: Spatially Balanced Sample Draws for Spatial Objects*. R Package Version 2.1.13. Available online at: <https://cran.r-project.org/web/packages/SDraw/index.html> (accessed September 30, 2020).
- Melosky, B., Juergens, R., Hirsh, V., Mcleod, D., Leighl, N., Tsao, M. S., et al. (2020). Amplifying outcomes: checkpoint inhibitor combinations in first-line non-small cell lung cancer. *Oncologist* 25:64. doi: 10.1634/theoncologist.2019-0027
- Milberg, O., Gong, C., Jafarnejad, M., Bartelink, I. H., Wang, B., Vicini, P., et al. (2019). A QSP model for predicting clinical responses to monotherapy, combination and sequential therapy following CTLA-4, PD-1, and PD-L1 checkpoint blockade. *Sci. Rep.* 9, 1–17. doi: 10.1038/s41598-019-47802-4
- Motzer, R. J., Penkov, K., Haanen, J., Rini, B., Albiges, L., Campbell, M. T., et al. (2019). Avelumab plus axitinib versus sunitinib for advanced renal-cell carcinoma. *N. Engl. J. Med.* 380, 1103–1115. doi: 10.1056/NEJMoa1816047
- Mrozek-Gorska, P., Buschle, A., Pich, D., Schwarzmayr, T., Fechtner, R., Scialdone, A., et al. (2019). Epstein–Barr virus reprograms human B lymphocytes immediately in the prelatent phase of infection. *Proc. Natl. Acad. Sci. U.S.A.* 116, 16046–16055. doi: 10.1073/pnas.1901314116
- Norton, K.-A., Gong, C., Jamalian, S., and Popel, A. S. (2019). Multiscale agent-based and hybrid modeling of the tumor immune microenvironment. *Processes* 7:37. doi: 10.3390/pr7010037
- Norton, K.-A., Jin, K., and Popel, A. S. (2018). Modeling triple-negative breast cancer heterogeneity: effects of stromal macrophages, fibroblasts and tumor vasculature. *J. Theor. Biol.* 452, 56–68. doi: 10.1016/j.jtbi.2018.05.003
- Norton, K.-A., Wallace, T., Pandey, N. B., and Popel, A. S. (2017). An agent-based model of triple-negative breast cancer: the interplay between chemokine receptor CCR5 expression, cancer stem cells, and hypoxia. *BMC Syst. Biol.* 11:68. doi: 10.1186/s12918-017-0445-x
- Pageon, S. V., Nicovich, P. R., Mollazade, M., Tabarin, T., and Gaus, K. (2016). Clus-DoC: a combined cluster detection and colocalization analysis for single-molecule localization microscopy data. *Mol. Biol. Cell* 27, 3627–3636. doi: 10.1091/mbc.E16-07-0478
- Pages, F., Kirilovsky, A., Mlecnik, B., Asslaber, M., Tosolini, M., Bindea, G., et al. (2009). In situ cytotoxic and memory T cells predict outcome in patients with early-stage colorectal cancer. *J. Clin. Oncol.* 27, 5944–5951. doi: 10.1200/JCO.2008.19.6147
- Pebesma, E., and Bivand, R. S. (2005). *S classes and Methods for Spatial Data: The Sp Package*. R Package Version 1.4-2. Available online at: <https://cran.r-project.org/web/packages/sp/index.html> (accessed September 30, 2020).
- Pourhasanzade, F., Sabzpooshan, S., Alizadeh, A. M., and Esmati, E. (2017). An agent-based model of avascular tumor growth: Immune response tendency to prevent cancer development. *Simulation* 93, 641–657. doi: 10.1177/0037549717699072
- Renner, K., Bruss, C., and Färber, S. (2020). *Monitoring Human T Cell Activation in the Context of Immunotherapeutic Approaches*. Bremen, Germany: OLS OMNI Life Science GmbH & Co KG. Available online at: [https://www.bulldog-bio.com/wp-content/uploads/2020/06/T-cell-activation-CASY-AppNote\\_OLS.pdf](https://www.bulldog-bio.com/wp-content/uploads/2020/06/T-cell-activation-CASY-AppNote_OLS.pdf)
- Rieger, T. R., Allen, R. J., Bystricky, L., Chen, Y., Colopy, G. W., Cui, Y., et al. (2018). Improving the generation and selection of virtual populations in quantitative systems pharmacology models. *Prog. Biophys. Mol. Biol.* 139, 15–22. doi: 10.1016/j.pbiomolbio.2018.06.002
- Rodríguez Casal, A., and Pateiro López, B. (2010). *Generalizing the Convex Hull of a Sample: the R Package Alphahull*. R Package Version 2.2. Available online at: <https://cran.r-project.org/web/packages/alphahull/index.html> (accessed September 30, 2020).
- Royet, J.-P. (1991). Stereology: a method for analyzing images. *Prog. Neurobiol.* 37, 433–474. doi: 10.1016/0301-0082(91)90009-P
- Santiago, I., Santinha, J., Ianus, A., Galzerano, A., Theias, R., Maia, J., et al. (2019). Susceptibility perturbation MRI maps tumor infiltration into mesorectal lymph nodes. *Cancer Res.* 79, 2435–2444. doi: 10.1158/0008-5472.CAN-18-3682
- Schmid, P., Rugo, H. S., Adams, S., Schneeweiss, A., Barrios, C. H., Iwata, H., et al. (2020). Atezolizumab plus nab-paclitaxel as first-line treatment for unresectable, locally advanced or metastatic triple-negative breast cancer (IMpassion130): updated efficacy results from a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet Oncol.* 21, 44–59. doi: 10.1016/S1470-2045(19)30689-8
- Schwen, L. O., Andersson, E., Korski, K., Weiss, N., Haase, S., Gaire, F., et al. (2018). Data-driven discovery of immune contexture biomarkers. *Front. Oncol.* 8:627. doi: 10.3389/fonc.2018.00627
- Shi, Z. Z., Wu, C.-H., and Ben-Arieh, D. (2014). Agent-based model: a surging tool to simulate infectious diseases in the immune system. *Open J. Model. Simul.* 2:12. doi: 10.4236/ojmsi.2014.21004
- Sová, R. J., Jafarnejad, M., Zhao, C., Wang, H., Ma, H., and Popel, A. S. (2020). QSP–IO: a quantitative systems pharmacology toolbox for mechanistic multi-scale modeling for immuno-oncology applications. *CPT Pharmacometrics Syst. Pharmacol.* doi: 10.1002/psp4.12546 [Epub ahead of print].

- Stamatelos, S. K., Bhargava, A., Kim, E., Popel, A. S., and Pathak, A. P. (2019). Tumor ensemble-based modeling and visualization of emergent angiogenic heterogeneity in breast cancer. *Sci. Rep.* 9, 1–14. doi: 10.1038/s41598-019-40888-w
- Strokotov, D. I., Yurkin, M. A., Gilev, K. V., Van Bockstaele, D. R., Hoekstra, A. G., Rubtsov, N., et al. (2009). Is there a difference between T-and B-lymphocyte morphology? *J. Biomed. Optics* 14:064036. doi: 10.1117/1.3275471
- Suzuki, H., Chikazawa, N., Tasaka, T., Wada, J., Yamasaki, A., Kitaura, Y., et al. (2010). Intratumoral CD8+ T/FOXP3+ cell ratio is a predictive marker for survival in patients with colorectal cancer. *Cancer Immunol. Immunother.* 59, 653–661. doi: 10.1007/s00262-009-0781-9
- Tanaka, U., Ogata, Y., and Stoyan, D. (2008). Parameter estimation and model selection for Neyman–Scott point processes. *Biom. J.* 50, 43–57. doi: 10.1002/bimj.200610339
- Tang, J., Liu, J., Zhang, M., and Mei, Q. (2016). “Visualizing large-scale and high-dimensional data,” in *Proceedings of the 25th International Conference on World Wide Web*, ed. J. Bourdeau, (New York, NY: Association for Computing Machinery), 287–297. doi: 10.1145/2872427.2883041
- Thomas, M. (1949). A generalization of Poisson’s binomial limit for use in ecology. *Biometrika* 36, 18–25. doi: 10.2307/2332526
- Tsakiroglou, A. M., Fergie, M., Oguejiofor, K., Linton, K., Thomson, D., Stern, P. L., et al. (2020). Spatial proximity between T and PD-L1 expressing cells as a prognostic biomarker for oropharyngeal squamous cell carcinoma. *Br. J. Cancer* 122, 539–544. doi: 10.1038/s41416-019-0634-z
- Tsourkas, P. K., Baumgarth, N., Simon, S. L., and Raychaudhuri, S. (2007). Mechanisms of B-cell synapse formation predicted by Monte Carlo simulation. *Biophys. J.* 92, 4196–4208. doi: 10.1529/biophysj.106.094995
- Turgeon, M. L. (2005). *Clinical Hematology: Theory and Procedures*. Philadelphia, PA: Lippincott Williams & Wilkins.
- Tyekucheva, S., Bowden, M., Bango, C., Giunchi, F., Huang, Y., Zhou, C., et al. (2017). Stromal and epithelial transcriptional map of initiation progression and metastatic potential of human prostate cancer. *Nat. Commun.* 8:420. doi: 10.1038/s41467-017-00460-4
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* 17, 261–272. doi: 10.1038/s41592-019-0686-2
- Waagepetersen, R. P. (2007). An estimating function approach to inference for inhomogeneous Neyman–Scott processes. *Biometrics* 63, 252–258. doi: 10.1111/j.1541-0420.2006.00667.x
- Wang, H., Milberg, O., Bartelink, I. H., Vicini, P., Wang, B., Narwal, R., et al. (2019). In silico simulation of a clinical trial with anti-CTLA-4 and anti-PD-L1 immunotherapies in metastatic breast cancer using a systems pharmacology model. *R. Soc. Open Sci.* 6:190366. doi: 10.1098/rsos.190366
- Wang, H., Sove, R. J., Jafarnejad, M., Rahmeh, S., Jaffee, E. M., Stearns, V., et al. (2020). Conducting a virtual clinical trial in HER2-negative breast cancer using a quantitative systems pharmacology model with an epigenetic modulator and immune checkpoint inhibitors. *Front. Bioeng. Biotechnol.* 8:141. doi: 10.3389/fbioe.2020.00141
- Wang, J., Zhang, L., Jing, C., Ye, G., Wu, H., Miao, H., et al. (2013). Multi-scale agent-based modeling on melanoma and its related angiogenesis analysis. *Theor. Biol. Med. Model.* 10:41. doi: 10.1186/1742-4682-10-41
- Weibel, E. R., Kistler, G. S., and Scherle, W. F. (1966). Practical stereological methods for morphometric cytology. *J. Cell Biol.* 30, 23–38. doi: 10.1083/jcb.30.1.23
- Wells, D. K., Chuang, Y., Knapp, L. M., Brockmann, D., Kath, W. L., and Leonard, J. N. (2015). Spatial and functional heterogeneities shape collective behavior of tumor-immune networks. *PLoS Comput. Biol.* 11:e1004181. doi: 10.1371/journal.pcbi.1004181
- Wong, P. F., Wei, W., Smithy, J. W., Acs, B., Toki, M. I., Blenman, K. R., et al. (2019). Multiplex quantitative analysis of tumor-infiltrating lymphocytes and immunotherapy outcome in metastatic melanoma. *Clin. Cancer Res.* 25, 2442–2449. doi: 10.1158/1078-0432.CCR-18-2652
- Yuan, Y. (2016). Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harb. Perspect. Med.* 6:a026583. doi: 10.1101/cshperspect.a026583
- Zhang, A. W., Mcpherson, A., Milne, K., Kroeger, D. R., Hamilton, P. T., Miranda, A., et al. (2018). Interfaces of malignant and immunologic clonal dynamics in ovarian cancer. *Cell* 173, 1755–1769.e22. doi: 10.1016/j.cell.2018.03.073

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Mi, Gong, Sulam, Fertig, Szalay, Jaffee, Stearns, Emens, Cimino-Mathews and Popel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.