

Systems biology

# spatialTIME and iTIME: R package and Shiny application for visualization and analysis of immunofluorescence data

Jordan H. Creed<sup>1†</sup>, Christopher M. Wilson<sup>1†</sup>, Alex C. Soupir<sup>1,2</sup>,  
Christelle M. Colin-Leitzinger<sup>3</sup>, Gregory J. Kimmel<sup>4</sup>, Oscar E. Ospina<sup>1</sup>,  
Nicholas H. Chakiryan<sup>5</sup>, Joseph Markowitz<sup>6</sup>, Lauren C. Peres <sup>3</sup>, Anna Coghill<sup>3</sup> and  
Brooke L. Fridley <sup>1,\*</sup>

<sup>1</sup>Department of Biostatistics and Bioinformatics, Tampa, FL, USA, <sup>2</sup>Department of Tumor Biology, Tampa, FL, USA, <sup>3</sup>Department of Cancer Epidemiology, Tampa, FL, USA, <sup>4</sup>Department of Integrated Mathematical Oncology, Tampa, FL, USA, <sup>5</sup>Department of Genitourinary Oncology, Tampa, FL, USA and <sup>6</sup>Department of Cutaneous Oncology, Moffitt Cancer Center, Tampa, FL 33612, USA

\*To whom correspondence should be addressed.

<sup>†</sup>The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Associate Editor: Anthony Mathelier

Received on May 25, 2021; revised on September 10, 2021; editorial decision on October 27, 2021; accepted on October 29, 2021

## Abstract

**Summary:** Multiplex immunofluorescence (mIF) staining combined with quantitative digital image analysis is a novel and increasingly used technique that allows for the characterization of the tumor immune microenvironment (TIME). Generally, mIF data is used to examine the abundance of immune cells in the TIME; however, this does not capture spatial patterns of immune cells throughout the TIME, a metric increasingly recognized as important for prognosis. To address this gap, we developed an R package spatialTIME that enables spatial analysis of mIF data, as well as the iTIME web application that provides a robust but simplified user interface for describing both abundance and spatial architecture of the TIME. The spatialTIME package calculates univariate and bivariate spatial statistics (e.g. Ripley's K, Besag's L, Macron's M and G or nearest neighbor distance) and creates publication quality plots for spatial organization of the cells in each tissue sample. The iTIME web application allows users to statistically compare the abundance measures with patient clinical features along with visualization of the TIME for one tissue sample at a time.

**Availability and implementation:** spatialTIME is implemented in R and can be downloaded from GitHub (<https://github.com/FridleyLab/spatialTIME>) or CRAN. An extensive vignette for using spatialTIME can also be found at <https://cran.r-project.org/web/packages/spatialTIME/index.html>. iTIME is implemented within a R Shiny application and can be accessed online (<http://itime.moffitt.org/>), with code available on GitHub (<https://github.com/FridleyLab/iTIME>).

**Contact:** [fridley.lab@moffitt.org](mailto:fridley.lab@moffitt.org) or [brooke.fridley@moffitt.org](mailto:brooke.fridley@moffitt.org)

**Supplementary information:** [Supplementary data](#) are available at *Bioinformatics* online.

## 1 Introduction

Multiplex immunofluorescence (mIF) staining combined with quantitative digital image analysis is a novel and increasingly used technique that allows for the assessment and visualization of distinct immune cell populations in the tumor immune microenvironment (TIME), as well as discrimination between tumor and stroma compartments. mIF has been recently applied to the study of many

cancer types, including oropharyngeal squamous cell carcinoma (Tsakiroglou *et al.*, 2020), gastric cancer (Huang *et al.*, 2019) and pancreatic cancer (Vayrynen *et al.*, 2021). An article was also recently published outlining the opportunities and challenges in the analysis of mIF data, including approaches for spatial analysis of mIF (Wilson *et al.*, 2021b).

Generally, mIF data has been used to examine the presence and abundance of immune cells in the TIME; however, this aggregate

measure assumes uniform spatial patterns of immune cells throughout the tissue sample and overlooks potential spatial heterogeneity of immune marker-positive cell populations. mIF technologies also provide position data for immune cells in the TIME, thus allowing for the assessment of the spatial architecture of the TIME. Therefore, to facilitate the spatial analysis and visualization of mIF data following the preprocessing of the image data, we have developed an R package, spatialTIME and a Shiny application, iTIME.

One commonly used technology for mIF data generation is the Vectra 3.0™/Polaris™ system. Images of immune marker-stained tissues are processed within InForm (Gorris *et al.*, 2018; Mezheyeuski *et al.*, 2018; Mori *et al.*, 2020; Shakya *et al.*, 2020) followed by analysis with the HALO Image Analysis Platform (Indica Labs, NM). In HALO, a supervised classifier is trained to classify tissue as tumor, stroma and glass (no tissue) regions (Amancio *et al.*, 2014; Breiman, 2001; Horai *et al.*, 2019). Cell segmentation and marker quantitation is performed by compartmental examination of fluorescent intensity thresholds, with each immune marker assigned a distinct fluorescence value (Mezheyeuski *et al.*, 2018; Mori *et al.*, 2020). The output from this preprocessing of the mIF data is a file for each tissue sample with the locations of all detected cells and information on whether the cell is positive or negative for each of the assayed markers. In addition to the sample specific file, a summary file is also produced with the descriptive summary for each tissue sample, including the number of cells detected, the number of positive cells for each marker, percentage positive for each marker and density estimates. These summary measures can also be computed across all cells and by tumor/stroma if a tumor-associated marker is included in the assayed panel of markers. The file inputs for the spatialTIME and iTIME package follow the formats of the HALO output; however, the required file formats are general in format (i.e. cell location, indicator variable for positivity for each of the markers or prespecified combinations or ‘phenotypes’) with output from other commonly used technologies able to be easily reformatted into this general format for input into the software tools. Future updates to the package will allow for inputs directly from a variety of other technologies used in mIF studies.

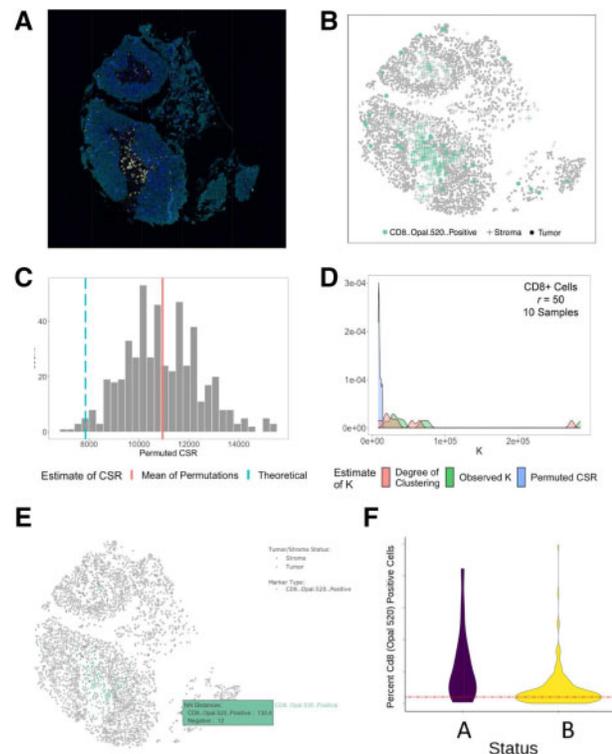
In spatialTIME, we have implemented a number of spatial measures of clustering including Ripley’s  $K$  (Ripley, 1976), Besag’s  $L$  (Besag, 1977) and G statistic (nearest neighbor distance) using spatstat R package (Baddeley *et al.*, 2016). To address a unique challenge presented in mIF data generated on tissue microarrays (TMAs) whereby cores often have areas where measurements are not able to be acquired (i.e. areas with tears, folded tissue or fibrous tissue) we have implemented a permutation approach for estimating the null distribution of no clustering for which the observed clustering statistic is compared (i.e. complete spatial randomness or CSR) (Wilson *et al.*, 2021a). That is, the measure to be used in downstream visualization and statistical association with clinical outcome is the degree of spatial clustering as defined as the observed value of measure—measure under assumption of CSR. Additional information on the various spatial measures can be found in [Supplemental Materials](#).

## 2 Implementation

spatialTIME is available as an R package from either GitHub (<https://github.com/FridleyLab/spatialTIME>) or CRAN (CRAN.R-project.org) and made publicly available under the MIT license. The package functionality is built upon the tidyverse principals and several of its packages along with the spatstat package. mIF experiments return two main datasets—a file with the summary-level information for the markers and individual files for each tissue sample with cell-level staining information. To combat potential confusion due to multiple samples per person, functions in spatialTIME use a custom input object called ‘mif.’ This object can be created by calling the create\_mif function and requires at a minimum for users to input the summary and clinical datasets. The resulting mif object is the basis for all other functions within the package.

spatialTIME provides functions for plotting mIF data for individual tissue samples and calculating univariate and bivariate

measures of degree of spatial clustering and co-localization of marker-positive cells. For spatial measures of clustering, we have implemented a number of spatial measures of clustering including Ripley’s  $K$  (Ripley, 1976), Besag’s  $L$  (Besag, 1977), Marcon’s  $M$  (Marcon *et al.*, 2015) and G statistic (nearest neighbor distance) using spatstat R package (Baddeley *et al.*, 2016). To address a unique challenge presented in mIF data generated on tissue microarrays (TMAs) whereby cores often have areas where measurements are not able to be acquired (i.e. areas with tears, folded tissue or fibrous tissue), we have implemented a permutation approach for estimating the null distribution of no clustering for which the observed clustering statistic is compared (i.e. complete spatial randomness or CSR) (Wilson *et al.*, 2021a). That is, the measure to be used in downstream visualization and statistical association with clinical outcome is the degree of spatial clustering as defined as the observed value of measure—measure under assumption of CSR. The clustering functions return data frames containing spatial clustering estimates for a range of neighborhood sizes, denoted by  $r$ , for each sample. Users also have the option of obtaining either the full empirical distribution under CSR or simply the mean of this empirical null distribution. The choice of the radius ( $r$ ) to use for estimating the spatial measures (i.e. size of neighborhood) should be based on the scale of clustering of interest (i.e. a small value of  $r$  will have clustering assess for small neighborhoods while a larger value of  $r$  would determine the level of clustering based on large-sized neighbors). It is recommended that the association analyses of spatial measure with endpoint of interested be conducted using a few different  $r$  values (i.e. sensitivity analysis). Future work is planned for conducting the



**Fig. 1.** mIF data generated from a TMA study of prostate cancer. (A) mIF image data for CD8 Positive cells measured on a prostate cancer tumor sample. (B) Plots showing CD8+ cells and locations, with illustrating a tissue sample with ‘holes’ or regions of unmeasured cells. (C) Empirical distribution of the estimate of Ripley’s  $K$  under CSR ( $N$  permutations = 500), with lines indicating the permuted (mean) and theoretical estimates under CSR. (D) Distribution of observed  $K$ , permuted estimate of  $K$  under CSR, and difference in these measurements (i.e. ‘degree of spatial clustering’) for the 10 core tissue samples included in spatialTIME. The degree of spatial clustering can be used in downstream association with the phenotype of interest. (E) Representation of cell locations, using the iTIME application with information in the box showing the nearest cell for each marker type. (F) Violin plot of the percent of CD8+ cells by a clinical endpoint (group A versus B) using the iTIME application

association of the endpoint with the entire spatial curve computed at several  $r$  values using functional data analysis approaches. Finally, the package allows for the creation of publication quality plots (one for each spatial sample provided) using ggplot2 which can be directly exported as a pdf. Detailed information on how to use spatialTIME can be found in the R package documentation and the vignette.

The iTIME Shiny application is available at <http://itime.moffitt.org/> and is a point and click version with much of the functionality found in the spatialTIME package. On the 'Importing Data' page, users upload clinical, summary- and cell-level data in csv files and specify the variable to use for the data merge (i.e. variable that links the data across the various files). The 'Univariate Summary' page provides a high-level overview of the summary and clinical data files, including plots and hypothesis testing based on a beta-binomial model to compare summary marker data by clinical variable of interest. Users can customize clinical variables, color schemes and plot type for effective data visualization, including a scatter plot, boxplot, violin plot or histogram. The 'Multivariate Summary' page provides a heatmap and principal component plot for selected immune cell populations and provides the option to perform a clustering analysis to understand if immune cell composition of the TIME is related to clinical features. The 'Spatial' page displays an interactive visual of the imported cell-level data for a sample, with tumor cells displayed as triangles and stromal cells as circles, and a separate color to represent positivity for each distinct immune marker. When hovering over each cell, a text window will display the nearest neighbor distance. A plot of the measures of marker-positive immune cell spatial clustering, such as Ripley's  $K$ , Besag's  $L$ , Marcon's  $M$  and nearest neighbor distance distribution  $G$ , over various  $r$  values is also generated and can be calculated specific to each immune marker. These calculations adjust for edge effects with the isotropic or translational edge ( $K$ ,  $L$ ,  $M$ ) and reduced sample or Hanisch ( $G$ ) correction methods (Baddeley et al., 2016), with a simulation-based envelope presented around the theoretical estimate under CSR. Currently, the iTIME application provides the theoretical distribution under the CSR assumption; however, future updates aim to also include the empirical distribution. To demonstrate the functionality of spatialTIME and iTIME, we examined two prostate tumor samples with differing TIME architecture, as presented in Figure 1.

### 3 Conclusions

The novel technique of mIF provides a cost-effective approach for studying the TIME in a large number of tissue samples simultaneously. This technology also provides the ability to measure not only the abundance but also the spatial location of multiple cell types within a tissue sample. The development of methods and approaches, such as those implemented in spatialTIME, that can

deeply characterize both marker abundance and the spatial heterogeneity of marker-positive cells in the TIME can inform treatment approaches (e.g. immunotherapy) by providing a clearer picture of therapeutic targets in tumor tissue.

**Financial Support:** This work has been supported in part by the Biostatistics and Bioinformatics Shared Resource at the H. Lee Moffitt Cancer Center & Research Institute, an NCI designated Comprehensive Cancer Center (P30-CA076292).

**Conflict of Interest:** Authors have no conflict of interest to declare.

### References

- Amancio,D.R. et al. (2014) A systematic comparison of supervised classifiers. *PLoS One*, **9**, e94137.
- Baddeley,A. et al. (2016) *Spatial Point Patterns: Methodology and Applications with R*. CRC Press, Boca Raton, FL.
- Besag,J. (1977) Comments on Ripley's paper. *J. R. Stat. Soc. Ser. A*, **B39**, 193–195.
- Breiman,L. (2001) Random forests. *Mach. Learn.*, **45**, 5–32.
- Gorris,M.A.J. et al. (2018) Eight-color multiplex immunohistochemistry for simultaneous detection of multiple immune checkpoint molecules within the tumor microenvironment. *J. Immunol.*, **200**, 347–354.
- Horai,Y. et al. (2019) Quantification of histopathological findings using a novel image analysis platform. *J. Toxicol. Pathol.*, **32**, 319–327.
- Huang,Y.K. et al. (2019) Macrophage spatial heterogeneity in gastric cancer defined by multiplex immunohistochemistry. *Nat. Commun.*, **10**, 3928.
- Marcon,E. et al. (2015) Tools to characterize point patterns: dbmss for R. *J. Stat. Softw.*, **67**, 1–15.
- Mezheyeuski,A. et al. (2018) Multispectral imaging for quantitative and compartment-specific immune infiltrates reveals distinct immune profiles that classify lung cancer patients. *J. Pathol.*, **244**, 421–431.
- Mori,H. et al. (2020) Characterizing the tumor immune microenvironment with tyramide-based multiplex immunofluorescence. *J. Mammary Gland Biol. Neoplasia*, **25**, 417–432.
- Ripley,B.D. (1976) The second-order analysis of stationary point processes. *J. Appl. Probab.*, **13**, 255–266.
- Shakya,R. et al. (2020) Immune contexture analysis in immuno-oncology: applications and challenges of multiplex fluorescent immunohistochemistry. *Clin. Transl. Immunol.*, **9**, e1183.
- Tsakiroglou,A.M. 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.
- Vayrynen,S.A. et al. (2021) Composition, spatial characteristics, and prognostic significance of myeloid cell infiltration in pancreatic cancer. *Clin. Cancer Res.*, **27**, 1069–1081.
- Wilson,C. et al. (2021a) Statistical framework for studying the spatial architecture of the tumor immune microenvironment. *medRxiv*, **2021**, 2004.2027.21256104. [10.1101/2021.04.27.21256104](https://doi.org/10.1101/2021.04.27.21256104).
- Wilson,C.M. et al. (2021b) Challenges and opportunities in the statistical analysis of multiplex immunofluorescence data. *Cancers Basel*, **13**, 3031.