Cytometry PART A
ISAC
Journal of Quantitative Cell Science

Robust Cell Detection and Segmentation for Image Cytometry Reveal Th17 Cell Heterogeneity

Takahiro Tsujikawa,[1,2,6†] Guillaume Thibault,[3†] Vahid Azimi,[4] Sam Sivagnanam,[4] Grace Banik,[2] Casey Means,[2] Rie Kawashima,[1] Daniel R. Clayburgh,[2,5] Joe W. Gray,[3,5] Lisa M. Coussens,[1,5] Young Hwan Chang[3,4*]

[1]Department of Cell, Developmental, and Cancer Biology, Oregon Health & Science University, Portland, Oregon, USA

[2]Department of Otolaryngology-Head & Neck Surgery, Oregon Health & Science University, Portland, Oregon, USA

[3]Department of Biomedical Engineering, Oregon Health & Science University, Portland, Oregon, USA

[4]Computational Biology Program, Oregon Health & Science University, Portland, Oregon, USA

[5]Department of Knight Cancer Institute, Oregon Health & Science University, Portland, Oregon, USA

[6]Department of Otolaryngology-Head and Neck Surgery, Kyoto Prefectural University of Medicine, Kyoto, Japan

Received 23 July 2018; Revised 30 October 2018; Accepted 14 January 2019

Grant sponsor: Grant-in-Aid for Scientific Research, Japan Society for the Promotion of Science, Grant number: 17H07016; Grant sponsor: National Center for Advancing Translational Sciences; Grant sponsor: NCI/NIH, Grant number: NCI U54CA209988; Grant sponsor: Oregon Clinical and Translational Research Institute, Grant number: UL1TR000128; Grant sponsor: Stand Up To Cancer - Lustgarten Foundation Pancreatic Cancer Convergence Dream Team Translational Research Grant, Grant number: SU2C-AACR-DT14-14; Grant sponsor: OHSU Center for Spatial Systems Biomedicine (OCSSB); Grant

Abstract

Image cytometry enables quantitative cell characterization with preserved tissue architecture; thus, it has been highlighted in the advancement of multiplex immunohistochemistry (IHC) and digital image analysis in the context of immune-based biomarker monitoring associated with cancer immunotherapy. However, one of the challenges in the current image cytometry methodology is a technical limitation in the segmentation of nuclei and cellular components particularly in heterogeneously stained cancer tissue images. To improve the detection and specificity of single-cell segmentation in hematoxylin-stained images (which can be utilized for recently reported 12-biomarker chromogenic sequential multiplex IHC), we adapted a segmentation algorithm previously developed for hematoxlin and eosin-stained images, where morphological features are extracted based on Gabor-filtering, followed by stacking of image pixels into *n*-dimensional feature space and unsupervised clustering of individual pixels. Our proposed method showed improved sensitivity and specificity in comparison with standard segmentation methods. Replacing previously proposed methods with our method in multiplex IHC/image cytometry analysis, we observed higher detection of cell lineages including relatively rare T_H17 cells, further enabling sub-population analysis into T_H1-like and T_H2-like phenotypes based on T-bet and GATA3 expression. Interestingly, predominance of T_H2-like T_H17 cells was associated with human papilloma virus (HPV)-negative status of oropharyngeal squamous cell carcinoma of head and neck, known as a poor-prognostic subtype in comparison with HPV-positive status. Furthermore, T_H2-like T_H17 cells in HPV-negative head and neck cancer tissues were spatiotemporally correlated with $CD66b^+$ granulocytes, presumably associated with an immunosuppressive microenvironment. Our cell segmentation method for multiplex IHC/image cytometry potentially contributes to in-depth immune profiling and spatial association, leading to further tissue-based biomarker exploration.

Key terms

cell segmentation; T_H17 cell phenotypes; tumor immune microenvironment

As emergence of immunotherapy has been revolutionizing cancer therapeutic strategies, the demands on biomarkers optimizing various treatment options have been constantly increasing (1–4). Profiling cellular complexities of tumor tissue provides a powerful platform to understand immune characteristics associated with clinical response to immunotherapy; thus, in situ tumor profiling and imaging have been increasingly important toward tissue-based biomarker development.

Single-cell segmentation is a key technology in digital image analysis, as it enables the quantitative assessment of cell frequency, localization, and phenotypes (5–7). Previously, based on single-cell segmentation, we built an image quantitation platform with multiplex immunohistochemistry (IHC) and image cytometry platform to evaluate in situ immune characteristics where 12 immune-related markers were quantitatively assessed by image cytometry (8). Image cytometry enables the

sponsor: Brenden-Colson Center for Pancreatic Care; Grant sponsor: National Cancer Insitute (NCI); Grant sponsor: Stand Up To Cancer—Lustgarten Foundation Pancreatic Cancer Convergence Dream Team Translational Research Grant; Grant sponsor: Brenden-Colson Center for Pancreatic Health; Grant sponsor: Japan Society for the Promotion of Science; Grant sponsor: National Center for Advancing Translational Sciences (NCATS) at the National Institutes of Health (NIH); Grant sponsor: Oregon Clinical and Translational Research Institute (OCTRI); Grant sponsor: Knight Cancer Institute, Grant number: P30 CA069533

Additional Supporting Information may be found in the online version of this article.

*Correspondence to: Young Hwan Chang, Department of Biomedical Engineering, Computational Biology Program, Oregon Health & Science University, 2720 SW Moody Ave., Mailcode KR-COM, Portland, OR 97201, USA Email: chanyo@ohsu.edu

†These authors contributed equally to the work as first authors.

Published online in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.23726

evaluation of cell morphology and signal intensity, analogous to flow cytometry, but where tumor architecture is maintained. However, a potential limitation of current image cytometry analysis is a technical challenge for single-cell segmentation in hematoxylin images, where CellProfiler (CP) (9) has been adopted as a standard. Although CP performs well using image-specific proper parameters, tweaking parameters is difficult, especially when dealing with clinical samples that show heterogeneity in shape, size, and intensity. Thus, a fully automated and robust single-cell segmentation method is required for not only providing a high-throughput cell segmentation but also enhancing image cytometry for multiplex IHC images.

In this study, we adapted our nuclei segmentation method (7) for hematoxylin and eosin (H&E) staining to provide stable cell segmentation and minimize the cost of frequent parameters tuning. In order to segment nuclei in hematoxylin-stained images, we extracted useful morphological features from the image using a set of Gabor filters (10). We mapped each image pixel to a point in an n-dimensional feature space and then clustered individual pixels with similar features in an unsupervised setting. By doing this, nuclei segmentation in hematoxylin-stained images can be effectively performed by partitioned groups. We validated our cell segmentation algorithm under the multiplex IHC setting and confirmed that the proposed method was able to detect higher cell numbers, although relative cellular ratios were maintained between CP-based segmentation in the previous study (8) and the proposed method. In addition, our method enabled the detection of relatively rare cell populations, such as T_H17 cells, a subset of T lymphocytes associated with autoimmune diseases but typically represents less than 1% of total immune cells (11). Use of this new segmentation method allowed the analysis of these cells. Analysis revealed that rare T_H17 cell populations can be further divided into T_H1-like and T_H2-like subpopulations, which correlated with differential immune characteristics of tissue. Development of an image analytic platform based on robust and effective single-cell segmentation can contribute to further tissue-based biomarker development.

Background

In situ profiling cellular complexities of tissues and tumors provide a unique opportunity to quantitatively evaluate characteristic features of disease pathogenesis. With regard to inflammation associated with cancer development and the advent of therapies targeting immune-modulatory programs in tumors, evaluating the complexity and functional status of immune cells infiltrating tumor tissue provides information not only for patient stratification but also for monitoring metrics to discern response and/or resistance to anticancer therapy. As such, multiplex IHC platforms have emerged enabling robust detection of immune cell lineage and phenotype on single-tissue sections. Rate limiting for these platforms is their capabilities for automated high-throughput and quantitative segmentation methods. Thus, to improve the detection and specificity of single-cell segmentation in hematoxylin-stained images and to provide capabilities for automated high-throughput and quantitative segmentation method, we developed a simple and effective methodology for a fully automated and robust single-cell segmentation approach that reduces the cost of parameter tuning.

Single-cell segmentation is vital for various fields of digital pathological evaluation despite technical obstacles such as morphological variation of tumor and stroma cells, and high cell density leading to back-to-back formations. As a consequence, many automatic segmentation, analysis, and computer-aided diagnosis methods have been proposed to alleviate pathologist's burden (12–16). Various approaches have been proposed, ranging from relatively simple thresholding methods (Otsu or adaptive) to more sophisticated methods including active contour, level set, watershedding with multi-scaling seeds, unsupervised Bayesian, fuzzy c-means classification, supervised methods using machine/deep learning, etc. For nuclear segmentation purpose, these methods performance range from 75 to 96% accuracy (14) and are sensitive to the image quality and tissue type. Moreover, as the most recent literature (17) still shows new developments of a robust, practically usable segmentation algorithm, this indicates that developing novel segmentation methods is still an ongoing task.

In the H&E staining method, cell nuclei is stained blue by hematoxylin, followed by counter-staining with eosin, which colors other structures in various shades of red and pink (18). Although cell segmentation in H&E images is a highly challenging task, it is more difficult for multiplex IHC, because multiplex IHC requires hematoxylin-only images without eosin, which is resistant to signal stripping protocol. This results in a dramatic loss of contrast. In this study, we proposed an unsupervised method for cell segmentation in

order to reduce an effort for parameter tuning for multiple experimental samples, enabling the exploration of phenotypes and spatial distribution of relatively rare T_H17 sub-populations. To do this, we utilize morphological features, which are particularly appropriate for texture representation and discrimination.

Materials and methods

Nuclei and Cell Segmentation for Hematoxylin-Only Image

Each pixel has an intensity which represents a part of morphological features. In order to provide robust nuclei and cell segmentation and minimize the cost of frequent parameter tuning, we simply cluster individual pixels based on the similarity of their features by extracting useful morphological features from the image. To do so, Gabor filtering was adopted with different frequencies and orientations, which are particularly appropriate for texture representation and discrimination, that is, edge detection in image processing (7). Also, various features such as intensities and Gabor filters' impulse responses were stacked, where these features can be chosen by users. To avoid over-segmented noisy pixels, 2D Gaussian-filtered images were added to the feature vector. Following mapping of each image pixel to a point in an *n*-dimensional feature space, each pixel is enhanced by chosen features. Then, *k*-means clustering was performed to differentiate foreground from background in different tissues, cells, or nuclei. Finally, after clustering and group selection, aggregated nuclei are discerned using mathematical morphology operations: first, an alternate sequential filter (consecutive openings and closings with structuring elements of increasing radii) reduced the noise and flattened the image by erasing small local gray level variation within the nuclei. Then, the resulting local minima (nuclei) were used into a seeded watershed on the gradient image.

Other Segmentation Methods for Comparison

Segmentation using CP with parameter tuning. With manual parameters tuning procedure, segmentation was performed in CP version 2.2.0 using the "Identify Primary Objects" module with the "Threshold strategy" parameter set to "manual." The primary parameters such as the cell diameter ("Typical diameter of objects, in pixel units [Min, Max]") and the object intensity minimum threshold ("Manual threshold") were tuned for each image. After manual tuning of parameters, the pipeline was executed and the results were inspected visually (using the "Convert Objects To Image" module). Parameters were re-tuned accordingly based on the visual inspection of segmentation results. This process was repeated with as much different iteration of parameters necessary until convergence to the best possible segmentation (as determined by visual inspection) was achieved. Additional parameters were used to achieve satisfactory segmentation for certain images, including suppression of local maxima within a minimum allowed distance (using the "Automatically calculate minimum allowed distance between local maxima" parameter in the "Identify Primary Objects" module) and objects filtering based on maximum allowed eccentricity (using the "Filter Objects" module).

ImageJ/FIJI-based segmentation. Segmentation in ImageJ/-FIJI (19) was performed with a custom macro created to produce a binary nuclei mask from a hematoxylin-stained image. This macro takes the hematoxylin-stained image and performs color deconvolution to separate the hematoxylin stain from the background. Next, several preprocessing steps are performed to enhance the nuclei stain including Gaussian blur, median filter, contrast enhancement by 0.25% saturation, and rolling ball background subtraction (19). An automatic Otsu threshold is applied on the resulting image to select foreground (20). The "Find Maxima" function is run with the "Segmented Particles" option and noise tolerance as an argument provided from the user to create the binary mask. The noise parameter was visually examined at multiple values to determine the best noise value for each image. The output assumes each maximum belongs to a particle and segments the foreground area by watershed (21), producing a binary image of particles identifying cells by each maximum point and surrounding area under foreground selection. The binarized image is post processed to remove remaining background noise with erosion, removing outliers, watershed, median filter, "Fill holes," dilation, and watershed again to produce the final nuclei separation and mask. All functions can be found with full descriptions at https://imagej.nih.gov/ij/docs/guide/.

Multiplex IHC and Image Acquisition

Chromogenic sequential multiplex IHC and digital image acquisition were conducted as described before (8). Briefly, 5 μm of formalin fixed paraffin-embedded tissue sections were stained by hematoxylin, followed by whole-tissue scanning using Aperio ImageScope AT (Leica Biosystems, Buffalo Grove, IL). Following endogenous peroxidase blocking and heat-mediated antigen retrieval, sequential IHC consisting of iterative cycles of staining, scanning, and antibody and chromogen stripping was performed. Acquired digital images were co-registered so that cell features overlap down to a single-pixel level, using CP pipeline (8,22). Then, pixel intensity and shape-size measurements were saved to a file format compatible with image cytometry data analysis software, FCS Express 5 Image Cytometry v.5.01.0029 (De Novo Software) (8).

Spatial Pattern Analysis

Extraction of spatial proximity and distance measurements. Positional data extracted from multiplex IHC images were used to measure distance from each individual neutrophil to the T_H1-like T_H17 cells and T_H2-like T_H17 cells. Using the Quickhull Algorithm (23), "dsearchn" function in MATLAB, the shortest distance between $CD66b^+$ granulocyte centroids and the nearest T_H17 cells, was measured to determine spatial proximity between $CD66b^+$ granulocyte distance to T_H1-like T_H17 cells and T_H2-like T_H17 cells, respectively. Scatter plot in Figure 4D represents the shortest distance from all neutrophil to the nearest T_H1-like T_H17 cells versus T_H2-like T_H17 cells.

Local density measurement by adopting radial distribution function. Since cell density can affect the spatial proximity of cell–cell distances, the radial distance function was adopted as previously reported (24). Instead of using $CD66b^+$ granulocytes as a reference point, each T_H1-like T_H17 or T_H2-like T_H17 cell was considered as a reference, enabling the calculation of the average density of $CD66b^+$ granulocytes at a distance *r*. To do this, first, *k*-nearest neighbors $CD66b^+$ granulocytes from individual T_H1-like T_H17 cells and T_H2-like T_H17 cells were counted across distance *r*. Then, only the *k*-nearest neighbor $CD66b^+$ granulocyte was counted, instead of density to handle boundary conditions. Second, an ensemble distribution of $CD66b^+$ granulocytes from T_H1-like T_H17 cells and T_H2-like T_H17 cells was individually assessed in differential *k* values, as shown in Figure 4E and Supporting Information Figure S5C.

Manual Cell Annotation

Hematoxylin images digitized at 20× magnification were utilized to manually count cell segments. By using the cell counter plugin in ImageJ, centroids of cell nucleus were manually marked by three physicians and recorded.

Patient Samples, IRB Description

All digital images from human cancer tissue were deidentically obtained under approval by institutional review board (IRB) (protocol 809, 3609, and 5886), and written informed consent was obtained.

Statistics

Mann–Whitney *U* and Wilcoxon signed rank tests were used to determine statistically significant differences in unpaired and paired data. Spearman rank correlation coefficient was used to assess correlations of cell percentages and densities among cell lineages. Mann–Whitney *U* and Wilcoxon signed rank tests were used for spatial relationship analysis, and a *P*-value of less than 0.05 was considered as statistically significant. All statistical calculations were performed by GraphPad Prism version 7.03. $P < 0.05$ was considered statistically significant.

Results

We proposed a simple but effective methodology for fully automated and robust single-cell segmentation with reduced cost of parameter tuning. To do this, we extracted useful morphological features from the image and group individual pixels using clustering based on similar features to segment nuclei. Notably, our segmentation has been optimized with images stained only by hematoxylin, which lack the cytoplasmic staining usually provided by eosin staining (Fig. 1A–C). This provides robust segmentation results across different cancer histological types with the same parameter setting. However, a limitation could be the dependence on image quality of hematoxylin staining, where low-quality images lead to failure in cell detection or over/under-segmentation even with our segmentation approach. Our data indicate that the proposed segmentation shows better detection of cells with higher sensitivity and specificity in three different cancer histological types. While we observed preserved cell ratio and composition (Supporting Information Fig. S4), our robust and effective cell detection enabled better detection of relatively rare T_H17 cells (Fig. 3A,B). Improved cell detection was particularly significant for subpopulation analysis, which depends on a sufficiently large cell population.

A Cell Segmentation Method Based on Gabor Filtering and Pixel-Based Clustering Outperforms Standard Cell Segmentation Methods

To improve the detection and specificity of single-cell segmentation, we adopted our previous method to provide better nuclei segmentation for multiplex IHC images (see "Materials and Methods" section). Next, we compared the new segmentation results with our previous CP-based method, which is composed of the following four steps: seeding, calculation of intensity gradient, watershed segmentation, and merging objects with weak borders. To perform a fair comparison, an expert user determined the optimal parameter by visual examination in order to achieve the best cell segmentation in CP. Our method effectively detected nuclei objects with fine edge identification, whereas the CP-based method occasionally failed to detect objects and showed over-segmentation (Fig. 1A–C and Supporting Information Fig. S1).

To statistically verify those observations, we utilized nuclei objects manually identified by three independent physicians as a comparative baseline using moderate agreement between three examiners' annotations, so-called ground truth (Supporting Information Fig. S2 and Table 1). Three examiners marked the centroids of all visually recognizable cells in assigned images derived from different histological cancer types of head and neck squamous cell carcinoma, malignant mesothelioma, and pancreatic ductal adenocarcinoma, where malignant mesothelioma had relatively homogenous cell size in contrast to moderate and high heterogeneity in head and neck cancer and pancreatic cancer, respectively (Fig. 1A–C). Based on the dilatation of annotated marks with various radii between one and six pixels, we evaluated the overlap of dilated annotations to see matching and discrepancy of cell identification across three examiners. The manual cell annotations by three observers were moderately consistent, where matched annotations (overlapping or no further than three pixels) by all three observers ranged between 70 and 76.7%, reflecting technical challenges in heterogeneous cancer tissues (Supporting Information Fig. S2 and Table S1).

Based on ground truth defined by manual cell identification, our proposed segmentation method was statistically compared with CP and additional ImageJ/FIJI-based segmentation methods (https://imagej.nih.gov/ij/) with manually tuned optimal parameters (see "Materials and Methods" section). When ground truth was set to annotations selected by two or all physicians in consideration of inter-observer variability of cell identification, our proposed segmentation showed 65–78% sensitivity in the three cancer types, whereas CP-based method remained 55% (Fig. 1D,E and

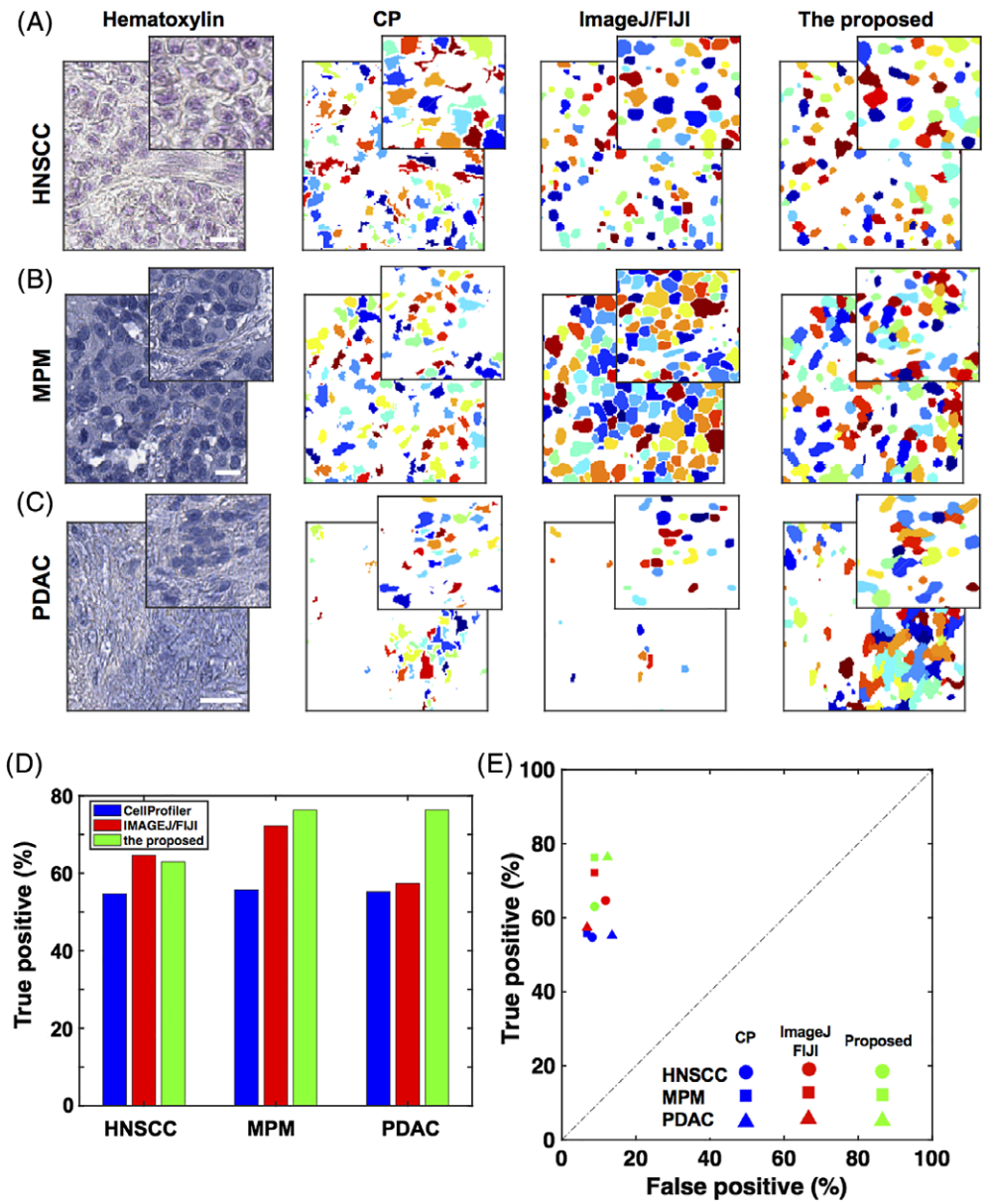

Figure 1. Improved cell segmentation for hematoxylin images. Cell segmentation by CellProfiler (CP), ImageJ/FIJI, and our proposed method were compared in three different histological cancer types of head and neck squamous cell carcinoma (HNSCC) (**A**), malignant pleural mesothelioma (MPM) (**B**), and pancreatic ductal adenocarnoma (PDAC) (**C**). Hematoxylin image and nuclei segmentation label masks based on the three methods are shown. Scale bars = 100 μm. (**D** and **E**) Segmentation methods were compared in terms of true positive ratio and specificity/sensitivity in representative cancer histological types of HNSCC, MPM, and PDAC. For a better visual identification, the selected magnified images are shown on the top right corner for each image where low magnified images are shown in Supporting Information Figure S1.

Supporting Information Table S2). While our method and ImageJ/FIJI-based methods showed equivalent sensitivity in malignant mesothelioma and head and neck squamous cell carcinoma tissues, our method outperforms ImageJ/FIJI-based method in heterogeneous pancreatic ductal adenocarcinoma tissues (Fig. 1D). Specificity of our method was approximately 90% when ground truth was set to annotations identified by one or more physicians (Fig. 1E and Supporting Information Table S2). Notably, our proposed segmentation method was applied without parameter tuning, whereas the other two methods were applied with the optimal or image-specific proper parameters by an expert user across different cancer tissue samples. Given technical challenges in heterogeneous cancer tissues and moderate consistency across three examiners (70–76.7%), these observations suggest that our method achieves reasonable performance among different segmentation methods regardless of heterogeneity of cancer tissues.

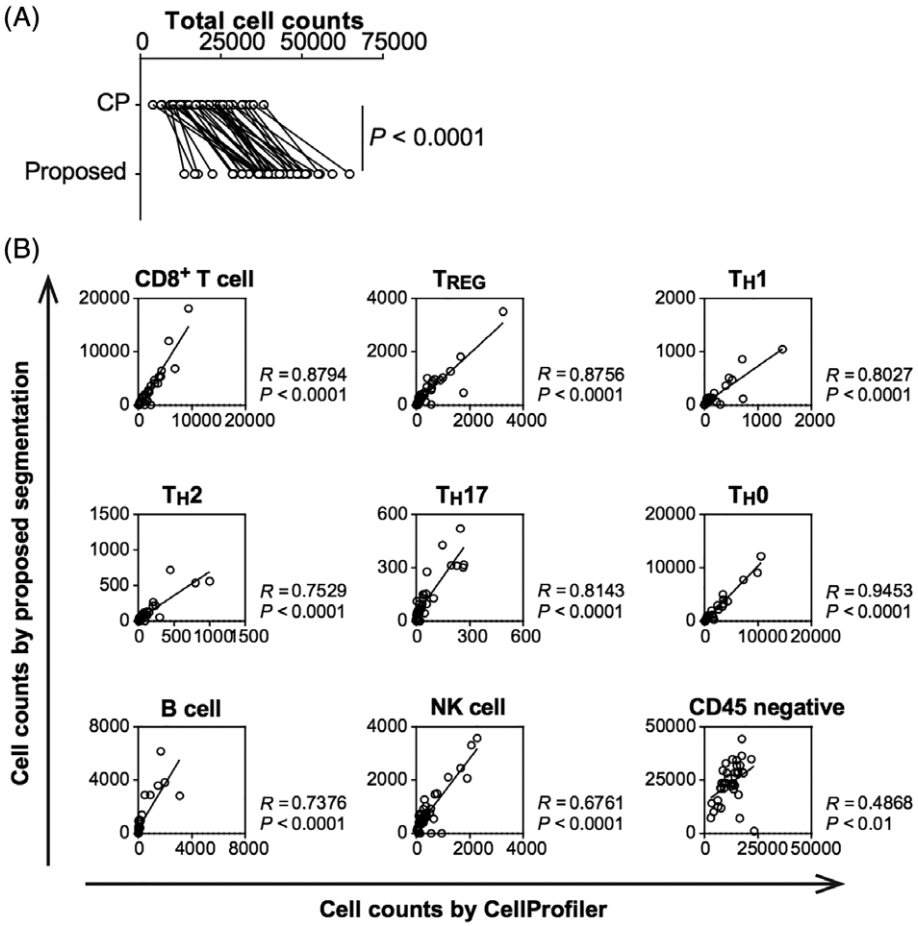

Figure 2. The proposed method improves detection of cell lineages in image cytometry analysis. (**A**) Total cell counts were compared between CellProfiler and the proposed method. Statistical significance was determined via a Wilcoxon signed rank test. (**B**) The numbers of detected cells across various cell lineages were compared between the proposed method (vertical axes) and CellProfiler (horizontal axes). The proposed approach showed better cell detection in most of cell lineages than CellProfiler-based segmentation. Spearman correlation coefficient was used to assess correlations of cell counts among cell lineages.

New Cell Segmentation Method Was Validated for Image Cytometry-Based Quantification

The new cell segmentation method was, then, tested for image cytometry-based quantification and was applied to a previously published tissue microarray of head and neck squamous cell carcinoma ($N = 38$), where immune cell lineages were quantitatively evaluated by CP-based cell segmentation and image cytometry (8). Image cytometry analysis with our new segmentation method was performed, according to matched gating strategies and lineage markers as CP-based segmentation was applied (Supporting Information Fig. S3). Cell percentages of total immune cells quantified by the two different segmentation methods were highly correlated (Supporting Information Fig. S4A), and cell compositions were also preserved regardless of cell segmentation methods (Supporting Information Fig. S4B). Importantly, our new segmentation method showed higher cell numbers in total cells (Fig. 2A) and most cell lineages (Fig. 2B) than CP-based segmentation, which was compatible with higher cell detection as observed in Figure 1. These results validate that the new method preserved cell composition together with better cell detection compared to the previous approach.

Image Cytometry Based on the New Cell Segmentation Method Enables Analysis for Subpopulations of T_H17 Cells

Since the improved cell detection could be advantageous for the detection of rare cell populations, we next focused on relatively rare T_H17 cells, whose role in the tumor microenvironment is still controversial (25). As shown in Figure 3A, the new segmentation method detected a higher frequency of T_H17 cells. The new segmentation method showed a statistically significant increase in T_H17 cell detection across 38 samples in tissue microarray (Fig. 3B), enabling subsequent analyses, including phenotyping and sub-classification. As the presence of subpopulations in T_H17 cells has been shown in a previous report (26), we sought to dissect T_H17 populations based on T-bet and GATA3 expressions, which represent T_H1 and T_H2 phenotypes in helper T-cell subsets, respectively. We observed that T-bet and GATA3 were expressed on T_H17 cells,

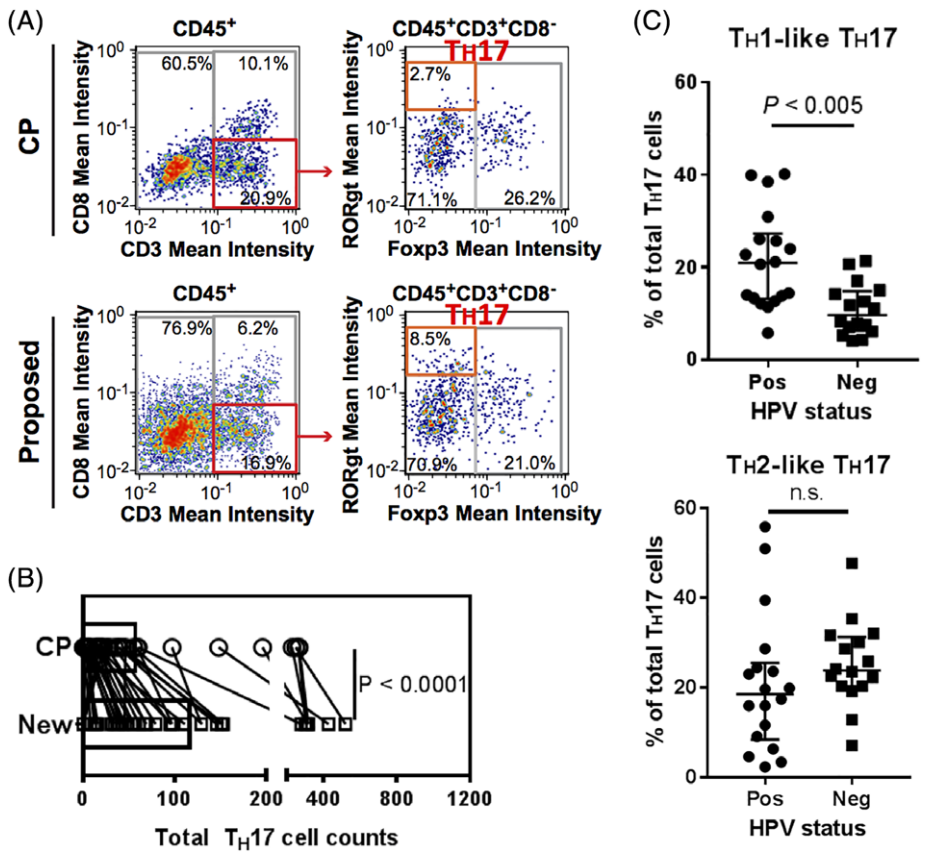

Figure 3. Robust cell segmentation in image cytometry enables subpopulation analysis for T_H17 cells. (**A**) Image cytometry plots are shown to compare frequency of T_H17 cells detected by CellProfiler (top) and the proposed method (new, bottom). (**B**) Total cell counts of T_H17 cells were compared between the two methods. Statistical significance was determined by a Wilcoxon signed rank test. (**C**) Percentages of T_H1-like and T_H2-like subpopulations in total T_H17 cells were compared in HPV-positive (n = 21) and HPV-negative (n = 17) tumors of head and neck squamous cell carcinoma. Statistical significance was determined via Mann–Whitney U test.

supporting the notion that T_H17 cells can be further classified into T_H1- and T_H2-like subpopulations as shown previously (26).

To assess the potential biological contexts of subpopulations of T_H17, we next evaluated the correlation between T_H1 versus T_H2-like phenotypes and the presence of oncogenic human papilloma virus (HPV), which is associated with immunogenic characteristics (27). We found that HPV-positive tumors (N = 21) showed a significantly higher frequency of T_H1-like T_H17 cells than HPV-negative tumors (N = 17), potentially related to differential immune profiles between HPV-positive and -negative tumors (Fig. 3C).

Robust Cell Segmentation Elucidates Imaging-Based T_H17 Cell Phenotypes in Tumor Immune Microenvironment

To further explore the significance of T_H17-associated immune characteristics, subpopulations of T_H17 cells were analyzed in tissue immunological contexts. We analyzed correlations between T_H17 subpopulation profiles and immune cell frequency of 14-different lineages quantified in a previous publication (Table 1). Without stratification into T_H1 versus T_H2 phenotypes, total T_H17 cells showed a positive correlation with $CD8^+$ T cells as well as a negative correlation with $CD163^-$ macrophages, which is suggested by a previous murine study (8). T_H2 cells correlated with T_H2-like T_H17 cells, suggesting an association with T_H2-driven cytokine profiles. Although T_H17 has been known to recruit granulocytes via cascades of cytokine production (28), total T_H17 cells did not correlate with $CD66b^+$ granulocyte infiltration. Interestingly, $CD66b^+$ granulocyte infiltration negatively correlated with T_H1-like T_H17 cells but positively correlated with T_H2-like T_H17 cells, indicating that granulocyte recruitment might be associated with T_H2-polarized phenotypes in T_H17 cells, rather than total T_H17 cells in the cancer microenvironment (Fig. 4A).

The potential relationship between T_H2-like T_H17 cells and $CD66b^+$ granulocytes was further investigated in terms of spatial distribution of the two cell populations in cancer tissues. For validation, we utilized another cohort of HPV-negative head and neck squamous cell carcinoma (N = 9), as we observed a high density of T_H2-like T_H17 cells in HPV-negative tumors (Fig. 3C). Interestingly, T_H2-like T_H17 cells were located close to $CD66b^+$ granulocytes (Fig. 4B,C and Supporting Information Fig. S5A,B). Compared with T_H1-like T_H17 cells, T_H2-like T_H17 cells were significantly closer to

Table 1. Spearman correlations of T_H17 cell densities versus immune cell percentages of total $CD45^+$ cells

	TOTAL TH17			TH1-LIKE TH17			TH2-LIKE TH17		
	SPEARMAN *R*	95% CONFIDENCE INTERVAL	*P* VALUE	SPEARMAN *R*	95% CONFIDENCE INTERVAL	*P* VALUE	SPEARMAN *R*	95% CONFIDENCE INTERVAL	*P* VALUE
vs. CD8	O.6137	0.3573 to 0.7841	<0.0001****	0.6878	0.4642 to 0.829	<0.0001****	0.3805	0.0595 to 0.6302	0.0184*
vs. Treg	0.2953	−0.03669 to 0.5686	0.0719^{ns}	0.308	−0.02273 to 0.578	0.0599^{ns}	0.08274	−0.2526 to 0.4003	0.6214^{ns}
vs. Thl	0.04298	−0.2896 to 0.3663	0.7978^{ns}	0.1935	−0.1441 to 0.4908	0.2444^{ns}	−0.2531	−0.5369 to 0.08216	0.1252^{ns}
vs. Th2	0.4043	0.08743 to 0.6468	0.0118*	0.3093	−0.02127 to 0.579	0.0588^{ns}	0.5279	0.2413 to 0.7298	0.0007***
vs. Th17	0.592	0.3272 to 0.7706	<0.0001****	0.5369	0.2531 to 0.7356	0.0005***	0.3754	0.0536 to 0.6266	0.0202*
vs. Th0	0.2447	−0.09104 to 0.5305	0.1386^{ns}	0.2471	−0.08858 to 0.5323	0.1348^{ns}	−0.01238	−0.3395 to 0.3174	0.9412^{ns}
vs. B cell	0.2825	−0.05066 to 0.5591	0.0857^{ns}	0.2786	−0.05486 to 0.5562	0.0903^{ns}	0.06616	−0.2681 to 0.3862	0.6931^{ns}
vs. NK	−0.3093	−0.579 to 0.02128	0.0588^{ns}	−0.4006	−0.6443 to −0.08305	0.0127*	0.03123	−0.3003 to 0.356	0.8523^{ns}
vs. $CD163^-$ TAM	−0.5171	−0.7228 to −0.2272	0.0009***	−0.4746	−0.6947 to −0.1731	0.0026**	−0.4429	−0.6734 to −0.1339	0.0054**
vs. $CD163^+$ TAM	−0.3116	−0.5807 to 0.01873	0.0568^{ns}	−0.2432	−0.5294 to 0.09263	0.1412^{ns}	−0.1745	−0.4757 to 0.1633	0.2948^{ns}
vs. DC^- $SIGN^+$	−0.2149	−0.5076 to 0.1221	0.195^{ns}	−0.1889	−0.4872 to 0.1487	0.2559^{ns}	−0.07157	−0.3908 to 0.2631	0.6694^{ns}
vs. $CD83^+$ DC	−0.02497	−0.3505 to 0.306	0.8817^{ns}	−0.05691	−0.3783 to 0.2767	0.7343^{ns}	0.1584	−0.1794 to 0.4627	0.3423^{ns}
vs. $CD66b^+$	0.05749	−0.2762 to 0.3788	0.7317^{ns}	−0.08761	−0.4044 to 0.248	0.6009^{ns}	0.372	0.04957 to 0.6242	0.0215*
vs. Mast cell	−0.1116	−0.4245 to 0.2251	0.5048^{ns}	−0.1408	−0.4485 to 0.1967	0.3991^{ns}	−0.1253	−0.4358 to 0.2119	0.4536^{ns}
vs. Other $CD45^+$	−0.5602	−0.7505 to −0.284	0.0003***	−0.4546	−0.6813 to −0.1483	0.0041**	−0.6055	−0.779 to −0.3458	<0.0001****

*: P < 0.05, **: P < 0.01, ***: P < 0.001, ****: P < 0.0001, ns: not significant.

$CD66b^+$ granulocytes (Fig. 4D). In order to investigate potential bias derived from cell frequency, we also evaluated local cell densities surrounding $CD66b^+$ granulocytes (Fig. 4E and Supporting Information Fig. S5C) with varying distance radius (*r*) and the *k*-nearest neighbors (*k*). Notably, two peaks were observed only in T_H2-like T_H17 cell density but not in T_H1-like T_H17 cells, suggesting the presence of certain biological interactions. The population density of $CD66b^+$ granulocytes is higher in T_H2-like T_H17 cell regions compared with T_H1-like T_H17 cell regions (first peak in Fig. 4E). As distance radius (*r*) increases, we can see the second peak of $CD66b^+$ granulocyte density in T_H2-like T_H17 cell regions, although high population densities of $CD66b^+$ granulocytes only exist at the small distance region (r = 1,000) in T_H1-like T_H17 cell regions. Supporting Information Figure S5C shows the population density of $CD66b^+$ granulocyte with different numbers of nearest neighbor cells (*k*). For small numbers of k (<100), the high density peak location of T_H2-like T_H17 cells is shorter than the T_H1-like T_H17 cells peak. For larger numbers of k (>100), within the same radius from each cell type, the overall $CD66b^+$ granulocyte density from T_H2-like T_H17 cells is higher than T_H1-like T_H17 cells' density. These suggest the presence of a spatial relationship between T_H2-like T_H17 cells and $CD66b^+$ granulocytes. Together, increased detection of rare T_H17 cell populations based on this new segmentation method potentially contributes to in-depth immune profiling and spatial association, leading to further tissue-based biomarker exploration.

Discussion

We proposed a simple but effective methodology for fully automated and robust single-cell segmentation with reduced cost of parameter tuning. To do this, we extract useful morphological features from the image and group individual pixels using unsupervised clustering based on similar features to segment nuclei. Notably, our segmentation has been optimized with images stained only by hematoxylin, which lack the cytoplasmic staining usually provided by eosin staining (Fig. 1A–C). This provides robust segmentation results across different cancer histological types with the same parameter setting. However, a limitation could be the dependence on image quality of hematoxylin staining, where low-quality images lead to failure in cell detection or over/under-segmentation even with our segmentation approach. Our data indicate that the proposed segmentation shows better detection of cells with higher sensitivity and specificity in three different cancer histological types. While we observed preserved cell ratio and composition (Supporting Information Fig. S4), our robust and effective cell detection enabled better detection of relatively rare T_H17 cells (Fig. 3A,B). Improved cell detection was particularly significant for sub-population analysis, which depends on a sufficiently large cell population.

Given that the immunological properties of T_H17 cells have not been fully elucidated, we sought to dissect the characteristics of T_H17 cells based on multiplex IHC and image cytometry analysis. Then, sub-populations reflecting T_H1

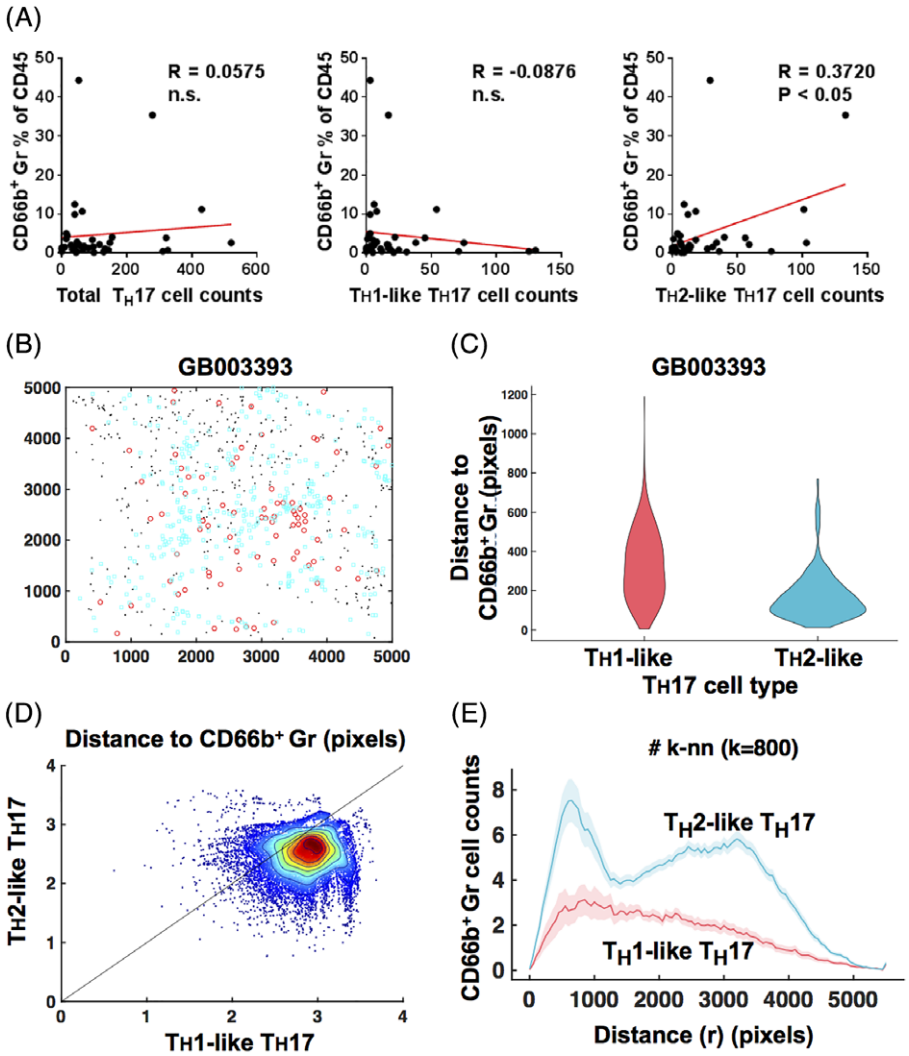

Figure 4. Image cytometry with robust cell segmentation reveals T_H17-associated tissue immunological characteristics. (**A**) Spearman correlations were analyzed between T_H17 cell density and cell percentages of CD66b$^+$ granulocytes. (**B**) A representative case depicts close proximity between CD66b$^+$ granulocytes (black dots) and T_H2-like (cyan circles) but not T_H1-like (red circles) T_H17 cells. (**C**) Cell-to-cell distances from CD66b$^+$ granulocyte to nearest T_H1-like T_H17 versus T_H2-like T_H17 cells were compared. Statistical significance was determined via Wilcoxon signed rank test. (**D**) A density plot presents overall distance distribution comparing distance from CD66b$^+$ granulocytes to nearest T_H1-like T_H17 (*x*-axis) and T_H2-like T_H17 (*y*-axis) across nine cases. (**E**) Microregional cell densities of T_H1-like T_H17 and T_H2-like T_H17 cells were shown depending on distance from nearest CD66b$^+$ granulocytes (*x*-axis).

versus T_H2 balance correlated with HPV status in head and neck cancer, where the immunogenic background of viral-related malignancy could be associated with T_H1-based anti-tumor immunity (Fig. 3C). Interestingly, our cell density analysis revealed that T_H2-like T_H17 cells correlated with granulocytes. This observation was further validated in another cohort of HPV-negative head and neck cancer tissues. A major advantage of image cytometry is that it retains tissue context with preserved tissue architectural information, enabling spatial relationship analysis. Thus, we demonstrated that T_H2-like but not T_H1-like T_H17 cells showed closer proximities to granulocytes (Fig. 4). In addition to previous experimental results, our data provide human tissue-based evidence of distinctive T_H17 subpopulations and spatial association with other immune subsets. Taken together, stratification of T_H2-like T_H17 cells can contribute to further understanding of the roles of T_H17 subpopulations in tumor immunity. Simultaneously, our data highlight the capability for improved cell segmentation and image cytometry, enabling better resolution of immune complexity analysis with spatial relationships.

In this study, we demonstrated a proof of concept where combinations of general image processing techniques such as Gabor/Gaussian filters and mathematical morphology operations could improve work flow for single-cell segmentation without dependence on manual parameter tuning. As improvement in usability comes after the early stage of proving concepts, we did not release a single package at this time. However, we are currently implementing our pipeline as a user command line tool and integrating it into Galaxy (scientific workflow system) to improve usability. This requires

software development so that a wide range of users can run our pipeline upon release.

In summary, this study proposed a significantly improved cell segmentation approach for image cytometry analysis based on multiplex IHC, which increased the detection of rare cell populations including T_H17 subpopulations. The image cytometry analyisis with improved cell detection may provide in-depth immune profiling with maintained spatial association, thereby leading to further tissue-based biomarker exploration.

Author Contributions

TT, GT, and YHC conceived and designed the study. GB performed the experiments, TT, RK, and CM annotated the sample images. TT, GT, and YHC analyzed the data and interpreted the results. VA and SS ran CP-based and ImageJ/-FIJI-based segmentation. TT, GT, and YHC drafted the manuscript. DRC reviewed manuscript, and JWG and LMC supervised and oversaw the study. All authors critically revised the manuscript and provided intellectual content.

Literature Cited

1. Yuan J, Hegde PS, Clynes R, Foukas PG, Harari A, Kleen TO, Kvistborg P, Maccalli C, Maecker HT, Page DB, Robins H, Song W, Stack EC, Wang E, Whiteside TL, Zhao Y, Zwierzina H, Butterfield LH, Fox BA. Novel technologies and emerging biomarkers for personalized cancer immunotherapy. J Immunother Cancer. 2016;4:3. Epub 2016/01/21. doi: https://doi.org/10.1186/s40425-016-0107-3. PubMed PMID: 26788324; PMCID: PMC4717548.
2. Scutti JAB. Importance of immune monitoring approaches and the use of immune checkpoints for the treatment of diffuse intrinsic pontine glioma: From bench to clinic and vice versa (Review). Int J Oncol. 2018;52(4):1041-56. Epub 2018/02/28. doi: https://doi.org/10.3892/ijo.2018.4283. PubMed PMID: 29484440; PMCID: PMC5843403.
3. Assi HI, Kamphorst AO, Moukalled NM, Ramalingam SS. Immune checkpoint inhibitors in advanced non-small cell lung cancer. Cancer. 2018;124(2):248-61. Epub 2017/12/07. doi: https://doi.org/10.1002/cncr.31105. PubMed PMID: 29211297.
4. Varn FS, Wang Y, Mullins DW, Fiering S, Cheng C. Systematic Pan-Cancer Analysis Reveals Immune Cell Interactions in the Tumor Microenvironment. Cancer Res. 2017;77(6):1271-82. Epub 2017/01/28. doi: https://doi.org/10.1158/0008-5472.CAN-16-2490. PubMed PMID: 28126714; PMCID: PMC5798883.
5. Dao D, Fraser AN, Hung J, Ljosa V, Singh S, Carpenter AE. CellProfiler Analyst: interactive data exploration, analysis and classification of large biological image sets. Bioinformatics. 2016;32(20):3210-2. Epub 2016/06/30. doi: https://doi.org/10.1093/bioinformatics/btw390. PubMed PMID: 27354701; PMCID: PMC5048071.
6. Caicedo JC, Singh S, Carpenter AE. Applications in image-based profiling of perturbations. Curr Opin Biotechnol. 2016;39:134–142. Epub 2016/04/19. https://doi.org/10.1016/j.copbio.2016.04.003 PubMed PMID: 27089218.
7. Young Hwan C, Thibault G, Azimi V, Johnson B, Jorgens D, Link J, Margolin A, Gray JW. Quantitative analysis of histological tissue image based on cytological profiles and spatial statistics. Conf Proc IEEE Eng Med Biol Soc. 2016;2016:1175-8. Epub 2016/01/01. doi: https://doi.org/10.1109/EMBC.2016.7590914. PubMed PMID: 28324942.
8. Tsujikawa T, Kumar S, Borkar RN, Azimi V, Thibault G, Chang YH, Balter A, Kawashima R, Choe G, Sauer D, El Rassi E, Clayburgh DR, Kulesz-Martin MF, Lutz ER, Zheng L, Jaffee EM, Leyshock P, Margolin AA, Mori M, Gray JW, Flint PW, Coussens LM. Quantitative Multiplex Immunohistochemistry Reveals Myeloid-Inflamed Tumor-Immune Complexity Associated with Poor Prognosis. Cell Rep. 2017;19(1):203-17. Epub 2017/04/06. doi: https://doi.org/10.1016/j.celrep.2017.03.037. PubMed PMID: 28380359; PMCID: PMC5564306.
9. Kamentsky L, Jones TR, Fraser A, Bray MA, Logan DJ, Madden KL, Ljosa V, Rueden C, Eliceiri KW, Carpenter AE. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. Bioinformatics. 2011;27(8):1179-80. Epub 2011/02/26. doi: https://doi.org/10.1093/bioinformatics/btr095. PubMed PMID: 21349861; PMCID: PMC3072555.
10. Fogel I, Sagi D. Gabor filters as texture discriminator. Biological cybernetics. 1989;61(2):103–113.
11. Kesselring R, Thiel A, Pries R, Trenkle T, Wollenberg B. Human Th17 cells can be induced through head and neck cancer and have a functional impact on HNSCC development. Br J Cancer. 2010;103(8):1245-54. Epub 2010/09/30. doi: https://doi.org/10.1038/sj.bjc.6605891. PubMed PMID: 20877351; PMCID: PMC2967064.
12. Veta M, Pluim JP, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: a review. IEEE Trans Biomed Eng. 2014;61(5):1400-11. Epub 2014/04/25. doi: https://doi.org/10.1109/TBME.2014.2303852. PubMed PMID: 24759275.
13. Ruiz P, Nathanson R, Kastner T. Pertussis immunization patterns in special care nursery graduates. J Dev Behav Pediatr. 1991;12(1):38–41. Epub 1991/02/01. PubMed PMID: 2016401.
14. Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B. Histopathological image analysis: a review. IEEE Rev Biomed Eng. 2009;2:147-71. Epub 2009/01/01. doi: https://doi.org/10.1109/RBME.2009.2034865. PubMed PMID: 20671804; PMCID: PMC2910932.
15. Aswathy M, Jagannath M. Detection of breast cancer on digital histopathology images: present status and future possibilities. Informatics in Medicine Unlocked. 2017;8:74–79.
16. Rujuta O, Vyavahare A. Review of Nuclei Detection, Segmentation in Microscopic Images. J Bioengineer & Biomedical Sci. 2017;7(227):2.
17. Wang Z, Li H. Generalizing cell segmentation and quantification. BMC bioinformatics. 2017;18(1):189.
18. Wang CW, Fennell D, Paul I, Savage K, Hamilton P. Robust automated tumour segmentation on histological and immunohistochemical tissue images. PLoS One. 2011;6(2):e15818. Epub 2011/03/10. doi: 10.1371/journal.pone.0015818. PubMed PMID: 21386898; PMCID: PMC3046129.
19. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. Nat Methods. 2012;9(7):676-82. Epub 2012/06/30. doi: https://doi.org/10.1038/nmeth.2019. PubMed PMID: 22743772; PMCID: PMC3855844.
20. Otsu N. A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics. 1979;9(1):62–66.
21. Najman L, Schmitt M. Geodesic saliency of watershed contours and hierarchical segmentation. IEEE Transactions on pattern analysis and machine intelligence. 1996;18(12):1163–1173.
22. Young Hwan C, Thibault G, Madin O, Azimi V, Meyers C, Johnson B, Link J, Margolin A, Gray JW. Deep learning based Nucleus Classification in pancreas histological images. Conf Proc IEEE Eng Med Biol Soc. 2017;2017:672-5. Epub 2017/10/25. doi: https://doi.org/10.1109/EMBC.2017.8036914. PubMed PMID: 29059962.
23. Barber CB, Dobkin DP, Huhdanpaa H. The quickhull algorithm for convex hulls. ACM Transactions on Mathematical Software (TOMS). 1996;22(4):469–483.
24. Paz P, Piqueras JR, Tizado E. A program for the application of the radial distribution function to cluster analysis in cell biology. Bioinformatics. 1992;8(4):307–309.
25. Bailey SR, Nelson MH, Himes RA, Li Z, Mehrotra S, Paulos CM. Th17 cells in cancer: the ultimate identity crisis. Front Immunol. 2014;5:276. Epub 2014/07/06. doi: https://doi.org/10.3389/fimmu.2014.00276. PubMed PMID: 24987392; PMCID: PMC4060300.
26. Geginat J, Paroni M, Maglie S, Alfen JS, Kastirr I, Gruarin P, De Simone M, Pagani M, Abrignani S. Plasticity of human CD4 T cell subsets. Front Immunol. 2014;5:630. Epub 2015/01/08. doi: https://doi.org/10.3389/fimmu.2014.00630. PubMed PMID: 25566245; PMCID: PMC4267263.
27. Thurlow JK, Pena Murillo CL, Hunter KD, Buffa FM, Patiar S, Betts G, West CM, Harris AL, Parkinson EK, Harrison PR, et al. Spectral clustering of microarray data elucidates the roles of microenvironment remodeling and immune responses in survival of head and neck squamous cell carcinoma. J Clin Oncol. 2010;28(17):2881–2888. Epub 2010/05/12. https://doi.org/10.1200/JCO.2009.24.8724 PubMed PMID: 20458058.
28. Romagnani S. Human Th17 cells. Arthritis Res Ther. 2008;10(2):206. Epub 2008/05/10. doi: https://doi.org/10.1186/ar2392. PubMed PMID: 18466633; PMCID: PMC2453756.