RESEARCH ARTICLE

Graph-based description of tertiary lymphoid organs at single-cell level

Nadine S. Schaadt[1], **Ralf Schönmeyer**[2], **Germain Forestier**[3], **Nicolas Brieu**[2], **Peter Braubach**[1], **Katharina Nekolla**[2], **Michael Meyer-Hermann**[4,5], **Friedrich Feuerhake**[1,6] *

1 Institute for Pathology, Hannover Medical School, Hannover, Germany, **2** Definiens AG, Munich, Germany, **3** IRIMAS, University of Haute Alsace, Mulhouse, France, **4** Systems Immunology and Integrated Centre of Systems Biology, Helmholtz Centre for Infection Research, Braunschweig, Germany, **5** Institute for Biochemistry, Biotechnology and Bioinformatics, TU Braunschweig, Braunschweig, Germany, **6** Institute for Neuropathology, University Clinic Freiburg, Freiburg, Germany

* Feuerhake.Friedrich@mh-hannover.de

Abstract

Our aim is to complement observer-dependent approaches of immune cell evaluation in microscopy images with reproducible measures for spatial composition of lymphocytic infiltrates. Analyzing such patterns of inflammation is becoming increasingly important for therapeutic decisions, for example in transplantation medicine or cancer immunology. We developed a graph-based assessment of lymphocyte clustering in full whole slide images. Based on cell coordinates detected in the full image, a Delaunay triangulation and distance criteria are used to build neighborhood graphs. The composition of nodes and edges are used for classification, e.g. using a support vector machine. We describe the variability of these infiltrates on CD3/CD20 duplex staining in renal biopsies of long-term functioning allografts, in breast cancer cases, and in lung tissue of cystic fibrosis patients. The assessment includes automated cell detection, identification of regions of interest, and classification of lymphocytic clusters according to their degree of organization. We propose a neighborhood feature which considers the occurrence of edges with a certain type in the graph to distinguish between phenotypically different immune infiltrates. Our work addresses a medical need and provides a scalable framework that can be easily adjusted to the requirements of different research questions.

Author summary

In this study, we developed a workflow to detect and classify immune infiltrates in gigapixel microscopy images. It allowed us to measure the degree of organization in lymphocyte clusters with graph-based features and finally to distinguish between tertiary lymphoid organs and other infiltrates. As a clinically relevant use case, we applied it to three different types of tissues and diseases. Our method addresses the need for observer-independent, large-scale evaluation of immune cell patterns and is a prerequisite to capture

spatial composition of immune cells that can be used to parameterize mathematical models in systems immunology.

This is a *PLOS Computational Biology* Methods paper.

Introduction

Visual evaluation of immune infiltration is the current state-of-the-art in many fields of translational research, such as transplant pathology and immuno-oncology. Despite efforts to benchmark immune cell evaluation, most approaches remain highly observer-dependent and large-scale quantitative slide consideration is limited by the time and availability of trained pathologists. A particularly difficult task is to distinguish between early stages of tertiary lymphoid organs (TLOs; also known as tertiary lymphoid structures, ectopic lymphoid tissue) and other immune infiltrations [1]. Nevertheless, these structures play an important role in medical biomarker research.

TLOs are highly organized structures that consist of distinct B-cell clusters and surrounding T-cell compartments with a supporting function. In contrast to primary and secondary lymphoid tissue in anatomically predefined locations (lymph node, tonsil) [2], these structures occur in ectopic sites of chronic inflammation [3]. Ectopic lymphoid tissues were termed dysmorphic [4] and simulations suggest that the diversity of shapes observed in histology is a result of highly dynamic shape of ectopic follicles [5]. Mature TLOs can form functional germinal centers [6, 7] with high endothelial venules (HEVs) and follicular dendritic cells (FDCs), where further lymphocytes can enter the lymphoid organs from blood [8]. They are implicated with B-cell maturation in chronic inflammation [3] and have different clinical relevance in cancer, transplantation, and inflammation.

In cancer, TLOs and corresponding HEVs, which are frequently occurring in tumor regions, are correlated to a longer disease-free survival [9–11]. Breast cancer cases with B-cell containing structures have a better prognosis than those with T-cell dominant immune infiltrates [12].

In transplantation medicine, TLOs are rarely observed in renal biopsies early after transplantation, but frequently in explanted non-functioning kidneys with chronic inflammation due to allograft rejection [13]. This suggests a role in chronic responses to persistent donor antigens. In routine, renal biopsies are graded by pathologists based on Banff, a descriptive and partially formalized classification system [14]. Immune cells are only scored in tubuli and blood vessels. However, immune infiltration in the interstitium can have a prognostic value and could improve the current evaluation [15, 16].

In the respiratory tract of human adults, TLOs are also called inducible bronchus-associated lymphoid tissue (iBALT) due to a similar structure as gut-associated lymphoid tissue (GALT) belonging to mucosa-associated lymphatic tissues (MALT). However, iBALT in human lungs is mediated by chronic inflammatory diseases and does not occur in healthy cases in contrast to BALT in other mammals [17–19].

In order to provide reproducible, observer-independent, and quantitative evaluation of immune infiltrates, image analysis approaches could automatically detect regions of interest (ROIs) [20–22] and classify cells based on immunohistochemistry [23]. The increasing medical need for evaluation of particular cell types in anatomically or immunologically defined ROIs falls into an era of massive advance in machine learning (ML), with deep learning and

pixel-based ML (as opposed to feature-based approaches to ML) holding great promise for identification, classification, and quantitative assessment of relevant patterns in medical images including microscopy [24–28]. Evolutionary algorithms have successfully been developed, and applied to different tasks including hyperspectral remote sensing image analysis, classification in different common benchmark data sets, and brain tumor medical imaging [29–32]. While the success and great potential of ML is undisputed, there is still a rationale to consider approaches of manually curated features according to domain-specific knowledge as they may help to biologically understand the result and consequently increase the acceptance by medical experts. Methods broadly considered by pathologists are descriptive, purely density-based, or include spatial characteristics by distance metrics [33, 34]. In contrast, graph-based concepts could include complex neighborhood characteristics [35–41].

To distinguish between infiltrates with varying B-cell composition, we herein combine image analysis in whole slide images (WSIs) with graph-based cluster analysis, enabling large-scale studies and robustness. Our concept captures neighborhood relations between T- and B-lymphocytes and classifies immune infiltrates into categories by their spatial organization. In particular, we developed a feature describing the neighborhood organization to detect TLOs. We applied the method to three examples of chronic adaptive immune response to persisting antigens; (1) breast cancer (2) renal transplantation (3) chronic lung infection.

Results

Workflow

Our workflow (Fig 1) detects and characterizes lymphocyte clusters in the full WSI aiming at overcoming the current lack of reproducible, quantitative, and context-specific methods of immune cell evaluation. For this, we constructed cell graphs based on Delaunay triangulation. We applied the method to (1) label immune infiltrates based on their degree of organization using a support vector machine (SVM) and to (2) compare their occurrence and characteristics in different diseases.

Graph-based description

Neighborhood graphs. For each ROI, neighborhood graphs $G = (V, E)$ in which a node $v = (x_v, y_v) \in V$ represents a certain cell with the coordinates (x_v, y_v) and $e = (v_1, v_2) \in E$ an edge between the nodes v_1 and v_2 are built (Fig 2). At first, a Delaunay triangulation is calculated for all cell coordinates independent of their specific cell type.

Our concept considers that pathologists define infiltrates based on the common practice to delineate grouped cells relying on visual perception. Therefore, we introduce a threshold t to define clusters correspondingly. Then, all sides of a Delaunay triangle are included as edges if the radius of its circumcircle is smaller than t. All resulting connected components are considered as independent neighborhood graphs, representing an infiltrate.

Features of infiltrates. An edge $e = (v_1, v_2)$ is labeled as α if v_1 and v_2 have the same cell type. All remaining edges are labeled based on the number of cells of a certain type A in the direct neighborhood of the corresponding A-cell (Fig 2E; A-type illustrated in red), i.e.

$$E_{\gamma_i} = \left\{ e = (v_1, v_2) \ \mid \ e \in E \backslash E_\alpha \wedge v_1 \in V_A \text{ has } i \cdot \alpha_A \text{ edges} \right\}.$$

In an application case, the predominant cell type could be chosen as A. We defined the edges as α or γ to investigate cell clusters with same type and their neighborhood for indications of organization. For example, this may differentiate between a B-cell zone of a TLO and other types of immune infiltrates. The following features characterize each neighborhood graph:
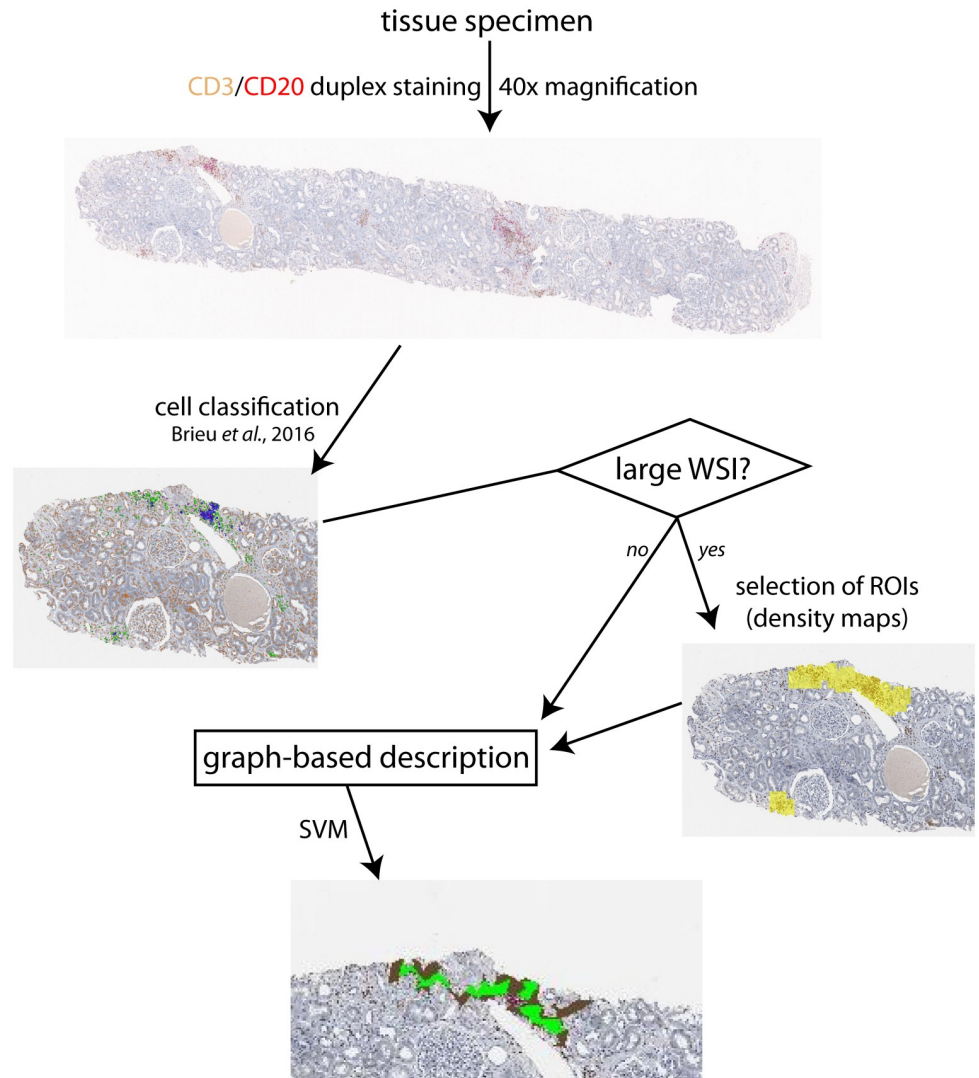
Fig 1. General workflow. CD3[+] T- and CD20[+] B-cells are identified in whole slide images (WSIs). In large WSIs, regions of interest (ROIs) could be selected based on lymphocyte density maps. In each ROI/small WSI, neighborhood graphs are built by detected cells, results visualized by concave hulls.

1. number of nodes $|V|$, $|V_A|/|V|$

2. number of edges with certain type $|E|$, $|E_\alpha|$,

3. relative number of edges between A-cells $|E_{\alpha_A}|/|E|$,

4. homogeneity $H = |E_\alpha|/|E|$,

5. TLO-like organization $\kappa(a)$,

$$\kappa(a) = \left(|E| - |E_\alpha| - \sum_{j=0}^{a} |E_{\gamma_j}| \right) \left(|E| - |E_\alpha| \right)^{-1} \tag{1}$$

κ is higher for graphs that form separated B- and T-cell zones as the case for TLOs (for A = B-cells).

Fig 2. Graph workflow. A: Blue and red dots represent cells of different phenotypes. B: Cells' Delaunay graph. C: Neighborhood graphs G_1, G_2, where triangles with large circumcircle radius were excluded. Edges between cells of different phenotype are illustrated as dashed lines. D: Classification based on the composition of nodes and edges. E: Illust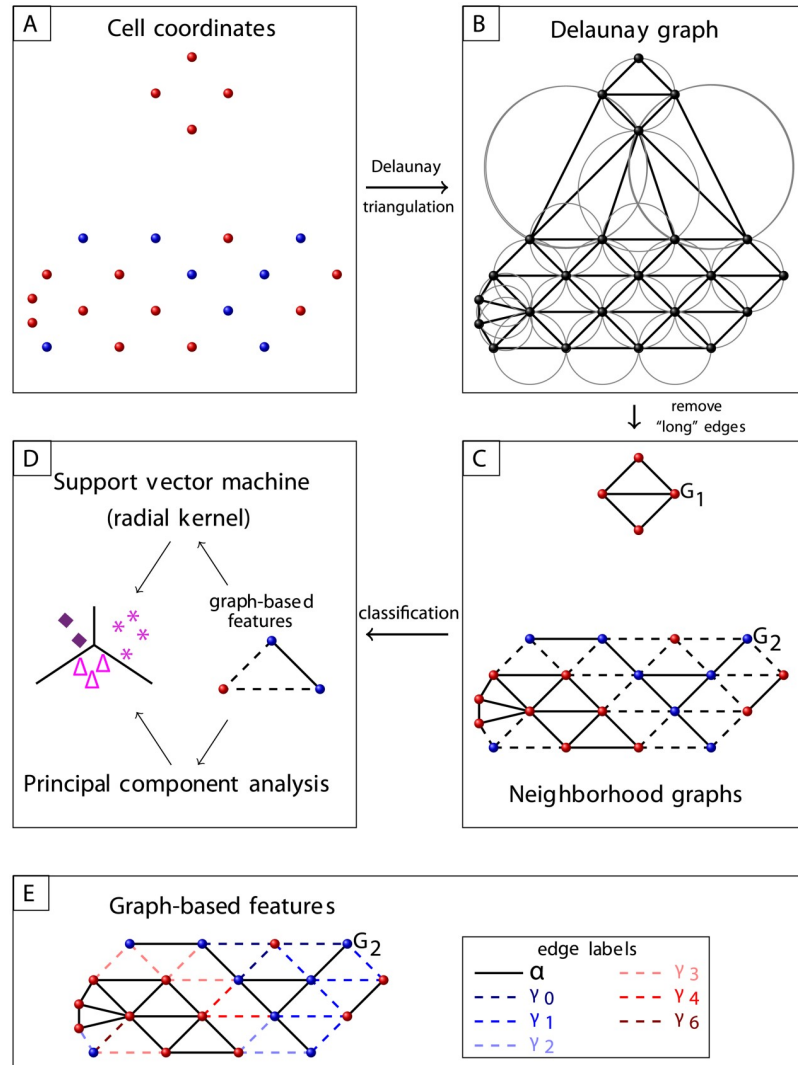ration of edge labels used in the feature $\kappa(a)$, see Section Features of infiltrates; nodes of type A are represented in red. For the illustrated graph $G_2 = (V, E)$, $\kappa(2) = \frac{|E_{\gamma_3}| + |E_{\gamma_4}| + |E_{\gamma_6}|}{|E| - |E_\alpha|} = 0.4$ and $\kappa(5) = \frac{|E_{\gamma_6}|}{|E| - |E_\alpha|} = 0.05$.

6. global clustering coefficient C,

$$C = \frac{1}{|V|} \sum_{l \in V} \frac{2 \cdot n_l}{|N_l| \cdot (|N_l| - 1)} \tag{2}$$

where n_l is the number of actual edges between the neighbors N_l of node l (i.e., $v \in N_l$ **if** $(v, l) \in E$) and $|N_l|$ is the number of neighbors of l [42].
We assume that density and connectivity of cell groups may potentially reflect biological interactions and therefore included C as a standard approach to analyze clusters.

7. degree distribution $P(k) = n_k/|V|$ with n_k number of nodes with k edges [42],

8. average degree $\langle K \rangle = 2|E|/|V|$ [42], and

9. average Euclidean distance between all nodes.

Clustering and classification of infiltrates

A pathologist labeled 700 automatically identified infiltrates based on the corresponding image area in the same staining into "T-cell area" with only a small number of B-cells, "mixed, unstructured" with no obvious connections between present B-cells, "intermediately organized" with some B-cell connections, or "TLO-like" with clustered B-cells. This ground truth is used to train and test classifier based on features extracted from a neighborhood graph. To select powerful features, we applied a principal component analysis (PCA) to study the data including $|V|$, $|E|$, $|E_{\alpha_A}|/|E|$, $|V_A|/|V|$, $\kappa(2)$, $\kappa(5)$, H, C, $\langle K \rangle$, and the average Euclidean distance between all nodes with B-cells designated as A as features. Then, an SVM with a radial kernel was used in a 5-fold cross-validation including $|E|$, $|E_{\alpha_A}|/|E|$, $|V_A|/|V|$, $\kappa(2)$, and $\kappa(5)$, see Eq (1), with B-cells designated as A as features.

For each class, we measured the sensitivity (true positive rate TPR), the specificity (true negative rate TNR),

$$TPR = \frac{TP}{TP + FN}, \qquad TNR = \frac{TN}{TN + FP},$$

the positive (precision PPV) and negative (NPV) predictive value,

$$PPV = \frac{TP}{TP + FP}, \qquad NPV = \frac{TN}{TN + FN},$$

the accuracy (ACC),

$$ACC = \frac{TP + TN}{TP + TN + FP + FN},$$

and the F_1 score

$$F_1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}$$

as average over each of the five test sets in the cross validation with the numbers of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Optimization of infiltrate detection

Our method relies on the classification of cells into clusters based on a critical distance that separates groups of cells into different clusters. This distance is associated with the circumcircle radii cutoff t of triangles defined by the cell centers (see Section Neighborhood graphs). As this critical distance is not know a priori, a phenomenological approach based on the experience of a pathologist is used to estimate its value. The pathologist delineated several immune infiltrates of each category in 14 images of different tissues, yielding 259 objects with an average area of 66533 μm². These drawn structures are in the following considered as ground truth for infiltrates. The rationale for this decision is that infiltrates delineated by a pathologist are generally based on the assumption that an immune infiltrate is a local accumulation of immune cells, characterized either as proliferation of motile immune cells or invasion induced by local signaling.

A sharp delineation of an infiltrate neglects the fact that in most cases they are surrounded by further lymphocytes (distance to closest lymphocyte inside the infiltrate in the same range

as between two lymphocytes inside the infiltrate). In some cases, these lymphocytes of the neighborhood belong to another annotated infiltrate. Obviously, a low radius threshold destroys infiltrates into small fragments, whereas high cutoffs consider large areas as a single infiltrate (S2).

We have chosen the smallest threshold such that inside a pathologist's cluster (ground truth) the graph is not disconnected into several connected components. The best approximation based on a comparison between graphs constructed based on different thresholds and the ground truth reached $t = 45$ pixels (≈ 11 μm), resulting in a median edge length of about 11 μm compared to a median length of about 32 μm for removed edges between cell centers. This refers to interactions that require a direct contact between the cells.

Validation of clustering and classification

We aim at detecting and comprehensively describing all infiltrates in a full WSI with an increasing B-cell organization from T-cell areas to TLO-like structures (Fig 3). TLOs are biologically defined as distinct B- and T-cell compartments with existence of HEVs and FDCs, i.e. by markers and cytokines that are not part of routine slide evaluation. In clinically established immune cell stainings, there are no commonly accepted criteria, especially for less defined early stages of dysmorphic forms of TLOs [4, 5]. Our approach includes a gold-standard labeling based on a state-of-the-art, visual assignment of infiltrates by a pathologist, which is used for training/testing of standard methods. Herein, we used PCA and/or SVM; whereas depending on the specific application case other classifiers or simple thresholding could be more useful (see S3).

A PCA confirmed a possible separation of the data into the independent classes T-cell area, mixed, unstructured regions, and TLO-like structures with a slight overlap at the borders (Fig 4). Intermediately organized structures showed an overlap with TLO-like areas. In general, the separation between the four classes was visible in all diseases but more pronounced in lung and breast cases than in kidney. Therefore, we considered the same classes in the SVM (see S4).

A 5-fold cross validation (including 101 infiltrates in breast cancer cases, 259 in kidneys, and 340 in lung specimens) for the SVM showed overall quite high F_1 scores, see Table 1. For intermediately organized structures, the TPR and PPV are remarkable small; both driven by a small number of TP. Besides the relative small size of training data for intermediately organized ($\sim 12\%$) and TLO-like ($\sim 11\%$) structures, this very low number of TP for intermediately organized structures reflects the difficulty to separate early stages from mature TLOs (about three-quarter of these FN are classified as mixed, unstructured and a quarter as TLO-like). Obviously, this effect also reduces the PPV of mixed, unstructured. An independent training and test phase for each tissue type resulted in similar overall accuracies (see S7).

Characteristics of the classes

We considered different tissue types with totally different underlying diseases as oncoimmunology, allograft rejection, and chronic inflammation in bronchiectasis. The above presented SVM successfully identified infiltrates in all these tissues. As expected, the major difference was in the amount (all breast and lung specimens and 128 of 160 renal cases contained infiltrates) and composition of inflammatory infiltrates regarding the four classes (Fig 5). In general, the frequency of T-cell areas was highest, however, bronchiectasis cases contained a larger fraction of TLO-like structures compared to the other tissue types.

To better understand the importance of each feature and the specific characteristics of the infiltrates, we inspected feature weights in SVMs and measured correlations between feature
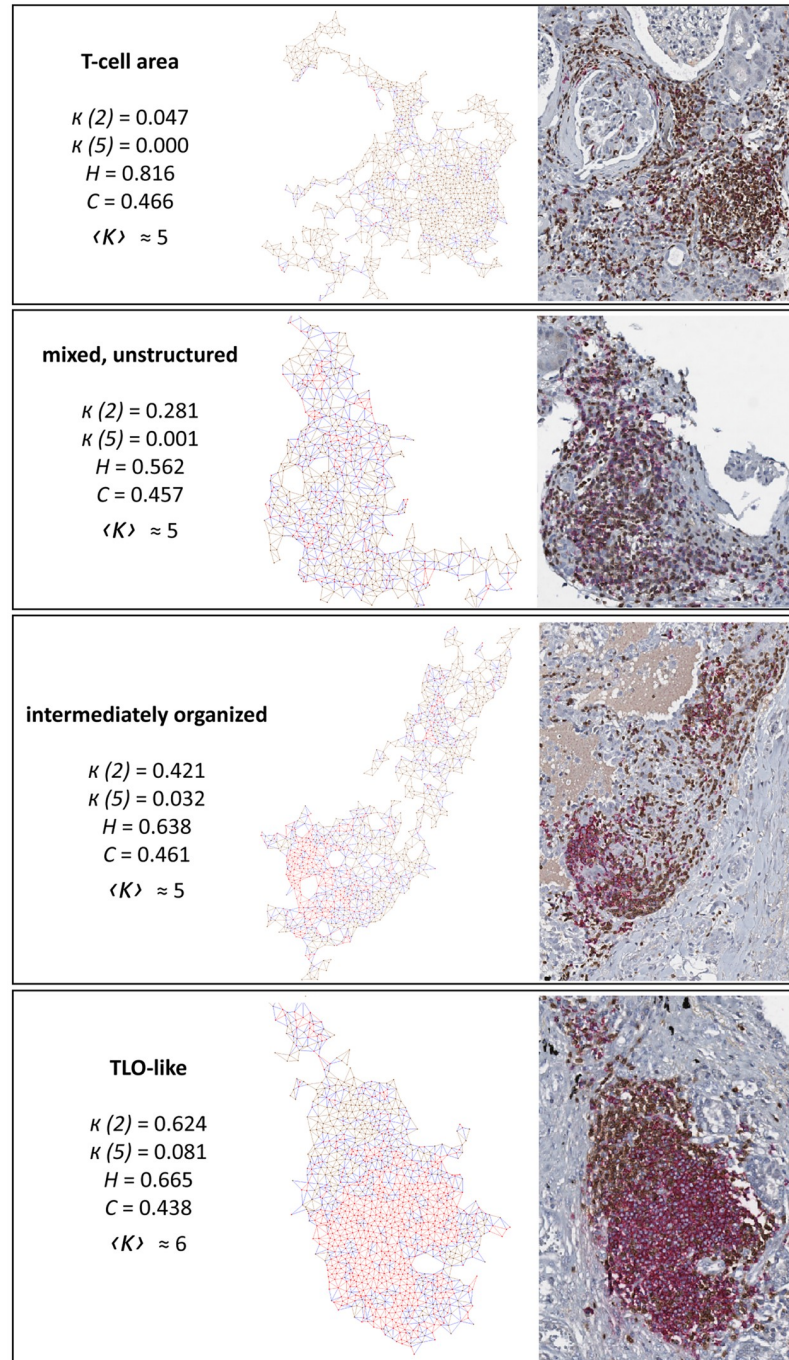
Fig 3. Examples of identified infiltrates. Left: as graph representation. Right: as histological image. T-cells and their links in brown, B-cells in red, edges between T- and B-cells in blue.

and ground truth as well as between features (see S3). The classification was mainly driven by the TLO organization features $\kappa(2)$ and $\kappa(5)$, in contrast the number of edges had only a very few impact. Further, we systematically analyzed the distribution of all features, explained in Section Features of infiltrates. Kruskal–Wallis tests showed significant differences between all classes for all tissue types based on all these features using a significance level of 0.001.
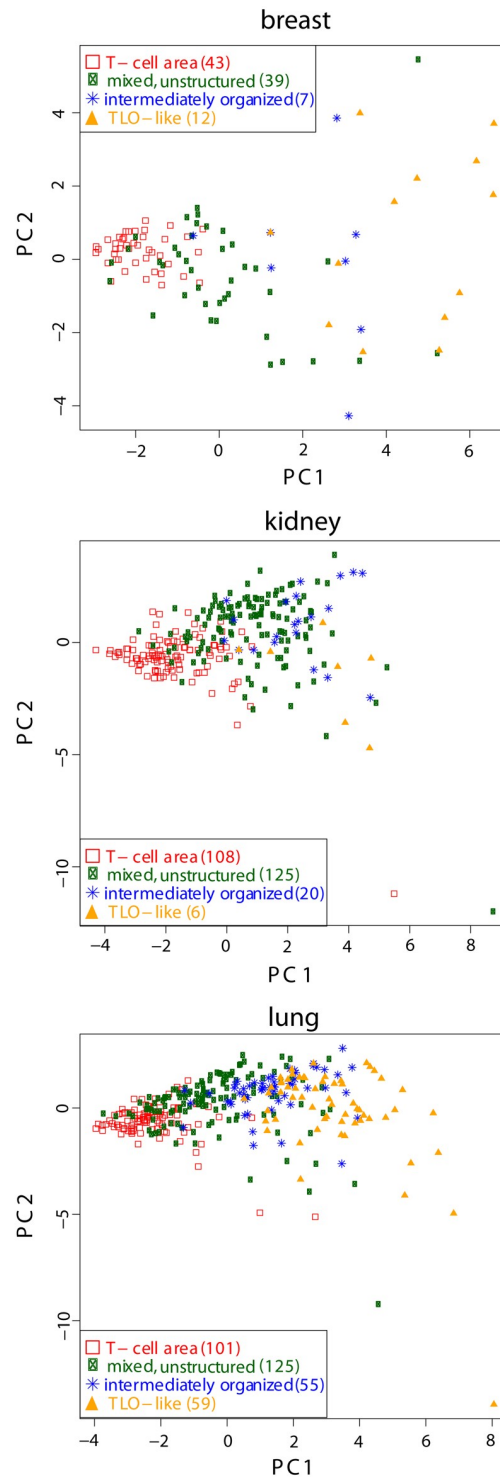
Fig 4. Principal component analysis between components with highest variance (PC1, PC2) for 700 samples in total (number of each class and tissue type listed in legend as brackets) that were visually classified by a pathologist. The number of nodes ($|V|$, $V_B/|V|$), edges ($|E|$, $|E_{\alpha_B}|/|E|$), TLO-like organization ($\kappa(2)$, $\kappa(5)$), homogeneity (H), clustering coefficient (C), average degree ($\langle K \rangle$), and the average Euclidean distance between all nodes are used as features, see Section Features of infiltrates.

https://doi.org/10.1371/journal.pcbi.1007385.g004

Table 1. Feasibility of support vector machine using 5-fold cross validation.

Measure	TPR	TNR	PPV	NPV	ACC	F_1
T-cell area	0.881	0.864	0.786	0.929	0.870	0.830
mixed, unstructured	0.709	0.742	0.660	0.785	0.729	0.682
intermediately organized	0.130	0.972	0.424	0.894	0.874	0.186
TLO-like	0.616	0.950	0.618	0.952	0.913	0.606
overall	0.584	0.882	0.622	0.890	0.846	0.576

measured in the test set of each fold and then averaged over all folds

Mann–Whitney U tests between a considered class and all remaining classes of the same tissue type confirmed significant differences for most but not all features.

The distributions of the relative number of edges between B-cells and $\kappa(2)$ are given in Fig 6. The TLO-like organization κ considers the B-cell cluster characteristics for the neighborhood of each edge between T- and B-cells. The number of B-cells and the organization $\kappa(5)$ showed a similar trend, see S5. The number of edges between B-cells increases with the number of B-cells starting with about 17 B-cells out of 100 cells in T-cell areas to 70 in intermediately organized structures. In the same way, the organization increases starting with $\kappa(5) = 0.001$ for T-cell areas to $\kappa(5) = 0.075$ for intermediately organized structures on average. This is in agreement with our expectation, as the probability to form B-cell clusters increases with the relative number of B-cells. The accumulation of B-cells is the core process of TLO genesis initiating B-cell clustering and CCR5–CXCL13 signaling [8], which may be the case in TLO-like and intermediately organized structures. Indeed, the differences between intermediately organized and TLO-like structures were less developed for both features (E_{α_B} and κ).

In comparison, the homogeneity and clustering coefficient showed less differences between the classes (S6). Obviously, areas consisting mainly of T-cells were quite homogeneous. In contrast, the homogeneity feature could not distinguish well between the other classes. In general, the clustering coefficient measures how dense a network is connected, suggesting that it could be useful in this context. Our observation that this feature was unsuitable to characterize an immune infiltrate can be explained by construction of Delaunay triangulation which avoids cliques. Further, the clustering coefficient neglects the information about the cell type. For example, there was no significant difference between mixed, unstructured and other areas in lung specimens.
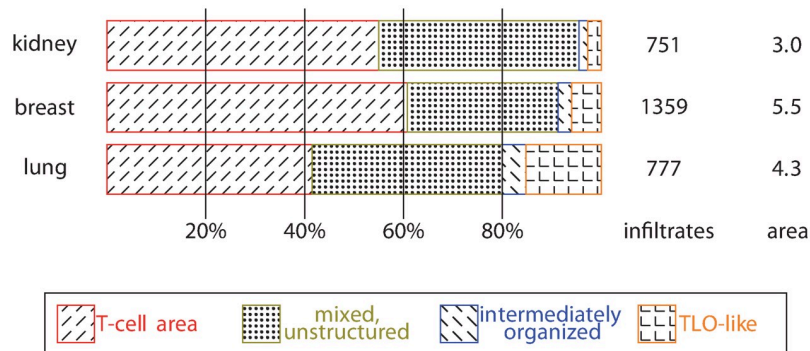
Fig 5. Composition of infiltrates: The absolute number of infiltrates and the estimated area considered to detect infiltrates (given in mm^2).
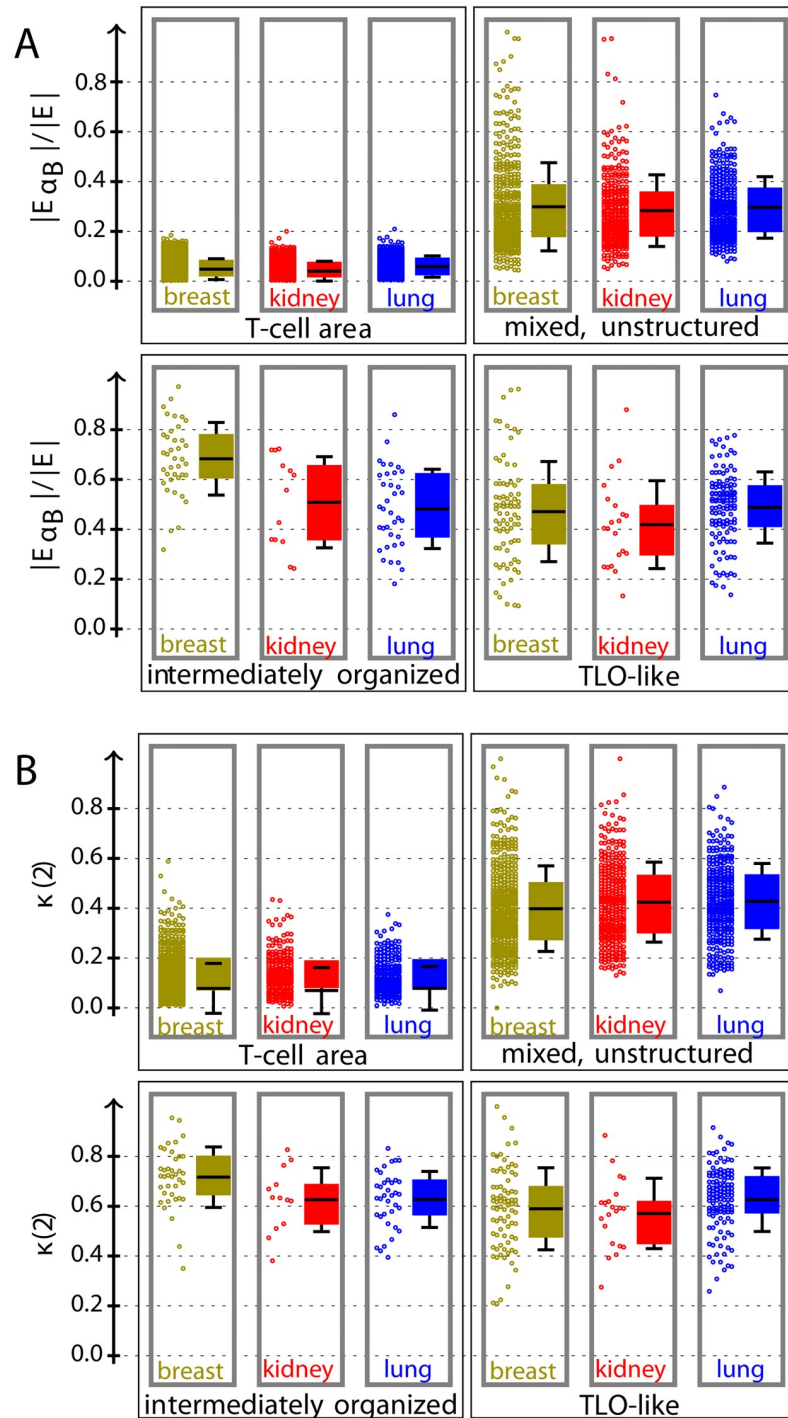
Fig 6. Distribution of relative number of edges between B-cells (A) and organization (B): Gray border represents significant difference between corresponding class and all remaining classes of same tissue type by Mann–Whitney *U* test using a significance level of 0.001. Each distribution is displayed as raw data (left) and a box plot (right) with mean (line), quartiles (color box), and standard deviation. The TLO-like organization κ measures the number of edges between T- and B-cells that are connected to a B-cell cluster, see Eq (1).

https://doi.org/10.1371/journal.pcbi.1007385.g006

The degree of the neighborhood graphs was Poisson distributed as expected by construction using Delaunay. Therefore, most nodes have a degree $\approx \langle K \rangle$, which was higher for more organized structures (4.5 for T-cell areas and 5.4 for TLO-like structures on average).

Overall, our results show that the organization feature κ differs clearly between the two classes with potentially higher degree of B-cell organization (intermediately organized and TLO-like) and other structures. It can be considered as "TLOness marker", with the caveat that it does not distinguish between early and mature forms.

Discussion

Analyzing immune infiltrates in WSIs has the clear advantage of considering the full information included in a tissue section in comparison to the common practice in histopathology to visually choose "representative" areas. A formalized way is less observer-dependent and is expected to increase concordance of immune cell scoring, for example between different clinical centers. Additionally, our approach avoids arbitrary cutting of infiltrates by tile borders. Nevertheless, some limitations should be considered: As immune infiltrates are composed of motile cells, they are not limited by predefined anatomical borders. Consequently, it is not possible to draw pixelwise exact outlines of infiltrates. In addition, even in better defined anatomical structures there can be a general interrater disagreement between pathologists in histological annotations [43]. Therefore, the manual annotation can only be an approximation to outline biologically relevant infiltrates. Even considering that these manually drawn outlines may not represent pixelwise exact ground truth, we nevertheless conclude that the approach is sufficient to detect immune infiltrates. We recommend that a transfer to different applications should carefully consider the importance of ground truth and aim for project-specific annotation strategies.

The recent advances in digital pathology show that robustness of cell detection and labeling is constantly increasing, even allowing large-scale studies with substantial variability of images. For graph-based approaches, the quality of the cell detection has a strong impact. To avoid false classifications, Raymond *et al.* in 1993 included only manually assigned cells into their cellular composition in germinal centers of lymph nodes [41]. As the quality of single cell detection is nowadays not a limiting factor, we used a robust cell detection [23] and ensured quality by a manual exclusion of artefact regions before hand.

We considered four categories of infiltrates that have been pragmatically chosen based on obvious differences in their histological appearance. They correspond to a hypothetical biological role in alignment with common practice of descriptive immune cell evaluation. It should be noted, that a specific staining defining early or even more mature TLO-like structures in histological images is currently not available (presence of HEVs, dendritic antigen-presenting cells, and lymph vessels can only detect late stages of TLOs; CXCL13/CCR5 signaling molecules are technically limited). Therefore, the functional biological meaning of the four categories, particularly for the intermediately organized structures, remains to be defined and may even vary between different infiltrates. One interpretation of the observed structures could be the underlying dynamics of forming and potentially also resolving TLO-like structures. Alternatively, these infiltrates could represent different immunological functions. In general, a certain degree of overlap between the four categories is expected because we are probably observing different temporal stages of variable cell interaction patterns in different microenvironments rather than a sharply defined phase transition of lymphoid cell differentiation like in primary lymphoid organs.

Our approach resulted in a proposed "TLOness marker" (κ) that specifically measures how many links between T- and B-cells are connected to a B-cell cluster; independent of overall

size and density of the infiltrate. κ was developed based on the biological description of TLOs and therefore characterizes the requirement of B-cell accumulations in the center and considers a surrounding T-cell zone. In addition, some dispersed T-cells in the middle of the B-cell predominant center occur frequently in TLOs and have putative supporting functions in B-cell maturation. They can be rate limiting for B-cell maturation in germinal centers [44–46]. The main advantage of κ compared to evaluation of B-cell quantities or classical clustering description is that this feature does not strongly penalize the single T-cells inside B-cell clusters.

Recent progress on deep learning approaches and increasing broader use of these methods holds promise to find ROIs in general more efficiently than based on conventional approaches [26]. In this specific application, our method could support such methods and our neighborhood graphs can be analyzed by graph neural networks [47] in the future.

Application to biomedical context

In breast cancer, TLO detection is recently becoming even more relevant as there are correlations between TLOs and other, therapeutically relevant infiltration patterns [48, 49].

In transplantation medicine, further evidence suggested that such "subclinical" patterns of inflammation are related to long-term decline of transplant function and associated with interstitial fibrosis and tubular atrophy (IF/TA) [50, 51]. Previous studies showed clear limitations for reliable TLO quantification in kidney biopsies [52] due to their tiny size and the large heterogeneity of B-cell areas. Since there is a growing interest in immune infiltrates occurring after transplantation in patients without evidence for rejection and therefore only small biopsies are available, our approach will nevertheless support more robust description of infiltrate with the risk of sampling error, missing larger or more dense structures elsewhere.

In lung tissue, TLOs are considered as a hallmark of autoimmune disease such as rheumatoid arthritis and Sjögren syndrome but can also be found in chronic infection and various other lung pathologies [53–55]. The exact role of iBALT is a field of active research and their exact composition may be pivotal in defining their function [56, 57]. Our approach plays an important role in translational research to better define, characterize, and by this understand the patterns of chronic inflammation in the lung.

Herein, we focused on these three relevant applications. However, the graph-based assessment could be applied to TLOs in other tissue types or diseases and beyond that could easily be adjusted to any cell accumulation with a larger number of different cell types.

Conclusion

Our workflow allows to describe relevant immune cell patterns by analyzing the spatial cellular organization. It could improve throughput, robustness, and objectivity of immune cell evaluation. Our results demonstrate that the immune infiltrates in the considered tissue specimens differ in their spatial composition and degree of organization. In order to distinguish between random immune infiltrates and actual TLOs related to the underlying disease, it is necessary to study possible correlations between graph-based features and clinical outcome.

Materials and methods

Image analysis

A nucleus detection algorithm (Tissue Phenomics framework, Definiens AG, Munich, Germany) using an auto-adaptive slide-specific visual context model based on random forests [23] was applied to deliver single cell segmentations and cell type classifications for CD3 (T-cells)

and CD20 (B-cells); for more details, see S8. For further analysis, we considered each full WSI as a single ROI.

Materials

We considered three data sets to validate our method containing (1) 23 breast cancer specimens including tumor area and distant normal glandular epithelium, (2) 160 renal transplantation biopsies from 54 patients, and (3) tissue from explanted lungs of 10 patients with cystic fibrosis and bronchiectasis with chronic and acute inflammation. The formalin-fixed, paraffin-embedded tissue was cut into 3 μm thick sections and immunohistochemically stained for CD3/CD20 (CD3 [EP449E]: abcam ab52959 colored in brown with 3'–3'–Diaminobenzidine, CD20: DAKO M0755 colored in red with Neufuchsin) on a Ventana Benchmark Ultra instrument. WSIs were scanned using Aperio AT2 (Leica Microsystems, Wetzlar, Germany) at a resolution of 0.253 μm/pixel (40x magnification). Staining artifacts or folds of tissue were excluded from analysis.

Ethics statement. Use of anonymized archival surplus tissue (formalin-fixed, paraffin-embedded tissue; FFPET) under a waiver for individual informed consent was approved by the institutional review board (ethics committee of Hannover Medical School); Approval Number 2063–2013.

Supporting information

S1 Fig. Visualization of compartments. All edges between cells of different phenotypes were removed from the neighborhood graphs (see Fig 2C). A: Then, the graphs were split into connected components of a single phenotype. B: Visualization by concave hull. Single cells/cell pairs do not span a compartment.
(PDF)

S2 Fig. Choice of parameter *t* for acceptance of Delaunay triangle according to circumcircle radius: Shown are two independent immune infiltrates (left and right columns) with varying *t* (rows). Blue bullets show the centre of the T- and B-cell nuclei, red lines represent α edges (between nuclei of same cell type) and green lines (between nuclei of different cell type) represent γ edges.
(PDF)

S3 Fig. Thresholding. Shown are Pearson's correlation coefficients ρ between ground truth (GT) and features (illustrated in blue) (**A**), selected parameters and features to distinguish between four considered classes (**B**), and corresponding F_1 scores (**C**). The number of nodes ($|V|$, $|V_B|$), edges ($|E|$, $|E_{\alpha_B}|/|E|$), TLO-like organization (κ), homogeneity (H), clustering coefficient (C), average degree ($\langle K \rangle$), and the average Euclidean distance between all nodes ($d_{avg}(v_l, v_m)$) are considered as features, see Section Features of infiltrates for definitions.
(PDF)

S4 Fig. Principal component analysis. Left: variances. Right: PC1 against PC3, where red rectangles represent T-cell areas, green dots mixed, unstructured areas, blue snowflakes intermediately organized structures, and yellow triangles TLO-like structures.
(PDF)

S5 Fig. Distribution. Gray border represents significant difference between corresponding class and all remaining classes of same tissue type by Mann–Whitney *U* test using a significance level of 0.001. A: relative number of B-cells. B: organization κ.
(PDF)

S6 Fig. Distribution. Gray border represents significant difference between corresponding class and all remaining classes of same tissue type by Mann–Whitney *U* test using a significance level of 0.001. A: homogeneity *H*. B: clustering coefficient *C*.
(PDF)

S1 Table. Feasibility of support vector machine using 5-fold cross validation, measured in the test set of each fold and then averaged over all folds.
(PDF)

S1 Appendix. Image analysis. Further information.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Nadine S. Schaadt, Michael Meyer-Hermann, Friedrich Feuerhake.

Formal analysis: Nadine S. Schaadt.

Funding acquisition: Nadine S. Schaadt, Ralf Schönmeyer, Michael Meyer-Hermann, Friedrich Feuerhake.

Methodology: Nadine S. Schaadt, Germain Forestier.

Resources: Ralf Schönmeyer, Nicolas Brieu, Peter Braubach, Katharina Nekolla, Friedrich Feuerhake.

Visualization: Nadine S. Schaadt.

Writing – original draft: Nadine S. Schaadt.

Writing – review & editing: Nadine S. Schaadt, Ralf Schönmeyer, Germain Forestier, Peter Braubach, Katharina Nekolla, Michael Meyer-Hermann, Friedrich Feuerhake.

References

1. Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G, et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. Annals Oncol. 2014; 26:259–271. https://doi.org/10.1093/annonc/mdu450

2. Beyer T, Meyer-Hermann M. Mechanisms of organogenesis of primary lymphoid follicles. Int Immunol. 2008; 20:615–623. https://doi.org/10.1093/intimm/dxn020 PMID: 18334502

3. Pitzalis C, Jones G, Bombardieri M, Jones S. Ectopic lymphoid-like structures in infection, cancer and autoimmunity. Nat Rev Immunol. 2014; 14:447–462. https://doi.org/10.1038/nri3700 PMID: 24948366

4. Krenn V, Schalhorn N, Greiner A, Molitoris R, König A, Gohlke F, et al. Immunohistochemical analysis of proliferating and antigen-presenting cells in rheumatoid synovial tissue. Rheumatology Int. 1996; 15:239–247. https://doi.org/10.1007/BF00290377

5. Beyer T, Meyer-Hermann M. Cell transmembrane receptors determine tissue pattern stability. Phys Rev Lett. 2008; 101:148102. https://doi.org/10.1103/PhysRevLett.101.148102 PMID: 18851578

6. MacLennan IC. Germinal centers. Annu Rev Immunol. 1994; 12:117–139. https://doi.org/10.1146/annurev.iy.12.040194.001001 PMID: 8011279

7. Victora GD, Nussenzweig MC. Germinal centers. Annu Rev Immunol. 2012; 30:429–457. https://doi.org/10.1146/annurev-immunol-020711-075032 PMID: 22224772

8. Stranford S, Ruddle N. Follicular dendritic cells, conduits, lymphatic vessels, and high endothelial venules in tertiary lymphoid organs: parallels with lymph node stroma. Front Immunol. 2012; 3. https://doi.org/10.3389/fimmu.2012.00350 PMID: 23230435

9. Martinet L, Garrido I, Filleron T, Le Guellec S, Bellard E, Fournie JJ, et al. Human solid tumors contain high endothelial venules: association with T-and B-lymphocyte infiltration and favorable prognosis in breast cancer. Cancer res. 2011; 71:5678–5687. https://doi.org/10.1158/0008-5472.CAN-11-0431 PMID: 21846823

10. Lee HJ, Park IA, Song IH, Shin SJ, Kim JY, Yu JH, et al. Tertiary lymphoid structures: prognostic significance and relationship with tumour-infiltrating lymphocytes in triple-negative breast cancer. J Clin Pathol. 2016; p. 422–430. https://doi.org/10.1136/jclinpath-2015-203089 PMID: 26475777

11. Dieu-Nosjean MC, Giraldo NA, Kaplon H, Germain C, Fridman WH, Sautès-Fridman C. Tertiary lymphoid structures, drivers of the anti-tumor responses in human cancers. Immunol Rev. 2016; 271:260–275. https://doi.org/10.1111/imr.12405 PMID: 27088920

12. Pimenta E, Barnes B. Role of tertiary lymphoid structures (TLS) in anti-tumor immunity: potential tumor-induced cytokines/chemokines that regulate TLS formation in epithelial-derived cancers. Cancers. 2014; 6:969–997. https://doi.org/10.3390/cancers6020969 PMID: 24762633

13. Xu X, Han Y, Wang Q, Cai M, Qian Y, Wang X, et al. Characterisation of Tertiary Lymphoid Organs in Explanted Rejected Donor Kidneys. Immunol Invest. 2016; 45:38–51. https://doi.org/10.3109/08820139.2015.1085394 PMID: 26709668

14. Loupy A, Haas M, Solez K, Racusen L, Glotz D, Seron D, et al. The Banff 2015 kidney meeting report: current challenges in rejection classification and prospects for adopting molecular pathology. Am J Transplantation. 2017; 17:28–41. https://doi.org/10.1111/ajt.14107

15. Mengel M, Gwinner W, Schwarz A, Bajeski R, Franz I, Bröcker V, et al. Infiltrates in protocol biopsies from renal allografts. Am J Transplantation. 2007; 7:356–365. https://doi.org/10.1111/j.1600-6143.2006.01635.x

16. Mengel M, Reeve J, Bunnag S, Einecke G, Jhangri G, Sis B, et al. Scoring total inflammation is superior to the current Banff inflammation score in predicting outcome and the degree of molecular disturbance in renal allografts. Am J Transplantation. 2009; 9:1859–1867. https://doi.org/10.1111/j.1600-6143.2009.02727.x

17. Pabst R, Gehrke I. Is the bronchus-associated lymphoid tissue (BALT) an integral structure of the lung in normal mammals, including humans? Am J Respir Cell Mol Biol. 1990; 3:131–135. https://doi.org/10.1165/ajrcmb/3.2.131 PMID: 2378747

18. Pabst R. Is BALT a major component of the human lung immune system? Immunol Today. 1992; 13:119–122. https://doi.org/10.1016/0167-5699(92)90106-H PMID: 1580992

19. Neyt K, Perros F, GeurtsvanKessel C, Hammad H, Lambrecht B. Tertiary lymphoid organs in infection and autoimmunity. Trends Immunol. 2012; 33:297–305. https://doi.org/10.1016/j.it.2012.04.006 PMID: 22622061

20. Apou G, Schaadt NS, Naegel B, Forestier G, Schönmeyer R, Feuerhake F, et al. Detection of lobular structures in normal breast tissue. Computers Biol Med. 2016; 74:91–102. https://doi.org/10.1016/j.compbiomed.2016.05.004

21. Kato T, Relator R, Ngouv H, Hirohashi Y, Takaki O, Kakimoto T, et al. Segmental HOG: new descriptor for glomerulus detection in kidney microscopy image. BMC Bioinformatics. 2015; 16:316. https://doi.org/10.1186/s12859-015-0739-1 PMID: 26423821

22. Grote A, Abbas M, Linder N, Kreipe H, Lundin J, Feuerhake F. Exploring the spatial dimension of estrogen and progesterone signaling: detection of nuclear labeling in lobular epithelial cells in normal mammary glands adjacent to breast cancer. Diagnostic Pathol. 2014; 9:S11. https://doi.org/10.1186/1746-1596-9-S1-S11

23. Brieu N, Pauly O, Zimmermann J, Binnig G, Schmidt G. Slide specific models for segmentation of differently stained digital histopathology whole slide images. Proc SPIE. 2016; 9784:978410–978417. https://doi.org/10.1117/12.2208620

24. Rogers MA, Aikawa E. Cardiovascular calcification: artificial intelligence and big data accelerate mechanistic discovery. Nature Rev Cardiology. 2018; p. 1.

25. Lan K, Wang Dt, Fong S, Liu Ls, Wong KK, Dey N. A survey of data mining and deep learning in bioinformatics. J Medical Systems. 2018; 42:139. https://doi.org/10.1007/s10916-018-1003-9

26. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. Ann Rev Biomed Eng. 2017; 19:221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

27. Dey N, Ashour A. Classification and clustering in biomedical signal processing. IGI global Hershey; 2016.

28. Suzuki K. Pixel-based machine learning in medical imaging. J Biomed Imaging. 2012; 2012:1.

29. Aguilar-Hidalgo D, Zurita AC, Fernández MCL. Complex networks evolutionary dynamics using genetic algorithms. Internat J Bifurcation Chaos. 2012; 22:1250156. https://doi.org/10.1142/S0218127412501568

30. Hemanth DJ, Anitha J. Modified Genetic Algorithm approaches for classification of abnormal Magnetic Resonance Brain tumour images. Appl Soft Computing. 2019; 75:21–28. https://doi.org/10.1016/j.asoc.2018.10.054

31. Li S, Wu H, Wan D, Zhu J. An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine. Knowledge-Based Systems. 2011; 24:40–48. https://doi.org/10.1016/j.knosys.2010.07.003

32. Sun Y, Xue B, Zhang M, Yen GG. Evolving deep convolutional neural networks for image classification. IEEE Transactions on Evolutionary Computation. 2019. https://doi.org/10.1109/TEVC.2019.2916183

33. Wieczorek G, Bigaud M, Pfister S, Ceci M, McMichael K, Afatsawo C, et al. Blockade of CD40–CD154 pathway interactions suppresses ectopic lymphoid structures and inhibits pathology in the NOD/ShiLtJ mouse model of Sjoegrens syndrome. Annals Rheumatic Diseases. 2019; 78:974–978. https://doi.org/10.1136/annrheumdis-2018-213929

34. Schiffer L, Wiehler F, Bräsen JH, Gwinner W, Greite R, Kreimann K, et al. Chemokine CXCL13 as a New Systemic Biomarker for B-Cell Involvement in Acute T Cell-Mediated Kidney Allograft Rejection. Internat J Molecular Sciences. 2019; 20:2552. https://doi.org/10.3390/ijms20102552

35. Cheikh BB, Elie N, Plancoulaine B, Bor-Angelier C, Racoceanu D. Spatial interaction analysis with graph based mathematical morphology for histopathology. Biomed Imaging, IEEE 14th Internat Symposium. 2017; p. 813–817.

36. Schäfer H, Schäfer T, Ackermann J, Dichter N, Döring C, Hartmann S, et al. CD30 cell graphs of Hodgkin lymphoma are not scale-free an image analysis approach. Bioinformatics. 2015; p. btv542.

37. Demir C, Gultekin SH, Yener B. Augmented cell-graphs for automated cancer diagnosis. Bioinformatics. 2005; 21:ii7–ii12. https://doi.org/10.1093/bioinformatics/bti1100 PMID: 16204128

38. Gunduz C, Yener B, Gultekin SH. The cell graphs of cancer. Bioinformatics. 2004; 20:i145–i151. https://doi.org/10.1093/bioinformatics/bth933 PMID: 15262793

39. Loménie N, Racoceanu D. Point set morphological filtering and semantic spatial configuration modeling: Application to microscopic image and bio-structure analysis. Pattern Recognition. 2012; 45:2894–2911. https://doi.org/10.1016/j.patcog.2012.01.021

40. Basavanhally A, Agner S, Alexe G, Bhanot G, Ganesan S, Madabhushi A; Citeseer. Manifold learning with graph-based features for identifying extent of lymphocytic infiltration from high grade, HER2+ breast cancer histology. Image Anal Appl Biol. 2008.

41. Raymond E, Raphael M, Grimaud M, Vincent L, Binet JL, Meyer F. Germinal center analysis with the tools of mathematical morphology on graphs. Cytometry. 1993; 14:848–861. https://doi.org/10.1002/cyto.990140803 PMID: 8287731

42. Barabasi A, Oltvai Z. Network biology: understanding the cell's functional organization. Nat Rev Genetics. 2004; 5:101–113. https://doi.org/10.1038/nrg1272 PMID: 14735121

43. Grote A, Schaadt NS, Forestier G, Wemmert C, Feuerhake F. Crowdsourcing of Histological Image Labeling and Object Delineation by Medical Students. IEEE TMI. 2018.

44. Meyer-Hermann ME, Maini PK, Iber D. An analysis of B cell selection mechanisms in germinal centers. Math Med Biol. 2006; 23:255–277. https://doi.org/10.1093/imammb/dql012 PMID: 16707510

45. Allen CD, Okada T, Cyster JG. Germinal-center organization and cellular dynamics. Immunity. 2007; 27:190–202. https://doi.org/10.1016/j.immuni.2007.07.009 PMID: 17723214

46. Victora GD, Schwickert TA, Fooksman DR, Kamphorst AO, Meyer-Hermann M, Dustin ML, et al. Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter. Cell. 2010; 143:592–605. https://doi.org/10.1016/j.cell.2010.10.032 PMID: 21074050

47. Wu Z, Pan S, Chen F, Long G, Zhang C, Yu PS. A comprehensive survey on graph neural networks. arXiv preprint arXiv:190100596. 2019.

48. Lee M, Heo SH, Song IH, Rajayi H, Park HS, Park IA, et al. Presence of tertiary lymphoid structures determines the level of tumor-infiltrating lymphocytes in primary breast cancer and metastasis. Modern Pathol. 2019; 32:70–80. https://doi.org/10.1038/s41379-018-0113-8

49. Boivin G, Kalambaden P, Faget J, Rusakiewicz S, Montay-Gruel P, Meylan E, et al. Cellular composition and contribution of tertiary lymphoid structures to tumor immune infiltration and modulation by radiation therapy. Frontiers Oncol. 2018; 8. https://doi.org/10.3389/fonc.2018.00256

50. Heilman RL, Devarapalli Y, Chakkera H, Mekeel K, Moss A, Mulligan D, et al. Impact of subclinical inflammation on the development of interstitial fibrosis and tubular atrophy in kidney transplant recipients. Am J Transplantation. 2010; 10:563–570. https://doi.org/10.1111/j.1600-6143.2009.02966.x

51. Park WD, Griffin MD, Cornell LD, Cosio FG, Stegall MD. Fibrosis with inflammation at one year predicts transplant functional decline. J Am Soc Nephrol. 2010; 21:1987–1997. https://doi.org/10.1681/ASN.2010010049 PMID: 20813870

52. Thaunat O. Pathophysiologic significance of B-cell clusters in chronically rejected grafts. Transplantation. 2011; 92:121–126. https://doi.org/10.1097/TP.0b013e31821f74fe PMID: 21555973

53. Frija-Masson J, Martin C, Regard L, Lothe MN, Touqui L, Durand A, et al. Bacteria-driven peribronchial lymphoid neogenesis in bronchiectasis and cystic fibrosis. Europ Respir J. 2017; 49. https://doi.org/10.1183/13993003.01873-2016

54. Rangel-Moreno J, Hartson L, Navarro C, Gaxiola M, Selman M, Randall TD. Inducible bronchus-associated lymphoid tissue (iBALT) in patients with pulmonary complications of rheumatoid arthritis. J Clin Investig. 2006; 116:3183–3194. https://doi.org/10.1172/JCI28756 PMID: 17143328

55. Schneider F, Gruden J, Tazelaar HD, Leslie KO. Pleuropulmonary pathology in patients with rheumatic disease. Arch Pathol Lab Med. 2012; 136:1242–1252. https://doi.org/10.5858/arpa.2012-0248-SA PMID: 23020730

56. Pabst R, Tschernig T. Bronchus-associated lymphoid tissue: an entry site for antigens for successful mucosal vaccinations? Am J Respir Cell Mol Biol. 2010; 43:137–141. https://doi.org/10.1165/rcmb.2010-0152RT PMID: 20508066

57. Gohy ST, Ladjemi MZ, Pilette C. Lung lymphoid neogenesis in cystic fibrosis: a model of adaptive responses to bacteria? Europ Respir J. 2017; 49. https://doi.org/10.1183/13993003.00380-2017