Spatial immunophenotyping using multiplexed imaging of immune follicles in secondary lymphoid tissues

Mayar Allam [ID][a,1], Thomas Hu[a,b,1], Zhou Fang[a], Michelle Pi [ID][c], Ankur Singh [ID][a,d,e,*] and Ahmet F. Coskun [ID][a,e,f,g,*]

[a]Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA 30332, USA
[b]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA
[c]School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA
[d]Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30318, USA
[e]Parker H. Petit Institute for Bioengineering and Bioscience, Georgia Institute of Technology, Atlanta, GA 30332, USA
[f]Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA
[g]Interdisciplinary Bioengineering Graduate Program, Georgia Institute of Technology, Atlanta, GA 30332, USA
*To whom correspondence should be addressed: Email: ahmet.coskun@bme.gatech.edu; ankur.singh@gatech.edu
[1]Equal authorship.
Edited By Richard Stanton

Abstract

Secondary lymphoid organs (SLOs), including tonsils (TS), lymph nodes (LN), and Peyer's Patches, exhibit complementary immune functions. However, little is known about the spatial organization of immune cells and extracellular matrix (ECM) in the SLOs. Traditional imaging is limited to a few markers, confining our understanding of the differences between the SLOs. Herein, imaging mass cytometry addressed this gap by simultaneously profiling 25-plex proteins in SLO tissues at subcellular resolution. The antibody panel targeted immune, stromal, chemokine, epigenetic, and functional markers. For robust cell identification, a computational workflow *SpatialVizPheno* was developed to spatially phenotype 999,970 cells using two approaches, including manual gating and semi-supervised gating, iterative clustering, and annotation. LN exhibited the highest density of B cells while the intestinal tissues contained the highest proportion of regulatory and follicular helper T cells. *SpatialVizPheno* identified the most prevalent interaction between follicular dendritic cells and stromal cells (SCs), plasmablasts/plasma cells, and the SCs across the lymphoid tissues. Collagen-enriched regions were associated with the spatial orientation of B cell follicles in both TS and LN tissues, but not in intestinal lymphoid tissues. Such spatial differences of immunophenotypes and ECM in different SLO tissues can be used to quantify the relationship between cellular organization and ultimate immune responses.

Significance Statement

Despite the significance of the secondary lymphoid organs (SLOs) and their diverse immune regulatory systems, the single-cell differences between the SLOs remain understudied. Further, the positional interactions among diverse cellular phenotypes in the different SLOs remain poorly understood. Herein, imaging mass cytometry with a panel of 25-plex markers revealed the differences in immune cell phenotype density, tissue neighborhoods, follicular organization, and stromal networks in tonsil, lymph node, and intestinal Peyer's Patches tissues.

Introduction

The secondary lymphoid organs (SLOs) are characterized by their rich immune microenvironments as they contain functionally specialized phenotypes of B and T lymphocytes and form the innate and adaptive immune response. They are distributed across the body and they function to expose their lymphocytes to different antigens and develop antigen-specific immune responses (1). The SLOs include the lymph nodes (LN), spleen (SP), tonsil (TS), and intestinal Peyer's patches (PP). Each of these organs exposes their lymphocytes to antigens at corresponding sites including the lymphatic fluid, the blood, the mouth or nose, and the digestive system respectively. B cells are a critical component of the secondary lymphoid tissues as they are responsible for the antibody-dependent functions of antigen presentation, regulation of T cell differentiation and survival, and secreting regulatory and proinflammatory cytokines (1, 2). Further, different SLOs also exhibit unique immune and stromal cell (SC) organization. For example, intestinal PP were shown to have distinct stromal infrastructures through the populations of fibroblastic reticular cells than LV, and hence they have unique immune cell interactions (3). In addition, TS SC composition was also found to be distinct from that of the LN as marked by the absence of marginal

reticular cells as part of the fibroblastic reticular cells network used for antigen sampling. This difference in the capture of the antigen materials also results in different immune regulatory trajectories (4). Within the SLOs, germinal centers (GCs) are the specialized microstructures for the B cell activation, wherein B cells proliferate and diversify their immunoglobulins in the dark zone (DZ) through somatic hypermutation, followed by affinity-based selection in the light zone (LZ) through their interactions with follicular dendritic cells (FDCs) and follicular T helper cells (5–8). However, recent work further suggests that events leading to GC responses and their structural organization are different in TS, LN, and PPs with an impact on clonal selection and GC kinetics (9). Given the differences in B cell maturation processes within the SLOs, it is essential to study the spatial organization of different phenotypes of B cells within these tissues and their proximity to other innate, adaptive immune cells and SCs, along with the ECM that defines a tissue microarchitecture (Fig. 1A and B). This information can expand our understanding of the immune microenvironment in the lymphoid tissues and can help infer cellular interactions and regulatory function. In turn, it can also help in understanding disease-relevant pathways and in designing more treatments.

Multiplexed tissue imaging is critical to capture the diverse immune phenotypes and decipher spatial distribution in SLOs (10). Fluorescence-based platforms including cyclic immunofluorescence (CycIF), multiplexed immunohistochemistry, and codetection by indexing (CODEX) can yield multiplex images of tissues through iterative cycles of antibody labeling, imaging, and antibody removal/bleaching (11). Experimentally, these techniques are labor-intensive and can take up to 8–12 h per cycle. They also require changes in temperatures of the tissue samples between the antibody labeling and removal which can change the tissue architecture between cycles or cause tissue loss (12, 13). Further, they often suffer from signal residuals between cycles that need to be corrected computationally in the data preprocessing. Formalin-fixed paraffin-embedded tissue samples often exhibit variant levels of autofluorescence which can further complicate downstream data processing and make it prone to artifacts (13). Cytometry by the time of flight (CyTOF) techniques, including imaging mass cytometry (IMC), pushes this limit by offering a high-throughput platform that combines the principles of flow cytometry and mass spectrometry to profile thousands of single cells using up to 40 markers. The masses of these metal tags can be used for the readout instead of relying on fluorescence tags, resulting in a highly sensitive measurement with minimal spillover and background noise. All the markers are detected in one round instead of relying on iterative cycles of labeling and marker removal. Further, the metal tags are rare to find in biological samples yielding more quantitatively accurate measurements (14, 15). Therefore, IMC was the best fit for this high-parameter spatial study due to the highly sensitive signal, the minimal spillover between mass channels, and the low batch effect. IMC was previously used to profile the spatial distribution of the marginal zone B cells and the classical memory B cell population in lymphoid tissues (16). The classical memory B cells were found to be localized close to the peripherals of the lymphoid tissue for potential pathogen exposures while the marginal zone B cells were positioned between the GCs and the epithelium around them (16). It was notable that B cell populations were present in spatially distinct microanatomical niches from one another (16). IMC was also previously utilized to compare the immune microenvironment in TS tissues from healthy donors and diseased donors with chronic tonsillitis (17). A significant correlation was identified between CD68$^+$ monocytes and the expression of granzyme B in diseased TS tissues (17). Further, CD3$^+$CD4$^+$ T cells were more predominant than CD3$^+$CD8$^+$ T cells in the diseased TSs, highlighting the central rule of the helper T cells in mediating the GC response in the case of infections (17). High-parameter imaging technologies investigated the spatial context of a few immune subsets; however, the number of annotated B cell and T cell phenotypes was limited, and understanding of the ECM organization in lymphoid tissues is still lacking.

Herein, we performed a comparative spatially resolved single-cell immunophenotyping analysis in multiple lymphoid tissue types using the IMC spatial proteomics approach and a companion computational tool: *SpatialVizPheno*. To dissect the GC organization, several immune (CD3, CD4, CD8, CD20, CD38, BCL6, ICOS, CD11b, CD11c, CD21, PD-1, CD27, CD138, CD86, CD83), stromal (Vimentin, Collagen I), chemokines (CXCR4, CXCR5), epigenetic (EZH2, H3K27me3), and functional markers (FoxP3, Ki67, C-Myc) markers were profiled in SLOs (Fig. 1A). This marker panel was used to understand the spatial difference of immune cells phenotypes and the stromal network in the SLOs. The annotated phenotypes included diverse B cells across their differentiation into the light zone (LZ B) and the dark zone (DZ B) phenotype found in GCs, where B cell maturation takes place in response to infection or vaccination (Fig. 1B). Further, several phenotypes of T cells were also annotated including helper T cells (Th), follicular helper T cells (Tfh), exhausted T cells (Tex), regulatory T cells (Treg), and cytotoxic T cells (Tc). It also segmented FDCs, SCs, and plasmablasts/plasma cells (PBs/PCs). To define the spatial organization of adjacent cells, the positional interactions and the crosstalk between cellular phenotypes were defined using neighborhood analysis in distinct lymphoid tissues. The spatial and organization of the ECM proteins (collagen and vimentin) were then associated with a spatially unique orientation of follicle polarization and ECM-rich image features in TS, LN, and intestinal PPs tissues.

Results

SpatialVizPheno enables spatially resolved, single-cell, multiplexed protein profiling of B and T cell phenotypes in lymphoid tissues using the IMC dataset

Herein, we designed an antibody panel of 25 markers (Table S1) spanning immune surface markers, cytokine markers, epigenetic regulators, and extracellular matrix (ECM) proteins. We limited ourselves to a panel of 25 that worked reliably and reproducibly across all three SLOs, although with other tissues it may be possible to go with a higher number of markers. The multiplexed antibody panel was designed using commercially available antibodies preconjugated to appropriate metal tags ($n = 17$) and custom-conjugated antibodies to unique metal tags ($n = 8$) following established protocols (18). We used this panel to profile three different SLO including TS, LN tissues, and intestinal PP (Figs. 1C and S1–S6 and Table S2).

Concurrently, we also developed *SpatialVizPheno* which is a highly modular computational workflow to spatially visualize phenotypes in the immune-rich microenvironments in the secondary lymphoid tissues using the high-parameter dataset generated using IMC. *SpatialVizPheno* spatially annotated the phenotypically similar immune phenotypes using two methods. Firstly, *SpatialVizPheno* leveraged the widely used FlowJo software
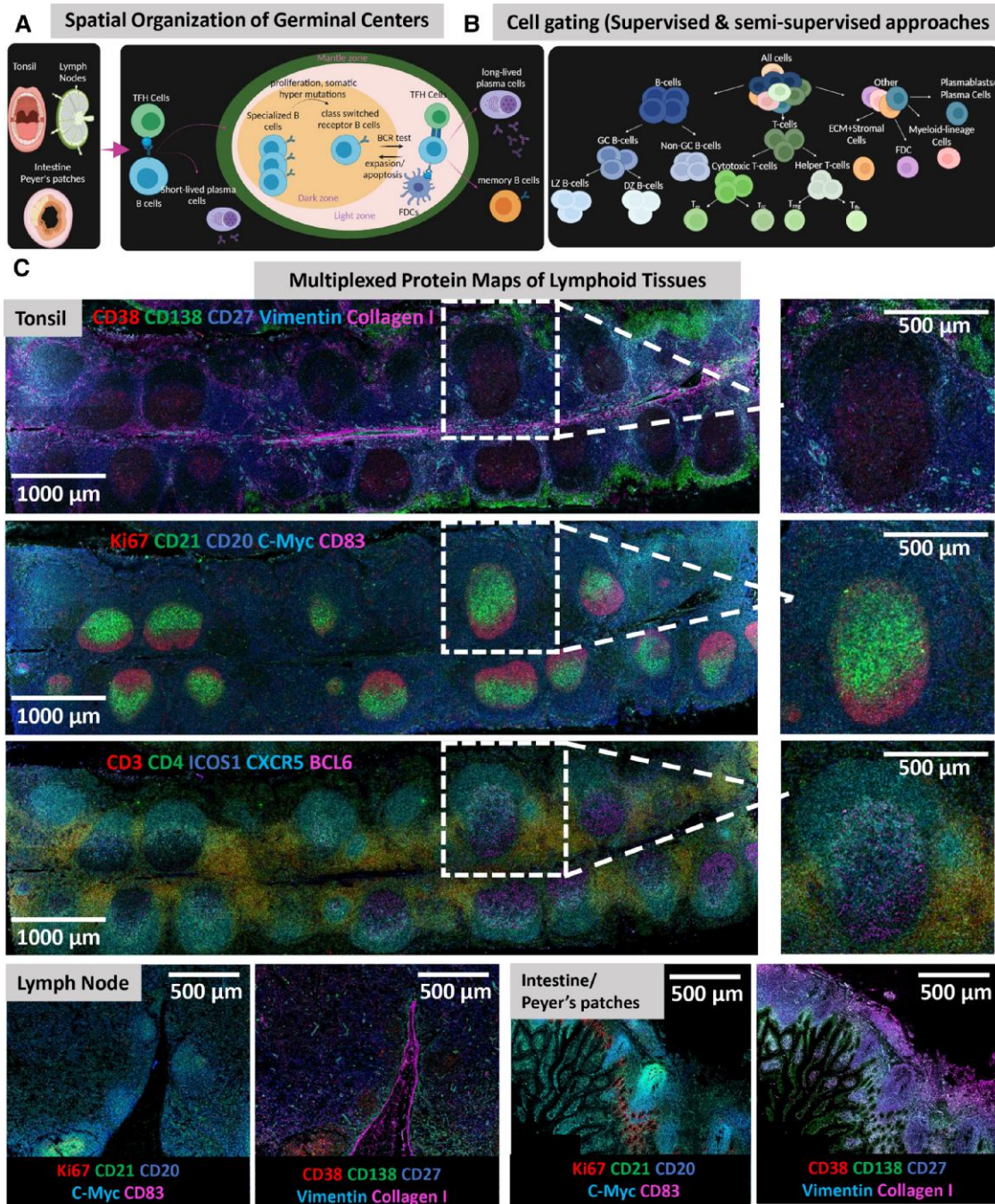
Fig. 1. Imaging mass cytometry enables deciphering of immune microenvironment in lymphoid tissues. A) Several human lymphoid tissues were used to understand the organization of B cell follicles and germinal centers including TS, LN, and intestinal PP. IMC can provide spatial information on the unique cellular phenotypes in distinct tissue regions including dark zone, light zone, and mantle zone. Created with BioRender.com. B) The cell gating approach consisted of classifying all cells in the datasets into either B cells, T cells, or others. Under the B cells tree, cells were further divided into the germinal center or nongerminal center B cells. The germinal center B cells were further split into light-zone B cells or dark-zone B cells. T cells were divided into cytotoxic or helper T cells. Cytotoxic T cells were further split into exhausted T cells or follicular T cells. Similarly, helper T cells were split into T-regulatory cells or follicular T cells. Finally, cells under the other category were split into stromal cells, follicular dendritic cells, myeloid-linage cells, or plasmablasts/plasma cells. Created with BioRender.com. C) A panel of 25 markers was used to decipher the follicle composition in three different lymphoid tissue types including TSs, LN, and intestinal PP.

to manually gate phenotypes on imaging datasets using sequential and Boolean gating methods. Secondly, *SpatialVizPheno* also used an iterative process of unsupervised clustering and manual annotation to separate or merge clusters resulting in clusters with matched biological phenotypes. *SpatialVizPheno* then used both methods simultaneously to compare the neighborhood among the cellular phenotypes in the different tissue types. It also analyzed the phenotypes infiltrating within or around stromal regions to shed more light on the importance of the stromal network in secondary lymphoid tissue immune functions. It also analyzed the effect of the stromal network on follicles' organizations and immune infiltration.

Manual cell gating distinguishes canonical cell types in the lymphoid tissues with the tradeoff of the total number of gated cells

The first step to understanding the spatial difference of the immune cell phenotypes in the SLOs is to annotate the high-parameter data generated using IMC and classify the cellular populations that make up the tissues. Herein, *SpatialVizPheno* incorporated the manual gating approach relying on the hierarchical and Boolean single-cell gating on FlowJo software (Figs. 2A and S7). A sequential gating approach was followed using the rationale in Fig. 1B. This approach starts by gating the entire cell population based on the expression of CD3 and CD20. The $CD3^-CD20^+$ population was further sub-divided into GC B cells based on the expression of $CD27^+CD38^+$. GC B cells were then divided into LZ B cells based on the expression of $CXCR5^+$, and DZ B cells based on the expression of $CXCR5^-Ki67^+$. Further, the $CD3^+CD20^-$ population was divided into Th cells based on the expression of CD4 and Tc based on the expression of CD8. Th cells were further sub-divided into Tfh based on the expression of $CD27^+CXCR5^+PD-1^+$, and Treg based on the expression of $FoxP3^+$. Tc cells were sub-divided into Tex based on the expression of $PD-1^+$ and follicular cytotoxic T (Tfc) cells based on the expression of $CXCR5^+$. Finally, the $CD3^-CD20^-$ population was divided into FDCs based on the expression of $C11b^+CD21^+$, myeloid-derived cells based on the expression of $CD11b^+CD11c^+$, SC based on the expression of collagen and/or vimentin, and PBs/PCs based on the expression of $CD27^+CD138^+$. The manual cell gating separated the tissue regions and the cellular phenotypes in the different types of lymphoid tissues resulting in 16 clusters and a total of 899,921 annotated cells (Figs. 2B and D and S8–S11).

Overall, this approach failed to capture the majority of DZ B cells in the follicles of the lymphoid tissues (Fig. 2B and D). This could be due to high cell density in this region which resulted in imperfect cell gating and later affected the phenotypic identification steps. The separation between the cellular phenotypes was distinct with minimal shared cells among the different phenotypes due to the nature of the sequential as well as Boolean gating applied to the dataset (Fig. 2C, E, and F). The lymphoid tissues showed a heterogeneous composition of the cellular phenotypes with a few interesting patterns (Fig. 2G). LN tissues showed a markedly higher density of B cells and Tc cells. Further, the intestinal PP showed a notably higher density of Treg, Tfh, Tex cells, and stromal regions with collagen and/or vimentin expression (Fig. 2G).

Using FlowJo to analyze high-content imaging data was introduced to quantify complex cellular populations on a large scale and across different experiments. It also offers a platform for a streamlined visualization of the cellular phenotypes on the tissue images for additional validation (19, 20). This workflow started with pixel classification using the software iLastik (21) to manually label pixels as "nuclear" or "background," and train a random-forest classifier for semantic segmentation. iLastik generated probability masks for each label which were later used for single-cell segmentation. The raw IMC data along with the probability masks generated from iLastik were then fed into the CellProfiler software to segment single cells and generate single-cell masks (22). The single-cell masks were then used to extract single-cell protein expression data. Following the pipeline used in specter (19), IMC data were converted to.fcs (flow cytometry standard) file format using the *x* and *y* positions as additional parameters. These files were then used for manual gating on FlowJo (Fig. 2A).

Semi-supervised cell gating distinguishes transitional cell types in the lymphoid tissues with the tradeoff of background noise detection

The tissue composition of the lymphoid tissue is complex due to the phenotypically similar immune cells and dense tissues. Thereby, *SpatialVizPheno* incorporated another approach to classify the cellular phenotypes and to identify the tissue regions. The automated semi-supervised approach was developed through iterative clustering using rounds of the Leiden clustering approach and cluster annotation based on preidentified marker expression (Fig. S7A). This process was repeated until the cellular phenotypes of interest can be detected on the lymphoid tissues which took three rounds for this dataset (Fig. 3A). This approach resulted in 18 clusters and showed superior tissue region separation compared to the manual gating and annotated more cells adding up to 999,970 cells (Figs. 3B and D and S12–S15). Several tissue regions were detected including follicular regions (LZ, DZ, T B cell boundary) as well as stromal regions including the crypted regions (Fig. 3B and D). The cell count of the classified phenotypes using the semi-supervised approach was higher than that of the manual gating approach which led to better spatial separation among the cellular phenotypes (Fig. 3C, E, and F). Similar to the manual gating results, the lymphoid tissue composition was heterogeneous with a few trends per tissue type. The intestinal PP showed a higher density of Tfh and SCs and the LN showed a higher density of T cells (Fig. 3G). Interestingly, we identified three additional clusters that were not assigned specific phenotypes. These clusters include $CXCR5^+$ B cells, $Ki67^+$ cells, and an unknown cluster. The $CXCR5^+$ B cells were abundant in TS and LN tissues whereas $Ki67^+$ cells were prevalent in the intestinal tissues. However, the unknown cluster was expressed in all tissue types.

These clusters were generated through automated iterative clustering and annotation to yield the best cluster separation and phenotype identification. The clusters generated using this approach were also compared with the clusters generated from the manual gating and ranked based on their single-cell phenotype overlap using cross-tabulation scoring. This process was done for the three rounds where the first round resulted in 24 clusters (Fig. 4A). The first round showed poor overlap with the manual gating clusters with the majority of the overlap being lower than 0.2 or 20% (Fig. 4B). It also resulted in poor separation between the LZ and the DZ and within the individual phenotypes in the LZ and the DZ (e.g. FDCs and Tfh) (Fig. 4C). A subset of clusters was further split into sub-clusters or merged into the same clusters which resulted in 20 clusters for the second round (Fig. 4D). This round showed improved overlap with the manual gating clusters with the highest overlap (Fig. 4E). It also showed enhanced tissue region identification on all lymphoid tissues; however, the LZ B cells, DZ B cells, and FDCs were not uniquely separated in the tissues (Fig. 4F). Thereby, another clustering round was performed resulting in 18 clusters (Fig. 4G) with the best overlap with the manual gating clusters (Fig. 4H) and tissue regions separation and cellular phenotype identification (Fig. 4I).

Tissue neighborhood analysis reveals unique interactions among cellular phenotypes in the lymphoid tissues

SLOs have been studied before in the context of health and disease; however, the spatial distribution and the neighborhood among immune phenotypes are not yet well studied. This can shed light on the similarities and differences among the different tissue types and help us develop a deeper understanding of
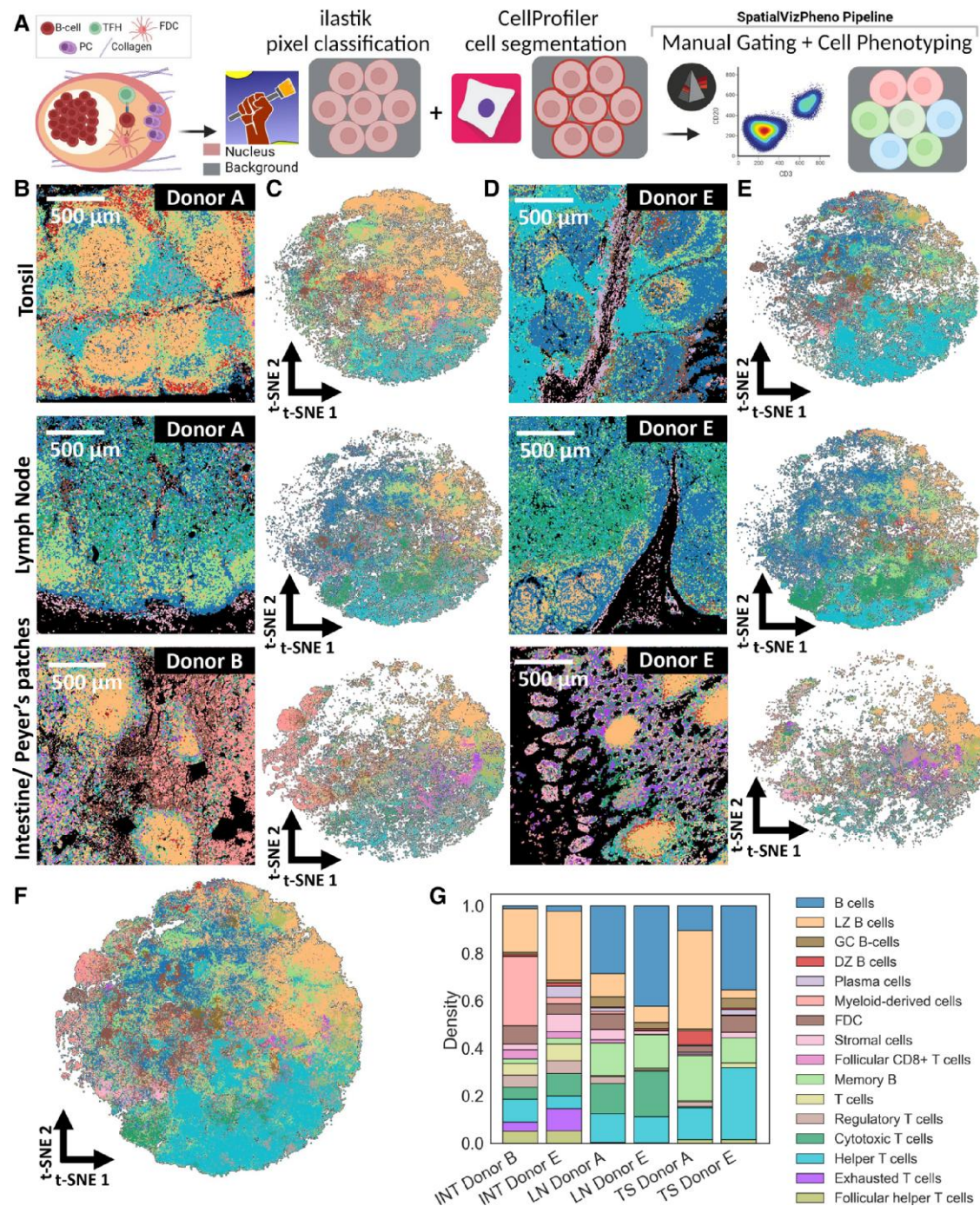
Fig. 2. Manual cell gating results in specific cell type separation with the tradeoff of the total number of gated cells. A) The workflow of the manual gating consisted of pixel classification using iLastik software with two labels including the nucleus and the background. The generated probability masks along with the lymphoid tissue IMC raw data were then added to CellProfiler software for cell segmentation. The single-cell data were then added to FlowJo software for cellular phenotype gating. Created with BioRender.com. B, D) Representative tissue images showing the distribution of the gated cellular phenotypes on tonsil, lymph node, and intestinal PP tissues. The scale bar is 500 μm. C, E) t-SNE plot showing the distribution of the gated phenotypes on the tonsil, LN, and intestinal PP tissues in panels b and d respectively. F) t-SNE plot showing the distribution of the gated phenotypes from all the six human tissue samples included in the study. G) Bar plot showing the composition of the human tissues from the gated phenotypes.

adaptive immunity. Thereby, *SpatialVizPheno* incorporated an additional analytical step in the workflow to quantify the spatial correlations between the identified phenotypes within TS, LN, and intestinal PPs. This was accomplished by quantifying the count of the cells that belong to a specific phenotype within a radius of 20-μm from each cell centroid and calculating their significance using a permutation-based approach (Fig. S16). This was performed for phenotypes identified from both the manual gating

approach and the semi-supervised approach to understand their similarities and differences.

For TS tissues, the manual gating showed that PBs/PCs were found to be neighboring DZ B and FDCs. Additionally, SCs were found to be spatially correlated with FDCs (Figs. 5A and C, S17, and S18). On the other hand, the semi-supervised gating approach revealed more interesting neighboring interactions such that it showed a high correlation between the FDCs and Tfh, as well as
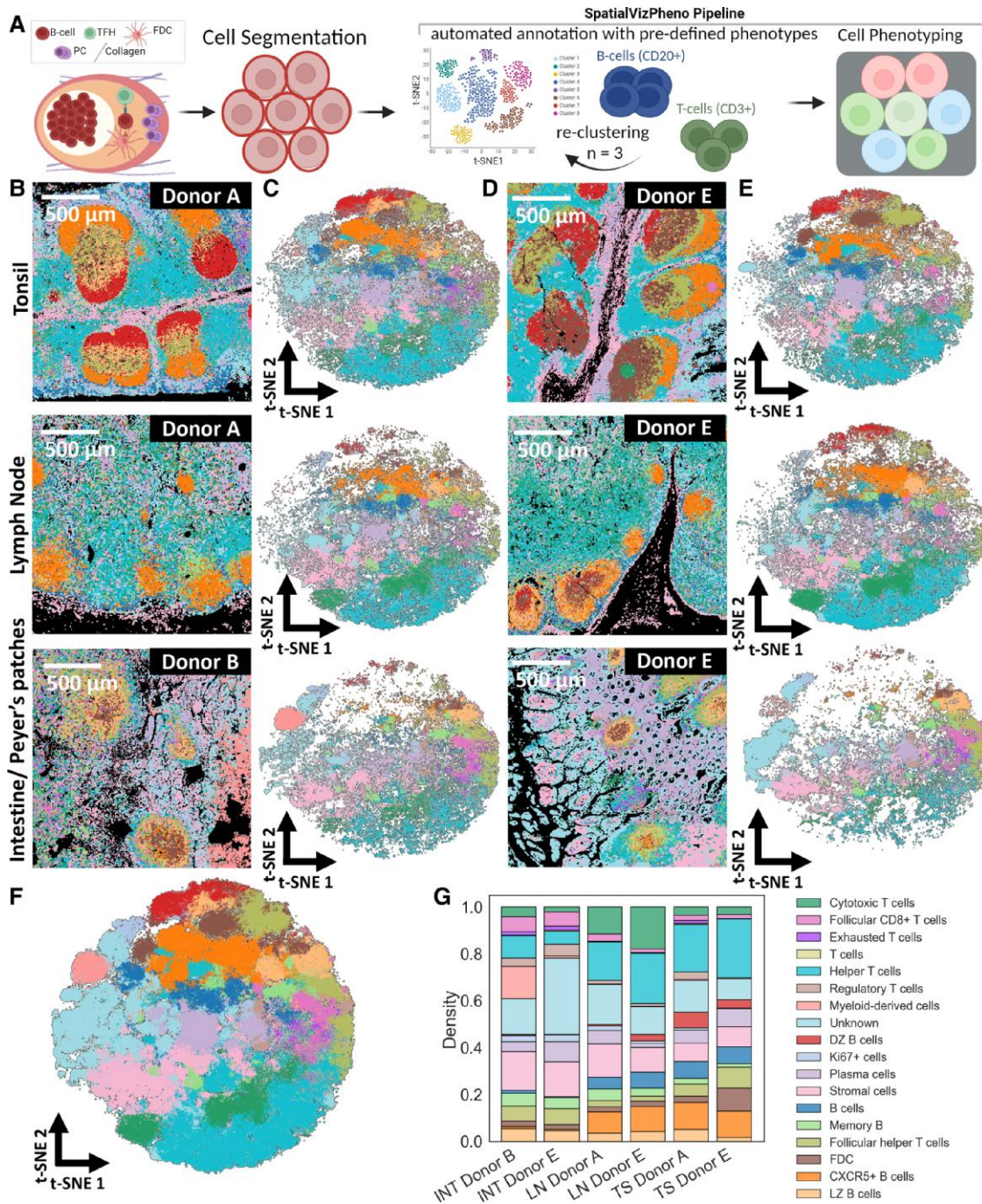
Fig. 3. Semi-supervised cell gating results in specific cell type separation with the tradeoff of background noise detection. A) The workflow of the automated semi-supervised gating began with cell segmentation, followed by iterative cycles of Leiden unsupervised clustering and manual annotation based on pre-established marker expression. This iterative process continued until the desired cellular phenotypes were detectable within the lymphoid tissues. In this dataset, it required three rounds of iteration to achieve this goal. Created with BioRender.com. B, D) Representative tissue images showing the distribution of the gated cellular phenotypes on tonsil, lymph node, and intestinal PP tissues. The scale bar is 500 µm. C, E) t-SNE plot showing the distribution of the gated phenotypes on the tonsil, LN, and intestinal PP tissues in (B) and (D), respectively. F) t-SNE plot showing the distribution of the gated phenotypes from all the six human tissue samples included in the study. G) Bar plot showing the composition of the human tissues from the gated phenotypes.

SCs cells. PBs/PCs and SCs as well as Tc and Th were also found to be spatially correlated (Figs. 5B and C, S19, and S20).

For the LN tissues, the manual gating showed close spatial neighboring between Tc cells with Th cells which are similar to TS tissues. Moreover, there was high spatial interaction between FDCs and SCs as well as PBs/PCs with SCs, matching with the findings from the TS tissues (Figs. 5A and C, S17, and S18). Similar to

manual gating, semi-supervised gating showed high spatial neighborhood interaction between Tc and Th as well as between FDCs and SCs. It also showed high interaction between DZ B cells and SCs as well as between Tfh and LZ B cells (Figs. 5B and C, S19, and S20).

For the intestinal tissues, the manual gating showed high neighborhood interaction between Tc and Th, between PBs/PCs
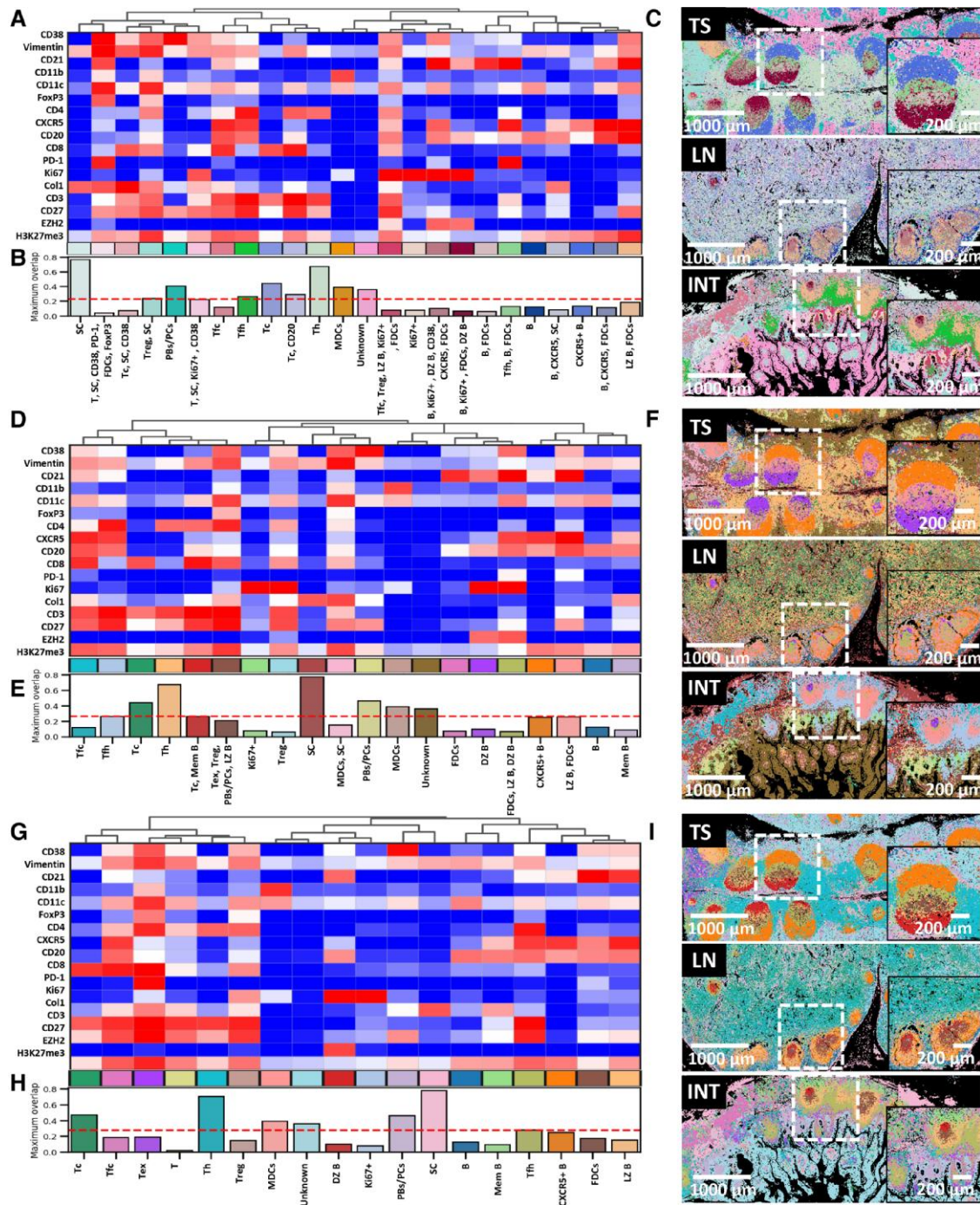
Fig. 4. The iterative clustering and annotation identify tissue regions and unique cellular phenotypes on the lymphoid tissues. A, D, G) Heatmap showing the marker expression profile of the results from the first, second, and third clustering and annotation rounds respectively using the unsupervised Leiden clustering. B, E, H) Bar plot showing the overlap between clusters generated from the manual gating and the clusters generated from the first, second, and third clustering and annotation rounds respectively. C, F, I) Representative tissue images showing the distribution of the clusters onTS, LN, and intestinal PP tissue images with the clusters generated from the first, second, and third clustering and annotation rounds, respectively. TS, tonsils; LN, lymph nodes; INT, intestinal Peyer's patch samples.

and SCs similar to the TS and LN tissues. Interestingly, there was also a high interaction between PBs/PCs and Tfc, FDCs and Treg, as well as Tfc and SCs which are all unique to the intestinal tissues (Figs. 5A and C, S17, and S18). On the other side, the semi-supervised gating showed high interaction with FDCs and Tfh, Tc, and Treg cells which matches with the manual gating results (Figs. 5B and C, S19, and S20). Overall, the most predominant neighborhood interaction across tissue types and donors was between PBs/PCs and SCs. Intestinal tissue showed the most variant

distribution of spatial neighborhoods among the immune pheno-types signifying their diverse immune microenvironment.

Collagen organization changes the follicles' organization and polarity

Given the importance of the stromal network in the SLOs, we sought to investigate the correlation of the stromal network with the organization of follicles and the polarization of GCs.
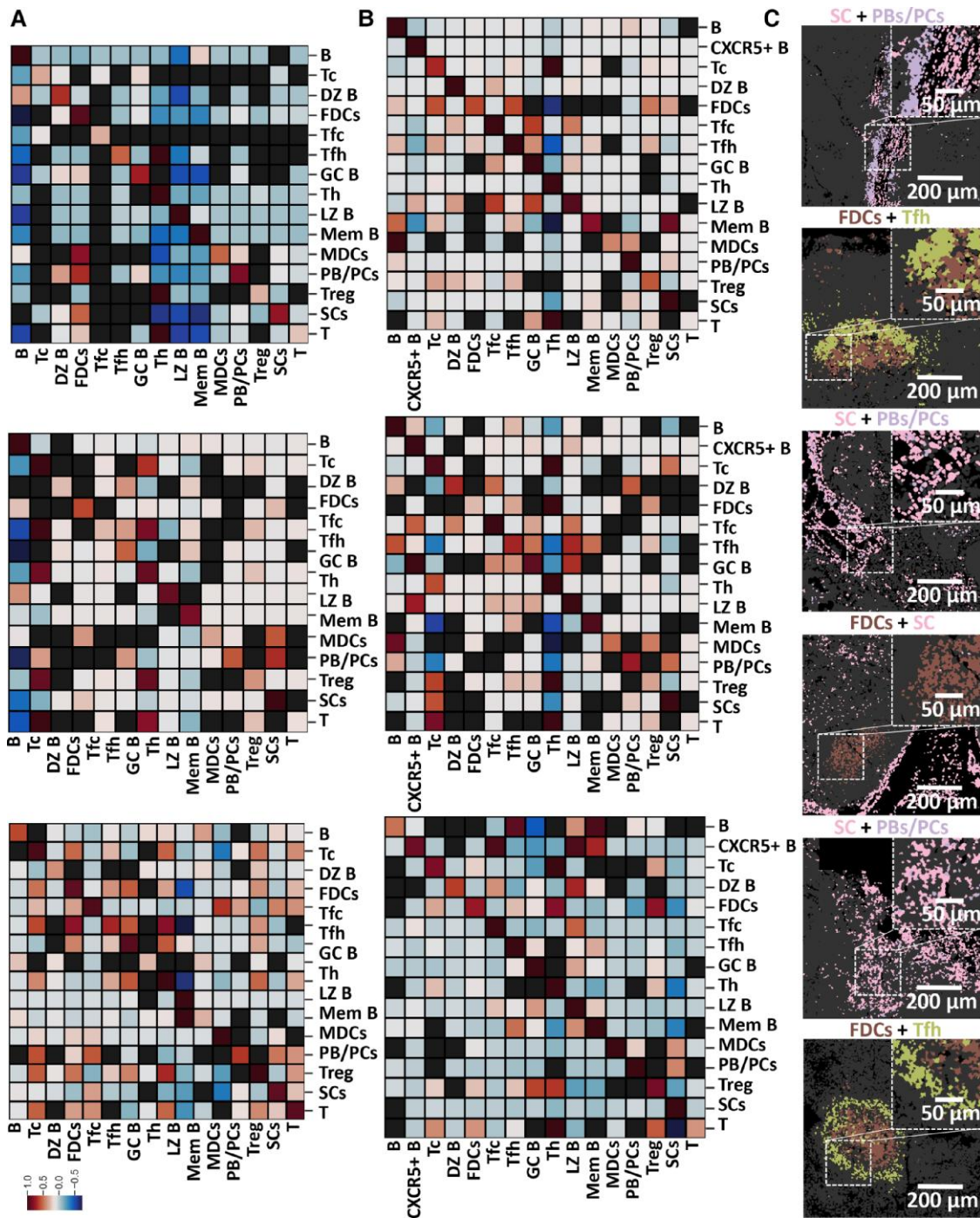
Fig. 5. Tissue neighborhood interaction analysis reveals the unique connections between cellular phenotypes using the manual gating and unsupervised clustering approaches. A) Heatmap showing the interaction between the resulting cellular phenotypes using the manual gating approach. The positional interaction is scaled between −0.05 and 1. B) Heatmap showing the interaction between the resulting cellular phenotypes using the automated semi-supervised approach. The positional interaction is scaled between −0.05 and 1. C) Visual representations of the spatial neighborhood interactions using IMC raw images and phenotypic clusters generated using both the manual and the automated semi-supervised gating.

This can shed light on the effect of the stromal structure on the organization as well as the function of GCs in response to infections and vaccinations. *SpatialVizPheno* analyzed the organization of the different phenotypes of immune cells around stromal regions by annotating collagen-expression regions as an individual zone and referred to as zone 1. Further, eight sequential additional areas were annotated away zone 1 such that each zone is 50-μm away from the former zone (Fig. 6A). Immune cell phenotypes in all nine zones were analyzed to understand their spatial proximity to stromal regions (Fig. 6A). This analysis was performed on the annotated B cell phenotypes (Fig. 6B) as well as T cell phenotypes (Fig. 6C) on all lymphoid tissues included in the study.

PBs/PCs show increased infiltration away from collagen-expressing regions in intestinal PP and in the LN tissues. However, the distribution of this population is variable in TS tissues (Fig. 6B and D). Memory B cells show variable infiltration in the stromal vicinity of TS tissues; however, their infiltration increases across the zones away from collagen regions in LN tissues.
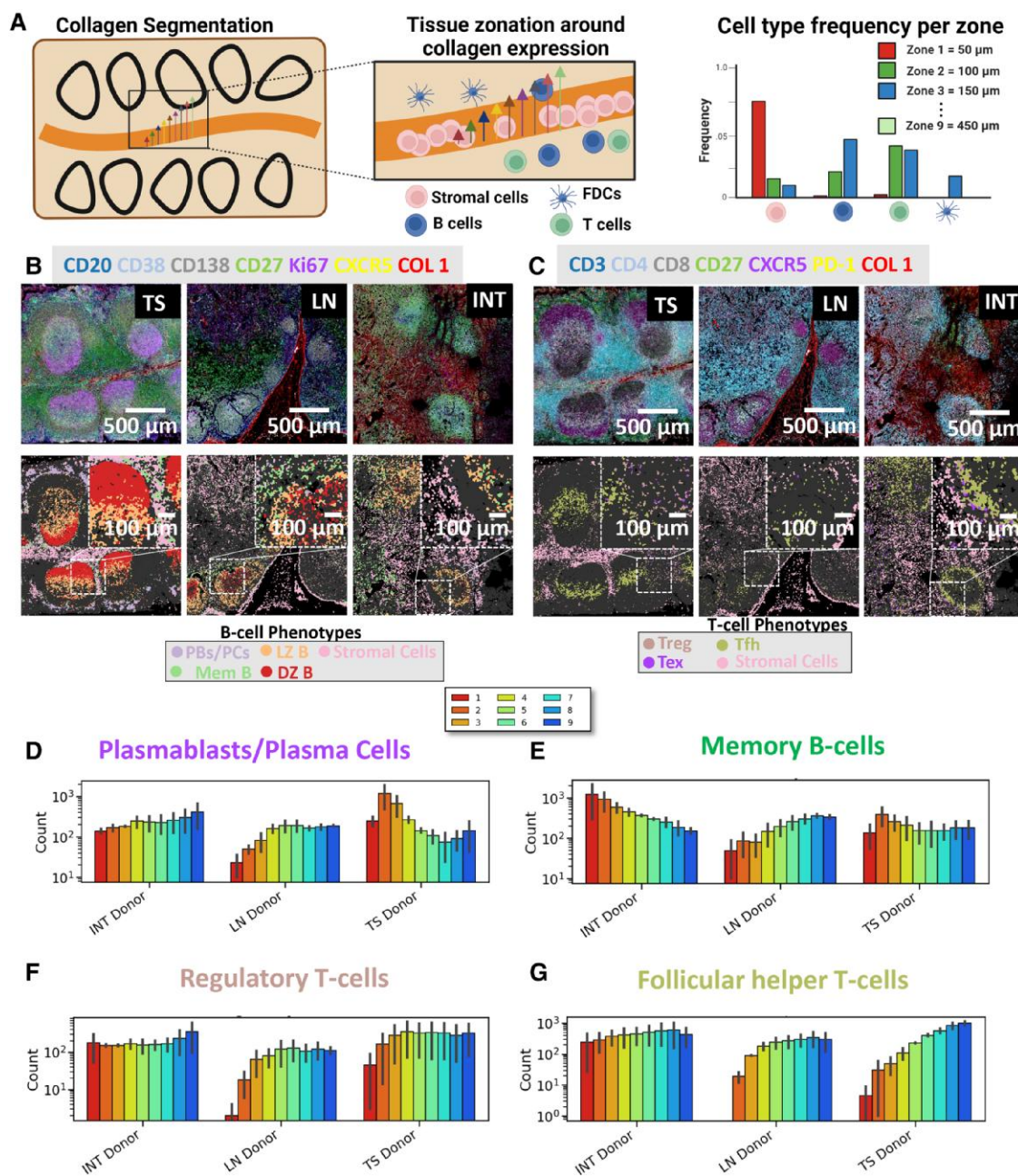
Fig. 6. Immune cell phenotypes show distinct distribution around stromal regions in secondary lymphoid tissues. A) The workflow of this analysis included annotating collagen-expressing regions and dividing the regions around the collagen into nine zones and each zone was 50-μm away from the previous zone. The T cells and the B cell phenotypes within the nine zones were then analyzed across the different secondary lymphoid tissues. Created with BioRender.com. B) IMC images showing the T cell markers in the dataset along with representative tissue images showing the annotated B cell phenotypes in the dataset. C) IMC images showing the B cells markers in the dataset along with representative tissue images showing the annotated T cell phenotypes in the dataset. D–G) Bar plot showing the count of D) plasmablasts/plasma cells, E) memory B cells, F) regulatory T cells, and G) follicular helper T cells along the nine annotated zones in the vicinity of collagen-expressing regions.

Intestinal tissues have the opposite distribution of this population (Fig. 6B and E). Treg cells were found to have a similar trend in TS and LN tissues where they show increased infiltration until zone 4 and then the permeation of this population subsides. Intestinal PPs, on the other side, show variable distribution of Treg (Fig. 6C and F). Tfh cells show a consistent trend across all lymphoid tissues such that the infiltration increases further away from stromal/collagen-expressing regions. However, this trend was more predominant in TS and LN tissues (Fig. 6C and G). Overall, our study highlighted the heterogeneity among lymphoid tissue types for the single-cell distribution concerning the stromal network.

To further understand the organization of the SLO, *SpatialVizPheno* also analyzed the polarization of the follicles in secondary lymphoid tissues. Herein, two methods were used to understand the correlation between the stromal network and follicles' morphology and organization. The follicles from the TS, LN, and intestinal tissues were annotated using a subset of phenotypes that are known to be prevalent in the follicles. Then, the follicles' segmentation was automated in all the lymphoid tissues, adding up to 95 follicles. Further, collagen-expressing regions were divided into square patches with the size of 10 μm × 10 μm. A circle with a radius of 500 μm was added to the center of each

follicle and collagen search was performed within the circle through angular sampling. Collagen-expressing patches were averaged and the direction with the highest collagen expression was used as the major axis. This technique was referred to as collagen-based orientation. On the other hand, another technique was used to define the major and the minor axes of the follicle which relied on the morphology of the single follicles. The longest radius of the follicle/ellipse was used as the major axis and referred to as shape-based orientation (Fig. 7A).

The collagen-based and shape-based orientations were significantly different for TS and LN tissues (Figs. 7B–E, S21, and S22). Conversely, the collagen-based orientation showed high similarity to that of the shape-based orientation for the intestinal tissues (Figs. 7B–E and S23). To quantify the difference between the two methods, we analyzed the angle difference between the major axes using both the collagen-based and the shape-based orientation (Fig. 7F–H). For TS tissues, 13 out of 44 follicles had an angle difference between $0.65\pi < x \le 1\pi$, 16 follicles had an angle difference between $0.35\pi < x \le 0.65\pi$, and 15 follicles had an angle difference between $0 < x \le 0.35\pi$ (Fig. 7F and H). For LN tissues, 15 out of 27 follicles had an angle difference between $0.65\pi < x \le 1\pi$, 5 follicles had an angle difference between $0.35\pi < x \le 0.65\pi$, and 7 follicles had an angle difference between $0 < x \le 0.35\pi$ (Fig. 7G). For intestinal tissues, 5 out of 24 follicles had an angle difference between $0.65\pi < x \le 1\pi$, 6 follicles had an angle difference between $0.35\pi < x \le 0.65\pi$, and 13 follicles had an angle difference between $0 < x \le 0.35\pi$ (Fig. 7H). TS and LN tissues exhibited collagen I-rich regions due to the collagen conduit system and the crypt epithelium, respectively. This rich stromal network changed the organization of follicles in both TSs and LN, but not in the intestinal tissues.

To further understand the effect of the stromal network, *SpatialVizPheno* also analyzed the gradient of the markers' expressions, and the frequencies of immune phenotypes across the major and the minor axes generated from both the collagen-based orientation and the shape-based orientation (Fig. 8A). For TS tissues, CXCR5 was found to have a decreasing expression across the major axis in the collagen-based orientation, but not in the shape-based orientation (Fig. 8B). This pattern was also visualized in the phenotypes for the LZ B cells that have a decreasing frequency across the major axis in the collagen-based orientation (Fig. 8E). The expression of Ki67 along the major axis showed a significantly greater increase in the collagen-based orientation compared to the shape-based orientation (Fig. 8B). Similarly, this trend was visualized for the DZ B cells in the collagen-based orientation to a higher extent than in the shape-based orientation (Fig. 8E).

For LN tissues, Ki67 had a decreasing gradient across the major axis in the collagen-based orientation but had an increasing gradient across the major axis in the shape-based orientation (Fig. 8C). The same trend was visualized for DZ B cells as well (Fig. 8F). For intestinal tissues, there was no observed difference in marker expression gradient or the immune phenotypes between the collagen-based orientation and the shape-based orientation (Fig. 8D–G). Therefore, the orientation of collagen expression is predictive of the LZ and the DZ polarization for the follicles in TSs and LN tissues, but not for intestinal tissues.

Discussion

Herein, the immune microenvironment in SLOs including TS ($n = 2$), LN ($n = 2$), and intestinal PP ($n = 2$) was studied with the high-parameter IMC technology using an antibody panel of 25 markers that spans B cell, T cell, SC, and key regulatory markers (e.g.

chemokines, proliferation, and epigenetic). To dissect the single cells and high-parameter IMC dataset, a highly modular computational workflow, *SpatialVizPheno*, was developed. Altogether, we have shown the rich immune cellular phenotypes among the tissues and highlighted the difference among the different SLOs in terms of the spatial neighborhoods, the stromal network, and the follicular organization (Fig. 1).

To account for the complex tissue composition, *SpatialVizPheno* incorporated two methods to identify phenotypically similar immune cells. The first approach is the manual gating approach that involves converting IMC data to a.fcs file format that is compatible with the widely used FlowJo software to allow for the cell phenotype gating in a fashion similar to flow cytometry data (19) (Fig. 2A). This approach resulted in the phenotypic identification for 16 cellular phenotypes on the SLOs tissues (Fig. 2B and D). Due to the nature of the sequential cell gating applied on FlowJo, there were minimal shared cells among the identified phenotypes which resulted in better segmentation of phenotypes even for nonprevalent populations (e.g. plasmablasts/plasma cells) (Fig. 2C, E, and F).

The second approach is the automated semi-supervised gating that involves iterations of unsupervised clustering followed by cluster annotations based on preidentified marker expressions for the unique phenotypes. This process was repeated ($n = 3$) until each of the clusters corresponded to a unique cellular phenotype (Fig. 3A). Some clusters were found to contain several phenotypes that needed to be split while other clusters showed individual marker expressions that needed to be merged (Fig. 4A, D, and G). This approach identified 18 cellular phenotypes that highlighted the difference between the SLOs (Fig. 3B and D). LN showed the highest density of B cells, and the intestinal PP showed the highest density of Tfh cells (Fig. 3G). Prior studies showed that the intestinal PP had an increased Tfh differentiation driven by the gut microbiota. This makes the immune microenvironment in the intestinal PP resemble the early stages of the GC response leading to the generation of diverse low-affinity B cell clones and antibodies to fight the microbial antigens (23).

To further understand the interactions among the cellular phenotypes, the spatial neighborhood was quantified as the probability of cell phenotypes interacting within a distance of 20 μm. This was performed for all the phenotypes generated using the manual gating, and the semi-supervised gating (Fig. 5A and B). In TS tissues, DZ B cells were found to be neighboring PBs/PCs using the manual gating approach (Figs. 5A and C, S17, and S18). Prior research showed that PBs as well as PCs are formed around the boundary between DZ B cells and the surrounding stroma. This is derived by Tfh cells secretion of IL-21 that derives the production of PBs, the migration of PCs, and fibroreticular cells secretion of TNFSF13 (APRIL) which further supports the proliferation of PCs (24, 25). While both PBs and PCs can localize at the GC T cell zone (GC T) near the DZ boundary, it was not possible to make this distinction due to various reasons. First, PBs and PCs exhibit similar protein expression profiles, making it difficult to distinguish between them with the antibody panel included in this study. For example, CD138 is a commonly used marker to identify PCs, but it could also be expressed at variant levels on PBs. CD27 and Ki67 are other markers that can be expressed by both PCs and PBs at different levels (26). Additionally, our dataset could be capturing the transition of phenotypes from PBs to PCs which is evident by the neighborhood interaction results. Further, FDCs were positioned close to SCs and PBs/PCs in TS tissues using the manual gating approach (Figs. 5A and C, S17, and S18). FDCs are a unique population of dendritic cells because they are not from bone
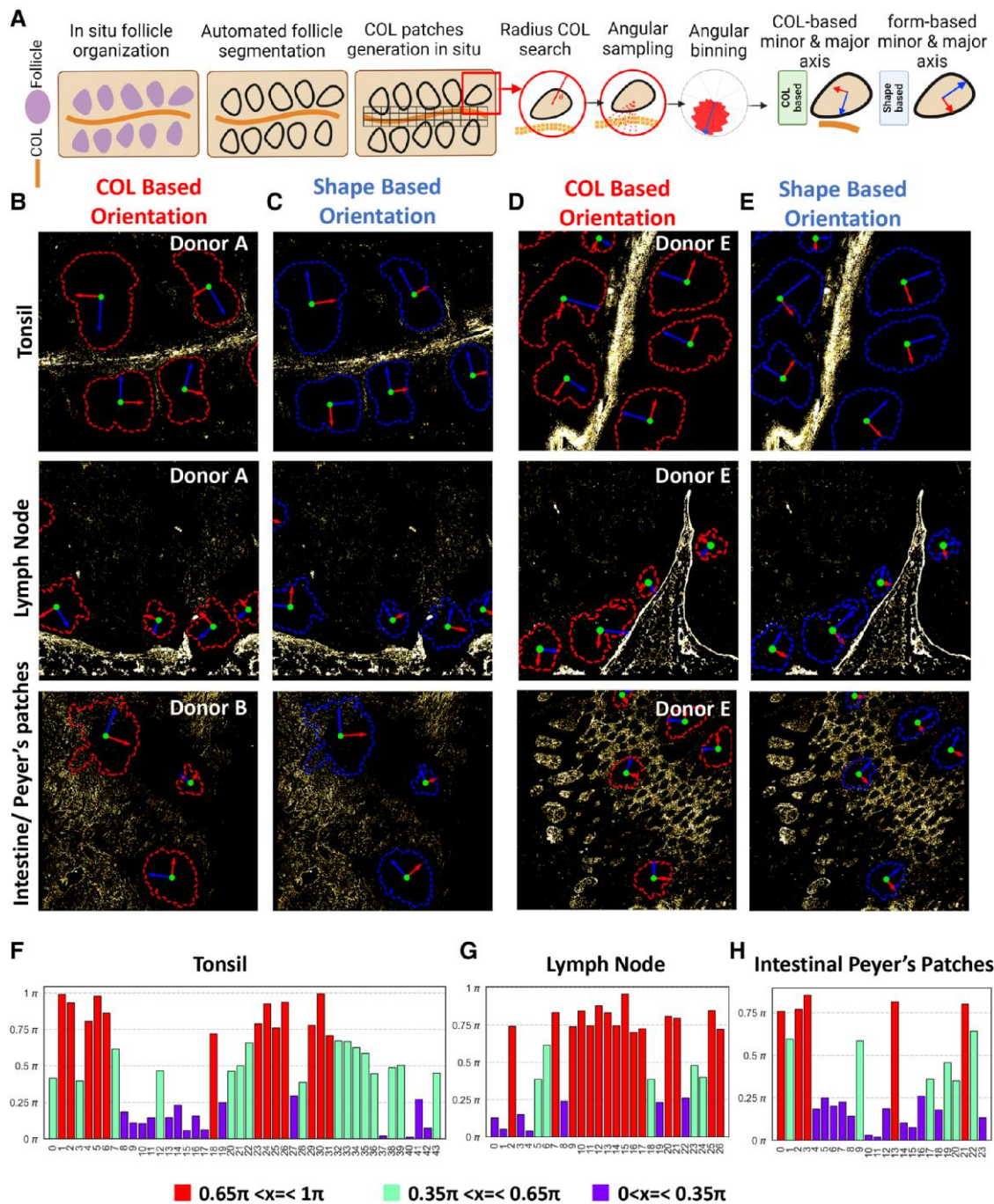
Fig. 7. Stromal cell network changes the morphology and the orientation of follicles in secondary lymphoid tissues. A) The workflow of this analysis included automated segmentation of the follicles and dividing the collagen-expressing regions into square patches with the size of 10 µm × 10 µm. Then, a circle with a radius of 500 µm was added to the center of the segmented follicle and the circle was divided into 40 sections where collagen+ patches were searched in each section for the angular sampling. The collagen expression of the patches in the sections was averaged for the angular binning. The direction with the highest collagen expression was used as the major axis for the collagen-based orientation. On the other hand, the shape-based orientation relied on the morphology of the follicles where the longest side of the follicle/ellipse was used as the major axis. Created with BioRender.com. B, D) Representative images show the collagen-based follicle orientation for the tonsil, lymph node, and intestinal tissues. Collagen expression is in yellow, the major axis in blue, and the minor axis in red. C, E) Representative images show the shape-based follicle orientation for the tonsil, lymph node, and intestinal tissues. Collagen expression, the major axis, and the minor axis are shown. F–H) Histogram showing the distribution of angle difference between the major axes generated using the collagen-based and shape-based orientation for the segmented follicles in F) the TS, G) the LN, and H) the intestine tissue samples. The angle differences in the range $0.65\pi < x \leq 1\pi$, $0.35\pi < x \leq 0.65\pi$, and $0 < x \leq 0.35\pi$ are presented.

marrow hematopoietic stem cell origin but rather from a mesenchymal origin. Thereby, they are considered to be SCs and part of their function is to maintain the microarchitecture of the follicles. This could explain its correlation with the stromal network (27).

Moreover, FDCs' affinity selection during the GC response was shown to determine the fate of B cells and the production of plasma cells. They also secrete critical cytokines required for B cell activation and survival including B cell activating factor (BAFF)
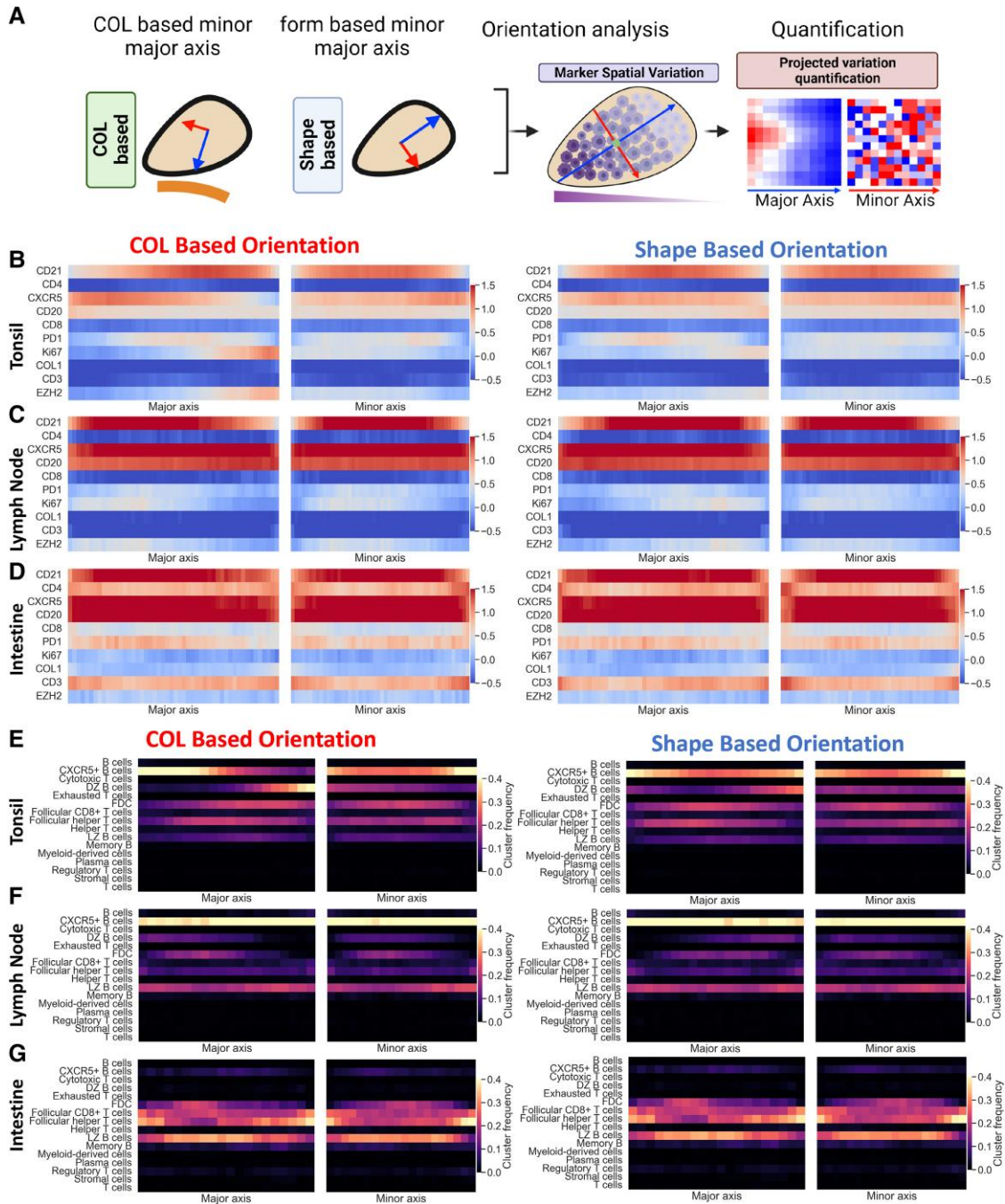
Fig. 8. Stromal cell network changes the organization of immune cell phenotypes in secondary lymphoid tissues. A) The workflow of this analysis included analyzing the gradient of marker expression, and the frequencies of immune cell phenotypes across the major and the minor axes generated from both the collagen-based orientation and the shape-based orientation. Created with BioRender.com. B–D) Heatmap showing the frequencies of marker expression gradient across the major and minor axes in the follicles using the collagen-based orientation and the shape-based orientation in (B) TS, (C) LN, and (D) intestinal tissues. E–G) Heatmap showing the frequencies of immune cell phenotypes across the major and minor axes in the follicles using the collagen-based orientation and the shape-based orientation in (E) TS, (F) LN, and (G) intestinal tissues.

which is directly linked to plasma cells' survival which can explain the neighboring between FDCs and PBs/PCs (28). Further, there was a high correlation between the FDCs and Tfh cells as well as between Tc and Th in TS tissues using the semi-supervised gating (Figs. 5B and C, S19, and S20). At the LZ, centrocytes were shown to receive the B cell receptor (BCR) stimulation by FDCs and repeated stimulation by Tfh to maintain the GC response which could explain the neighboring between FDCs and Tfh (25). Further, it has been shown that the presence of CD8$^+$ T cells (Tc)

is colocalized with Tfh cells in the GCs to limit their rapid expansion and to keep the GC response in check which justifies the neighboring between Tc and Th (29).

LN tissues exhibited several shared neighborhood interactions with the TS tissues. However, the semi-supervised gating also showed a high interaction between DZ B cells and SCs as well as between Tfh and LZ B cells (Figs. 5B and C, S19, and S20). The rapidly dividing DZ B cells are known to express CXCR4 which attracts them to the ligand CXCL12 expressed by the stromal network.

This interaction is critical to localizing the centroblasts within the DZ and maintaining the GC response (5). Further, Tfh cells are localized near LZ B cells to maintain their activation along with FDCs (25). Overall, it can be noted that the immune cells distribution within TS and LN tissues is very similar suggesting the same sequence of events during the GC response.

Intestinal tissues showed several unique neighborhood interactions including PBs/PCs and Tfc, FDCs and Treg, as well as Tfc and SCs using the manual gating approach (Figs. 5A and C, S17, and S18). Further, the semi-supervised gating showed several unique interactions including FDCs and Tfh, Tc, and Treg cells which matches with the manual gating results (Figs. 5B and C, S19, and S20). In contrast to TSs and LN that produce antigen-specific GC response, intestinal tissues are under chronic inflammation to induce an ongoing response to food antigens. This continuous immune response needs to be tightly regulated to prevent systemic autoimmunity. This is controlled by Treg cells and Tc cells which could explain the higher prevalence of these cell types and their neighboring with other cells. Even though the exact neighboring interactions cannot be explained, the higher density of suppressive T cell phenotypes can explain this observation (30).

The most observed neighborhood interaction in all tissue types was between FDCs and SCs and between PBs/PCs and SCs. In response to antigen, the stromal network along with FDCs is known to expand and polarize the GC into the LZ and the DZ. This ensures the interactions between immune cells in the GC and the maintenance of the immune response (5). Further, plasma cells were shown to rely on colocalization with SCs for both migration and survival. For example, the reticular SCs form a plasma cells niche through their expression of CXCL12 which acts to both maintain plasma cells' survival and to derive plasma cells migration to the bone marrow (31). Further, intestinal tissue showed the most unique distribution of spatial neighborhoods among the immune phenotypes signifying their diverse immune microenvironment. LN tissues showed the highest spatial neighborhood correlations with several phenotypes of T cells which could potentially be due to the higher T cell population (Fig. 5A and B) (5).

Inspired by the spatial neighborhood results, we further sought to understand the distribution of SCs in the SLO and their neighboring phenotypes. The stromal network in the secondary lymphoid structures acts as a scaffold for the organization of immune cells. It is responsible for transporting antigens, facilitating cellular migrations or adhesion, and maintaining chemokines gradients (32). Collagen-expressing regions were segmented from all tissues in the study and annotated as zone 1. The neighboring regions were segmented into eight sequential areas away from zone 1 where each zone is 50 μm away from the preceding one (Fig. 6A). The cellular phenotypes existing within the nine zones were later quantified (Fig. 6D–G). Plasmablasts/plasma cells were found to have a similar pattern to memory B cells in TS and LN tissues (Fig. 6D and E). This spatial neighborhood can be attributed to memory B cells differentiating into the plasmablasts/plasma cells upon antigen re-exposure (33). Further, plasmablasts/plasma cells were found to increase in infiltration away from collagen-expressing regions (Fig. 6D). Collagen is predominantly expressed in the submucosal layer of the intestinal PPs and PBs/PCs were previously shown to be populated in the subepithelial domes which is distal to the submucosa. This could explain the increased infiltration of PBs/PCs away from collagen-expressing regions (9, 10).

Memory B cells showed a variable distribution across the zone in TS tissues. On the other side, LN tissues have a consistent increase of memory B infiltration across the zones, and the opposite is observed for intestinal tissues (Fig. 6E). Memory B cells generated in TS tissues primarily remain localized within the lymphoid tissues, whereas memory B cells generate in LN tissues migrate to peripheral tissues throughout the body. TS tissues are part of the mucosal-associated lymphoid tissues (MALT) which act as the primary site for the immune response against pathogens. MALT has chemokines and cell adhesion molecules which can help in retaining the memory B cells. On the other hand, LN tissues act as an immune hub and it lacks homing receptors for the memory B cells, leading to the migration of memory B cells to other peripheral tissues (34, 35). Regulatory T cells showed an increased infiltration away from collagen-expression regions in TS and LN tissues, whereas they showed a consistent infiltration across all zones in the intestinal PP (Fig. 6F). This could be attributed to the importance of Treg cells in maintaining homeostasis under the chronic inflammation in the gut induced by the ingested antigens (36). Follicular helper T cells showed a similar pattern to memory B cells in LN tissues signifying the importance of Tfh in the maintenance and the generation of memory B cells in these tissues (Fig. 6E and G) (37).

The follicles' morphologies were later quantified concerning the collagen expression through the identification of the major and the minor axes of the follicles. Two approaches were tested including using collagen expression as the identifying factor for the major axis or using the shape of the follicles/ellipses (Fig. 7A). TS and LN tissues showed different follicle orientations between the collagen-based orientation, and the shape-based orientation. This could be due to the crypt epithelium in the TS tissues and the collagen conduit system in the LN tissues (Fig. 7B–E). This rich stromal network changed the organization of follicles in both TS and LN, but not in the intestinal tissues (Fig. 7F–H). To further confirm this, the marker expression gradient as well as the cellular phenotypes were quantified in both the major and the minor axes using the collagen-based orientation and the shape-based orientation (Fig. 8A). Collagen expression was found to be a predictor of the LZ and the DZ polarization for the follicles in TS and LN tissues, but not for intestinal tissues (Fig. 8E–G).

The distribution pattern of the collagen-expressing network around the follicles can help in inferring the role of these fibers in lymphocyte trafficking and antigen transport. In TS and LN tissues, the collagen fibers were oriented parallel to the T cell zone of the GCs where Tfh cells interact with B cells for activation. The activated B cells might have migrated toward the DZ where they proliferate rapidly which could explain why DZ B cells and Ki67 are closer to collagen-expressing fibers in both TS (5). After undergoing somatic hypermutation, B cells migrate toward the LZ to test their interaction with FDCs and their uptake of the antigen which could explain why LZ B cells, CXCR5, and CD21 expression is further from collagen-expressing regions in TS tissues (5). The opposite is true for the LN tissues such that Ki67 and DZ B cells are expressed away from collagen-expressing regions whereas CXCR5 and LZ B cells are closer to the stromal network. Even though we cannot explain this trend, it is notable to observe the differences in immune cells' distribution around the stromal structure in the lymphoid tissues. On the other hand, antigen presentation is different in the intestinal PPs because they have a continuous GC response opposite to the stimulation-driven GC response in TS and LN tissues. Antigens are transported from the gut lumen through a specialized class of epithelial cells which is the follicle-associated epithelium (9). Activated B cells then migrate toward the preformed GCs. This can explain the lack of a noticeable pattern for marker expressions or cell phenotypes with collagen-expressing regions in the intestinal PPs.

Although this is a powerful workflow to investigate the differences between SLOs (Fig. S24), there are limitations. For example, a few of the markers included in the IMC panels resulted in a weak signal which could be due to the intrinsic nature of the IMC technology that lacks signal amplification. Thereby, we tested supplementing the IMC dataset with a multiplexed immunofluorescence (IF) experiment with additional markers. We also tested a data augmentation pipeline to merge the IMC dataset with the IF dataset (Fig. S25A). Some of the markers used within the IF panel (e.g. VCAM1, GLUT1, IgD, and CD138) resulted in an enhanced separation between GC zones (LZ and DZ) (Fig. S25B and C). Future direction for this panel will include developing a pipeline to perform a pixel-level matching between the two different technologies (IMC and IF) such that datasets generated from either technology can be incorporated into the same dataset. Here, we show a proof-of-concept for this approach to further supplement faulty or missing data from large-scale imaging multiplexed experiments.

Another limitation of the study is that all of the tissues were acquired from third-party vendors as healthy donors without justifying the reason for the surgical removal of the tissues. Further, the difference between the tissues could be due to a recent infection, vaccination, or any immunological stress the donors had faced before donating the tissues. The tissues were from different sexes and variant ages, and we did not have access to the ethnic background. Therefore, all these factors can affect the GC response in the SLOs. This study shed light on the differences between SLOs, but future studies can incorporate more tissues from a specific age group, disease model, or ethnicity to answer more targeted questions related to adaptive immunity in humans.

Materials and methods
Secondary lymphoid tissues
All of the tissues used in this study are formalin-fixed, paraffin-embedded, and with a thickness of 5 μm. They were all acquired from third-party vendors including TissueArray.com (previously: Biomax) and PrecisionMed. Tissue samples were collected with informed consent from patients, adhering to stringent ethical and medical standards. Before distribution to researchers, all tissue samples were de-identified. The human TS tissue section and the LN tissue sections were both from TissueArray.com under the IDs HuFPT161 and HuFPT210, respectively. TS sample 1 had tissue ID SU1 and TS sample 2 had tissue ID SM2. LN sample 1 had tissue ID lly06N022A9 and LN sample 2 had tissue ID lly50N001A4. The human normal small bowel tissue sections were acquired from PrecisionMed. Human intestine sample 1 had tissue ID S09-10377 and human intestine sample 2 had tissue ID S09-11278.

SpatialVizPheno workflow introduction
SpatialVizPheno workflow starts by performing pixel classification using commercially available software, iLastik. This software was used to manually label and annotate cellular regions vs. background on several regions of interest on the tissue images. This step trained the software to generate probability masks that delineate nucleated regions on the tissue images, enhancing the capability to perform single-cell segmentation and subsequent data analysis. These probability masks generated using iLastik along with the raw intercalator (191Ir/193Ir) IMC images were then loaded into another commercially available software, CellProfiler where single-cell segmentation was performed and

cell masks were generated. Then, the single-channel raw IMC images along with the cell masks were fed into a custom Python code and converted to a single.fcs file format following a previously published method (19). These files were then with the widely used software in the field of flow cytometry, FlowJo where single-cell phenotypes were gated for the manual gating approach. The gated cells were exported from FlowJo in the form of .csv files that also included their x and y locations on the tissues. This information was later used for downstream processing. Simultaneously, another approach was explored to identify cellular phenotypes. This involved iterative cycles of clustering, manual annotations of clusters, and re-clustering using custom Python codes. These iterations were performed automatically using preidentified marker expressions for each phenotype until each of the clusters was associated with a unique phenotype and this approach was referred to as semi-supervised clustering. Using phenotypes/clusters generated from both techniques, tissue neighborhoods were evaluated in all SLOs included in the study. The effect of the SLOs' stromal network organization was then explored solely based on the phenotypes generated from the semi-supervised gating, given their superior performance in tissue region identification. The data can be found at https://zenodo.org/record/8303732, and all accompanying data analysis pipelines can be found at https://github.com/coskunlab/SpatialVizPheno.

Pixel classification and single-cell segmentation
iLastik (v 1.4.0) was used to predict the pixels corresponding to the nuclei and the background using the Intercalator (191Ir/193Ir) images. Two labels were created corresponding to nuclei and background, and the software was manually trained to detect the two labels by annotating a few regions for each label. The software then generated probability maps for each label. These probability maps along with the raw images were then added to the CellProfiler (v 4.2.1) software for the single-cell segmentation. This step is critical because it generates the single-cell masks that were later used for all the downstream analysis. The workflow starts with identifying the primary objecting or the nuclei using the "IdentifyPrimaryObjects" module and setting the diameter limits to 10–16 pixels. Then, the cellular membranes were identified using the "IdentifySecondaryObjects" module by expanding the primary objects by two pixels. Then, both the primary and the secondary objects were used to generate the single-cell masks that were later exported as tiff images.

FlowJo gating and phenotype identification
The single-channel images along with their single-cell masks were converted to a single .fcs file format following a previously published method (19). It starts by converting the dataset into single-cell protein expression data using the single-cell masks. Then, the dataset was transformed using ArcSinh to enhance the signal over the background and to convert the IMC data to a linear space. Further, both the x and y locations were used as variables as well to be later added to the .fcs files. The tissues were initially gated in a similar strategy as the flow cytometry data using the FlowJo software (v10.8.1). The entire cell population in the dataset was first gated based on their expression of CD3 and CD20. T cells as CD3$^+$CD20$^-$, B cells as CD3$^-$CD20$^+$, and other phenotypes as CD3$^-$CD20. T cells were then further divided into cytotoxic T cells based on the expression of CD8α and helper T cells based on the expression of CD4. Cytotoxic T cells were then further divided into exhausted T cells based on the expression of PD-1 and

follicular cytotoxic T cells as CXCR5$^+$. On the other hand, helper T cells were divided into follicular helper T cells as CXCR5$^+$CD27$^+$PD-1$^+$ and regulatory T cells as FoxP3$^+$.

Further, B cells were divided into GC B cells as CD27$^+$CD38$^+$ and nongerminal center B cells as CD27$^+$CD38$^-$. The GC B cells were divided into LZ B cells as CXCR5$^+$ and DZ B cells as CXCR5$^-$Ki67$^+$. The other cell population was divided into plasma cells as CD27$^+$CD38$^+$, and SCs as being either single positive by vimentin and collagen or double positive. Finally, the other cell population was divided into myeloid lineage immune cells as CD11b$^+$CD11c$^+$ and FDCs as CD11c$^+$CD21$^+$.

Semi-supervised iterative gating

An iterative semi-supervised clustering method was used to determine cell types. First, a single-cell unsupervised clustering was performed using the Leiden algorithm, a graph-based community detection algorithm. From each segmented cell region, the mean intensity of each marker expression was calculated. The resulting feature matrix consisted of n rows of the total number of cells and p columns of marker expression. Each column of the feature matrix was z-score normalized, and batch correction between samples was performed using the Scanorama pipeline. The neighborhood graph in the embedding space was constructed and used for unsupervised community detection. First, a set of study-wide target phenotypes is defined by a combination of markers from the experiment panel. This step is performed with a .csv file in a table format with the expected positive marker expressed in each defined cell type as outlined in Fig. S7A. The phenotypes recapture the tree-like information provided in Fig. S7B with each cell type derived from a parent class.

Each cluster from the Leiden clustering step is matched with the user-defined table using their cluster median expression to automatically give each cluster an annotation based on the table. Each annotated cluster can be assigned to a unique annotation or multiple annotations based on the combination of markers. To assign the positivity of the markers, we first perform a Gaussian mixture model (GMM) model to find a desirable threshold for each marker automatically. After finding the desirable threshold, we rescale the expression level between 0 and 1 with 0.5 the threshold found using the GMM. The annotation is performed by matching the user-defined markers from each phenotype with the Leiden cluster median expression level. To account for clusters without enough separation power, we allow multiple phenotype annotations. Moreover, we also excluded all individual markers that are not part of the Leiden clusters, and they are not considered in the phenotype-defining markers. All clusters that correspond to one annotation are considered final whereas the other phenotypes are run through the same process to separate their annotations that contain multiple phenotypes. Finally, we run the same pipeline at the single-cell level when the number of cells considered is low. This is due to considering Leiden clustering not capturing the whole distribution of each marker and leading to biased data. On the other hand, we do not run this directly with marker expression from each cell from the beginning to reduce the batch effect visualized in the manual gating pipeline.

To check for associations amongst phenotypes generated from both the manual gating and the semi-supervised gating, a cross-tabulation of phenotypes was extracted. The two factors used for the cross-tabulation were the phenotypes generated from the manual gating, and the clusters generated from the semi-supervised gating across all single cells. The maximum overlap was calculated as the maximum cross-tabulation score or highest count of matching cell types between the two labeling approaches.

Spatial distance between cell types

To quantify the spatial distance between cell types, for each cell we extract the minimum distance to each cell present in the cell type subset. That is for each cell type $A_1, \ldots, A_n$ present in our data, we extract for cell i the distance $d_1, \ldots, d_n$ representing the minimum distance of cell i to each cell in the cell type $A_1, \ldots, A_n$. Then for each imaging region, we aggregate the mean minimum distance by incorporating each cell type therefore creating a n by n matrix representing the average minimum distance of cell type i to cell type j.

Phenotype neighborhood interaction

To quantify the phenotype neighborhood interaction between cell types, we used a permutation-based approach to consider the difference in frequency of each cell type. For each cell in the dataset, we extracted its local neighbor's phenotypes by looking at a radius of 20 μm from the cell centroid in its spatial domain. A cell–cell neighborhood matrix is generated by looking at all the cells in the imaging region. Then, we performed random permutation (1,000 times) while maintaining the frequency of cell phenotypes in the region and extracted 1,000 randomly permutated cell–cell neighborhood matrices. Finally, we compared the real neighborhood matrix with the 1,000 randomly permutated neighborhood matrices by calculating $z_{ij} = (n_{ij} - \mu_{ij})/\sigma_{ij}$, with n_{ij} the number of neighbors of cell type j found in the proximity of the cell type i from the real neighborhood matrix, μ_{ij} and σ_{ij} the mean number and standard deviation of the number of neighbors of cell type j found in the proximity of the cell type i from the randomly permutated matrices. The corresponding P-value is calculated using the following formula: $p_{ij} = \mathrm{erfc}(z_{ij}/\sqrt{2})$ with erfc the complementary cumulative distribution function.

Collagen zonation analysis

We defined collagen zones to quantify the variation of cell type frequency as distance to the collagen-positive regions in the image. First, we define the collagen-positive region in the image by applying Gaussian blur and thresholding from the collagen image. From the collagen-positive region, we expand the mask incrementally by 200 μm (user-defined). For each zone, we extract the frequency of each cell type present in our dataset.

Follicles major/minor axis using shape and collagen

For each segmented follicle, we define the major axis in two ways:

- Shape-based: The major axis is defined as the angle major axis of the ellipse that has the same second moments as the segmented follicle.
- Collagen-based: The collagen marker image is transformed into patches of size 10 μm by taking the average in each patch. Thresholding is applied to extract collagen-positive patches. For each segmented follicle, a radius circle of 1,000 μm radius (user-defined) is searched and binned into 40 (user-defined) sectors. For each sector, we count the number of collagen-positive patches. The major axis is then defined as the average angle around the radius circle weighted by the number of collagen-positive patches in each sector.

The minor axis is defined as 90° clockwise from the major axis. For each follicle, we calculated the minimum angle difference when calculating clockwise and counter-clockwise between the major axis defined by collagen-based and shape-based.

Single cells inside a segmented follicle are projected onto the major and minor axis based on their spatial location with the origin defined by the centroid of the follicle. By binarizing along the major and minor axis, we extracted the single-cell variation and cell type variation across the major and minor axis for each follicle.

Acknowledgments

Supplementary Material

Supplementary material is available at *PNAS Nexus* online.

Funding

Author Contributions

Conceptualization: A.F.C., A.S., M.A., and T.H.; Methodology: A.F.C., A.S., M.A., T.H., Z.F., and M.P.; Investigation: A.F.C., A.S., M.A., and T.H.; Visualization: A.F.C., A.S., M.A., and T.H.; Supervision: A.F.C. and A.S.; Writing—original draft: M.A. and T.H.; Writing—review and editing: M.A., T.H., A.F.C., and A.S.

Data Availability

The data can be found at https://zenodo.org/record/8303732, and all accompanying data analysis pipelines can be found at https://github.com/coskunlab/SpatialVizPheno.

References

1   Ruddle NH, Akirav EM. 1950. 2009. Secondary lymphoid organs: responding to genetic and environmental cues in ontogeny and the immune response. *J Immunol Baltim Md*. 183(4):2205–2212.

2   Matsuno K, *et al.* 2010. The microstructure of secondary lymphoid organs that support immune cell trafficking. *Arch Histol Cytol*. 73(1):1–21.

3   Prados A, *et al.* 2021. Fibroblastic reticular cell lineage convergence in Peyer's patches governs intestinal immunity. *Nat Immunol*. 22(4):510–519.

4   De Martin A, *et al.* 2023. PI16$^+$ reticular cells in human palatine tonsils govern T cell activity in distinct subepithelial niches. *Nat Immunol*. 24(7):1138–1148.

5   Stebegg M, *et al.* 2018. Regulation of the germinal center response. *Front Immunol*. 9:2469.

6   Stoler-Barak L, *et al.* 2019. B cell dissemination patterns during the germinal center reaction revealed by whole-organ imaging. *J Exp Med*. 216(11):2515–2530.

7   Victora GD, Nussenzweig MC. 2012. Germinal centers. *Annu Rev Immunol*. 30(1):429–457.

8   Singh A. 2021. Eliciting B cell immunity against infectious diseases using nanovaccines. *Nat Nanotechnol*. 16(1):16–24.

9   Biram A, *et al.* 2019. BCR affinity differentially regulates colonization of the subepithelial dome and infiltration into germinal centers within Peyer's patches. *Nat Immunol*. 20(4):482–492.

10 Sanz I, *et al.* 2019. Challenges and opportunities for consistent classification of human B cell and plasma cell populations. *Front Immunol*. 10:2458.

11 Tan WCC, *et al.* 2020. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. *Cancer Commun*. 40(4):135–153.

12 Allam M, Cai S, Coskun AF. 2020. Multiplex bioimaging of single-cell spatial profiles for precision cancer diagnostics and therapeutics. *NPJ Precis Oncol*. 4:11.

13 Kim EN, *et al.* 2023. Dual-modality imaging of immunofluorescence and imaging mass cytometry for whole-slide imaging and accurate segmentation. *Cell Rep Methods*. 3(10):100595.

14 Bandura DR, *et al.* 2009. Mass cytometry: technique for real time single cell multitarget immunoassay based on inductively coupled plasma time-of-flight mass spectrometry. *Anal Chem*. 81(16):6813–6822.

15 Bendall SC, *et al.* 2011. Single-cell mass cytometry of differential immune and drug responses across a human hematopoietic continuum. *Science*. 332(6030):687–696.

16 Zhao Y, *et al.* 2018. Spatiotemporal segregation of human marginal zone and memory B cell populations in lymphoid tissue. *Nat Commun*. 9(1):3857.

17 Allam M, *et al.* 2021. Spatially visualized single-cell pathology of highly multiplexed protein profiles in health and disease. *Commun Biol*. 4(1):632.

18 Han G, Spitzer MH, Bendall SC, Fantl WJ, Nolan GP. 2018. Metal-isotope-tagged monoclonal antibodies for high-dimensional mass cytometry. *Nat Protoc*. 13(10):2121–2148.

19 Ashhurst TM, *et al.* 2022. Integration, exploration, and analysis of high-dimensional single-cell cytometry data using Spectre. *Cytometry A*. 101(3):237–253.

20 Gerner MY, Kastenmuller W, Ifrim I, Kabat J, Germain RN. 2012. Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes. *Immunity*. 37(2):364–376.

21 Berg S, *et al.* 2019. Ilastik: interactive machine learning for (bio) image analysis. *Nat Methods*. 16(12):1226–1232.

22 Carpenter AE, *et al.* 2006. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol*. 7(10): R100.

23 Kato LM, Kawamoto S, Maruya M, Fagarasan S. 2014. Gut TFH and IgA: key players for regulation of bacterial communities and immune homeostasis. *Immunol Cell Biol*. 92(1):49–56.

24 Zhang Y, *et al*. 2018. Plasma cell output from germinal centers is regulated by signals from Tfh and stromal cells. *J Exp Med*. 215(4): 1227–1243.

25 Zhang Y, Toellner KM. 2022. Germinal center derived B cell memory without T cells. *J Exp Med*. 219(3):e20220012.

26 Fink K. 2012. Origin and function of circulating plasmablasts during acute viral infections. *Front Immunol*. 3:78.

27 El Shikh MEM, Pitzalis C. 2012. Follicular dendritic cells in health and disease. *Front Immunol*. 3:292.

28 Good-Jacobson KL, Shlomchik MJ. 2010. Plasticity and heterogeneity in the generation of memory B cells and long-lived plasma cells: the influence of germinal center interactions and dynamics. *J Immunol*. 185(6):3117–3125.

29 Vaccari M, Franchini G. 2018. T cell subsets in the germinal center: lessons from the macaque model. *Front Immunol*. 9:348.

30 Jacobse J, Li J, Rings EHHM, Samsom JN, Goettel JA. 2021. Intestinal regulatory T cells as specialized tissue-restricted immune cells in intestinal immune homeostasis and disease. *Front Immunol*. 12:716499.

31 Roth K, *et al*. 2014. Tracking plasma cell differentiation and survival. *Cytom Part J Int Soc Anal Cytol*. 85(1):15–24.

32 Cinti I, Denton AE. 2021. Lymphoid stromal cells—more than just a highway to humoral immunity. *Oxf Open Immunol*. 2(1):iqab011.

33 Akkaya M, Kwak K, Pierce SK. 2020. B cell memory: building two walls of protection against pathogens. *Nat Rev Immunol*. 20(4): 229–238.

34 Allie SR, Randall TD. 2020. Resident memory B cells. *Viral Immunol*. 33(4):282–293.

35 Lee CM, Oh JE. 2022. Resident memory B cells in barrier tissues. *Front Immunol*. 13:953088.

36 Cosovanu C, Neumann C. 2020. The many functions of foxp3[+] regulatory T cells in the intestine. *Front Immunol*. 11:600973.

37 Zhang X, *et al*. 2013. Follicular helper T cells: new insights into mechanisms of autoimmune diseases. *Ochsner J*. 13(1):131–139.