Follicular Lymphoma Transformation is Characterized by Cytokine-associated Remodeling of Stromal and Macrophage Compartments

Authors:

Nicholas J Haradhvala[1]*, Stephanie Pei Tung Yiu[1,2]*, Laura Beckmann[1,3,4]*, Sam Sadigh[3,5], Zachary Leibowitz[4], Erik Landolsi[1,6], Santiago Rivero[4], Stephanie L. Deng[4], Sabrina Hu[4], Mingzeng Zhang[4], Daniel Filip[7,8], Amber Pospistle[1], Yao Yu Yeo[2,9], Katie Maurer[1,3,4], Johan Gustafsson[1,10,11], Kai Stewart[1], Vignesh Shanmugam[5], McKayla Van Orden[12], Wesley S. Lu[12], Shuqiang Li[1,12], Kenneth J. Livak[4,12], Huaying Qiu[2], Ankit Basak[1,13,14,15,16], Alex K. Shalek[1,13,14,15,16], Brian P. Danysh[1], Mikaela McDonough[4], Satyen H. Gohil[17], Yue Ren[18], Yinglu Zhou[18], Donna S. Neuberg[18], Svitlana Tyekucheva[18], Scott J. Rodig[3,5,19], Eric D. Jacobsen[3,4], Marek Mraz[7,8], Mark A. Murakami[3,4], Habibe Kurt[3,5], Catherine J. Wu[1,3,4], Gad Getz[1, 20,21]**, Sizun Jiang[1,2,3,5,9,22]**, Erin M. Parry[1,3,4]**#

* The authors contributed equally and share first authorship.

** The authors contributed equally and share last authorship.

corresponding author

Affiliations

[1]Broad Institute of Harvard and MIT, Cambridge, MA, United States.

[2]Center for Virology and Vaccine Research, Beth Israel Deaconess Medical Center, Boston, MA, United States

[3]Harvard Medical School, Boston, MA

[4]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA

[5] Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States.

[6]Chalmers University of Technology, Gothenburg (Göteborg), Sweden.

[7]Central European Institute of Technology, Masaryk University, Brno, Czech Rep.

[8]Department of Internal Medicine, Hematology and Oncology, University Hospital Brno and Faculty of Medicine, Masaryk University, Brno, Czech Rep

[9]Program in Virology, Division of Medical Sciences, Harvard Medical School, Boston, MA

[10]TIMM Laboratory at Sahlgrenska Center for Cancer Research, University of Gothenburg, Gothenburg, Sweden

[11]Department of Medical Biochemistry and Cell Biology, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

[12]Translational Immunogenomics Laboratory, Dana-Farber Cancer Institute, Boston, MA

[13]Department of Chemistry, Massachusetts Institute of Technology; Cambridge, MA, USA.

[14]Ragon Institute of MGH, MIT, and Harvard; Cambridge, MA, USA.

[15]Institute for Medical Engineering and Science, MIT; Cambridge, MA, USA.

[16]Koch Institute for Integrative Cancer Research, MIT; Cambridge, MA, USA.

[17]University College London Hospitals NHS Foundation Trust, London, United Kingdom

[18]Department of Data Science, Dana-Farber Cancer Institute, Boston, MA

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

[19]Center for Immuno-Oncology, Dana-Farber Cancer Institute, Boston, MA
[20]Krantz Family Center for Cancer Research and Dept. of Pathology, Massachusetts General Hospital, Boston, MA 02115, USA
[21]Department of Pathology, Harvard Medical School, Boston, MA 02115
[22]Department of Pathology, Dana-Farber Cancer Institute, Boston, MA, United States.

Running title: Cytokine-linked TME populations define FL transformation

Corresponding author:
Erin M Parry
Dana-Farber Cancer Institute
450 Brookline Avenue
Boston, MA, 02215
erinm_parry@dfci.harvard.edu

Conflicts of interest

C.J.W: Research Funding from Pharmacyclics, holds equity in BioNTech, Inc.. C.J.W is on the SAB of Repertoire, Adventris, Nature's Toolbox, Aethon Therapeutics

G.G. receives research funds from IBM, Pharmacyclics/Abbvie, Bayer, Genentech, Calico, Ultima Genomics, Inocras, Google, Kite, and Novartis and is also an inventor on patent applications filed by the Broad Institute related to MSMuTect, MSMutSig, POLYSOLVER, SignatureAnalyzer-GPU, MSEye, MinimuMM-seq, and DLBclass. He is a founder, consultant, and holds privately held equity in Scorpion Therapeutics; he is also a founder of, and holds privately held equity in, Predicta Biosciences. He was also a consultant to Merck.

S.J.R.: receives research funding from Bristol-Myers Squibb, KITE/Gilead, and Coherus. S.J.R is a member of the SAB of Immunitas Therapeutics.

K.J.L. holds equity in Standard BioTools, inc. and is on the scientific advisory board for MBQ Pharma, Inc.

D.S.N. has stock ownership in Madrigal Pharmaceuticals.

S.J.: Research Funding from Roche, Novartis. Founder and consultant for Elucidate Bio, Inc, and serves on its Board of Directors. Consultant for Quanterix Inc.

A.K.S.: consulting and/or SAB membership from Honeycomb Biotechnologies, Cellarity, Bio-Rad Laboratories, Fog Pharma, Passkey Therapeutics, Ochre Bio, Danaher, Janssen, Quotient Therapeutics, Relation Therapeutics, IntrECate biotherapeutics, and Dahlia Biosciences unrelated to this work; research funding from IBM, Wellcome Leap, Break Through Cancer, the Coca Cola Company, the Gates Foundation, the Chan Zuckerberg Initiative, the Gates Medical Research Institute, and St. Jude's Children's Hospital, as well as in-kind support from Becton, Dickinson, and Company (BD).

M.A.M. received research funding from Roche/Genentech, Genmab, BeiGene, and Kite/Gilead and is an advisory board member of CancerModels.org.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

N.J.H, S.P.T.Y., L.B., S.S., Z.L. E.L., S.R., S.L.D., S.H., M.Z., D.F., Y.Y.Y., A.P., K.M., J.G., K.S., V.S., M.O., A.B., M.MD W.S.L., H.Q., S.L., B.P.D., S.G., Y.R., Y.Z., S.T., E.D.J., M.M., H.K., E.M.P.: None

Additional Information:

Word count: 13839 (Main Text: 6062, Methods: 7777)

Number of Figures: 7

Number of Supplementary Figures: 9

Number of Supplementary Tables: 11

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

ABSTRACT

Across cancer, one of the most frequent examples of histologic transformation is the evolution of follicular lymphoma (FL) to an aggressive large cell lymphoma. Despite recent progress, understanding of the molecular and cellular underpinnings of transformation remains incomplete. Here, we dissect the interplay of tumor and microenvironment cell populations across transformation through a multimodal investigation of 95 FL and transformed FL (tFL) samples, including single-cell and bulk RNA-sequencing alongside spatial transcriptomics and proteomics, and validate findings across independent FL-tFL pairs. Upon transformation, fibroblasts and *GPNMB*+ macrophages increase while lymph-node organizing follicular dendritic and *CCL21*+ fibroblastic reticular cells were lost, resulting in an altered spatial distribution of cytokines that impacts T cell infiltration and macrophage differentiation and function. Secreted stromal and macrophage signals were further evident by non-invasive plasma proteomics. Taken together, our data reveal expansion of macrophages and fibroblasts as key features of transformation with potential diagnostic and therapeutic implications.

Statement of Significance:

Single-cell and spatial profiling of indolent and transformed follicular lymphoma (tFL) samples reveals transformation-associated microenvironment alterations, including increased *GPNMB*+ macrophages and fibroblasts, and decreased follicular dendritic and *CCL21*+ fibroblastic reticular cells. The altered chemokine landscape of tFL is detectable through plasma proteomics, affording potential diagnostic and therapeutic opportunities.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Introduction

Histologic transformation (HT) is a well-recognized phenomenon of accelerated cancer progression observed across diverse malignancies. Transformation of indolent follicular lymphoma (FL) to an aggressive large B cell lymphoma (transformed FL [tFL]) is one of the most common occurrences of malignant HT, occurring in 10-25% of patients with FL, and a leading cause of disease related mortality (1–4). Despite tFL being defined as the loss of follicular architecture with a diffuse growth of predominantly large cells on tissue biopsy (5,6), the altered cell populations and mechanisms that govern FL transformation and its associated nodal reorganization are incompletely understood.

Cell-intrinsic pathways have been identified to contribute to the evolution of HT from an ancestral FL precursor clone, including through acquisition of genetic alterations (7–9), and processes that increase memory-like B cell FL phenotypes (10,11). Recent work demonstrates that the tFL histologic phenotype is not entirely explained by genetic evolution, with similar genomic and transcriptomic profiles observed in proximally sampled FL and tFL cells (12), implicating a role of the microenvironment in facilitating transformation onset. While increased T cell exhaustion has been identified in tFL, little is known about the myeloid and stromal compartments during transformation evolution, as existing single-cell datasets in tFL largely fail to capture these populations, are restricted to few cases or markers, and lack spatial context (12–16).

To identify the microenvironment populations and signaling pathways altered upon HT of FL to tFL, we performed an integrated analysis of single-cell and bulk transcriptomic data and demonstrated findings in an expanded cohort, including biopsies with composite FL and tFL histologies, using spatial transcriptomics and proteomics. Independent collections of paired FL-tFL samples further confirmed key tFL-altered cell populations. During the process of transformation, our analysis identified replacement of nodal-organizing follicular dendritic cells (FDCs) and cytokine-producing fibroblastic reticular cells (FRCs) with a new fibroblast phenotype marked by matrix remodeling genes alongside expansion of *GPNMB*+ macrophages. Furthermore, we highlight co-occurrence and communication between these tFL-emergent cell types and demonstrate that markers of this microenvironment evolution are detectable in peripheral blood through non-invasive means.

Results

Transcriptome analysis of FL and tFL identifies immune cell types across disease stages

To investigate the cellular composition of the lymph node in FL and tFL, we assembled independent single-cell (n=10) and bulk (n=19) transcriptomic cohorts of unpaired FL and tFL samples for cellular population identification and comparative analysis (**Fig. 1A- top left**). We corroborated and localized the identified altered cell populations using spatial transcriptomics and proteomics across two tissue microarrays (TMAs), with transcriptomic capture obtained for 52 unique pathology biopsies, of which 18 provided 9 FL/tFL sample pairs derived from the same patient, and proteomic assessment by CODEX acquired for 59 biopsies, of which 24 provided 12 FL/tFL pairs (**Supplementary Table 1A-C**). Notably, we had available four diagnostic biopsy samples

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

containing distinct adjacent regions of FL and tFL histology, providing us the opportunity to identify the dynamic immune remodeling occurring at the transformation interface by application of high-resolution spatial transcriptomics, and, in 2 of 4 cases, spatial proteomics (**Fig. 1A - top right**). Lastly, findings were validated in two external transcriptomic cohorts (**Fig. 1A - bottom**) including 11 paired FL-tFL transcriptomes (17) and GeoMx tumor rich transcriptomic samplings of 20 FL cases, 4 of which had paired tFL samples.

To first identify FL to tFL nodal cell types, we focused on a dataset of single-cell RNA-sequencing (scRNA-seq) derived from disaggregated excisional biopsy samples from 10 patients with FL, of which 8 had a previous or subsequent diagnosis of tFL (**Supplementary Table 1A**). At the time of study biopsy, 6 patients displayed FL and 4 tFL histology (2 with diffuse large B cell lymphoma, not otherwise specified [DLBCL, NOS] histology and 2 with grade 3B FL following prior history of low grade FL). Among samplings, one biopsy (FL/tFL33) contained both indolent FL and tFL histology regions, while another (FL27) was derived from a patient with rapidly advancing disease diagnosed with tFL on repeat biopsy following progression on therapy (**Fig. 1B, Supplementary Table 1A**).

A total of 24,417 non-B cells and 47,345 B cell lineage cells were captured. Initial clustering of non-B cells identified 13 populations (**Fig. 1C**), including cell types often underrepresented in droplet-based scRNA-seq datasets, such as lymph node (LN) stromal cells and macrophages(10,12,16,18). Each patient's malignant B cells clustered separately from other patient's malignant B cells, while normal B cells, defined by polyclonal B cell receptor (BCR) sequences, and non-B cell types clustered together across patients (**Fig. 1D, Supplementary Fig. S1A-B**). Given the observed heterogeneity within immune cell populations, application of finer resolution clustering further divided non-malignant cells into a total of 22 clusters (**Fig. 1E, Supplementary Table 2A-C**). Plasmacytoid dendritic cell (pDC) and FDC populations were less frequent in tFL cases compared to FL (FDR-corrected rank-sum q=0.12), the latter confirming the expected loss of the FDC network characterizing HT (5,6) (**Fig. 1F**). The proximal-to-transformation sample of FL27 showed a composition similar to tFL, including a loss of FDCs along with expanded monocytes/macrophages, CD8+ cytotoxic T lymphocytes (CTLs), and FRCs/fibroblasts, suggesting that FL approaching clinical diagnosis of tFL may already acquire an altered microenvironment (**Supplementary Fig. S1C**).

Consistent with recent findings (12), tFL and the proximal-to-tFL samples showed expansion of varied CTLs **(Supplementary Fig. S1D)**; however, unlike those previous findings, we detected no significant increase in frequency of exhausted CTLs, nor higher expression of *PDCD1* or *LAG3* in CTLs at tFL (**Supplementary Fig. S1E**)**.** T cell receptor (TCR) repertoires were diverse, with the most abundant clones found at a frequency of less than 2 percent of all T cells, and generally contained cells corresponding to the dominant T cell cluster in the sample (**Supplementary Fig. S1F**). Among FL samples, the 3 samples with downregulation of MHC II (FL 6, 8 and 31) (18) showed a trend of reduced exhausted CTLs (CTL Ex) and increased *TGFβ1*+ CD8+ CTLs (**Supplementary Fig. S1D and G**). Malignant B cell clusters varied from normal B cells by patient-

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

specific gene expression patterns with no consistent FL versus tFL malignant cell changes detected (**Supplementary Fig. S1H**).

Increased macrophage and fibroblast gene-programs distinguish FL from tFL

To confirm microenvironmental differences between FL and tFL, and test whether these differences persist at lymphoma relapse, bulk transcriptome analysis was conducted on 19 (7 FL, 12 tFL histology) formalin-fixed paraffin embedded (FFPE) biopsies from patients that were referred for chimeric antigen receptor (CAR)-T cell therapy (**Supplementary Table 1A** for patient characteristics). The majority of these samples were collected at the time of relapsed/refractory disease (median of 2 lines of prior systemic therapy [range 0-6]), but 4 cases (3 tFL, 1 FL) only had initial diagnostic samples available. By principal component analysis (PCA), FL and tFL biopsy transcriptomes separated (Hotelling's T^2 p=0.0019) with tFL cases higher in PC1. The top positively associated genes for PC1 were related to stromal populations (*COL1A1, COL2A1, POSTN*) and macrophages (*CD163, FCGR3A, MMP2, CHIT1, APOC2, VCAN*), and negatively associated genes included the FDC marker *FDCSP* (**Fig. 2A, Supplementary Fig. S2A, Supplementary Table 3**). The three tFL samples lower in PC1 represented a case of admixed FL and tFL histology, an early post-allogeneic stem cell transplant relapse, and a biopsy containing muscle and soft tissue fragments. CIBERSORTx deconvolution of bulk transcriptomic data of the 7 FLs and 12 tFLs, using scRNA-seq–derived cell type signatures, inferred increased monocytes/macrophages (FDR-corrected rank-sum q=0.02) and FRCs/fibroblasts (q=0.02), as well as decreased T follicular helper (T_{FH}) cells (q=0.09), FDCs (q=0.01) and pDCs (q=0.18, **Fig. 2B**). Comparison to relapsed/refractory DLBCL-derived transcriptomes (6 germinal center B [GCB], 9 non-GCB cell type) showed similarity of tFL to DLBCL, with no differences (FDR-corrected rank-sum q>0.25) detected for any cell population (**Supplementary Fig. S2B)**.

Validation of immune and stromal composition in malignant B cell regions

Given the disruption of FL nodal structure associated with transformation, we assembled a TMA for spatial transcriptomic analysis from a collection of FL and tFL FFPE samples to localize the identified cellular populations within the tissue architecture (**Methods**). Contained within the captured TMA area were tissue samples derived from 52 unique lymphoid tissue biopsies (33 FL and 19 tFL) from 38 patients with an established history of tFL. These included paired FL and tFL biopsies (9 cases, 18 biopsies) as well as unpaired biopsies from 29 patients (20 FL, 9 tFL) (**Supplementary Table 1C, Supplementary Fig. S2C**). In addition, serial biopsies for 4 FL and 1 tFL cases and 17 multiple (7 duplicates and 1 triplicate) samplings were included to account for sample heterogeneity. The majority of the transformed cases showed histology consistent with DLBCL (n=12), with additional cases demonstrating either double hit lymphoma (DHL) (n=2) or grade 3B FL involvement (n=4) (**Supplementary Table 1A**).

Following pathology-guided selection of representative histology regions, antibody staining of PAX5 (B cells), CD3 (T cells), CD68 (monocytes/macrophages), and SYTO13 (nuclei) was performed to mark cell types of interest and create phenotype masks to guide pseudo-bulk cell type-specific transcriptome extraction (**Fig. 2C-D**). In total, 70 unique lymphoma and 2 tonsil control

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

regions of interest (ROI) were acquired using the GeoMx spatial transcriptomics platform across the TMAs (**Supplementary Fig. S2D, Supplementary Table 4**). Pathways of MYC, mTORC1 signaling, and oxidative phosphorylation were upregulated in the B cell regions of tFL compared to FL, consistent with prior studies of transformed lymphoma, including transformation of chronic lymphocytic leukemia (CLL) (19–21) and FL (12) (**Supplementary Fig. S2E, Supplementary Table 5**). Gene set enrichment analysis (GSEA) of the GeoMx B cell samplings did not reveal any particular B cell signature (10,22,23) to meet significance between FL and tFL, with near-significant enrichments linked to differences in tumor density and proliferation rather than distinguishing B cell phenotypic differences (**Supplementary Fig. S2F, Supplementary Table 5**). A higher cell cycle score was observed in FL biopsies obtained following a prior tFL diagnosis (p=0.001 random-intercepts model, **Supplementary Fig. S2G**).

Differentially expressed genes between tFL and FL in B cell-proximal regions were concordant with the prior bulk RNA-seq analysis (**Fig. 2E**). The cellular origin of the differentially expressed genes was found to correspond to several unique microenvironmental or malignant populations in the scRNA-seq data (**Fig. 2F**) including FDCs (*FDCSP*), FRCs/fibroblasts (*CCL21, TIMP1, FN1, COL3A1, POSTN*), macrophages/monocytes (*IFI30, CTSB, NAMPT, CSTB, APOE, CD14, GPNMB, APOC1, VEGFA, SOD2, LGALS3*), and T_{FH} and regulatory T cells (*SPOCK2, TCF7, PBXIP1, SEMA4D*) (**Fig. 2F, inset**); in addition, cell cycle and tumor-related cell cycle genes (*PCNA, CDK1*) were also observed.

Stromal cell evolution at transformation redefines cytokine-linked cellular neighborhoods

Given that FDC- and FRC-linked genes were among the top differentially expressed genes, we further investigated these populations using the B cell-proximal samplings from GeoMx across all unique ROIs (**Supplementary Fig. S3**). As expected, an FDC score derived from the scRNA-seq data was enriched in FL, lost upon tFL, and partially retained in grade 3B FL cases (**Fig. 3A, top**). A fibroblast score, derived from scRNA-seq cluster marker genes of *CCL21*- FRCs/fibroblasts, was higher in tFL than in FL (**Fig. 3A bottom panel**). The FDC-restricted chemokine *CXCL13* correlated with the inferred fraction of FDCs in biopsies and thus decreased at tFL. Despite the increase in the fibroblast score, T cell chemoattractants *CCL19* and *CCL21* were largely lost in tFL, demonstrating a clear phenotypic difference in the stromal cells at tFL (**Fig. 3B**).

To further characterize FRC/fibroblast phenotypes across histologies, sub-clustering of scRNA-seq LN stromal cells was performed, resulting in the identification of 7 different cell populations, including clusters defined by differential expression of chemokines (*CCL19, CCL21, CXCL12, CXCL9, CXCL10, CCL5*) (**Fig. 3C, Supplementary Fig. S4A-B**). While all FRC clusters expressed collagens **(Supplementary Fig. S4A-B***), CCL21*- clusters displayed a gene program resembling solid tumor cancer-associated fibroblasts, with expression of *CD34* and matrix remodeling or extracellular matrix proteins (*TIMP1, POSTN, DCN, VCAN, MGP)*, while *CCL21*+ FRCs co-expressed *CCL19* and *CLU* (**Fig. 3D-E**, **Supplementary Fig. S4A-B)**. In line with spatial data, both *CCL21*+ FRCs and FDCs were lost in transformed or near-transformed scRNA-seq samples (**Supplementary Fig. S4C**). Given that chemoattractant-linked stromal cell populations were

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

altered at transformation, we examined the expression of an expanded set of genes associated with cytokine activity (**Fig. 3F**, **Supplementary Fig. S4D**). At transformation, FRC- and FDC-derived chemokine signals were lost, and new secreted signals were seen, including macrophage-derived *SPP1* and *NAMPT*, fibroblast-derived *VEGFA* and *TIMP1*, and tumor-derived *MIF* **(Fig. 3F-G, Supplementary Fig. S4E**). We hypothesized that loss of nodal-organizing chemokines in tFL impacts infiltration of T cell populations expressing the corresponding surface ligands. Indeed, *CXCL13* strongly correlated with *CXCR5* (R^2=0.24, p<0.01), consistent with the loss of FDCs and T_{FH} cells at tFL observed in initial cohorts (**Fig. 3H - top, Supplementary Fig. S4F**). Similarly, *CCL21* expression showed correlation with *CCR7* in T cell regions (expressed by CD8+ T naive (Tn)/central memory (Cm), CD4+ Tn/CM, Treg and *CXCL8*+ pDCs, **Fig. 1E**) (R^2=0.25, p<0.01, **Fig. 3H - bottom, Supplementary Fig. S4F**), with both decreasing at tFL (limma log_2FC=-1.6, q=1.8x10^{-6} and log_2FC=-0.76, q=4.3x10^{-5} for *CCL21* and *CCR7*, respectively).

***GPNMB*+ macrophages are a hallmark of transformation**

Sub-clustering of the myeloid-derived cells from scRNA-seq identified three populations (**Fig. 4A**), including monocytes and two populations of *MRC1(CD206)*-expressing macrophages: (1) *GPNMB*+ and (2) *CXCL9*+ macrophages. *GPNMB*+ macrophages expressed genes associated with lipid metabolism (*APOE, APOC1*) and tissue remodeling (*MMP9, SPARC, SPP1*) compared to *CXCL9*+ macrophages, which co-expressed *CXCL10* **(Fig. 4B**). Across B cell and CD68-guided GeoMx spatial transcriptome capture, markers of *GPNMB*+ macrophages were increased in tFL compared to FL (B cell-rich region: *APOE*: log_2FC=0.78, q=9.72x10^{-5}; *GPNMB:* log_2FC=1.02, q=3.11x10^{-6}; SPP1: log_2FC=0.68, q=0.02; and CD68-rich region: *APOE*: log_2FC=0.42, q=0.04; *GPNMB:* log_2FC=0.72, q=0.002, SPP1: log_2FC=1.00 q=0.03), (**Fig. 4C**). These findings were confirmed in bulk transcriptomes (*GPNMB,* p<0.01; *APOE,* p=0.35; *APOC1,* p=0.11; *SPP1*, p=0.03) (**Fig. 4D**). Receptors for *SPP1*, including *CD44* and several integrins ($\alpha_v\beta_3$, $\alpha_4\beta_1$, $\alpha_9\beta_1$), were expressed by LN cell types with the strongest interaction scores between monocytes (via *CD44*) and pericytes (via $\alpha_4\beta_1$), and lower but detectable receptor expression observed in malignant B-cells (**Fig. 4E, Supplementary Fig. S4G)**. We next probed GeoMx samples for association of *SPP1* with ligands and markers of these cell types. We observed *SPP1* correlation with its ligands *ITGAV* and *ITGB6* (**Supplementary Fig. S4H**) as well as with fibroblast-associated marker genes *POSTN (*p<0.01*), COL1A1* (p=0.01), and pericyte-linked *MCAM* (p<0.01), suggesting co-localization of *GPNMB*+*SPP1*+ macrophages, fibroblasts, and pericytes within B cell regions (**Fig. 4F, Supplementary Fig. S4I**). For *CXCL9*+ macrophages, we investigated whether their abundance correlated with the CTL infiltrate, as our single-cell data previously identified *CXCR3*, which encodes the receptor for *CXCL9*, as most highly expressed in CTLs compared to other cell types. Indeed, macrophage-rich *CXCL9* expression correlated with T cell region-associated *CXCR3* expression in both FL and tFL histology (**Fig. 4G**). However, correlation of *CXCL9* with CTL-associated *GZMB* (R^2=0.27*,* p=0.02) as well as *PDCD1* (R^2=0.30, p<0.01) was more pronounced in tFL (**Fig. 4G)**, with *CXCL9* in FL instead correlating with *NKG7* (R^2=0.25*,* p<0.01) and *TGFB1* (R^2=0.21*,* p<0.01) (**Supplementary Fig. S4J)**. Thus, our data suggests that while *CXCL9*+ macrophages drive an inflammatory environment, *GPNMB*+ macrophages may serve an

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

immunosuppressive role via stromal communication, as has been seen in other malignancies and during tissue repair (24).

Fibroblast and *GPNMB*+ macrophage gene signatures increase at transformation in independent FL-tFL sample pairs.

To confirm our findings, we examined bulk transcriptomes derived from 11 FL-tFL sample pairs (17). In these cases, tFL followed an initial diagnosis of FL (range 8-116 months, median 52 months) and the majority displayed DLBCL NOS histology, although 3 were admixed with low-grade and/or grade 3B FL regions. As demonstrated above, gene programs associated with *GPNMB*+ macrophages and *CCL21*- fibroblasts increased at transformation (**Fig. 4H**). Furthermore, across sample pairs, normalized gene expression showed a consistent decrease in FDC/FRC-derived genes (e.g. *CXCL13,* q=0.04*; CCL21,* q=0.06*)* alongside an increase in fibroblast (e.g. *COL3A1,* q=0.05*; POSTN,* q=0.01*; FN1,* q=0.2; *TIMP1,* q=0.04) and macrophage (e.g. *GPNMB,* q<0.1) genes at tFL (**Supplementary Fig. S5A, Supplementary Table 6**).

In addition, we analyzed spatial transcriptome GeoMx samples from hematopathologist-identified tumor dense regions across 2 additional, independent FL-tFL TMAs. Applying our cell type scorings to 74 samplings (FL=61, tFL=13) derived from 20 FL patients, including 4 patients with paired FL-tFL samplings, we again found decreased FDC and increased *CCL21*- fibroblast and *GPNMB*+ macrophage signatures (**Fig. 4I,** p<0.01). Of note, the grade 3B FL case maintained evidence of expected preserved FDC meshwork, as previously observed.

Within this additional independent TMA dataset were 8 patients with high-risk FL that went on to have progression of disease within 24 months after initial chemoimmunotherapy (POD24). To determine whether tFL-expanded macrophages and fibroblasts varied by clinical history, association with early progression (i.e. POD24) and transformation were probed. In our initial GeoMx expansion cohort, these populations were detected at similar abundances in tFL biopsies regardless of chemotherapy history or timing of the transformation event (**Supplementary Fig. 5B**). Similarly, these signatures remained low in FL and were not enriched in baseline FL samples that went on to develop early FL progression (as defined by POD24) in a recently published FL dataset (25) (**Supplementary Fig. S5C**) or in the independent GeoMx sample cohort (**Supplementary Fig. S5D**). Thus, these signatures predominantly separated cases by histology and did not appear to be associated with high-risk FL characteristics.

To compare the tFL-enriched macrophage and stromal cell populations to a recently published large scRNA cohort of DLBCL, we scored our single-cell populations with marker genes from each reported DLBCL microenvironment cell type (26). *CCL21*- fibroblasts displayed strong overlap in gene expression with DLBCL cancer-associated fibroblasts and *POSTN* and *ABCA8*-expressing non-endothelial stromal cells (NESC) (**Supplementary Fig. S5E**), while tFL enriched *GPNMB*+ macrophages were most similar to DLBCL-identified tissue resident macrophages (TRM) and *SPP1* and *CHI3L1* macrophage populations (**Supplementary Fig. S5F**). Application of the LymphoMap (26) classifier across our bulk transcriptomes illustrated that FL samples largely

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

resided within the lymph node subtype marked by intact chemokine signaling. In contrast, tFL biopsies were enriched for the FMAC phenotype defined by expanded fibroblasts and macrophages, and displayed a similar microenvironment archetype distribution as non-transformed DLBCL (**Supplementary Fig. S5G**). Overall, this data validated expanded macrophage and fibroblast networks and loss of residual LN chemokine-signaling networks in the progression of FL to tFL, prompting further investigation into the spatial reorganization of these populations.

Reorganization of stromal and monocyte neighborhoods across the FL to tFL transition

To identify the tissue localization of tFL-enriched macrophage and stromal populations across the FL to tFL transition, Visium HD was applied to 4 cases of FL to tFL in which both indolent and aggressive histology were encompassed within an excisional biopsy (**Fig. 5A, Supplementary Fig. S6A, Methods**). Each ROI was placed to capture tFL histology with inclusion of adjacent low grade FL (3 cases) or a region with increased large B cells and fibrosis that did not yet meet diagnostic criteria for transformed FL (case FL137) (**Supplementary Fig. S6A**).

Differential expression analysis between adjacent FL and tFL regions identified genes related to *GPNMB*+ macrophage populations increased at transformation (*GPNMB, CHIT1, CHI3L1, APOD, C1QB*) in the 3 biopsies with clearly demarcated tFL histology, with genes reflecting fibroblasts increased in 2 of these 3 biopsies (*POSTN, CILP, DCN, FN1*) (**Fig. 5B, Supplementary Table 7**) Application of our cell type gene expression signatures derived from scRNA-seq revealed increases in *GPNMB*+ macrophages and fibroblasts alongside loss of lymph-node organizing FDCs and *CCL21*+ FRCs in association with transformation **(Fig. 5C-D)**. In FL137, the one case that did not display an increase in *GPNMB* expression or macrophage/monocyte gene signature, dense *GPNMB*+ macrophage populations were observed surrounding residual follicles in the unclassifiable adjacent tissue as well as the histology-confirmed region of tFL. To probe the co-localization of these populations, we performed clustering of individual 8x8 μm bins, and classified clusters into cell types based on marker genes (**Supplementary Fig. S6B**). Focusing on the biopsy with clear increase in macrophage and fibroblast signatures (FL14), neighborhood enrichment analysis identified co-localized *CCL21*- fibroblasts, *GPNMB*+ macrophages, and pericytes that shared closer proximity to each other than the tFL cells, visually forming islands within the transformed region (**Fig. 5E-F, Supplementary Fig. S6C**).

In biopsy FL33, we observed the gene *POSTN*, a marker of our *CCL21*- fibroblasts, to be among the most highly differentially upregulated genes (**Supplementary Fig. S6D**) despite lack of increased expression of other fibroblast genes. Examination of paired single-cell data confirmed low fibroblast abundance (**Fig. 1F**), and demonstrated that this sample had the unique quality of expressing *POSTN* in its tumor cells (**Supplementary Fig. S6D-E**). Thus, these data suggest that there may be convergent evolution of *POSTN* upregulation through means of different cell populations. Across our two admixed cases with clear demarcation of FL and tFL regions (FL14, FL33), we did observe patterns of memory (MEM)-like and GCB-signatures (10,22) aligning with the indolent follicular tissue structures (**Supplementary Fig. S6F**). In line with our previous results,

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

overall abundance of these signatures was not different between FL and tFL (**Supplementary Fig. S2F**), though the follicular GC/MEM organization observed in low grade FL was lost at tFL.

Macrophage differentiation and function is impacted by co-localizing fibroblasts.

Based on the colocalization of fibroblasts and *SPP1+GPNMB+* macrophages in tFL, we hypothesized that interactions between these cell types result in macrophage phenotypic evolution and impaired phagocytic function. First, we queried whether soluble signals may induce migration of peripheral blood–derived monocytes or T cells to tissue. Neither SPP1 nor POSTN attracted monocytes or T cells, while control chemokines CXCL12 and CCL21 induced migration (**Supplementary Fig. S6G**). Next, we tested whether fibroblast secreted signals impact macrophage phenotype or function by exposing CD206+ macrophages, differentiated from healthy donor peripheral blood monocytes, to conditioned media and/or soluble POSTN or SPP1 (**Fig. 5G**). SPP1 protein expression was higher in CD206+ macrophages exposed to conditioned media derived from two separate fibroblast cell lines, Hs895.T ($p=0.026$) and HS-5 ($p=0.031$), but not from conditioned media from germinal center (GC)-derived large cell lymphoma cell line or to soluble signals (POSTN, SPP1, CCL21) (**Fig. 5H, Supplementary Fig. S6H-I**). In addition, phagocytosis of tFL cell line DOHH2 was significantly impaired in macrophages pre-exposed to HS-5 conditioned media (**Fig. 5I, top,** adj.$p=0.007$, $n=4$), with a similar trend observed for the FL cell line SC1(**Fig. 5I, bottom**). Increased SPP1 levels and impaired phagocytic activity in macrophages exposed to soluble fibroblast-derived signals were validated using blood-derived autologous monocytes and primary FL cells (**Fig. 5J**). Together, this data provides evidence that tFL expanded fibroblasts impact local macrophage state and function, contributing to malignant B cell immune evasion.

Microenvironmental remodeling during FL transformation is marked by stromal and macrophage expansion

To validate our identified cell types and further define neighborhoods, CODEX was applied to adjacent sections of our FL-tFL TMAs, capturing a total of 80 biopsy samples. As before, duplicate and serial samplings were included to lend robustness and account for heterogeneity (**Supplementary Table 1C** for TMA details). A panel of 13 antibody markers (PAX5, CD21, CD8, CD3, CD4, CD68, CD11c, CD20, CD31, CD163, CD206, Podoplanin, FOXP3) enabled robust annotation for major cell types of interest including $CD68^+$ macrophages, $CD68^+CD163^+$macrophages, $CD11c^+$ cells (encompassing both $CD11c^+$ macrophages and DCs), and $PDPN^+$ lymphoid stromal cells ([LSC], representing fibroblasts). (**Fig. 6A, Supplementary Fig. S7A-B, Methods**). Providing independent protein-level validation of transcriptomic findings, analysis of cell type composition demonstrated an increase in the proportion of $PDPN^+$ LSCs (rank-sum permutation test, $q=0.007$), $CD68^+$ macrophages ($q=0.036$) and $CD11c^+$ cells $q<0.001$), along with a reduction in T cells ($CD3^+$, $CD4^+$ and regulatory T cell populations, $q<0.001$, $q=0.007$, $q=0.031$, respectively) in tFL compared to FL (**Fig. 6B, Supplementary Fig. S7C**).

CD68+ macrophages and PDPN+ LSCs spatially associate in tFL

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

To validate the spatial association between macrophages and fibroblasts identified in transcriptomic analyses (**Fig. 5E-F**), we performed spatial proximity analyses across TMA biopsies. We quantified the fraction of PDPN$^+$LSCs and tumor B cells located within 50 μm and 100 μm of CD68$^+$ or CD68$^+$CD163$^+$ macrophages on a per-core basis (**Fig. 6C**). The proximity enrichment fraction of PDPN$^+$ LSCs surrounding CD68$^+$macrophages was higher in tFL compared with indolent FL at both distances (rank-sum permutation test, **Methods**, 50 μm: q=0.057; 100 μm: q=0.076). In contrast, no enrichment was observed for tumor B cells (50 μm: q=0.259; 100 μm: q=0.259) (**Fig. 6D, Supplementary Fig. S7D**) or for PDPN$^+$ LSCs surrounding CD68$^+$CD163$^+$macrophages at either distance (**Fig. 6E, Supplementary Fig. S7E**). Collectively these findings demonstrate that increased macrophage-stromal spatial coupling in tFL is preferentially associated with CD68$^+$ macrophages and is not a general feature of CD68$^+$CD163$^+$ macrophage subsets.

Inflammatory *CXCL9*+ macrophages define T cell infiltrate

Identification of both *SPP1*+*GPNMB*+ and *CXCL9*+ macrophages within tFL biopsies prompted us to further investigate SPP1:CXCL9 macrophage polarity and association with T cell density, as described in solid tumors (27). Using macrophages as anchor cells, spatial graphs were constructed by connecting surrounding cells located within 20 μm (1-hop) and 40 μm (2-hop) of each anchor macrophage within 17 tFL cores with available data for GeoMx and CODEX and sufficient capture of CD68-guided transcriptional signatures (**Fig. 6F, left**), thus building local subgraphs capturing the immediate cellular neighborhood of each macrophage (**Fig. 6F, right**). While *SPP1*+*GPNMB*+ macrophages were largely restricted to tFL, *CXCL9*+ macrophages were seen across both histologies, with tFL displaying higher variability in *CXCL9* gene expression (**Supplementary Fig. 7F**). This allowed the stratification of each tFL ROI into either $CXCL9^{high}$ or $CXCL9^{low}$ categories. Comparison of macrophage-anchored neighborhoods revealed that, $CXCL9^{high}$ cores exhibited a higher overall CD4$^+$ T cell fraction compared to $CXCL9^{low}$ (Mann-Whitney U test, p=0.042) (**Fig. 6G**). Furthermore, a higher mean count of CD4$^+$ T cells was present within the immediate proximity of macrophages, as measured by 1-hop, 2-hop, and combined 1+2-hop neighborhoods (Mann-Whitney U test, p=0.071, p=0.042, p=0.042, respectively) (**Fig. 6H, Supplementary Fig. S7G**). To evaluate whether this spatial association was specific to *CXCL9*, we performed the same hop analysis stratifying tFL cores by *SPP1* expression (**Supplementary Fig. S7H**). In contrast, no enrichment in CD4$^+$ T cells was observed in macrophage-anchored neighborhoods of either $SPP1^{high}$ or $SPP1^{low}$ cores (**Fig. 6I**, **Supplementary Fig. S7I**). Together, these findings demonstrate that local recruitment or retention of CD4$^+$ T cells in macrophage neighborhoods within tFL is specifically associated with *CXCL9* expression.

FL in transition to tFL is marked by spatially defined inflammatory macrophage-T cell niches

We next focused on our two cases (FL14 and FL33) with FL and tFL histology regions clearly demarcated, enabling study of both histologies and the associated zones of histologic transition via single-cell whole-tissue imaging with CODEX **(Fig. 7A, Supplementary Fig. S8A-E**). To validate that PDPN$^+$LSCs are spatially associating with macrophages, we performed a macrophage-anchored cell type enrichment analysis on one of these representative cases (FL14). From the in-transition region, we sampled CD68$^+$ macrophage-centered neighborhoods at successive distances,

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

moving towards either indolent FL or tFL regions in 500 μm increments (**Fig. 7B**). Anchoring on $CD68^+$ macrophages and using a fixed 100 μm window, we observed a progressive enrichment of $PDPN^+$LSC accompanied by a decrease in $CD8^+$ T cell representation in regions farther from the in-transition region in tFL (**Fig. 7C**). This indicates that $CD68^+$ macrophage niches undergo a gradual transition from a T cell-enriched environment near in-transition regions towards a $PDPN^+$LSC-enriched microenvironment at tFL.

At the transitional histology region, both FL14 and FL33 displayed prominent $CD8^+$ T cell infiltrates that co-localized with macrophages along the tFL border (**Fig. 7D-E**). To quantitatively assess the spatial organization of immune and stromal cell populations across FL, transitional and tFL regions, we performed a spatial expansion analysis anchored on the in-transition zones. Regions were expanded outward in 500 μm increments toward both indolent and transformed areas to capture cellular composition across the full tissue landscape (**Fig. 7F**). In FL14, tFL regions demonstrated an expansion of $PDPN^+$LSCs, consistent with previously observed upregulation of fibroblast-associated gene programs (**Fig. 7F**). Across both cases, we observed a progressive increase in $CD68^+$ macrophages toward the transformed regions. Notably, $CD8^+$ T cell infiltration was most prominent at the tFL border, where $CD8^+$ T cells co-localized with $CD68^+CD163^+$ macrophages as well as $CD4^+$ T cells, and progressively declined deeper within the transformed tissue region (**Supplementary Fig. S8F, green transition zones**). Upon closer examination, early signs of increasing macrophage and $CD8^+$ T cells, accompanied by $CD4^+$ T cells (**Fig. 7E-F**) were also observed within the T cell reticular zone of the indolent regions approaching the tFL border.

Consistent with these observations, the spatial transcriptomic profiling on adjacent sections demonstrated that these T cell-enriched in transition regions corresponded to *CXCL9*^{high} macrophage-rich zones, supporting a *CXCL9*-associated T cell recruitment mechanism previously identified in the TMA dataset (**Fig. 6**). Providing orthogonal validation, probabilistic non-negative matrix factorization (PNMF) signature analysis (**Methods**) was applied to each of the 3 composite FL and tFL spatial transcriptomic cases (**Supplementary Fig. S8G-I, Supplementary Table 8**). Mapped signatures closely approximated single-cell proteomic results and validated tFL-altered populations. A histology border zone enriched in T/NK cell genes was identified in all cases, with co-localization with *CXCL9* seen in the 2 cases (FL14, FL33) analyzed by CODEX, supporting the emergence of spatially organized inflammatory macrophage-T cell niche at the transition interface. Additionally, all 3 cases were characterized by *CCL21*+ FRC-rich and T cell dense areas combining into a single signature in indolent histology that was lost at transformation as well as an increased *GPNMB*+ macrophage signature in tFL.

Evidence of altered immune remodeling is detectable non-invasively in plasma

Lastly, we queried whether these soluble signals could be detectable in blood and distinguish cases of tFL from FL. We examined a panel of 86 soluble proteins, largely associated with fibroblasts and macrophages, guided by our single-cell transcriptomic data (**Methods, Supplementary Table 9**), in 13 plasma samples (6 FL, 5 tFL-DLBCL, 2 tFL-grade 3B FL) from 11 patients. By PCA, the difference between FL and tFL was driven by PC1 (**Fig. 7G),** with tFL low in FDC- and FRC-

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

linked chemokines (**Supplementary Fig. S9A**). Differential expression analysis of plasma proteins confirmed key differences in chemokines from our transcriptomic studies in the plasma, with loss of CCL22 (p=0.033, Welch's t-test FL vs. tFL[DHL/DLBCL]), FDC-linked CXCL13 (p=0.0499), and FRC-derived CCL19 (p=0.0271), and CCL21 (p=0.145) in samples from patients with tFL (DHL/DLBCL) (**Fig. 7H**). Plasma levels of CCL19 and CCL21 correlated (r=0.71, p=0.0073, Pearson) and separated by histology (**Fig. 7I**). Furthermore, SPP1 levels correlated with detection of fibroblast-associated analytes (**Supplementary Fig. 9B**), supporting co-occurrence of these populations within tFL. Comparing the top analytes driving PC1 from this cohort to an independent dataset from untreated FL patients (n=18) and healthy donors (n=3) demonstrated that chemokines CCL19, CCL21, CXCL13 and CCL22 were also higher in indolent FL as compared to healthy controls (p<0.01) (**Fig. 7J**, **Supplementary Fig. S9C**). Together, this data suggests transformation-associated signals, including macrophage-fibroblast networks and loss of FL cytokine-linked architecture, are detectable in plasma, opening up a possibility for future efforts to detect tFL-associated signals non-invasively.

Discussion

A more complete understanding of the molecular signals that underlie FL transformation has the potential to enable advances in patient management, such as developing diagnostics to non-invasively detect transformation, informing improvements to transformation risk stratification, and identifying the most appropriate therapeutic interventions to prevent transformation or improve transformation outcomes. Prior work has identified tumor intrinsic features (7–9,28) and T cell exhaustion in FL transformation (12), yet gaps in mechanistic understanding remain (4). Here, we used a combination of single-cell RNA, bulk RNA, and spatial analysis approaches on distinct, largely independent cohorts to demonstrate striking shifts in frequency and phenotype of the stromal and macrophage compartments at transformation, which result in a loss of lymph-node organizing chemokines, formation of new cell interactions within the malignant lymph node, extracellular matrix remodeling, and altered T cell infiltration.

Despite evidence linking stromal populations to FL nodal polarization and even early natural progression of FL (29,30), the role of stromal cell types in histologic transformation remains obscured. This may be due in part to underrepresentation of stromal populations in single-cell or bulk transcriptomic datasets that rely upon CD45 selection or non-enzymatic dissociations. Here we demonstrate, across cohorts and methodologies, that FL transformation is distinctly marked by stromal remodeling. Specifically, tFL is not only defined by the loss of FDCs but also by the replacement of *CCL21*+ FRCs with *CCL21*- fibroblasts. This transition is associated with a sharp decline in *CXCR5* and *CCR7*-expressing T cell and pDC populations and a shift towards a macrophage-rich infiltrate. While the origin of the tFL fibroblast-like cells is not defined, these cell types are found in normal lymph nodes at low frequency, primarily lining the trabecular structures (31) **(Supplementary Fig. 9D**), which may expand at transformation. Alternatively, fibroblasts may arise through altered differentiation of *CCL21*+ FRCs or pericyte cells.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Regardless of origin, we found these expanded tFL fibroblast-like populations to co-localize with *GPNMB*+ macrophages and pericytes, and demonstrate these fibroblasts drive macrophage-linked *SPP1* expression and negatively impact macrophage phagocytic activity *in vitro*. CXCL9:SPP1 macrophage polarity has previously been linked to T cell infiltration, tertiary lymphoid structure formation, and immunotherapy responses across diverse types of solid tumors (27,32). In the present work, we demonstrate that macrophages similarly influence the density and localization of the T cell infiltrate in the malignant lymph node of a hematologic malignancy, with *CXCL9* rich macrophage residing in niches with $CD4^+$ T cells. Furthermore, our multimodal spatial data revealed that these regions rich in CTLs and inflammatory macrophages primarily reside on the periphery of the transformed tissue with limited penetration into deeper tFL regions. The dense tFL areas are instead intermixed with *GPNMB*+*SPP1*+ macrophages in close proximity to fibroblasts, with few T cells, perhaps limited by immunosuppressive signals, reduced chemoattractants, or impaired mobility due to extracellular matrix composition. As immunotherapies, such as CAR-T cells and bispecific engagers, increasingly become part of standard-of-care therapy for FL and tFL, macrophage polarity and distribution of the resulting CTL infiltrate may have major prognostic implications for disease response. Emerging studies suggest a critical microenvironment role for response to these therapies in DLBCL (26,33,34), and this is an important focus of ongoing research.

tFL remains a morphologic diagnosis made by expert hematopathologists; however, obtaining an appropriate tissue sample adequate for diagnosis may be challenging due to a number of clinical factors, such as difficult-to-sample lymphoma, rapid clinical progression, or limited biopsy samplings that may miss focal HT, resulting in risk of improper therapy selection (4). As such, the ability to detect transformation-associated signals in the blood would be a useful tool to improve identification of tFL, prompt expedited re-biopsy, monitor response to therapy, or distinguish FL versus tFL relapse. Here, we demonstrate that microenvironmental-linked soluble signals can be detectable in a non-invasive way using sensitive plasma proteomics and reflect nodal transcriptomic results. Given that the immune microenvironment changes are consistent across diverse cases of FL and tFL, this approach may provide potential advantage over, or added benefit to, cell-free DNA-based assays due to challenges posed by genetic heterogeneity, FL-tFL driver overlap, and DNA admixture. Our data does suggest this may be more difficult in the rare cases of grade 3B FL histology given retained FDC meshworks and *CCL21*+ FRCs. Therefore, this is a promising area of future research for study in larger cohorts.

Looking forward, an important focus of future studies will be to examine how malignant clonal evolution or other inciting events may trigger onset of macrophage and stromal changes. Our study did not fully elucidate how the microenvironmental changes impact B cell phenotypes in FL and tFL, potentially due to limitations in resolution of spatial platforms or interpatient heterogeneity. Thus, how the transformation process affects the malignant B cell phenotype and relates to transformation risk or prognosis should be investigated in future studies. Furthermore, a major remaining clinical challenge in the care of patients with FL is to identify and stratify patients into those who are at risk for transformation and those who are unlikely to experience transformation.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

In our study, cases were selected due to history of histologic transformation, providing us with a cohort to define transformation-specific changes. Across our initial and validation cohorts, these immune features did not clearly separate early adverse-risk FL progression, or POD24, status, however, this analysis was limited by small sample groups and should be further explored in larger cohorts, especially given potential challenges surrounding indolent FL heterogeneity. Thus, a future direction will be to examine whether early features of the tFL microenvironmental remodeling can be used to distinguish patients with low versus high risk of transformation, or predict the onset of subsequent HT. Our study suggests that macrophage expansion may be an early clue to impending risk of transformation, but this requires further investigation. Furthermore, additional studies in larger cohorts will be needed to investigate whether the stromal and immune populations linked to transformation are prognostic, particularly in the context of emerging immunotherapies.

METHODS

Human Samples and Ethics Statement

Follicular lymphoma samples, plasma, and PBMCs were obtained from clinical trial and research protocols at Dana-Farber Cancer Institute in accordance with the Declaration of Helsinki, with written informed consent obtained from all participants. All studies were reviewed and approved by the Dana-Farber Cancer Institute Review Board. For the independent external datasets, studies were reviewed and approved by the respective local review boards. For the 11 paired FL or tFL transcriptomes, the institutional review board (Ethics Committee of University Hospital, Brno, Czechia) approved the study, and samples were obtained with written, informed consent according to the Declaration of Helsinki (17). POD24 transcriptomes were obtained from NCBI Gene Expression Omnibus GSE272403 (25). The independent GeoMx study used for validation was approved by Lifespan IRB.

Lymph node biopsy specimens for single-cell sequencing were banked at DFCI on research protocol and processed into single-cell suspension according to standard operating protocol in the Pasquarello Tissue Bank through mechanical dissociation followed by enzymatic dissociation with collagenase in a Stomacher lab blender. Single-cell suspensions were cryopreserved with FBS and 10% DMSO before being stored in vapor-phase liquid nitrogen until the time of analysis.

All specimens were collected between 2005 and 2022. Histologic diagnosis of each FPPE sample was confirmed by secondary expert hematopathology review prior to study inclusion to ensure diagnoses were compatible with current pathology guidelines and standards. For this study, histologically low grade Follicular Lymphoma (FL) includes grades 1-2 and grade 3A (WHO 4), which is encompassed within classical FL (WHO 5(5)). Transformed Follicular Lymphoma (tFL) includes histologies of DLBCL NOS, Double/Triple Hit Lymphoma with MYC/BCL2 and/or BCL6, FL grade 3B (WHO 4, now recognized as Follicular large B cell lymphoma [FLBCL, WHO 5]) and large cell transformation of FL with unavailable FISH. All tFL cases had documented low grade FL.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Specimens were selected based on availability, given the rarity of the disease. Across cohorts I-III, the patient population was slightly skewed towards males with 43 out of 67 patients being male and 24 out of 67 patients being female.

Tissue microarray construction

Representative areas of appropriate histology were circled by an expert hematopathologist on hematoxylin and eosin (H&E) stained sections for targeted tissue microarray (TMA) core sampling. The TMA was constructed by S.S. and S.J.R. with IRB approval and includes 1-2 cores from each case, with each core measuring 1.5 mm in diameter. The final assembled TMA was again reviewed by an expert hematopathologist to ensure sampled regions indeed had appropriate histology for each case and to guide selection of representative appropriate capture of regions of interest (ROI) for GeoMx.

Sample processing for single-cell analysis

Frozen aliquots of cells were thawed into RPMI with 10% FBS. Samples were washed with media and a dead cell depletion (Miltenyi Biotec 130-090-101) was performed prior to staining of all viable cells with cell hashtag antibodies (TotalSeq-C, Biolegend). After washing, samples were pooled and stained with cell surface antibodies (TotalSeq-C, Biolegend, **Supplementary Table 10**) according to manufacturer's protocol prior to single-cell RNA-sequencing (10X genomics, 5' kit with TCR and BCR). A total of 40,000 cells were loaded onto eight 10x channels. About 20,000 cells per sample were loaded onto a 10× Genomics Chromium™ instrument (10× Genomics) according to the manufacturer's instructions. The scRNA-seq libraries were processed using Chromium Next GEM Single Cell 5' Kit v2, and the TCR/BCR libraries were done using Chromium single cell human TCR/BCR amplification kits, and the hashing libraries were processed using Chromium 5' feature barcode kit (10× Genomics). Quality controls for amplified cDNA libraries and final sequencing libraries were performed using Bioanalyzer High Sensitivity DNA Kit (Agilent). All the sequencing libraries were normalized to 4nM concentration and pooled. The pooled sequencing libraries were sequenced on Illumina NovaSeq S4 300 cycle platform. The sequencing parameters used were: Read 1 of 26bp, Read 2 of 90bp, Index 1 of 10bp and Index 2 of 10bp. Given lower initial relative cell numbers from FL35, this sample was reprocessed individually using similar methods and sequenced on a single lane without cell hashing to improve representation of cells from this sample. Samples were pooled and sequenced on a NovaSeqS4 (Illumina) as described.

Demultiplexing and sample analysis

Raw fastq files for all 10x runs were aligned and quantified using Cellranger version 5.0.0 and reference package GRCh38 3.0.0 for the gene expression libraries. TCR sequencing data for all runs was processed using cellranger vdj (v5.0.0) and the GRCh38_vdj_v5.0.0 reference package.

For all cell multiplexed samples, we ran souporcell v2020.7 to identify genotype clusters using the common SNP sites from the 1000 Genomes Project. To match genotype clusters across the eight different channels, we performed hierarchical clustering of all inferred genotypes using a Hamming

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

distance and single linkage calculated at all sites with covered loci. Genotype clusters with fewer than 100 cells or fewer than 1000 variants in a given pool were filtered due to lack of confidence. Clustering of genotypes into 10 clusters was observed by dendrogram as expected and used to group cells from the same individual across different 10x channels. Finally, the identity of each individual was labeled based on the average expression of the hash antibodies across all cells in a given individual.

Separation of B cells and microenvironment

An initial clustering of the cells was performed using scanpy (35) v1.9.8 and a standard single cell workflow. Doublets were filtered by running scrublet (36) v0.2.3 on each 10x channel individually for PBMC samples. Cells were filtered for a minimum of 200 genes and <15% mitochondrial RNA transcripts. Data was log-normalized as $\ln(1+10^4\, n_g/N)$, where n_g is the number of UMIs in gene g and N are the total number of UMIs for a given cell. Variable genes were selected using scanpy's highly_variable_genes process with a minimum dispersion of .2 and mean expression values between 0.005 and 3. Removed from clustering consideration were sex-biased genes XIST and RPS4Y1. The effects of total UMI counts and mitochondrial RNA percentages were regressed out and then the expression values with a maximum of 10 standardized. PCA was performed, then harmony used to integrate the pooled and non-pooled samples. Nearest-neighbors were computed using the top 40 harmony-adjusted PCs and 10 neighbors, and data was clustered using a resolution parameter of 2. To identify B cell clusters, the maximum expression of all immunoglobulin light and heavy chains for each cell was calculated and then averaged by cluster. Clusters with average expression >1 were labeled as B cells, with the exception of one cluster corresponding to plasmacytoid dendritic cells, which was identified based on *LILRA4* expression. B cell clusters were then separated from immune and stromal clusters for downstream analysis.

Immune and stromal analysis

To further define immune and stromal populations, clustering analysis subset on these populations was performed. Data normalization as before was repeated, and the top 3,014 highly variable genes were selected excluding immunoglobulin variable genes, *XIST*, and Y chromosome genes. Cell cycle scores for S and G2M phases were assigned using the score_genes_cell_cycle routine in scanpy using the cell cycle gene signature published in Tirosh et al (37). These scores were regressed out in addition to UMI counts and mitochondrial percentages. Harmony batch correction was again performed, and the top 20 PCs were used for subsequent clustering with the Leiden clustering algorithm. First, a coarse clustering was performed with resolution parameter 0.5 and broad cell type labels were assigned. Next, each coarse cell type was isolated and clustering at varying resolutions was performed, and cells were further subdivided into a finer categorization where a biologically meaningful signal could be ascertained. In several cases, clusters corresponding to doublets missed by the prior scrublet filtering were removed at this stage.

RNA-sequencing

RNA was extracted from FFPE biopsy samples using standard methods (Qiagen Allprep FFPE kit). RNA extracted from FFPE tissues was assessed for quality using the Agilent Bioanalyzer producing

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

a DV200 metric. Samples with >30% of fragments and >200 base pairs had the highest chance of yielding high quality sequencing data. 100ng of RNA was input for cDNA synthesis using Superscript III reverse transcriptase and then processed with Illumina's TruSeq RNA Access Sample Prep Kit. Synthesis of the second strand of cDNA was followed by indexed adapter ligation with unique molecular index adapters. Subsequent PCR amplification enriched for adapted fragments. Amplified libraries were quantified, normalized, pooled, and hybridized with exome targeting oligos. Following hybridization, a bead-based clean-up was performed, followed by an elution step. PCR was performed to prepare library pools for sequencing on Illumina HiSeq2500 sequencers utilizing sequencing-by-synthesis kits to produce 76bp paired reads. Output from Illumina software was processed by the Picard data-processing pipeline to yield CRAM or BAM files containing demultiplexed, aggregated aligned reads. Sequencing yielded approximately 50M paired reads. This transcriptome capture targeted 21,415 genes, representing 98.3% of the RefSeq exome.

RNA sequencing data analysis

Raw fastq files were aligned with STAR v2.5.3a to hg19 with a gencode v26 annotation. QC and gene-level quantification were performed using RNA-SeQC2 (38). Only samples with detection with >10k genes were included for analysis. Expression units normalized to transcripts per million were utilized for downstream analysis. To perform principal component analysis, data was filtered for coding genes, excluding the immunoglobulin variable loci, normalized using the DE-SeQC2 size factors, transformed with $\log_{10}(x+1)$, and genes were mean-centered. PCA was then performed using the sklearn package.

RNA deconvolution

Single-cell data was used to derive an FL-specific cell type signature matrix. For each cell type, differentially upregulated genes comparing that cell type to all other cells in the dataset were identified with the scanpy rank_genes_groups routine employing a t-test. Genes with $\log_2$-fold change greater than 1 were sorted by their p value, and the top 50 genes were taken for each cell type. A pseudobulk cell type x marker gene expression matrix was computed by summing the UMIs across cells of a given cell type and computing a transcripts-per-million normalization. This matrix was then passed to the "Create Signature Matrix" module of the CIBERSORTx (39) web portal (cibersortx.stanford.edu), and the resulting signature matrix was used for deconvolution of the bulk RNA sequencing matrix using the "Impute Cell Fractions" routine.

Data visualization

Swimmer plots were generated in RStudio (version 2021.09.1) using the packages swimplot (version 1.2.0) and ggplot2 (version 3.5.1). All boxplots presented here show the median, inter-quartile range (IQR), and minima/maxima of underlying samples. Overlaid dotplots show individual samples.

GeoMx Methods

FFPE tissue rehydration and antigen retrieval

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

FFPE tissue slides were baked in an oven (VWR, 10055-006) at 70°C for 1 hr and thoroughly deparaffinized by immersing in xylene for 2× 5 minutes. The slides were then rehydrated by a series of graded solutions using a linear stainer (Leica Biosystems, ST4020), with each step proceeding for 3 minutes: 3× xylene, 2× 100% ethanol (EtOH), 2× 95% EtOH, 1× 80% EtOH, 1× 70% EtOH, 3× UltraPure water (Invitrogen 10977-023), and finally left in UltraPure water (Invitrogen 10977-023). Unless otherwise specified, all washes mentioned below were with gentle agitation. Antigen retrieval was performed at 97°C for 20 minutes with pH 9 Target Retrieval Solution (Agilent, S236784-2) using a PT Module (ThermoFisher, A80400012), after which the slides were cooled to room temperature on the benchtop and washed once in 1× PBS for 5 minutes.

RNA probe hybridization, GeoMx antibody staining and ROI selection

The RNA probe staining cocktail was prepared according to the Nanostring RNA slide Prep kit (Nanostring, 121300313) using the Nanostring Human Whole Transcriptome Atlas detection probe set (Nanostring, 121401102). The RNA probe cocktail was applied to the tissue slides, covered with a hybridization cover slip (EMS Diasum, 70329-40), and incubated at 37ºC for 18 hrs. After hybridization, the tissue slides were first washed twice in Stringent Wash Buffer (mixing 4x SSC (Millipore Sigma, S6639) 1:1 with 100% formamide (Millipore Sigma, 344206-1L-M)) for 5 mins each at 37ºC, and twice with 2x SSC for 5 mins each at room temperature (RT) on an orbital shaker.

Tissues were then stained with antibodies against PAX5, CD3 and CD68 for 1 hr at RT, in dark, followed by SYTO13 (100nM) for 10 mins at RT, in dark. The slides were then washed twice in 2x SSC for 2 mins each at RT and scanned on the GeoMx for ROI selection. One 660x660 μm square ROI was drawn in each core, resulting in a total of 72 ROI across the two TMA. Fibrotic areas were excluded from ROI selection for each core. ROIs were selected to achieve the maximum possible coverage of both intra- and inter-follicular regions in cores with distinct follicular structures, while in cores lacking clear follicular structures, ROIs were chosen based on the most representative antibody staining patterns. The spatial coordinates of each ROI were then recorded and used to register the ROI segmentation mask.

Cell segmentation and single cell feature extraction

Cell segmentation was performed using the MESMER model of DeepCell version 0.12.2 (40,41), with maxima_threshold set to 0.075 and interior_threshold set to 0.05. The nuclear channel input of MESMER was SYTO13. The membrane channel input of MESMER was a summation of CD3 and CD68. For each marker, the pixel value within the area of each cell, determined by the segmentation mask, was summed and then divided by the area of each cell, and the resulting cell-size scaled sum was set as the expression value for a given marker. Each cell from each core was identified by unique cell label and tissue ID.

Cell phenotyping

Cell phenotyping was performed through iterative rounds of clustering and annotating processes with PhenoGraph based on the phenotyping marker enrichment profiles(42). Clustering results were plotted as heatmaps showing the Z-score of each marker within each cluster and were annotated based on the combination of markers that had the highest Z-score in that cluster. PAX5, CD3 and

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

CD68 were used to annotate the general cell types of B/Tumor, T and myeloid cells, respectively. Cells that showed unclear marker enrichment patterns were annotated as "Other" cells.

After an initial round of clustering with PhenoGraph was performed, clusters with clear enrichment patterns were annotated, while clusters with mixed patterns underwent additional rounds of iterative clustering and annotation until all identifiable cells were annotated. To visualize and confirm the assigned annotations, Mantis Viewer (Schiemann R et al., Zenodo 2020, available from: https://doi.org/10.5281/zenodo.4909620) was used to overlay the annotations onto the segmentation mask and the marker image, for visual inspection. A segmentation mask for each annotated cell population was iteratively generated, and these masks were re-uploaded to the GeoMx to guide cell-type specific genome-wide transcriptome extraction.

NGS library preparation

After transcriptome capture, unique molecular barcodes for the RNA probes were aspirated from each cell population to 96-well collection plates (Nanostring, 100473). Collection plates were then dried according to Nanostring instructions. Sequencing libraries were prepared by first resuspending each aspirate in 10 µL UltraPure water (Invitrogen 10977-023) and uniquely indexed with the Illumina i5×i7 dual indexing system as part of the Nanostring NGS library preparation kits (Nanostring, 121400201 & 121400202 & 121400203 & 121400204). PCR reactions were performed as follows: 37ºC for 30 mins, 50ºC for 10 mins, 95ºC for 3 mins, followed by 18 cycles of 95ºC for 15 s, 65ºC for 50 s, 68ºC for 30s, followed by a final extension of 68ºC for 5 mins and hold indefinitely at 12ºC. Next, 4 µL of PCR product from each well of the same plate was pooled and purified with AMPure XP beads (Beckman Coulter A63881) according to Nanostring instructions. Finally, QC was performed for the library with the Agilent TapeStation. Equal concentrations of each sub-library were pooled and 5% of PhiX sequencing control (Illumina FC-110-3002) was added to the library. The mixture was then subjected to paired-end sequencing with the NextSeq 2000 platform.

GeoMx analysis

Data processing

The demultiplexed FASTQ output files from next-generation sequencing were used to map and quantify the human probes through the GeoMx Data Analysis software pipeline (43). The .dcc files produced were then uploaded onto the GeoMx to generate gene counts tables using the default "QC" and "Biological probe QC" settings without filtering out any genes.

GeoMX probes were then quality controlled following established guidelines. The limit of quantification (LOQ) for each sample was determined as a two standard deviation fold increase over the geometric mean of signal from the negative control probes. Probes with <10% LOQ across all areas of interest (AOIs) were considered low quality. Data was then upper-quartile normalized by calculating the upper-quartile gene expression value for each AOI and then normalizing to the mean UQ value.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Differential expression analysis was performed using the limma-voom package. A linear model was fit using the formula "expression ~ 0 + SegmentLabel + SegmentLabel:disease" where SegmentLabel encodes the AOI region (B cell, T cell, Macrophage, other) and disease encodes FL or tFL. Repeat sampling from the same patient was accounted for using the DuplicateCorrelation function in the limma package. Differential expression for each cell type region was then performed by testing the contrast "SegmentLabel:tFL - SegmentLabel:FL" for each value of SegmentLabel. For the cytokine-specific analysis, genes were subset to those with annotated "cytokine activity" in the Gene Ontology database, and Benjamini-Hochberg FDR correction was re-applied.

Scoring scRNA-derived genesets

To score FDC and *CCL21*-negative fibroblast signatures, we established genesets using the single-cell data. Upregulated genes in each cell type were determined by using the rank_genes_groups routine pairwise against all other cell types, then taking the maximum p value across comparisons to identify the genes most specific to the particular cell type. The top 20 most significant genes were taken, then used to score cell type signatures in the GeoMX data using the score_genes routine in scanpy. This procedure normalizes the expression of genes sets to a control set with similar average expression levels to the test set, and thus corrects for any situations where a given gene may be more or less strongly detected in a given modality.

Analysis of cell-cell interactions

Potential receptors for differentially expressed cytokines were identified using the CellPhoneDB (44) database. Correlations in expression values between cytokine-receptor pairs were quantified by fitting a linear hierarchical random intercepts model using the cytokine as an explanatory variable and the patient and sample labels as a blocking variable. The strength of the association was determined as the marginal R^2 (45) and significance by a t-statistic compared to an intercept-only model.

CODEX methods

Antibody panel selection, conjugation, and titration

Antibodies used in the tFL/FL studies were obtained from their respective commercial sources and were conjugated either in-house or through collaboration with Cell Signaling Technologies (CST), both of which include previously validated clones (46–48). For antibodies conjugated in-house, the specificity of antibody candidates was first validated by immunohistochemistry (IHC) on FFPE lymphoid tissues to ensure robustness of staining. The selected antibody clones were then conjugated by maleimide-based chemistry with minor modifications from a previously published protocol (49). Briefly, 50 or 100 μg of carrier-free antibody was concentrated using a PBS-T pre-wetted 50kDa filter (Sigma Millipore, UFC5050BK) and then incubated with 0.9 μM TCEP (Sigma, C4706-10G) for 10-30 minutes in a 37°C water bath to reduce the thiol groups for conjugation. Reduction was quenched by two washes with Buffer C (1mM Tris pH 7.5, 1mM Tris pH 7.0, 150mM NaCl, 1mM EDTA) supplemented with 0.02% NaN3. Maleimide oligos were resuspended in Buffer C supplemented with 250mM NaCl. The reduced antibody was then

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

incubated with 100 or 200 µg (for 50 or 100 µg of antibody, respectively) of maleimide oligos (Biomers, 5'-Maleimide) in a 37°C water bath for 2 hrs. The resulting conjugated antibody was purified by washing five times with the 50kDa filter with high-salt PBS (1× DPBS, 0.9M NaCl, 0.02% NaN3). The conjugated antibody was quantified in IgG mode at A280 using a NanoDrop (Thermo Scientific, ND-2000). The final concentration was adjusted by adding >30% v/v Candor Antibody Stabilizer (FisherScientific, NC0414486) supplemented with 0.2% NaN3, and the conjugated antibody was stored at 4°C. Each conjugated antibody was titrated and validated via immunofluorescence on FFPE lymphoid tissues.

Tissue preparation, blocking and antibody staining

FFPE tissue slides were prepared according to the previous section "GeoMx methods: FFPE tissue rehydration and antigen retrieval." Tissue regions were circled with a hydrophobic barrier pen (Vector Laboratories, H-4000) and then washed in 1× TBS-T for 5 minutes. The slides were then equilibrated by washing in S2 Buffer (0.5× DPBS, 0.25% BSA, 0.02% NaN3, 250 mM NaCl, 61 mM Na2HPO4, 39 mM NaH2PO4, 2.5mM EDTA) for 20 minutes. Tissue slides were then subjected to blocking with Blocking Buffer that contains BBDG (5% normal donkey serum, 0.05% NaN3 in 1× TBS-T wash buffer (Sigma, 935B-09)) supplemented with 50 µg/mL mouse IgG (diluted from 1 mg/mL stock (Sigma, I5381-10mg) in S2), 50 µg/mL rat IgG (diluted from 1 mg/mL stock (Sigma, I4141-10mg) in S2), 500 µg/mL sheared salmon sperm DNA (ThermoFisher, AM9680) and 50 nM oligo block (diluted from stock with 500 nM of each oligo in 1× TE pH 8.0 (Invitrogen, AM9849)). Blocking was performed in a humidity chamber on ice during which the tissue was photobleached using Happy Lights (Verilux, VT22) for 1 hr, with temperature kept below 40°C.

Each conjugated antibody in the panel was first centrifuged at 12,500 rpm for 8 minutes at 4ºC to remove aggregates. They were then diluted in a single tube containing 50 µL Antibody Diluent (5% normal donkey serum, 0.05% NaN3 in 1x TBS-T) at concentration based on prior titration. The antibody mixture was then filtered through a S2 Buffer pre-wetted 50kDa filter (Sigma Millipore, UFC5050BK) at 12,500 rpm for 8 minutes at 4ºC. The filtered mixture was adjusted to the desired final volume by adding 25% v/v total volume with FFPE Block (50 µg/mL mouse IgG (Sigma, I5381-10mg), 50 µg/mL rat IgG (Sigma, I4141-10mg), 500 µg/mL sheared salmon sperm DNA (ThermoFisher, AM9680), 50 nM oligo block (diluted from stock with 500 nM of each oligo in 1× TE pH 8.0 (Invitrogen, AM9849) in S2)) and topping up the remainder with Antibody Diluent. The antibody mixture was then filtered through a S2 Buffer pre-wetted 0.1 µm pore size filter (Millipore, UFC30VV00) at 12,500 rpm for 1 minute at 4ºC. Blocked/photobleached tissues were then incubated with the filtered antibody mixture at 4ºC in a humidity chamber overnight.

After incubation, tissues were washed twice in S2 Buffer for 5 minutes each at RT. The slides were fixed with 1.6% PFA (diluted from 16% stock (EMS Diasum, 15740-04)) in PBS in a humidity chamber protected from light, at RT for 10 minutes. The slides were then rinsed quickly with PBS twice and washed once with PBS for 2 minutes. The slides were fixed with ice-cold methanol for 5 minutes on ice, followed by rinsing with PBS twice and washed once with PBS for 2 minutes. The

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

slides were then fixed in 4 µg/µL of BS3 Final Fixative (diluted from 200 µg/µL stock (ThermoFisher, 21580) in GlycoBuffer) twice for 15 minutes each at RT protected from light. The slides were washed once in 1X CODEX Buffer (10mM Tris pH 7.5, 0.02% NaN3, 0.1% Triton X-100, 10 mM MgCl2-6H2O, 150mM NaCl). A cotton-tipped applicator was used to scrape off the hydrophobic barrier on the slides and the slides were stored in 1X CODEX Buffer. Details regarding the antibody clones, vendors, conjugated channels, titers, exposure times, and assigned channels throughout the study are in **Supplementary Table 11**.

Imaging

CODEX image capture was performed using two platforms. For FL14 and FL33 cases, imaging was conducted using the automated PhenoCycler Fusion platform (Akoya Biosciences). For TMA cases, imaging was performed using the CODEX Keyence platform (Akoya Biosciences). For PhenoCycler Fusion imaging, each tissue slide was mounted with a flow cell (Akoya Biosciences, 240205) and incubated in 1× CODEX Buffer for 10 minutes at room temperature (RT). Flow cells were not required for TMA slides imaged using the CODEX Keyence platform. A reporter plate was prepared for each tissue slide with each well corresponding to each imaging cycle. Briefly, each well of a 96-well black reporter plate (BRAND Tech, 781607) was filled with Plate Buffer (500 µg/mL sheared salmon sperm DNA in 1X CODEX Buffer) supplemented with 54.11 mM Hoechst 33342 (ThermoFisher, H3570) and 100 nM each of ATTO550 or AlexaFluor647-conjugated antibody corresponding to complementary reporter oligos (GenScript, HPLC purified). The wells were then sealed using aluminum plate seal (ThermoFisher, AB0626) and stored at 4ºC until use.

CODEX analysis

Image processing

For all antibody markers in the FL14 and FL33 datasets, the 16-bit intermediate QPTIFF files generated by the Phenocycler Fusion were used to ensure optimal dynamic range of data. The QPTIFFs were first processed by subtracting the last blank cycle scaled by the ratio between current channel cycle and total cycle number, i.e.,

$$X'_{i,j} = X_{i,j,0} - \left(\frac{i}{N}\right) \times X_{\varepsilon}$$

where $X'_{i,j}$ is the blank-subtracted image of marker *j* in cycle *i*; $X_{i,j,0}$ is the 16-bit intermediate image of marker *j* in cycle *i*; and Xε is the last blank cycle. These images were then processed in imageJ using the "Math" and "Subtract Background" functions under "Process" to 1) subtract the mean pixel value of the image to remove most of the "salt and pepper" noise, and 2) subtract the background generated by the sliding paraboloid algorithm with a 5 pixel radius. Finally, binary masks were generated for these images using Otsu binarization. Then the final images were generated by taking the Hadamard product of the binary masks and the background subtracted images.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

For the TMA dataset, final 16-bit QPTIFF files generated following image stitching and background subtraction were directly used for downstream analysis.

Cell segmentation and single cell feature extraction

For the FL14, FL33 and TMA datasets, cell segmentation was performed as described in the **GeoMx methods** section only with the following changes: the nuclear channel input of MESMER was DAPI while that for membrane channel was a summation of HLA-1, CD31, CD3, CD68 and podoplanin.

Data processing

Cell sizes that were below 0.001 (i.e., non-specific dots or tissue debris) and over 0.995 percentile (i.e., due to flipping of tissue during imaging) were first removed. All the markers were then scaled by their respective median nuclear DAPI signal of each tissue sample to adjust for different binding efficiency of markers. Lastly, a local min-max scaling within each tissue was applied to all the markers to scale the marker expression levels to be within [0,1].

Cell phenotyping

Cell phenotyping was performed on the processed datasets through iterative rounds of clustering and annotation using PhenoGraph based on phenotyping marker enrichment profiles (42). The markers used for phenotyping includedPax5/CD20, CD21, Podoplanin, CD3, CD8, CD4, FOXP3, CD68, CD163, CD206, CD11b, CD11c and CD31. Clustering results were visualized as heatmaps showing Z-score-normalized expression of each marker across clusters. Clusters were annotated based on the combination of markers with the highest Z-score within each cluster. The following cell populations were identified: B/Tumor, $PDPN^{+}$Lymphoid stromal cells (LSC), CD8 T cell, CD4 T cell, T regulatory (Treg), $CD68^{+}$ macrophage, $CD68^{+}CD163^{+}$ macrophage, $CD11c^{+}$ cells, $CD31^{+}$ LSC and $PDPN^{+}CD31^{+}CD206^{+}$ LSCs. Monocytes were additionally annotated in the FL14 and FL33 samples, where the larger tissue area compared with TMA cores enabled more reliable identification of this population. Within follicular regions, differential expression patterns of CD21, podoplanin, and Pax5/CD20 were observed across the TMA, FL14, and FL33 samples based on multiplexed imaging data. To improve annotation accuracy, follicular regions were further classified according to marker expression patterns. In the TMA dataset, these regions were annotated as $PDPN^{lo}CD21^{+}CD20^{+}$ and $PDPN^{hi}CD21^{+}CD20^{+}$ (**Supplementary Fig. 7A-B**). In the FL14 sample, regions were annotated as $PDPN^{lo}CD21^{+}Pax5^{+}$ and $PDPN^{hi}CD21^{+}Pax5^{+}$ (**Supplementary Fig. 8B-C**). In the FL33 sample, regions were annotated as $PDPN^{lo}CD21^{+}Pax5^{+}$ and $PDPN^{hi}CD21^{+}Pax5^{lo}$ (**Supplementary Fig. 8D-E**). Cells that showed unclear marker enrichment patterns were annotated as "Other" cells. After an initial round of clustering with PhenoGraph was performed, clusters with clear enrichment patterns were annotated, while clusters with mixed patterns underwent additional rounds of iterative clustering and annotation until all identifiable cells were annotated. To visualize and confirm the assigned annotations, Mantis Viewer (Schiemann R et al., Zenodo 2020, available from: https://doi.org/10.5281/zenodo.4909620) was used to overlay the annotations onto the segmentation mask and the marker image, for visual inspection.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Marker enrichment heatmap

The marker enrichment heatmap showed the Z-score of a given tuple (e.g., marker, cell type, disease state). In other words, it showed how many standard deviations away the mean of marker A expression of cell type B given a disease state is from the population mean of marker A expression:

$$Z_{i,j,k} = \frac{(\mu_{i,j,k} - \mu_i)}{\sigma_i}$$

where $Z_{i,j,k}$ stands for the Z-score for marker *i*, cell type *j* and disease status *k*; $\mu_{i,j,k}$ stands for the mean expression for marker *i*, cell type *j* and disease status *k*; μ_i stands for the population mean of marker *i*, and σ_i stands for the population standard deviation of marker *i*. Heatmaps were plotted with R package ggplot2.

Relative proportion of cell populations

Relative proportion of cell type was presented as the ratio between the total number of cells within a given cell type with the total number of cells within the tissue sample of a disease state:

$$\frac{\sum_i}{\sum_j}$$

where $\sum_i$ stands for the total number of cells in cell type *i*, and $\sum_j$ stands for the total number of cells within the tissue sample *j*. Violin plots were plotted with R package ggplot2. Differences in cell proportions were determined using a Wilcoxon rank-sum test statistic, with significance assessed by a permutation test taking into account the partially paired observations. 1,000,000 iterations were performed randomly swapping FL/tFL labels of paired samples and randomly permuting FL/tFL labels amongst unpaired samples to generate the null distribution of rank-sum statistics and used to calculate two-tailed empirical p-values, which were then corrected for multiple hypotheses with a Benjamini-Hochberg procedure.

Spatial anchor expansion analysis

Binary masks demarcating the in-transition regions of FL14 and FL33 based on pathological review from S.S. were initially generated using ImageJ. These masks were then used to create concentric expansion zones at 500 μm intervals, which were overlaid onto the cell segmentation mask. Each cell falling within a given zone was assigned a corresponding step number, enabling analysis of relative cell type proportions as described above.

Macrophage-anchored hop-analysis (cytokine)

CXCL9^{high} and *CXCL9*^{low} cores were defined based on gene expression profiles obtained from GeoMx analysis performed on adjacent tissue sections. Regions of interest (ROIs) were selected on tissue microarray (TMA) cores, which were spatially matched to the corresponding adjacent sections, allowing inference of *CXCL9* expression levels for each core based on ROI-derived transcriptomic data. *CXCL9* expression status (high vs. low) was determined using the median *CXCL9* expression value across all analyzed cores, including both FL and tFL samples. All cores were subsequently classified into *CXCL9*^{high} or *CXCL9*^{low} groups. Downstream analyses were restricted to tFL cores. Using known core identities, multiplexed imaging data generated by

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

CODEX on the adjacent section were leveraged. Cell populations were previously annotated within the TMA dataset. Hop analysis was performed using macrophages as anchor cells within tFL cores. Hop distances were defined based on centroid-to-centroid distances between anchor macrophages and neighboring cells, where 1-hop corresponded to 20 μm, 2-hop corresponded to 40 μm, and combined 1+2 hop represented cells located within 0–40 μm from the anchor macrophage centroid. Within each hop distance, $CD4^+$ T cells, $CD8^+$ T cells, and Tregs were quantified per core and summarized as mean cell counts. The same analysis pipeline was repeated using SPP1 expression for core stratification and downstream spatial hop analysis. Statistical differences between FL and tFL cores were assessed using the Mann-Whitney U test accounting for duplicate sampling within the same category by averaging.

Macrophage-anchored window analysis (TMA)

Individual $CD68^+$ macrophages or $CD68^+CD163^+$ macrophages were designated as anchor cells. Macrophage-centered windows were generated by expanding 25 μm or 50 μm in all directions from the centroid of each anchor cell, resulting in macrophage windows with diameters of 50 μm and 100 μm, respectively. Within each window, the number of $PDPN^+$ LSCs or tumor B cells was quantified. The ratio of this number to the total number of $PDPN^+$ LSCs or tumor B cells within the corresponding core was calculated and defined as the proximity fraction of $PDPN^+$ LSCs or tumor B cells relative to the anchor macrophage. Results were visualized as violin plots, where each data point represents the proximity fraction calculated per core. Statistical comparison between FL and tFL was done using a rank-sum permutation test as described above.

Macrophage-anchored spatial cell type enrichment analysis in FL14

To characterize spatial cellular composition surrounding macrophages, an anchor-centered window analysis was performed with the FL14 sample. $CD68^+$ macrophages were designated as anchor cells. For each anchor macrophage, a square spatial window was generated by expanding ±50 μm in both X and Y directions from the centroid of the anchor cell, corresponding to a 100 × 100 μm macrophage-centered window. Overlapping windows were permitted. Windows containing fewer than 10 cells were excluded from downstream analysis. Within each window, the fraction of each annotated cell type was calculated relative to the total number of cells in that window. To account for differences in overall cellular composition between disease groups, window-level cell-type fractions were converted to $\log_2$ enrichment scores by normalizing to the baseline fraction of each cell type across all cells within the corresponding disease group. Each window was assigned to a step based on the step annotation of the anchor macrophage generated from the previous **Spatial anchor expansion analysis.** Enrichment scores were then averaged across windows within each step to generate step-level enrichment profiles for each cell type. For visualization, enrichment values were Z-score normalized across steps for each cell type. To facilitate direct comparison between FL and tFL, step indices from indolent samples were represented as negative values and transformed samples as positive values, and combined into a unified step heatmap.

Visium HD methods

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

For Visium HD, tissue deparaffinization and H&E staining was done according to the Visium HD manual (CG000684) using the staining solutions recommended by 10x Genomics. Tissue sections were deparaffinized and stained with hematoxylin (Sigma Aldrich, cat. GHS232) and eosin (Sigma Aldrich, cat. HT110116-500ML) (H&E) according to manufacturer protocol (Visium CytAssist Protocol CG000520). After H&E staining, slides were coverslipped with VWR Micro Cover Glasses (VWR, cat. 48393-241) and glycerol solution (Sigma Aldrich, cat. 49781) as mounting medium and whole slides were imaged on a Grundium Ocus 40 scanner. Slides were then de-coverslipped according to manufacturer protocol and stored at 4°C with dessicator until being processed on the CytAssist. RNA was extracted from a separate slice of tissue for each case and DV200 was >30% for all used samples. After assembling two slides with the deparaffinized and stained tissue sections into a Tissue Slide Cassette, sections were de-stained and de-crosslinked, the human whole transcriptome probe panel was added to slides, and probe pair hybridization to their complementary target RNA proceeded overnight. Following washing, hybridized probe pairs were ligated. Following placement of the tissue slides and Visium slide in the Visium CytAssist instrument, probe release and probe capture on Visium slide were performed. The Visium slide was assembled into a Visium Cassette to perform probe extension. Eluted barcoded ligation products were transferred to PCR tubes for subsequent amplification, bead clean-up, and sequencing library preparation. The sequencing libraries were quantified using the Bioanalyzer DNA High Sensitivity Kit (Agilent 5067-4626). The libraries were sequenced on the Illumina NextSeq 1000 using P2-100cycle kits or Illumina NovaSeq SP 100cycle kit. The sequencing parameters for Visium HD libraries were: Read 1: 43nt, Read 2: 50nt, Index 1: 10nt, Index 2: 10nt.

Visium HD data analysis

Visium HD libraries were pre-processed using spaceranger v3.0.1. Manual alignment was performed for 3/5 samples using loupe browser. Clustering solutions from the spaceranger pipeline using the 8x8 μm bin resolution were selected by manual review to best capture known cell populations (Graph-based clustering for FL14, and K-means clustering with K=10 for FL33, FL58, FL137), and clusters were manually annotated in loupe browser. Biopsy regions were annotated as indolent or transformed by expert pathologist. Differentially expressed genes between indolent and transformed regions were performed using loupe browser and visualized by volcano plot in python. For FL 137, no truly indolent region could be determined within Visium ROI. Differential expression was done between the transformed region and the adjacent, transitional region. Neighborhood analysis was performed using squidpy 1.2.3. Neighborhood enrichment score quantifies the number of co-occurring of cluster labels within the nearest 20 neighboring spots, z-normalized by the null mean and standard deviation under label permutation as determined by squidpy (50). Cell type scoring was done using our scRNA-derived signatures described previously by summing expression across genes in the signature.

PNMF Analysis

Spatial gene expression programs were analyzed using probabilistic non-negative matrix factorization (PNMF) with a Poisson expression distribution model, inspired by Townes and Engelhardt (51). 16 μm-binned Visium HD data was used with the number of components set to

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

15. The most relevant components were selected and manually analyzed by comparing the expression profile of each component with classical markers as well as markers derived from our scRNA-seq data. Figure titles indicate the cell types with the most prominent marker contribution within each component.

Migration assay

T cells and monocytes were isolated using EasySep Human T cell Isolation Kit and EasySep Human Monocyte Enrichment Kit without CD16 Depletion (both StemCell, cat. 17911 and cat. 19058) according to the manufacturer's instructions. Purity of isolated cells was confirmed by flow cytometry. Migration assay was performed in a 96 Transwell Assay Plate with 5 μM pore size (Corning, cat. 3387) with 50,000 cells per well in RPMI + 10% FBS + 1% P.S. Cytokines (SPP1, cat. 1433-OP; POSTN, cat. 3548-F2; CCL21, cat. 366-6C; SDF1alpha/CXCL12, cat. 350-NS, all R&D) were added to the lower chamber at a final concentration of 100 ng/ul, and equal amounts of the solvent PBS+0.1% BSA served as control. Supernatant of Oci-Ly-1 and Hs895.T cell lines (1:1 ratio with RPMI + 10% FBS + 1% P.S.) was used for conditioned media conditions with equal amounts of IMDM + 10% FBS + 1% P.S. or DMEM + 10% FBS + 1% P.S. as controls. Assay was performed in triplicates, and cells that migrated to the lower chamber were counted on a Fortessa flow cytometer (BD).

Cell Lines

Cell lines were used for production of conditioned media in experiments. Hs895.T (CRL-7637,) and HS-5 (CRL-3611) cell lines were obtained from ATCC (Manassas, VA, USA). SC1 (ACC 558) and DOHH2 (52) (ACC 47) were obtained from DSMZ (Heidelberg, Germany). Early passages of these cell lines (<10 after thawing) were used for experiments. OCI-Ly-1 was provided by the Wu Lab as reported in (33), cell line authentication was done through STR profiling (ATCC) in May 2026. Hs895.T were cultured in DMEM + 10% FBS + 1% P.S., OCI-Ly-1 were cultured in IMDM + 10% FBS + 1% P.S., all other cell lines were cultured in RPMI + 10% FBS + 1% P.S. All media and media supplements were obtained from Gibco Life Technologies/Thermo Fisher. All cell lines were routinely tested negative for mycoplasma contamination.

Isolation of monocytes and FL cells from primary blood samples

PBMCs from healthy donor leukoreduction collars were isolated by density centrifugation using Ficoll Paque (Global Life Sciences, cat. 1714400). Monocytes from PBMCs were isolated using the EasySep Human Monocyte Enrichment Kit without CD16 depletion (StemCell, cat. 19058) or Mojosort Human Pan Monocyte Isolation Kit (Biolegend, cat. 480059). For the primary FL patient blood sample with circulating FL, PBMCs were isolated from EDTA-blood using density centrifugation with Ficoll Paque. CD19+ cells were depleted from the PBMC using the REAlease CD19 microbead kit (Mitenyi, cat. 130-117-034) to obtain label-free CD19+ cells. CD19+ cells were frozen in 90% FBS + 10% DMSO until used in the phagocytosis assay. After depletion of CD19+ cells, negative selection of monocytes was done using the Mojosort Human Pan Monocyte Isolation Kit (Biolegend, cat. 480059). Purity was confirmed by flow cytometry.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Monocyte-to-macrophage Differentiation

After isolation of monocytes as described above, 500,000 monocytes/ml were plated in RPMI+10%FBS+1%P.S supplemented with 40 ng/ml M-CSF (R&D, cat. 216-MC). The media was changed every 2-3 days until day 5 or 6. To assess SPP1 levels in macrophages, all supernatant was removed and fresh media + M-CSF (final concentration 40 ng/ml) with or without 100 ng/ml SPP1/POSTN/CCL21(R&D, cat.1433-OP, 3548-F2, 366-6C) or 50 % supernatant from GC B cell lymphoma cell line OCI-Ly-1 or from stromal cell lines Hs895.T or HS-5, respectively, was added. IMDM+ 10% FBS + 1% P.S or DMEM 10% FBS + 1% P.S served as controls for conditioned media. 48 hours later, all supernatant was removed, and macrophages were detached using 10X TrypLe (Gibco Life Technologies/Thermo Fisher, cat. A1217702), harvested and analyzed by flow cytometry.

Flow Cytometry

After harvesting, monocytes/macrophages were washed and resuspended in PBS. Zombie Violet Fixable Viability Kit (Biolegend, cat. 423114) was used for viability staining. TruStain FcX (Biolegend, cat. 422302) and True-Stain monocyte blocker (Biolegend, cat. 426102) were added to prevent nonspecific staining. Surface staining (CD206 APC, clone 15-2, cat. 321110; CD163 Per-Cp-Cy 5.5, clone RM3/1, cat.326512; CD14 BV605, clone M5E2, cat. 301834; CD16 AF488, clone 3G8, cat. 302019 30201; CD19 BV711, clone HIB19, cat. 302246, all Biolegend) was performed according to the manufacturer's protocol. Fixation and Permeabilization for intracellular staining were performed using the Cyto Fast Fix/Perm Buffer Set (Biolegend, cat. 426803) following the manufacturer's protocol. For intracellular staining, PE-conjugated anti-SPP1 antibody (R&D, clone 223126, cat. IC14331P) and IgG2A PE-conjugated isotype control (R&D, clone 223126, cat. IC003P) were used. Flow analysis was performed on a Fortessa analyzer (BD) at Dana-Farber Flow Cytometry Core Facility. Analysis of all flow cytometry experiments was done in FlowJo version 10.

Western Blot

After harvesting the macrophages as described above, cells were washed twice with ice cold PBS and lysed using RIPA buffer (Sigma-Aldrich, cat. R0278) containing protease inhibitor (Roche, cat. 11836170001) according to the manufacturer's instructions. Protein concentration was determined using the Pierce Dilution-Free Rapid Gold BCA Assay (Thermo Fisher, cat. A55860). Gel electrophoresis was performed on NuPage Bis-Tris Protein Gels following the manufacturer's protocol and blotting was performed using the iBlot3 system (Invitrogen). After blocking the membrane in TBS-T (Cell Signaling Technology, cat. 9997) containing 5% BSA (Sigma-Aldrich, cat. 8914) membranes were incubated in anti-SPP1(R&D, cat. MAB1433) and anti-GAPDH (Cell Signaling Technology, cat. 5174) overnight, followed by secondary antibody incubation in anti-mouse or anti-rabbit IgG HRP-linked antibody (Cell Signaling Technology, cat. 7076 and 7074). The SuperSignal West Pico Plus Chemiluminescence Substrate (Thermo Fisher, cat. 34580) was used for chemiluminescence-based analysis on the Amersham ImageQuant 800 System.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Phagocytosis Assay

PBMC-derived monocytes were differentiated into macrophages as described above. After 7-10 days, 40.000 macrophages per well were plated into 96-well plates (Falcon, cat. 353072) in 70 µl full media. For HS-5 conditioned media, 30 µl of the full media was replaced by supernatant from HS-5 cells during their growth phase. All media was supplied with 40 ng/ml M-CSF. After 24 hours, SC1 or DOHH2 cells were labelled with 15 ng/ml pHrodo Red (Sartorius, cat. 4649) according to the manufacturer's instructions and were added to the macrophages at an effector:tumor ratio of 1:4 to a total well volume of 100 µl. POSTN and SPP1 were added to wells at a concentration of 100 ng/ml, with PBS+0.1%BSA was used as control. All conditions were analyzed in duplicates or triplicates. For the primary FL samples, frozen FL cells were thawed and incubated for four hours before labelling with 15 ng/ml pHrodo Red, which was performed as described above. Live-cell imaging was performed on the Incucyte S3 for 4 hours, with 4 images per well imaged every 30 minutes. For analysis of the cell lines and healthy donor-derived macrophages, total red color unit (RCU) integrated intensity was determined at the timepoint where phagocytic activity reached its plateau in the control samples of each cell line (DOHH2: 2 hours, SC1: 1.5 hours). Due to higher variability in staining at baseline due to background cell debris, the phagocytosis assays with FL patient macrophages and primary FL B cells were analyzed by normalizing the RCU integrated intensity of each timepoint and well to its respective baseline value.

SomaScan

14 EDTA-plasma samples from 12 patients with active disease at time of transformation or with indolent FL and history of transformation was submitted to SomaLogic. A panel of 100 reagents targeting 86 different proteins was analyzed on the SomaScan platform. Adaptive Normalization by Maximum Likelihood (ANML) normalized data received from SomaLogic was used for analysis. One patient/sample with FL histology 8 weeks prior to the sampling yet clinical deterioration suggestive of potential transformation, was excluded from the analysis after second review of the patient's charts as ultimate histology could not be determined. Data analysis was performed in Visual Studio Code with python using pandas, seaborn, scipy and statsmodels for analysis and data visualization. Welch's t-test of $\log_2$ transformed data was performed to compare FL and tFL (DLBCL, DHL). FDR was calculated using the Benjamini-Hochberg method. Due to small sample size leading to high FDR values, uncorrected p values are shown for differential protein levels confirming transcriptomic results (**Supplementary Table 9**). Additional Somascan data for 31 analytes corresponding to 23 unique proteins was obtained from an independent cohort of Streck-tube-derived plasma samples from untreated FL patients and healthy donors. Data analysis was performed as described as above for the EDTA-plasma FL-tFL cohort.

Validation cohort I - Bulk transcriptome. For mRNA profiling, 22 FFPE samples (11 pairs) were used for library preparation with QuantSeq 3' mRNA-Seq Library Prep Kit FWD (Lexogen, Catalog number: 015) according to manufacturer's protocol (without the modifications for FFPE samples). The amplification step of library preparation was performed as qRT-PCR by adding 1 ul of EVAGREEN dye (Biotium) to the reaction, which allowed for the determination of the correct

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

amount of amplification cycles to avoid under or overamplification. 500 ng of total RNA was used as input and the Lexogen i5 6 nt Unique Dual Indexing Add-on Kit together with UMI Second Strand Synthesis Module for QuantSeq FWD was used during the library preparation according to protocol. The analysis was done with DESeq2NGS data has been deposited at the European Genome-phenome Archive (EGA), under accession number EGAD50000001384.

Validation cohort II - GeoMx samples. TMAs were assembled under the guidance of expert hematopathologists with GeoMx guided bulk samplings performed of tFL/DLBCL tumor, FL follicular tumor or interfollicular regions. Data was pre-processed and analyzed as above.

Data sharing statement:

Raw scRNA-seq and bulk transcriptomic sequencing data are available at dbgap upon publication at https://www.ncbi.nlm.nih.gov/gap/ with study ID phs004175.v1. GeoMx and Visium HD data as well as scRNA-seq count matrices and Scanpy objects are available at NCBI Geo (https://www.ncbi.nlm.nih.gov/geo/) under GSE303544 (scRNA, processed data only), GSE303545 (GeoMX, raw and processed), GSE303546 (Visium, raw and processed), GSE303548 (bulk RNA, processed only). CODEX image files and related data files are available at Zenodo (10.5281/zenodo.15690757).

Code availability statement:

Code to reproduce the analyses in this manuscript is available on Github under https://github.com/getzlab/FL_transformation_paper. Code to reproduce CODEX analysis is available at Zenodo (10.5281/zenodo.15690757) along with the CODEX data.

Author contributions:

N.J.H. analyzed single-cell transcriptome data of cohort 1 with S.G., C.J.W., E.M.P., and G.G.
E.M.P. processed single cell samples and generated data with S.L., K.L., and W.S.L..
V.S. collected and reviewed pathology for discovery cohort 1. S.S. collected and reviewed pathology samples and supervised selection of cases for TMA and Visium HD. S.J.R. provided technical support and advised and supervised TMA creation with S.S. TMA was organized by E.M.P., S.J.R., S.S., S.L.D. and S.P.T.Y..
S.R., S.L., K.L., M.O., E.M.P. and S.P.T.Y. generated Visium HD data. N.J.H. performed Visium HD analysis with A.P. with input from L.B., E.M.P, and S.P.T.Y..
S.P.T.Y. performed GeoMx and analyzed data with N.J.H. with input from Y.Y.Y., H.Q., Y.Z., S.J.R., E.M.P., S.S., S.J. and G.G. A.B. and A.K.S. performed sequencing of GeoMx libraries.
S.P.T.Y. performed CODEX and analyzed data with input from H.Q. and S.J..
S.G., E.M.P., K.M. and L.B. obtained clinical information.
S.G. and K.M. collected discovery cohort II tissue for bulk transcriptome analysis.
N.J.H. and K.S. analyzed the bulk transcriptomes with E.M.P, G.G. and C.J.W..
L.B. performed and analyzed migration assays, macrophage differentiation assays and phagocytosis assays with support from Z.L.
S.H. helped with cell culture and preparing functional assays.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

L.B. analyzed SomaScan data with N.J.H..
E.D.J. and M.MD. designed and organized sample biobanking.
M.M. and D.F. generated and analyzed RNA-sequencing data for the 11 paired FL and tFL samples and shared TPM data
HK collected, reviewed, annotated all TMA samplings and conducted GeoMx sequencing for the validation spatial transcriptomic dataset.
M.A.M. and M.Z. generated Somascan data for the FL-healthy donor cohort.
E.L. performed PNMF and spatial transcriptomic analysis together with N.J.H. and J.G.
B.P.D. provided project operations oversight.
E.M.P., N.J.H., S.P.T.Y. and L.B. wrote the manuscript with G.G. and S.J. and input from all co-authors.
S.J., G.G., and E.M.P. jointly co-supervised the study and co-designed the study with N.J.H., S.P.T.Y. and L.B.
Y.Y. and Y.Z. provided statistical support under the supervision of S.T. and D.S.N. provided initial statistical support and aided in study design.
All authors reviewed the manuscript.

Acknowledgements:
E.M.P. received support from NIH/NCI K08 CA270085, is a 2023 LRF LSRMP scholar and received research funding from an ASH Fellow to Faculty Scholar Award and a Damon Runyon Clinical Investigator Award.
S.J. is supported by in part by NIH DP2AI171139, R01AI149672, R01GM152585, U24CA224331, R01NS139479, a Gilead's Research Scholars Program in Hematologic Malignancies, a Sanofi iAward, the Broad Next Generation Award, the Dye Family Foundation, and the Bridge Project, a partnership between the Koch Institute for Integrative Cancer Research at MIT and the Dana-Farber/Harvard Cancer Center.
N.J.H. was partially funded by the Landry Cancer Biology Consortium fellowship.
C.J.W. received support from CIMAC. CJW is a member of the Parker Institute for Cancer Immunotherapy at DFCI. Her work is supported, in part, by the Parker Institute for Cancer Immunotherapy. C.J.W. is also the Lavine Family Chair for Preventative Cancer Therapies at DFCI.
G.G. was partially funded by the Paul C. Zamecnik Chair in Oncology at the Mass General Cancer Center.
S.L. is supported by NCI R50 Research Specialist Award (R50CA251956).
S.P.T.Y. is a MacMillan Family Foundation Awardee of the Life Sciences Research Foundation.
L.B. is funded by a Walter Benjamin fellowship from the German Research Foundation.
K.M. was supported by the Lubin Family Foundation Scholar Award
A.B. is supported by a David H. Koch fellowship from the Koch Institute of Integrative Cancer Research, MIT.
A.K.S. was supported, in part, by the Bridge Project, a partnership between the Koch Institute for Integrative Cancer Research at MIT and the Dana-Farber/Harvard Cancer Center.
M.M. and D.F. were supported by the Ministry of Health of the Czech Republic in cooperation with the Czech Health Research Council under project No. NU22-03-00117, and by the Ministry of

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Education, Youth and Sports within programme Inter-Excellence II, INTER-ACTION (LUAUS26082).
M.A.M. and M.Z. gratefully acknowledge funding support from N. and T. Colatosti. M.Z. is supported by the 2023 AACR-Incyte Immuno-oncology Research Fellowship, Grant Number 23-40-46-ZHAN and the National Cancer Center Postdoctoral Fellowship.
H.K. received support from the Rhode Island Foundation (Grant #8442_20210960).

Sequencing was partially supported by the Broad/IBM Cancer Resistance Research Project.
We acknowledge Oriol Olive Noguen and Mikaela McDonough for assistance with IRB protocols and Fanny Dao and Sam Pollock for data management. Kristie Defalco and Dan McDermott are thanked for grant management. Natalie Bavli and Katie Tooley for data and sample collection.
Ashley Pelton for TMA construction.
Mikaela McDonough, Jamie Dela Cruz, Siang Dim, Dr. Jerry Ritz and members of the Lymphoma Biobank and the Pasquarello Tissue Bank Teams.
For facilitating TMA construction with Habibe Kurt, we thank Dr. Adam Olszewski

Bibliography

1. Wagner-Johnston ND, Link BK, Byrtek M, Dawson KL, Hainsworth J, Flowers CR, et al. Outcomes of transformed follicular lymphoma in the modern era: a report from the National LymphoCare Study (NLCS). Blood. 2015 Aug 13;126(7):851–7.

2. Al-Tourah AJ, Gill KK, Chhanabhai M, Hoskins PJ, Klasa RJ, Savage KJ, et al. Population-based analysis of incidence and outcome of transformed non-Hodgkin's lymphoma. J Clin Oncol. 2008 Nov 10;26(32):5165–9.

3. Conconi A, Ponzio C, Lobetti-Bodoni C, Motta M, Rancoita PMV, Stathis A, et al. Incidence, risk factors and outcome of histological transformation in follicular lymphoma. Br J Haematol. 2012 Apr;157(2):188–96.

4. Parry EM, Okosun J. An updated understanding of follicular lymphoma transformation. Blood. 2025 Oct 9;146(15):1812–23.

5. Alaggio R, Amador C, Anagnostopoulos I, Attygalle AD, Araujo IB de O, Berti E, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid neoplasms. Leukemia. 2022 July;36(7):1720–48.

6. Campo E, Jaffe ES, cook JR, et al. The international consensus classification of mature lymphoid neoplasms: A report from the clinical advisory committee. Blood. 2022;140(11):1229-1253. Blood. 2023 Jan 26;141(4):437.

7. Okosun J, Bödör C, Wang J, Araf S, Yang C-Y, Pan C, et al. Integrated genomic analysis identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nat Genet. 2014 Feb;46(2):176–81.

8. Pasqualucci L, Khiabanian H, Fangazio M, Vasishtha M, Messina M, Holmes AB, et al. Genetics of follicular lymphoma transformation. Cell Rep. 2014 Jan 16;6(1):130–40.

9. Kridel R, Chan FC, Mottok A, Boyle M, Farinha P, Tan K, et al. Histological transformation

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

and progression in follicular lymphoma: A clonal evolution study. PLoS Med. 2016 Dec;13(12):e1002197.

10. Wang X, Nissen M, Gracias D, Kusakabe M, Simkin G, Jiang A, et al. Single-cell profiling reveals a memory B cell-like subtype of follicular lymphoma with increased transformation risk. Nat Commun. 2022 Nov 9;13(1):6772.

11. Barisic D, Chin CR, Meydan C, Teater M, Tsialta I, Mlynarczyk C, et al. ARID1A orchestrates SWI/SNF-mediated sequential binding of transcription factors with ARID1A loss driving pre-memory B cell fate and lymphomagenesis. Cancer Cell. 2024 Apr 8;42(4):720–2.

12. Sarkozy C, Wu S, Takata K, Aoki T, Neriah SB, Milne K, et al. Integrated single cell analysis reveals co-evolution of malignant B cells and tumor micro-environment in transformed follicular lymphoma. Cancer Cell. 2024 Sept 9;42(9):1630.

13. Kotlov N, Bagaev A, Revuelta MV, Phillip JM, Cacciapuoti MT, Antysheva Z, et al. Clinical and biological subtypes of B-cell lymphoma revealed by microenvironmental signatures. Cancer Discov. 2021 June;11(6):1468–89.

14. Smeltzer JP, Jones JM, Ziesmer SC, Grote DM, Xiu B, Ristow KM, et al. Pattern of CD14+ follicular dendritic cells and PD1+ T cells independently predicts time to transformation in follicular lymphoma. Clin Cancer Res. 2014 June 1;20(11):2862–72.

15. Abe Y, Sakata-Yanagimoto M, Fujisawa M, Miyoshi H, Suehara Y, Hattori K, et al. A single-cell atlas of non-haematopoietic cells in human lymph nodes and lymphoma reveals a landscape of stromal remodelling. Nat Cell Biol. 2022 Apr;24(4):565–78.

16. Roider T, Seufert J, Uvarovskii A, Frauhammer F, Bordas M, Abedpour N, et al. Dissecting intratumour heterogeneity of nodal B-cell lymphomas at the transcriptional, genetic and drug-response levels. Nat Cell Biol. 2020 July;22(7):896–906.

17. Filip D, Litzmanova K, Michaelou A, Kledus F, Devan J, Boudny M, et al. Repression of miR-29 via MYC leads to increased CD40 signaling in transformed follicular lymphoma. Leukemia. 2026 Feb 19;1–14.

18. Han G, Deng Q, Marques-Piubelli ML, Dai E, Dang M, Ma MCJ, et al. Follicular lymphoma microenvironment characteristics associated with tumor cell mutations and MHC class II expression. Blood Cancer Discov. American Association for Cancer Research (AACR); 2022 Sept 6;3(5):428–43.

19. Parry EM, Leshchiner I, Guièze R, Johnson C, Tausch E, Parikh SA, et al. Evolutionary history of transformation from chronic lymphocytic leukemia to Richter syndrome. Nat Med. 2023 Jan;29(1):158–69.

20. Nadeu F, Royo R, Massoni-Badosa R, Playa-Albinyana H, Garcia-Torre B, Duran-Ferrer M, et al. Detection of early seeding of Richter transformation in chronic lymphocytic leukemia. Nat Med. 2022 Aug;28(8):1662–71.

21. Broséus J, Hergalant S, Vogt J, Tausch E, Kreuz M, Mottok A, et al. Molecular characterization of Richter syndrome identifies de novo diffuse large B-cell lymphomas

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

with poor prognosis. Nat Commun. 2023 Jan 19;14(1):309.

22. Laurent C, Trisal P, Tesson B, Seth S, Beyou A, Roulland S, et al. Follicular lymphoma comprises germinal center-like and memory-like molecular subtypes with prognostic significance. Blood. 2024 Dec 12;144(24):2503–16.

23. Massoni-Badosa R, Aguilar-Fernández S, Nieto JC, Soler-Vila P, Elosua-Bayes M, Marchese D, et al. An atlas of cells in the human tonsil. Immunity. 2024 Feb;57(2):379-399.e18.

24. Patsalos A, Halasz L, Oleksak D, Wei X, Nagy G, Tzerpos P, et al. Spatiotemporal transcriptomic mapping of regenerative inflammation in skeletal muscle reveals a dynamic multilayered tissue architecture. J Clin Invest [Internet]. 2024 Aug 27;134(20). Available from: https://doi.org/10.1172/JCI173858

25. Bai B, Wise JF, Vodák D, Nakken S, Sharma A, Blaker YN, et al. Multi-omics profiling of longitudinal samples reveals early genomic changes in follicular lymphoma. Blood Cancer J. 2024 Aug 27;14(1):147.

26. Li X, Singhal K, Deng Q, Chihara D, Russler-Germain D, Harkins RA, et al. Large B cell lymphoma microenvironment archetype profiles. Cancer Cell. 2025 July 14;43(7):1347-1364.e13.

27. Bill R, Wirapati P, Messemaker M, Roh W, Zitti B, Duval F, et al. CXCL9:SPP1 macrophage polarity identifies a network of cellular programs that control human cancers. Science. 2023 Aug 4;381(6657):515–24.

28. Kridel R, Mottok A, Farinha P, Ben-Neriah S, Ennishi D, Zheng Y, et al. Cell of origin of transformed follicular lymphoma. Blood. 2015 Oct 29;126(18):2118–27.

29. Mourcin F, Verdière L, Roulois D, Amin R, Lamaison C, Sibut V, et al. Follicular lymphoma triggers phenotypic and functional remodeling of the human lymphoid stromal cell landscape. Immunity. 2021 Aug 10;54(8):1788-1806.e7.

30. Radtke AJ, Postovalova E, Varlamova A, Bagaev A, Sorokina M, Kudryashova O, et al. Multi-omic profiling of follicular lymphoma reveals changes in tissue architecture and enhanced stromal remodeling in high-risk patients. Cancer Cell. 2024 Mar 11;42(3):444-463.e10.

31. Thul PJ, Lindskog C. The human protein atlas: A spatial map of the human proteome. Protein Sci. 2018 Jan;27(1):233–44.

32. Qi J, Sun H, Zhang Y, Wang Z, Xun Z, Li Z, et al. Single-cell and spatial analysis reveal interaction of FAP+ fibroblasts and SPP1+ macrophages in colorectal cancer. Nat Commun. 2022 Apr 1;13(1):1742.

33. Tumuluru S, Godfrey JK, Cooper A, Yu J, Chen X, MacNabb BW, et al. Integrative genomic analysis of DLBCL identifies immune environments associated with bispecific antibody response. Blood. 2025 May 22;145(21):2460–72.

34. Stahl D, Gödel P, Balke-Want H, Gholamipoorfard R, Segbers P, Tetenborg L, et al.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

CSF1R+ myeloid-monocytic cells drive CAR-T cell resistance in aggressive B cell lymphoma. Cancer Cell. 2025 Aug 11;43(8):1476-1494.e10.

35. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. Genome Biol [Internet]. 2018 Dec;19(1). Available from: http://dx.doi.org/10.1186/s13059-017-1382-0

36. Wolock SL, Lopez R, Klein AM. Scrublet: Computational identification of cell Doublets in Single-cell transcriptomic data. Cell Syst. 2019 Apr 24;8(4):281-291.e9.

37. Tirosh I, Izar B, Prakadan SM, Wadsworth MH 2nd, Treacy D, Trombetta JJ, et al. Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. Science. 2016 Apr 8;352(6282):189–96.

38. Graubert A, Aguet F, Ravi A, Ardlie KG, Getz G. RNA-SeQC 2: efficient RNA-seq quality control and quantification for large cohorts. Bioinformatics. 2021 Sept 29;37(18):3048–50.

39. Newman AM, Steen CB, Liu CL, Gentles AJ, Chaudhuri AA, Scherer F, et al. Determining cell type abundance and expression from bulk tissues with digital cytometry. Nat Biotechnol. 2019 July;37(7):773–82.

40. Van Valen DA, Kudo T, Lane KM, Macklin DN, Quach NT, DeFelice MM, et al. Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. PLoS Comput Biol. 2016 Nov;12(11):e1005177.

41. Greenwald NF, Miller G, Moen E, Kong A, Kagel A, Dougherty T, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. Nat Biotechnol. 2022 Apr;40(4):555–65.

42. Levine JH, Simonds EF, Bendall SC, Davis KL, Amir E-AD, Tadmor MD, et al. Data-driven phenotypic dissection of AML reveals progenitor-like cells that correlate with prognosis. Cell. 2015 July;162(1):184–97.

43. Merritt CR, Ong GT, Church SE, Barker K, Danaher P, Geiss G, et al. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. Nat Biotechnol. 2020 May;38(5):586–99.

44. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc. 2020 Apr;15(4):1484–506.

45. Nakagawa S, Schielzeth H. A general and simple method for obtainingR2from generalized linear mixed-effects models. Methods Ecol Evol. 2013 Feb;4(2):133–42.

46. Jiang S, Chan CN, Rovira-Clavé X, Chen H, Bai Y, Zhu B, et al. Combined protein and nucleic acid imaging reveals virus-dependent B cell and macrophage immunosuppression of tissue microenvironments. Immunity. 2022 June 14;55(6):1118-1134.e8.

47. Ji AL, Rubin AJ, Thrane K, Jiang S, Reynolds DL, Meyers RM, et al. Multimodal analysis of composition and spatial architecture in human squamous cell carcinoma. Cell. 2020 Sept 17;182(6):1661–2.

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

48. Phillips D, Schürch CM, Khodadoust MS, Kim YH, Nolan GP, Jiang S. Highly multiplexed phenotyping of immunoregulatory proteins in the tumor microenvironment by CODEX tissue imaging. Front Immunol. 2021 May 19;12:687673.

49. Black S, Phillips D, Hickey JW, Kennedy-Darling J, Venkataraaman VG, Samusik N, et al. CODEX multiplexed tissue imaging with DNA-conjugated antibodies. Nat Protoc. 2021 Aug;16(8):3802–35.

50. Palla G, Spitzer H, Klein M, Fischer D, Schaar AC, Kuemmerle LB, et al. Squidpy: a scalable framework for spatial omics analysis. Nat Methods. 2022 Feb;19(2):171–8.

51. Townes FW, Engelhardt BE. Nonnegative spatial factorization applied to spatial genomics. Nat Methods. 2023 Feb;20(2):229–38.

52. Kluin-Nelemans HC, Limpens J, Meerabux J, Beverstock GC, Jansen JH, de Jong D, et al. A new non-Hodgkin's B-cell line (DoHH2) with a chromosomal translocation t(14;18)(q32;q21). Leukemia. 1991 Mar;5(3):221–4.

Figure Legends

Figure 1. Identification of nodal populations across FL transforming to tFL. A. Schema showing experimental approaches and sample cohorts used for discovery and validation. FL and tFL samples from the same patient are reflected by connecting line. Created with www.biorender.com. **B**. Swimmer plot showing biopsy timepoints (triangles) and histology relative to FL diagnosis (time 0) for 10 patients with FL in discovery cohort I. Triangle color reflects biopsy histology (blue, FL; red, tFL; half blue and half red, admixed FL and tFL) and timepoints used for scRNA-seq (yellow outline). **C.** UMAP labelled with non-B cell nodal microenvironment populations recovered from scRNA-seq. Black outlines identify 13 coarse clustering populations. Corresponding fine clustering annotations are color coded in legend on right. **D.** UMAP of B cell populations isolated in scRNA-seq colored by patient. **E.** Dot plot of marker genes (bottom) that identify coarse (top) and fine cell type clusters (left) in scRNA-seq of FL and tFL samples (red shading, mean expression in units of log UMI counts per 10,000 [cp10k]; dot size, fraction of cells in cluster expressing transcript). **F.** Heatmap showing coarse cell populations from scRNA-seq by FL and tFL histology. Top bar represents histology (blue, FL; red, tFL; purple, FL close to tFL). Heatmap values represent Z-score of cell fraction. Line graph (top) reflects years to next FL transformation event (y-axis) for each biopsy sample (x-axis). Cell types with significant ($q<0.25$) FDR-corrected Wilcoxon text comparing FL and tFL samples are annotated.

Figure 2. tFL nodal composition is different from FL. **A.** Principal component (PC) analysis in bulk RNA-sequencing. Scatter plot (top) shows tFL (n=12, red) and FL (n=7, blue) biopsy samples separated by PC analysis (PC1, x-axis; PC2, y-axis; variance in parenthesis). Half blue and red dot indicates a tFL case that contained admixed FL. Loadings for the top weighted PC1 genes are plotted across bottom. **B.** Box plots showing cell fractions estimated by CIBERSORTx deconvolution of bulk expression data using signatures derived from the scRNA-seq data. Shown on box plots is median (bar) and inter-quartile range (IQR) with whiskers showing maximum and minimum values. Q values shown at top represent a Wilcoxon rank-sum test with Benjamini-Hochberg FDR correction. **C.** Schema of the GeoMx spatial transcriptomics design and analysis

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

workflow. Created with www.biorender.com. **D.** Representative images of FL (top row) and tFL (bottom row) regions of interest (ROI) with markers (left panel - nuclei, SYTO13 in gray; B/tumor cells, PAX5 in blue; macrophages, CD68 in yellow; T cells, CD3 in pink), corresponding phenotype maps (center panel - other cells, gray; tumor cells, blue; myeloid cells, yellow; and T cells, pink) and tissue H&E staining (right panel - hematoxylin, purple; eosin, pink). Individual images and ROI phenotype maps for TMAs are in **Supplementary Fig. 3**. **E.** Correlation of bulk-RNA PC1 loading (x-axis) versus differential expression (DE) t-statistic for tFL vs FL in GeoMx B cell guided samplings (y-axis). Pearson correlation coefficient is shown, with a p-value testing the null hypothesis that variables are uncorrelated using a beta test. Dots are colored by DE significance (orange, q less than 0.1; gray, q greater than or equal to 0.1) with top differential genes labelled on plot. **F.** Heat map of top transformation-associated genes. Values represent min-max scaled pseudobulk expression (transcripts-per-million) in scRNA data. Two separate dendrograms are shown for genes down (left) or up (right) in tFL compared to FL. Immune cell clusters and malignant cells per patient are at heatmap right y-axis and transformation-associated genes labelled on the x-axis. Breakout plots (T cells, blue; FRC/fibroblasts, red; monocytes/macrophages, green) show fine cluster annotations (y-axis) and associated genes (x-axis).

Figure 3. Transformation is associated with expansion of cytokine-negative stromal cells. A. Cell type scoring (**Methods**) in GeoMx-other samplings (FL, blue; tFL-DLBCL, red; tFL-Gr3B/FLBL, light blue). Box plots show median and IQR and whiskers show min/max. Connecting lines show paired samples. P-values are computed using a linear mixed-effect model. **B.** FDC and fibroblast cell type scores (x-axis) plotted against cytokine relative gene expression (y-axis, log UQ-normalized counts) in GeoMX B-rich regions. **C.** UMAP plot showing sub-clustered stromal cell populations identified in scRNA-seq. **D.** Violin plots showing expression (units, log cp10k) of marker genes across stromal populations. Blue shading reflects median expression in cluster group. **E.** Heatmap showing differential expression of *CCL21-CD34+* fibroblasts/FRCs as compared to the *CCL21+* FRCs. Heatmap values reflect expression in units of log cp10k. **F.** Volcano plot showing genes with GO annotation "Cytokine activity" differentially expressed between FL (left, blue arrow) and tFL (right, red arrow). X-axis is $\log_2$ fold change and y-axis is negative $\log_{10}$ p value as computed by limma testing. Points are color coded by significance (gray, q value $\geq$ 0.1; orange, q value < 0.1). **G.** Expression of *CCL21-CD34+* fibroblast and macrophage-restricted cytokine genes in bulk RNA-seq (top row: y-axis, transcripts per million or tpm; p values reflect rank-sum) and GeoMx (bottom row: y-axis, UQ-normalized counts or UQ; q values derive from limma analysis; lines show 9 paired FL and tFL samples). **H.** Correlation of stromal-produced cytokines (y-axis) and corresponding receptors (x-axis) in T cell-rich regions of GeoMx samplings (FL samples, blue; tFL, red; units, log UQ-normalized counts). P values and marginal R^2 statistics are derived from a linear-mixed effects model using the receptor and ligand as dependent and independent variables, respectively, and patient and sample labels as blocking variables.

Figure 4. Expanded tFL macrophages shape T cell infiltration. A. UMAP showing fine resolution monocyte and macrophage populations captured across FL and tFL. **B.** Heatmap showing top differential genes between macrophage and monocyte populations. Heatmap values reflect expression in logcp10k. **C.** tFL-enriched *GPNMB+* macrophage-associated genes in B cell-rich (left) and CD68+ macrophage(mac)-rich (right) GeoMx TMA regions. q values represent limma-adjusted p values and gray lines represent 9 paired FL and tFL samples. **D.** tFL-enriched *GPNMB+* macrophage-associated genes in bulk RNA samples. p values reflect independent Wilcoxon test. Boxplots in both **C** and **D** show median (horizontal line) and IQR with whiskers showing minima/maxima. **E.** SPP1 is predicted to interact with cell types through its ligands by

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

CellphoneDB. Heatmap values indicate relative expression of ligand pair within each scRNA-seq cluster. Line thickness represents the interaction score (product of ligand-receptor expression) **F.** *SPP1* expression in B cell-rich GeoMx samplings (x-axis) correlates with expression of fibroblast and pericyte marker genes (y-axis). Units are log UQ-normalized counts. **G.** Correlation of *CXCL9* gene expression (x-axis) in CD68-guided mac-rich TMA samplings with cytotoxic T cell gene expression (y-axis) in CD3-guided (T cell) TMA samplings. Units are log UQ-normalized counts. Marginal R^2 and a p-value testing parameter significance are derived from a linear mixed effects model accounting for repeat sampling (**Methods**). For **F** and **G,** FL samples are plotted in blue and tFL samples in red. **H.** Heatmap of scRNA-derived marker genes from cell types s differentially abundant between FL and tFL (top four sections), as well as top overall differentially expressed genes previously found in GeoMx analysis (bottom). Heatmap shows Z scored gene expression for 11 paired sequential FL and tFL samples(left) and average $\log_2$ fold change in tFL vs. FL calculated by limma (right column). **I.** Boxplots showing cell type scores across a validation cohort of FL (n=61 samplings from 20 unique biopsies) and tFL (n=13 samplings from 4 biopsies). Box plots show median and IQR and whiskers show minima/maxima across all ROIs. Connecting lines show average scores for paired samples (n=4FL/tFL pairs) with 95% confidence interval bands. P-values are computed using a linear mixed-effect model with histology as a fixed effect and random intercepts for each patient. * highlights one FL grade 3B case with preserved FDC signature.

Figure 5. Fibroblasts co-occur with macrophages in tFL regions and impact macrophage function. A. Visium HD ROI selection on H&E sections of four admixed FL and tFL cases (uc=unclassifiable histology) **B.** Volcano plot showing differentially regulated genes between tFL and FL/adjacent region up to tFL border in Visium HD by negative binomial test. **C.** H&E staining (left panel) of magnified ROI with dotted line separating tFL (red) from FL (blue; FL14, FL33, FL58) or transitional/unclassifiable region (FL137). Cell type specific scores (right panels) for FDCs, *CCL21*+ FRCs, *CCL21*- FRCs and *GPNMB*+ macrophages are depicted in green shading. **D.** Inferred cell types enriched in tFL regions (right) compared to FL or adjacent region (left) for Visium HD cases. Gray shading highlights monocyte/macrophage and stromal cell types. **E.** Clustering of cell types shows population co-localization in tFL regions in FL14. **F.** Neighborhood enrichment scoring across representative case (FL14). **G**. Schema showing macrophage assay workflow. Created with biorender.com. **H.** Bar plots representing Median Fluorescence intensity (MFI) of SPP1 in macrophages exposed to Hs895.T or HS-5 conditioned media (CM), OCI-Ly-1 CM or cytokine-enriched media relative to respective media-only control (DMEM, IMDM or RPMI) (n=3-4, one-sample t-test). Dotted lines mark separate experiments with different media controls. **I.** Representative images showing phagocytosis of DOHH2 (top) and SC1 (bottom) cells by macrophages exposed to CM derived from stromal cell line HS-5 (right) compared to RPMI control (left) after 2 hours/1.5 hours of coculture. SC1/DOHH2 cells engulfed by macrophages are depicted in yellow. Right: Quantification of phagocytosis of DOHH2 (top) and SC1 (bottom) by total red color unit (RCU) integrated intensity x μm^2 per image (y-axis) after 2 hours (DOHH2) or 1.5 hours (SC1), respectively (n= 4 healthy macrophage donors). Barplots show mean plus SD. One-way Anova with repeated measures and with Dunnett's multiple comparison test. **J.** Top: Barplot representing absolute MFI of SPP1 in macrophages derived from peripheral blood of a patient with circulating FL after exposure to HS-5 CM compared to RPMI control. Barplots show mean of technical duplicate plus standard deviation. Bottom: Line graph representing phagocytic rate of patient-derived FL cells by autologous macrophages in RCU integrated intensity x μm^2 per

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

image relative to baseline (y-axis), x axis shows time in hours. Data curves show individual lines for each technical replicate.

Figure 6. Distinct macrophages subsets spatially associate with fibroblasts and T cells. A. Top: Representative multiplexed images from TMA cores of patient 88 showing Pax5 (blue), CD3 (cyan), CD31 (orange), CD21 (yellow), and Podoplanin (magenta). Bottom: Corresponding phenotype maps of the same core showing cell types tumor (blue), $CD4^+$ T cells (purple), $CD8^+$ T cells (black), Treg (green), $CD11c^+$ cells (light green), $CD31^+$LSCs (orange), $PDPN^{low}CD21^+CD20^+$ (yellow), $PDPN^{hi}CD21^+CD20^+$ (pink), $PDPN^+$LSCs (red), $PDPN^+CD31^+CD206^+$LSCs (dark orange), $CD68^+$macrophage (purple-blue), $CD68^+CD163^+$macrophage (bright green) and others (white). Scale bar is 100 μm. **B.** Violin plots showing the proportions of indicated cell types in the TMA dataset. Each dot represents an individual core (FL=55, tFL=25). **C.** Schematic diagram showing the proximity enrichment fraction analysis. Created with www.biorender.com. **D.** Violin plots showing the proximity-enriched fractions of $PDPN^+$ LSCs and tumor B cells within 50 μm of $CD68^+$ macrophage anchors in FL and tFL cores. **E.** Violin plots showing the proximity-enriched fractions of $PDPN^+$ LSCs and tumor B cells within 50 μm of $CD68^+CD163^+$ macrophage anchors in FL and tFL cores. **F.** Left: schematic diagram showing the hop analysis in tFL cores (n=17). Right: example subgraphs generated from *CXCL9^{high}* and *CXCL9^{low}* groups. **G.** Violin plots showing the proportion of T cells per core in *CXCL9^{high}* and *CXCL9^{low}* groups. **H.** Violin plots showing the mean counts of T cells within 1+2-hop distances (0-40 μM) from macrophage anchors in *CXCL9^{high}* and *CXCL9^{low}* groups. **I.** Violin plots showing the mean counts of T cells within 1+2-hop distances (0-40 μM) from macrophage anchors in *SPP1^{high}* and *SPP1^{low}* groups. Statistical comparisons between the indicated groups in panels A-E was performed using a rank-sum permutation test. For panels F-I, a Mann-Whitney U test was used.

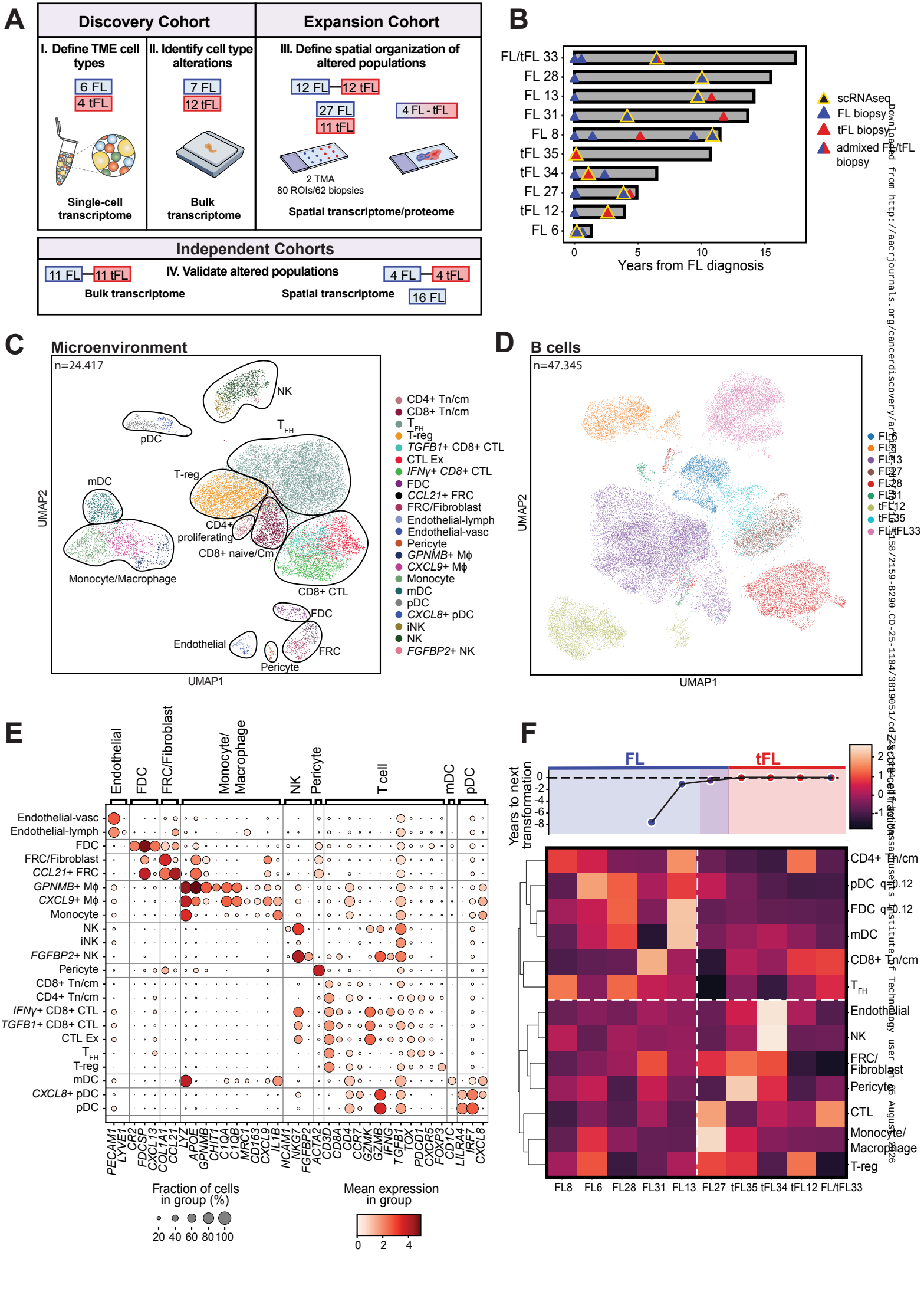
Figure 7. Macrophages and fibroblasts shape zones of tFL evolution and are detectable non-invasively. A. Left: representative CODEX multiplexed images of indolent FL and transformed tFL regions in FL14 (left) with markers for nuclei (DAPI, gray), BCL6 (blue), CD206 (green), CD31 (orange), CD21 (yellow) and CD3 (magenta) Right: corresponding phenotype map showing the same region in FL14 with cell types B cell/Tumor (blue), $CD4^+$ T cell (purple), Treg (teal), $CD68^+$ macrophage (purplish blue), $CD68^+CD163^+$ macrophage (green), $PDPN^{hi}CD21^+Pax5^+$(yellow), $PDPN^{lo}CD21^+Pax5^+$, (pink), $PDPN^+$LSC (cyan), $PDPN^+CD31^+CD206^+$LSC (red), $CD31^+$LSC (orange), DC (brown), $CD8^+$ T cell (black), monocytes (light green), others (gray). Scale bar: 100μm on larger images and 20μm for inset. Full phenotype maps of FL14 and FL33 are in **Supplementary Fig.8C&E**. **B.** Schematic diagram showing the macrophage-anchored cell type enrichment analysis across distance from the transition region. Created with www.biorender.com. **C.** Heatmap showing Z-score–normalized mean enrichment of cell types (columns) within $CD68^+$ macrophage–anchored windows (100 μm, rows) across 500 μm spatial increments extending from the transition region toward FL and tFL areas. Multiplexed images from FL14 (**D**) and FL33 (**E**) showing markers for nuclei (gray), CD21 (cyan), Pax5 (blue), CD4 (magenta), CD8 (green), and CD68 (red). Right panels show magnified transition zones highlighting abundant T cell and macrophage infiltrates at the border of regions with overt large-cell morphological changes, with macrophage-T cell niche retention/recruitment. Scale bars: 1 mm (whole tissue) and 100 μm (insets). **F.** Left: Scheme for spatial analysis anchored from in-transition regions (yellow and green boxes) for FL14 (top) and FL33 (bottom). Right: Proportion changes of $PDPN^+$ LSC, $CD68^+$ macrophages and $CD4^+$ T cells between 0-2500 μm (FL14) and 0-4500 μm (FL33) from the in-transition anchor. **G.** Scatterplot showing PC1 and PC2

Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

inferred from principal component analysis of targeted plasma proteomics (n=13, FL, blue; tFL-DLBCL, red; tFL-Gr3B/FLBL, light blue). **H.** Boxplots for cytokines detected in targeted plasma proteomic analysis. Boxplots show median (horizontal line) and IQR, whiskers represent minima/maxima. Units are log_2 transformed relative fluorescence units (RFU). p values reflect Welch's t-test. **I.** Scatterplot showing correlation of log_2 transformed RFU of CCL19 (x-axis) and CCL21 (y-axis) in plasma samples. R and p values reflect Pearson correlation coefficient. **J.** Scatterplot showing log_2 fold-change of plasma analytes in relative fluorescence units (RFU) in tFL to FL (y-axis) against RFU log_2 fold-change of plasma analytes in RFU in FL (x-axis) to healthy donor plasma.

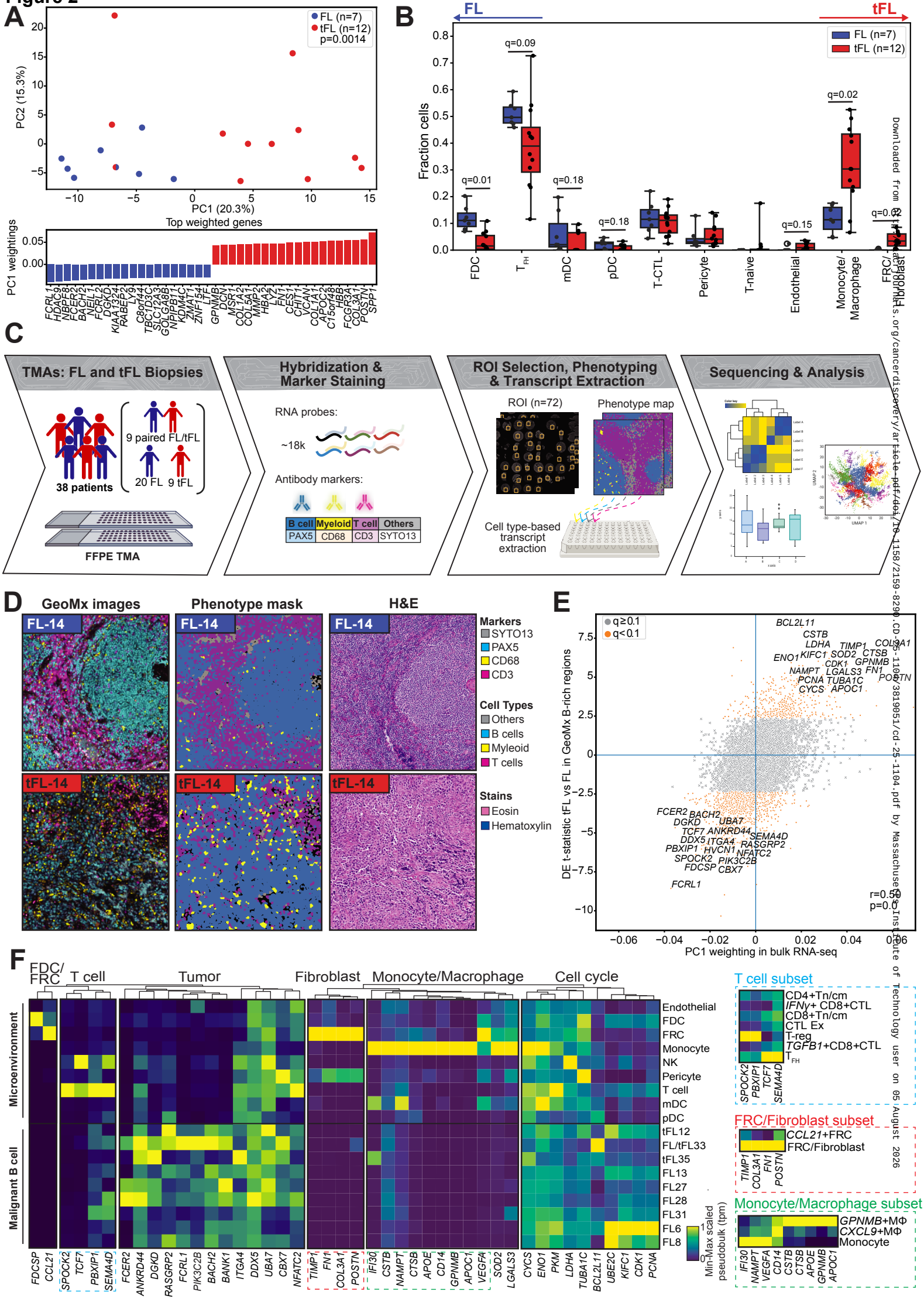
Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Figure 1



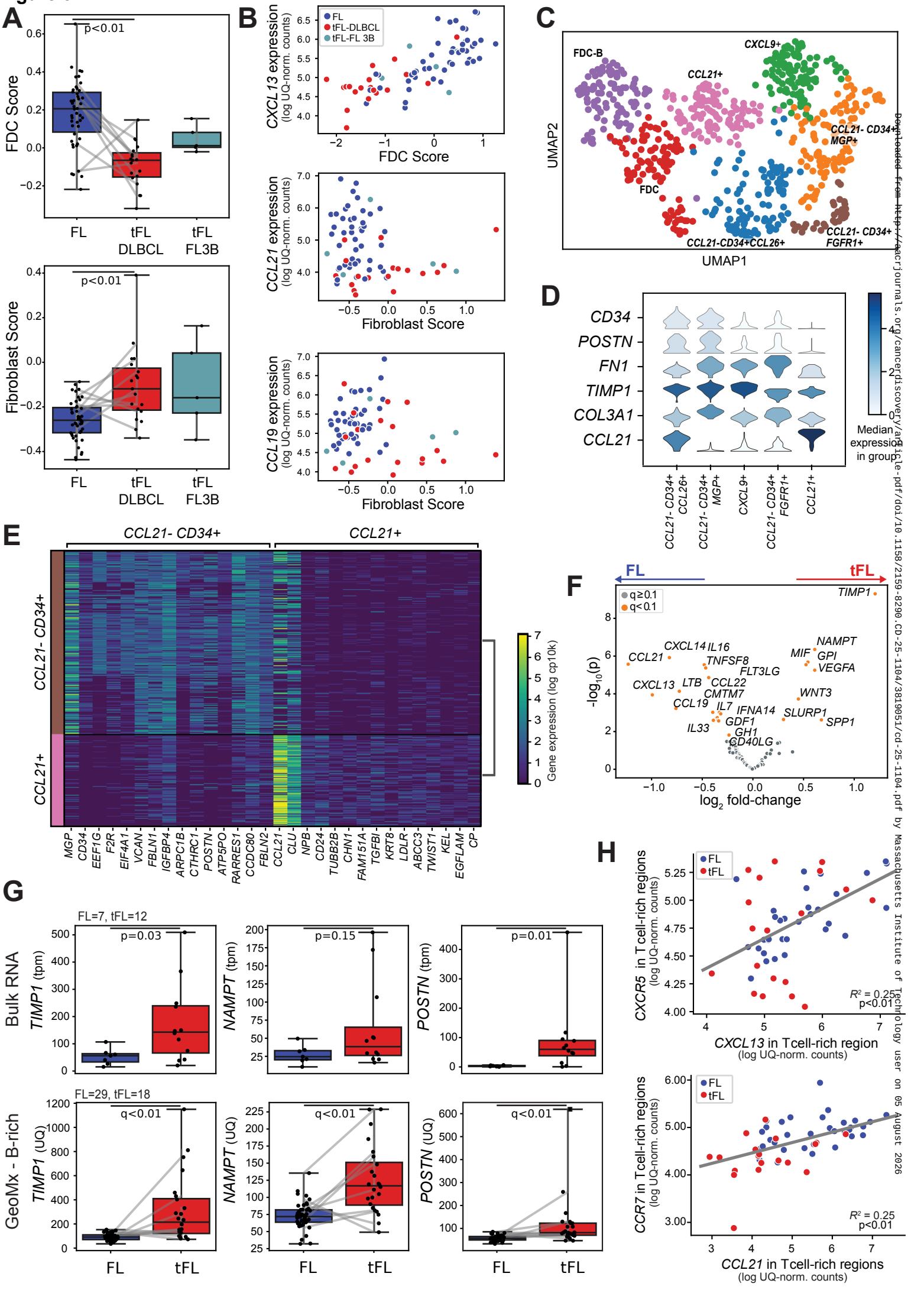
Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/11/8/2158/2159-8290.CD-25-1104/3819051/cd25-1104.pdf by Massachusetts Institute of Technology user on 3 August 2026

Figure 2

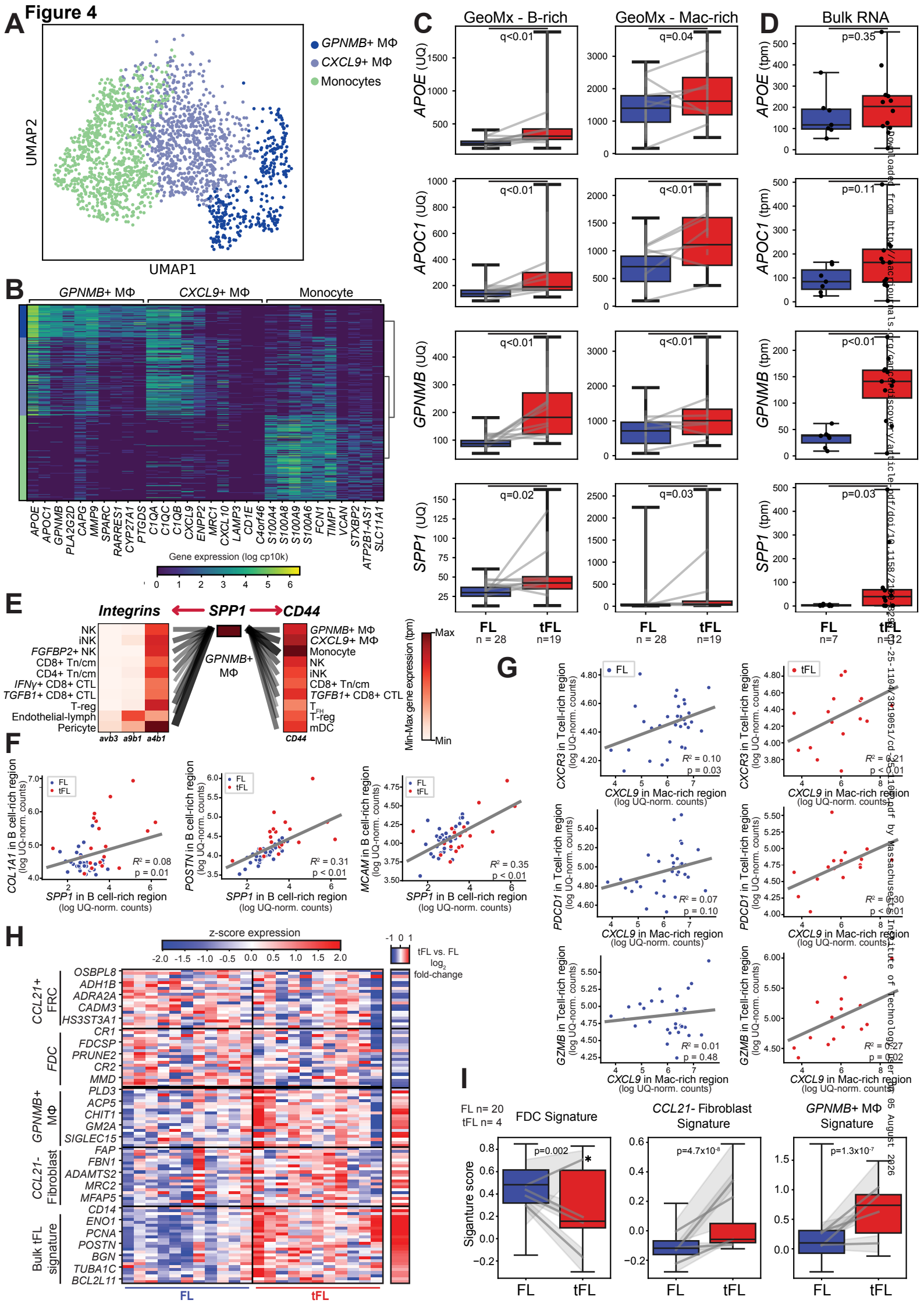


Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-1104-2025/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Figure 3

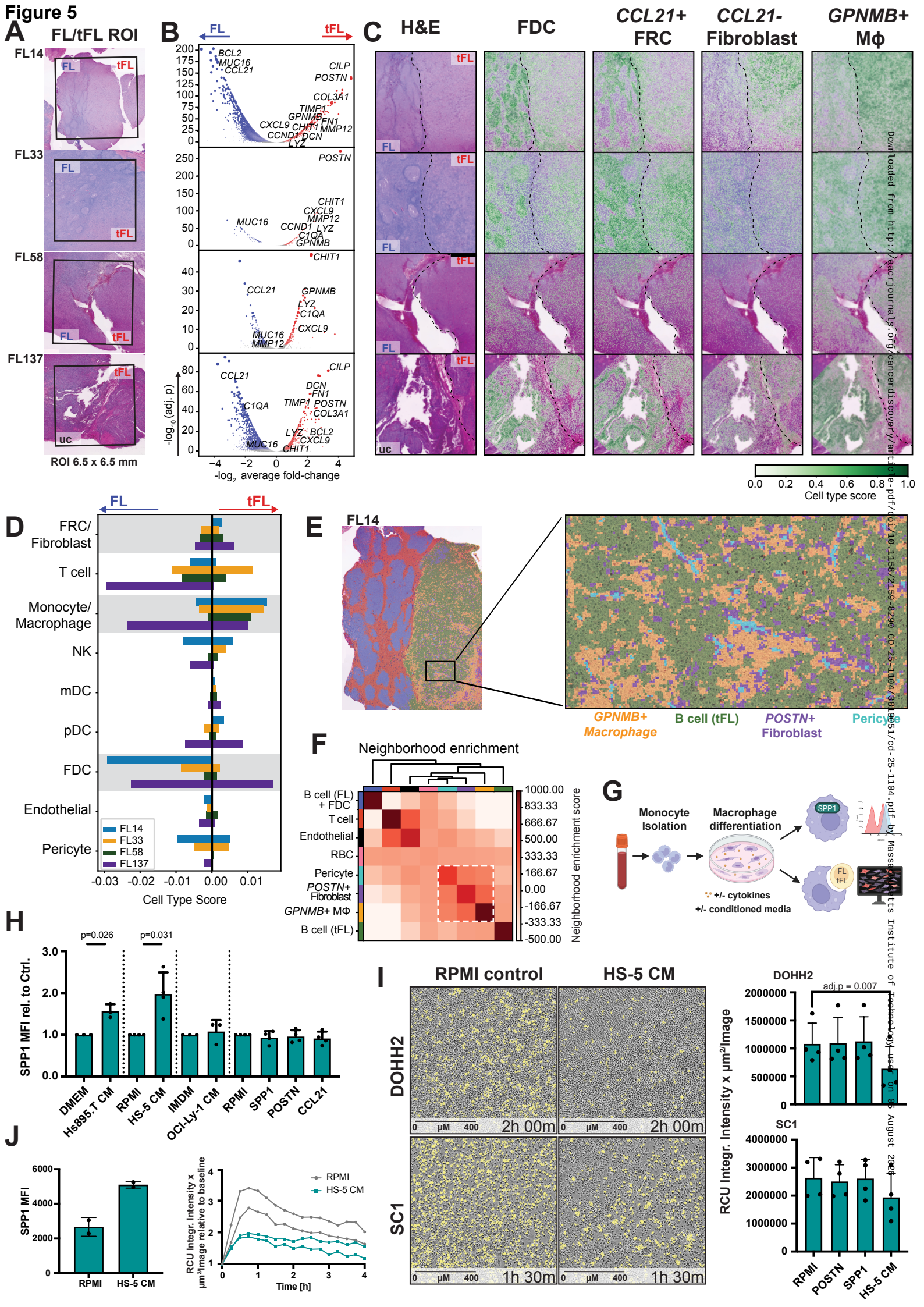


Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026



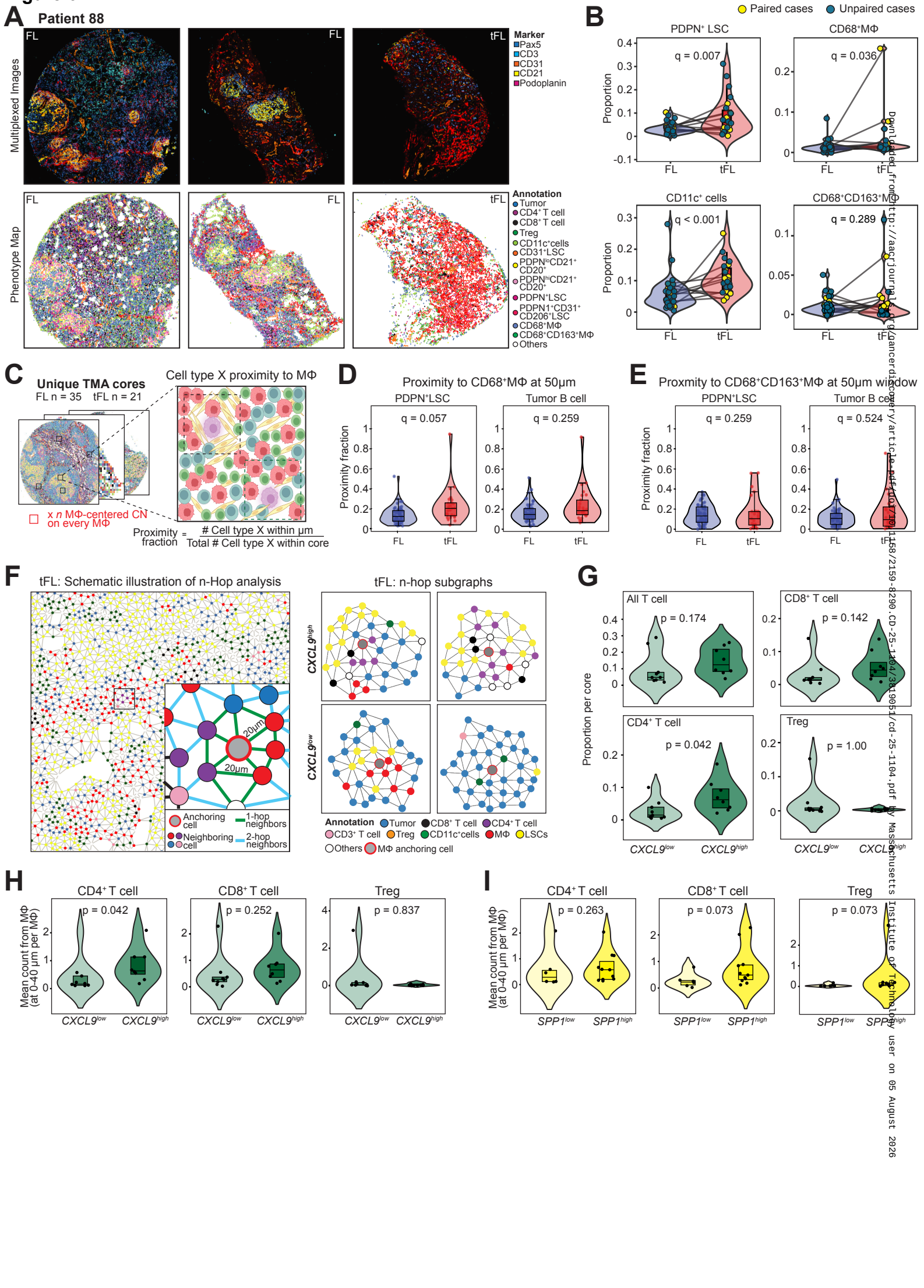
Downloaded from http://aacrjournals.org/cancerdiscov/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology User on 05 August 2026

Figure 5



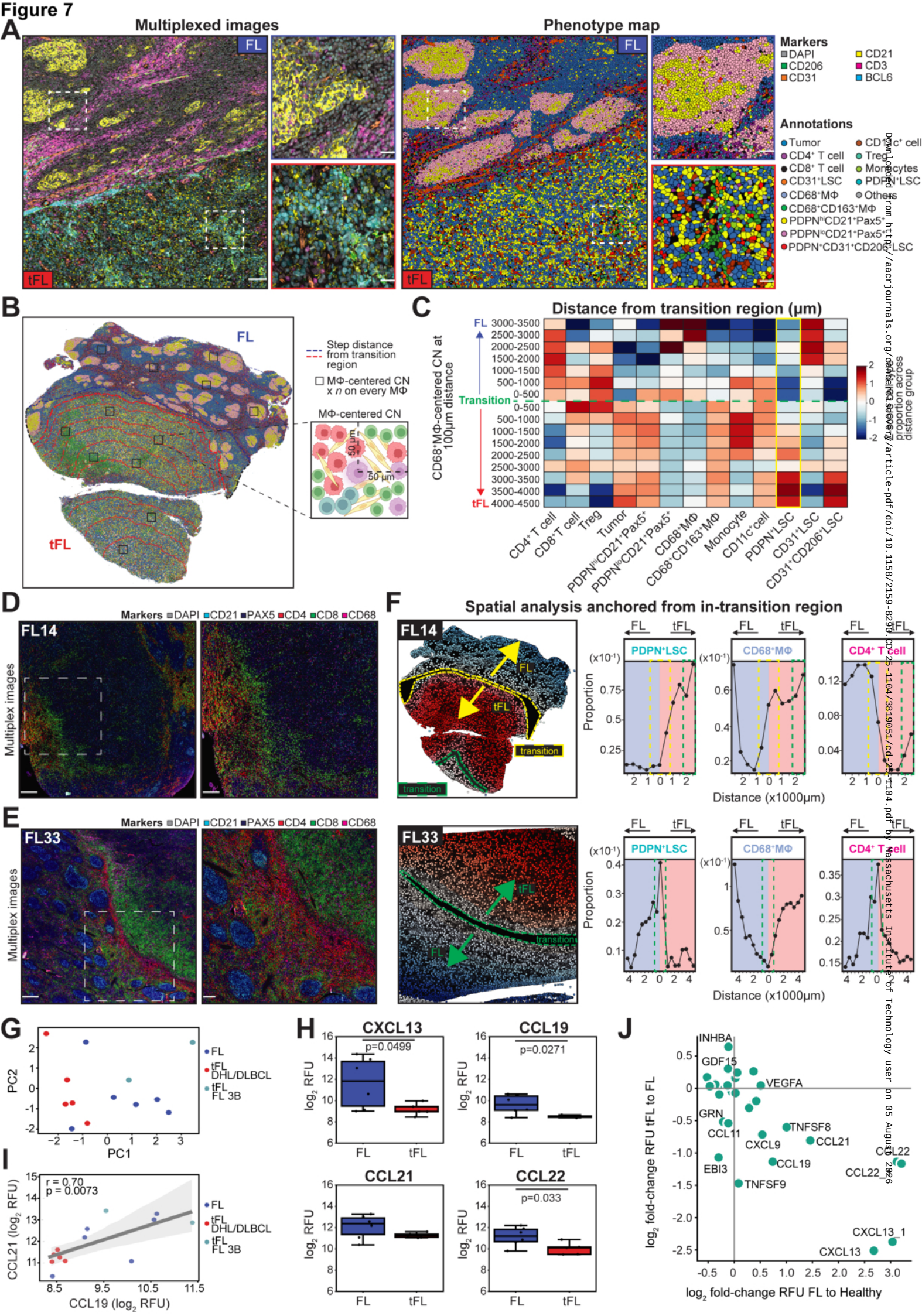
Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Figure 6



Downloaded from http://aacrjournals.org/cancerdiscovery/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3319051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026

Figure 7



Downloaded from http://aacrjournals.org/cancerres/article-pdf/doi/10.1158/2159-8290.CD-25-1104/3819051/cd-25-1104.pdf by Massachusetts Institute of Technology user on 05 August 2026