Single-cell RNA sequencing analyses of primary cutaneous B-cell disorders reveal distinct molecular patterns consistent with clinical behavior

Short Title: scRNA-seq analyses of primary cutaneous B-cell disorders

Johannes Griss MD, PhD,[1] Mathias C. Drach MD,[1] Vy Nguyen MSc,[1] Jasmine P. Levine BA,[2,3] Felix M. Thaler,[1] Marco A. Medjimorec,[1] Lisa E. Shaw MSc,[1] Ulrike Mann,[1] Wolfgang Weninger MD,[1] Christine Wagner,[1] Stephan N. Wagner MD,[1] Ingrid Simonitsch-Klupp MD,[4] Matthias Farlik PhD,[1] Constanze Jonak MD,[1] Patrick M. Brunner MD, MSc[1,2]

[1] Department of Dermatology, Medical University of Vienna, Austria

[2] Department of Dermatology, Icahn School of Medicine at Mount Sinai, New York, NY, USA

[3] New York Medical College, Valhalla, NY, USA

[4] Department of Pathology, Medical University of Vienna, Austria

Corresponding Authors:

Patrick M. Brunner, MD, MSc

Associate Professor

Department of Dermatology

Icahn School of Medicine at Mount Sinai

One Gustave L Levy Place

New York, NY 10029

Email: patrick.brunner@mountsinai.org

Johannes Griss, MD, PhD
Associate Professor
Department of Dermatology
Medical University of Vienna
Währinger Gürtel 18-20
1090 Vienna, Austria
Email: johannes.griss@meduniwien.ac.at
Telephone: +43 (0) 1 40400 77020
FAX: +43 (0) 1 40400 76990

Word Count: 3,999
Tables: 1
Figures: 7
Reference Count: 42

Abbreviations

MZLPD	Marginal zone lymphoproliferative disorder
MZL	Marginal zone lymphoma
FCL	Follicle center lymphoma
DLBCL-LT	Diffuse large B-cell lymphoma, leg type
scRNA-seq	Single-cell RNA sequencing
TCR	T-cell receptor
BCL	B-cell lymphoma
PSEUDO	Pseudolymphoma
HC	Healthy control

Key Points

- This is the first comprehensive molecular map of primary cutaneous B-cell lymphoproliferative disorders.
- Differences between entities were primarily geared by the expanded B-cell clone and less by other immune or stromal cells.

Abstract

Cutaneous B-cell lymphomas (CBL) comprise a diverse group of diseases with variable clinical course and prognosis. The cellular and molecular mechanisms underlying the pathogenesis of CBL remain ill-defined. Here we performed single-cell RNA sequencing combined with B-cell and T-cell receptor sequencing in three distinct CBL entities. In contrast to benign cutaneous pseudolymphomas and healthy skin, each CBL variant contained a single dominant B-cell clone. In marginal zone lymphoproliferative disorder (MZLPD), expanded clones generally showed terminal differentiation along a physiological trajectory towards plasma cells and displayed transcriptomic profiles indistinguishable from the benign polyclonal bystander B-cells. Follicle center lymphoma (FCL) clones also exhibited a canonical B-cell marker profile but consisted of more homogeneous populations of proliferating and non-proliferating germinal center (GC) B-cell-like phenotypes. In the most aggressive CBL, diffuse large B-cell lymphoma, leg type (DCBCL-LT), cells revealed multiple non-physiological marker profiles as well as aberrant proliferative activity beyond conventional GC-like cells. Fibroblasts in MZLPD, FCL and pseudolymphomas, but not DLBCL-LT, expressed the classical chemotactic chemokine *CXCL13.* These data suggest that in cutaneous B-cell disorders other than DLBCL-LT, physiological immune circuits are primarily operative, consistent with their overall indolent behavior.

Introduction

Primary cutaneous B-cell lymphomas comprise a heterogeneous group of extranodal non-Hodgkin lymphomas.[1] The three most common subtypes include primary cutaneous marginal zone lymphoproliferative disorder[2] (MZLPD, also named primary cutaneous marginal zone lymphoma),[3] primary cutaneous follicle center lymphoma (FCL), and primary cutaneous diffuse large B-cell lymphoma, leg type (DLBCL-LT).[1] MZLPD and FCL are highly indolent with a 5-year disease-specific survival of >95%, while DLBCL-LT is associated with an aggressive disease course as reflected by a 5-year survival rate of less than 60%.[4]

Previous gene expression analyses identified characteristic B-cell phenotypes in each of these diseases. MZLPD B cells were found to be most similar to plasma cells, FCL showed a more germinal center (GC)-like phenotype, and DLBCL-LT lesions were most consistent with activated B-cells.[5-7] A key challenge is that, in B-cell lymphomas, clonality is not always detectable using routine molecular methods, or when a surrogate such as light-chain restriction is used.[8] In those instances, the differentiation from so-called pseudolymphomas can be difficult, as they are usually defined as non-clonal B-cell proliferations.[9] Additionally, some authors suggest that especially MZLPD may develop from pre-existing, non-neoplastic lymphoid infiltrates,[10] adding another level of complexity to the diagnosis of cutaneous B-cell lymphomas. To distinguish these from pseudolymphomas, skin flow cytometry was found to have superior sensitivity than immunohistochemistry, *in situ* hybridization or immunoglobulin heavy chain gene rearrangement examinations. Nevertheless, its routine use is difficult as it requires fresh tissue, is prone to technical challenges during processing, and is not widely available.[9] Thus, we urgently need a better cellular and molecular characterization of primary cutaneous B-cell neoplasms and

pseudolymphomas to improve our understanding of their pathophysiology, and identify diagnostic and prognostic markers for improved patient care.

Here, we present an in-depth, single-cell RNA sequencing-based characterization of primary cutaneous B-cell lymphomas and pseudolymphomas. By combining this with B and T-cell receptor sequencing, we were able to clearly differentiate each condition on a molecular level.

Materials & Methods

Patient recruitment and sample processing

Patients were recruited at the Medical University of Vienna, Austria. Skin punch biopsies were taken after obtaining written informed consent, under a protocol approved by the Ethics Committee of the Medical University of Vienna (EK 1360/2018). After processing using the Skin Dissociation Kit by Miltenyi Biotech (Bergisch Gladbach, Germany),[11-13] single-cell suspensions were subjected to scRNA-seq using the Chromium Single Cell Controller and Single Cell 5' Library & Gel Bead Kit v1.1 (10X Genomics, Pleasanton, CA) according to the manufacturer's protocol. TCR and BCR sequences were enriched from cDNA using the VDJ Kit workflow by 10X Genomics. Sequencing was performed using the Illumina NovaSeq platform and the 150bp paired-end configuration.

Data analysis

Preprocessing of the scRNA-seq data was performed using Cell Ranger version 6.1.2 (10x Genomics). Raw sequencing files were demultiplexed using the Cell Ranger command "mkfastq." Each sample was aligned to the human reference genome assembly "refdata-gex-GRCh38-2020-A" using the Cell Ranger command "count".

TCR and BCR sequences were processed using CellRanger's "vdj" command. CellRanger's clonotype assignments were used as a basis. Clonotypes that shared the same CDR3 sequence in one chain and did not differ in any other chain were merged.

Expression data was processed using R (version 4.2.1),[14] Seurat (version 4.0.5),[15] and monocle3 (version 1.0.0).[16] Plots were generated using Seurat and ggplot2 (version 3.3.5).[17] As initial quality control, cells with <500 genes or >20% mitochondrial genes were removed. Subsequently, *MALAT1*, mitochondrial and ribosomal genes, as well as immunoglobulin variable cluster and TCR variable genes were removed. Next, basic normalization and sample-specific bias removal was performed using Seurat's scTransform on all samples from one disease group. For this step, the "method" parameter was set to "glmGamPoi" and "conserver.memory" was enabled. Variables to regress were set to % mitochondrial genes and the sample identifier. Disease group data was subsequently integrated using Seurat's reciprocal principal component approach.[15] Integration markers were identified using the first 30 principal components and the number of anchors ("k.anchor") set to 20. Dataset integration was subsequently performed using the first 30 principial components. Clustering was performed using Seurat's "FindNeighbours" function to construct the nearest neighborhood graph, followed by the "FindClusters" function to identify distinct clusters. For the completely integrated dataset, this was done using the first 15 principial components for the non-lymphocytic subclustering and 30 principal components for the B-cell and T-cell subclustering, and a resolution of 0.5 for the T-cell clustering and 0.8 for the B cell and non-lymphocytic ones. Certain cell populations were further subclustered to identify smaller populations. These included NK cells/ILCs, proliferating T cells, *CD4*+ T cells, and all major non-lymphocytic cell types. Here, the subclustering was performed by extracting the respective cells and re-running the PCA, the neighborhood graph creation, and the identification of clusters. This step used the first 15 principal components and a resolution of 0.1-0.5 depending on subset size.

Clustering results were then re-integrated into the original subclustering. Differential expression analysis was performed using Seurat's FindMarkers function and the data in the "SCT" slot.

The trajectory/pseudotime analysis was performed using monocle3.[16] Here, the Seurat object was converted into monocle's cds object. Clusters were identified based on the UMAP embedding. Subsequently terminally differentiated plasma cells were used as the starting point to derive the pseudotime, which was then inverted in order to match the canonical B-cell differentiation path.

Multiplex Immunohistochemistry

Multiplex immunostainings were conducted as previously described.[18] Briefly, 4µm sections were deparaffinized and antigen retrieval was performed in heated citrate buffer (pH 6.0) and/or Tris-EDTA buffer (pH 9) for 30 min. Thereafter, sections were fixed with 7.5% neutralized formaldehyde (SAV Liquid Production GmbH, Flintsbach am Inn, Germany). Each section was subjected to 6 successive rounds of antibody staining, consisting of protein blocking with 20% normal goat serum (Dako, Glostrup, Denmark) in PBS, incubation with primary antibodies, biotinylated anti-mouse/rabbit secondary antibodies and Streptavidin-HRP (Dako), followed by TSA visualization with fluorophores Opal 520, Opal 540, Opal 570, Opal 620, Opal 650, and Opal 690 (PerkinElmer, Waltham, MA, USA) diluted in 1X Plus Amplification Diluent (PerkinElmer), Ab-TSA complex-stripping in heated citrate buffer (pH 6.0) and/or Tris-EDTA buffer (pH 9) for 30 minutes, and fixation with 7.5% neutralized formaldehyde. Thereafter, nuclei were counterstained with DAPI (PerkinElmer), and sections were mounted with PermaFluor fluorescence mounting medium (Thermo Fisher Scientific, Waltham, MA, USA). Multiplexed slides were scanned on a Vectra

Multispectral Imaging System version 2 (Perkin Elmer). Briefly, a library from spectral peaks emitted by each fluorophore from single stained slides was created with the Advanced Image Analysis software (InForm 2.4, PerkinElmer) and used for spectral unmixing of multispectral images allowing for identification of all markers of interest. Autofluorescence was determined on an unstained representative study sample. After adaptive cell segmentation, cut-off values for each fluorophore/antibody staining were defined. All phenotyping and subsequent quantifications were performed blinded to the sample identity.

Data Sharing Statement

Complete gene level expression and VDJ sequencing data of all newly analyzed samples will be made available on GEO. Data from healthy control samples in a previously published dataset from our laboratory are available under GEO accession number GSE173205.[11]

Results

Single-cell RNA sequencing map of primary cutaneous B-cell lymphomas, pseudolymphomas and healthy control individuals

Using single-cell RNA sequencing (scRNA-seq), we profiled 178,647 cells from skin biopsies of patients with clinical and histopathological diagnoses of primary cutaneous B-cell lymphomas, namely MZLPD (97,680 cells, n=7 patients), FCL (14,107 cells, n=1 patient), DLBCL-LT (13,387 cells, n=3 biopsies from 2 patients), in comparison to patients with idiopathic nodular B-cell pseudolymphomas[19] (PSEUDO, 36,163 cells, n=4 biopsies from 3 patients) and healthy control subjects (NHS, 17,310 cells, n=4 individuals, Table 1, Table S1). Clustering analyses revealed diverse skin cell populations according to expression of canonical cell markers, including large numbers of B and T cells, as well as smaller populations of NK cells, myeloid dendritic cells (DCs), plasmacytoid dendritic cells (pDCs), macrophages (Macro), fibroblasts (FB), keratinocytes (KRT), smooth muscle cells (SMC), melanocytes (Mel), and blood and lymphatic endothelial cells (EC, LEC), (Figure 1A-C). While most cell types were found at varying numbers in all 5 groups, B cells and pDCs were essentially absent in NHS samples (Figure 1D). Further subclustering of all B cells (Figure 1E) revealed a diverse pool of differentiation stages, from naïve B cells (*IGHD*+), GC cells (*CD38+AICDA*+) to fully differentiated plasma cells (*CD27+CD38+SDC1+IGHG1+IGHG4+)* (Figure 1E-F).

B cells from DLBCL-LT showed distant, separate clustering from other entities, while FCL was largely confined to GC-like cells (Figure 1G). By contrast, MZLPD B cells were found across various different cell types, largely overlapping with pseudolymphomas (Figure 1G). In order to characterize B-cell clonality of our

samples, we combined scRNA-seq profiling with B-cell receptor (BCR) sequencing and depicted the single most abundant clone for each sample, colored for each disease group (Figure 1H). High top clone numbers were found in DLBCL-LT, FCL and MZLPD, consistent with a clonal disease, while numbers were much lower in pseudolymphoma (Figure 1H). While the top clones of DLBCL-LT and FCL showed distinct clustering, those from MZLPD and pseudolymphomas were largely overlapping (Figure 1H). These data suggest a close transcriptomic relation between MZLPD and pseudolymphomas, when compared to FCL and DLBCL-LT.

MZLPD shows distinctively expanded clones in contrast to pseudolymphoma samples

To better delineate potential differences between MZLPD and pseudolymphomas, we clustered each of these samples separately (Figure 2A-B). In contrast to pseudolymphomas (Figure 2A), MZLPD showed distinct clusters of the single most abundant clone in all samples (Figure 2B). Consistently, top clone numbers were considerably higher in MZLPD than in pseudolymphomas (Figure 2C), and the gap from the first to the second most abundant clone was much larger in MZLPD than pseudolymphoma samples (Figure 2D). This confirms the concept that MZLPD is a clonal process, while pseudolymphomas show a generally polyclonal B-cell receptor pattern, despite otherwise strong transcriptomic similarities.

Expanded clones of MZLPD span several B-cell subsets

Given these similarities beyond the clonal expansion, and to facilitate a direct comparison, we separately subclustered B-cells of MZLPD and pseudolymphomas (Figure 3A). This subclustering allowed us to phenotypically identify eleven different

B-cell subtypes (Figure 3A-B, Table S2). We found five clusters of naïve B cells, as defined by expression of *CD19, MS4A1* (CD20), and *IGHD*, while being negative for *CD27* and *CD38* (Figure 3B). All naïve B cells expressed low, but noticeable levels of the B-cell activation gene *CD83*. Naïve 1 and 2 were further characterized by up-regulation of MHC II genes, indicating early activation.[20] Naïve 3 showed an additional up-regulation of heat shock proteins and thus likely represent stressed activated B cells in the dataset.[21] Naïve 4 exhibited strong *VPREB3* expression, a known B-cell differentiation marker[22] that we also found in some GC B cells (Figure 3B). Therefore, naïve 4 might be most consistent with activated cells differentiating towards GC B cells. Naïve 5 were *ATM*+ and partly *BCL6*+, both associated with V(D)J recombination.[23] Their embedding in close vicinity to memory cells (Figure 3A) suggests very late *IGHD*+ cells in the process of class switching. Besides this large pool of naïve B cells, we observed activated B cells with stronger *CD83* expression that further harbored the anti-apoptotic NF-κB target transcription factor gene *BCL2A1* and the chemokine ligand *CCL4*.[24] We were able to identify two distinct populations of *CD38*+*AICDA*+ GC B cells: *BCL6*+ cells (GC1), and an *MKI67*+ population of proliferating B cells (GC2)[25] (Figure 3A-B). We also found a comparably large fraction of *CD27*+, *IGHD*+ and *IGHG1*^{low} B cells which most likely represent unswitched memory B cells, and, together with switched memory B cells (*CD27*+*IGHE*+*IGHG4*+) represent marginal zone B cells.[24] Finally, plasma cells (*CD27*+*CD38*+*IGHG1*+*IGHG4*+*SDC1/CD138*+) formed a clearly separated cluster (Figure 3A-B). Importantly, clonally expanded B cells from MZLPD samples did not form a distinct population but spanned several cell clusters (Figure 3C). These top abundant MZLPD clones were primarily found in (i) *CD27*+ memory B cells, (ii) plasma cells and (iii) GC B cells (Figure 3C). While frequencies of expanded clones

increased towards more differentiated B-cell clusters, the opposite was found within the polyclonal infiltrate (Figure 3D). However, the cluster distribution of these clones showed considerable patient-to-patient variation (Figure 3E, Table S3), a finding consistent with the pathological observation that some, but not all, MZLPD lesions are dominated by plasma cell-like B cells, which have previously been classified as cutaneous immunocytomas.[26] While most MZLPD clones showed a switch towards IgG isotypes, sample 92 was characterized by *IGHM* and *IGHE* expression (Figure 3F), consistent with the previously described rare form of IgM-dominant / uncommon class-switched MZLPD.[2] The expanded clone of sample 92 formed three distinct groups: two as part of the “switched memory” cluster and one as part of plasma cells (Figure 3E). The cluster adjacent to “GC2” had a clearly proliferative component that was *MS4A1* (CD20) negative (Figure 3G), a phenotype that is not typically seen in physiological B-cell development and differentiation.[27] Similarly, the second group in the “switched memory” fraction also contained an *MS4A1* negative portion (Figure 3G). This kind of aberrant phenotype was not present in any other MZLPD sample (not shown), and might explain the potentially more aggressive behavior of non-switched or non-IgG switched MZL.[28,29] Generally, though, differential gene expression analysis between expanded clones and the polyclonal B-cell infiltrate in MZLPD samples did not reveal major differences. The analysis resulted in only 24 differentially expressed genes (DEGs) within the GC1 cluster, followed by as little as 13 DEGs in switched memory B cells (Figure 3H, Table S4). These DEGs consisted of down-regulated MHC II genes (*HLA-DRB1, HLA-DRA, HLA-DQA1, HLA-DQB1*) and an up-regulation of *IGLC3* within the expanded clone, corresponding to the observed dominance of lambda chain positive lymphomas in our dataset (Figure 2B,

Table S4). Taken together, these minimal differences indicate that clonally expanded B cells were very similar to the respective subsets of polyclonal B cells.

To further assess the distribution of clonally expanded B cells on this differentiation path, we used Monocle3 to place all B cells on a pseudotime scale (Figure 3I). This put naïve 1 (MHC II^{high}) and activated B cells at the beginning of the differentiation scale, followed by naïve 2. Naïve 4 cells were similarly positioned as GC cells, followed by proliferating GC cells (Figure 3I). Here, clustering positioned some switched memory B cells close to late GC B cells, suggesting that these late GC types differentiate towards memory B cells / early plasma cells. Naïve 5 were put on a similar pseudotime section as some memory B cells, which matches their assumed up-regulation of genes associated with V(D)J recombination.[23] Therefore, this pseudotime trajectory approximated the known B-cell differentiation path.[27] Importantly, cells of the expanded clone tightly followed this trajectory, with increasing canonical differentiation markers such as *CD27, IGHE, IGHG1,* and *IGHG4* during pseudotime (Figure 3J). In each sample, these clonally expanded cells were distributed over a wide range of the pseudotime trajectory and ranged from naïve B cells to terminally differentiated plasma cells (Figure 3K). This indicates that expanded clones of MZLPD generally follow the normal B-cell differentiation path and terminally differentiate within the skin.

In contrast to MZLPD, expanded FCL clones only consisted of GC-like B cells

We next assessed whether other forms of primary cutaneous B-cell lymphomas also followed conventional B-cell differentiation trajectories. Clustering analysis of a biopsy from a primary cutaneous follicle center lymphoma (FCL) patient showed GC-like cells to dominate the B-cell population (Figure 4A-C). In contrast to MZLPD,

clonally expanded B cells were strictly confined to proliferating and non-proliferating *CD38+, AICDA+* GC B cells (Figure 4B). By contrast, activated B cells only consisted of polyclonal cells (Figure 4B). Similarly, a published FCL sample from a study on the effect of oncolytic virotherapy in cutaneous B-cell lymphoma[30] revealed the dominant clone also to be limited to proliferating and non-proliferating GC clusters, corroborating our results (Figure 4D-G). From these data, we conclude that the expanded clones from primary cutaneous FCLs are confined to the histopathologically observed GC forming B cells.[31]

DLBCL-LT showed multiple aberrant B-cell phenotypes

We further profiled 3 samples from 2 DLBCL-LT patients (Table 1, Figure 5A). In contrast to expanded clones of MZLPD (Figure 3) and FCL (Figure 4), DLBCL-LT cells showed marker combinations that did not follow physiological B-cell phenotypes (Figure 5B). Virtually all clonally expanded cells (Figure 5C) showed expression of *CD27,* with widespread loss of *MS4A1* (CD20, Figure 5D). There was no expression of late-stage differentiation markers like *IGHE, IGHG1, IGHG4*, and *SDC1* (CD138) (Figure 5D). The large fraction of proliferating (*MKI67+)* cells included *MS4A1-* and *CD38-* cells (Figure 5D), which do not typically exist under physiological conditions.[27] These findings were replicated in a published DLBCL-LT sample[30] (Figure S1A-D). There, malignant cells were again *CD27+,* with fractions showing a loss of *MS4A1*, and proliferating cells spanning several B-cell phenotypes (Figure S1D). Plasma cells and canonical *IGHD+* naïve B cells were only found in the polyclonal B-cell fraction (Figure S1D). In both datasets, malignant clones were not confined to specific cell clusters, but were found across all B-cell clusters at high proportions, while polyclonal cell counts were generally much lower in number (Figure 5E, S1).

This was in stark contrast to FCL and MZLPD, where polyclonal B cells were consistently present at higher frequencies than the top expanded B-cell clone (Figure 5F). The presence of considerably proliferating and dedifferentiated cells that do not correspond to canonical B-cell stages is very consistent with the clinical observation that DLBCL-LT is a more aggressive entity than FCL or MZLPD.

In contrast to B cells, major T-cell populations did not differ between disease groups

Since T cells constituted the largest group of cells in our dataset, even outnumbering B-cell counts in most samples (Table S1), we further subclustered this population (Figure S2A). T-cell clusters consisted of diverse subpopulations of *CD4*+ helper T cells, *CD8A*+ cytotoxic T cells including a *CCL4*+*IFNG*+ population, NKT cells, smaller clusters of *FOXP3*+ regulatory T cells, *NR4A2*+*TNFRSF18*+*TNFRSF4*+ activated T cells, exhausted *CD4*+*PDCD1*+*TIGIT*+ T cells, *MKI67*+ proliferating T cells, *CD3D-KLRD1*+ NK cells and *CD3D-KLRD1-KLRB1*+ innate lymphoid cells (ILCs) (Figure S2B, Table S5). In contrast to B cells that showed several non-overlapping clusters depending on the disease type (Figure 1G), T lymphoid clusters were generally present in all conditions, albeit at lower levels and with less T-cell proliferation in NHS samples (Figure S2C). Overall, differential gene expression did not reveal major differences among groups, except for increased expression of cytotoxic molecules in DLBCL-LT compared to all other conditions (Table S6). When combining scRNA-seq with TCR sequencing, we found a generally polyclonal TCR pattern, with frequencies of top expanded clones in the range of NHS (Figure S2D). In contrast to the *CD4* dominance in NHS, lymphomas and pseudolymphomas showed the top clones to be primarily present in *CD8A*+ cytotoxic populations, which

might be consistent with a cytotoxic anti-lymphomatous immune reaction[32] (Figure 7D). Nevertheless, we did not detect T-cell populations uniquely expanded in individual disease groups, as opposed to B cells.

Multiplex immunohistochemistry showed a diverse B-cell infiltrate in MZLPD and pseudolymphomas in contrast to DLBCL-LT

As current scRNA-seq methods can be prone to sampling bias, we assessed B cell and T cell populations by using a multiplex immunohistochemistry (IHC) approach previously established at our center.[18] This panel allowed us to simultaneously characterize naive-like B cells (CD19+/CD20+, CD27-, CD38-, CD138-), GC-like B cells (CD19+/CD20+, CD27-, CD38+, CD138-), memory-like B cells (CD19+/CD20+, CD27+, CD38-, CD138-), plasma cells (CD19+, C20-, CD138+), and T cells (CD19-/CD20-, CD5+). Overall, MZLPD and pseudolymphoma samples harbored a diverse lymphocytic infiltrate, as opposed to DLBCL-LT (Figure 6A-B). While we observed all B-cell subtypes in pseudolymphomas and MZLPD samples, several DLBCL-LT samples only consisted of naive-like CD19+ and CD20+ B cells. We further quantified phenotypic diversity of the infiltrate using the Shannon diversity index (Figure 6C). This confirmed a significantly larger heterogeneity of lymphocyte composition in pseudolymphomas and MZLPD samples compared to DLBCL-LT, paralleling indolent *vs.* aggressive clinical disease behavior.[33]

Indolent B-cell lymphomas and pseudolymphomas are characterized by a *CXCL13*+ fibroblast population

We also assessed non-lymphocytes and their distribution across diseases (Figure 7A-B, Table S7). As expected, DCs and macrophages were significantly less

frequent in NHS, consistent with a relative increase of fibroblast and melanocyte counts (Figure 7C). While total fibroblast numbers were comparable across lymphomas and pseudolymphomas, *CXCL13*+ fibroblasts were the only population consistently absent in all DLBCL-LT samples (Figure 7D-E, S3), cells known to be key to B-cell recruitment in the tumor microenvironment[34] and infection.[35,36] All other subsets showed no significant difference in cell number between disease groups (Figure S3).

We further performed differential gene expression analysis comparing each cell type of one disease against the respective cells of the other three subtypes (Figure 7F-G). Most differentially expressed genes in MZLPD and pseudolymphomas were found in keratinocytes (Figure 7F-G). However, these were driven by single samples and correlated with clinically visible inflammation rather than the disease subset (not shown). FCL was dominated by down-regulated genes (Figure 7F) which were rather caused by up-regulated genes in other samples (Table S8). The comparably large number of significantly up-regulated genes in DLBCL-LT (Figure 7G) was driven by B cell and cell division associated genes, most likely originating from lysed tumor cells leading to ambient RNA contamination.[37] The largest number of non-B cell-associated genes was found in DCs, macrophages, and fibroblasts (Figure 7F-G), mainly including inflammation-associated genes (Table S8). On closer inspection, we found that these differences were largely driven by single samples (not shown). Therefore, the major consistent disease specific difference on a cell population level was the lack of *CXCL13*+ fibroblasts in DLBCL-LT, suggesting a potentially critical role in disease-specific B cell recruitment for these cells.

Discussion

By using scRNA-seq analyses of primary cutaneous B-cell disorders, we were able to detect disease-specific features, primarily on the level of expanded B-cell clones. Transcriptomic properties of other cells, including T cells and stromal cells, were surprisingly homogeneous, without consistent differences across disease groups. In line with DLBCL-LT's aggressive clinical behavior, DLBCL-LT clones consisted of highly aberrant B cells, in contrast to all other conditions investigated, which largely showed cellular phenotypes mimicking physiological cell populations. Specifically, the B-cell infiltrate in FCL was dominated by an expanded clone of relatively homogeneous GC B cells. By contrast, MZLPD clones spanned multiple B-cell phenotypes, indicating that these cells follow a physiological B-cell differentiation path towards end-differentiated plasma cell-like B cells that no longer proliferate. This observation supports the recently suggested re-classification of IgG class-switched marginal zone lymphoma as a lymphoproliferative disorder,[2] except for a small subset of IgM+/non-canonically class-switched MZLPD, which might still be regarded as a lymphoma due to frequent extracutaneous involvement.[28,29,38] In line, we found a single *IGHM*+ MZLPD patient to show aberrant phenotypic features that were absent in class-switched MZLPD. Nevertheless, expanded clones of MZLPD could phenotypically not be differentiated from polyclonal B cells, further suggesting their close correlation with physiological phenotypes.[28,29,38-41] This might explain why many studies struggled to identify the putatively malignant clone in MZLPD,[30] which can make the differentiation from so-called pseudolymphomas particularly difficult.[19] For this reason, there is an ongoing debate whether some pseudolymphomas might merely represent lymphomas that have not yet been correctly diagnosed due to lack of sensitivity of routine diagnostic methods.[42] However, we consistently found that

pseudolymphomas harbored a clearly polyclonal B-cell pattern, in contrast to MZLPD. Nevertheless, longitudinal studies will be necessary to investigate whether these patients will eventually develop overt lymphomas, or whether they retain their polyclonal, non-malignant characteristics.

The fact that expanded clones in FCL and MZLPD samples did not differ significantly from physiological, polyclonal B cells might also suggest a generally physiological mechanism of recruitment of these cells to the skin. Interestingly, we detected a distinct population of *CXCL13*+ fibroblasts in FCL, MZLPD, and pseudolymphomas, but not in DLBCL-LT or healthy control skin. By expressing this prominent B-cell attracting chemokine, these fibroblasts might have a role in chemotaxis of the observed B cell infiltrates in pseudolymphomas and indolent lymphoma groups,[34-36] but not DCBCL-LT. In line, DLBCL-LT showed a much less diverse B-cell infiltrate than pseudolymphomas or MZLPD and was rather dominated by monomorphous B-cell populations. Taken together, this investigation is the first comprehensive molecular map of primary B-cell lymphoproliferative disorders in humans, which might help guide future treatment approaches for these skin conditions.

Acknowledgements

This work was funded by a research grant to PMB from the Austrian Science Fund (grant number KLI 849-B). SNW was supported by a research grant from the Austrian Science Fund (grant number P31127). JG was supported by a research grant from the Austrian Science Fund (P35937).

Authorship contributions

Conceptualization, C.J., J.G., P.M.B.; Software and Formal Analysis, J.G., V.N.; Visualization: J.G., J.P.L., F.M.T.; Investigation, C.J., C.W., J.G., I.S.K., L.E.S., M.A.M., M.C.D., M.F., P.M.B., S.N.W., U.M.; Resources, P.M.B., M.F., W.W.; Writing – Original Draft C.J., J.G., P.M.B.; Writing – Review & Editing by all co-authors; Project Administration & Supervision, C.J., J.G., M.F.; P.M.B., Funding Acquisition, P.M.B.

Disclosure of Conflicts of Interest

JG received personal fees from AbbVie and Novartis. CJ has received personal fees from Boehringer Ingelheim, LEO, Pfizer, Recordati Rare Diseases, Eli Lilly, Novartis, Takeda, Kyowa Kirin, Amgen, Sandoz, UCB, BMS, AbbVie, Janssen, and Almirall. CJ is an investigator for Eli Lilly, Novartis, 4SC, and Innate. WW has received personal fees from LEO Pharma, Pfizer, Sanofi Genzyme, Eli Lilly, Novartis, Boehringer Ingelheim, AbbVie, and Janssen. PMB has received personal fees from Almirall, Sanofi, Janssen, Amgen, LEO Pharma, AbbVie, Pfizer, Boehringer Ingelheim, GSK, Regeneron, Eli Lilly, Celgene, Arena Pharma, Novartis, UCB Pharma, Biotest and BMS. PMB is an investigator for Pfizer. MF is involved in an

ongoing collaboration with Böhringer Ingelheim. The rest of the authors declare that they have no relevant conflicts of interest.

References

1. Willemze R, Cerroni L, Kempf W, et al. The 2018 update of the WHO-EORTC classification for primary cutaneous lymphomas. Blood 2019;133(16):1703-1714. DOI: 10.1182/blood-2018-11-881268.
2. 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. DOI: 10.1182/blood.2022015851.
3. 14]. (WHO classification of tumours series, 5th ed.; vol. 11). Available from: https://tumourclassification.iarc.who.int/chapters/63. WCoTEBHtIbvaopLFIAfRoCcD.
4. Vitiello P, Sica A, Ronchi A, Caccavale S, Franco R, Argenziano G. Primary Cutaneous B-Cell Lymphomas: An Update. Front Oncol 2020;10:651. DOI: 10.3389/fonc.2020.00651.
5. Lima M. Cutaneous primary B-cell lymphomas: from diagnosis to treatment. An Bras Dermatol 2015;90(5):687-706. DOI: 10.1590/abd1806-4841.20153638.
6. Storz MN, van de Rijn M, Kim YH, Mraz-Gernhard S, Hoppe RT, Kohler S. Gene expression profiles of cutaneous B cell lymphoma. J Invest Dermatol 2003;120(5):865-70. DOI: 10.1046/j.1523-1747.2003.12142.x.
7. Hoefnagel JJ, Vermeer MH, Jansen PM, et al. Primary cutaneous marginal zone B-cell lymphoma: clinical and therapeutic features in 50 cases. Arch Dermatol 2005;141(9):1139-45. DOI: 10.1001/archderm.141.9.1139.
8. Nicolay JP, Wobser M. Cutaneous B-cell lymphomas - pathogenesis, diagnostic workup, and therapy. J Dtsch Dermatol Ges 2016;14(12):1207-1224. DOI: 10.1111/ddg.13164.

9. Nakagawa Y, Hamada T, Takahashi T, et al. Analysis of clonality in cutaneous B-cell lymphoma and B-cell pseudolymphoma using skin flow cytometry: Comparison of immunophenotyping and gene rearrangement studies. J Dermatol 2021. DOI: 10.1111/1346-8138.16057.

10. Hristov AC, Comfere NI, Vidal CI, Sundram U. Kappa and lambda immunohistochemistry and in situ hybridization in the evaluation of atypical cutaneous lymphoid infiltrates. J Cutan Pathol 2020;47(11):1103-1110. DOI: 10.1111/cup.13858.

11. Rindler K, Jonak C, Alkon N, et al. Single-cell RNA sequencing reveals markers of disease progression in primary cutaneous T-cell lymphoma. Mol Cancer 2021;20(1):124. DOI: 10.1186/s12943-021-01419-2.

12. Bangert C, Rindler K, Krausgruber T, et al. Persistence of mature dendritic cells, TH2A, and Tc2 cells characterize clinically resolved atopic dermatitis under IL-4Ralpha blockade. Sci Immunol 2021;6(55). DOI: 10.1126/sciimmunol.abe2749.

13. Rojahn TB, Vorstandlechner V, Krausgruber T, et al. Single-cell transcriptomics combined with interstitial fluid proteomics defines cell type-specific immune regulation in atopic dermatitis. J Allergy Clin Immunol 2020;146(5):1056-1069. DOI: 10.1016/j.jaci.2020.03.041.

14. R: A Language and Environment for Statistical Computing RCT, 2022, https://www.R-project.org.

15. Hao Y, Hao S, Andersen-Nissen E, et al. Integrated analysis of multimodal single-cell data. Cell 2021;184(13):3573-3587 e29. DOI: 10.1016/j.cell.2021.04.048.

16. Qiu X, Mao Q, Tang Y, et al. Reversed graph embedding resolves complex single-cell trajectories. Nat Methods 2017;14(10):979-982. DOI: 10.1038/nmeth.4402.

17. Wickham H. ggplot2 - Elegant Graphics for Data Analysis. ISBN 9783319242774 2016;2nd Edition.

18. Griss J, Bauer W, Wagner C, et al. B cells sustain inflammation and predict response to immune checkpoint blockade in human melanoma. Nat Commun 2019;10(1):4186. DOI: 10.1038/s41467-019-12160-2.

19. Mitteldorf C, Kempf W. Cutaneous pseudolymphoma-A review on the spectrum and a proposal for a new classification. J Cutan Pathol 2020;47(1):76-97. DOI: 10.1111/cup.13532.

20. Lankar D, Vincent-Schneider H, Briken V, Yokozeki T, Raposo G, Bonnerot C. Dynamics of major histocompatibility complex class II compartments during B cell receptor-mediated cell activation. J Exp Med 2002;195(4):461-72. DOI: 10.1084/jem.20011543.

21. Jonak C, Klosner G, Trautinger F. Heat shock proteins in the skin. Int J Cosmet Sci 2006;28(4):233-41. DOI: 10.1111/j.1467-2494.2006.00327.x.

22. Felizola SJ, Katsu K, Ise K, et al. Pre-B Lymphocyte Protein 3 (VPREB3) Expression in the Adrenal Cortex: Precedent for non-Immunological Roles in Normal and Neoplastic Human Tissues. Endocr Pathol 2015;26(2):119-28. DOI: 10.1007/s12022-015-9366-7.

23. Mensah KA, Chen JW, Schickel JN, et al. Impaired ATM activation in B cells is associated with bone resorption in rheumatoid arthritis. Sci Transl Med 2019;11(519). DOI: 10.1126/scitranslmed.aaw4626.

24. King HW, Orban N, Riches JC, et al. Single-cell analysis of human B cell maturation predicts how antibody class switching shapes selection dynamics. Sci Immunol 2021;6(56). DOI: 10.1126/sciimmunol.abe6291.

25. Cui C, Wang J, Fagerberg E, et al. Neoantigen-driven B cell and CD4 T follicular helper cell collaboration promotes anti-tumor CD8 T cell responses. Cell 2021;184(25):6101-6118 e13. DOI: 10.1016/j.cell.2021.11.007.

26. Duncan LM, LeBoit PE. Are primary cutaneous immunocytoma and marginal zone lymphoma the same disease? Am J Surg Pathol 1997;21(11):1368-72. DOI: 10.1097/00000478-199711000-00012.

27. Weisel NM, Joachim SM, Smita S, et al. Surface phenotypes of naive and memory B cells in mouse and human tissues. Nat Immunol 2022;23(1):135-145. DOI: 10.1038/s41590-021-01078-x.

28. Fernandez-Flores A. Is there a narrow connection between the two subsets of cutaneous MALT lymphomas and the dynamics of the follicle? Am J Dermatopathol 2013;35(2):283-4. DOI: 10.1097/DAD.0b013e31825aa352.

29. Carlsen ED, Swerdlow SH, Cook JR, Gibson SE. Class-switched Primary Cutaneous Marginal Zone Lymphomas Are Frequently IgG4-positive and Have Features Distinct From IgM-positive Cases. Am J Surg Pathol 2019;43(10):1403-1412. DOI: 10.1097/PAS.0000000000001363.

30. Ramelyte E, Tastanova A, Balazs Z, et al. Oncolytic virotherapy-mediated anti-tumor response: a single-cell perspective. Cancer Cell 2021;39(3):394-406 e4. DOI: 10.1016/j.ccell.2020.12.022.

31. Schrader AMR, de Groen RAL, Willemze R, et al. Cell-of-origin classification using the Hans and Lymph2Cx algorithms in primary cutaneous large B-cell

lymphomas. Virchows Arch 2022;480(3):667-675. DOI: 10.1007/s00428-021-03265-5.

32. Jiang X, Xu J, Liu M, et al. Adoptive CD8(+) T cell therapy against cancer:Challenges and opportunities. Cancer Lett 2019;462:23-32. DOI: 10.1016/j.canlet.2019.07.017.

33. Tsakiroglou AM, Astley S, Dave M, et al. Immune infiltrate diversity confers a good prognosis in follicular lymphoma. Cancer Immunol Immunother 2021;70(12):3573-3585. DOI: 10.1007/s00262-021-02945-0.

34. Rodriguez AB, Peske JD, Woods AN, et al. Immune mechanisms orchestrate tertiary lymphoid structures in tumors via cancer-associated fibroblasts. Cell Rep 2021;36(3):109422. DOI: 10.1016/j.celrep.2021.109422.

35. Denton AE, Innocentin S, Carr EJ, et al. Type I interferon induces CXCL13 to support ectopic germinal center formation. J Exp Med 2019;216(3):621-637. DOI: 10.1084/jem.20181216.

36. Kogame T, Nomura T, Kataoka T, et al. Possible inducible skin-associated lymphoid tissue (iSALT)-like structures with CXCL13(+) fibroblast-like cells in secondary syphilis. Br J Dermatol 2017;177(6):1737-1739. DOI: 10.1111/bjd.15349.

37. Young MD, Behjati S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. Gigascience 2020;9(12). DOI: 10.1093/gigascience/giaa151.

38. Edinger JT, Kant JA, Swerdlow SH. Cutaneous marginal zone lymphomas have distinctive features and include 2 subsets. Am J Surg Pathol 2010;34(12):1830-41. DOI: 10.1097/PAS.0b013e3181f72835.

39. van Maldegem F, van Dijk R, Wormhoudt TA, et al. The majority of cutaneous marginal zone B-cell lymphomas expresses class-switched immunoglobulins and develops in a T-helper type 2 inflammatory environment. Blood 2008;112(8):3355-61. DOI: 10.1182/blood-2008-01-132415.

40. Swerdlow SH. Cutaneous marginal zone lymphomas. Semin Diagn Pathol 2017;34(1):76-84. DOI: 10.1053/j.semdp.2016.11.007.

41. Guitart J. Rethinking primary cutaneous marginal zone lymphoma: shifting the focus to the cause of the infiltrate. J Cutan Pathol 2015;42(9):600-3. DOI: 10.1111/cup.12509.

42. Charli-Joseph YV, Gatica-Torres M, Pincus LB. Approach to Cutaneous Lymphoid Infiltrates: When to Consider Lymphoma? Indian J Dermatol 2016;61(4):351-74. DOI: 10.4103/0019-5154.185698.

Tables

Subject ID		Age	Sex	Race	Histopathological diagnosis	Location	Disease duration (years)	Ongoing treament	Previous treamtents	Disease stage
112		51	F	White	Healthy control	Trunk	n.a.	n.a.	n.a.	n.a.
115		48	M	White	Healthy control	Trunk	n.a.	n.a.	n.a.	n.a.
116		56	F	White	Healthy control	Trunk	n.a.	n.a.	n.a.	n.a.
121		44	F	White	Healthy control	Trunk	n.a.	n.a.	n.a.	n.a.
92		72	M	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Upper arm	8 years	None	Rituximab	T1bN0M0
104		33	M	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Trunk	3 years	None	Clarithromycin	T1bN0M0
110		47	M	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Trunk	13 years	None	Rituximab	T3aN0M0
129		38	M	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Shoulder	14 years	None	Clarithromycin, surgery	T3aN0M0
146		47	M	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Upper arm	1 year	None	Surgery	T1aN0M0
172		72	W	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Lower thigh	2 years	None	Surgery, radiotherapy	T1bN0M0
WB3		60	M	White	Primary cutaneous marginal zone lymphoma/lymphoproliferative disorder	Neck	1 year	None	None	T1bN0M0
99 + 159		30	M	White	Nodular cutaneous pseudolymphoma	Upper arms	1 year	None	None	T3aN0M0
134A		73	M	White	Nodular cutaneous pseudolymphoma	Scalp	11 years	None	Rituximab	T2cN0M0
166		47	F	White	Nodular cutaneous pseudolymphoma	Face	16 years	None	Glucocorticosteroids, hydroxychloroquine, apremilast	T2cN0M0
198		77	M	White	Primary cutaneous follicle center lymphoma	Scalp	1 year	None	None	T2aN0M0
206		74	M	White	Primary cutaneous diffuse large cell lymphoma, leg type	Lower leg	<1 year	None	None	T2bN0M0
117 + 207		78	F	White	Primary cutaneous diffuse large cell lymphoma, leg type	Forearm and lower leg	<1 year	None	Radiotherapy, R-CHOP	T3aN0M0

Table 1. Patient baseline characteristics at time of sampling.

Figure Legends

Figure 1. Fig. 1: scRNA-seq map of B-cell lymphoma, pseudolymphoma and healthy control samples. (A) UMAP of unsupervised clustering of 178,647 cells integrated from primary cutaneous marginal zone lymphoproliferative disorder (MZLPD, n=7 patients), primary cutaneous follicle center lymphoma (FCL, n=1 patient), primary cutaneous diffuse large B-cell lymphoma leg-type (DLBCL-LT, n=3 biopsies from 2 patients), in comparison to patients with idiopathic nodular B-cell pseudolymphoma (PSEUDO, n=4 biopsies from 3 patients) and healthy control subjects (NHS, n=4 individuals) according to similarity of transcriptome, resulting in 14 different color-coded clusters. **(B)** Feature plots of individual patient groups. **(C)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(D)** Relative distribution of major cell types across disease groups. **(E)** Sub-clustering of the entire B-cell population. **(F)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers of the subclustered B-cell population. **(G)** Feature plot of the B-cell population colored according to the disease group. **(H)** Feature plot displaying the top abundant B-cell clone of each sample, color coded for each disease group. UMAP: Uniform Manifold Approximation and Projection.

Figure 2. Distribution and quantification of the single most abundant clone within individual MZLPD and pseudolymphoma samples. (A-B) Feature plots of (A) pseudolymphoma and (B) MZLPD samples clustered individually, without sample integration. Red: cells with the single most abundant BCR from each sample; single

point mutations of the BCR sequence were considered to originate from the same clone; blue: cells with polyclonal BCR; grey: no BCR detected. **(C)** Absolute numbers of the single most abundant clone per sample and group, horizontal line denotes median. **(D)** Absolute numbers of the first 10 most abundant clones per denoted sample. UMAP: Uniform Manifold Approximation and Projection. BCR: B-cell receptor.

Figure 3. Subclustering of B cells from MZLPD, pseudolymphoma and NHS samples. (A) UMAP of unsupervised sub-clustering of 51,723 cells integrated from initial B-cell clusters of primary cutaneous marginal zone lymphoma (MZL, n=7), pseudolymphoma (n=4) and healthy control individuals (n=4) according to similarity of transcriptome, resulting in 11 different color-coded clusters. **(B)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(C)** Combined UMAP of cell clusters from MZLPD samples colored according to clonality of BCR. Red: cells with most abundant monoclonal BCR from each sample; blue: cells with polyclonal BCR; grey: no BCR detected. **(D)** Absolute numbers of clonally expanded B cells (red) and polyclonal B cells (blue) per cell cluster of MZLPD samples. **(E)** Feature plots of cell clusters colored according to clonality of BCR, showing individual MZLPD samples. **(F)** Dot plots displaying average gene expression (red color) and frequency (circle size) of selected immunoglobulin markers within the top abundant clone of each MZLPD sample. **(G)** Feature plots of sample #92 showing expression distribution for selected marker genes. Expression levels for each cell are color-coded and overlaid onto UMAP plots. Red color intensity represents the respective level of expression. **(H)** Number of DEGs within each cluster, comparing clonally expanded with

polyclonal B cells from MZLPD samples. Genes that were expressed in at least two samples per group and had an adjusted p value ≤0.01 and an absolute average log2 fold-change ≥0.8 were considered significantly regulated. **(I)** Monocle3 pseudotime trajectory within all B cell subclusters. **(J)** Normalized expression of selected canonical B cell developmental genes within top expanded clones of MZLPD samples according to pseudotime buckets from 1.99 to 12.4; cells are color-coded according to clustering. **(K)** Positioning of individual cells of the most abundant clone for MZLPD samples according to the pseudotime trajectory; cells are color-coded according to clustering. UMAP: Uniform Manifold Approximation and Projection. BCR: B-cell receptor.

Figure 4. Follicle center lymphoma (FCL) shows germinal center clustering of the expanded clone. (A) UMAP of unsupervised clustering of 14,107 cells from a follicle center lymphoma sample according to similarity of transcriptome resulting in 12 different color-coded clusters. **(B)** Combined UMAP of cells colored according to clonality of BCR. Red: cells from the top expanded BCR; blue: cells with polyclonal BCR; grey: no BCR detected. **(C)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(D)** UMAP of unsupervised clustering of 11,110 cells from the P12 follicle center lymphoma sample from Ramelyte et al.[30] **(E)** UMAP of cells from sample P12 with cells colored according to the clonality of the BCR. **(F)** Dot plots of canonical markers for the identified clusters from sample P12. **(G)** Absolute number of BCR positive cells from sample P12 per cell type, split by cells from the dominant clone and the remaining polyclonal infiltrate.

Figure 5. Diffuse large B cell lymphoma samples show aberrant B cell phenotypes. (A) UMAP of unsupervised clustering of 13,387 cells integrated from 3 samples of diffuse large B-cell lymphoma, leg type, resulting in 18 different color-coded clusters. **(B)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(C)** Combined UMAP of cell clusters colored for the most abundant BCR+ clone per sample. **(D)** Feature plots of expression distribution for selected marker genes. Expression levels for each cell are color-coded and overlaid onto UMAP plots. Red color intensity represents the respective level of expression. **(E)** Absolute numbers of clonally expanded B cells (red, yellow, purple) and polyclonal B cells (blue) per B-cell cluster and sample. **(F)** Percentages of expanded clones (red) vs the polyclonal B-cell infiltrate (blue) for each depicted sample.

Figure 6. Multicolor immunofluorescence assessment of diversity of B-cell subpopulations in pseudolymphoma, MZLPD, and DLBCL-LT samples. (A) Relative number of B-cell phenotypes per sample based on multicolor immunofluorescence quantification. **(B)** Representative pictures of multicolor immunohistochemistry staining of pseudolymphoma, MZLPD and DLBCL-LT samples. **(C)** Shannon entropy per sample to estimate the diversity of the cell compositions between the disease groups. DLBCL-LT samples showed a significantly reduced diversity compared to pseudolymphoma and MZLPD.

Figure 7. *CXCL13*+ fibroblasts are absent in DLBCL-LT samples. (A) UMAP of unsupervised clustering of the 35,457 non-lymphoid cells integrated from all samples resulting in 19 different color-coded clusters. **(B)** Dot plots of cell clusters displaying

average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(C)** Combined UMAP of cell clusters split by sample group. **(D)** Absolute numbers of *CXCL13*+ fibroblasts for all disease groups. **(E)** Absolute numbers of the total fibroblast population for all disease groups. Upper and lower hinges correspond to the first and third quartiles. Upper and lower whiskers extend to the largest value within a maximum of 1.5 times the inter-quartile range. Colored points represent individual values. **(F-G)** Numbers of differentially expressed genes (F) downregulated or (G) upregulated per disease group within the depicted cluster.

Single-cell RNA sequencing analyses of primary cutaneous B-cell disorders reveal distinct molecular patterns consistent with clinical behavior

Supplementary Figure Legends

Figure S1. Published data of a DLBCL-LT sample confirms aberrant B-cell phenotypes within the malignant clone. (A) UMAP embedding of B cells extracted from sample P11 from Ramelyte et al. Cells are colored based on the BCR data with clonally expanded cells in red, polyclonal cells in green, and cells without detectable BCR in grey. **(B)** UMAP embedding with cells colored based on the clustering results and B cell phenotypes. **(C)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(D)** UMAP embedding with cells colored based on the normalized expression of *IGHD, CD38, MKI67*, and *CD27*.

Figure S2. T cell, NK and ILC populations across patient groups. (A) UMAP of unsupervised clustering of the 74,493 T lymphoid cells integrated from all samples, resulting in 10 different color-coded clusters. **(B)** Dot plots of cell clusters displaying average gene expression (red color) and frequency (circle size) of selected canonical cell type markers. **(C)** Combined UMAP of cell clusters split by sample group. **(D)** UMAP embedding with cells color coded based on TCR sequencing. Red points represent the most frequent clone, blue points represent polyclonal cells, and grey points represent cells with missing TCR information.

Figure S3. Absolute cell numbers of non-lymphoid cells per disease group. Upper and lower hinges correspond to the first and third quartiles. Upper and lower

whiskers extend to the largest value within a maximum of 1.5 time the inter-quartile range. Values outside this range are shown as individual points. Colored points represent individual values.

Supplementary Table Legends

Table S1. Absolute numbers of cells per sample and cell type, corresponding to Figure 1A.

Table S2. Absolute numbers of B-cell subtypes from MZLPD, pseudolymphomas, and NHS per sample and subtype, corresponding to Figure 3A.

Table S3. Absolute numbers of clonally expanded B cells from MZLPD samples per B-cell subtype, corresponding to Figure 3C.

Table S4. Differentially expressed genes per B-cell subtype comparing clonally expanded with polyclonal B cells from MZLPD samples, corresponding to Figure 3H. Genes that were expressed in at least two samples per group and had an adjusted p value ≤0.01 and an absolute average log2 fold-change ≥0.8 were considered significantly regulated.

Table S5. Cluster marker genes of individual T-cell/NK/ILC populations, corresponding to Figure S2.

Table S6. Differentially expressed genes per T-cell/NK/ILC subtype comparing sample groups, corresponding to Figure S2.

Table S7. Cluster marker genes of individual non-lymphoid populations, corresponding to Figure 7.

Table S8. Differentially expressed genes per non-lymphoid cell subtype comparing sample groups, corresponding to Figure 7.

Figure 1

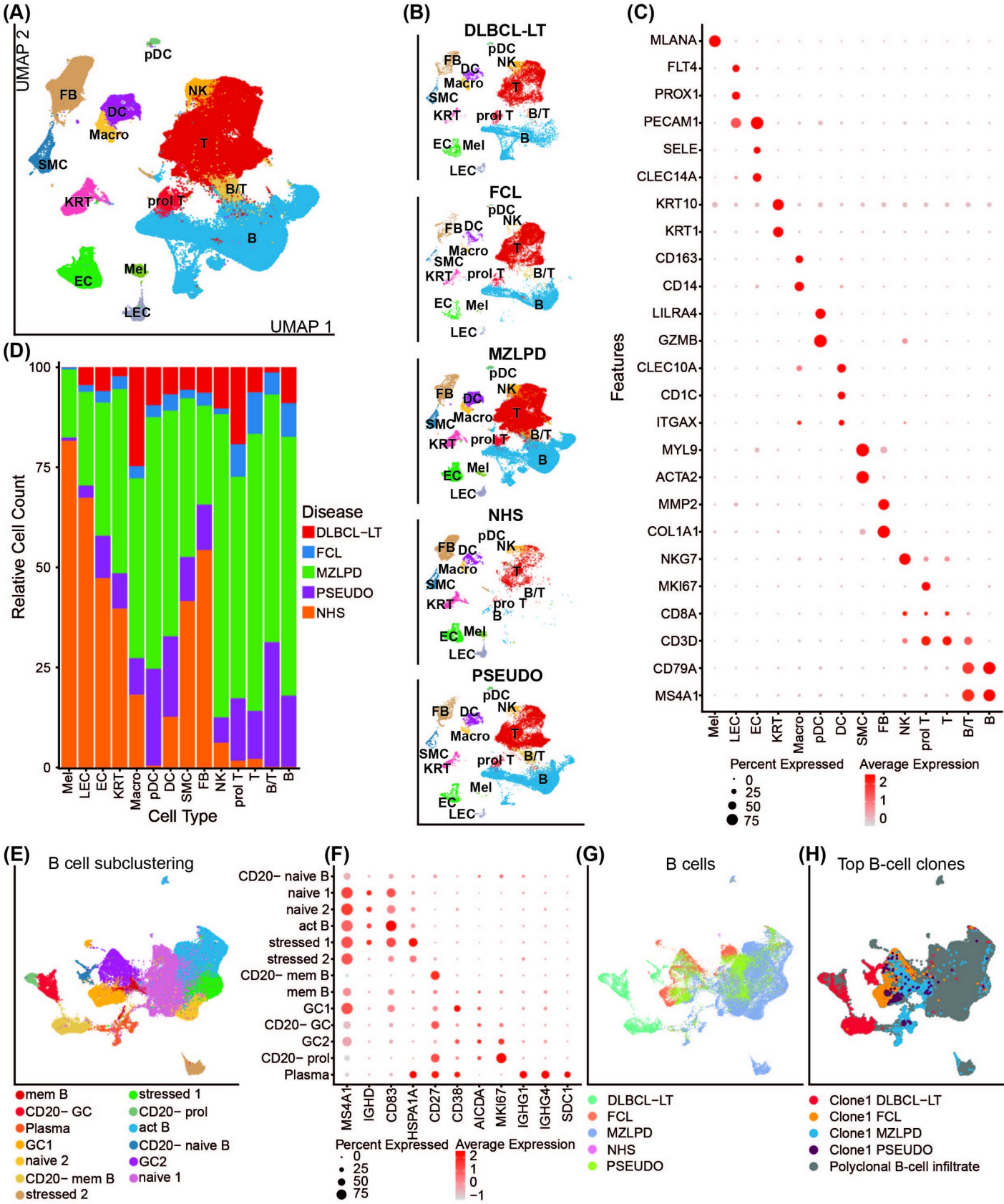

(A)
UMAP 2
UMAP 1
pDC
FB
NK
DC
Macro
T
SMC
KRT
prol T
B/T
B
EC
Mel
LEC
(B)
DLBCL-LT
FCL
MZLPD
NHS
PSEUDO
(C)
MLANA
FLT4
PROX1
PECAM1
SELE
CLEC14A
KRT10
KRT1
CD163
CD14
LILRA4
GZMB
CLEC10A
CD1C
ITGAX
MYL9
ACTA2
MMP2
COL1A1
NKG7
MKI67
CD8A
CD3D
CD79A
MS4A1
Features
Mel LEC EC KRT Macro pDC DC SMC FB NK prol T T B/T B
Percent Expressed
0 25 50 75
Average Expression
2 1 0
(D)
Relative Cell Count
100 75 50 25 0
Disease
DLBCL−LT
FCL
MZLPD
PSEUDO
NHS
Mel LEC EC KRT Macro pDC DC SMC FB NK prol T T B/T B
Cell Type
(E)
B cell subclustering
mem B
CD20− GC
Plasma
GC1
naive 2
CD20− mem B
stressed 2
stressed 1
CD20− prol
act B
CD20− naive B
GC2
naive 1
(F)
CD20− naive B
naive 1
naive 2
act B
stressed 1
stressed 2
CD20− mem B
mem B
GC1
CD20− GC
GC2
CD20− prol
Plasma
MS4A1 IGHD CD83 HSPA1A CD27 CD38 AICDA MKI67 IGHG1 IGHG4 SDC1
Percent Expressed
0 25 50 75
Average Expression
2 1 0 −1
(G)
B cells
DLBCL−LT
FCL
MZLPD
NHS
PSEUDO
(H)
Top B-cell clones
Clone1 DLBCL−LT
Clone1 FCL
Clone1 MZLPD
Clone1 PSEUDO
Polyclonal B-cell infiltrate

Figure 2

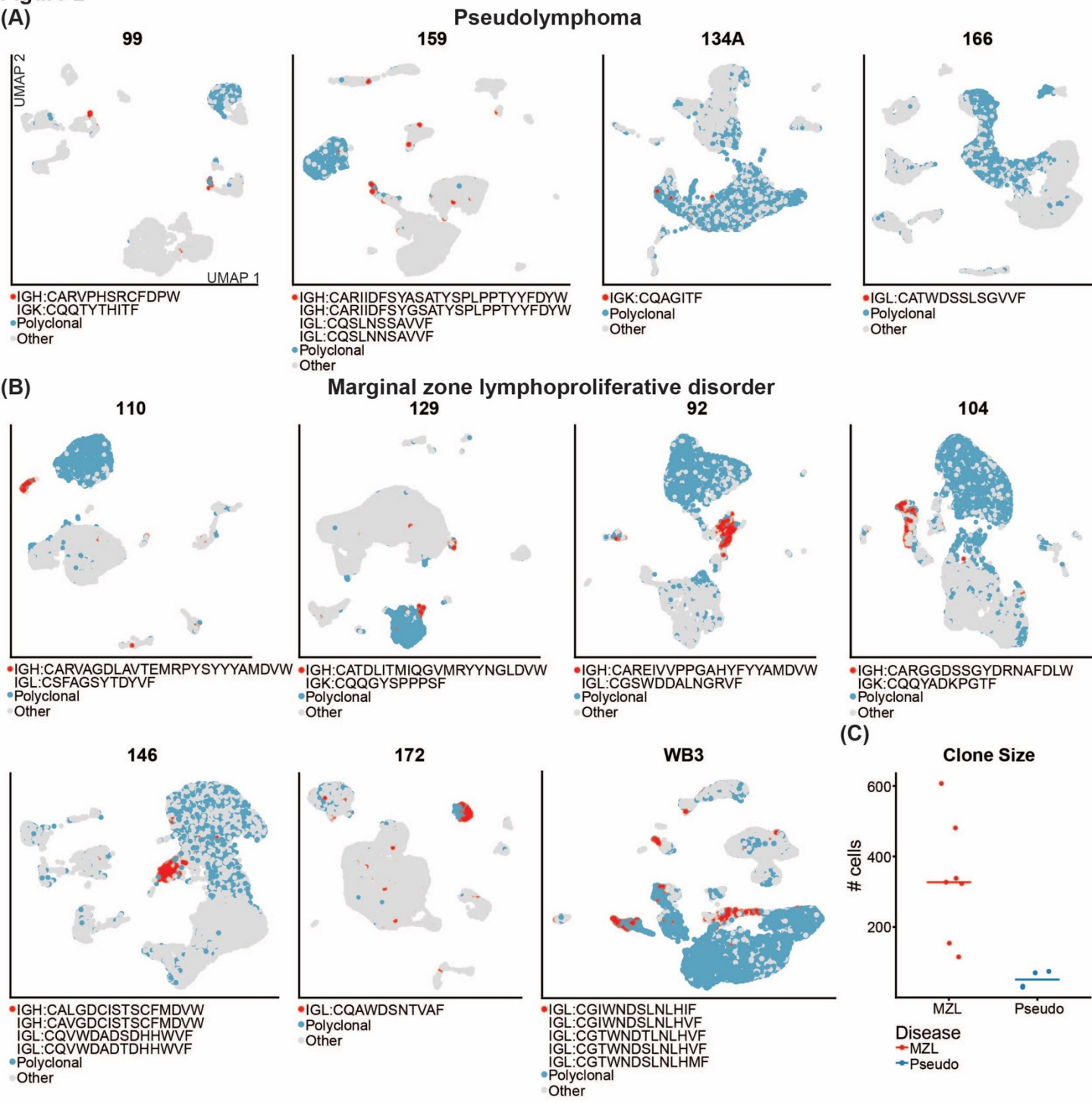

(A)
Pseudolymphoma
99
159
134A
166
UMAP 2
UMAP 1
IGH:CARVPHSRCFDPW
IGK:CQQTYTHITF
Polyclonal
Other
IGH:CARIIDFSYASATYSPLPPTYYFDYW
IGH:CARIIDFSYGSATYSPLPPTYYFDYW
IGL:CQSLNSSAVVF
IGL:CQSLNNSAVVF
Polyclonal
Other
IGK:CQAGITF
Polyclonal
Other
IGL:CATWDSSLSGVVF
Polyclonal
Other
(B)
Marginal zone lymphoproliferative disorder
110
129
92
104
IGH:CARVAGDLAVTEMRPYSYYYAMDVW
IGL:CSFAGSYTDYVF
Polyclonal
Other
IGH:CATDLITMIQGVMRYYNGLDVW
IGK:CQQGYSPPPSF
Polyclonal
Other
IGH:CAREIVVPPGAHYFYYAMDVW
IGL:CGSWDDALNGRVF
Polyclonal
Other
IGH:CARGGDSSGYDRNAFDLW
IGK:CQQYADKPGTF
Polyclonal
Other
146
172
WB3
IGH:CALGDCISTSCFMDVW
IGH:CAVGDCISTSCFMDVW
IGL:CQVWDADSDHHWVF
IGL:CQVWDADTDHHWVF
Polyclonal
Other
IGL:CQAWDSNTVAF
Polyclonal
Other
IGL:CGIWNDSLNLHIF
IGL:CGIWNDSLNLHVF
IGL:CGTWNDTLNLHVF
IGL:CGTWNDSLNLHVF
IGL:CGTWNDSLNLHMF
Polyclonal
Other
(C)
Clone Size
600
400
200
cells
MZL
Pseudo
Disease
MZL
Pseudo
(D)
MZLPD 104
MZLPD 110
MZLPD 129
MZLPD 146
MZLPD 172
MZLPD 92
MZLPD WB3
PSEUDO 134A
PSEUDO 159
PSEUDO 166
PSEUDO 99
600
400
200
0
cells

Figure 3

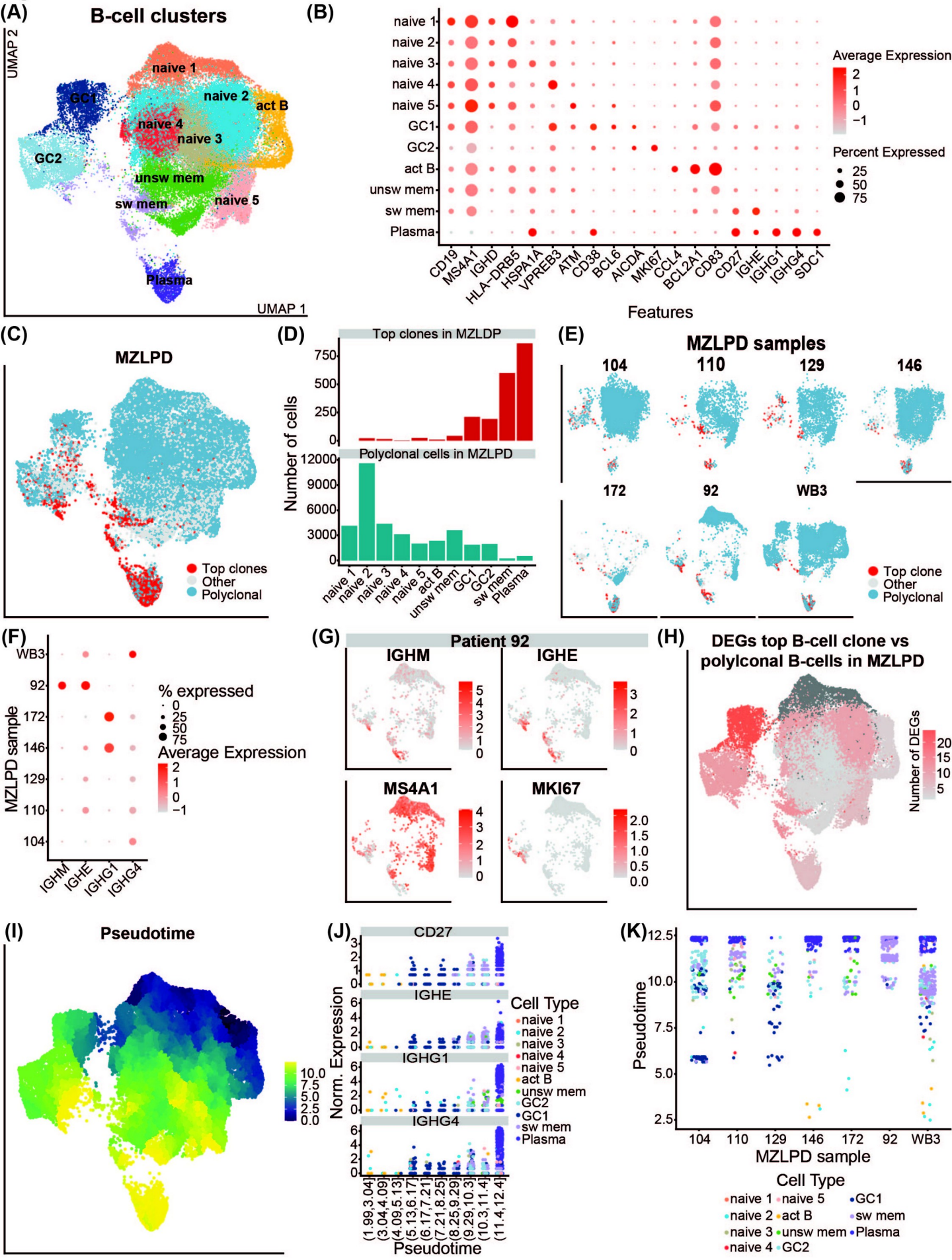

(A)
B-cell clusters
UMAP 2
UMAP 1
naive 1
naive 2
naive 3
naive 4
naive 5
GC1
GC2
act B
unsw mem
sw mem
Plasma
(B)
naive 1
naive 2
naive 3
naive 4
naive 5
GC1
GC2
act B
unsw mem
sw mem
Plasma
CD19
MS4A1
IGHD
HLA−DRB5
HSPA1A
VPREB3
ATM
CD38
BCL6
AICDA
MKI67
CCL4
BCL2A1
CD83
CD27
IGHE
IGHG1
IGHG4
SDC1
Features
Average Expression
Percent Expressed
(C)
MZLPD
Top clones
Other
Polyclonal
(D)
Top clones in MZLDP
Polyclonal cells in MZLPD
Number of cells
naive 1
naive 2
naive 3
naive 4
naive 5
act B
unsw mem
GC1
GC2
sw mem
Plasma
(E)
MZLPD samples
104
110
129
146
172
92
WB3
Top clone
Other
Polyclonal
(F)
MZLPD sample
WB3
92
172
146
129
110
104
IGHM
IGHE
IGHG1
IGHG4
% expressed
Average Expression
(G)
Patient 92
IGHM
IGHE
MS4A1
MKI67
(H)
DEGs top B-cell clone vs polylconal B-cells in MZLPD
Number of DEGs
(I)
Pseudotime
(J)
CD27
IGHE
IGHG1
IGHG4
Norm. Expression
Pseudotime
Cell Type
naive 1
naive 2
naive 3
naive 4
naive 5
act B
unsw mem
GC2
GC1
sw mem
Plasma
(K)
Pseudotime
MZLPD sample
104
110
129
146
172
92
WB3
Cell Type
naive 1
naive 2
naive 3
naive 4
naive 5
act B
unsw mem
GC2
GC1
sw mem
Plasma

Figure 4

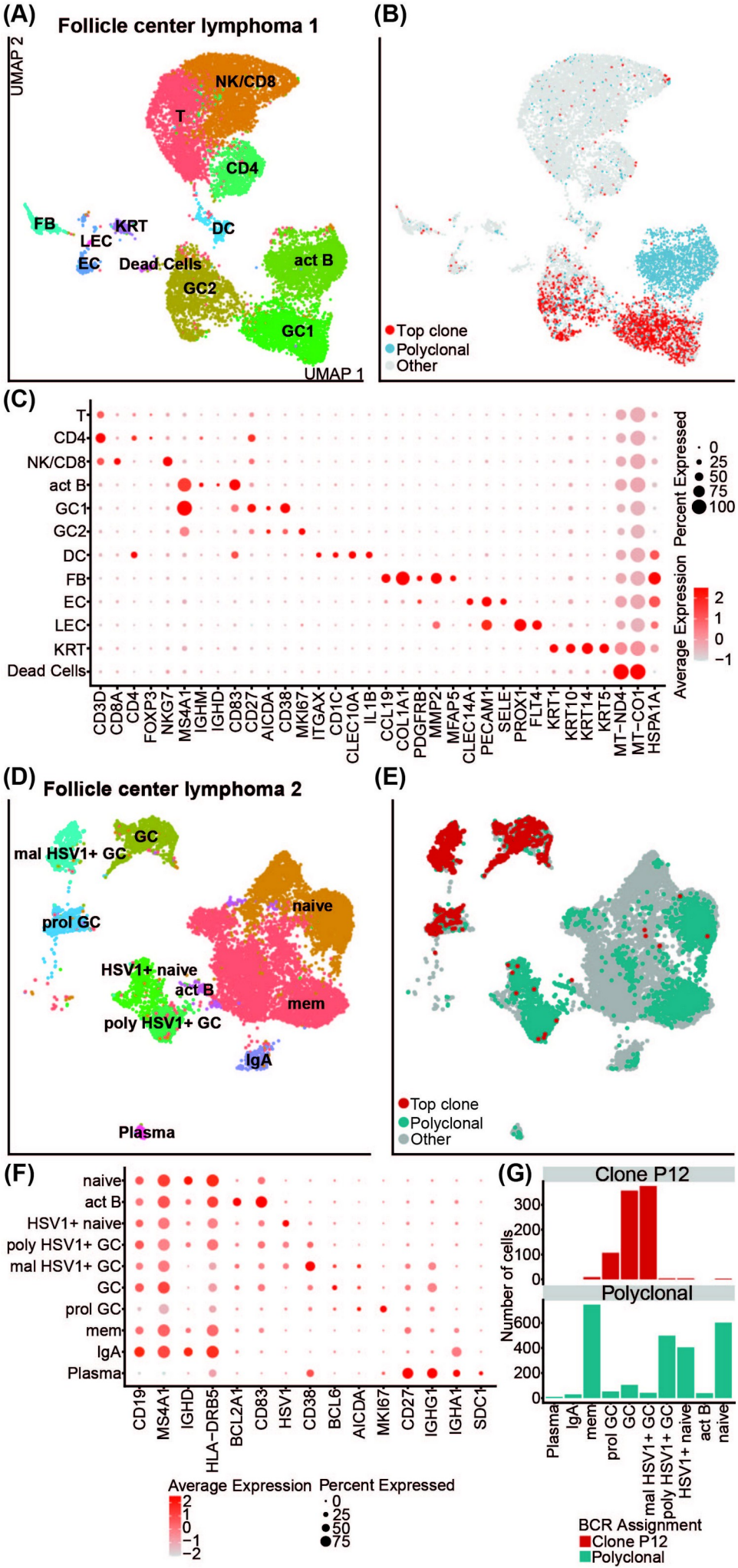

(A)
Follicle center lymphoma 1
UMAP 2
UMAP 1
NK/CD8
T
CD4
FB
KRT
LEC
EC
Dead Cells
DC
act B
GC2
GC1
(B)
Top clone
Polyclonal
Other
(C)
T
CD4
NK/CD8
act B
GC1
GC2
DC
FB
EC
LEC
KRT
Dead Cells
CD3D
CD8A
CD4
FOXP3
NKG7
MS4A1
IGHM
IGHD
CD83
CD27
AICDA
CD38
MKI67
ITGAX
CD1C
CLEC10A
IL1B
CCL19
COL1A1
PDGFRB
MMP2
MFAP5
CLEC14A
PECAM1
SELE
PROX1
FLT4
KRT1
KRT10
KRT14
KRT5
MT-ND4
MT-CO1
HSPA1A
Percent Expressed
0
25
50
75
100
Average Expression
2
1
0
−1
(D)
Follicle center lymphoma 2
GC
mal HSV1+ GC
prol GC
naive
HSV1+ naive
act B
mem
poly HSV1+ GC
IgA
Plasma
(E)
Top clone
Polyclonal
Other
(F)
naive
act B
HSV1+ naive
poly HSV1+ GC
mal HSV1+ GC
GC
prol GC
mem
IgA
Plasma
CD19
MS4A1
IGHD
HLA-DRB5
BCL2A1
CD83
HSV1
CD38
BCL6
AICDA
MKI67
CD27
IGHG1
IGHA1
SDC1
Average Expression
2
1
0
−1
−2
Percent Expressed
0
25
50
75
(G)
Clone P12
Polyclonal
Number of cells
300
200
100
0
600
400
200
0
Plasma
IgA
mem
prol GC
GC
mal HSV1+ GC
poly HSV1+ GC
HSV1+ naive
act B
naive
BCR Assignment
Clone P12
Polyclonal

Figure 5

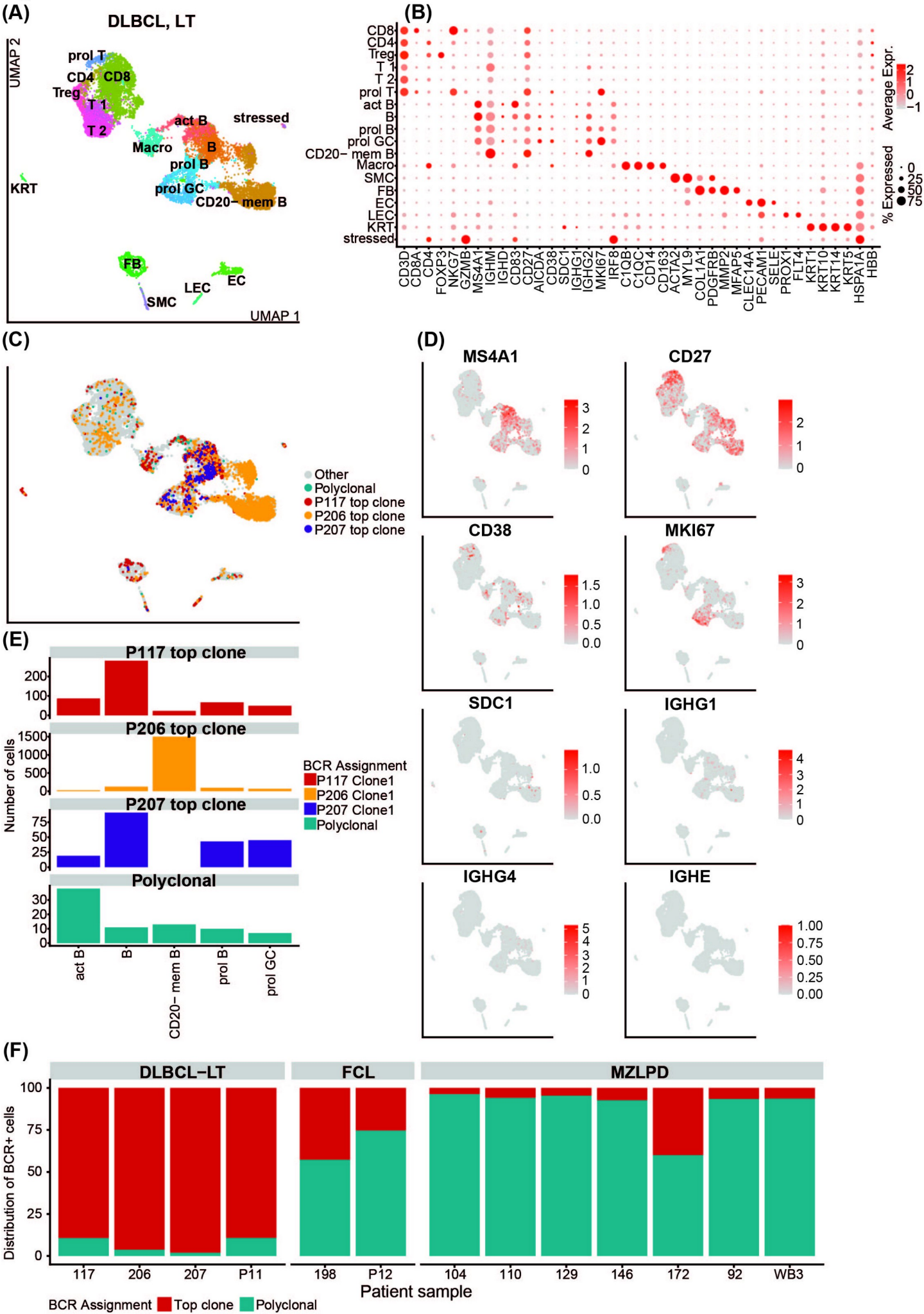
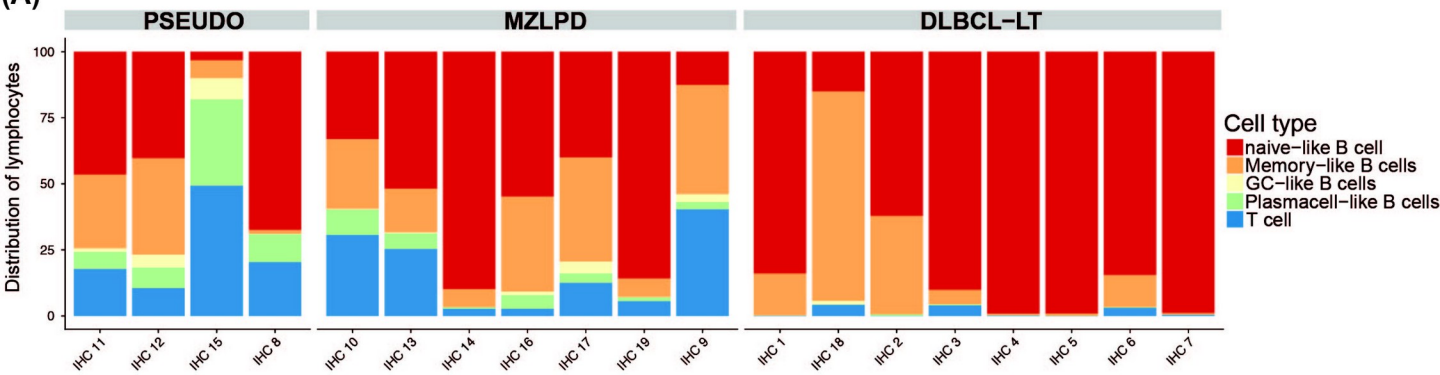
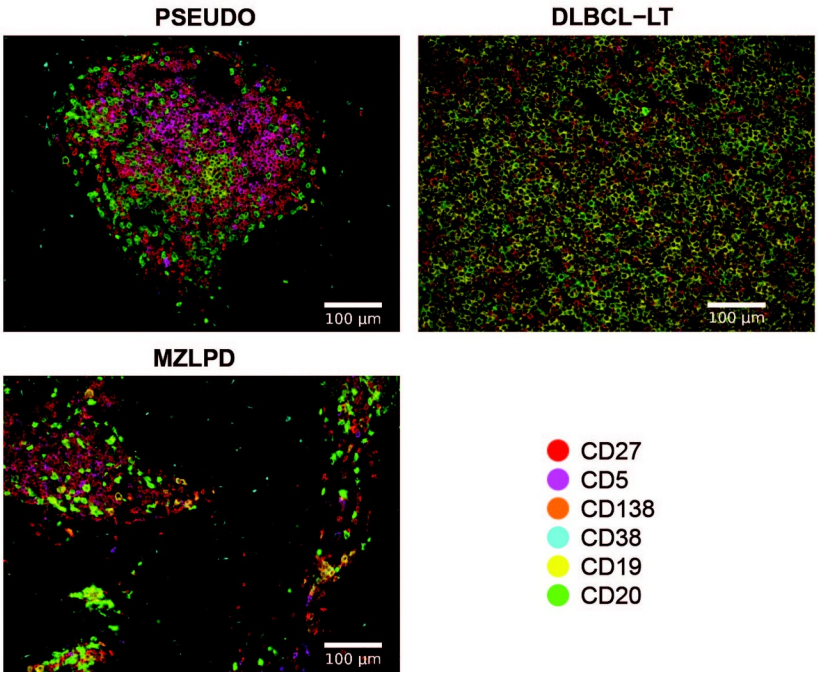

(A)
DLBCL, LT
UMAP 2
UMAP 1
prol T
CD4
CD8
Treg
T 1
T 2
act B
stressed
Macro
B
prol B
KRT
prol GC
CD20− mem B
FB
LEC
EC
SMC
(B)
CD8
CD4
Treg
T 1
T 2
prol T
act B
B
prol B
prol GC
CD20− mem B
Macro
SMC
FB
EC
LEC
KRT
stressed
CD3D
CD8A
CD4
FOXP3
NKG7
GZMB
MS4A1
IGHM
IGHD
CD83
CD27
AICDA
CD38
SDC1
IGHG1
IGHG2
MKI67
IRF8
C1QB
C1QC
CD14
CD163
ACTA2
MYL9
COL1A1
PDGFRB
MMP2
MFAP5
CLEC14A
PECAM1
SELE
PROX1
FLT4
KRT1
KRT10
KRT14
KRT5
HSPA1A
HBB
Average Expr.
% Expressed
(C)
Other
Polyclonal
P117 top clone
P206 top clone
P207 top clone
(D)
MS4A1
CD27
CD38
MKI67
SDC1
IGHG1
IGHG4
IGHE
(E)
P117 top clone
P206 top clone
P207 top clone
Polyclonal
Number of cells
BCR Assignment
P117 Clone1
P206 Clone1
P207 Clone1
Polyclonal
act B
B
CD20− mem B
prol B
prol GC
(F)
DLBCL−LT
FCL
MZLPD
Distribution of BCR+ cells
117
206
207
P11
198
P12
104
110
129
146
172
92
WB3
Patient sample
BCR Assignment
Top clone
Polyclonal

Figure 6

(A)


PSEUDO
MZLPD
DLBCL−LT
Distribution of lymphocytes
100
75
50
25
0
IHC 11
IHC 12
IHC 15
IHC 8
IHC 10
IHC 13
IHC 14
IHC 16
IHC 17
IHC 19
IHC 9
IHC 1
IHC 18
IHC 2
IHC 3
IHC 4
IHC 5
IHC 6
IHC 7
Cell type
naive−like B cell
Memory−like B cells
GC−like B cells
Plasmacell−like B cells
T cell

(B)


PSEUDO
DLBCL−LT
MZLPD
100 µm
100 µm
100 µm
CD27
CD5
CD138
CD38
CD19
CD20

(C)

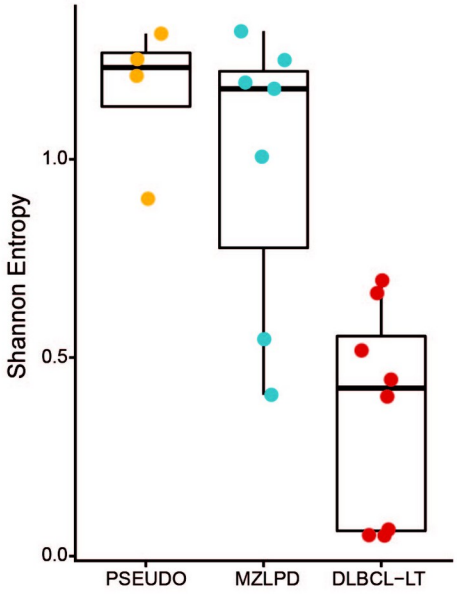

Shannon Entropy
1.0
0.5
0.0
PSEUDO
MZLPD
DLBCL−LT

Figure 7

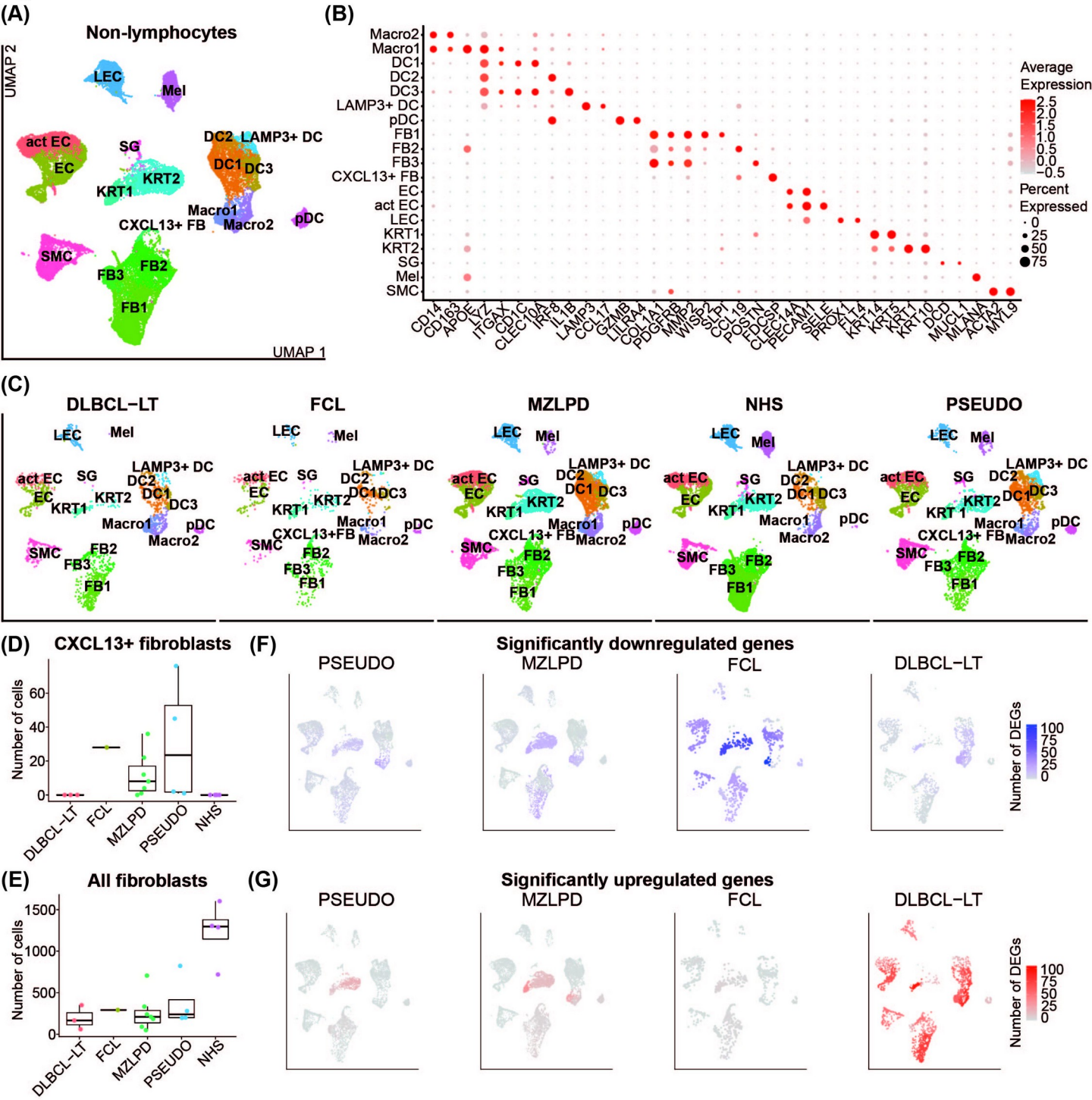

(A)
Non-lymphocytes
UMAP 2
UMAP 1
LEC
Mel
act EC
EC
SG
KRT2
KRT1
DC2
LAMP3+ DC
DC1
DC3
Macro1
Macro2
CXCL13+ FB
pDC
SMC
FB3
FB2
FB1
(B)
Macro2
Macro1
DC1
DC2
DC3
LAMP3+ DC
pDC
FB1
FB2
FB3
CXCL13+ FB
EC
act EC
LEC
KRT1
KRT2
SG
Mel
SMC
CD14
CD163
APOE
LYZ
ITGAX
CD1C
CLEC10A
IRF8
IL1B
LAMP3
CCL17
GZMB
LILRA4
COL1A1
PDGFRB
MMP2
WISP2
SLPI
CCL19
POSTN
FDCSP
CLEC14A
PECAM1
SELE
PROX1
FLT4
KRT14
KRT5
KRT1
KRT10
DCD
MUCL1
MLANA
ACTA2
MYL9
Average Expression
2.5
2.0
1.5
1.0
0.5
0.0
−0.5
Percent Expressed
0
25
50
75
(C)
DLBCL−LT
FCL
MZLPD
NHS
PSEUDO
(D)
CXCL13+ fibroblasts
Number of cells
60
40
20
0
DLBCL−LT
FCL
MZLPD
PSEUDO
NHS
(E)
All fibroblasts
Number of cells
1500
1000
500
0
DLBCL−LT
FCL
MZLPD
PSEUDO
NHS
(F)
Significantly downregulated genes
PSEUDO
MZLPD
FCL
DLBCL−LT
Number of DEGs
100
75
50
25
0
(G)
Significantly upregulated genes
PSEUDO
MZLPD
FCL
DLBCL−LT
Number of DEGs
100
75
50
25
0

Figure S1

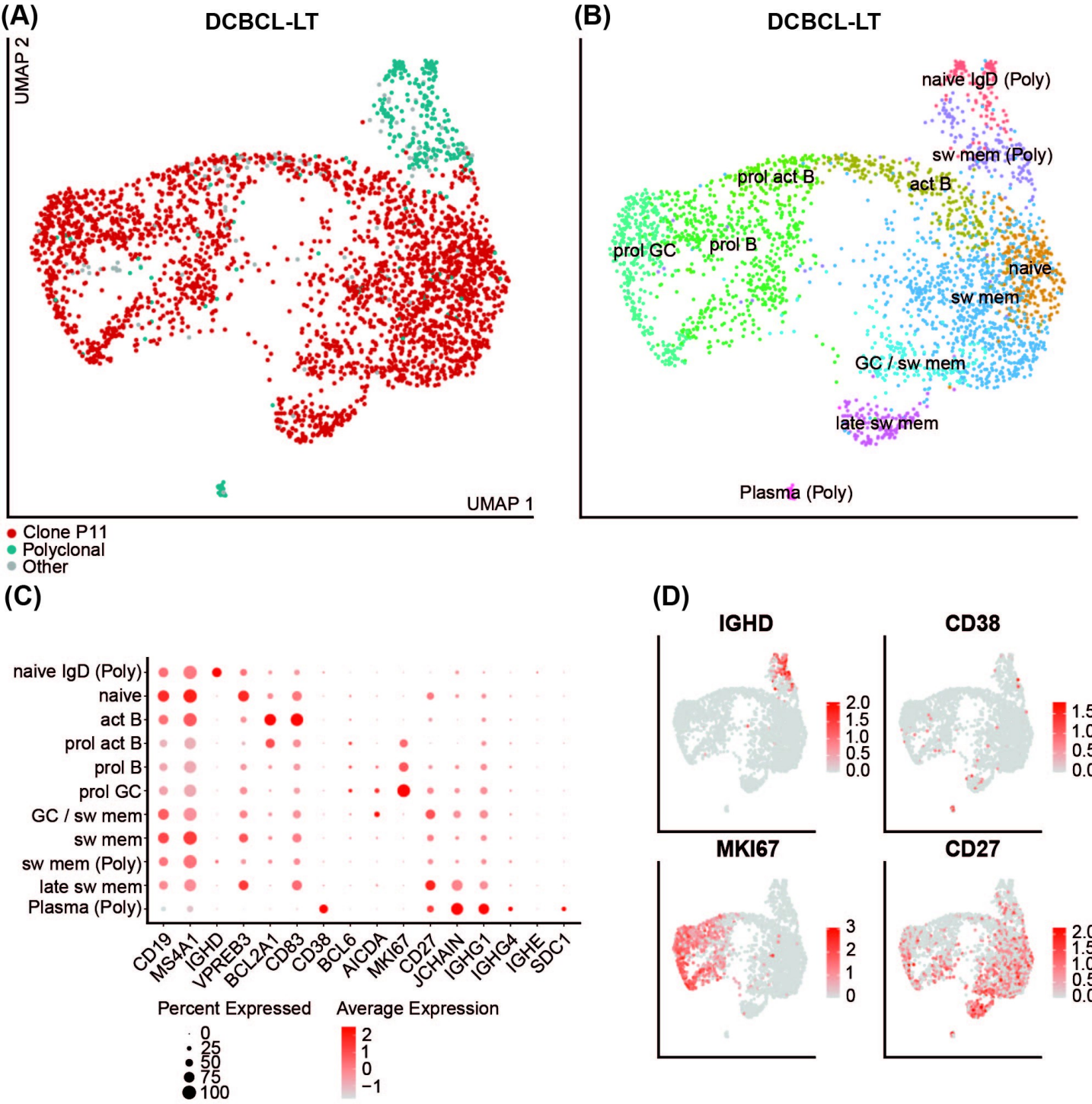

(A)
DCBCL-LT
UMAP 2
UMAP 1
Clone P11
Polyclonal
Other
(B)
DCBCL-LT
naive IgD (Poly)
sw mem (Poly)
prol act B
act B
prol GC
prol B
naive
sw mem
GC / sw mem
late sw mem
Plasma (Poly)
(C)
naive IgD (Poly)
naive
act B
prol act B
prol B
prol GC
GC / sw mem
sw mem
sw mem (Poly)
late sw mem
Plasma (Poly)
CD19
MS4A1
IGHD
VPREB3
BCL2A1
CD83
CD38
BCL6
AICDA
MKI67
CD27
JCHAIN
IGHG1
IGHG4
IGHE
SDC1
Percent Expressed
0
25
50
75
100
Average Expression
2
1
0
−1
(D)
IGHD
2.0
1.5
1.0
0.5
0.0
CD38
1.5
1.0
0.5
0.0
MKI67
3
2
1
0
CD27
2.0
1.5
1.0
0.5
0.0

Figure S2

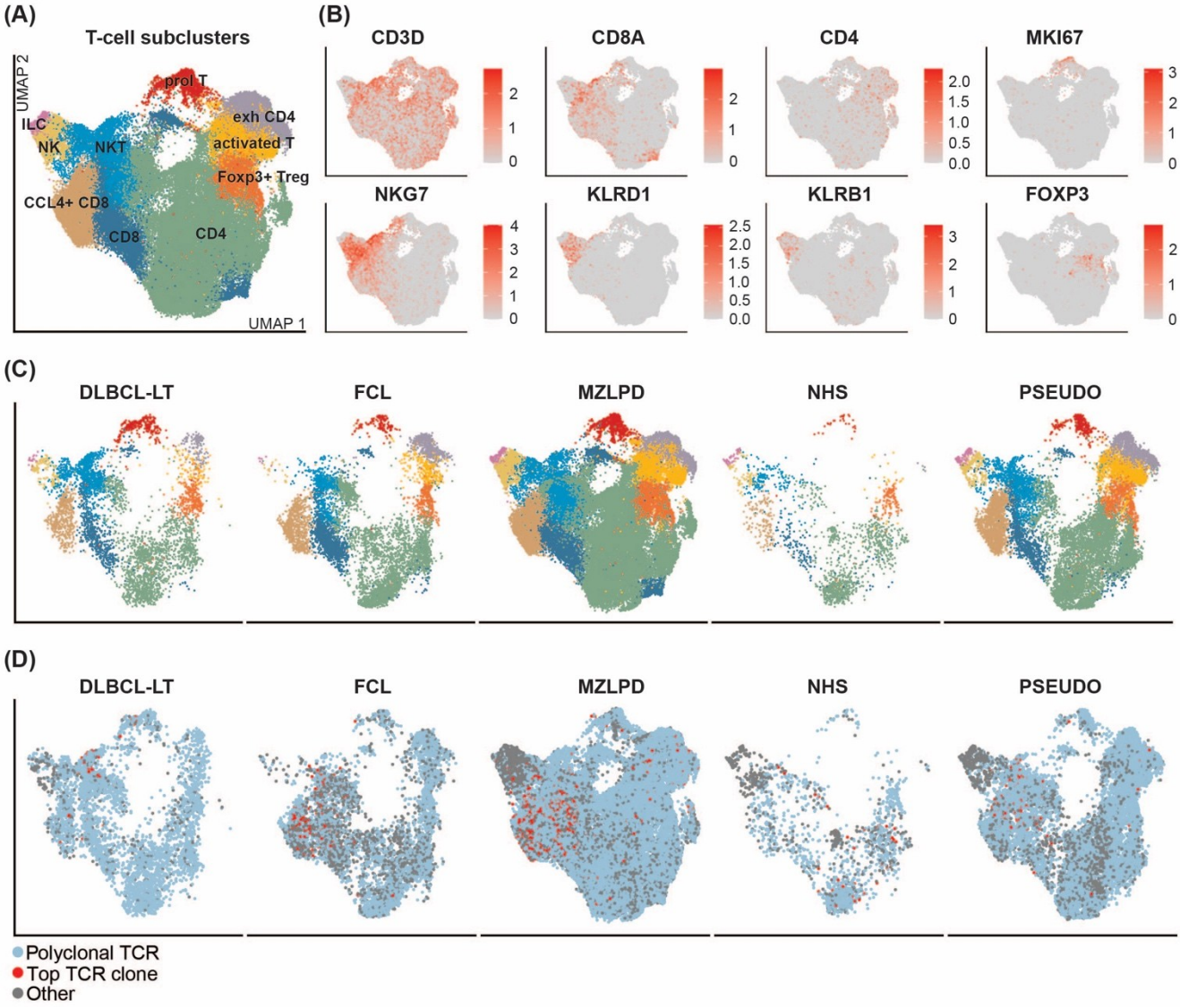
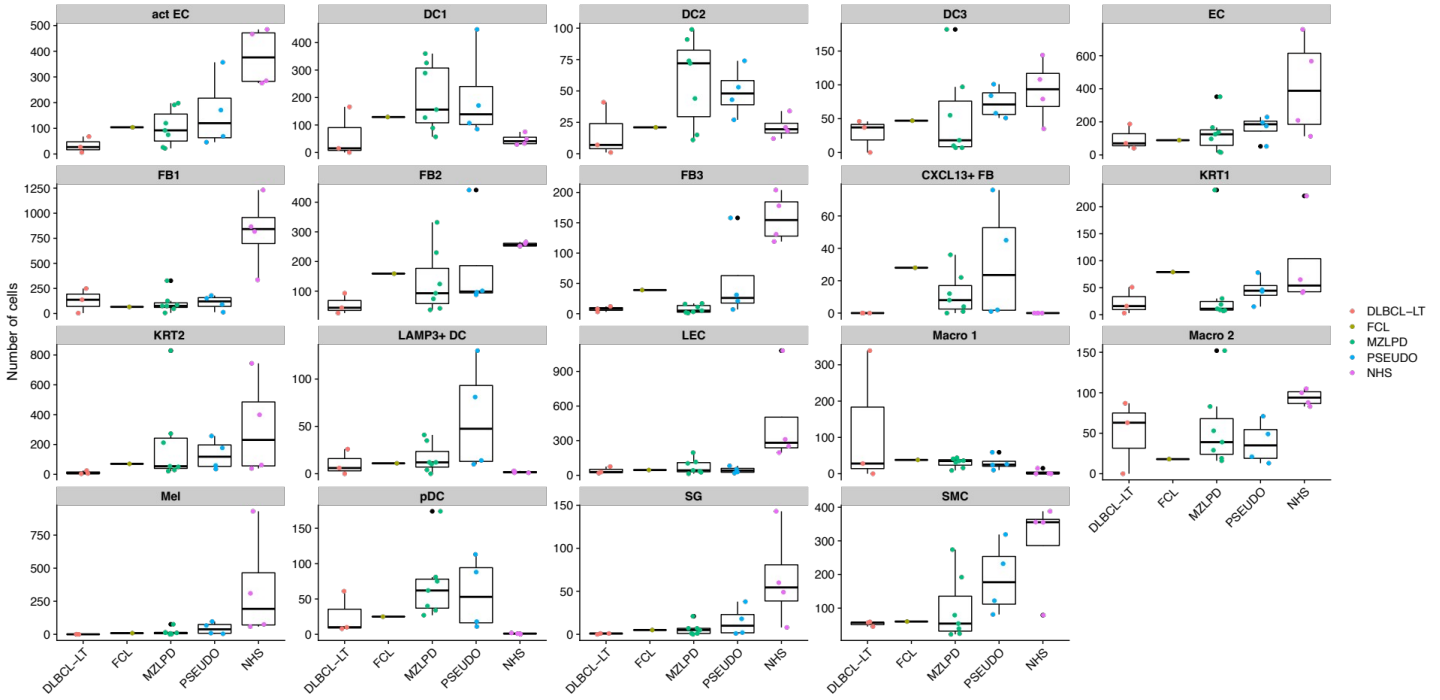

(A)
T-cell subclusters
UMAP 2
UMAP 1
prol T
exh CD4
activated T
Foxp3+ Treg
ILC
NK
NKT
CCL4+ CD8
CD8
CD4
(B)
CD3D
CD8A
CD4
MKI67
NKG7
KLRD1
KLRB1
FOXP3
(C)
DLBCL-LT
FCL
MZLPD
NHS
PSEUDO
(D)
DLBCL-LT
FCL
MZLPD
NHS
PSEUDO
Polyclonal TCR
Top TCR clone
Other

Figure S3


act EC
DC1
DC2
DC3
EC
FB1
FB2
FB3
CXCL13+ FB
KRT1
KRT2
LAMP3+ DC
LEC
Macro 1
Macro 2
Mel
pDC
SG
SMC
Number of cells
DLBCL−LT
FCL
MZLPD
PSEUDO
NHS