haematologica
Journal of the Ferrata Storti Foundation

Gradient of CD79B expression in diffuse large B-cell lymphoma corresponds to stages of germinal center B-cell differentiation

by Yusuke Naoi, Ryota Chijimatsu, Tomohiro Urata, Kazutaka Sunami, Toshi Imai, Yuichiro Nawa, Yasushi Hiramatsu, Kazuhiko Yamamoto, Soichiro Fujii, Isao Yoshida, Tomofumi Yano, Kazuhiro Ikeuchi, Hiroki Kobayashi, Hiroyuki Murakami, Hideki Ujiie, Katsuma Tani, Kaho Kondo, Hirofumi Inoue, Shuta Tomida, Akira Yamamoto, Keisuke Seike, Hideaki Fujiwara, Noboru Asada, Keiko Fujii, Nobuharu Fujii, Ken-ichi Matsuoka, Merrill Boyle, Aixiang Jiang, Pedro Farinha, Katsuyoshi Takata, Yasuharu Sato, Tadashi Yoshino, Yoshinobu Maeda, David W. Scott and Daisuke Ennishi

Received: November 18, 2025.
Accepted: June 18, 2026.

Citation: Yusuke Naoi, Ryota Chijimatsu, Tomohiro Urata, Kazutaka Sunami, Toshi Imai, Yuichiro Nawa, Yasushi Hiramatsu, Kazuhiko Yamamoto, Soichiro Fujii, Isao Yoshida, Tomofumi Yano, Kazuhiro Ikeuchi, Hiroki Kobayashi, Hiroyuki Murakami, Hideki Ujiie, Katsuma Tani, Kaho Kondo, Hirofumi Inoue, Shuta Tomida, Akira Yamamoto, Keisuke Seike, Hideaki Fujiwara, Noboru Asada, Keiko Fujii, Nobuharu Fujii, Ken-ichi Matsuoka, Merrill Boyle, Aixiang Jiang, Pedro Farinha, Katsuyoshi Takata, Yasuharu Sato, Tadashi Yoshino, Yoshinobu Maeda, David W. Scott and Daisuke Ennishi.
Gradient of CD79B expression in diffuse large B-cell lymphoma corresponds to stages of germinal center B-cell differentiation.
Haematologica. 2026 June 25. doi: 10.3324/haematol.2025.300248 [Epub ahead of print]

Gradient of CD79B expression in diffuse large B-cell lymphoma corresponds to stages of germinal center B-cell differentiation

*Yusuke Naoi[1,2], *Ryota Chijimatsu[2], Tomohiro Urata[1], Kazutaka Sunami[3], Toshi Imai[4], Yuichiro Nawa[5], Yasushi Hiramatsu[6], Kazuhiko Yamamoto[7], Soichiro Fujii[8], Isao Yoshida[9], Tomofumi Yano[10], Kazuhiro Ikeuchi[1], Hiroki Kobayashi[1], Hiroyuki Murakami[1], Hideki Ujiie[11], Katsuma Tani[1], Kaho Kondo[1], Hirofumi Inoue[12], Shuta Tomida[2], Akira Yamamoto[13], Keisuke Seike[13], Hideaki Fujiwara[13], Noboru Asada[13], Keiko Fujii[13], Nobuharu Fujii[13], Ken-ichi Matsuoka[13], Merrill Boyle[14], Aixiang Jiang[14], Pedro Farinha[14], Katsuyoshi Takata[15], Yasuharu Sato[16], Tadashi Yoshino[16], Yoshinobu Maeda[13,17], David W. Scott[14], Daisuke Ennishi[2,13#]

[1] Department of Hematology, Oncology and Respiratory Medicine, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan

[2] Center for Comprehensive Genomic Medicine, Okayama University Hospital, Okayama, Japan

[3] Department of Hematology, National Hospital Organization Okayama Medical Center, Okayama, Japan

[4] Division of Hematology and Blood Transfusion, Kochi Health Sciences Center, Kochi, Japan

[5] Division of Hematology, Ehime Prefectural Central Hospital, Matsuyama, Japan

[6] Division of Hematology and Oncology, Japanese Red Cross Society Himeji Hospital, Hyogo, Japan

[7] Department of Hematology and Oncology, Okayama City Hospital, Okayama, Japan

[8] Department of Hematology, Japanese Red Cross Okayama Hospital, Okayama, Japan

[9] Department of Hematologic Oncology, National Hospital Organization Shikoku Cancer Center, Matsuyama, Japan

[10] Department of Internal Medicine, Okayama Rosai Hospital, Okayama, Japan

[11] Department of Molecular Hematopathology, Okayama University Graduate School of Health Sciences, Okayama, Japan

[12] Clinical Genomic Medicine, Okayama University Graduate School of Medicine, Dentistry, and Pharmaceutical Science, Okayama, Japan

[13] Department of Hematology and Oncology, Okayama University Hospital, Okayama, Japan

[14] Centre for Lymphoid Cancer, BC Cancer, Vancouver, British Columbia, Canada

[15] Division of Molecular and Cellular Pathology, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan

[16] Department of Molecular Hematopathology, Okayama University, Okayama, Japan

[17] Department of Hematology, Oncology and Respiratory Medicine, Okayama University Faculty of Medicine, Dentistry and Pharmaceutical Sciences, Okayama, Japan

*YN and RC contributed equally as co-first authors.

#Corresponding Author

Daisuke Ennishi

Mailing Address: 2-5-1, Shikata-cho, Kita-ku, Okayama city, Japan

Email Address: daisukeennishi@okayama-u.ac.jp

Running Title: Clinicopathological significance of CD79B in DLBCL

Prior presentation: This study was presented at the 17th International Conference on Malignant Lymphoma (ICML) Lugano, on Jun 16th 2023, and at the 65th American Society of Hematology (ASH) annual meeting 2023, San Diego, on Dec 10th 2023.

Data Availability Protein expression data used for the analyses are found in the data supplement available in the online version of this article. Counts matrix generated using Cell Ranger v7.0.0 and processed Seurat objects of CITEseq experiment are available in the European Genome-phenome Archive (EGA; EGAD50000000538) via controlled access. Scripts used for data analysis are available on GitHub (https://github.com/ennishilab/ynaoi_CD79B). Any additional data and scripts are available upon request.

Contribution: Y.N., R.C. and D.E. designed and performed the research, analyzed and interpreted data, and wrote the manuscript; K.S., T.I., Y.N., Y.H., K.Y., S.F., I.Y., T.Y., A.Y., K.S., H.F., N.A., K.F., N.F., and K.M. provided clinical data; K.T., Y.S., T.Y., and P.F. performed pathological experiments and review.; Y.N., R.C., T.U., K.I., H.K.,

H.M., H.U., K.T., K.K., H.I., and S.T. performed experiments and analyzed data; A.J., M.B., Y.M., D.W.S., and D.E. reviewed the manuscript.

Conflicts of interest disclosure: I.Y. declares research funding and speakers' bureau from Kyowa Kirin and Chugai; speakers' bureau from Janssen, Eisai, Dainippon Sumitomo, Bristol Myers Squibb, Meiji Seika, Nippon Shinyaku, and AstraZeneca. K.S. declares research funding and honoraria from Ono, Bristol Myers Squibb, Janssen, and Sanofi; research funding from Celgene, AbbVie, Takeda, GSK, Chugai, Otsuka, MSD, Novartis, Pfizer, and Kyowa Kirin. S.T. declares honoraria from Illumina, Inc. and NanoString Technologies, Inc. Y.S. declares honoraria from Chugai, Kyowa Kirin, and Takeda; current holder of stock options in a privately held company and honoraria from Daiichi Sankyo. Y.M. declares research funding and honoraria from Astellas, Chugai, Nippon Shinyaku; research funding and speakers' bureau from Novartis; research funding from AstraZeneca; honoraria from Eisai, Kyowa Kirin, Takeda, and Otsuka. D.W.S. declares consultancy for AbbVie, AstraZeneca, Incyte, and Veracyte; research funding from Janssen and Roche. D.E. declares research funding from Nippon Shinyaku, Chugai, and Eisai; honoraria from Eisai, Kyowa Kirin, Chugai, SymBio, Bristol Myers Squibb, and Nippon Shinyaku.

Acknowledgments: We thank Yuka Gion, Misa Sakamoto, Yuria Egusa, Azusa Fujita, Jin Kiyama, Yukina Maekawa, Sayako Yoshida, Kanna Maehama and Shoma Kato for histology expertise in FFPE block processing and tissue microarray construction. We also thank Toshiya Tsubaki, Yukari Kawai, Yuri Tamada, Mutsumi Okabe and Takehiro Matsubara for expert technical assistance.

Funding: This work was supported by Grants-in-Aid for Scientific Research (KAKENHI: 19K07774, 22K19546, 23K27625 and 23H02934 to D.E.), JST FOREST Program (JPMJFR210L to D.E.) and Japan Agency for Medical Research and Development (AMED) under grant number 22ama221516h0001 (to D.E.).

Abstract

A CD79B-targeted antibody-drug conjugate, polatuzumab-vedotin has significantly improved outcomes in DLBCL, particularly in the activated B-cell-like (ABC) subtype. This indicates that CD79B expression varies among COO subtypes, although the expression level was not necessarily higher in ABC-DLBCL. Here, we evaluated CD79B protein expression by immunohistochemistry (IHC) in 590 *de novo* DLBCL cases, observing a CD79B gradient according to COO subtypes, with ABC-DLBCL showing the lowest expression, followed by GCB and dark-zone signature-positive ($DZsig^{pos}$) cases in ascending order ($P < 0.0001$). This observation was fully validated in an independent population-based cohort of 272 cases derived from BC Cancer registry. Given that COO classification reflects the stages of normal germinal center (GC) B-cell differentiation, we further explored CD79B expression dynamics during B-cell maturation by performing single-cell proteomic and transcriptomic analyses. Analyses of 2,447 normal GC B-cells demonstrated a progressive decrease in CD79B expression as GC B-cells transitioned toward terminal differentiation. Collectively, our findings reveal a novel COO-dependent gradient of CD79B expression, particularly with lower expression in ABC-DLBCL.

Introduction

Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma and displays considerable clinical and biological heterogeneity. Based on gene expression profiles (GEP), DLBCL is broadly classified into two main cell-of-origin (COO) subtypes: germinal center B-cell-like (GCB) and activated B-cell-like (ABC) DLBCL[1,2]. These COO subtypes correlate with different stages of normal germinal center (GC) B-cell differentiation, with ABC-DLBCL resembling post-GC B-cells and most GCB-DLBCLs reflecting light zone (LZ) GC B-cells[3]. More recently, a "dark-zone signature" (DZsig) was identified, further refining GCB-DLBCL by distinguishing tumors that transcriptionally resemble dark zone (DZ) GC B-cells[4,5].

CD79B is a component of the immunoglobulin-associated heterodimer CD79A-CD79B within the proximal B-cell receptor signaling complex and serves as the target of polatuzumab-vedotin, an antibody-drug conjugate based on an anti-CD79B monoclonal antibody[6]. The POLARIX study[7,8], which evaluated polatuzumab-vedotin in combination with rituximab, cyclophosphamide, doxorubicin, and prednisone (Pola-R-CHP) in previously untreated patients with DLBCL, demonstrated a significant improvement in progression-free survival (PFS) compared with the conventional R-CHOP regimen (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone). Subgroup analyses of POLARIX suggested that Pola-R-CHP conferred greater PFS benefits in patients with the ABC-DLBCL subtype. Similar clinical study findings have reinforced the notion of superior efficacy of polatuzumab-vedotin in ABC-DLBCL compared to GCB-DLBCL[8–11]. These observations raise a critical question of whether CD79B expression varies across COO subtypes and whether such variation might underlie the increased therapeutic efficacy observed in ABC-DLBCL. However, this issue has not been systematically addressed to date.

Here, we conducted CD79B immunohistochemistry (IHC) analysis on a large cohort of *de novo* DLBCL patients to elucidate the relationship between CD79B expression and the clinicopathological and molecular characteristics of the disease. We further integrated single-cell protein and transcriptome

analyses of normal B-cells to investigate the association between CD79B expression and stages of B-cell differentiation. These findings delineate a COO-specific gradient of CD79B expression in DLBCL and provide biological context to support future efforts in refining molecular classification and optimizing therapeutic stratification.

Materials and Methods

Patient cohort

We retrospectively analyzed patients with *de novo* DLBCL treated with R-CHOP or R-CHOP-like treatments with curative intent between 2008 and 2018 at the Okayama Hematology Study Group (OHSG) in Japan[5]. All cases had refined COO subtype information determined by the DLBCL90 assay[12]. This study was conducted in accordance with the Declaration of Helsinki and approved by the ethics board of each participating institution. The characteristics of the patients in our cohort are presented in Table S1.

Immunohistochemistry (IHC) data

We used a tissue microarray (TMA) of formalin-fixed paraffin-embedded (FFPE) tissues generated in our previous study[5]. CD79B (clone AT107-2, Serotec, Hercules, CA, USA; dilution used, 1:750) staining was performed on a BOND-Rx platform (Leica Biosystems, Wetzlar, Germany) according to routine staining protocols. To estimate tumor content, CD20 (clone L26, Leica Biosystems, Wetzlar, Germany) and CD79A (clone JCB117, DAKO, Santa Clara, CA, USA) were stained and the H-score[5,10,13] was calculated. CD79B staining was visually assessed by two expert hematopathologists (KT and YS) to define staining positivity, where cases without cytoplasmic and membranous staining are classified as CD79B-negative. Of the 623 cases available for CD79B IHC data with sufficient clinical information, 4 cases were excluded due to low tumor content (< 20%), and 29 cases were excluded due to the following

reasons: extremely small specimens, widespread tissue damage, overstaining, and negative staining without internal positive controls. As the distribution of CD79B H-scores differed between cohorts, cohort-specific median values were used for dichotomization to enable relative comparisons within each cohort (Figure S1). In addition, IHC using alternative CD79B antibody clones (B29/123 and EPR6861) and other B-cell receptor (BCR) components (CD79A, IgM and IgG) was performed in a subset of cases to validate staining consistency and assess their relationships (Table S2).

External cohort analyses

CD79B (AT107-2) IHC staining was performed on 347 *de novo* DLBCL in the BC Cancer population-based registry ("BCC cohort")[4,12,14]. Data for 272 patients had available CD79B IHC and gene expression data, along with sufficient clinical information (Table S3).

Additionally, we analyzed 517 DLBCL cases from "NCI cohort"[15]. Both the BCC and NCI cohorts were assigned to genetic subgroups (LymphGen)[16], as well as to a taxonomy reflecting different stages of GC B-cell differentiation, as detected by single-cell RNA sequencing (scRNA-seq) and hereafter referred to as the "scGC cluster"[17].

Single-cell analysis

We conducted cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq)[18] on reactive lymphoid hyperplasia tissues collected from two patients (OKU_S010 and OKU_S073) following standard protocols using the Chromium Single Cell 5' Kit v2 and the Chromium Controller (10x Genomics). In addition, we analyzed six recent representative scRNA-seq datasets (22 patients) of non-malignant lymphoid tissues[17,19–23] (Table S4). BCR signaling activity was quantified using R UCell package[24] based on the KEGG_B_CELL_RECEPTOR_SIGNALING_PATHWAY gene set from MSigDB (75 genes). *CD79A* and *CD79B* were excluded from the gene set prior to scoring.

Statistical analysis

For categorical data, Fisher's exact test and the Chi-square test were used, where applicable. Wilcoxon rank-sum test or Kruskal-Wallis test followed by Dunn's test was applied for continuous variables. Multiple test corrections were performed using the Benjamini-Hochberg procedure when necessary. The Kaplan-Meier method was used to estimate overall survival (OS) and PFS, and the log-rank test was used for comparison. Univariate and multivariate Cox proportional hazard regression models were used to evaluate prognostic factors. All p-values were calculated from two-sided tests, and a significance level of 0.05 was used. The analyses were performed using R version 4.3.2 and Python version 3.10.13.

Results

Distinct clinicopathological characteristics according to CD79B expression in DLBCL

Of the 1050 *de novo* patients with DLBCL who were included in our previous study, CD79B IHC results were interpretable for 590 patients. To analyze the association between CD79B protein expression and COO subtypes, we used gene-expression-based COO subtypes defined by the DLBCL90 assay generated in our previous study[5]. This assay assigned the present cohort to 302 ABC (51%), 187 GCB (32%), 33 DZsigpos (6%), and 68 unclassified (11%). Figure 1A shows the representative CD79B IHC images at different H-scores. CD79B H-scores showed significant differences according to the COO, with ABC-DLBCL showing the lowest values, followed by GCB-DLBCL and DZsig^{pos}-DLBCL ($P < 0.0001$; Figure 1B,C). Consistent COO-specific differences in CD79B H-score were observed when unclassified cases were included (Figure S2). Similarly, the CD79B H-score was significantly lower in the non-GCB-DLBCL cohort than in the GCB-DLBCL cohort, as defined by the Hans' classification ($P < 0.0001$; Figure 1D and Figure S2C). Of note, out of 590 patients, 60 (10%) did not express the protein in either membrane or cytoplasm and these CD79B-negative cases were more common in ABC-DLBCL (Fisher's exact test, $P = 0.007$; Table S5). The patients were divided into CD79B^{high}- and CD79B^{low}-groups based

on the median CD79B H-score, there were no significant differences in baseline characteristics between the groups (Table 1). Consistent with the association of CD79B H-score with COO, CD79B^{low}-DLBCL was more prevalent in ABC-DLBCL (57%) compared to GCB-DLBCL (39%) and DZsig^{pos}-DLBCL (18%) groups (Chi-square test, $P < 0.001$; Table 1). Additional validation using two alternative CD79B antibody clones (B29/123 and EPR6861) in a subset of OHSG cases showed high concordance with respect to positivity across clones (90-93%), supporting the reproducibility of the staining results (Figure S3 and Table S6). Furthermore, correlation analyses integrating protein and transcriptomic data in the BCC cohort revealed that CD79B protein expression was positively associated with *CD79B* and *CD79A* mRNA expression, whereas only a weak association was observed with *KLHL6* mRNA expression (Figure S4).

Independent cohort analysis validated the association of CD79B protein expression with molecular subtypes

To assess the differential expression of CD79B across the COO groups in an independent DLBCL cohort, we performed an external validation experiment using the BCC cohort derived from the BC Cancer population-based registry[12,14]. Out of the 272 patients with CD79B IHC and molecular classification data, 88 had ABC-DLBCL (32%), 117 had GCB-DLBCL (43%), 38 had DZsig^{pos}-DLBCL (14%), and 29 were unclassified (11%)[4]. Consistent with the results in our cohort, the CD79B H-score was lowest in ABC-DLBCL and highest in DZsig^{pos}-DLBCL ($P < 0.001$; Figure 2A). These COO-specific variations in CD79B H-score were preserved regardless of unclassified case inclusion (Figure S5A). The CD79B H-score was also significantly lower in non-GCB-DLBCL compared to GCB-DLBCL, as defined by Hans' classification ($P < 0.001$; Figure 2B and S5B). No significant differences in baseline characteristics were found in median age, sex, or international prognostic index (IPI) score between CD79B^{high}- and CD79B^{low}-DLBCL (Table S7).

Since mutation and whole transcriptome data were available for the BCC cohort, we analyzed the association of CD79B protein expression with these molecular profiles. First, we examined the association of CD79B H-score with *CD79B* mutations and LymphGen classification[16]. As LymphGen classification is independent of COO classification, cases remaining unclassified when using COO were assigned to LymphGen subgroups. In the BCC cohort, 26 of 29 unclassified cases were assigned to non-MCD subtypes and 3 to the MCD subtype. CD79B H-score showed no apparent difference based on the *CD79B* mutation status in the entire cohort and in ABC-DLBCL (Figure 2C). Notably, there was no significant association between *CD79B* mutation status and CD79B-negative IHC staining was observed (Fisher's exact test, $P = 1.00$; Table S8). However, the CD79B H-score was significantly lower in the MCD subtype across the whole cohort, and a similar trend was observed even when the analysis was restricted to ABC-DLBCL cases within the MCD subtype ($P = 0.0034$ and $P = 0.097$, respectively; Figure 2C and Figure S6). Indeed, the CD79B^{low} group was significantly enriched in the MCD subtype both in the entire cohort and in ABC-DLBCL (Fisher's exact test, $P < 0.001$ and $P = 0.018$, respectively; Table S9). In contrast, CD79B positivity was not significantly associated with the MCD subtype in either the entire cohort or ABC-DLBCL ($P = 0.15$ and $P = 0.63$, respectively; Table S9).

To further characterize the biological association between CD79B expression and COO, we applied a publicly available GC-based classification developed by Holmes et al., which allowed for the identification of gene signatures reflecting the differentiation of normal GC B-cells at the single-cell level[17]. The 140 patients in the BCC cohort were assigned to one of the five scGC clusters: dark zone B-cell like (DZ-like; n = 19), light zone B-cell like (LZ-like; n = 23), intermediate zone B-cell like (INT-like; n = 57), plasmablast like (PBL-like; n = 16) and precursor memory B-cell like (PreM-like; n = 25). The distribution of scGC clusters significantly differed among COO subtypes, aligning with their presumed cellular origins: DZsig^{pos} DLBCL was enriched for DZ-like clusters, ABC for post-GC-like clusters, and GCB displayed an intermediate INT-like pattern (Fisher's exact test, $P < 0.0001$; Table S10). Consistent with the relationship between CD79B expression and COO observed in our cohort and the

BCC cohort, the CD79B H-score exhibited a notable difference according to the scGC cluster, with DZ-like DLBCL displaying the highest values, followed by INT-like-, LZ-like-, PBL-like-, and PreM-like DLBCL, in descending order ($P = 0.013$; Figure 2D). These findings collectively indicate the association of GC B-cell differentiation with CD79B expression.

The clinical relevance of CD79B expression

To evaluate the clinical effect of CD79B expression in patients with DLBCL treated with R-CHOP, we conducted survival analyses using Kaplan-Meier methods based on CD79B H-score subgroups. In survival analyses, the "non-ABC" group included GCB, $DZsig^{pos}$, and unclassified cases. $CD79B^{low}$ patients were found to have an inferior overall outcome in the entire cohort (log-rank test, $P < 0.0001$ for OS and $P = 0.018$ for PFS) and in ABC-DLBCL (log-rank test, $P < 0.001$ for OS and $P = 0.21$ for PFS; Figure 3A,B). A similar trend was observed using an alternative dichotomization based on IHC positivity (Figure S7). After adjusting for IPI and COO subtypes, reduced CD79B expression remained significantly associated with shorter OS and PFS across multiple analytical approaches, including median H-score dichotomization, continuous modeling, and IHC positivity (Table S11). In contrast, in the BCC cohort, CD79B H-score was not significantly associated with survival, regardless of whether it was analyzed as a dichotomized or continuous variable, whereas IPI and COO remained significant prognostic factors (Figure 3C,D, Figure S8, and Table S12).

Association of *CD79B* gene expression and other molecular profiles in DLBCL

Next, we investigated the relationship between *CD79B* gene expression and molecular subtypes of DLBCL using two representative datasets: the BCC cohort (n = 272) and the NCI cohort (n = 571). The $CD79B^{low}$-group showed significantly lower *CD79B* gene expression than the $CD79B^{high}$-group in the BCC cohort (Figure 4A). In the NCI cohort, *CD79B* mRNA levels were analyzed as a continuous variable

without defining high or low expression groups. A gradient of *CD79B* gene expression was observed according to COO subtypes, with the lowest levels in ABC-DLBCL and the highest in DZsig^{pos}-DLBCL, in both the BCC and NCI cohorts ($P = 0.0019$ and $P = 0.079$, respectively; Figure 4B). DZ-like DLBCL exhibited significantly higher *CD79B* gene expression than PBL-like and PreM-like subtypes, which reflect terminally differentiating GC B-cells, in both cohorts ($P < 0.0001$ for both; Figure 4C). These findings support the association between the *CD79B* gene expression gradient and B-cell differentiation-related DLBCL subtypes. Lastly, *CD79B* gene expression did not significantly differ according to *CD79B* mutation status (Figure 4D) or between MCD and non-MCD DLBCL (Figure 4E) in either cohort.

Lower CD79B expression is associated with terminal differentiation of germinal center B-cells

The observed gradient of CD79B expression according to COO subtypes prompted us to further investigate the association of CD79B expression with normal GC B-cell differentiation. First, to assess the surface protein expression of CD79B of normal B-cells at the single-cell level, we performed simultaneous protein and transcriptome analyses using CITE-seq on two reactive lymphoid samples (Figure S9). The expression data of 2,447 B-cells underwent dimensionality reduction, and each cell was scored by the R UCell package[24] using the previously discovered 13 GC B-cell signatures (Figure S10)[17]. By applying non-negative matrix factorization (NMF) clustering, 13 clusters identified by Holmes et al. were re-clustered into seven clusters (Figure 5A). Remarkably, both CD79B protein and mRNA expression levels showed a similar decreasing trend in GC-exiting B-cell subsets, particularly in PBL, compared with DZ B-cells ($P < 0.0001$; Figure 5B,5C and Figure S11).

Next, we analyzed publicly available scRNA-seq datasets from normal lymphoid tissues (Figure S9)[17,19–23]. By applying UCell score-based NMF clustering to 75,771 B-cells, we identified eight clusters (Figure 5D and Figure S12). Consistent with the results from the above CITE-seq analysis, PreM/INTd- and PBLa/PBLB-cells showed significantly lower *CD79B* mRNA levels than cells exhibiting DZ B-cell signatures (Figure 5E,F). Collectively, these findings demonstrate that CD79B expression is tightly linked

to the differentiation status of normal B-cells, particularly the transition from GC to post-GC states. In addition, analysis of BCR signaling activity using a KEGG BCR signaling signature (excluding *CD79A* and *CD79B*) showed lower activity in PBL-like populations compared to other clusters, although substantial overlap was observed across scGC clusters, broadly paralleling the reduction in *CD79B* mRNA and protein expression (Figure S13).

Discussion

The present study highlighted the association of CD79B expression with the clinicopathological and molecular characteristics of DLBCL using large patient cohorts. Notably, their findings were validated in large external cohorts with different demographic features. Furthermore, integrative single-cell protein and transcriptome analyses of normal lymphoid tissues showed that CD79B expression peaks in DZ B-cells and declines as B-cells exited the GC. Collectively, our findings provide a foundational framework for future translational and mechanistic studies on CD79B expression in DLBCL.

The greater benefits of polatuzumab-vedotin in ABC-DLBCL compared to GCB-DLBCL have been repeatedly observed in clinical trials and real-world settings[8–11,25]. In contrast, lower CD79B expression in ABC-DLBCL or non-GCB-DLBCL was observed in our study, consistent with earlier reports[26,27]. Indeed, limited correlation between IHC-based CD79B expression and therapeutic response has been described[10,11,28], suggesting that IHC alone may not fully capture functionally relevant surface CD79B expression. In line with this interpretation, previous experimental studies using flow cytometry (FCM) quantification have shown a significant association between CD79B levels and polatuzumab-vedotin sensitivity[26,29]. This suggests that complementary approaches to assess CD79B expression may provide more informative insights.

Recent mechanistic studies have highlighted the complex regulation of CD79B. In particular, KLHL6 has been identified as a post-translational regulator of BCR components, modulating CD79B protein levels through ubiquitin-mediated degradation[29,30]. In addition, biological consequences of *KLHL6* alterations may vary according to mutation patterns, with distinct BTB-domain and Kelch-domain mutations affecting different functional protein groups[30]. Because *KLHL6* mutation data for the present cohort were not available, the relationship between specific *KLHL6* alterations and CD79B protein expression remains unclear. Furthermore, additional IHC analyses of BCR components showed that CD79B protein expression was positively associated with CD79A, whereas relationships with IgM and IgG expression levels were less pronounced. This pattern differs from previous reports suggesting coordinated regulation of BCR components, including immunoglobulin heavy chain expression, in the context of BCR silencing[31]. These findings suggest that CD79B expression cannot be explained by a single regulatory mechanism and is likely influenced by multiple factors, including post-translational regulation and differentiation-associated processes.

Single-cell multi-omics analyses of normal lymphoid tissues showed that CD79B expression was high in DZ B-cells but decreased as B-cells transitioned toward terminal differentiation, particularly in PBLs. These findings suggest that differential CD79B expression among COO subtypes mirrors a physiological gradient of CD79B expression during normal B-cell differentiation. Consistent with this finding, we also observed a gradual decrease in BCR signaling activity toward GC exit, paralleling the reduction in CD79B expression, suggesting that CD79B downregulation may reflect changes in BCR signaling capacity during B-cell differentiation. Similar subtype-specific protein expression gradients have been reported in previous studies[32–34]. Lower MHC-II expression in DLBCL tumors derived from centroblasts in the GC DZ corresponds to the normal gradient of MHC-II expression during GC B-cell maturation. Here, centroblasts express lower MHC-II levels than centrocytes[35]. Our data indicate that *CD79B* mutations do not influence CD79B expression levels. Consistent with previous reports, our data indicate

that *CD79B* mutations are not necessarily associated with CD79B expression levels, suggesting the involvement of additional regulatory mechanisms underlying subtype-specific CD79B expression in DLBCL, potentially related to B-cell differentiation.

Moreover, in our cohort, complete loss of CD79B protein expression occurred in approximately 10% and was significantly more frequent in ABC-DLBCL. Notably, IHC analyses using independent CD79B antibody clones showed high concordance in positivity, indicating that CD79B-negative cases do not simply reflect technical artifacts but rather represent biologically relevant loss of expression. Consistent with this finding, previous studies have reported loss of CD79B expression in a subset of DLBCL cases by IHC[26] or by FCM[36]. Recently, unique mechanisms for CD20 loss in B-cell lymphomas were elucidated, such as transcriptional downregulation and mutations acquired during rituximab resistance[37,38], suggesting that multiple molecular pathways may contribute to CD79B loss. Taken together, these findings underscore the need for further investigation into the biological mechanisms underlying complete CD79B loss, which may have critical implications for resistance to polatuzumab-vedotin.

Our analysis showed that the CD79B^{low}-group was associated with poor outcomes in DLBCL patients treated with R-CHOP. Cases with complete loss of CD79B expression had significantly shorter OS and PFS compared to CD79B-positive cases. However, these associations were not consistently observed in the external BCC cohort, indicating that the prognostic impact of CD79B expression may be context-dependent. The lack of reproducibility across cohorts may be partly explained by differences in clinical background. Taken together, these findings suggest that the clinical relevance of CD79B expression in R-CHOP-treated DLBCL may be limited. In this context, evaluation of CD79B expression in patients treated with polatuzumab-based regimens will be warranted to elucidate its potential predictive value.

Several other limitations of our study should be acknowledged when interpreting the findings. Although we analyzed a relatively large dataset of CD79B protein expression with an independent validation cohort, IHC results were unavailable for approximately one-third of the patients in our previous study. In addition, CD79B expression was assessed by IHC using FFPE tissue samples, which does not specifically quantify surface CD79B expression. Therefore, the functional relevance of CD79B expression levels may not be fully captured by this approach. Furthermore, flow cytometric validation was not feasible in this retrospective cohort due to the lack of viable samples suitable for such analyses. Although clinical studies have suggested differential efficacy of polatuzumab-vedotin across COO subtypes, our study did not directly investigate the biological mechanisms underlying therapeutic response, including potential contributions from BCR signaling dynamics, cell-cycle regulation, or DNA damage response pathways.

This study demonstrated that CD79B expression varies across DLBCL subtypes in association with molecular features and B-cell differentiation states. Integrative single-cell analyses further revealed a distinct association between GC B-cell physiology and CD79B expression patterns, suggesting that similarities to normal B-cell differentiation stages may contribute to the variation in CD79B expression among DLBCL subtypes. Despite several limitations, our findings provide novel insights into the biological regulation of CD79B expression in DLBCL and improve our understanding of disease heterogeneity.

References

1. Alizadeh AA, Eisen MB, Davis RE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503-511.

2. Venturutti L, Melnick AM. The dangers of deja vu: Memory B cells as the cells of origin of ABC-DLBCLs. Blood. 2020;136(20):2263-2274.

3. Victora GD, Dominguez-Sola D, Holmes AB, Deroubaix S, Dalla-Favera R, Nussenzweig MC. Identification of human germinal center light and dark zone cells and their relationship to human B-cell lymphomas. Blood. 2012;120(11):2240-2248.

4. Alduaij W, Collinge B, Ben-Neriah S, et al. Molecular determinants of clinical outcomes in a real-world diffuse large B-cell lymphoma population. Blood. 2023;141(20):2493-2507.

5. Urata T, Naoi Y, Jiang A, et al. Distribution and clinical impact of molecular subtypes with dark zone signature of DLBCL in a Japanese real-world study. Blood Adv. 2023;7(24):7459-7470.

6. Deeks ED. Polatuzumab vedotin: First global approval. Drugs. 2019;79(13):1467-1475.

7. Song Y, Tilly H, Rai S, et al. Polatuzumab vedotin in previously untreated DLBCL: An asia subpopulation analysis from the phase 3 POLARIX trial. Blood. 2023;141(16):1971-1981.

8. Tilly H, Morschhauser F, Sehn LH, et al. Polatuzumab vedotin in previously untreated diffuse large B-cell lymphoma. N Engl J Med. 2022;386(4):351-363.

9. Morschhauser F, Flinn IW, Advani R, et al. Polatuzumab vedotin or pinatuzumab vedotin plus rituximab in patients with relapsed or refractory non-Hodgkin lymphoma: Final results from a phase 2 randomised study (ROMULUS). Lancet Haematol. 2019;6(5):e254-e265.

10. Sehn LH, Herrera AF, Flowers CR, et al. Polatuzumab vedotin in relapsed or refractory diffuse large B-cell lymphoma. J Clin Oncol. 2020;38(2):155-165.

11. Terui Y, Rai S, Izutsu K, et al. A phase 2 study of polatuzumab vedotin + bendamustine + rituximab in relapsed/refractory diffuse large B-cell lymphoma. Cancer Sci. 2021;112(7):2845-2854.

12. Ennishi D, Jiang A, Boyle M, et al. Double-hit gene expression signature defines a distinct subgroup of germinal center B-cell-like diffuse large B-cell lymphoma. J Clin Oncol. 2019;37(3):190-201.

13. Pfeifer M, Zheng B, Erdmann T, et al. Anti-CD22 and anti-CD79B antibody drug conjugates are active in different molecular diffuse large B-cell lymphoma subtypes. Leukemia. 2015;29(7):1578-1586.

14. Ennishi D, Mottok A, Ben-Neriah S, et al. Genetic profiling of MYC and BCL2 in diffuse large B-cell lymphoma determines cell-of-origin-specific clinical impact. Blood. 2017;129(20):2760-2770.

15. Schmitz R, Wright GW, Huang DW, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. N Engl J Med. 2018;378(15):1396-1407.

16. Wright GW, Huang DW, Phelan JD, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. Cancer Cell. 2020;37(4):551-568.e14.

17. Holmes AB, Corinaldesi C, Shen Q, et al. Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. J Exp Med. 2020;217(10):e20200483.

18. Stoeckius M, Hafemeister C, Stephenson W, et al. Simultaneous epitope and transcriptome measurement in single cells. Nat Methods. 2017;14(9):865-868.

19. Aoki T, Chong LC, Takata K, et al. Single-cell transcriptome analysis reveals disease-defining T-cell subsets in the tumor microenvironment of classic Hodgkin lymphoma. Cancer Discov. 2020;10(3):406-421.

20. 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):eabe6291.

21. Roider T, Seufert J, Uvarovskii A, et al. Dissecting intratumour heterogeneity of nodal B-cell lymphomas at the transcriptional, genetic and drug-response levels. Nat Cell Biol. 2020;22(7):896-906.

22. Siu JHY, Pitcher MJ, Tull TJ, et al. Two subsets of human marginal zone B cells resolved by global analysis of lymphoid tissues and blood. Sci Immunol. 2022;7(69):eabm9060.

23. Steen CB, Luca BA, Esfahani MS, et al. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. Cancer Cell. 2021;39(10):1422-1437.e10.

24. Andreatta M, Carmona SJ. UCell: Robust and scalable single-cell gene signature scoring. Comput Struct Biotechnol J. 2021;19:3796-3798.

25. Cliff ERS, Pelaez GD, Wan F, et al. Cell-of-origin subtype predicts response to polatuzumab vedotin in large B-cell lymphoma. Clin Cancer Res. 2026;32(1):159-168.

26. Dornan D, Bennett F, Chen Y, et al. Therapeutic potential of an anti-CD79b antibody-drug conjugate, anti-CD79b-vc-MMAE, for the treatment of non-Hodgkin lymphoma. Blood. 2009;114(13):2721-2729.

27. Kaimi Y, Shibata M, Fukuhara S, et al. Correlation with CD79B expression and clinicopathological parameters including cell of origin, CD79A and CD19 expression, and CD79B mutation in diffuse large B-cell lymphoma. Hum Pathol. 2025;160:105852.

28. Tilly H, Morschhauser F, Bartlett NL, et al. Polatuzumab vedotin in combination with immunochemotherapy in patients with previously untreated diffuse large B-cell lymphoma: An open-label, non-randomised, phase 1b-2 study. Lancet Oncol. 2019;20(7):998-1010.

29. Corcoran SR, Phelan JD, Choi J, et al. Molecular determinants of sensitivity to polatuzumab vedotin in diffuse large B-cell lymphoma. Cancer Discov. 2024;14(9):1653-1674.

30. Meriranta L, Sorri S, Huse K, et al. Disruption of KLHL6 fuels oncogenic antigen receptor signaling in B-cell lymphoma. Blood Cancer Discov. 2024;5(5):331-352.

31. Varano G, Lonardi S, Sindaco P, et al. B-cell receptor silencing reveals the origin and dependencies of high-grade B-cell lymphomas with MYC and BCL2 rearrangements. Blood Cancer Discov. 2025;6(4):364-393.

32. Basso K, Dalla-Favera R. Germinal centres and B cell lymphomagenesis. Nat Rev Immunol. 2015;15(3):172-184.

33. Ennishi D, Hsi ED, Steidl C, Scott DW. Toward a new molecular taxonomy of diffuse large B-cell lymphoma. Cancer Discov. 2020;10(9):1267-1281.

34. Pasqualucci L. Molecular pathogenesis of germinal center-derived B cell lymphomas. Immunol Rev. 2019;288(1):240-261.

35. Ennishi D, Takata K, Beguelin W, et al. Molecular and genetic characterization of MHC deficiency identifies EZH2 as therapeutic target for enhancing immune recognition. Cancer Discov. 2019;9(4):546-563.

36. Maiolo E, Bellesi S, Campana F, et al. Heterogeneous surface CD79b expression in aggressive B-cell lymphomas assessed by flow cytometry on lymph node biopsies. Cancers (Basel). 2024;16(23):3968.

37. Ang Z, Paruzzo L, Hayer KE, et al. Alternative splicing of its 5'-UTR limits CD20 mRNA translation and enables resistance to CD20-directed immunotherapies. Blood. 2023;142(20):1724-1739.

38. Schuster SJ, Huw L-Y, Bolen CR, et al. Loss of CD20 expression as a mechanism of resistance to mosunetuzumab in relapsed/refractory B-cell lymphomas. Blood. 2024;143(9):822-832.

Table 1. Comparison of baseline clinical and pathological characteristics according to CD79B H-score class in this study cohort

Characteristic	CD79B H-score status			*P**
	ALL, N = 590	High, N = 295	Low, N = 295	
Age, Median(range,years)	72.0 (23.0-95.0)	73.0 (29.0-90.0)	71.0 (23.0-95.0)	**0.015†**
Sex (n, %)				
Female	264 (45%)	141 (48%)	123 (42%)	0.159
Male	326 (55%)	154 (52%)	172 (58%)	
IPI score (n, %)				
<=2	287 (49%)	140 (47%)	147 (50%)	0.621
>2	303 (51%)	155 (53%)	148 (50%)	
ECOG PS score (n, %)				
<=2	535 (91%)	261 (88%)	274 (93%)	0.089
>2	55 (9.3%)	34 (12%)	21 (7.1%)	
Ann Arbor stage (n, %)				
<=2	242 (41%)	121 (41%)	121 (41%)	>0.999
>2	348 (59%)	174 (59%)	174 (59%)	
Lactate dehydrogenase level (n, %)				0.341
Elevated	386 (65%)	199 (67%)	187 (63%)	
Normal	204 (35%)	96 (33%)	108 (37%)	
Biopsy Site (n, %)				
Extranodal	230 (39%)	112 (38%)	118 (40%)	0.673
Nodal	360 (61%)	183 (62%)	177 (60%)	
No. of extranodal sites (n, %)				
<2	468 (79%)	239 (81%)	229 (78%)	0.36
>=2	122 (21%)	56 (19%)	66 (22%)	
Treatment_regimen (n, %)				
R-CHOP	339 (57%)	170 (58%)	169 (57%)	
R-THP-COP	198 (34%)	106 (36%)	92 (31%)	**0.005**
R-EPOCH	44 (7.5%)	12 (4.1%)	32 (11%)	
Other chemotherapy	9 (1.5%)	7 (2.4%)	2 (0.7%)	
Primary refractory (n, %)				
No	418 (71%)	203 (69%)	215 (73%)	0.319
Yes	172 (29%)	92 (31%)	80 (27%)	
Early refractory (n, %)				
No	282 (48%)	124 (42%)	158 (54%)	**0.007**
Yes	308 (52%)	171 (58%)	137 (46%)	
CNS relapse (n, %)				
No	571 (97%)	283 (96%)	288 (98%)	0.351
Yes	19 (3.2%)	12 (4.1%)	7 (2.4%)	
Cell-of-origin (n, %)				
DZsig-pos	33 (5.6%)	6 (2.0%)	27 (9.2%)	
GCB	187 (32%)	73 (25%)	114 (39%)	**<0.001**
ABC	302 (51%)	171 (58%)	131 (44%)	
Unclassified	68 (12%)	45 (15%)	23 (7.8%)	

***P* values are derived from Chi-square test or Fisher's exact test. Bold values indicate statistical significance. †Wilcoxon rank sum test. $DZsig^{pos}$: dark zone signature positive; GCB: germinal center B-cell; ABC: activated B-cell; IPI: International Prognostic Index; ECOG PS: Eastern Cooperative Oncology Group Performance Status; R-CHOP: rituximab, cyclophosphamide, doxorubicin, vincristine and prednisone; R-THP-COP: rituximab, pirarubicin, cyclophosphamide, vincristine and prednisone; R-EPOCH: rituximab, etoposide, prednisone, vincristine and cyclophosphamide; CNS: Central nervous system.**

Figure Legends

Figure 1: Expression of CD79B protein in the study cohort. (A) Representative IHC result of CD79B at high magnification (original magnification ×400, bars represent 50 μm). Values indicate H-scores. The images were captured using a color camera (DP23, Olympus). (B) Frequency density histogram of CD79B H-scores across the DLBCL90 cell-of-origin (COO). (C, D) Boxplots of CD79B H-score among COO subgroups defined by the DLBCL90 assay (C) and Hans' classification (D). Pie charts below the boxplots describe the proportion of CD79B H-score class subpopulations across the COO. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. ** $P < 0.01$, **** $P < 0.0001$. $DZsig^{pos}$: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell.

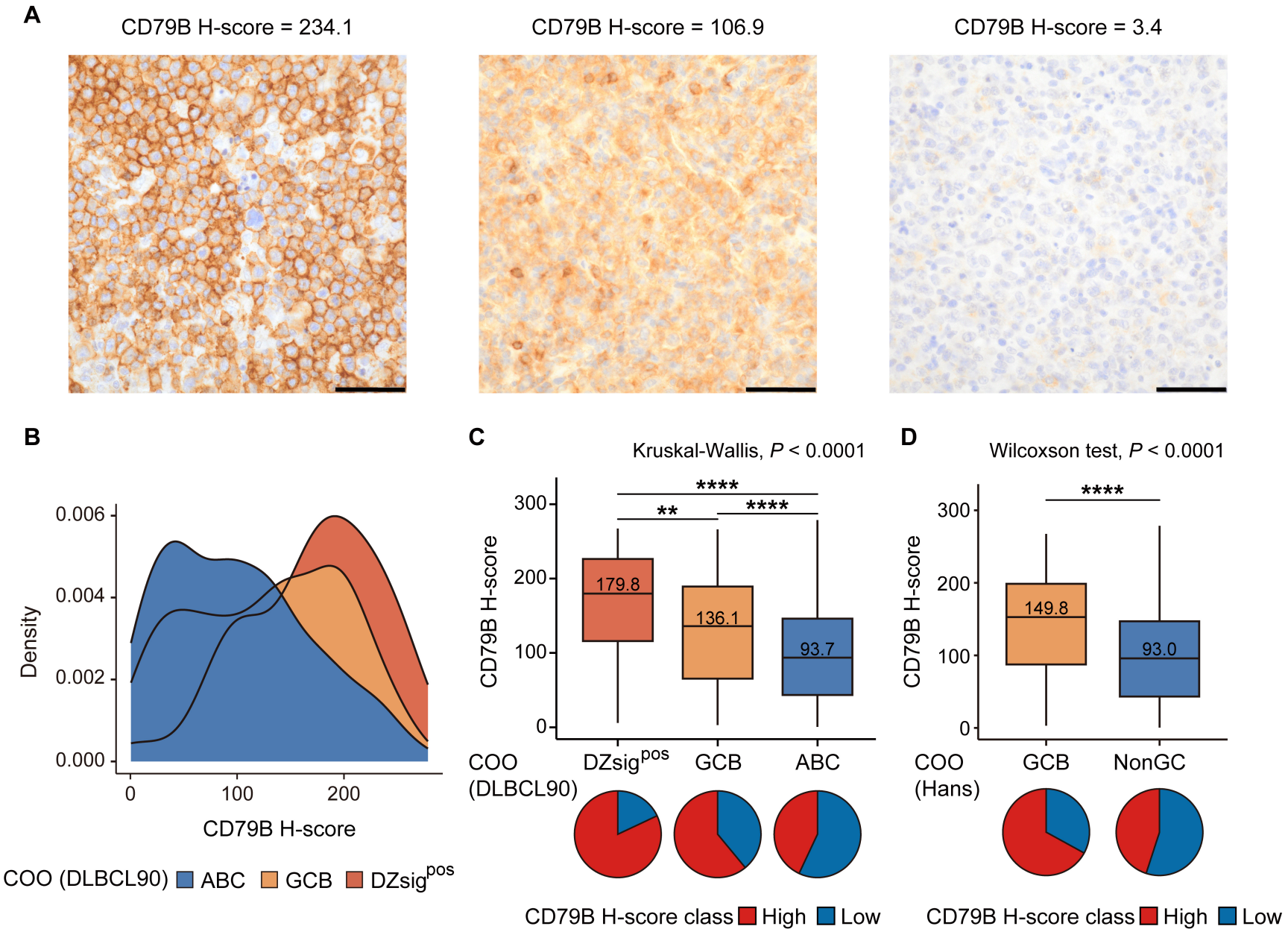
Figure 2: Expression of CD79B protein in the BCC cohort. (A, B) Boxplots of CD79B H-scores among cell-of-origin (COO) subgroups defined by the DLBCL90 assay (A) and Hans' classification (B). Pie charts below the boxplots describe the proportion of CD79B H-score class subpopulations across the COO. (C, D) Boxplots of CD79B H-scores among *CD79B* mutation status (C) and LymphGen subgroups (D) in the entire cohort and in ABC-DLBCL. (E) Boxplots of CD79B H-scores among scGC clusters. Bar charts below the boxplot describe the proportion of CD79B H-score class subpopulations across the scGC cluster. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. $DZsig^{pos}$: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell; GC: germinal center

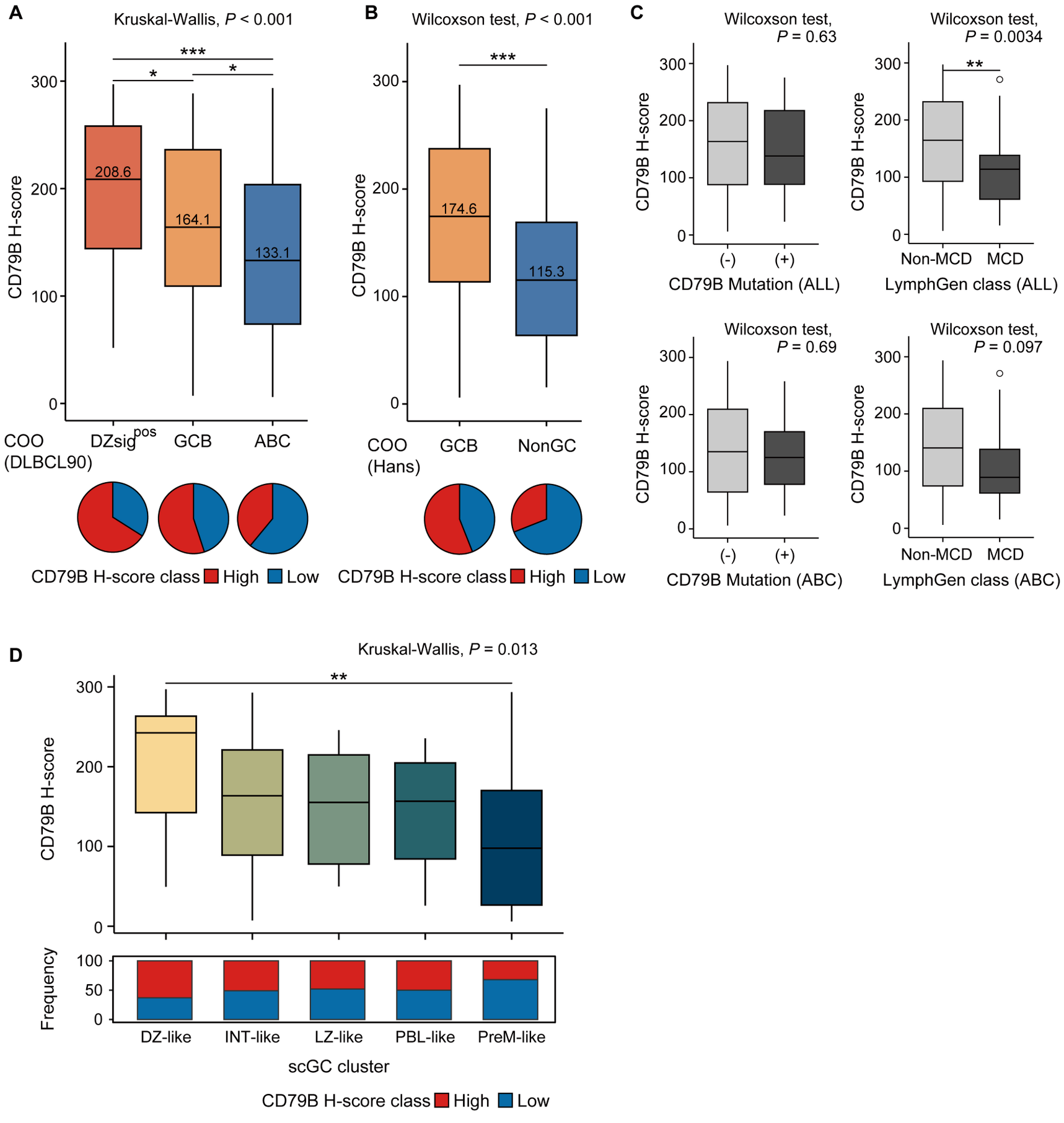
Figure 3: Clinical significance of CD79B expression. (A, B) Kaplan-Meier analyses of overall survival and progression-free survival according to CD79B H-score class in the present study cohort, including the entire cohort (ABC+GCB+DZsig^{pos}), ABC, and non-ABC (GCB+DZsig^{pos}) subgroups. (C, D) Corresponding analyses in the BCC cohort. ABC: activated B-cell.

Figure 4: Relationship between CD79B protein and gene expression in DLBCL. (A) Boxplots of *CD79B* mRNA levels between CD79B H-score classes in the entire cohort. (B) Boxplots of *CD79B* mRNA levels among cell-of-origin (COO) subgroups in BCC and NCI cohorts. (C-E) Boxplots of *CD79B* mRNA level among scGC clusters (C), *CD79B* mutation status (D), and LymphGen subgroups (E) in the BCC cohort and NCI cohort. Note: mRNA levels are log normalized. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, DZsig^{pos}: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell; DZ: dark zone B-cell, INT: intermediate zone B-cell, LZ: light zone B-cell, PreM: precursor memory B-cell, PBL: plasmablast.

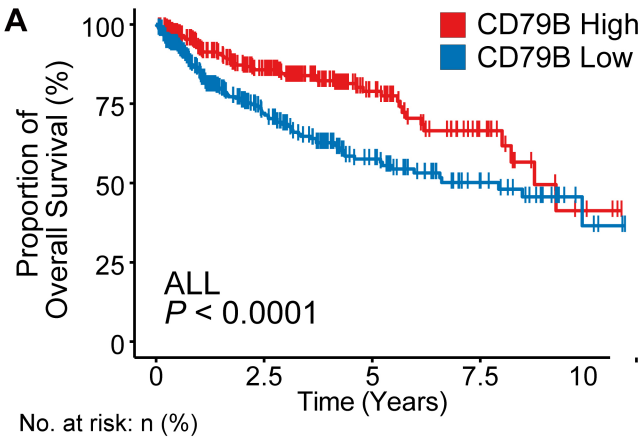
Figure 5: CD79B expression in normal germinal center B-cells at a single-cell level using scRNA-seq and CITE-seq analyses. (A,B) Supervised UMAP projection and cluster labeling in the CITE-seq experiment of 2,447 B-cells from two donors (A) and feature plots showing CD79B gene and protein expression (B). (C) Boxplots of *CD79B* gene expression across scGC B-cell clusters in the CITE-seq dataset. (D) Boxplot of *CD79B* gene expression across scGC B-cell clusters in the merged public scRNA-seq dataset. (E) UMAP projection shows the cluster identification from scRNA-seq profiles of 75,771 cells from the merged public dataset. Eight scGC B-cell clusters were identified by NMF clustering. (F) The feature plot describes *CD79B* mRNA levels. Note: Statistical significance was assessed using the

Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. DZ: dark zone B-cell, INT: intermediate zone B-cell, LZ: light zone B-cell, PreM: precursor memory B-cell, PBL: plasmablast.

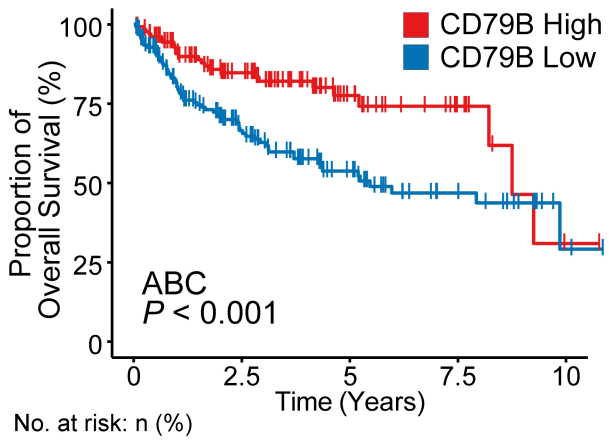

A
CD79B H-score = 234.1
CD79B H-score = 106.9
CD79B H-score = 3.4
B
Density
0.006
0.004
0.002
0.000
0
100
200
CD79B H-score
COO (DLBCL90) ABC GCB DZsig^pos
C
Kruskal-Wallis, P < 0.0001
CD79B H-score
300
200
100
0
****
**
****
179.8
136.1
93.7
COO (DLBCL90)
DZsig^pos
GCB
ABC
CD79B H-score class High Low
D
Wilcoxson test, P < 0.0001
****
149.8
93.0
COO (Hans)
GCB
NonGC
CD79B H-score class High Low


A
Kruskal-Wallis, P < 0.001
CD79B H-score
208.6
164.1
133.1
COO (DLBCL90)
DZsig pos
GCB
ABC
B
Wilcoxson test, P < 0.001
174.6
115.3
COO (Hans)
GCB
NonGC
CD79B H-score class High Low
C
Wilcoxson test, P = 0.63
CD79B Mutation (ALL)
Wilcoxson test, P = 0.0034
LymphGen class (ALL)
Non-MCD
MCD
Wilcoxson test, P = 0.69
CD79B Mutation (ABC)
Wilcoxson test, P = 0.097
LymphGen class (ABC)
D
Kruskal-Wallis, P = 0.013
Frequency
DZ-like
INT-like
LZ-like
PBL-like
PreM-like
scGC cluster
CD79B H-score class High Low

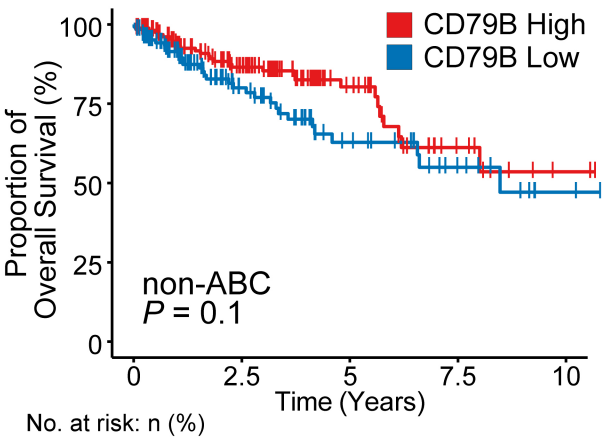
A



No. at risk: n (%)					
High	295 (100)	160 (54)	60 (20)	19 (6)	3 (1)
Low	295 (100)	128 (43)	60 (20)	26 (9)	4 (1)

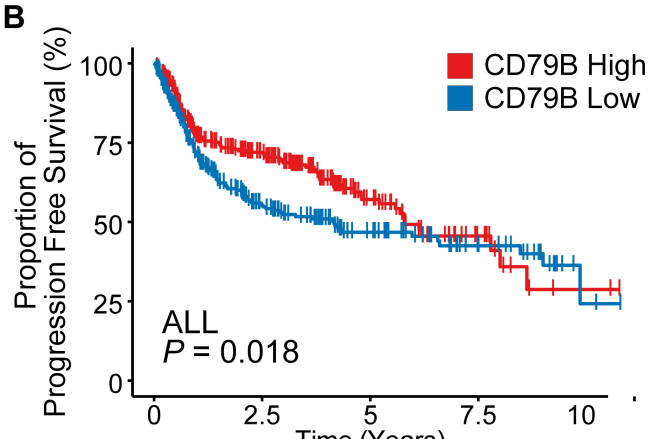


No. at risk: n (%)					
High	131 (100)	70 (53)	25 (19)	9 (7)	1 (1)
Low	171 (100)	75 (44)	37 (22)	16 (9)	2 (1)

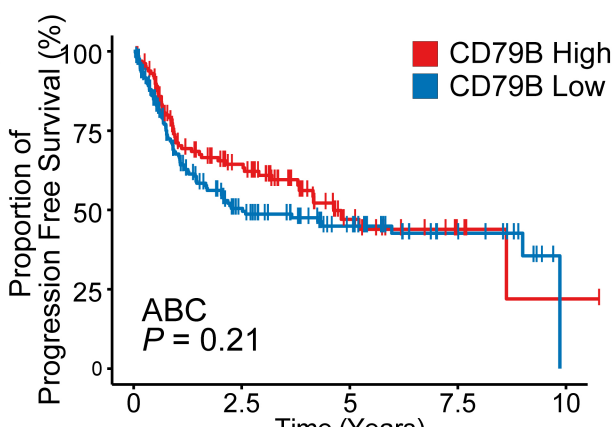


No. at risk: n (%)					
High	164 (100)	90 (55)	35 (21)	10 (6)	2 (1)
Low	124 (100)	53 (43)	23 (19)	10 (8)	2 (2)

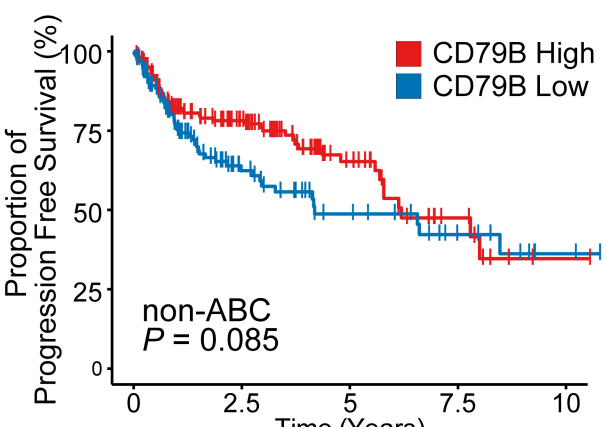
B



No. at risk: n (%)					
High	295 (100)	140 (47)	46 (16)	13 (4)	2 (1)
Low	295 (100)	96 (33)	48 (16)	21 (7)	2 (1)

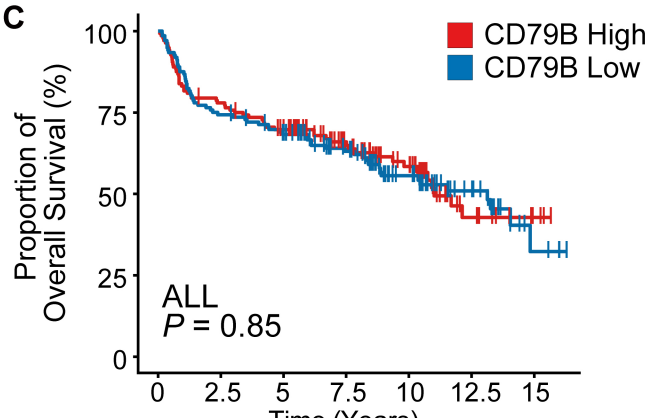


No. at risk: n (%)					
High	131 (100)	57 (44)	16 (12)	4 (3)	1 (1)
Low	171 (100)	56 (33)	29 (17)	12 (7)	0 (0)

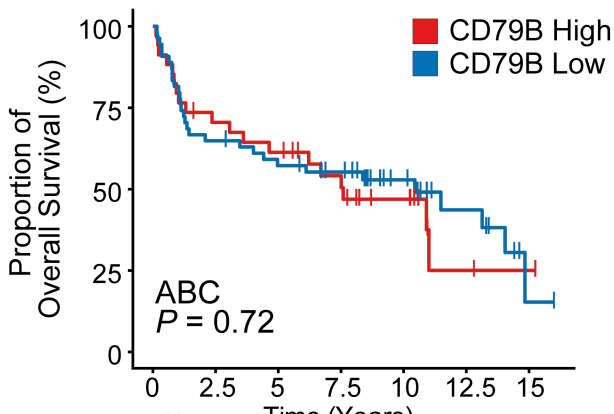


No. at risk: n (%)					
High	164 (100)	83 (51)	30 (18)	9 (5)	1 (1)
Low	124 (100)	40 (32)	19 (15)	9 (7)	2 (2)

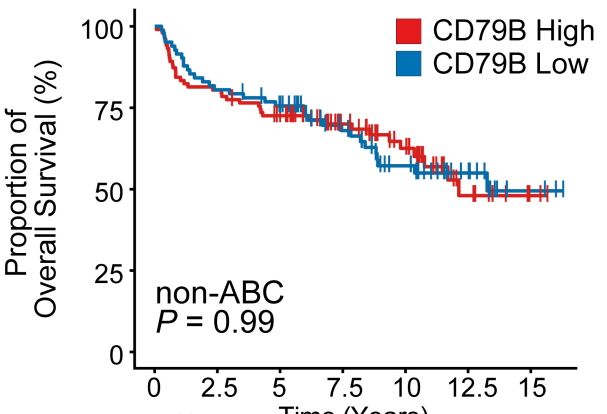
C



No. at risk: n (%)							
High	136(100)	105(77)	91(67)	60(44)	38(28)	12(9)	3(2)
Low	136(100)	101(74)	89(65)	68(50)	42(31)	22(16)	4(3)

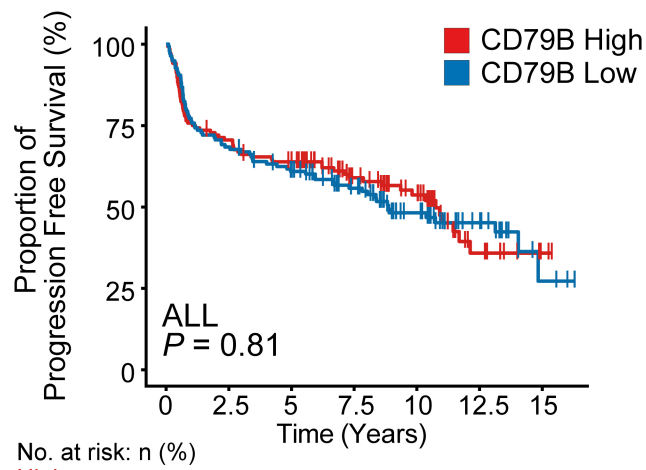


No. at risk: n (%)							
High	34(100)	23(68)	20(59)	15(44)	9(26)	2(6)	1(3)
Low	54(100)	35(65)	30(56)	26(48)	15(28)	8(15)	1(2)

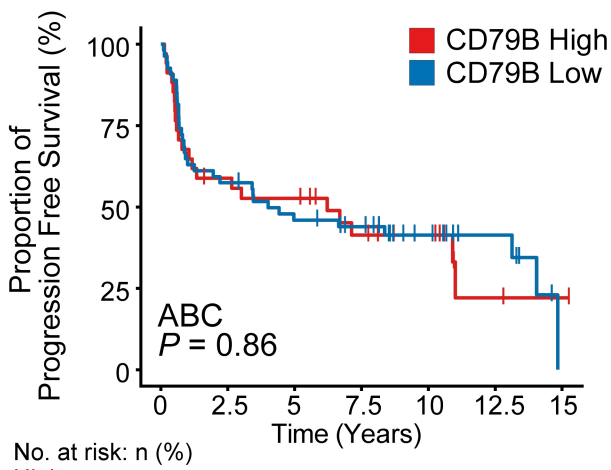


No. at risk: n (%)							
High	102(100)	82(80)	71(70)	45(44)	29(28)	10(10)	2(2)
Low	82(100)	66(80)	59(72)	42(51)	27(33)	14(17)	3(4)

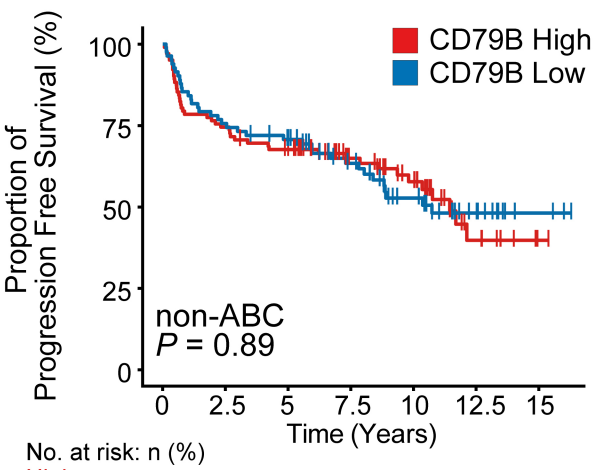
D



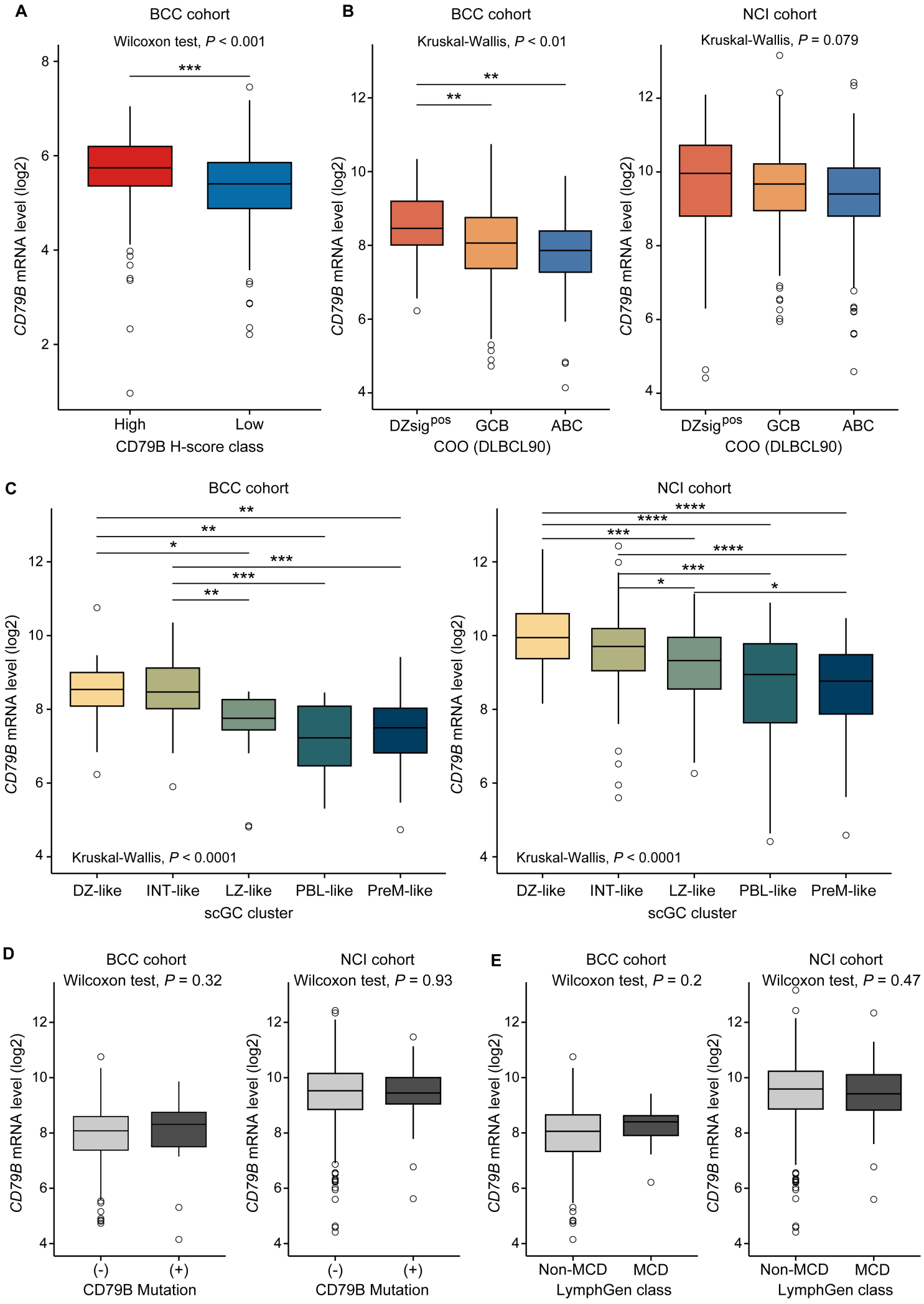
No. at risk: n (%)							
High	136(100)	95(70)	84(62)	52(38)	36(26)	10(7)	2(1)
Low	136(100)	93(68)	79(58)	60(44)	36(26)	19(14)	3(2)

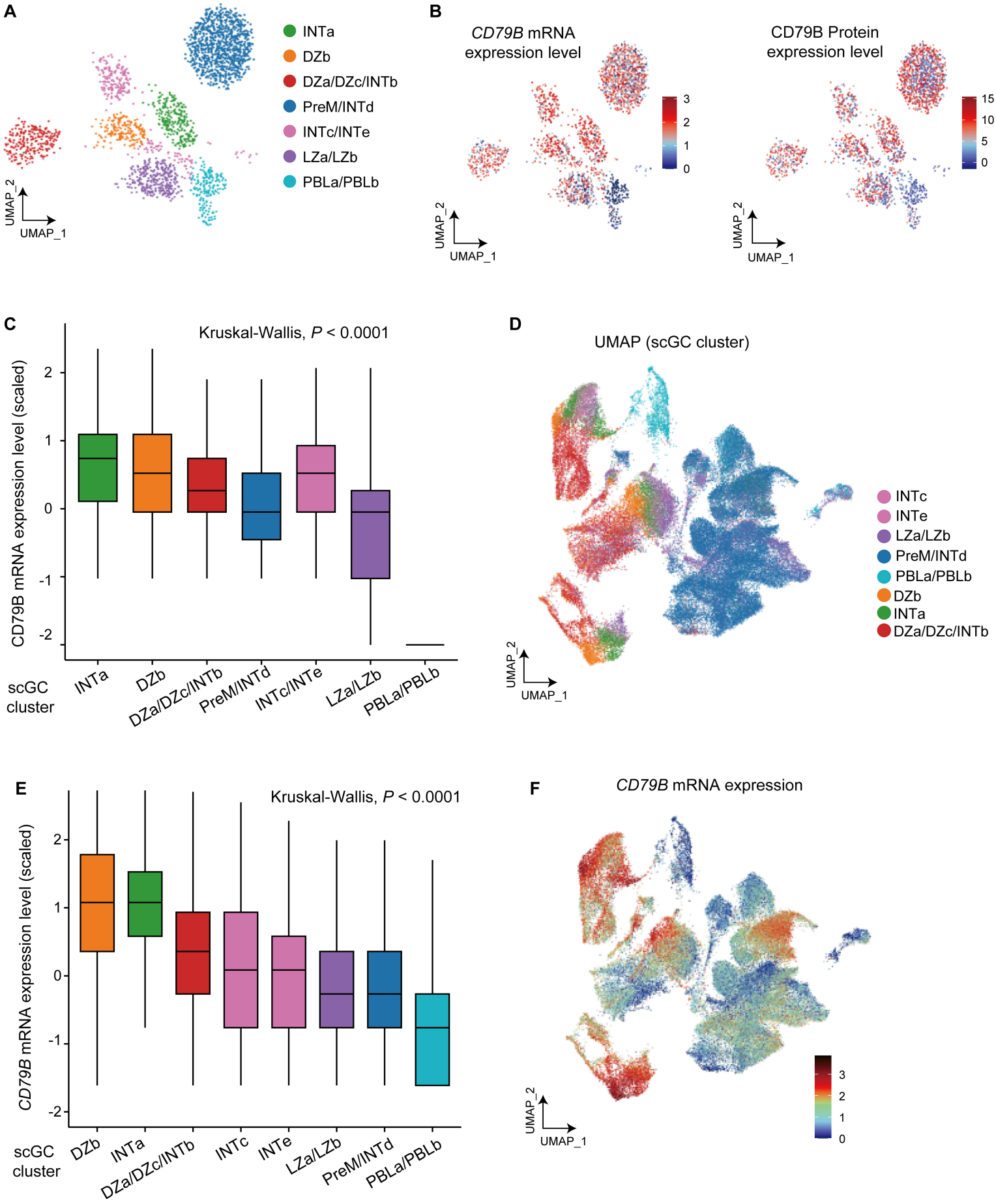


No. at risk: n (%)							
High	34(100)	19(56)	17(50)	11(32)	8(24)	2(6)	1(3)
Low	54(100)	31(57)	24(44)	20(37)	11(20)	6(11)	0(0)



No. at risk: n (%)							
High	102(100)	76(75)	67(66)	41(40)	28(27)	8(8)	1(1)
Low	82(100)	62(76)	55(67)	40(49)	25(30)	13(16)	3(4)


A
BCC cohort
Wilcoxon test, P < 0.001
CD79B mRNA level (log2)
High
Low
CD79B H-score class
B
BCC cohort
Kruskal-Wallis, P < 0.01
DZsig pos
GCB
ABC
COO (DLBCL90)
NCI cohort
Kruskal-Wallis, P = 0.079
C
BCC cohort
Kruskal-Wallis, P < 0.0001
DZ-like
INT-like
LZ-like
PBL-like
PreM-like
scGC cluster
NCI cohort
D
BCC cohort
Wilcoxon test, P = 0.32
NCI cohort
Wilcoxon test, P = 0.93
(-)
(+)
CD79B Mutation
E
BCC cohort
Wilcoxon test, P = 0.2
NCI cohort
Wilcoxon test, P = 0.47
Non-MCD
MCD
LymphGen class


A
INTa
DZb
DZa/DZc/INTb
PreM/INTd
INTc/INTe
LZa/LZb
PBLa/PBLb
UMAP_2
UMAP_1
B
CD79B mRNA expression level
CD79B Protein expression level
C
Kruskal-Wallis, P < 0.0001
CD79B mRNA expression level (scaled)
scGC cluster
D
UMAP (scGC cluster)
E
Kruskal-Wallis, P < 0.0001
CD79B mRNA expression level (scaled)
F
CD79B mRNA expression

Supplementary Data

Gradient of CD79B expression in diffuse large B-cell lymphoma corresponds to stages of germinal center B-cell differentiation

Yusuke Naoi, et al.

Contents

1. **Supplementary Methods**
2. **Supplementary Tables (legends)**
3. **Supplementary Figures**

1. Supplementary Methods

Immunohistochemistry (IHC) staining and analysis

The diameter of each tissue microarray (TMA) core was 2.0 mm. All samples were obtained from formalin-fixed, paraffin-embedded (FFPE) blocks. Staining intensities using different clones were confirmed using the antibodies specific for CD79B (EPR6861, Abcam, Cambridge, UK) and CD79B (B29/123, Santa Cruz Biotechnology, Dallas, Texas; Supplemental Figure 1). All the stained slides were scanned using an Axio Scan Z1 (Carl Zeiss) at 200-times magnification with a pixel size of 0.22. All scanned images were analyzed using QuPath software(1). Single-cell detection was conducted using StarDist(2), a machine-learning method. Each cell was scored from 0 (negative) to 3 (strongly positive) according to 3,3'-Diaminobenzidine (DAB) intensity. We calculated the H-score(3-5), which is a continuous value ranging from 0 to 300. In the following equation, *H* represents the H-score of a target antigen for each TMA core *X*:

$$H(X)_{target} = \sum_{i=0}^{3}(iP_i) \quad [1]$$

where *i* represents the DAB intensity score, a discrete value ranging from 0 to 3 (increasing by 1), and *Pi* represents the cell percentage corresponding to each DAB intensity score *i*.

Additional immunohistochemical analyses for IgM and IgG were performed in a subset of cases, and staining was evaluated using the same image analysis pipeline described above (Table S2).

Public bulk DLBCL dataset

Raw count data were normalized using the edgeR package R(6) before performing downstream analysis. Patients with insufficient clinical information were excluded. We labeled *CD79B* mutational status as "(+)" if the sample had any type of *CD79B* mutation, whereas "(-)" indicates patients without *CD79B* mutation. *CD79B* mutation was observed in 24 (11%) of the 223 patients available for mutational data in the BCC cohort. In the NCI cohort, 63 patients (16%) harbored *CD79B* mutations out of 388 individuals available for *CD79B* mutation data. To simplify comparison, we defined the binary classes of the LymphGen cluster as follows: "MCD" and "Non-MCD." For the BCC cohort, MCD and MCD/A53 were classified as "MCD" (n = 20), and the others were classified as "Non-MCD" (n = 201). Within the NCI cohort, MCD, MCD/A53, and MCD/ST2 were labeled as "MCD" (n = 87), and the rest were labeled as "Non-MCD" (n = 370).

Single-cell level analysis of reactive lymphoid tissues

In the CITE-seq experiment, mononuclear cells were isolated as cell suspensions by dissociation in a Tumor Tissue Dissociation Reagent (BD Biosciences) using a gentleMACS dissociator (Miltenyi Biotec). A totalseq-C antibody cocktail (Biolegend) was used, and the cell suspensions were diluted to a concentration of 1,000 cells/μL. Library sequencing was performed using the NovaSeq6000 System (Illumina). The FASTQ files were aligned to the 10x provided reference genome using 10x Genomics Cell Ranger software v7.0.0 to create unique molecular identifier count tables of gene expression of the samples.

The computational steps used to analyze the scRNA-seq data are shown in Figure 5A. Briefly, cells expressing fewer genes (≤ 200) or exhibiting higher levels of mitochondrial transcripts (> 5%) were

excluded. For the downstream analysis, count samples were normalized and scaled using the "SCTransform v2" function in the Seurat R package(7, 8). Protein expression levels were normalized using the "dsb" function in the Seurat R(9). Linear dimensional reduction was performed using principal component analysis (PCA). The top 50 principal components were further summarized using UMAP to present the data in a two-dimensional panel. Clusters were identified by an SNN modularity optimization-based clustering algorithm(10) using the "FindClusters" function in the Seurat R package. Non-negative matrix factorization (NMF)(11) was performed using UCell scores(12) as a reference. To calculate UCell scores according to single-cell germinal center B-cell subclusters, we used the signature gene set reported by Holmes et al., which is available online at https://doi.org/10.1084/jem.20200483. For visualization, supervised UMAP projection was applied in Fig. 5B, C, and Fig. S6 using the "umap-learn" package (version 0.5.5) in Python.

References

1. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):16878.
2. Uwe Schmidt MW, Coleman Broaddus, Gene Myers. Cell Detection with Star-convex Polygons. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2018.
3. Sehn LH, Herrera AF, Flowers CR, Kamdar MK, McMillan A, Hertzberg M, et al. Polatuzumab Vedotin in Relapsed or Refractory Diffuse Large B-Cell Lymphoma. J Clin Oncol. 2020;38(2):155-65.
4. Terui Y, Rai S, Izutsu K, Yamaguchi M, Takizawa J, Kuroda J, et al. A phase 2 study of polatuzumab vedotin + bendamustine + rituximab in relapsed/refractory diffuse large B-cell lymphoma. Cancer Sci. 2021;112(7):2845-54.
5. Tilly H, Morschhauser F, Bartlett NL, Mehta A, Salles G, Haioun C, et al. Polatuzumab vedotin in combination with immunochemotherapy in patients with previously untreated diffuse large B-cell lymphoma: an open-label, non-randomised, phase 1b-2 study. Lancet Oncol. 2019;20(7):998-1010.

6. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139-40.
7. Choudhary S, Satija R. Comparison and evaluation of statistical error models for scRNA-seq. Genome Biol. 2022;23(1):27.
8. Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. Genome Biol. 2019;20(1):296.
9. Mule MP, Martins AJ, Tsang JS. Normalizing and denoising protein expression data from droplet-based single cell profiling. Nat Commun. 2022;13(1):2099.
10. Waltman L, van Eck NJ. A smart local moving algorithm for large-scale modularity-based community detection. The European Physical Journal B. 2013;86(11):471.
11. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. BMC Bioinformatics. 2010;11:367.
12. Andreatta M, Carmona SJ. UCell: Robust and scalable single-cell gene signature scoring. Comput Struct Biotechnol J. 2021;19:3796-8.

2. Supplementary Tables (legends)

Table S1. Clinical and phenotypic characteristic of the OHSG cohort (n = 576). See Excel file.

Table S2. Antibodies used for immunohistochemical analyses. See Excel file.

Table S3. Clinical and phenotypic characteristic of the entire BCC cohort with CD79B IHC result. See Excel file.

Table S4. Overview of bulk and single-cell gene expression datasets analyzed in this study. See Excel file.

Table S5. Comparison of CD79B staining pattern among cell-of-origin (COO) in the present cohort. See Excel file.

Table S6. Concordance of CD79B staining positivity across different antibody clones in the OHSG cohort subset. See Excel file.

Table S7. Baseline clinical and pathological characteristics according to CD79B H-score class in the BCC cohort. See Excel file.

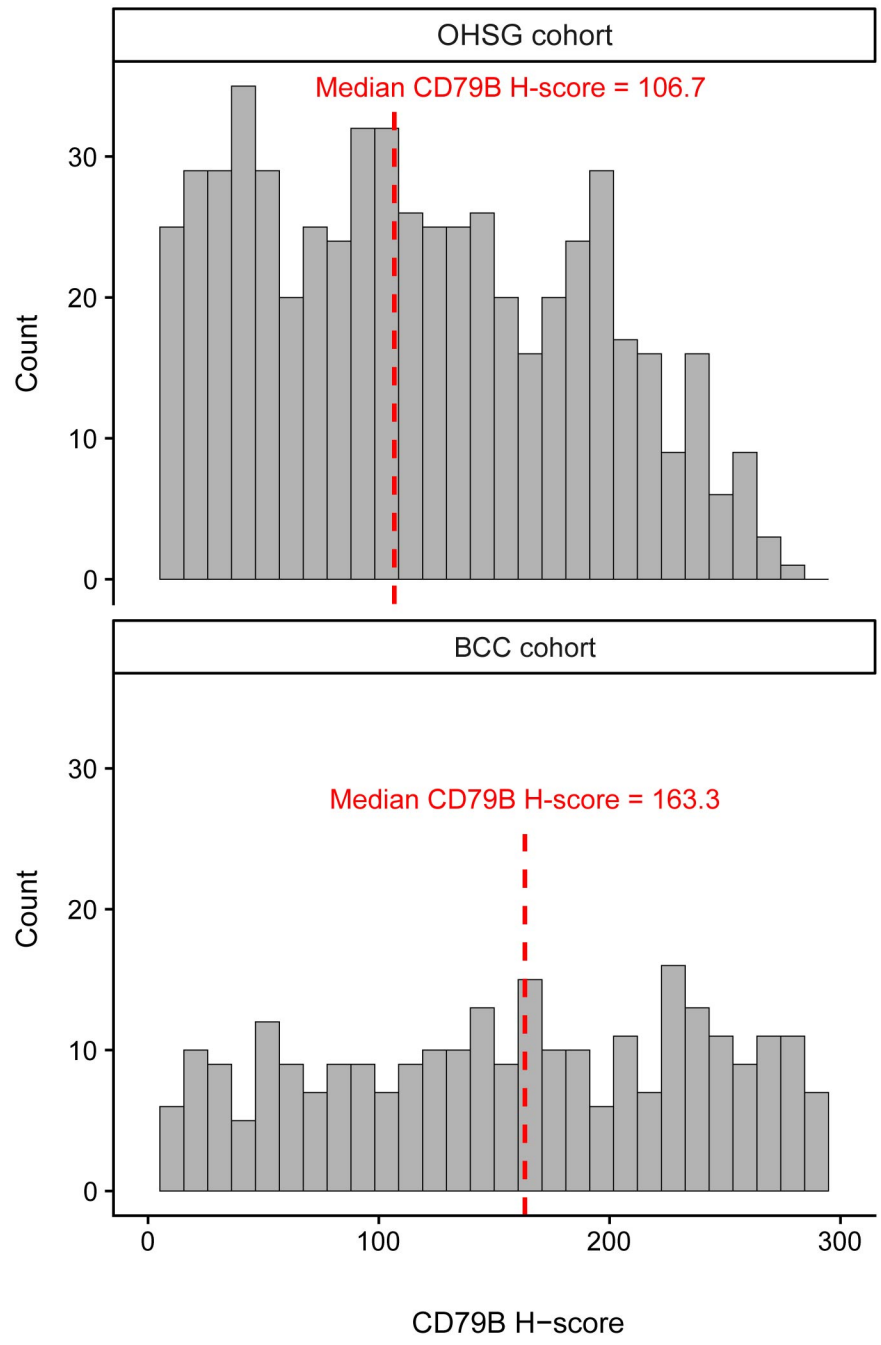
Table S8. Association between CD79B mutation status and CD79B immunohistochemical positivity in the BCC cohort. See Excel file.

Table S9. Comparison of CD79B H-score class across LymphGen subclass in BCC cohort. See Excel file.

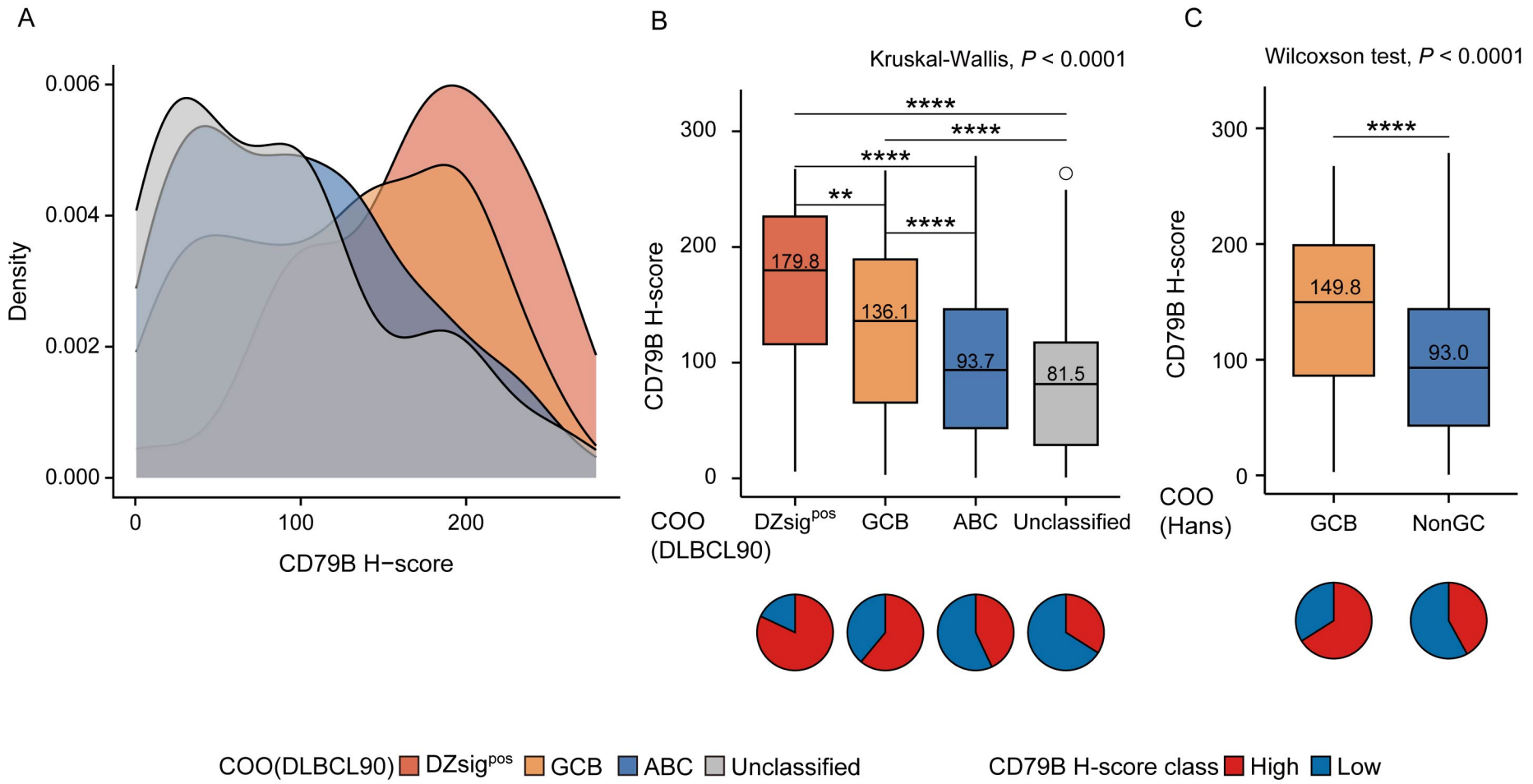
Table S10. Comparison of CD79B H-score class across scGC cluster in BCC cohort. See Excel file.

Table S11. Cox regression analysis for overall survival and progression-free survival in the overall cohort and ABC-DLBCL according to CD79B expression defined by H-score (continuous and dichotomized) and staining positivity in the OHSG cohort. See Excel file.

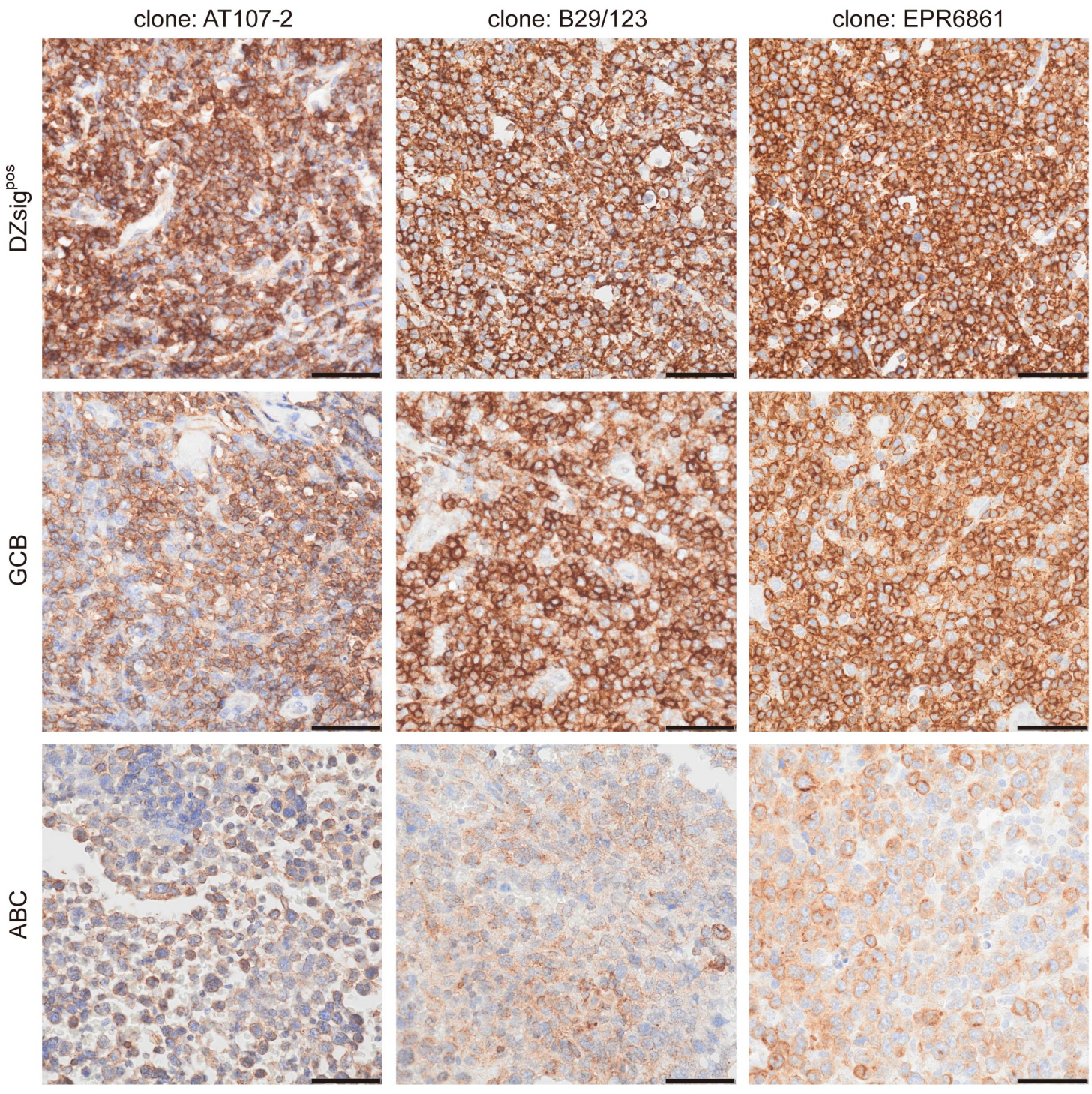
Table S12. Cox regression analysis for overall survival and progression-free survival in the overall cohort and ABC-DLBCL according to CD79B expression defined by H-score (continuous and dichotomized) and staining positivity in the BCC cohort. See Excel file.



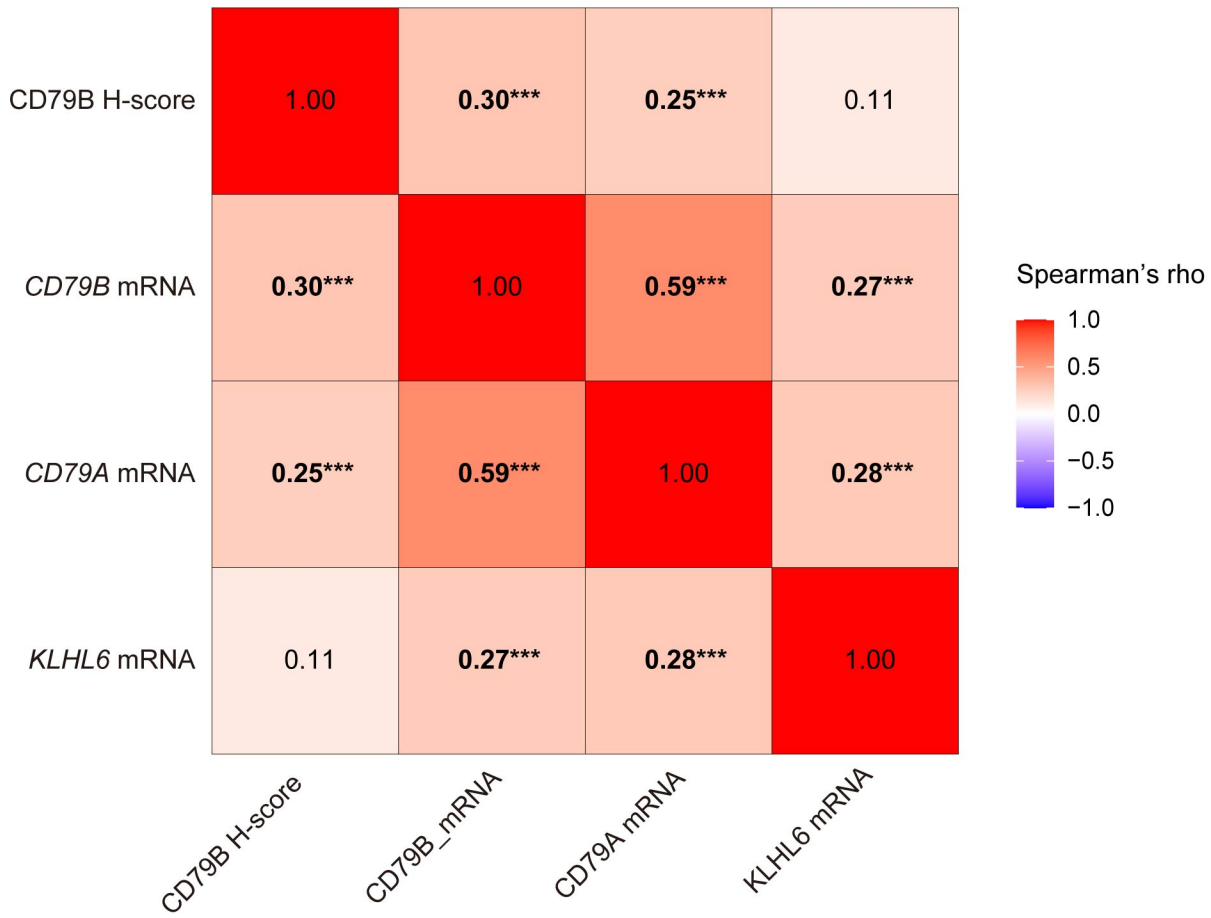
Supplemental Figure 1: Distribution of CD79B H-scores in the OHSG and BCC cohorts. Histograms showing the distribution of CD79B H-scores in the OHSG (top) and BCC (bottom) cohorts. Red dashed lines indicate the cohort-specific median values (OHSG: 106.7; BCC: 163.3), which were used as cutoffs for dichotomization into $CD79B^{high}$ and $CD79B^{low}$ groups.



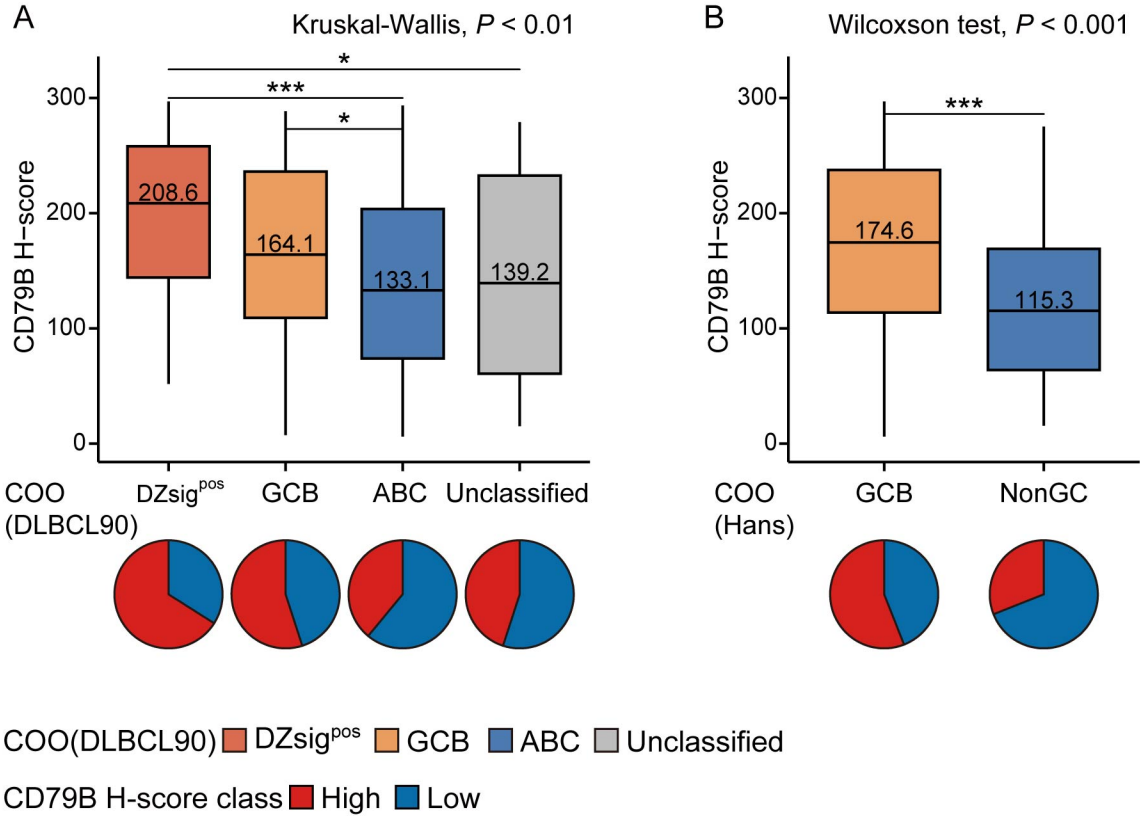
Supplemental Figure 2: Expression of CD79B protein in the study cohort including Unclassified COO. (A) Frequency density histogram of CD79B H-scores across the DLBCL90 cell-of-origin (COO). (B, C) Boxplots of CD79B H-score among COO subgroups defined by the DLBCL90 assay (B) and Hans' classification (C). Pie charts below the boxplots describe the proportion of CD79B H-score class subpopulations across the COO. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. ** $P < 0.01$, **** $P < 0.0001$. DZsig^{pos}: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell.



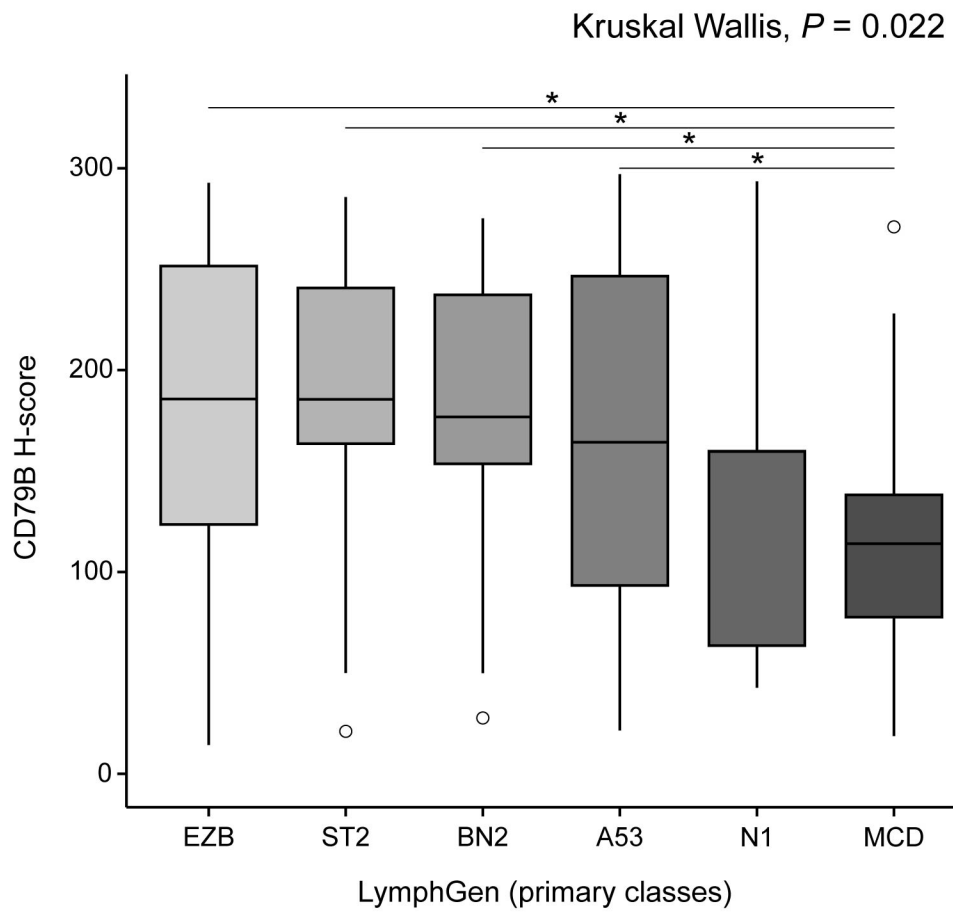
Supplemental Figure 3: Immunohistochemical results of different clones of CD79B antibodies. Clone AT107-2, which was primarily used in our study, was verified using two other clones, specifically B29/123 and EPR6861. Original magnification is ×400 and bars indicate 50 μm. All images were captured by a color camera (DP23, Olympus).



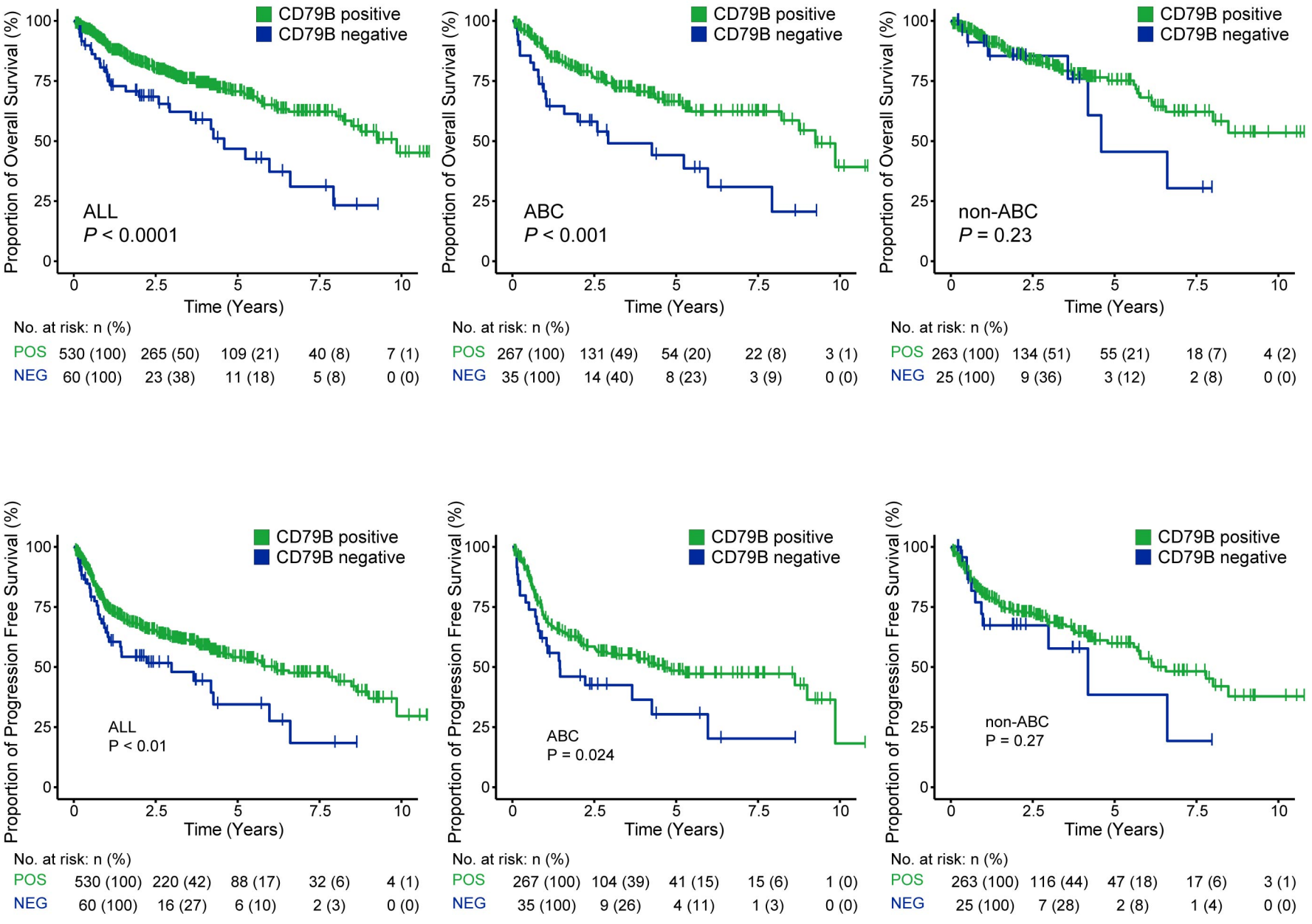
Supplemental Figure 4: Correlation matrix of CD79B H-score, *CD79B* mRNA, *CD79A* mRNA, and *KLHL6* mRNA expression in the BCC cohort. Heatmap showing Spearman's correlation coefficients between protein and transcriptomic expression levels. Numbers indicate Spearman's rho values. *** $P < 0.001$.



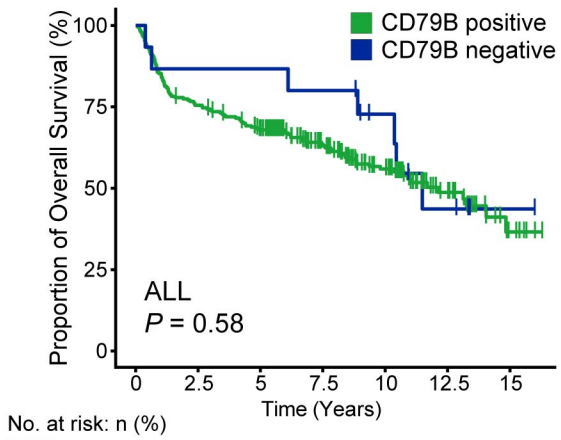
Supplemental Figure 5: Expression of CD79B protein in the BCC cohort including Unclassified COO. (A, B) Boxplots of CD79B H-score among COO subgroups defined by the DLBCL90 assay (A) and Hans' classification (B). Pie charts below the boxplots describe the proportion of CD79B H-score class subpopulations across the COO. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. DZsig^{pos}: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell.



Supplemental Figure 6: CD79B H-score in BCC cohort among LymphGen classification. A boxplot of CD79B H-scores among LymphGen subgroups in the entire BCC cohort. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed, and only statistically significant results are shown for clarity. * $P < 0.05$.

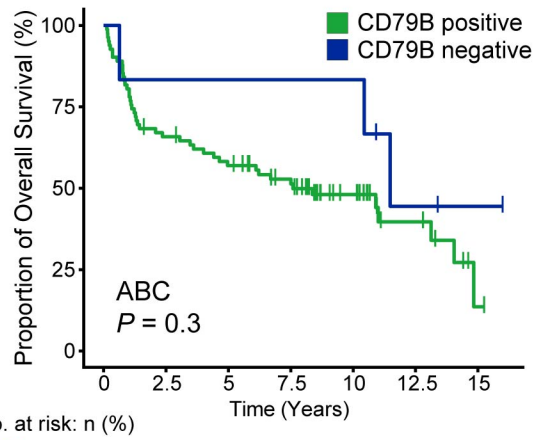


Supplemental Figure 7: Kaplan-Meier curves of OS and PFS based on CD79B staining positivity in the OHSG cohort. Results are shown for the entire cohort, as well as for ABC and non-ABC subtypes (GCB+DZsig^{pos}+Unclassified). DZsig^{pos}: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell.



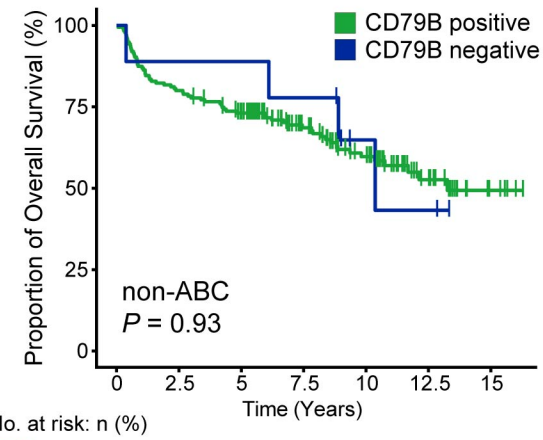
No. at risk: n (%)

POS	257(100)	193(75)	167(65)	116(45)	72(28)	30(12)	6(2)
NEG	15(100)	13(87)	13(87)	12(80)	8(53)	4(27)	1(7)



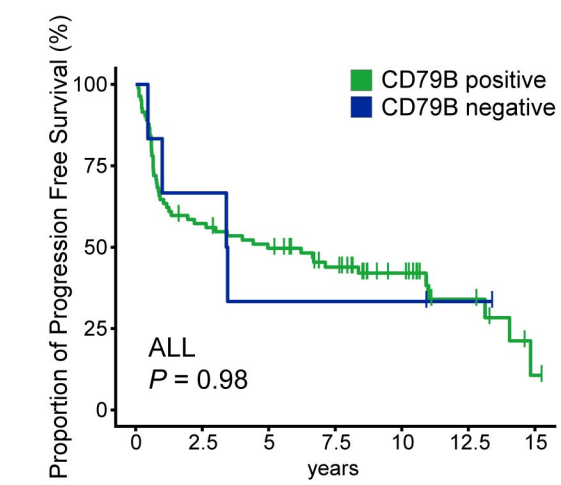
No. at risk: n (%)

POS	82 (100)	53(65)	45(55)	36(44)	19(23)	8(10)	1(1)
NEG	6 (100)	5(83)	5(83)	5(83)	5(83)	2(33)	1(17)



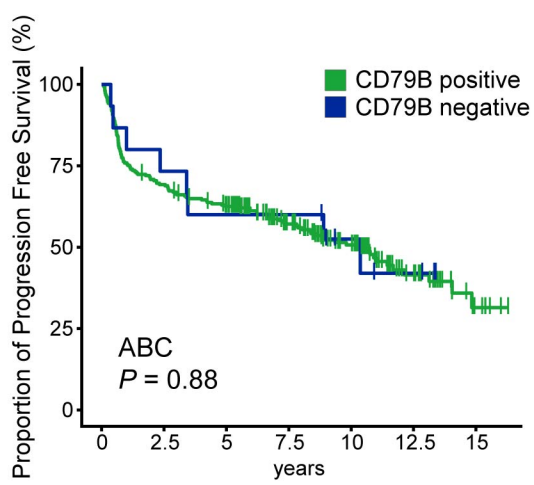
No. at risk: n (%)

POS	175(100)	140(80)	122(70)	80(46)	53(30)	22(13)	5(3)
NEG	9(100)	8(89)	8(89)	7(78)	3(33)	2(22)	0(0)



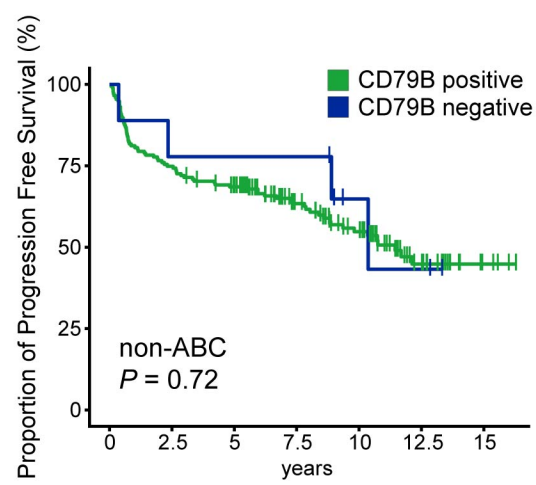
No. at risk: n (%)

POS	82(100)	46(56)	39(48)	29(35)	17(21)	7(9)	1(1)
NEG	6(100)	4(67)	2(33)	2(33)	2(33)	1(17)	0(0)



No. at risk: n (%)

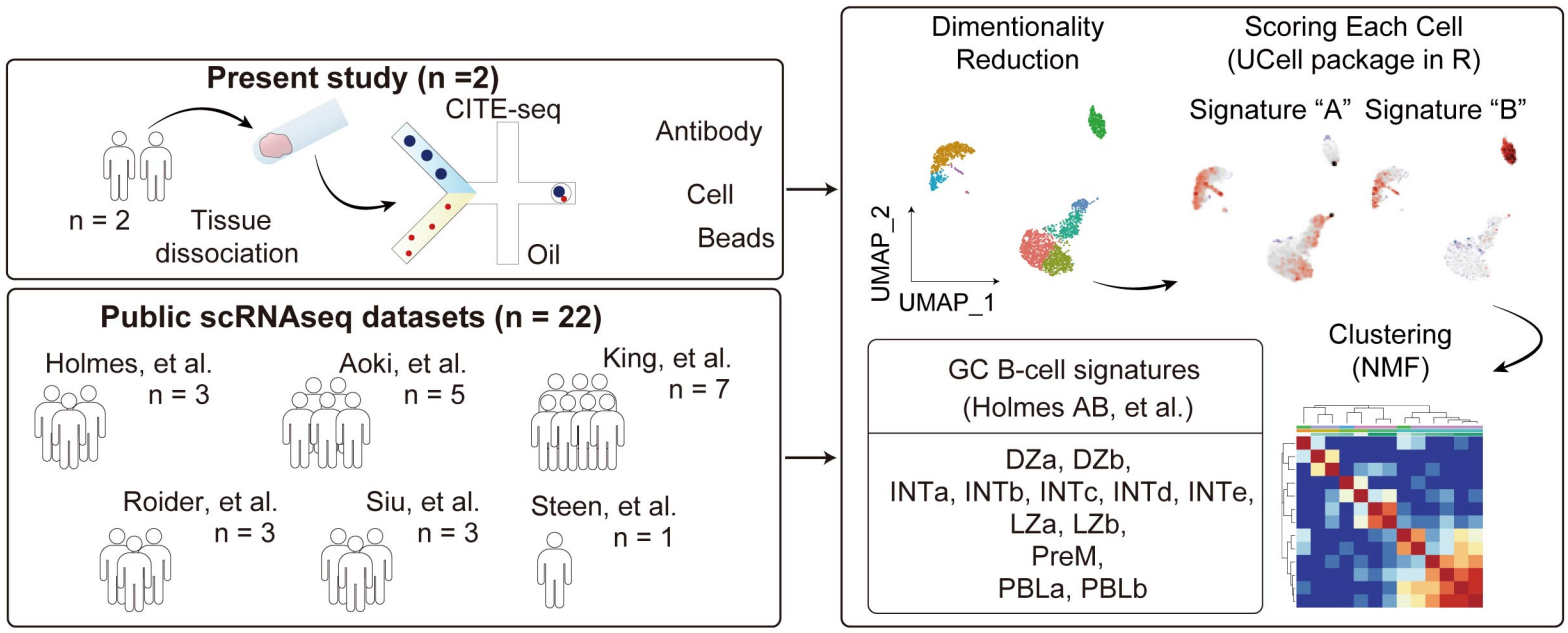
POS	257(100)	177(69)	154(60)	103(40)	67(26)	26(10)	5(2)
NEG	15(100)	11(73)	9(60)	9(60)	5(33)	3(20)	0(0)



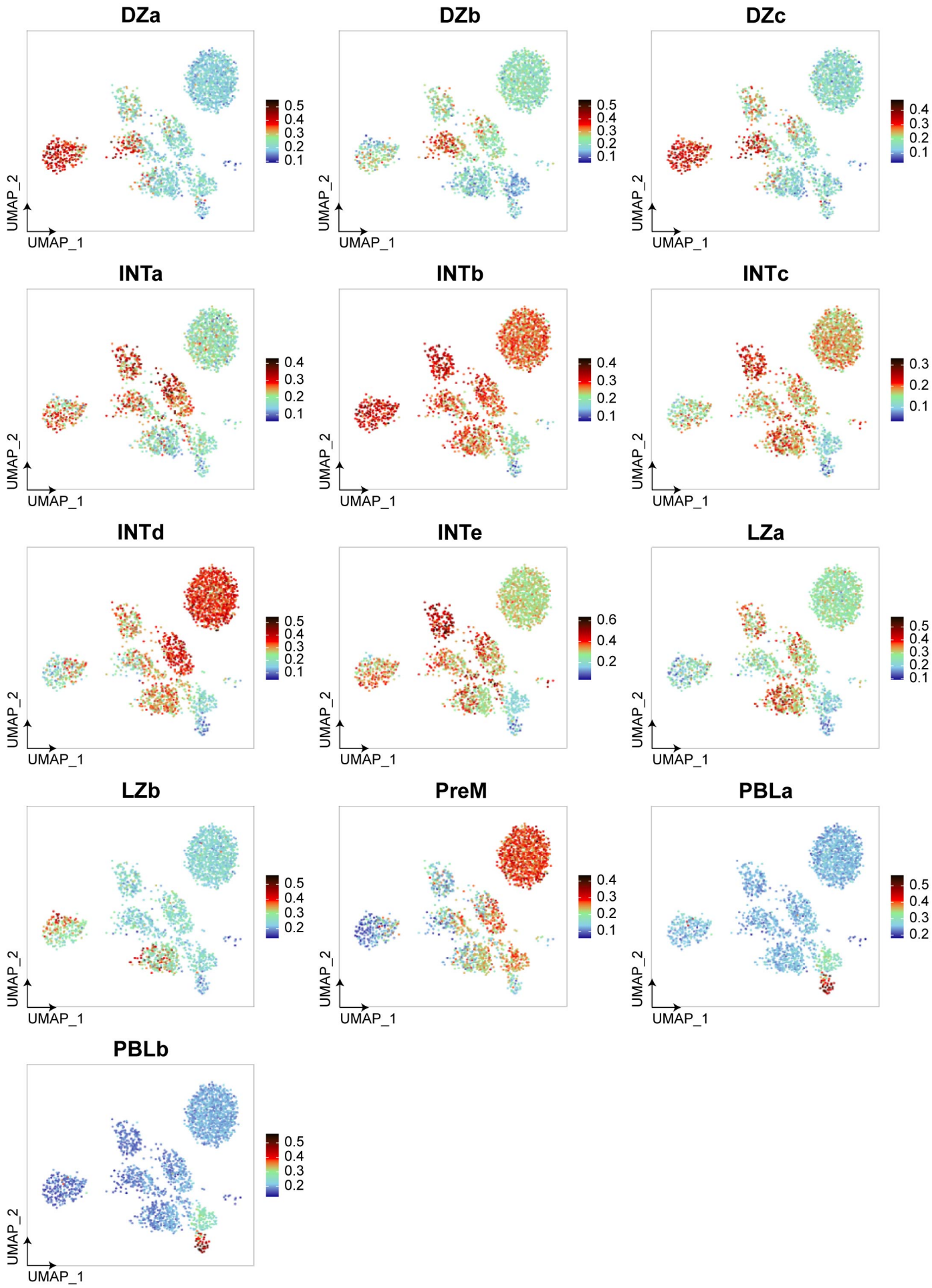
No. at risk: n (%)

POS	175(100)	131(75)	115(66)	74(42)	50(29)	19(11)	4(2)
NEG	9(100)	7(78)	7(78)	7(78)	3(33)	2(22)	0(0)

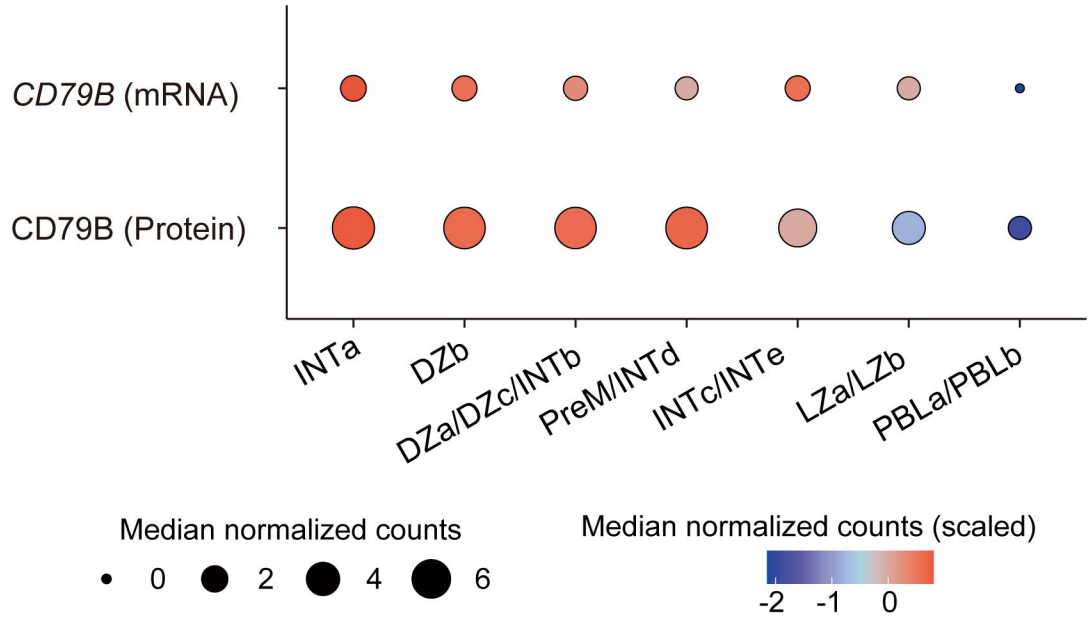
Supplemental Figure 8: Kaplan-Meier curves of OS and PFS based on CD79B staining positivity in the BCC cohort. Results are shown for the entire cohort, as well as for ABC and non-ABC subtypes (GCB+DZsig^{pos}+Unclassified). DZsig^{pos}: dark-zone signature-positive; GCB: germinal center B-cell; ABC: activated B-cell.



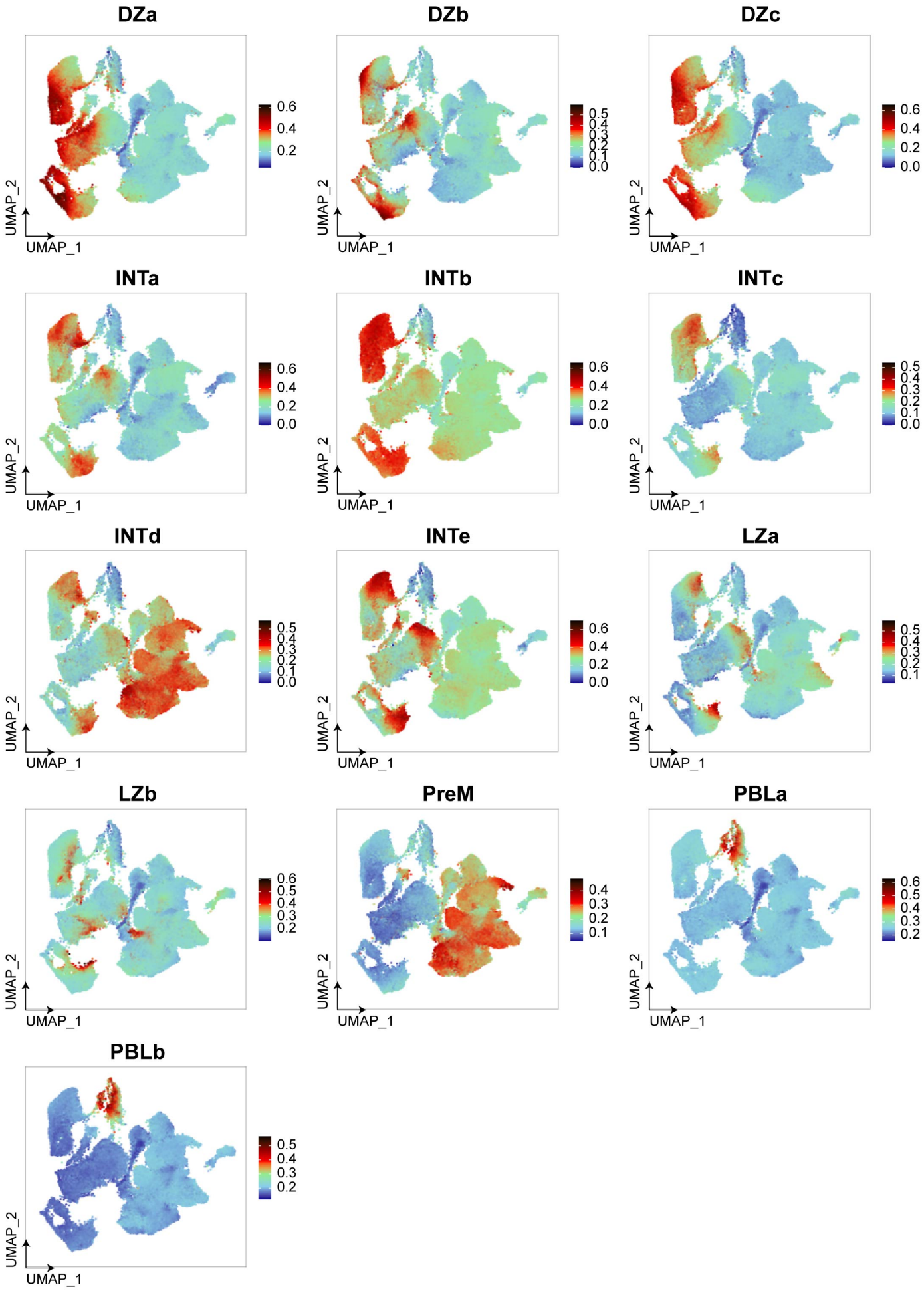
Supplemental Figure 9: Schematic workflow of single-cell experiments and downstream analyses. We conducted cellular indexing of transcriptomes and epitopes by sequencing (CITE-seq) using samples from two patients with reactive lymphoid hyperplasia. In addition, 22 single-cell RNA-sequencing (scRNA-seq) samples from six different publicly available datasets were analyzed.



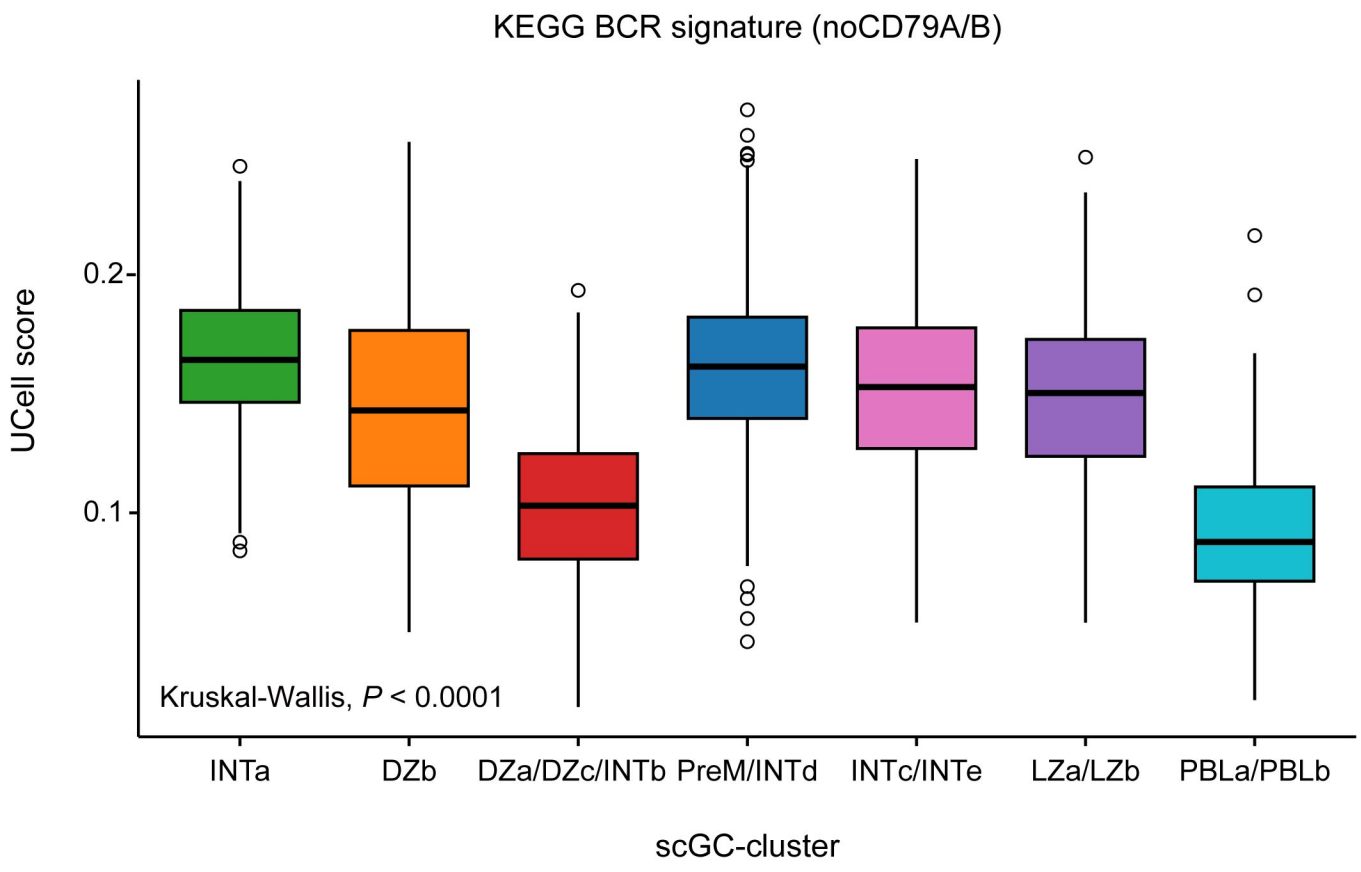
Supplemental Figure 10: UCell scores using 13 germinal center B-cell signatures in our CITE-seq cohort. Each cell on the feature plots is scored according to 13 germinal center B-cell signatures using the R UCell package. DZ: dark zone B-cell, INT: intermediate zone B-cell, LZ: light zone B-cell, PBL: plasmablasts, PreM: precursor memory B-cells.



Supplemental Figure 11. CD79B expression measured using the in-house CITE-seq dataset. A dot plot demonstrates the comparison of CD79B protein expression and *CD79B* gene expression across the scGC B-cell clusters. The size of each dot represents the median normalized counts, and the color indicates the scaled median normalized counts. DZ: dark zone B-cell, INT: intermediate zone B-cell, LZ: light zone B-cell, PBL: plasmablasts, PreM: precursor memory B-cells.



Supplemental Figure 12: UCell scores using 13 germinal center B-cell signatures in public datasets. Each cell on the feature plots is scored according to 13 germinal center B-cell signatures using the R UCell package. DZ: dark zone B-cell, INT: intermediate zone B-cell, LZ: light zone B-cell, PBL: plasmablasts, PreM: precursor memory B-cells.



Supplemental Figure 13: BCR signaling activity across GC B-cell subsets. BCR signaling activity was quantified using a KEGG BCR signaling gene signature excluding CD79A and CD79B. UCell scores are shown across scGC clusters. Statistical significance was assessed using the Kruskal-Wallis test followed by Dunn's multiple comparison test with Benjamini-Hochberg correction. All pairwise comparisons were performed. DZ: dark zone B-cell, INT: intermediate zone B-cell, LZ: light zone B-cell, PreM: precursor memory B-cell, PBL: plasmablast.