Neoplasia 60 (2025) 101119

Contents lists available at ScienceDirect

Neoplasia

journal homepage: www.elsevier.com/locate/neo

Original Research

Comparison of differences in transcriptional and genetic profiles between intra-central nervous system and extra-central nervous system large B-cell lymphoma

Shu Wang [a,1], Hong Chen [b,1], Bo Dai [a,1], Kang Zheng [c,1], Jiajun Zheng [c], Yuqi Zhu [d], Yan Yuan [a], Tianling Ding [a], Qian Wang [a], Liqian Xie [c], Rui Feng [c,e,f], Fengping Zhu [c], Jianbin Xiang [g], Weiqun Ding [h], Hong Ding [i], Yuan Li [d], Xiaodong Gu [j], Kunpeng Wu [a], Yifan Yuan [c,e], Jianping Song [c,e,f,k,l], Dongxiao Zhuang [c], Haoshu Zhong [a], Hanfeng Wu [m,***], Ying Mao [c,**], Tong Chen [a,*]

[a] Department of Hematology, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[b] Department of Pathology, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[c] Department of Neurosurgery, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[d] Department of Radiology, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[e] National Center for Neurological Disorders, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[f] Neurosurgical Institute of Fudan University, Huashan Hospital, Shanghai 200040, PR China
[g] Department of General Surgery, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[h] Department of Gastroenterology, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[i] Department of Ultrasound, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[j] Department of Gastrointestinal Surgery, Huashan Hospital, Fudan University, Shanghai 200040, PR China
[k] Shanghai Clinical Medical Center of Neurosurgery, Shanghai 200040, PR China
[l] Shanghai Key Laboratory of Brain Function Restoration and Neural Regeneration, Shanghai 200040, PR China
[m] Department of Neurosurgery, Shanghai Gamma Hospital, Shanghai 200235, PR China

ARTICLE INFO

Keywords:
Diffuse large B cell lymphoma
Central nervous system
Single-cell RNA sequencing
Genomics
Tumor Microenvironment

ABSTRACT

Primary central nervous system diffused large B-cell lymphoma (PCNS-DLBCL) is a rare type of non-Hodgkin lymphoma restricted to the central nervous system (CNS). To explore its specific pathogenesis and therapeutic targets, we performed multi-omics sequencing on tumor samples from patients diagnosed with PCNS-DLBCL, secondary CNS-DLBCL or extracranial (ec) DLBCL.By single-cell RNA sequencing, highly proliferated and dark zone (DZ)-related B cell subclusters, MKI67_B1, PTTG1_B2 and BTG1_B3, were predominant significantly in PCNS-DLBCL. Compared to SCNS-DLBCL and ecDLBCL, an immune-suppressive tumor microenvironment was observed in PCNS-DLBCL by analysis of immune-stimulating/inhibitory ligand–receptor (L-R) pairs. By performing whole-exome sequencing in 93 patients, mutations enriched in BCR-NFkB and TLR pathways and the cooperation of these two pathways were found to be predominant in PCNS-DLBCL comparing to nonGCB-ecDLBCL. In summary, our study provides comprehensive insights into the transcriptomic and genetic characteristics of PCNS-DLBCL in contrast to ecDLBCL and will help dissect the oncogenic mechanism of this disease.

Introduction

Primary central nervous system (CNS) lymphoma (PCNSL) is an aggressive extranodal non-Hodgkin lymphoma involving the brain, spinal cord, cranial nerves, eye and leptomeninges, most of which are pathologically classified as diffuse large B-cell lymphoma (DLBCL) of

* Corresponding author: Department of Hematology, Huashan Hospital, Fudan University. 12 Wulumuqi Middle Road, Shanghai 200040, PR China.
** Corresponding author: Department of Neurosurgery, Huashan Hospital, Fudan University. 12 Wulumuqi Middle Road, Shanghai 200040, PR China.
*** Corresponding author: Department of Neurosurgery, Shanghai Gamma Hospital, 518 Wuzhong East Road, Shanghai 200235, PR China.
E-mail addresses: wolf2yy@qq.com (H. Wu), maoying@fudan.edu.cn (Y. Mao), chentong@fudan.edu.cn (T. Chen).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.neo.2024.101119
Received 11 June 2024; Accepted 19 December 2024

immune-privileged sites [1–3]. The majority PCNS-DLBCL cases are of the activated B-cell-like (ABC) subtype, whose prognosis is far worse than that of extracranial DLBCL (ecDLBCL) due to the blockage of drug delivery, molecular diversity of tumor cells and immune privileging status of the tumor microenvironment (TME) [4,5].

Based on normal B cell development process, the deconvolution of malignant components in DLBCL cell-of-origin (COO) classification was widely applied in clinical practice, including germinal center B-like (GCB) or activated B-cell-like (ABC) DLBCL subtypes with distinct survival outcomes [6,7]. Furthermore, phenotype involving centroblast and centrocyte subtypes was also proposed by gene expression profiling [8]. The developmental trajectory model of DLBCL B cell states was expanded by single-cell RNA sequencing technology (scRNA-seq) [9], and further prognostic sc-COO classification in DLBCL corresponding to detailed B cell states was also proposed [10]. Comparing to in-depth analysis of malignant B cells in DLBCL, the information of PCNSL cell states was relative absent. Except PCNSL was thought to be more familiar to ABC subtype but insufficient COO classification in predicting survival [11], predominant constitution of activated CD27+CD38+ *B* cells and potentially inferior survival indicator plasmablast-like MP2 cluster were uncovered [12–14].

Recently, the immunosuppressive TME has been reported to be critical for oncogenesis and metastasis of various neoplasms [15,16]. To discover clinically relevant cell states and TME ecosystems in ecDLBCL, EcoTyper was implemented by integrating transcriptome deconvolution and scRNA sequencing. The results verified that TME-comprising cellular components have distinct cell-state interactions beyond COO classification [9]. For PCNS-DLBCL, another prognostically significant cluster model, RBraLymP (RNA-based Brain Lymphoma Profiler) was developed. It identified four subtypes with their own oncogenic pathways, TMEs, clinical outcomes and potential therapeutic targets [4]. The difference in TME composition between PCNS-DLBCL and systematic ecDLBCL has not been clarified.

Given the heterogeneous outcomes of DLBCL cases, a variety of classifications have been applied to facilitate therapeutic decisions. The classification of malignant COO has been conventionally applied for pathological diagnosis in clinical practice for decades [6,7]. Recently, a few genetic classifiers were developed based on genetic events and transcriptome characteristics. For example, LymphGen algorithms seeding six genetic classes (MCD, BN2, N1, EZB, A53, ST2 subtypes) or the molecular division into subtypes C1-C5 have been established with distinct molecular characteristics and prognoses [9,17–19]. Although mutational pattern of PCNSL had been described in distinct sizes of cohorts, the targeted interpretation of differences between inter- and extra- cranial DLBCL had not been performed before [4,20,21].

In this study, we compared the transcriptome and TME stereotypes of tumor B-cells from PCNS-DLBCL, secondary CNS (SCNS)-DLBCL and ecDLBCL, respectively. A mutational profile in a large PCNSL cohort was also proposed accompanied by a comparison to nonGCB-ecDLBCL cohort. Our results provided comprehensive insights into the transcriptomic and genomic features of PCNS-DLBCL and will help to dissect the oncogenic mechanism for this rare disease entity.

Results

Study outline

To uncover the difference in transcriptome profiles between intracranial DLBCL (icDLBCL) and ecDLBCL, scRNA-seq was performed on a cohort of six PCNS-DLBCL, two SCNS-DLBCL, and five ecDLBCL (4 from lymph nodal and 1 from tumor tissue) samples, all of which were Epstein–Barr virus (EBV)-negative by EBV-encoded RNA (EBER) analysis. Ten were from newly diagnosed patients (6 PCNS-DLBCL, 1 SCNS-DLBCL and 3 ecDLBCL), while 3 samples were from relapsed patients at least half a year after discontinuing treatment (1 SCNS-DLBCL and 2 ecDLBCL). Detailed information on the patients with scRNA-seq is listed in Supplementary Table 1. The mutation landscape of PCNS-DLBCL was described by whole-exome sequencing (WES) in a cohort of 93 cases with fresh or formalin-fixed paraffin-embedded (FFPE) samples. The schematic procedure of this study is shown in Fig. 1A.

Classification of B-cell subclusters by scRNA-seq

Altogether, 134,710 cells were obtained from 13 samples after a strict filtering process (described in Methods). The samples subjected to scRNA-seq were made up of $CD19^+$ *B* cells (80.32 %−88.29 %), $CD3^+$ *T* cells (3.73 %−10.77 %), $CD14^+$ myeloid cells (5.17 %−7.39 %) and few gliocytes/endothelial cells (1.23 %−3.74 %) (Supplementary Fig. 1A-B) [10,22].

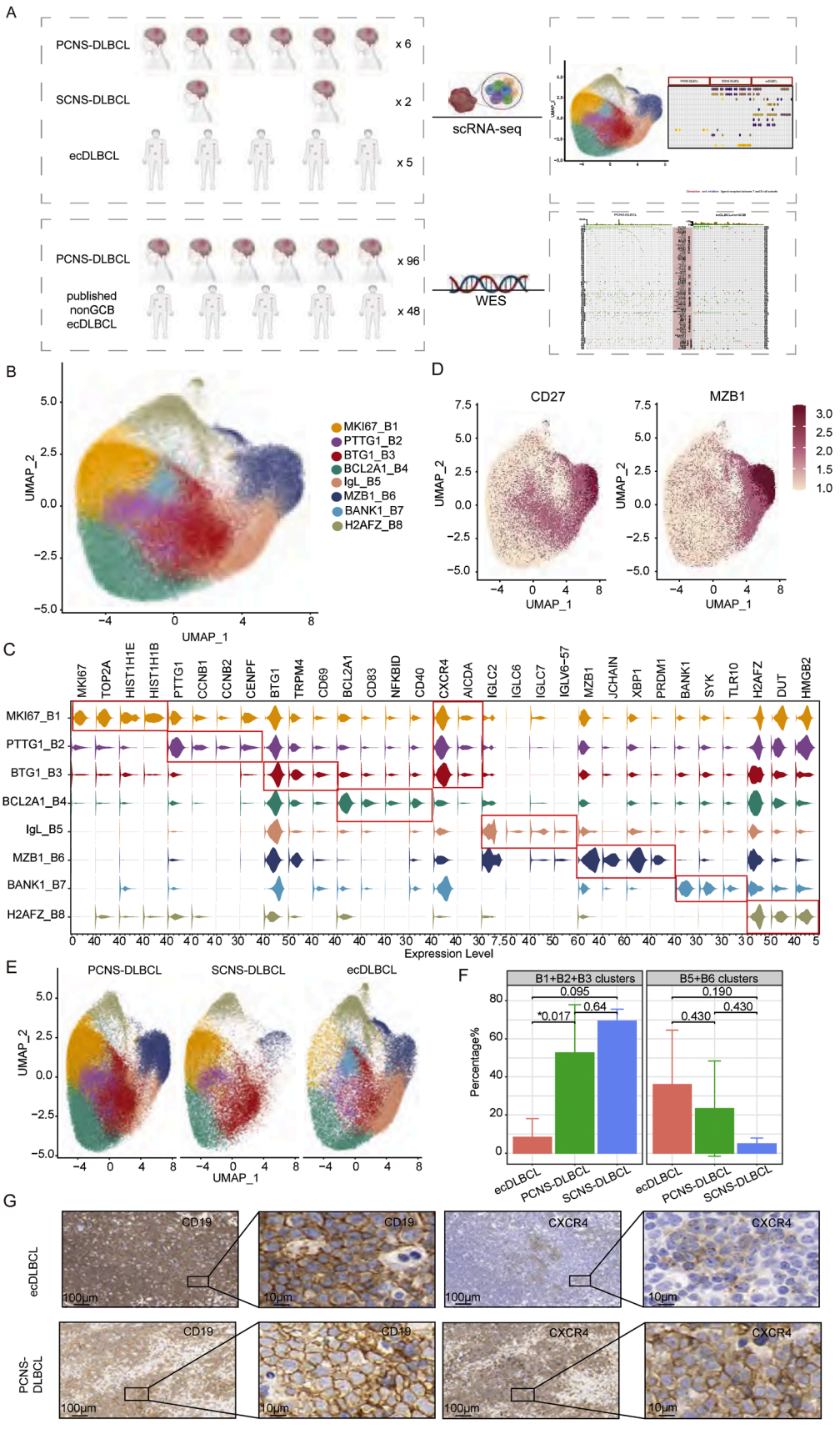
B cell subgroups were divided based on specifically expressed markers combined with canonical makers of B cell development (Fig. 1B-C, Supplementary Table 2-3). The MKI67_B1 subtype was characterized by the expression of proliferative markers (MKI67 and TOP2A) [23], DNA replication contributors histone H1 clusters (HIST1H1E, HIST1H1B) [24] and scattered immunoglobulin heavy Mu (IGHM) expression (Supplementary Fig. 1C & Supplementary Table 4). Similarly, the PTTG1_B2 subtype, representing an activated cell division pattern, was distinguished by the expression of genes linked to G2/M chromatid separation and mitosis, including PTTG1, CCNB1, CCNB2 and CENPF [25]. BTG1_B3 subcluster expressed higher levels of B-cell developing regulators, BTG1 and CD69, with a mixture of IGHM and IGHA1 expression. BCL2A1_B4 subcluster specifically expressed signatures of GC light zone B cells (LZB), BCL2A1 and CD83, with mixed expression of IGHM, IGHG1 and IGHA1[10]. Additionally, CD40-stimulation markers (NFKBID, CD40) were also identified in this subgroup [26]. Furthermore, the markers of dark zone (DZ) B cells in GC, CXCR4 and AICDA were mainly expressed in MKI67_B1, PTTG1_B2 and BTG1_B3, which represented the signature of proliferation and somatic hypermutation (SHM) [10,27].

Among $CD27^+$ post-GC memory B-cell subclusters, the IgL_B5 subcluster was determined by variable immunoglobulin light chains (IGLC2) and immunoglobulin heavy chains (IGHG1, IGHM), representing a relative mature B-cell population experiencing antigen stimulation and preparing for antibody production (Fig. 1D). In the MZB1_B6 subcluster, canonical plasma markers (MZB1, JCHAIN, XBP1, and PRDM1) were identified with IGHM, IGHA1 and $IGHG1^{dim}$. Activation of the BCR and TLR pathways was demonstrated in BANK1_B7 subcluster cells by the expression of BCR members (BANK1 and SYK) and Toll-like receptor 10 (TLR10). The H2AFZ_B8 subcluster exhibited DNA replication markers (H2AFZ, DUT, HMGB2) [28].

We then compared the distribution of different subpopulations between three DLBCL entities (Fig. 1E). Interestingly, the DZ-related clusters MKI67_B1, PTTG1_B2 and BTG1_B3 accumulated more dominantly in PCNS-DLBCL (median 59.09 %, range 12.05 %−78.81 %) than in ecDLBCL (median 7.92 %, range 1.45 %−20.87 %) ($P = 0.017$) (Fig. 1F). This distinct distribution of DZ clusters between different entities was also confirmed by the higher expression of the DZ marker CXCR4 in PCNS-DLBCL samples by immunohistochemistry (IHC) (Fig. 1G). In contrast, no significant difference was found in distribution of relative mature B cell clusters (IgL_B5 and MZB1_B6 subclusters) between PCNS-DLBCL (median 13.48 %, range 4.08 %−65.78 %) and ecDLBCL (median 43.08 %, range 7.36 %−64.49 %).

Diversity of B-cell subclusters in icDLBCL and ecDLBCL

To further investigate the heterogeneity of B-cell distribution and malignant pathogenesis, we performed WES analysis to detect the mutational profiles of different DLBCL entities. In PCNS-DLBCL samples, 4/6 patients (P3-P6) exhibited a predominant DZ B cells (DZB) population, while 1/6 patients (P1) predominantly displayed mature IgL_B5 and MZB1_B6 clusters, and 1/6 patients (P2) had identical percentages of the DZB and mature subclusters (Fig. 2A). Of the patients with



(caption on next page)

Fig. 1. Schematic overview and identification of B-cell subsets in icDLBCL and ecDLBCL
(A) Schematic outline of the study design.
(B) UMAP plot of scRNA-seq profiled B cells from 13 samples. Each cluster is indicated by distinct colors.
(C) Violin plot of selected signatures in B-cell clusters. The expression level of each marker gene is indicated by the size of the violin and is labeled at the bottom. Markers in the red box represent specific gene subsets.
(D) UMAP plot of the marker genes for different B-cell lineages.
(E) UMAP plot of B cell subsets in each type of DLBCL
(F) Bar plot represents comparison of DZ related B cell subsets (B1+B2+B3) and CD27+ *B* cell subsets (B5+B6) distribution in each type of DLBCL, respectively. *P* value is labeled on top of bars with asterisk indicates *P* value<0.05.
(G) IHC staining of CD19 and CXCR4 in PCNS-DLBCL and ecDLBCL FFPE samples (Scale bar=100 μm and 10 μm, respectively).

available mutational profiles, 4/5 patients (P1, P2, P4, P6) had mutations of proximal BCR (*PLCG2* and *BLK*) or the NF-kB signaling pathway (*CARD11*). Among them, 3/3 DZB-dominant cases (P4-P6) carried mutations of $MYD88^{L265P}$, $CD79B^{Y196}$, *PIM1* and histone H1, such as *HIST1H1E (H1E)* and *HIST1H1B (H1B)*. In P1 with fewer DZB and dominant mature B-clusters, wild-type *MYD88, CD79B, PIM1*, BCR members and mutated *CARD11* were identified (Fig. 2B). A similar grouping pattern was seen in two SCNS-DLBCL patients with major DZB subclusters and minor mature B-cell clusters at >60 % and <10 %, respectively. Instead of *MYD88* and *CD79B*, mutations in NF-kB pathway members (*TNFAIP3, CARD11, BCL10*) were detected in SCNS-DLBCL. In ecDLBCL samples, 3/5 patients (D1-D3) had prevalent mature IgL_B5 and MZB1_B6 subclusters (43.07 %−64.69 %), whereas the other two patients exhibited dominant BANK1_B7 (D4) or H2AFZ_B8 (D5), at 33.29 % and 58.32 %, respectively. Mutations in *PIM1*, the *NF-kB* pathway or H1 clusters were identified in 4/4 cases whose WES information was available. The gene expression profiling (GEP) of these 13 samples is shown in Supplementary Fig. 2 [6].

The pathways that were enriched in PCNS-DLBCL, SCNS-DLBCL and ecDLBCL were further compared. In PCNS-DLBCL, the SHM-related pathway, antigen recognition and B-cell activation signaling were upregulated in DZ-related B subclusters and IgL_B5 subclusters (Supplementary Fig. 3, Supplementary Table 5 & 6). Through pseudotime analysis performed by Monocle3, the evolutionary trajectory from MKI67_B1, PTTG1_B2 and BTG1_B3 to BCL2A1_B4 was identified in the cohorts with PCNS-DLBCL (Supplementary Fig. 4A-D) and SCNS-DLBCL (Supplementary Fig. 4E-F). No evolutionary tendency relating to IgL_B5 or MZB1_B6 was observed in icDLBCL, while a trajectory from BCL2A1_B4 to IgL_B5 was observed in ecDLBCL (Supplementary Fig. 4G-H).

Copy number variants of B-cell clusters

Specific copy number variants (CNVs) were detected in all B-cell clusters in PCNS-DLBCL, SCNS-DLBCL and ecDLBCL, indicating that all B cells in these samples were malignant per se (Fig. 2B-D). CNVs of chromosome 6 were validated by WES in 26 tumor-blood paired samples (Supplementary Fig. 5). Copy loss of chromosome 6 (6p21.32-6p22.1) involving surveillance molecular-MHC I/II was observed in each PCNS-DLBCL sample. Specifically, in P2 and P4, there was large-scale copy loss of 6q15-6q27 involving B-cell maturation (PRDM1) and NF-KB signaling genes (*IFNGR1, TNFAIP3*) (Fig. 2B). On chromosome 12, large-scale copy gains of 12q13.13-12q24.33 involving cell cycling-related genes (CDKN2, CDKN4, TUBA1A, BTG1) and transcription factors (STAT2, STAT6) were identified in 4/6 PCNS-DLBCL samples. All of the above CNV events on chromosomes 6 and 12 were shared by all subclusters. However, copy gains on chromosome 22 (22q11.1-22q12.1) involving *IGLL5* were only observed in the MZB1_B6 cluster of the P1 and P2 . CNVs of chromosomes 6, 12 and 22 were also observed in SCNS-DLBCL cases. Interestingly, gains of chromosome 9 with immune escape genes CD274 (*PD-L1*) were specifically found in both SCNS-DLBCL cases (Fig. 2C).

The CNVs in the ecDLBCL cohort were more heterogeneously. Similar to published literature [29], genes involved in cell cycling (*CDK11A* on chromosome 1, *CDKN1A* on chromosome 6, and *CDK6* and *CDK14* on chromosome 7), the NF-kB pathway (*IRF4* and *MAP3K7* on chromosome 6) and epigenomic modifiers (*EZH2* and *KMT2C* on chromosome 7) were found recurrently in ecDLBCL. Two cases (D3 and D4) carried relatively simple CNVs involving mainly two chromosomes, while the other three involved >5 chromosomes (D1, D2, D5) (Fig. 2D).

T-cell subclusters in icDLBCL and ecDLBCL

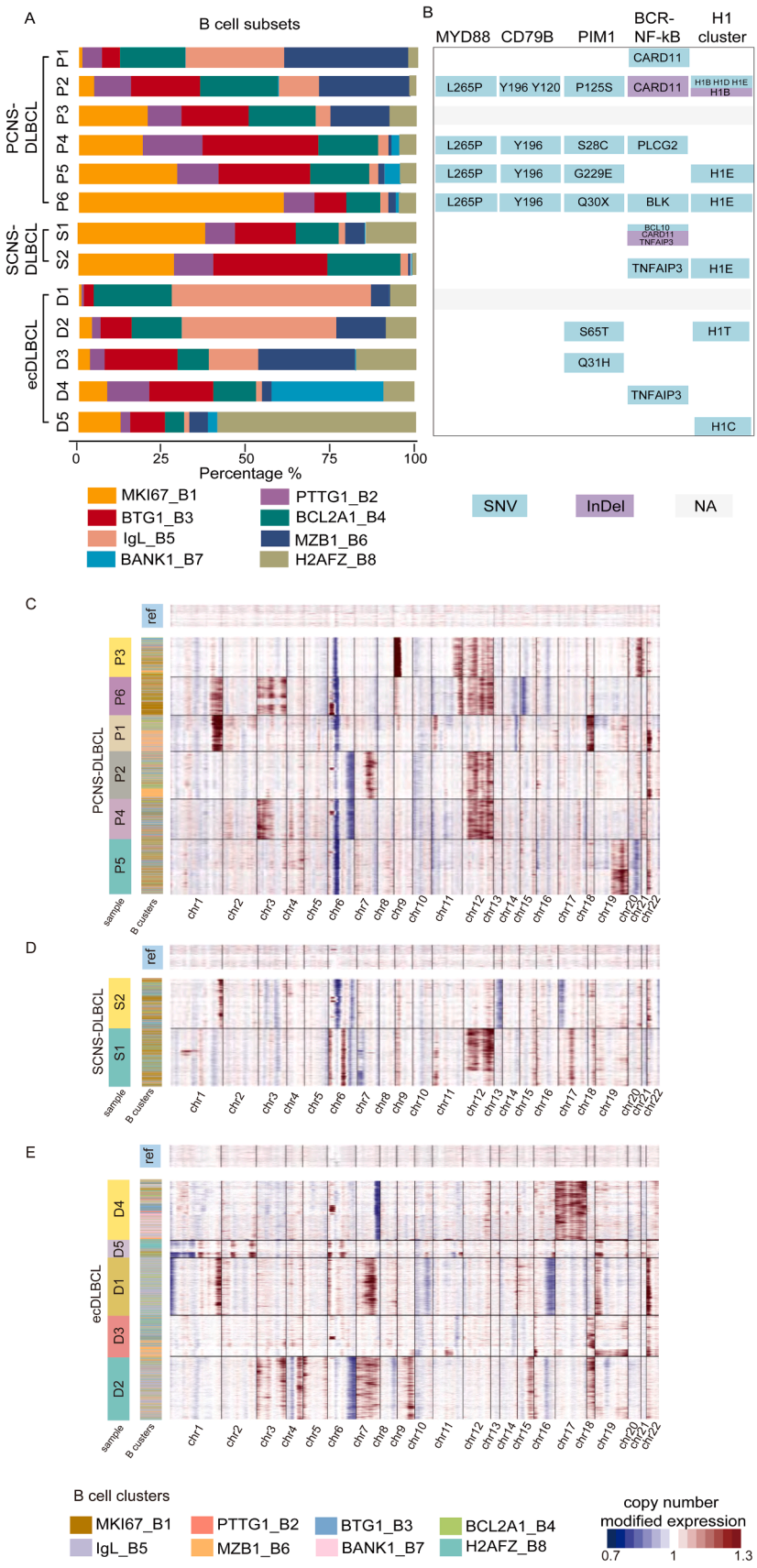
We further analyzed tumor-infiltrating T lymphocytes to characterize the TME in DLBCLs. They were collectively divided into CD8_T lymphocytes, CD4_T lymphocytes and NKT cells by their expression of CD8A, CD8B, CD4 and FCGR3A, respectively (Fig. 3A-B). Specifically, all CD8_T cells expressed effective cytotoxic molecules (GZMB, GZMH, PRF1, EOMES) and exhaustion markers (HAVCR2, LAG3, TIGIT, PDCD1) (Fig. 3C). The CD8_T1 subcluster was associated with a common effector T-cell signature (GZMA, PLEKHF1), and the CD8_T2 subcluster was associated with antigen-responsive markers (CRTAM, TNFRSF9) and costimulatory molecules (TNFRSF18, TNFRSF4). As a subpopulation with proliferative capability, the CD8_T3 subpopulation was characterized by MKI67 expression. CD4_T lymphocytes and NKT subclusters were identified with IL7R, CCR7 and nature killer signatures (KLRD1, GNLY) (Fig. 3C, Supplementary Table 4 & 7).

Totally, CD8 T cells tended to be more predominant in PCNS-DLBCL (median 84.53 %, range 61.54 %−96.29 %) than in ecDLBCL (median 65.59 %, range 58.12 %−69.69 %) but without significance ($P = 0.052$) (Fig. 3D-E). The median percentage of CD4 T cells was 25.48 % in ecDLBCL (range 16.71 %−33.18 %) and 10.63 % in PCNS-DLBCL (range 0 %−34.62 %), respectively ($P = 0.082$).

Myeloid cell and macrophage subclusters in icDLBCL and ecDLBCL

Myeloid-derived cells serve as another key component of TME-comprising immune cells in tumor progression [30]. In the CNS, microglia are CNS-residing macrophages with distinct myeloid features [31]. We grouped myeloid clusters by their CD14 expression, which were subdivided into $CD68^{+}$ macrophages, $CD68^{-}FCGR3A^{+}$monocytes and $TREM2^{+}$ microglia. The $HLA\text{-}DPB1^{+}$ population was classified as dendritic cells at a relatively lower frequency (Fig. 4A-B). In addition, macrophage clusters were subgrouped into an immunosuppressive M2-like Macro1 subcluster expressing CD163, an alveolar macrophage-like Macro2 subcluster expressing UQCRH and HMGB1, and a B-cell attracting Marco3 subcluster expressing CXCL13 (Fig. 4C, Supplementary Table 4 & 8). Additionally, by comparing different entities of DLBCL, the Macro2 subcluster was mostly enriched in ecDLBCL (median 8.93 %, range 6.52 %−21 %) than in PCNS-DLBCL (median 44.49 %, range 10.96 %−74 %) ($P = 0.017$) (Fig. 4D-E). A peripheral monocyte-like subcluster expressing chemotactic factor CXCL8 was identified.

All microglia expressed disease-associated microglia (DAM) markers (GPNMB, APOE), inflammatory chemokines (CCL3, CCL4) and an M2-like signature (CD163), indicating an inflammatory ecosystem in icDLBCL (Fig. 4C) [31]. Two microglial subclusters, the Micro1 subpopulation with a protumor signature (CCL18) and Micro2 with a neurodegenerative signature (SPP1, C3), were identified [32,33]. According to the canonical DC classification, dendritic cells were identified



(caption on next page)

Fig. 2. Identification of B-cell subsets in different types of DLBCL and Copy number variants in different types of DLBCL
(A) The bar plot indicates the percentage of B-cell subsets in every sample from all 13 patients. Each B-cell cluster is represented by a distinct color and corresponding legend.
(B) The diamond shows the mutational state of *MYD88, CD79B, PIM1*, BCR-NFkB pathway genes and histone H1 clusters in individual samples. SNVs and InDels are indicated in blue and purple, respectively. The sample without available data is indicated in gray. H1B: *HIST1H1B*; H1C: *HIST1H1C*; H1D: *HIST1H1D*; H1E: *HIST1H1E*; H1T: *HIST1H1T*.
(C-E) Analysis of CNVs in PCNS-DLBCL (C), SCNS-DLBCL (D) and ecDBLCL (E) by scRNA-seq with the labeled chromosome region at the bottom. Copy gains are colored red, and copy loss is colored blue. Infiltrated T lymphocytes and myeloid cells in each type of DLBCL were analyzed as references (ref) on the top. Each B-cell cluster is indicated by a distinct color, and the case barcode is labeled on the left.

as DC1 (CLEC9A, XCR1) and DC2 (CD1C) subclusters across intracranial and extracranial tumors [34].

Cellular interaction in the tumor microenvironment

Communication between infiltrated immune cells and B-cell subsets was analyzed using CellphoneDB. Among the three entities, there were fewer cellular interactions in PCNS-DLBCL than in SCNS-DLBCL or ecDLBCL (PCNS-DLBCL vs. SCNS-DLBCL vs. ecDLBCL: 23.48 % vs. 39.78 % vs. 36.73 %). Not surprisingly, the ligand−receptor (L-R) crosstalk between T-B cells in PCNS-DLBCL was the lowest (PCNS-DLBCL vs. SCNS-DLBCL vs. ecDLBCL: 18.78 % vs. 34.97 % vs. 46.24 %), while that between myeloid-B cells (M-B) in SCNS-DLBCL was the most activated (PCNS-DLBCL vs. SCNS-DLBCL vs. ecDLBCL: 28.96 % vs. 40.63 % vs. 30.41 %, Supplementary Fig. 6 A-C).

Stimulating interactions of T-B and M-B cells were dramatically silenced in PCNS-DLBCL (Fig. 5). For example, several T-B-related antitumor l-R pairs, CD40LG-CD40, ICOS-ICOSLG, IL7 receptor-IL7 and CD28-CD80/CD86, were only detected in ecDLBCL (Fig. 5A). As for M-B crosstalk, proinflammatory l-R pairs, including TNFRSF1A-LTA, MRC1-PTPRC and TNFRSF1A-TNF, were found only in SCNS-DLBCL and ecDLBCL (Fig. 5B).

The immune inhibitory l-R counterparts were expressed differently in different DLBCLs. Some immune checkpoint or T-cell-inhibiting l-R pairs, PDCD1-CD274, HLA-E-CD94:NKG2A and KLRC1-HLA-E, were only identified in SCNS-DLBCL (Fig. 5A). Some myeloid-inhibitory pairs, APP-CD74, HBEGF-CD44 and C5AR1-RPS19, were observed in the non-PCNS-DBLCL cases (Fig. 5B). Interestingly, the IL10 receptor–IL10 interaction of either T-B or M-B was exclusively observed in icDLBCL (Fig. 5A-B).

CNS-specific l-R expression in T-myeloid (T-M) populations was also identified (Supplementary Fig. 6D). For example, FASLG-FAS, CD6-ALCAM, LTBR-LTB and SPN-SIGLEC1 were expressed in SCNS-DLBCL and ecDLBCL. Chemokine l-R pairs, CXCR3-CXCL10, CCR1-CCL18, CCR1-CCL8 and CXCR3-CCL20, were mostly detected in SCNS-DLBCL (Supplementary Fig. 6D). Interestingly, the exclusive expression of SPP1-CD44 in icDLBCL cells indicated a potential role in the oncogenesis of CNS lymphoma (Fig. 5B & Supplementary Fig. 6D).

Mutational pattern of PCNS-DLBCL compared to nonGCB-ecDLBCL

To construct a mutational profile of PCNS-DLBCL, we performed WES on a large cohort of newly diagnosed patients, including 31 freshly collected tumors with 26 paired PB samples and 62 FFPE tumor specimens without matched blood. The clinical characteristics of the included patients are listed in Supplementary Table 9.

After a strict filtering workflow, a total of 25,126 mutations of 10,642 genes in exons or splice sites were detected, of which somatic nucleotide variants (SNVs) and insertions and deletions (InDels) were 88.29 % and 11.71 %, respectively. The mutation types of each sample are presented in Supplementary Fig. 7A, where missense mutations were the absolute leading prevalence at 80.69 %. The dominant base substitution event in the somatic mutation spectrum was C > T (62.76 %), followed by C > A (20.23 %) and T > C (17.01 %) (Supplementary Fig. 7B). The median mutation number was 234/case (53-2548), and the median nonsynonymous tumor mutation burden (TMB) in FFPE, fresh tumor and fresh-paired samples was 5.04/Mb, 4.69/Mb and 2.50/Mb, respectively.

Nineteen mutated genes were detected recurrently in ≥20 % of patients. Among them, mutations of *MYD88* (77.42 %), P*IM1* (53.76 %), *CD79B* (45.16 %), *KMT2D* (44.09 %), *IGLL5* (36.56 %), *OSBPL10* (32.26 %), *HIST1H1E* (32.26 %), *KLHL14* (32.26 %) and *SETD1B* (30.11 %) had predominant frequencies >30 % (Fig. 6A-B). The details of the mutated loci of these genes are shown in Supplementary Fig. 7C. Specifically, 87.50 % (63/72) of *MYD88*-mutated loci were L265P, and two cases were observed as multiple hits, one of which had coexisting mutations of L265P and the Toll/interleukin-1 receptor (TIR) domain . Multiple *PIM1* mutations were observed in 46 % of the cases, suggesting an aberrant somatic hypermutation (aSHM) target in PCNS-DLBCL. Among *CD79B* mutation carriers, 78.57 % had variants residing in the immune receptor tyrosine-based activation motif (ITAM) domain at Y196 or E197. Notably, the ibrutinib-resistant variant C481 was not detected in BTK mutations, indicating no common resistance to BTK inhibitors in our newly diagnosed PCNS-DLBCL cohort [35]. Gene mutations with frequencies >10 % are listed in Supplementary Table 10.

To simplify the gene mutation profile for translation to the clinic, we combined the genes related to DLBCL lymphomagenesis and cellular survival as different functional gene sets, including epigenetic modifiers, B-cell development, BCR-NFkB pathway, *MYD88*-included TLR pathway, tumor suppressors & immunity, *JAK/STAT3* pathway, *Notch* signaling pathway and *PI3K* signaling pathway in the sequential order of 87.09 %−21.51 % (Fig. 6C, Supplementary Table 12) [18,36,37]. In epigenetic modification sets, recurrent mutations of the H1 cluster (e.g., *HIST1H1E, HIST1H1B, HIST1H1C*) were identified in 74.19 % of PCNSL cases (Supplementary Fig. 7D), showing the importance of chromatin instability in this disease. In the BCR signaling pathway, mutated *CD79B* was more prevalent than other members. Dysregulation of the signaling pathway involving the BCR-NFkB/TLR and BCR-NFkB/PI3K genes was found in about 68.81 % and 19.35 % of cases, respectively, illustrating a genetic cooperative mode during PCNS-DLBCL oncogenesis (Fig. 6D).

For comparison, we also explored the mutational characteristics in a 48 cases of EBV negtive nonGCB-ecDLBCL cohort from previous research [19], based on previous finding that genomic pattern of PCNS-DLBCL was close to nonGCB-DLBCL subtype [11,17,36]. The frequencies of top mutated genes in PCNS-DLBCL such as *MYD88, CD79B* and *KMT2D*, were lower in nonGCB-ecDLBCL at 33.33 %, 25 % and 14.58 % respectively (Fig. 6A-B, Supplementary Table 11), with less involving TLR pathway (35.42 %) and BCR-NF-kB pathway (47.92 %) (Fig. 6C, Supplementary Table 12). Interactions between pathways of BCR-NF-kB and TLR or BCR-NF-kB and PI3K were also decreased in nonGCB-ecDLBCL than in PCNS-DLBCL (Fig. 6D). However, mutational percentage of functional gene sets including epigenomic modifier, B cell development and tumor suppression & immunity were similar in both entities. All the above comparison indicated a specific TLR and BCR-NFkB pathway-depending genetic landscape of PCNS-DLBCL, which was quite different from nonGCB-ecDLBCL.

Discussion

The biological heterogeneity of DLBCL has been indicated by dissecting the tumorous COO, TME and associated genomic diversities in depth [6–10]. As a subtype exclusively involving the central nervous

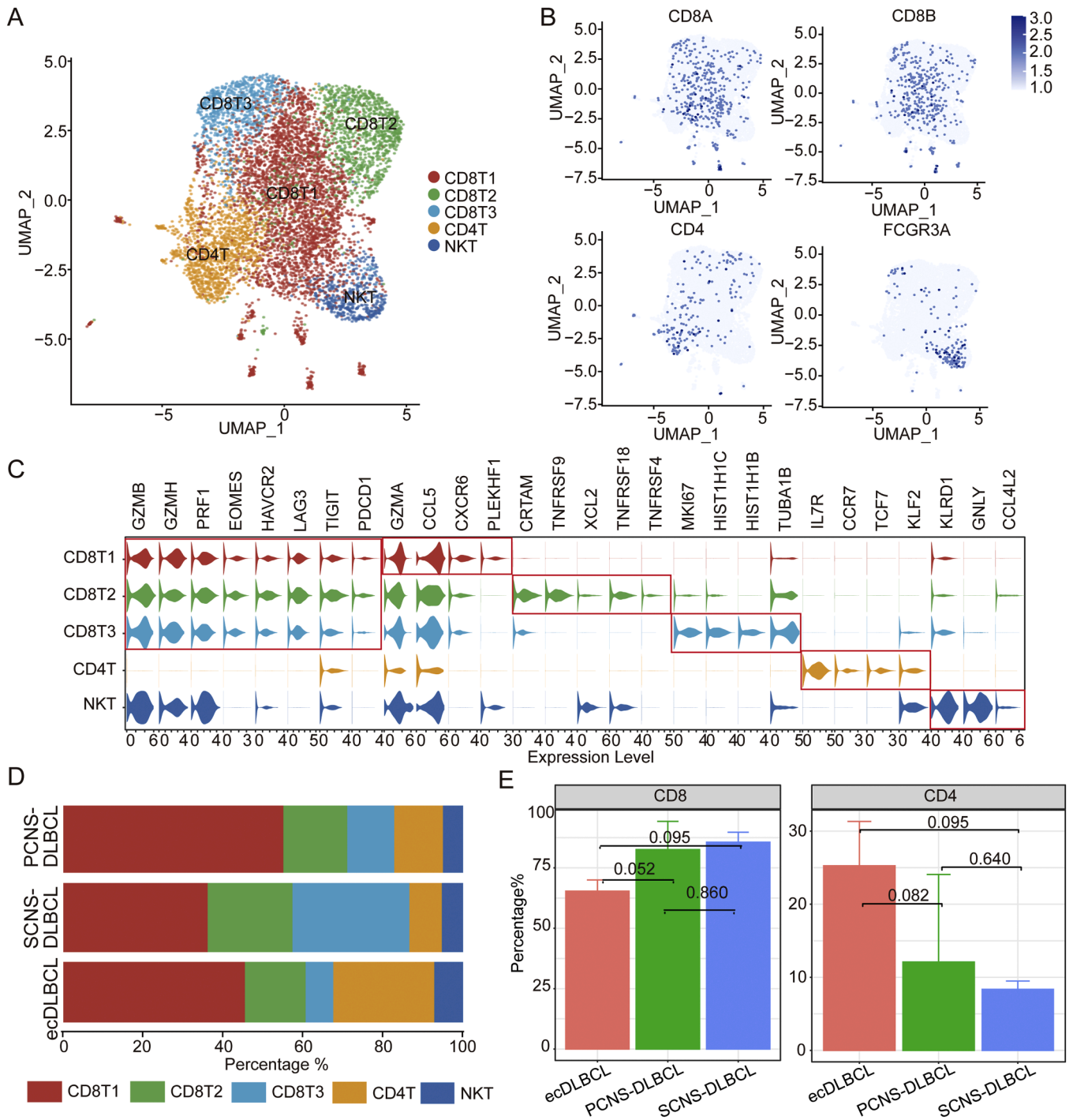


Fig. 3. Identification of T-cell subsets in different types of DLBCL

(A) UMAP plot of scRNA-seq profiled T-cell subsets from all 13 samples. Each cluster is indicated by a distinct color.

(B) UMAP plot of marker genes for different T-cell lineages.

(C) Violin plot of selected signatures in T-cell clusters. The expression level of each marker gene is indicated by the size of the violin and is labeled at the bottom. Markers in the red box represent specific gene subsets.

(D) Bar plot represents comparison of CD8+ and CD4+ *T* cell subsets distribution in each type of DLBCL, respectively. *P* value is labeled on top of bars with asterisk indicates *P* value<0.05.

(E) Bar plot represens the percentage of different T-cell populations in each type of DLBCL. *P* value is labeled on top of bars with asterisk indicates *P* value<0.05.

system, PCNS-DLBCL was recently described in terms of its oncogenic pathways, gene expression phenotypes, methylation profiles, TME and outcomes by multiomic analysis [4,14,38]. However, the difference between icDLBCL and ecDLBCL has not been fully explored. In this study, we compared the transcriptome, ecosystem and mutational profile of DLBCL inside and outside the CNS, providing a comprehensive perspective for further exploring pathogenesis of PCNS-DLBCL.

Most PCNS-DLBCL cases are roughly staged as GC-experiencing or 'time-slot' overlaps of late GC and early post-GC by bulk-RNA seq [11, 39]. In this study, we deconvoluted the diversity of PCNS-DLBCL cells by scRNA-seq and demonstrated that PCNS-DLBCL tumors were composed of a mixture of B-cell subsets, providing an explanation for the

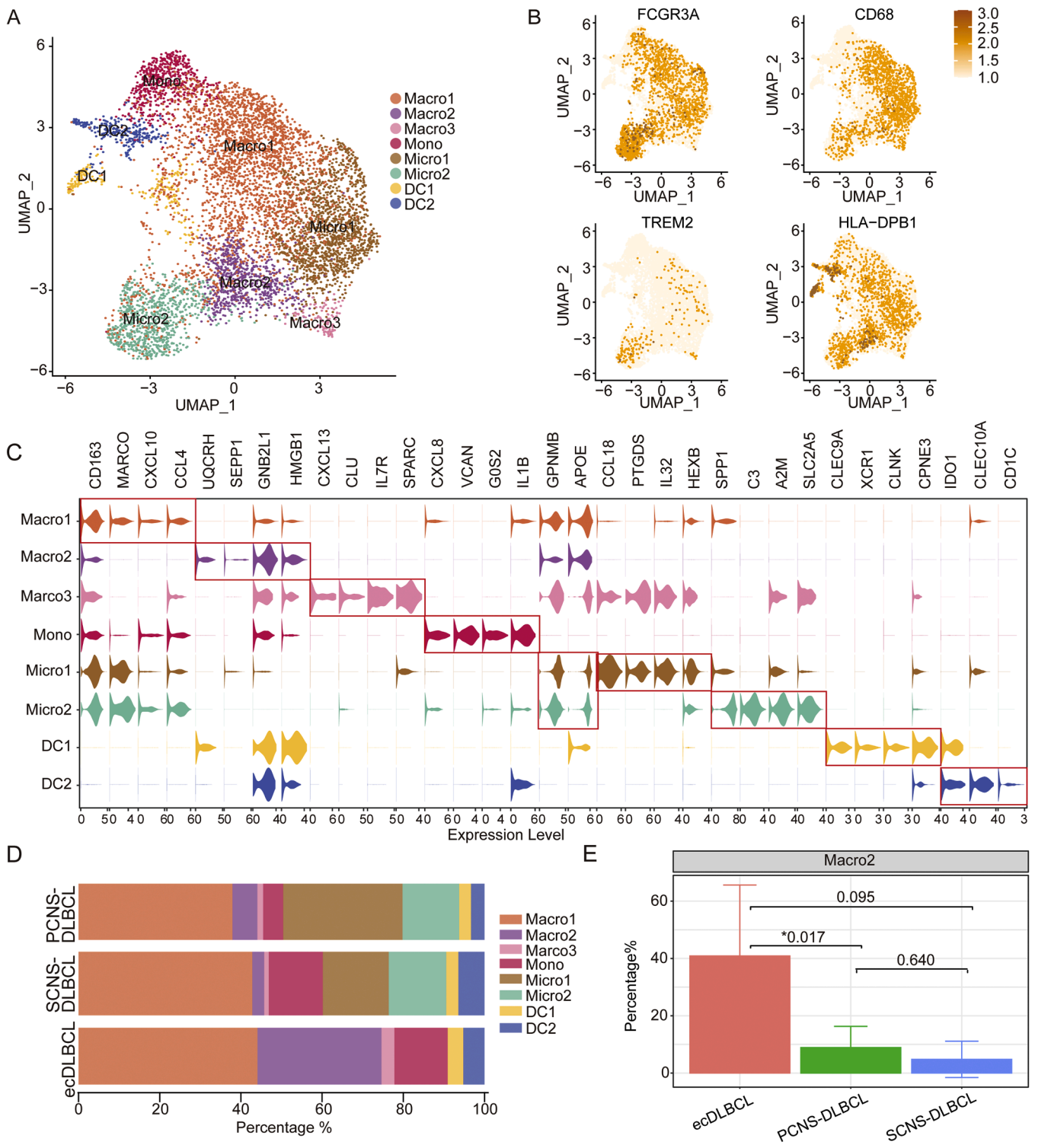


Fig. 4. Identification of myeloid subsets in different types of DLBCL

(A) UMAP plot of sc-RNAseq profiled myeloid cells from all 13 samples. Each cluster is indicated by a distinct color.

(B) UMAP plot of marker genes for different myeloid lineages.

(C) Violin plot of selected signatures in myeloid clusters. The expression level of each marker gene is indicated by the size of the violin and labeled at the bottom. Markers in the red box represent specific gene subsets.

(D) Percentage of total infiltrated myeloid cells in three types of DLBCL.

(E) Bar plot represents comparison of Marco2 cell subset distribution in each type of DLBCL, respectively. *P* value is labeled on top of bars with asterisk indicates *P* value<0.05.

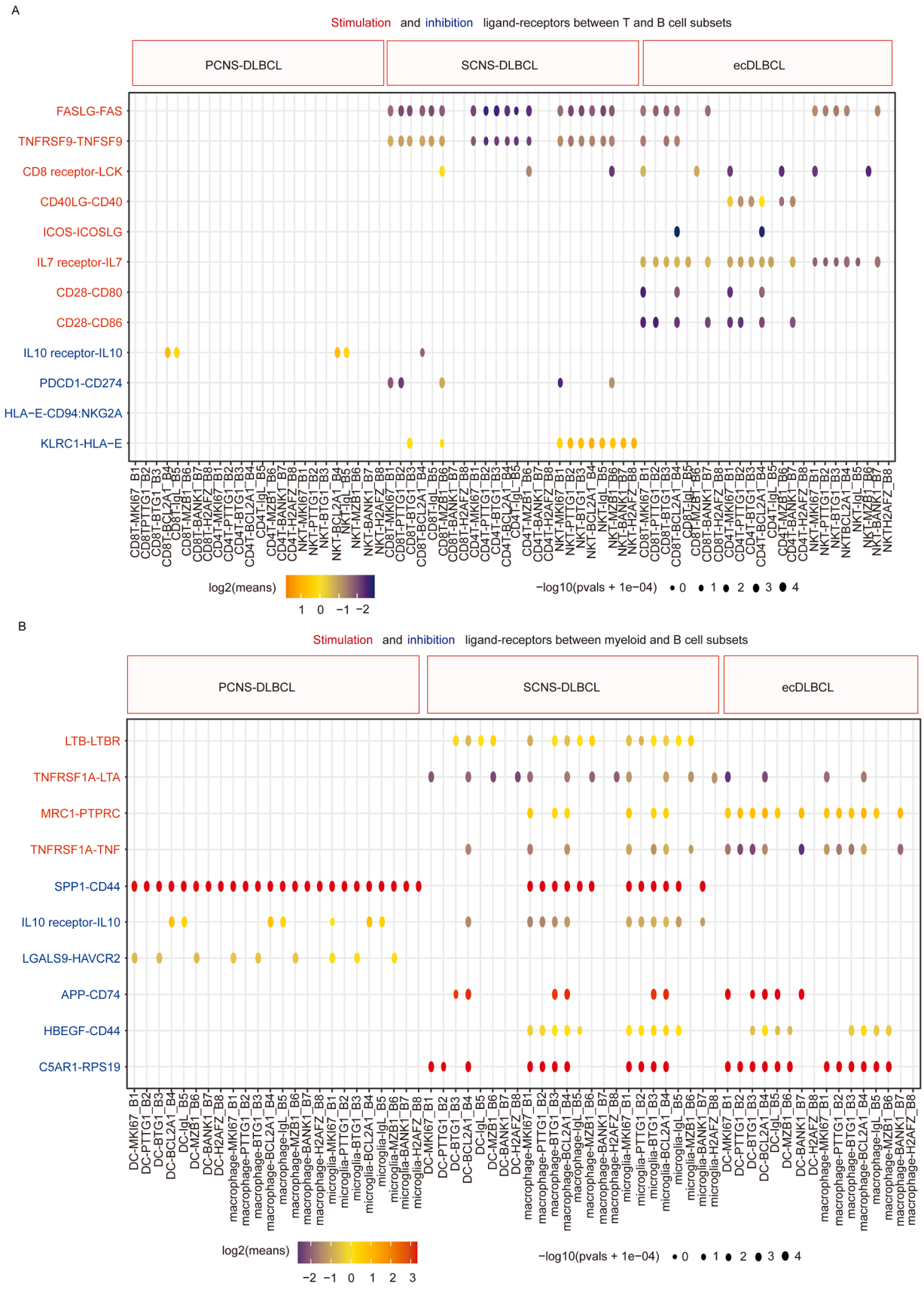
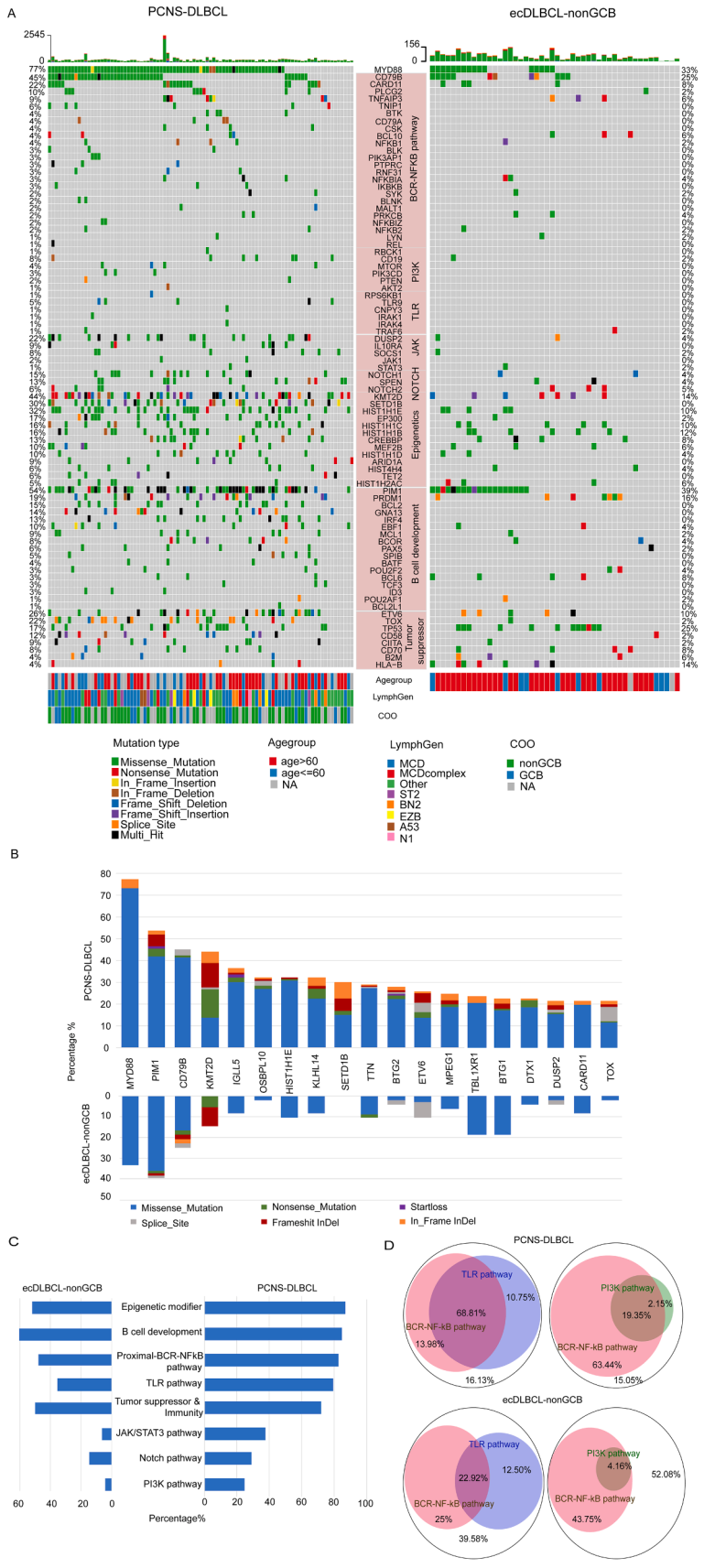


Fig. 5. Cellular crosstalk between infiltrated immune cells and B cells in different types of DLBCL
Bubble heatmap showing the mean strength of selected ligand−receptor pairs between T lymphocytes-B cells (A) and myeloid cells-B cells (B) in each type of DLBCL. Relating immune stimulation pairs are labeled in red, and inhibition is labeled in blue. Dot size indicates a p value and is colored with mean expression levels. The interacting cell clusters are listed at the bottom.



(caption on next page)

Fig. 6. Somatic mutation profile of PCNS-DLBCL vs nonGCB-ecDLBCL
(A) Mutational profile of functional gene sets and key signaling pathways in PCNS-DLBCL (left panel) and nonGCB-ecDLBCL(right panel). The bars on the top show the TMB of each sample. Signaling pathways or functional gene sets are labeled in middle, and the mutational frequencies of each gene are listed on each side. The color of the diamonds represents a distinct type of mutation. Clinical features, including COO, LymphGen and age group, are presented by different colors at the bottom.
(B) Mutated genes (>20 %) in PCNS-DLBCL (upper panel) and nonGCB-ecDLBCL (lower panel) are depicted. The height of the bars represents the mutation frequency, and the colors represent different mutation types.
(C) Frequencies of mutated functional gene sets or key signaling pathways in PCNS-DLBCL(right panel) and nonGCB-ecDLBCL(left panel).
(D) Co-occurrence of mutations in BCR-NFkB (pink), TLR (blue) and PI3K (green) signaling pathways in PCNS-DLBCL (upper panel) and nonGCB-ecDLBCL(lower panel). The size of the circle and overlap area indicate the percentage of each pathway.

ineffectiveness of COO subtyping in PCNS-DLBCL. Compared to systemic DLBCL, we found that the enrichment of DZ-related B cells, exhibited markedly proliferative features in CNS lymphoma. This finding is consistent with previous knowledge that GC-staged centroblast-B cells experience rapid division with fierce SHM and increased affinity to specific antigens [40,41]. Due to the genetic vulnerability of DZB cells, tumorigenic SHM can accumulate under particular environmental stresses, resulting in an increased likelihood of tumor formation [42,43]. The finding promoted the understanding of tumor pro-survival programs by previous PCNSL study [14]. Tumorous heterogeneities from proliferative GC staging to the mature plasmablast-like phenotype were also seen among the individuals in our study [14].

In immune-privileged DLBCL, loss of MHC expression is a fundamental phenomenon in tumor cell escape of immune monitoring [44–46]. Hypo-MHC expression was demonstrated in all B-cell subsets of our PCNS-DLBCL cohort, providing evidence that immune unresponsiveness contributes to PCNS-DLBCL pathogenesis. Consistent with the exhausted or cold signature TME in PCNS-DLBCL [12,14,38,47], the absence of immune cell infiltration and the presence of immune-suppressive cell interactions highlight the fact that the immune-privileged TME as an ecosystem supports PCNS-DLBCL growth.

Recently, it has been revealed that multiple downstream signaling pathways of the supercomplex MYD88–TLR9–BCR (My-T-BCR) orchestrate the oncogenesis of ABC-DLBCL [36]. The existence of a higher expression of IgM-TLR9 complex members in PCNS-DLBCL than in other mature B-lymphocytic malignancies indicates a role of pro-survival My-T-BCR signaling in PCNS-DLBCL pathogenesis [36]. In our study, higher frequent co-occurrence of TLR, BCR and downstream NF-kB alternations in PCNS-DLBCL than in ecDLBCL implied that this pathogenic co-operative mode was in-played to a different extent between two DLBCL entities. This might also explain the ineffectiveness of COO classification or LymphGen prognosis stratification in PCNS-DLBCL [4].

Conclusions

In summary, our study explored different status of PCNS-DLBCL and ecDLBCL from comprehensive transcriptome and genome perspectives. Genetically unstable and highly proliferated DZB cells are more sensitive to oncogenic stimulation, and the collective mutations in TLR and BCR-NFkB pathways facilitate accumulated malignant B-cell proliferation in the immune-privileged central nervous system. More profound anatomization of pathogenic mechanisms and clinical relationship is needed to devise a state-of-art strategy to deal with this rare but highly deadly disease in the era of precision medicine.

Material and method

Patient information

All the included patients were diagnosed and treated in Huashan Hospital, Fudan University. All applied samples were determined to have >60 % tumorous B cells by immunohistochemistry (IHC) staining and were diagnosed by experienced pathologists. Samples of PCNS-DLBCL ($n = 6$), SCNS-DLBCL ($n = 2$) and ecDLBCL ($n = 5$) for single-cell RNA sequencing (scRNA-seq) were obtained from 2020 to 2021. For whole-exon sequencing (WES) of PCNS-DLBCL, 31 freshly collected tumors with 26 paired PB samples and 62 FFPE tumor specimens without matched blood were analyzed ($n = 93$). Among them, newly diagnosed PCNS-DLBCL patients with qualified sequencing data, while those having ≤60 % tumorous B cells or without qualified sequencing data were excluded.

The study was approved by the Institutional Review Board of Huashan Hospital, Fudan University (KY2020-879), and was registered at Chinese Clinical Trial Registry (ChiCTR2100043151). Informed consent was obtained individually from the included patients.

Tissue dissociation and cell purification

Tumor tissues were digested in 0.25 % trypsin (Thermo Fisher, Cat. no. 25200-072) and 10 μg/mL lDNase I (Sigma, Cat. no. 11284932001) with 5 % fetal bovine serum (FBS; Thermo Fisher, Cat. no. SV30087.02) at 37°C for 30 min (icDLBCL) or 40 min (ecDLBCL). The dissociated cells were collected and filtered using a 40 μm nylon cell strainer on removal of red blood cells with 1× red blood cell lysis solution (Thermo Fisher, Cat. no. 00-4333-57). Samples with cellular viability >90 % were submitted to scRNA-seq, except for the sample obtained from the intestine, which had a threshold of >80 %.

ScRNA-seq on 10x platform

A 10× library and sequencing beads were prepared with unique molecular identifiers, with saturated loading cell barcodes to pair each cell with a bead in a gel-beads-in-emulsion. Randomly interrupted whole-transcriptome amplification products were used to prepare the sequencing library to enrich 3′-end transcripts with cell barcodes and UMIs. The remaining library construction procedures were performed according to the standard manufacturer's protocol (CG000206 RevD). Quantification of sequencing libraries was performed by a high-sensitivity DNA chip (Agilent) on a Bioanalyzer 2100 and the Qubit high-sensitivity DNA assay (Thermo Fisher Scientific). scRNA-seq was performed using 10x Chromium Single Cell 3′ Reagent Kits v3, according to the manufacturer's protocol (CG000206 RevD). The paired-ended (2 × 150 bp) sequencing of libraries was performed on a NovaSeq 6000 (Illumina). Using the STAR algorithm, the raw data were processed with the Cell Ranger 3.0.1 pipeline and were converted into FASTQ files aligned to the human reference genome hg38. Gene–barcode matrixes were generated by counting UMIs and filtering non-cell-related barcodes.

Data analysis of scRNA-seq

Two preliminary steps were performed to remove potential interference for the reads. The ambient RNA was corrected by setContaminationFraction and adjustCounts function of R package soupX, and the doublets were detected by performing a software scrublet. The output data were then processed by the R package Seurat (v4.1.1) for quality control and normalization. The counts of genes and UMIs per cell were computed using Seurat by excluding the genes expressed in fewer than one cell.

The included cells satisfied the following criteria: (1) the number of

expressed genes ranged from 200 to 5000; (2) the percentage of mitochondrial transcripts was <15 %; (3) the percentage of ribosome transcripts was <55 %; and (4) the scrublet_score was <0.35. For the intestinal sample, the range of gene numbers was 200-6000, and the percentage of mitochondria was <35 %. The median cell count was 10,445 (6516-12,905) after the filtering process.

The variables, including percentage of mitochondrial genes, percentage of ribosomal genes, and cell cycle scores (CellCycleScoring function for computing the scores of phases S and G2-M) were all regressed using the ScaleData function. To perform integrated analysis and remove batch effects of the individual expression matrix, the FindIntegrationAnchors function with parameter k.filter was set to 200, and then IntegrateData was applied. Principal component analysis was performed using the RunPCA function, where functions FindNeighbors, FindClusters, and FindIntegrationAnchors were run for clustering. The first round of clustering was performed with a resolution parameter of 0.9, with uniform manifold approximation and projection (UMAP) applied to visualize the clustering results. The resolution parameters of FindClusters for the clusters of B cells, T cells and myeloid cells were 0.55, 0.6 and 0.4, respectively. Cells coexpressing markers of B, T or myeloid cells were considered doublets and were removed.

To identify the top markers of each subcluster, the FindAllMarkers function was run with an adjusted two-sided p value≤ 0.05. Log fold changes for B cells, T cells, and myeloid cells were 0.5, 0.3 and 0.5, respectively. Dimplot from the Seurat package and VlnPlot from the MyseuratWrappers package were used to further visualize the clustering and expression markers.

Analysis of trajectory and ligand–receptor crosstalk

The expression data from the Seurat object were subjected to Monocle3 to identify the differentiation trajectory among different populations. Genes expressed in >10 cells with minimal expression >0.1 were selected, and marker genes with adjusted two-sided p value <0.05 and log fold change >0.3 in each subpopulation were used for sequential setOrderingFilter analysis. The DDRTree method was applied for dimension reduction. CellPhoneDB was applied to identify significant ligand–receptor pairs between different cell subclusters. A two-sided p value <0.05 was significant [48].

Analysis of gene expression and pathway enrichment

The differentially expressed genes between the three types of DLBCL were calculated by the FindMarkers function of Seurat, with a log fold change of 0.5 and min.pct of 0.25. By the GSEA function of the clusterProfiler package, the top 200 genes were enriched in the Molecular Signatures Database (MSigDB v7.5.1) with a two-sided p value <0.05 [49]. The 15 most enriched pathways were visualized using the enrichplot package.

DNA extraction and library construction

Genomic DNA was extracted from FFPE and fresh tumor specimens using a Qiagen kit. The Agilent SureSelect Human All Exon v8 library (Agilent Technologies, USA) was used to capture genomic DNA samples following the manufacturer's protocol. Briefly, the genomic DNA was sheared into short purified fragments and ligated onto the polished ends. Then, the libraries were amplified and hybridized with custom probes and sequenced on an Illumina sequencing platform (NovaSeq 6000, Illumina, Inc., San Diego, CA), where 150 bp paired-end reads were generated.

Preprocessing of WES data

MarkDuplicates, FixMateInformation and BaseRecalibrator of The Genome Analysis Toolkit (GATK version 4.1.2) were used to prepare BAM files. Polymerase chain reaction duplicates were removed by the REMOVE_DUPLICATES parameter. The median testing depths of fresh-frozen tumors, peripheral blood and FFPE samples were 127× (98-262×), 75× (59-159×), and 85× (48-204x), respectively. Paired-end FASTQ reads were mapped to the UCSC human genome hg19 using BWA software [50]. Mutations were called by The Genome Analysis Toolkit (GATK version 4.1.2) and Mutect (version 2) workflow [51]. For FFPE samples, the Mutect2 LearnReadOrientationModel was used to revise potential artifacts. All mutations were annotated by ANNOVAR software [52].

Pipeline of mutation filtering

The filtering pipeline of freshly collected paired tumor samples was as follows: (1) Variant allele frequency (VAF) was ≥5 % and altered reads at the mutation site were ≥3, with at least 15× coverage. (2) Variants in exons or splices were retained. (3) The variant was a non-synonymous somatic nucleotide variation (SNV). And (4) variants of frequency >1 % in the 1000 G East Asian or East Asian of Genome Aggregation Database (gnomAD_EAS) sites but not in the COSMICv92 database were excluded [53–55].

The criteria for FFPE samples were as follows: (1) VAF ≥5 % and altered reads ≥5, with at least 20× reads at the mutation site. (2) For variants whose reads ranged from 8 to 20×, VAF must be higher than 0.15 and altered reads ≥3. (3) Variants in exons or splice sites were retained, and SNVs were nonsynonymous; (4) Variants of frequency >1 % in the 1000 G East Asian or East Asian of Genome Aggregation Database (gnomAD_EAS) were filtered as germline variants if they harbored variant allele frequency >80 %, or included in avsnp150 database, or not in COSMICv92 database; (5) C > T and G > *T* substitutions with VAF <15 % were filtered out to avoid artifacts brought by formalin preparation [56,57].

Copy number alteration analysis by WES

BAM files from paired tumor-PB samples were used as inputs to calculate somatic copy number variation by Facets software [58]. Recurrent focal gain and loss regions were estimated by GISTIC 2.0 (Genomic Identification of Significant Targets in Cancer v2.0.23) with a two-sided p value of 0.25 and confidence of 0.95 [59].

Immunohistochemical staining

Rabbit antibodies against human CD19 (Abcam, ab134114), CXCR4 (Abcam, ab124824) were used in IHC on FFPE sections (4 μm thick). HRP conjugated goat anti-rabbit IgG (Servicebio, GB23303) was used as secondary antibody with DAB Horseradish Peroxidase Color Development Kit (Servicebio, G1212). EBV-encoded RNA in situ hybridization was used to determine EBV infection status as reported [60].

Statistical Analysis

Kruskal test was performed to determine whether differences of cell distribution existed among groups. Wilcox test method was then used to estimate the significance (P value) between different groups.

CRediT authorship contribution statement

Shu Wang: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Software, Visualization, Writing – original draft. **Hong Chen:** Methodology, Project administration, Resources. **Bo Dai:** Methodology, Project administration. **Kang Zheng:** Project administration, Resources. **Jiajun Zheng:** Project administration, Resources. **Yuqi Zhu:** Methodology. **Yan Yuan:** Methodology. **Tianling Ding:** Formal analysis. **Qian Wang:** Formal analysis. **Liqian Xie:** Resources. **Rui Feng:** Resources. **Fengping Zhu:** Resources.

Jianbin Xiang: Resources. **Weiqun Ding:** Resources. **Hong Ding:** Resources. **Yuan Li:** Resources. **Xiaodong Gu:** Resources. **Kunpeng Wu:** Formal analysis. **Yifan Yuan:** Resources. **Jianping Song:** Resources. **Dongxiao Zhuang:** Resources. **Haoshu Zhong:** Formal analysis. **Hanfeng Wu:** Data curation, Formal analysis, Project administration, Resources. **Ying Mao:** Formal analysis, Investigation, Methodology, Resources, Supervision. **Tong Chen:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Visualization, Writing – original draft.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the National Natural & Science Foundation of China (No. 82330004), National Key R&D Program of China (2022YFA1103500), Science and Technology Innovation Plan of Shanghai Science and Technology Commission (23Y21900300), and Program of Shanghai Municipal Human Resources for Outstanding Leadership to Dr. Tong Chen. The multiomic sequencing was supported by Medical Science Data Center in Shanghai Medical College of Fudan University. The authors would like to thank all the patients for their contributing and donating samples, and thank Prof. Xiaolei Lin (Fudan University) for her remarkable suggestion on statistical analysis. All authors read and approved the final manuscript.

Ethics approval and consent to participate

The study was approved by the Institutional Review Board of Huashan Hospital, Fudan University (KY2020-879). Informed consent was obtained individually from the included patients.

Data Availability

Additional source data involved in this research are presented in the source data files. Raw sequence data were deposited into the National Genomics Data Center (https://ngdc.cncb.ac.cn/) as of the project number (PRJCA019473). The raw sequencing data will be opened until August 30, 2025 under the Rules for the Implementation of Regulations on the Management of Human Genetic Resources from Chinese Ministry of Science and Technology.

Code availability

No custom algorithms were used in this study. The R code used to perform analysis of scRNA-seq and WES data can be obtained from the corresponding author upon request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neo.2024.101119.

References

[1] C. Grommes, L.M. DeAngelis, Primary CNS lymphoma, J. Clin. Oncol. 35 (21) (2017) 2410–2418.
[2] A. Korfel, U. Schlegel, Diagnosis and treatment of primary CNS lymphoma, Nat. Rev. Neurol. 9 (6) (2013) 317–327.
[3] R. Alaggio, C. Amador, I. Anagnostopoulos, et al., The 5th edition of the world health organization classification of haematolymphoid tumours: lymphoid neoplasms, Leukemia 36 (7) (2022) 1720–1748.
[4] I. Hernández-Verdin, E. Kirasic, K. Wienand, et al., Molecular and clinical diversity in primary central nervous system lymphoma, Ann. Oncol. 34 (2) (2023) 186–199.
[5] C. Grommes, L. Nayak, H.W. Tun, et al., Introduction of novel agents in the treatment of primary CNS lymphoma, Neuro Oncol. 21 (3) (2019) 306–313.
[6] A.A. Alizadeh, M.B. Eisen, R.E. Davis, et al., Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling, Nature 403 (6769) (2000) 503–511.
[7] A. Rosenwald, G. Wright, W.C. Chan, et al., The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma, N. Engl. J. Med. 346 (25) (2002) 1937–1947.
[8] K. Dybkær, M. Bøgsted, S. Falgreen, et al., Diffuse large B-cell lymphoma classification system that associates normal B-cell subset phenotypes with prognosis, J. Clin. Oncol. 33 (12) (2015) 1379–1388.
[9] C.B. Steen, B.A. Luca, M.S. Esfahani, et al., The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma, Cancer Cell 39 (10) (2021), 1422-37.e10.
[10] A.B. Holmes, C. Corinaldesi, Q. Shen, et al., Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome, J. Exp. Med. 217 (10) (2020).
[11] S. Camilleri-Broët, E. Crinière, P. Broët, et al., A uniform activated B-cell-like immunophenotype might explain the poor prognosis of primary central nervous system lymphomas: analysis of 83 cases, Blood 107 (1) (2006) 190–196.
[12] M. Heming, S. Haessner, J. Wolbert, et al., Intratumor heterogeneity and T cell exhaustion in primary CNS lymphoma, Genome Med. 14 (1) (2022) 109.
[13] I. Hernández-Verdin, E. Kirasic, K. Wienand, et al., Molecular and clinical diversity in primary central nervous system lymphoma, Ann. Oncol. (2022).
[14] N. Liu, C. Jiang, X. Yao, et al., Single-cell landscape of primary central nervous system diffuse large B-cell lymphoma, Cell Discov. 9 (1) (2023) 55.
[15] F. Klemm, R.R. Maas, R.L. Bowman, et al., Interrogation of the microenvironmental landscape in brain tumors reveals disease-specific alterations of immune cells, Cell 181 (7) (2020), 1643-60.e17.
[16] T. Risom, D.R. Glass, I. Averbukh, et al., Transition to invasive breast cancer is associated with progressive changes in the structure and composition of tumor stroma, Cell 185 (2) (2022), 299-310.e18.
[17] G.W. Wright, D.W. Huang, J.D. Phelan, et al., A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications, Cancer Cell 37 (4) (2020), 551-68.e14.
[18] R. Schmitz, G.W. Wright, D.W. Huang, et al., Genetics and pathogenesis of diffuse large B-Cell lymphoma, N. Engl. J. Med. 378 (15) (2018) 1396–1407.
[19] B. Chapuy, C. Stewart, A.J. Dunford, et al., Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes, Nat. Med. 24 (5) (2018) 679–690.
[20] K. Fukumura, M. Kawazu, S. Kojima, et al., Genomic characterization of primary central nervous system lymphoma, Acta Neuropathol. 131 (6) (2016) 865–875.
[21] B. Chapuy, M.G. Roemer, C. Stewart, et al., Targetable genetic features of primary testicular and primary central nervous system lymphomas, Blood 127 (7) (2016) 869–881.
[22] D. Morgan, V. Tergaonkar, Unraveling B cell trajectories at single cell resolution, Trends. Immunol. 43 (3) (2022) 210–229.
[23] O. Bannard, R.M. Horton, C.D. Allen, et al., Germinal center centroblasts transition to a centrocyte phenotype according to a timed program and depend on the dark zone for effective selection, Immunity. 39 (5) (2013) 912–924.
[24] D.V. Fyodorov, B.R. Zhou, A.I. Skoultchi, et al., Emerging roles of linker histones in regulating chromatin structure and function, Nat. Rev. Mole. Cell Biol. 19 (3) (2018) 192–206.
[25] J.E. Noll, K. Vandyke, D.R. Hewett, et al., PTTG1 expression is associated with hyperproliferative disease and poor prognosis in multiple myeloma, J. Hematol. Oncol. 8 (2015) 106.
[26] U. Klein, R. Dalla-Favera, Germinal centres: role in B-cell physiology and malignancy, Nat. Rev. Immunol. 8 (1) (2008) 22–33.
[27] N.S. De Silva, U. Klein, Dynamics of B cells in germinal centres, Nat. Rev. Immun. 15 (3) (2015) 137–148.
[28] Z. Li, Q. Yang, X. Tang, et al., Single-cell RNA-seq and chromatin accessibility profiling decipher the heterogeneity of mouse γδ T cells, Sci. Bull. 67 (4) (2022) 408–426.
[29] L. Cascione, L. Aresu, M. Baudis, et al., DNA Copy number changes in diffuse large b cell lymphomas, Front. Oncol. 10 (2020) 584095.
[30] K.E. de Visser, J.A. Joyce, The evolving tumor microenvironment: from cancer initiation to metastatic outgrowth, Cancer Cell 41 (3) (2023) 374–403.
[31] T. Masuda, R. Sankowski, O. Staszewski, et al., Microglia heterogeneity in the single-cell era, Cell Rep. 30 (5) (2020) 1271–1281.
[32] A. Aravind, A. Palollathil, D.A.B. Rex, et al., A multi-cellular molecular signaling and functional network map of C-C motif chemokine ligand 18 (CCL18): a chemokine with immunosuppressive and pro-tumor functions, J. Cell Commun. Signal. 16 (2) (2022) 293–300.
[33] Wang H. Microglia, Heterogeneity in Alzheimer's disease: insights from single-cell technologies, Front. Synaptic. Neurosci. 13 (2021) 773590.
[34] S. Cheng, Z. Li, R. Gao, et al., A pan-cancer single-cell transcriptional atlas of tumor infiltrating myeloid cells, Cell 184 (3) (2021) 792–809.e23.
[35] E. Wang, X. Mi, M.C. Thompson, et al., Mechanisms of resistance to noncovalent bruton's tyrosine kinase inhibitors, N. Engl. J. Med. 386 (8) (2022) 735–743.
[36] J.D. Phelan, R.M. Young, D.E. Webster, et al., A multiprotein supercomplex controlling oncogenic signalling in lymphoma, Nature 560 (7718) (2018) 387–391.
[37] Y. Miao, L.J. Medeiros, Y. Li, et al., Genetic alterations and their clinical implications in DLBCL, Nat. Rev. Clinc. Oncol. 16 (10) (2019) 634–652.

[38] Y. Xia, T. Sun, G. Li, et al., Spatial single cell analysis of tumor microenvironment remodeling pattern in primary central nervous system lymphoma, Leukemia 37 (7) (2023) 1499–1510.
[39] M. Montesinos-Rongen, A. Brunn, S. Bentink, et al., Gene expression profiling suggests primary central nervous system lymphomas to be derived from a late germinal center B cell, Leukemia 22 (2) (2008) 400–405.
[40] G.D. Victora, T.A. Schwickert, D.R. Fooksman, et al., Germinal center dynamics revealed by multiphoton microscopy with a photoactivatable fluorescent reporter, Cell 143 (4) (2010) 592–605.
[41] C.D. Allen, T. Okada, H.L. Tang, et al., Imaging of germinal center selection events during affinity maturation, Science (New York, NY) 315 (5811) (2007) 528–531.
[42] R. Küppers, Mechanisms of B-cell lymphoma pathogenesis, Nat. Rev. Cancer 5 (4) (2005) 251–262.
[43] A. Nussenzweig, M.C. Nussenzweig, Origin of chromosomal translocations in lymphoid cancer, Cell 141 (1) (2010) 27–38.
[44] S.A. Riemersma, E.S. Jordanova, R.F. Schop, et al., Extensive genetic alterations of the HLA region, including homozygous deletions of HLA class II genes in B-cell lymphomas arising in immune-privileged sites, Blood 96 (10) (2000) 3569–3577.
[45] M. Booman, K. Szuhai, A. Rosenwald, et al., Genomic alterations and gene expression in primary diffuse large B-cell lymphomas of immune-privileged sites: the importance of apoptosis and immunomodulatory pathways, J. Pathol. 216 (2) (2008) 209–217.
[46] M. Booman, J. Douwes, A.M. Glas, et al., Mechanisms and effects of loss of human leukocyte antigen class II expression in immune-privileged site-associated B-cell lymphoma, Clin. Cancer Res. 12 (9) (2006) 2698–2705.
[47] M. Alcantara, J. Fuentealba, C. Soussain, Emerging landscape of immunotherapy for primary central nervous system lymphoma, Cancers. (Basel) 13 (20) (2021).
[48] M. Efremova, M. Vento-Tormo, S.A. Teichmann, et al., CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes, Nat. Protoc. 15 (4) (2020) 1484–1506.
[49] A. Liberzon, C. Birger, H. Thorvaldsdóttir, et al., The Molecular Signatures Database (MSigDB) hallmark gene set collection, Cell Syst. 1 (6) (2015) 417–425.
[50] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics. 25 (14) (2009) 1754–1760.
[51] A. McKenna, M. Hanna, E. Banks, et al., The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res. 20 (9) (2010) 1297–1303.
[52] K. Wang, M. Li, H. Hakonarson, ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data, Nucleic. Acids. Res. 38 (16) (2010) e164.
[53] G.R. Abecasis, A. Auton, L.D. Brooks, et al., An integrated map of genetic variation from 1,092 human genomes, Nature 491 (7422) (2012) 56–65.
[54] J. Zou, G. Valiant, P. Valiant, et al., Quantifying unobserved protein-coding variants in human populations provides a roadmap for large-scale sequencing projects, Nat. Commun. 7 (2016) 13293.
[55] S.A. Forbes, D. Beare, H. Boutselakis, et al., COSMIC: somatic cancer genetics at high-resolution, Nucleic. Acids. Res. 45 (D1) (2017). D777-d83.
[56] S. Devarakonda, F. Rotolo, M.S. Tsao, et al., Tumor mutation burden as a biomarker in resected non-small-cell lung cancer, J. Clin. Oncol. 36 (30) (2018) 2995–3006.
[57] S.Q. Wong, J. Li, A.Y. Tan, et al., Sequence artefacts in a prospective series of formalin-fixed tumours tested for mutations in hotspot regions by massively parallel sequencing, BMC. Med. Genomics. 7 (2014) 23.
[58] R. Shen, V.E. Seshan, FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing, Nucleic. Acids. Res. 44 (16) (2016) e131.
[59] C.H. Mermel, S.E. Schumacher, B. Hill, et al., GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers, Genome Biol. 12 (4) (2011) R41.
[60] J. Morscio, D. Dierickx, J.F. Ferreiro, et al., Gene expression profiling reveals clear differences between EBV-positive and EBV-negative posttransplant lymphoproliferative disorders, Am. J. TransPlant 13 (5) (2013) 1305–1316.