Front. Med. 2023, 17(5): 889–906
https://doi.org/10.1007/s11684-023-0994-x

RESEARCH ARTICLE

Analysis of the genomic landscape of primary central nervous system lymphoma using whole-genome sequencing in Chinese patients

Xianggui Yuan[1,*], Teng Yu[1,*], Jianzhi Zhao[1,*], Huawei Jiang[1], Yuanyuan Hao[1], Wen Lei[1], Yun Liang[1], Baizhou Li (✉)[2], Wenbin Qian (✉)[1,3,4]

[1]*Department of Hematology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China;* [2]*Department of Pathology, the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China;* [3]*National Clinical Research Center for Hematologic Diseases, the First Affiliated Hospital of Soochow University, Suzhou 215006, China;* [4]*Cancer Institute (Key Laboratory of Cancer Prevention and Intervention, China National Ministry of Education), the Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou 310009, China*

© Higher Education Press 2023

Abstract **Primary central nervous system lymphoma (PCNSL) is an uncommon non-Hodgkin's lymphoma with poor prognosis. This study aimed to depict the genetic landscape of Chinese PCNSLs. Whole-genome sequencing was performed on 68 newly diagnosed Chinese PCNSL samples, whose genomic characteristics and clinicopathologic features were also analyzed. Structural variations were identified in all patients with a mean of 349, which did not significantly influence prognosis. Copy loss occurred in all samples, while gains were detected in 77.9% of the samples. The high level of copy number variations was significantly associated with poor progression-free survival (PFS) and overall survival (OS). A total of 263 genes mutated in coding regions were identified, including 6 newly discovered genes (*ROBO2*, *KMT2C*, *CXCR4*, *MYOM2*, *BCLAF1*, and *NRXN3*) detected in ⩾ 10% of the cases. *CD79B* mutation was significantly associated with lower PFS, *TMSB4X* mutation and high expression of TMSB4X protein was associated with lower OS. A prognostic risk scoring system was also established for PCNSL, which included Karnofsky performance status and six mutated genes (*BRD4*, *EBF1*, *BTG1*, *CCND3*, *STAG2*, and *TMSB4X*). Collectively, this study comprehensively reveals the genomic landscape of newly diagnosed Chinese PCNSLs, thereby enriching the present understanding of the genetic mechanisms of PCNSL.**

Keywords primary central nervous system lymphoma; whole-genome sequencing; *TMSB4X*; copy number variation; gene mutation

Introduction

Primary central nervous system lymphoma (PCNSL) is an uncommon non-Hodgkin's lymphoma confined to the central nervous system. Approximately 95% of PCNSLs are diffuse large B cell lymphoma (DLBCL), with the remaining 5% belonging to Burkitt, marginal zone, lymphoblastic, or T cell lymphomas [1,2]. This disease is most commonly detected among immunocompromised patients. However, its incidence has increased among healthy elderly individuals over the past decade [3]. High-dose methotrexate (HD-MTX)-based chemotherapy is the main strategy for treating PCNSL, but this treatment only has a 30% 5-year overall survival (OS) rate, with many patients relapsing [4]. Therefore, the genetic characteristics of PCNSL should be understood to develop effective treatments.

Unlike systemic DLBCL, studies focusing on PCNSL are limited due to its low incidence and difficulty in sample obtainment. With the development of modern high-throughput molecular techniques, several recurrent chromosomal abnormalities and mutations have been identified in PCNSL, including inactivating the mutation of *PRDM1* [5] and activating the mutation of the oncogene *CARD11*, which may induce the activation of

Received November 23, 2022; accepted March 6, 2023
Correspondence: Wenbin Qian, qianwb@zju.edu.cn;
Baizhou Li, alexlibz@126.com

*These three authors contributed equally to this work.

the NF-κB pathway [6] and Ig-H-BCL6 translocation [7]. Braggio *et al*. [8] found the biallelic inactivation of the *TOX* and *PRKCD* genes in PCNSL samples using whole-exome sequencing (WES) but not in systemic DLBCL. Additionally, growing evidence shows that gene mutations are closely linked to the prognosis of PCNSL. For instance, using WES, Fukumura *et al*. [9] demonstrated that the focal deletions or somatic mutations of HLA genes are associated with the poor prognosis of PCNSL patients, while Zhou *et al*. [10] revealed that the mutations of the *CD79B* and *GNA13* genes (members of the NF-κB signaling pathway) were related to inferior progression-free survival (PFS) in PCNSL patients. These findings provide novel insights for understanding the pathogenesis of PCNSL. However, the sample size of these studies was small due to the rare incidence of PCNSL. In addition, whole-genome sequencing (WGS) has rarely been applied due to its high cost. Therefore, larger sample sizes, especially Chinese PCNSLs, should be analyzed using WGS.

This study aimed to reveal the genetic characteristics of Chinese PCNSLs using WGS with a focus on structural variations (SVs), copy number variations (CNVs), and gene mutation profiles, and then to explore the relationship of these molecular aberrations with the patients' clinicopathological features, response to HD-MTX-based polychemotherapies, and prognosis, and finally to explore the differences in gene mutation profiles between Chinese and other cohorts,

Materials and methods

Clinical samples

A total of 95 tumor specimens and 21 paired bone marrow tissues were obtained from 95 PCNSL patients between March 25, 2009 and June 22, 2020 for use in WGS. Following a quality assessment, 68 tumor specimens and 8 paired bone marrow specimens were sequenced successfully. All patients were newly diagnosed with PCNSL (DLBCL) without immune deficiencies, and the samples were collected prior to any type of treatment. Bone marrow tissues obtained from 8 of the 68 patients served as negative controls, and these tissues were not infiltrated. All patients were monitored up to March 1, 2021. General clinical characteristics, including age, sex, Hans' classification, number of lesions, Eastern Cooperative Oncology Group (ECOG), Karnofsky performance status (KPS), International Extranodal Lymphoma Study Group scores (IELSG), Memorial Sloan–Kettering Cancer Center scores (MSKCC), lactic dehydrogenase (LDH) level, number of lesions, deep involvement, BCL2/MYC expression, lymphocyte count, therapeutic response, and survival time, were collected.

Experiments involving human clinical samples were conducted according to the *Helsinki Declaration* and were approved by the Institutional Ethical Standard Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine. Informed consent forms were signed by each patient.

Whole-genome sequencing (WGS)

Genomic DNA was extracted using a formalin-fixed and paraffin-embedded (FFPE) tissue kit (No. 56404, Qiagen, NV, Venlo, Netherlands) following the manufacturer's instructions. The genomic DNA was cut into fragments of ~200 bp using a focused ultrasonicator (No. M220, Covaris, Woburn, MA, USA). After quality control and equimolar pooling, WGS was performed on the DNBSEQ-T7 sequencing instruments (MGI, Guangdong, China) with a sequencing data volume of 90 G and a mean sequencing depth of 58.82× (range: 23.23×–96.61×). FastQC (version 1.11.4) was used to assess the quality of the raw sequencing data. These data were then processed using Trimmomatic (version 3.6) to remove sequencing adapters and low-quality reads, that is, the joint sequence fragments of the 3′ end and low-quality fragments with a Q value of < 25 and fragments of < 35 bp. To prepare read alignments for analysis, all sequencing data were processed through the Broad Institute's data processing pipeline. Reads were aligned to the Human Genome Reference Consortium build 37 (GRCh37) using BWA (version 0.5.9-tpx). The BAM files contained reads aligned to the human genome, and their quality scores were recalibrated using the Table Recalibration tool in the Genome Analysis Toolkit (version 4.1.4.0). Variant detection and analysis of the BAM files were performed using the Broad Institute's Cancer Genome Analysis infrastructure program Mutect 2.

Data processing

SVs, deletions, duplications, inversions, and translations were annotated using Lumpy-0.2.13 and were screened based on the following filtering conditions: (1) deletions, duplications, and inversions sharing a common overlap of > 10% with the 8 negative control samples were removed; (2) translations involving the same two sites were filtered out; (3) SVs with a mutation frequency of < 15% were filtered out; and (4) SVs with no side or one side located at the gene sites were filtered out. The circos image (circlizeR, R software, R Core Team, Vienna, Austria) demonstrated fusions in different genes identified in ⩾ 2 cases.

Following a comparison with GRCh37, the CNVs were analyzed using the CNV kit software package and Nexus software version 5 (Biodiscovery, El Segundo, CA). A copy number of 0 indicated loss, and a copy number of

⩾ 3 indicated gain. The "Copynumber" package (version 1.26.0) of the R software was used to draw the frequency plot and heatmap.

Variants of the coding regions were screened based on the following filtering conditions: (1) mutations with mutation allele frequency (MAF) ⩾ 0.001 in the 1000 Genomes Project or 1000 Genomes East Asian databases or with a MAF ⩾ 0.001 in the Exome Aggregation Consortium (ExAC) or ExAC East Asian databases were removed; (2) mutations detected in the 8 negative controls were removed; (3) mutations with MAF < 0.001 in GenomAD and MAF < 0.001 in ExAC were reserved; (4) mutations with sites that were functionally annotated in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database were retained; and (5) mutations in genes not included in the union set or were derived from the full genetic database for cancer (MSK-Impact, Foundation One) or DLBCL-related gene lists were filtered out (Supplementary File 1). The "Maftools" package (version 2.2.10) of the R software was used to draw the horizontal histogram, which showed the genes with high mutation frequencies, and to perform the co-mutation analysis. The mutations in coding regions were annotated using Annovar (version 2017-07-17).

Functional enrichment analysis

Gene ontology (GO) analysis, which covered biological processes, cellular components, and molecular function terms, was performed using DAVID to evaluate the enriched functions of the CNVs and mutated genes. Fisher's exact test was applied to detect the overlap between the genes and the GO annotation list beyond that which would be expected by chance. The KEGG database was used to understand the high-level functions and effects of biological systems. DAVID was also applied to assess the KEGG pathway enrichment of the mutated genes. A P value of < 0.05 was considered significantly enriched.

Establishment of a prognostic risk scoring system for PCNSL

To establish the training and test cohorts, the 68 samples were randomly divided into a training set and a test set with 51 (75%) and 17 samples (25%), respectively. The training set was subjected to 10-fold cross-validation to account for variability and provide risk estimates. All mutated genes and four clinicopathologic features (age, KPS, deep involvement, and LDH level) were included in the models.

Conventional logistic regression was used to train the models for assessing mortality risk. The methods included logistic regression with a forward selection of variables, and logistic regression with lasso regularization was carried out using the "Glmnet" package. Receiver operating characteristic (ROC) curves were used to estimate model discrimination by calculating the area under the curve (AUC). The test set was divided into low- and high-risk groups according to the cutoff value determined by the ROC curve. Afterward, the OS of the two risk groups was analyzed by Kaplan–Meier (K–M) curves with log-rank tests. A nomogram model was also established according to previously described methods [11] to visually determine the value of this prognostic risk scoring system in predicting the OS of each PCNSL patient.

Immunohistochemistry

A total of 68 paraffin-embedded newly diagnosed PCNSLs were studied using a rabbit anti-TMSB4X polyclonal antibody (Ag13914, Proteintech) at 1:400 dilution on an automated immunostainer (Bond, Leica, Wetzlar, Germany).

Statistical analysis

The Maftools ("clinicalEnrichment") R package was applied to assess the relationship between the mutated genes and clinical characteristics via one-tailed Fisher's exact tests. K–M curves with log-rank tests were used to analyze the relationship between the mutation profiles and the PFS and OS of patients with PCNSL. A P value of < 0.05 was considered significant.

Results

Patient characteristics and sequencing data

A total of 68 patients with newly diagnosed PCNSL were sequenced successfully and included in this study. All patients were identified as DLBCL subtype, with 26.5% (18/68) of germinal center B cell-like (GCB) subtype and 70.6% (48/68) of non-GCB subtype. Their clinicopathologic features are summarized in Table 1. Their median age at diagnosis was 58 years (range: 24–80), 58.8% (40/68) were male and 41.2% (28/68) were female, 55.9% (38/68) had unifocal and 44.1% (30/68) had multifocal lesions, 54.4% (37/68) had an ECOG score of 1 to 2, and 45.6% (31/68) had an ECOG score of 3 to 4. After the diagnosis, 37 patients received HD-MTX-based chemotherapy only, 28 received HD-MTX-based chemotherapy combined with radiotherapy, and the other 3 were given HD-MTX-based chemotherapy followed by auto-transplantation. The genomic DNA isolated from tumor samples and 8 paired bone marrow control samples was subjected to WGS with a mean sequencing depth of 58.82× (range: 23.23×–96.61×).

Table 1 Characteristics of patients with PCNSL (n = 68)

Factor	n (%)
Age	
⩽ 60, year	35 (51.5)
> 60, year	33 (48.5)
Sex	
Male	40 (58.8)
Female	28 (41.2)
ECOG PS	
1	4 (5.9)
2	33 (48.5)
3	18 (26.5)
4	13 (19.1)
KPS	
< 70	41 (60.3)
⩾ 70	27 (39.7)
Number of lesions	
Unifocal	38 (55.9)
Multifocal	30 (44.1)
Deep involvement	
Yes	35 (51.7)
No	33 (48.5)
LDH	
High	8 (11.8)
Normal	60 (88.2)
Lymphocyte count	
⩽ 2 10^9/L	44 (64.7)
> 2 10^9/L	24 (35.3)
IELSG score	
Low risk (0–1)	12 (17.6)
Middle risk (2–3)	41 (60.3)
High risk (4–5)	15 (22.1)
MSKCC score	
Low risk (age ⩽ 50)	8 (11.8)
Middle risk (age > 50, KPS ⩾ 70)	26 (38.2)
High risk (age > 50, KPS < 70)	34 (50)
Pathology	
DLBCL	68 (100)
Hans' classification	
GCB	18 (26.5)
Non-GCB	48 (70.6)
NA	2 (2.9)
BCL2/MYC double expression	
Yes	14 (20.6)
No	54 (79.4)
Treatment method	
MTX-based chemotherapy	37 (54.4)
MTX-based chemotherapy and radiotherapy	28 (41.2)
MTX-based chemotherapy and auto-transplantation	3 (4.4)
Efficacy following chemotherapy	
CR	31 (45.6)
PR	15 (22.1)
SD	1 (1.5)
PD	15 (22.1)
NA	6 (8.8)

CR, complete remission; DLBCL, diffuse large B cell lymphoma; ECOG PS, Eastern Cooperative Oncology Group performance status; GCB, germinal center B cell-like; IELSG, International Extranodal Lymphoma Study Group; LDH, lactate dehydrogenase; MTX, methotrexate; MSKCC, Memorial Sloan–Kettering Cancer Center prognostic score; PR, partial remission; PD, progressive disease; SD, stable disease.

SVs of the 68 PCNSL samples identified by WGS

SVs, including deletions, duplications, inversions, and translocations (break end), are common causes of gene functional alterations and have been identified in several kinds of cancers [12–14]. In this study, 23 338 SVs were identified in the 68 PCNSL samples (mean: 343, range: 1–858), with deletions being the most frequent events (Fig. 1A). To explore how SVs drive PCNSL occurrence, those genes related to SV breakpoints were identified across the whole genome (Fig. 1B). The genes marked with red were located in the frequent SV region and included *CACNA1B*, *BCLAF1*, *OPRM1*, and *RSPO1*. The top 35 genes containing SV breakpoints are shown in Fig. 1C, with the top 10 genes being *RAD51B*, *CHD2*, *PTPRD*, *BCLAF1*, *ETV6*, *FGF12*, *MS12*, *NRXN3*, *TP63*, and *MAP3K13*. Given that translations of *BCL6, BCL2, and MYC* were reported in PCNSL, these rearrangements were also screened in this cohort. *BCL6* and *BCL2* rearrangements were found in 5 (7.4%) and 1 (1.5%) of the 68 samples, respectively, while no *MYC* rearrangement was identified. This difference may be due to the use of a different study cohort.

Further analysis showed that *CLCNKB-FAM131C*, *FAM234B-MIR7641-2*, *MYOF-WBP1L*, and *SUSD6-CCDC50* were the most frequent fusions, which were detected in 28, 19, 17, and 11 samples, respectively. The fusions of different genes detected in ⩾ 2 samples are shown in Fig. S1. The other frequent fusions included *SLC2A5-BTBD7*, *MRPS18A-B4GALT1*, *FAM238B-FAM238C*, *LLPH-DT-TMBIM4*, *KCNMA1-AS3*, and *KCNMA1*.

To explore the clinical value of SVs in PCNSL, the associations between SV numbers and the OS or PFS of patients with PCNSL were assessed. These 68 patients were divided into an SV-high group (median SVs ⩾ 359)

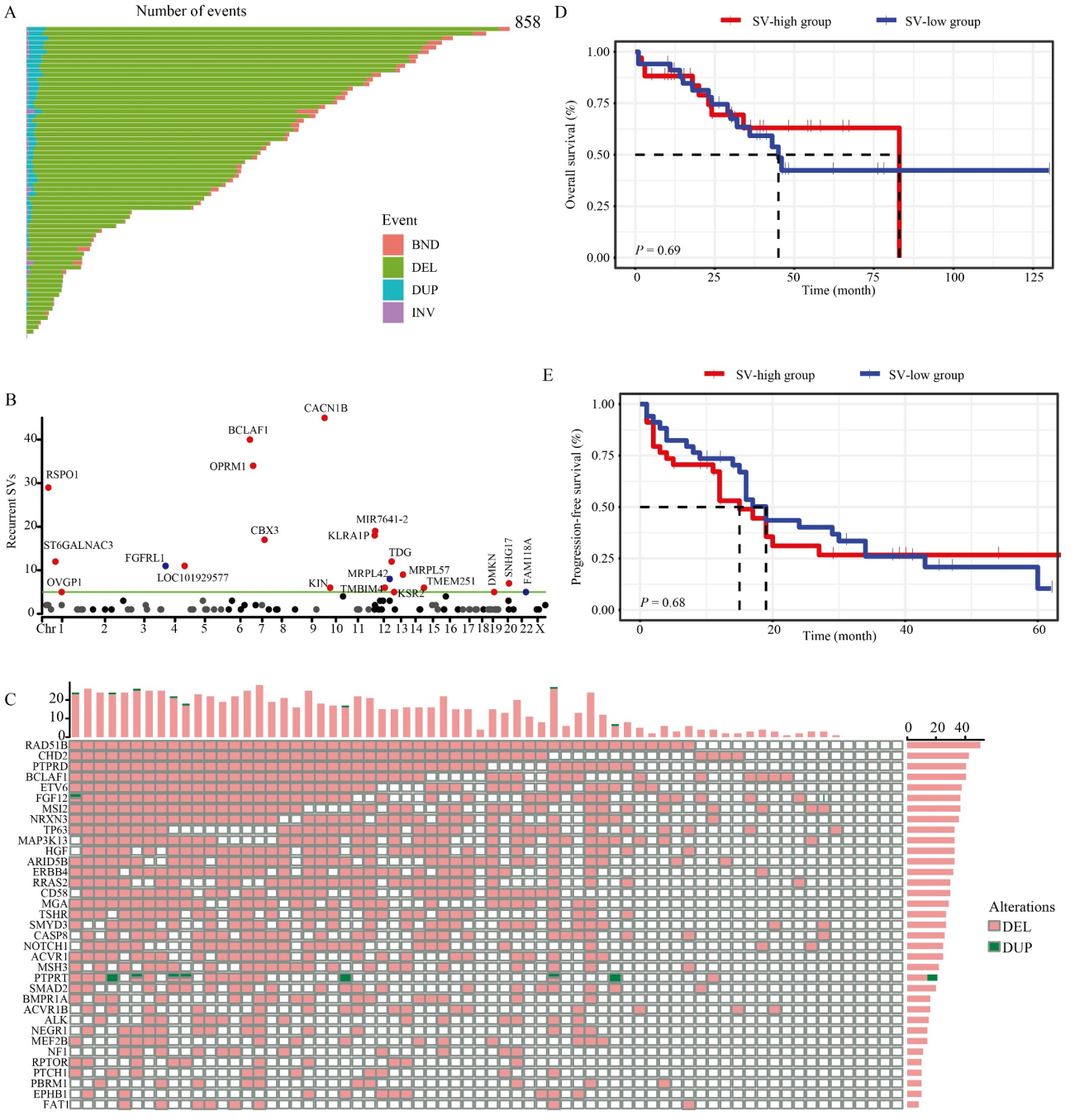

Fig. 1 Structural variation (SV) analysis of the 68 PCNSL patients. (A) Bar chart showing the SVs in each case of PCNSL. (B) Significantly altered genes induced by the SVs (⩾ 5%). Each dot represents an SV-affected gene. Red represents gene duplication, and blue represents gene deletion. (C) Summary of the significantly altered genes (top 35) affected by SV breakpoints across the whole genome. K–M curves were used to analyze the relationship of SV number with the overall survival (D) and progression-free survival (E) of the 68 patients with PCNSL.

and a SV-low group (median SVs < 359), with 34 patients in each group. The SV number did not significantly influence either the OS or PFS of PCNSL patients (Fig. 1C and 1D).

CNV profiles across the 68 PCNSL samples

The CNVs of the 68 samples were profiled to identify significant peaks that might contain potential driver genes. A total of 2952 CNVs were found in the samples (mean: 43, range: 8–493). Copy loss occurred in all 68 samples, while gains were detected in 53 samples (77.9%). The most common losses were those of 16p11.2 (97.1%), 6p11.2 (97.1%), 22q11.1 (91.2%), 2q11.1 (91.2%), 9q13 (89.7%), 2q11.2 (89.7%), 9p11.2 (86.8%), 1p36.21 (82.4%), 22q11.22 (77.9%), and 15q11.2

(75.0%), whereas the most common gains were those of 6q13 (22.1%), 16q13 (22.1%), 17q11.2 (22.1%), 3p21.31 (20.1%), 7p13 (19.1%), 9q13 (17.6%), 2q13 (17.6%), 19p13.2 (17.6%), 9p13.2 (17.6%), and 17p13.2 (17.6%) (Fig. 2A). Previously reported CNVs, such as deletions of 6p21, 6q, and 9p21.3 and gains of 7q, 11q, and 9p24.1 [8,15–17], were also found in the cohort, with frequencies of 41.2%, 20.6%, 11.8%, 25.0%, 23.5%, and 2.9%, respectively.

GO enrichment analysis was performed to assess the biological functions of the 7791 genes related to the CNVs. The genes were significantly enriched in the biological process (BP) terms of innate immune response, natural killer cell activation involved in immune response, B cell proliferation, humoral immune response, and positive regulation of peptidyl−serine phosphorylation of STAT protein, the cellular component (CC) terms of nucleosome and intermediate filament, and the molecular function (MF) terms of type I interferon receptor binding and ribonuclease activity (Fig. 2B).

To reveal the role of CNVs in predicting PCNSL prognosis, the 68 samples were divided into a CNV-high group (median CNV: ⩾ 24.5) and a CNV-low group (median CNV: < 24.5). Interestingly, patients with high CNV levels had lower OS (Fig. 2C) and PFS (Fig. 2D) compared with patients with low CNV levels, hence underscoring the vital role of CNV number in PCNSL.

Recurrently mutated protein-coding genes across the 68 PCNSL samples

The mutation profiles of coding regions in the 68 PCNSL patients were evaluated based on the WGS results. A total of 1424 mutated sites in 263 genes were detected (Supplementary File 2). The median number of gene alterations per patient was 2 (range: 1–46). The most common mutation type was missense mutation (Fig. 3A). The number of detected single-nucleotide-variants (SNVs) was approximately 11–12-fold the number of insertions and deletions (Fig. 3B). Among the 263 mutated genes, 36 frequently mutated in ⩾ 10% of the samples, with the top 10 genes being *IGLL5* (68%), *PIM1* (68%), *MYD88* (53%), *CD79B* (47%), *BTG2* (40%), *KMT2D* (37%), *TBL1XR1* (31%), *PCLO* (28%), *HIST1H1E* (26%), and *BTG1* (25%) (Fig. 3C).

By checking the 263 mutation genes in the cancer hotspots, Genomic Data Commons, intogene, and OncoKB databases, 263 genes, except for *LRRN3*, *SLITRK3*, *RAG2*, *ZNF608*, *HIST2H2AB*, and *TLDC2* were identified as cancer drivers, among which 81 (*AR, B2M, BCL2, BTG1, BTG2, CARD11, CD79A, CD79B, CDKN2A, CREBBP, DTX1, DUSP2, HIST1H1E, HLA-B, IGLL5, IRF4, ITPKB, KLHL6, KMT2D, MPEG1, MYC, MYD88, PAX5, PIK3C2G, PIM1, TBL1XR1, TMSB4X, TP53, UBE2A, XPO1, TET2, SPOP, EP300, DDX3X, CXCR4, ATM, AMER1, EZH2, BRCA1, SMARCA4, BTK, PRDM1, MET, JAK3, CDK12, TBX3, ERBB4, MED12, ATR, FGFR4, PTCH1, NF1, STAG2, PBRM1, KDR, ARID1A, FAS, STAT3, FGFR2, BRAF, APC, CD58, EBF1, SPEN, ID3, GNA13, NOTCH2, CIITA, DOT1L, SOCS1, TNFAIP3, CDKN1B, NOTCH1, MEF2B, SGK1, CCND3, BCL6, STAT6, PTPN6, MAP2K1,* and *CDH1*) were DLBCL driver genes.

GO enrichment and KEGG pathway analysis were then applied to assess the enriched signaling pathways of the 263 mutated genes. The GO analysis showed that these mutated genes were enriched in “BP” terms of regulation of phosphatidylinositol 3 kinase (PI3K) signaling and positive regulation of ERK1 and ERK2 cascades, MAPK cascade, and B cell receptor signaling pathway, the “CC” terms of nucleus and nucleoplasm, and the “MF” terms of PI3K activity, Ras guanyl nucleotide exchange factor activity, ATP binding, and protein binding (Fig. S2A). The KEGG analysis demonstrated that the mutated genes were enriched in the Rap1 signaling pathway, PI3K-Akt signaling pathway, signaling pathways regulating the pluripotency of stem cells, ErbB signaling pathway, and B cell receptor signaling pathway (Fig. S2B). Genes in the BCR/PI3K signaling pathway were significantly mutated, including the previously reported genes *CD79B*, *CARD11*, and *BTK* [18] (Fig. S2C). Therefore, the mutated genes might participate in PCNSL pathogenesis by modulating these pathways.

Differential mutation profiles of Chinese PCNSLs

The mutated genes identified in the present cohort using WGS were compared with those identified in previous studies that used WES, including the articles of Vater *et al*. [19], Bruno *et al*. [20], and Fukumura *et al*. [9]. A total of 195 genes were exclusively found in the present cohort compared with the above three cohorts (Fig. 4A). After integrating the genes exclusively detected in the present cohort with other 3 reports [8,10,21], 138 genes were retained (Fig. 4B), of which 6 (*ROBO2*, *KMT2C*, *CXCR4*, *MYOM2*, *BCLAF1*, and *NRXN3*) were detected in ⩾ 10% of patients in the present cohort. By comparing the mutation profile of the present cohort with that of 2 other Chinese cohorts [10,22], mutations in 72 genes were exclusively detected in the present cohort, and 28 genes were commonly mutated in the 3 Chinese cohorts (Fig. 4C). The mutation frequencies of *PIM1*, *CD79B*, *KMT2D*, *IRF4*, *CARD11*, *PRDM1*, *GNA13*, *CREBBP*, and *TP53* were ⩾ 10% in the present cohort.

The mutation landscapes detected in Asian patients (including 3 Chinese cohorts [10,22] and 2 Japanese cohorts [9,21]) and non-Asian patients (including 1 American cohort [20] and 1 German cohort [19]) with PCNSL were then compared. A total of 406 genes were exclusively detected among Asian PCNSL patients

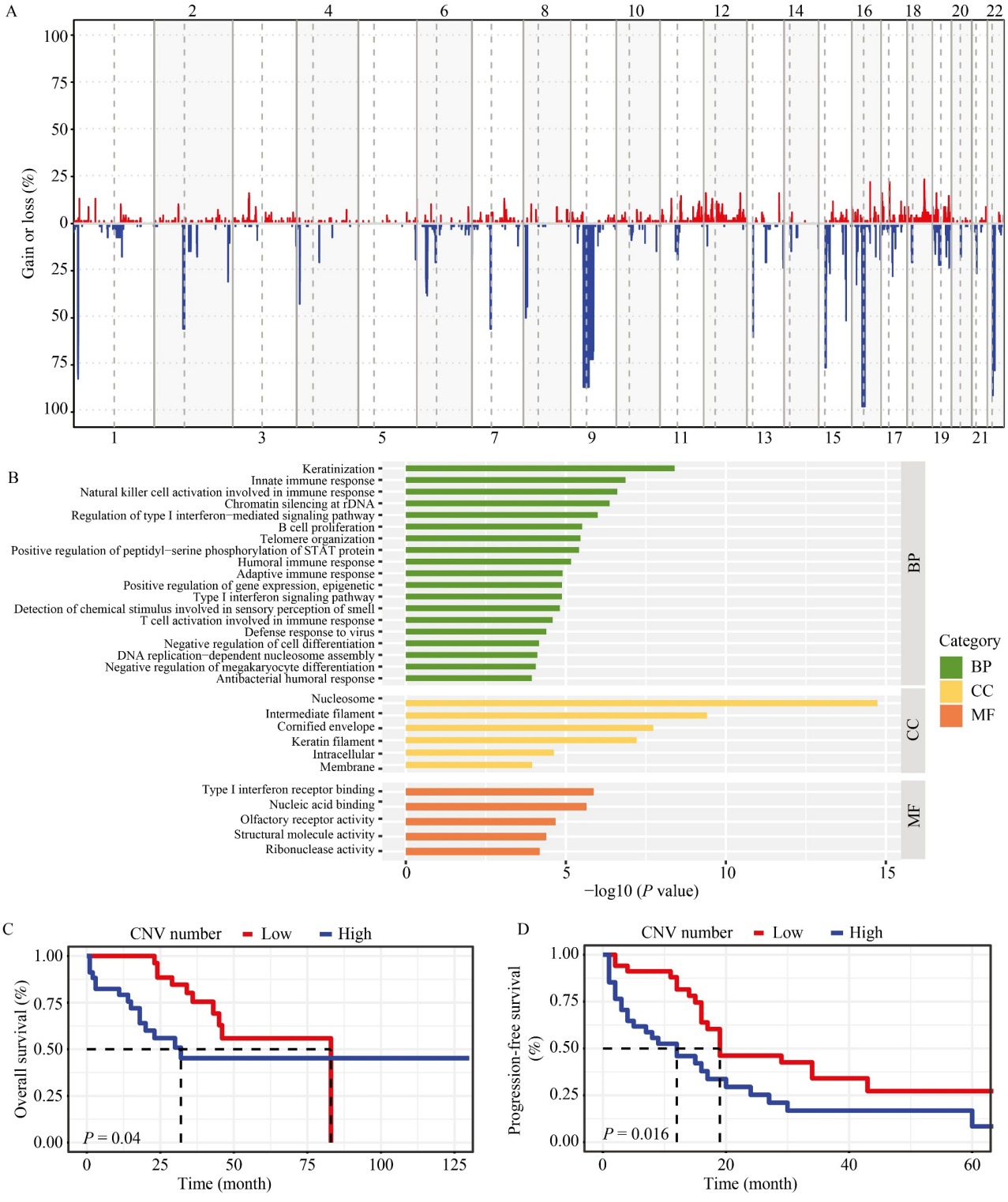

Fig. 2 Copy number variation (CNV) profile and clinical significances of the 68 PCNSL patients. (A) Frequency plot of CNVs in the 68 samples. Chromosomes 1 to 22 are represented from left to right. Red blocks represent chromosome gains, whereas blue blocks represent chromosome losses. The amplitude in each abnormal region represents the incidence of each copy-number abnormality in the studied cohort. (B) GO pathway analysis of the 7791 genes related to the CNV identified in the 68 PCNSL samples (CC, cellular component; MF, molecular functions; BP, biological process). KM curves were used to analyze the relationship of CNV number with the (C) overall survival and (D) progression-free survival of the 68 patients with PCNSL.

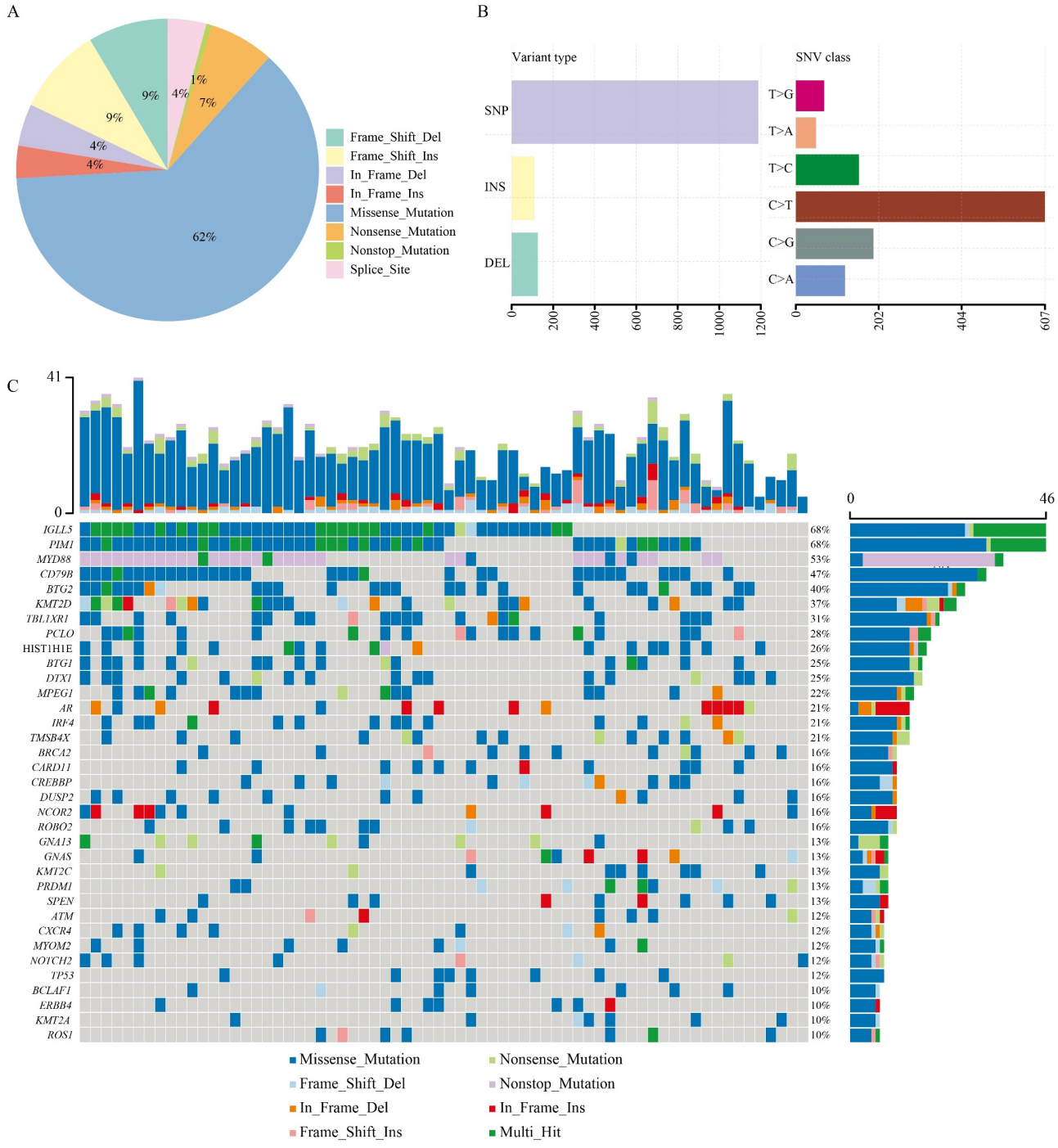

Fig. 3 Coding mutated genes identified by WGS in the 68 PCNSL samples. (A) Pie chart showing the percentages of different types of mutations in PCNSL. (B) Number of single-nucleotide polymorphisms (SNPs), insertions (INSs), and deletions (DELs). (C) Top 35 mutated genes found in the 68 PCNSL samples.

(Fig. 4D), with 38 co-mutated genes identified in both the Japanese and Chinese cohorts, among which *KMT2A*, *AR*, and *ROS1* had a mutation frequency of ⩾ 10% in the present cohort.

The mutated genes in the present cohort of PCNSLs were also compared with those in three reports targeting the system DLBCL [23–25] (Fig. 4E). A total of 19 genes (*RAG1*, *FGF4*, *MYCN*, *BCORL1*, *TGM7*, *HSD3B1*, *MTAP*, *ESR1*, *GATA4*, *MPL*, *FGF12*, *FOXL2*, *MAP3K13*, *FGF6*, *FGF10*, *FGF3*, *FGF14*, *SOX2*, and *FGF19*) were found exclusively in the present cohort. These 19 genes were then compared with other common brain tumors,

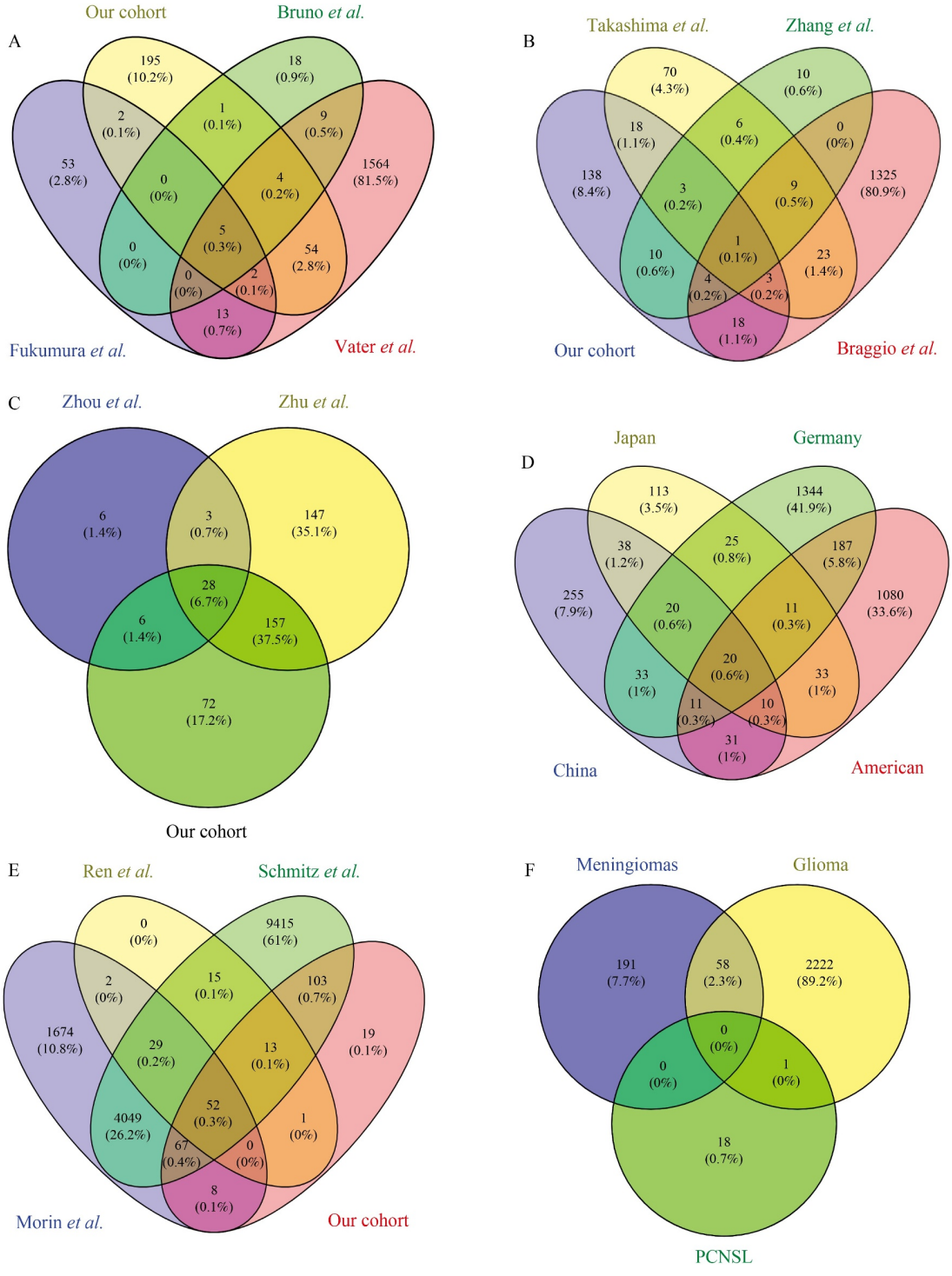

Fig. 4 Differential mutation profiles of Chinese PCNSLs. (A, B) Mutated genes in the present cohort and six published PCNSL cohorts. (C) Mutated genes in the present cohort and other three published Chinese PCNSL cohorts. (D) Mutated genes in Asian and non-Asian cohorts with PCNSL. (E) Mutated genes in the present cohort and three published system DLBCL cohorts. (F) Mutated genes detected in the present cohort, the published system DLBCL cohorts, glioma, and meningioma.

such as glioma [26] and meningioma [27,28]. The mutation of only 1 of these genes (*BCORL1*) was detected in glioma (Fig. 4F), thereby suggesting that the remaining 18 genes (*RAG1*, *FGF4*, *MYCN*, *TGM7*, *HSD3B1*, *MTAP*, *ESR1*, *GATA4*, *MPL*, *FGF12*, *FOXL2*, *MAP3K13*, *FGF6*, *FGF10*, *FGF3*, *FGF14*, *SOX2*, and *FGF19*) might be used for the differential diagnosis of PCNSL from system DLBCL, glioma, and meningiomas.

Mutation profiles associated with patients' clinical features

The possible link between the mutated genes and the clinical features of PCNSL patients was then tested. The mutation frequencies of *HIST1H1E*, *CARD11*, and *ROBO2* in the GCB group were significantly higher than those in the non-GCB group, while the mutation frequencies of *MPEG1*, *DUSP2*, and *PRDM1* in the GCB group were significantly lower than those in the non-GCB group (Fig. 5A). The mutation frequencies of *PCLO*, *CREBBP*, *PRDM1*, *ATM*, and *NOTCH2* showed significant differences among the high-, middle-, and low-risk groups based on their IELSG scores (Fig. 5B). In risk groups defined based on MSKCC scores, the *BRCA2*, *CREBBP*, and *ERBB4* mutation frequencies showed noticeable differences across different risk groups (Fig. 5C). In addition, the mutation frequencies of *MPEG1* and *AR* in the multifocal group were significantly higher than those in the unifocal group (Fig. 5D). The mutation frequencies of *PCLO* and *AR* were higher among patients with deep involvement than among those patients without deep involvement (Fig. 5E). *IRF4* mutation was less common among patients with double expression of both BCL2 and MYC, while *TMSB4X* and *BTG1* mutations were more common among these patients (Fig. 5F). An increased AR mutation frequency was significantly related to a lower lymphocyte count (⩽ 2.0) (Fig. 5G). *KMT2D* mutation frequency was positively associated with LDH level (Fig. 5H). Additionally, *KMT2A* mutation was only detected in the KPS < 70 group (Fig. 5I). However, no significant differences in gene mutation frequency were observed based on age and ECOG score. These results linked the gene mutation profiles to the clinical features of PCNSL patients.

Mutation profiles associated with patients' treatment response and prognosis

The relationship between the mutation profiles and the patients' treatment response and progression was then analyzed. The *CARD11* mutation was detected in the response group (CR + PR) (21.3%) but not in the non-response group (SD + PD) (P = 0.0379) (Fig. 6A), whereas the *KMT2A* mutation frequency was lower in patients with two-year progression (4.1%) than in those without progression (26.3%) (P = 0.016) (Fig. 6B). Therefore, *CARD11* mutation might be a potent marker for treatment response, and *KMT2A* mutation could predict the two-year progression of PCNSL.

The relationship between the mutation profiles and patients' prognosis was then examined. A total of 36 genes with a mutation frequency of ⩾ 10% were evaluated using K–M curves to understand their effects on the OS and PFS of PCNSL patients. Only *TMSB4X* mutation was significantly associated with lower OS compared with the wild-type *TMSB4X* group (Fig. 6D), while *CD79B* mutation was significantly associated with lower PFS (Fig. 6E). The mutation frequencies of 10 genes, namely, *PCLO*, *PAX5*, *ZNF217*, *MSH6*, *EZH2*, *B2M*, *SOCS1*, *TNFAIP3*, *CDKN1B*, and *BCL6*, showed significant differences among patients with and without *TMSB4X* mutation (Fig. 6C).

High expression of TMSB4X protein associated with poor prognosis

The expression of *TMSB4X* and its association with the prognosis of systemic DLBCL patients was initially examined. The expression of *TMSB4X* was higher among DLBCL patients than those with other common tumors both at the transcriptional level as shown by the GEPIA database (Fig. 7A) and the translation level as retrieved from the ProteinAtlas database (Fig. 7B). The *TMSB4X* transcriptional level of DLBCL in the GEPIA database was higher than that in the normal control group (Fig. 7C). According to the GEO database (GSE11392 cohort), the *TMSB4X* transcriptional level in PCNSL patients was obviously higher than that of nodal DLBCL (Fig. 7D), but no significant difference was found between PCNSL patients and extranodal DLBCL (Fig. 7E). To reveal the effect of *TMSB4X* mutation on its expression in PCNSL, the TMSB4X protein expression levels in PCNSL patients carrying mutated *TMSB4X* and wild-type *TMSB4X* were compared. Figure 7F shows the mutation schematic diagram of the *TMSB4X* gene in its domain region. The expression of TMSB4X protein in patients carrying *TMSB4X* mutation was slightly higher than that in patients carrying the wild-type *TMSB4X* (Fig. 7G). Typical IHC images of the TMSB4X protein were shown in Fig. 7H and 7I. On the basis of the average integrated option density (AOD) of TMSB4X staining, the 68 PCNSL patients were divided into the low *TMSB4X* expression group (AOD ⩽ 0.08) and high-risk (AOD > 0.08) group according to the cutoff value determined by the ROC curve. High TMSB4X expression was associated with lower OS in PCNSL patients compared with the low TMSB4X group (Fig. 7J).

Establishment of a new prognostic risk scoring system for PCNSL

A prognostic model was built to further clarify the clinical value of gene mutation profiles in PCNSL. Lasso regression and 10-fold cross-validation were used to select risk factors among the 263 mutated genes and 4 clinicopathologic features (age, KPS, deep involvement, and LDH level) (Fig. 8A). Following the training and validation, KPS and 6 genes (*BRD4*, *EBF1*, *BTG1*, *CCND3*, *STAG2*, and *TMSB4X*) were incorporated into

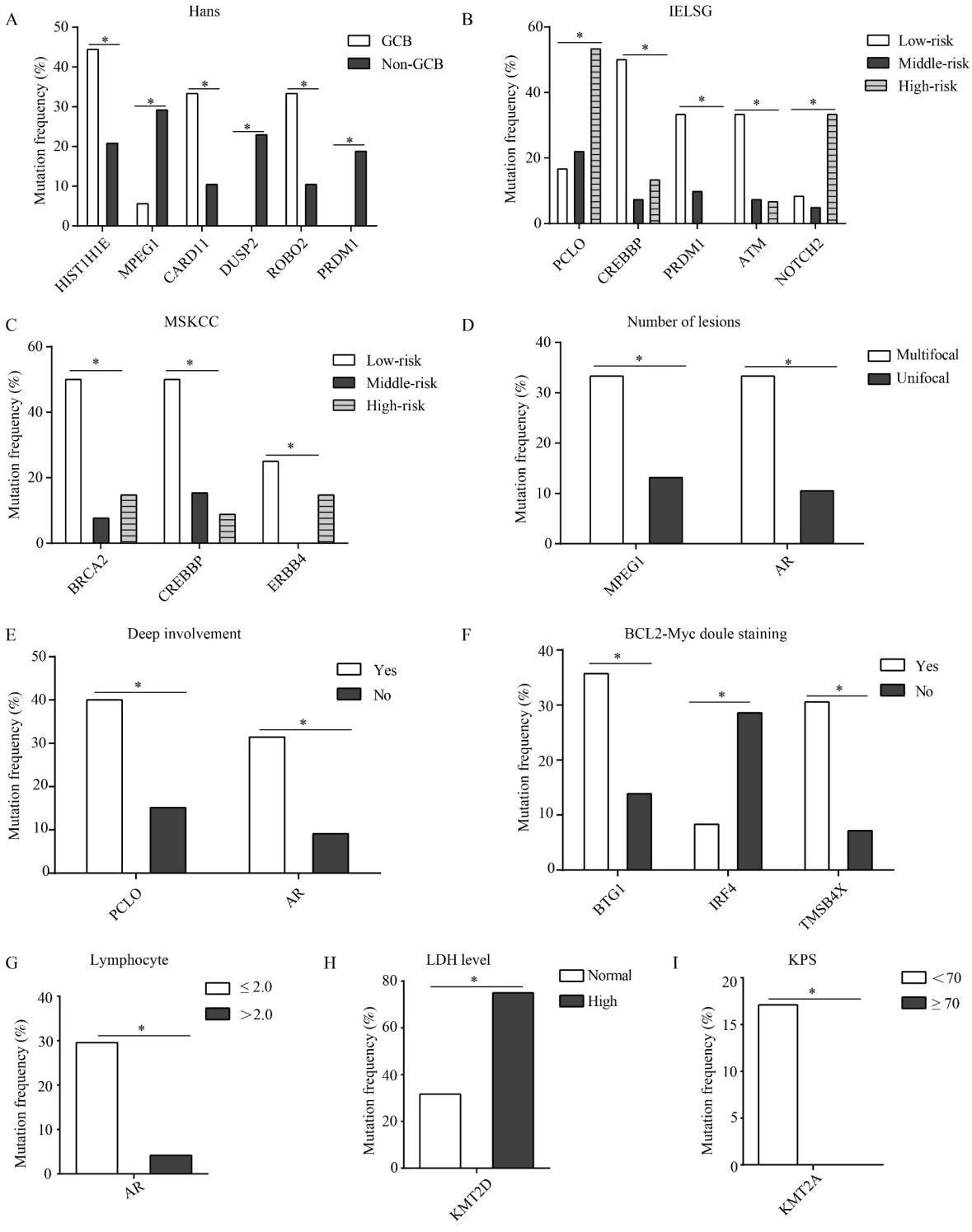

Fig. 5 Gene mutations are associated with patients' clinicopathologic features. The mutation frequencies of genes with significant differences between various items were determined. (A) Hans' classifications. (B) IELSG score (low-, middle-, and high-risk). (C) MSKCC score (low-, middle-, and high-risk). (D) Number of lesions. (E) Deep involvement. (F) BCL2 and MYC double expression. (G) Lymphocyte count. (H) LDH level. (I) KPS. **P* < 0.05.

the prognostic risk scoring system, which was validated by ROC analysis (AUC, 0.8182; 95% CI, 0.6356–0.9854; kappa, 0.7213) (Fig. 8B). The probability of high risk was calculated as follows = 1/[1+exp(–0.37993+0.1624KPS +0.63449*BRD4*+0.52019*EBF1*–0.41161*BTG1*+ 0.04921*CCND3*+0.23689*STAG2*–0.61347*TMSB4X*)]. The test set was divided into the low-risk (≤ 0.504) and high-risk (> 0.504) groups according to the cutoff value

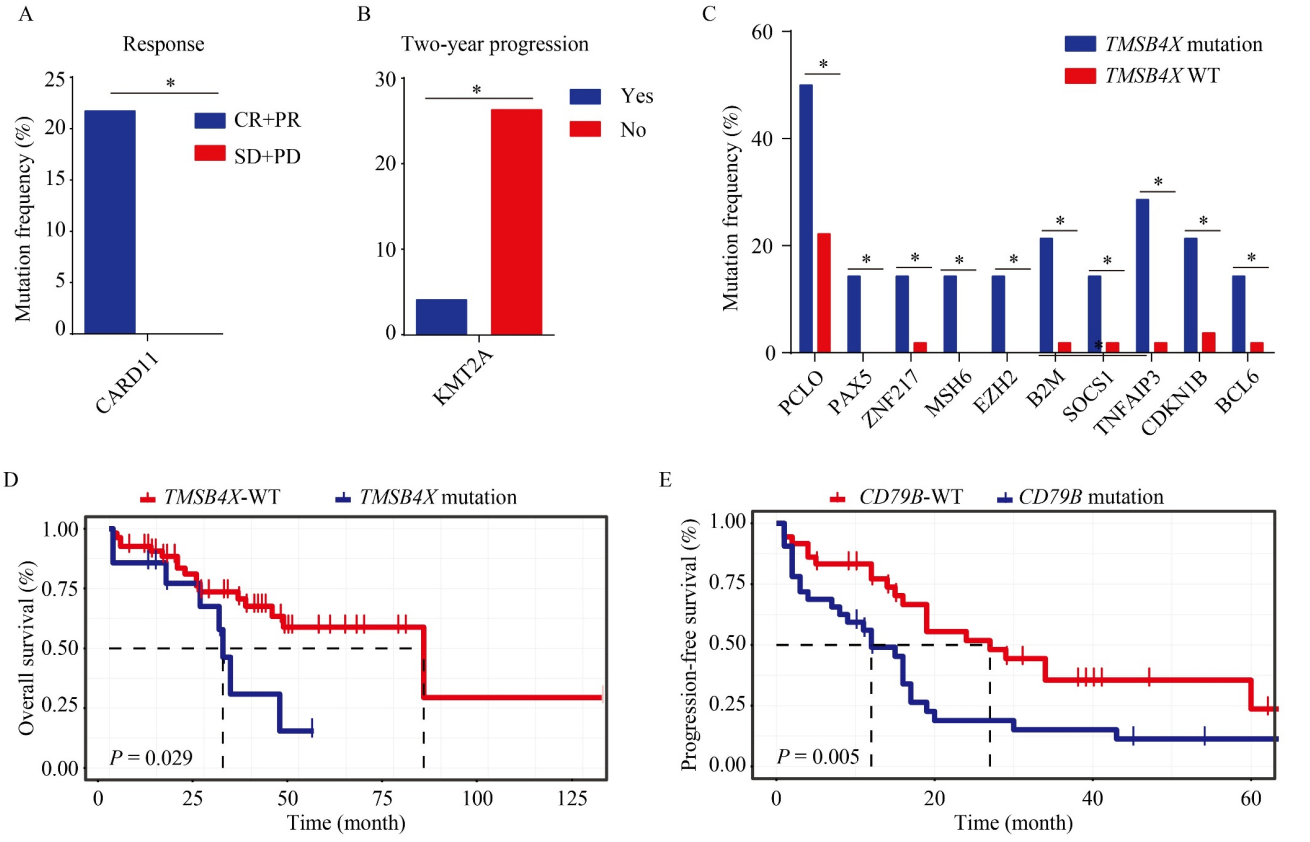

Fig. 6 Mutation profiles are associated with the chemotherapy response and prognosis of PCNSL patients. The significantly mutated genes in the (A) chemotherapy response (CR + PR) group and non-response group (SD + PD) and in (B) patients with or without two-year progression. (C) The mutated genes with significantly different mutation frequencies among patients with or without *TMSB4X* mutation. (D) K–M curves showing that *TMSB4X* mutation was linked to lower overall survival among patients with PCNSL. (E) K–M curves showing that *CD79B* mutation was linked to lower progression-free survival among patients with PCNSL. **P* < 0.05; WT, wild type.

determined by the ROC curve. The K–M curve showed that the patients in the high-risk group had significantly lower OS than those in the low-risk group (Fig. 8C). Subsequently, KPS and the above six genes were applied to construct a nomogram model for predicting the 1-, 2-, and 3-year OS rates of PCNSL patients (Fig. 8D). The mutations of three of the six genes (*BTG1*, *STAG2*, and *BRD4*) showed interactions with other high-frequency mutations (Fig. S3).

The clinical value of the IELSG [29] and MSKCC scoring systems [30], which are two prognostic scoring systems developed specifically for PCNSL, was also assessed. The IELSG system showed a higher prognosis value in both OS and PFS (Fig. S4A and S4B) compared with the MSKCC system (Fig. S4C and S4D) but was only confined to predicting the low-risk group.

Discussion

With the development of precision medicine, the genetic features underlying PCNSL have received increasing attention. Next-generation sequencing provides a path to explore the genetic profiles of PCNSL but is far from sufficient. Recently, Radke *et al*. [31] from Germany compared the genomic and transcriptional landscapes of 51 CNS lymphomas (42 cases of PCNSL, 6 cases of secondary CNS lymphomas, and 3 cases of EBV + CNS lymphomas) with those of 39 follicular lymphoma and 36 DLBCL cases outside the CNS via WGS and RNA sequencing and concluded that PCNSL can be clearly distinguished from DLBCL due to its distinct expression profiles and IG expression and translocation patterns. In this study, the genetic landscape of 68 Chinese patients with PCNSL was revealed using WGS, which can completely profile the status of the genome. To the best of the authors' knowledge, this study is the largest population study of PCNSL in China.

Mutations in the *IGLL5*, *PIM1*, *MYD88*, *CD79B*, *BTG2*, *KMT2D*, *TBL1XR1*, *PCLO*, *HIST1H1E*, *BTG1*, *DTX1*, and *CARD11* genes were the most common coding region variations in the present cohort. In addition to the common mutations previously identified in PCNSL [9,15,32], the mutations in the *ROBO2*, *KMT2C*, *CXCR4*, *MYOM2*, *BCLAF1*, and *NRXN3* genes were exclusively detected in the present cohort with a ⩾ 10% frequency rate, thus highlighting the high heterogeneity and

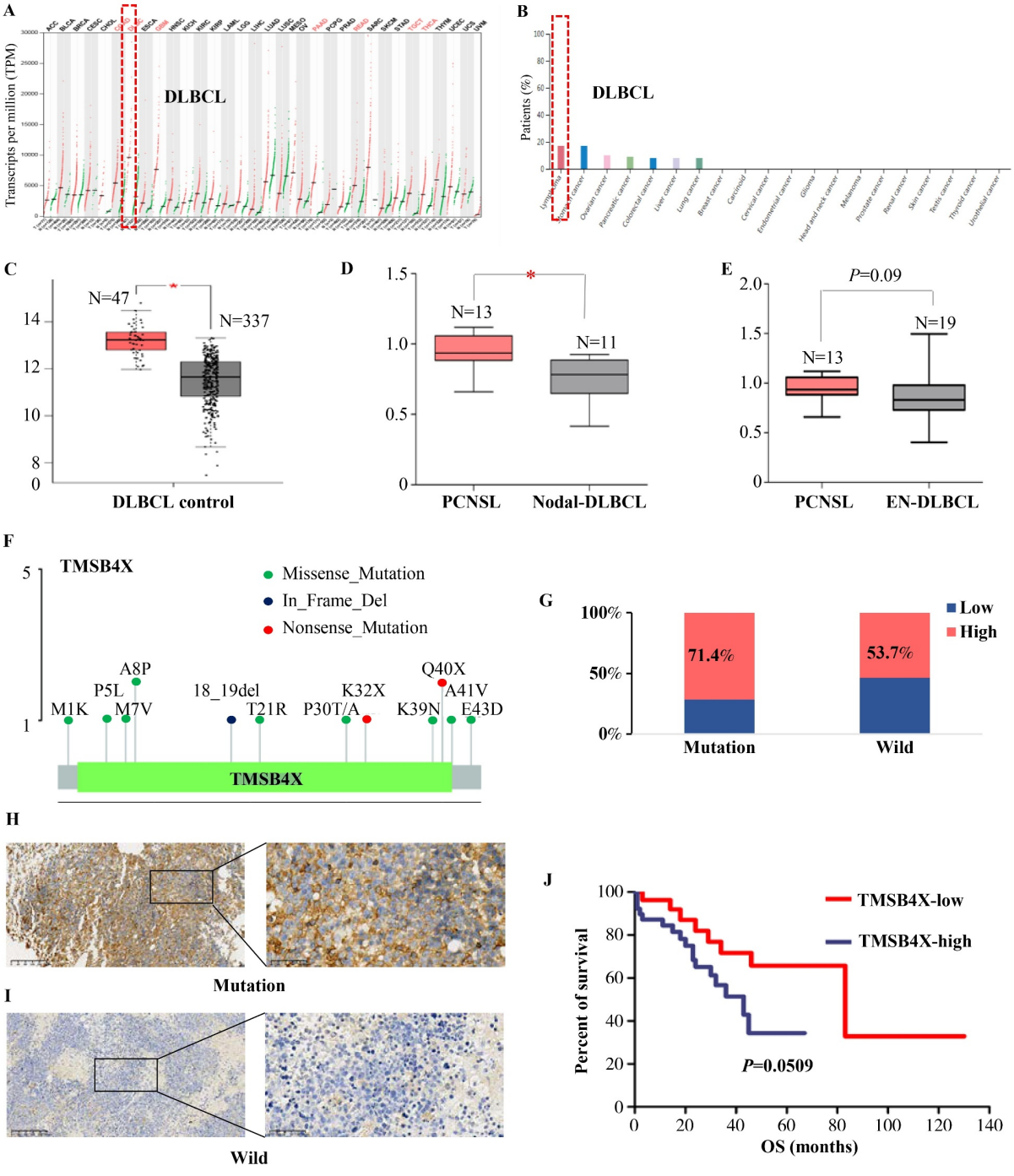

Fig. 7 High expression of TMSB4X protein was associated with poor prognosis. (A) GEPIA database showing *TMSB4X* transcriptional levels in DLBCL and other kinds of cancers. (B) ProteinAtlas database showing TMSB4X protein expression levels in DLBCL and other kinds of cancers. (C) The transcriptional level of *TMSB4X* in DLBCL and normal control group as shown in the GEPIA database. GEO database (GSE11392 cohort) showing the transcriptional levels of *TMSB4X* in PCNSL and nodal DLBCL patients (D) and in PCNSL and extranodal DLBCL patients (E). (F) TMSB4X amino acid mutation sites in PCNSL patients. (G) Patients with *TMSB4X* mutations had a slightly higher proportion of high TMSB4X protein expression. (H) Typical immunohistochemistry (IHC) images of TMSB4X expression in patients carrying *TMSB4X* mutations. (I) Typical IHC images of TMSB4X protein expression in patients carrying wild-type *TMSB4X*. (J) K–M curves showing that the high expression of TMSB4X was associated with lower OS in PCNSL patients.

complexity of PCNSL. The mutation landscapes detected in Asian patients (3 Chinese cohorts [10,22] and 2 Japanese cohorts [9,21]) and non-Asian patients (1 American cohort [20] and 1 German cohort [19]) with PCNSL were also compared. A total of 38 co-mutated genes were identified in the Japanese and Chinese

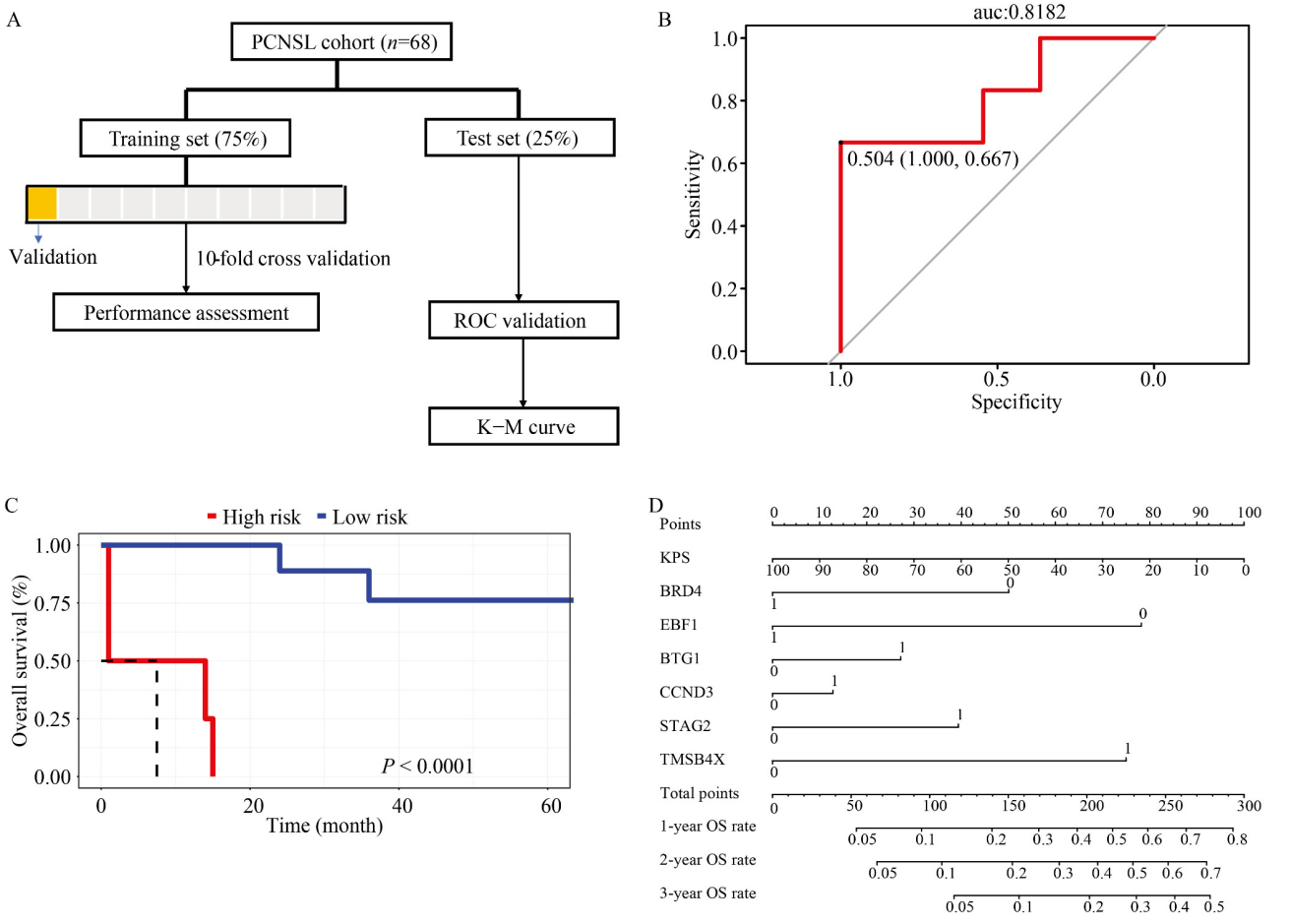

Fig. 8 New prognostic model for predicting the prognosis of PCNSL. (A) Flow chart for the establishment of a prognostic risk scoring system for PCNSL. (B) ROC and (C) K–M curve analyses for testing the prognostic model. (D) The nomogram model, which includes a combination of Karnofsky performance status (KPS) and six mutated genes, and the overall survival rate of PCNSL patients. After the nomogram assigned a score for the KPS plus mutations of the *BRD4*, *EBF1*, *BTG1*, *CCND3*, *STAG2*, and *TMSB4X* genes for each patient, the total scores were obtained and used to predict the overall survival rates of patients.

cohorts, among which *KMT2A*, *AR*, and *ROS1* had a mutation frequency of ⩾ 10% in the present cohort. The mutated genes detected in PCNSL, DLBCL, glioma and meningioma were also compared. A total of 18 genes, namely, *RAG1*, *FGF4*, *MYCN*, *TGM7*, *HSD3B1*, *MTAP*, *ESR1*, *GATA4*, *MPL*, *FGF12*, *FOXL2*, *MAP3K13*, *FGF6*, *FGF10*, *FGF3*, *FGF14*, *SOX2*, and *FGF19*, were detected only in PCNSL, of which 6 (*FGF3*/*6*/*10*/*12*/*14*/*19*) belonged to the fibroblast growth factors (FGFs) family, which contains > 20 members that encode secreted polypeptides and act through tyrosine kinase receptors [33]. Accumulated evidence shows that FGFs are involved in the progression of multiple types of tumors, such as non-small cell lung cancer (NSCLC) [34], ovarian cancer [35], and pancreatic cancer [36]. Noticeably, targeting FGF signaling is a promising therapeutic strategy [37,38]. Results indicated that the mutations in the genes of the FGF family might be used for the differential diagnosis of PCNSL from DLBCL and other brain tumors.

Evidence also shows that CNVs and mutation profiles are associated with the prognosis of many kinds of cancers [39,40]. In this study, the *CD79B* mutation and a high level of CNV were significantly associated with the lower PFS of PCNSL. Similarly, Zhou *et al*. [10] revealed that the *CD79B* mutation was related to inferior PFS among PCNSL patients. In addition, the *TMSB4X* mutation and a high level of CNV were significantly associated with lower OS. *TMSB4X* encodes a 43-amino-acid short peptide implicated in multiple biological activities, including inflammation, cell survival, and metastasis [41,42]. The mutation of *TMSB4X* was also previously found to be a recurrent variation in PCNSL [19] and a target of aberrant somatic hypermutation in systemic DLBCL [43]. However, whether or not *TMSB4X* affects the prognosis of PCNSL remains unknown. Therefore, the role of *TMSB4X* in PCNSL warrants further research.

The IELSG [29] and MSKCC scoring systems [30] are two prognostic scoring systems developed specifically for

PCNSL. However, these scoring systems do not include genetic profiles. In the present cohort, the IELSG system showed a higher prognosis value than the MSKCC system but was only confined to predicting the low-risk group. In this study, a prognostic model combining KPS with six mutated genes (*BRD4*, *EBF1*, *BTG1*, *CCND3*, *STAG2,* and *TMSB4X*) was developed. Among these genes, *EBF1* (early B cell factor 1) and *BTG1* are tumor suppressors [44,45]. *EBF1* belongs to a family of four highly conserved DNA binding transcription factors that are involved in the differentiation and maturation of B-progenitor lymphoblasts. Specifically, *EBF1* is required for the expression of *PAX5* (paired box 5) during B cell development, and the deletion of *EBF1* leads to the pathogenesis, drug resistance, and relapse of B-progenitor acute lymphoblastic leukemia (ALL) [45]. Hodkinson *et al*. [46] recently revealed that *EBF1* mutation was not detected in responders (ibrutinib plus nivolumab) in relapsed DLBCL, follicular lymphoma, or Richter's transformation patients, thereby suggesting that *EBF*1 mutation may be linked to poor response to ibrutinib and nivolumab. *CCND3* is a potential driver of Burkitt lymphomagenesis [47] and mantle cell lymphoma (MCL) [48], a treatment-refractory subtype of malignant lymphoma. *BRD4* (bromodomain containing 4) is a transcriptional and epigenetic regulator that is preferentially located at the super-enhancer regions of several crucial enhances, such as *C-MYC*, *BCL-XL*, and *BCL-6*, and can lead to cancer development [49,50]. *I-BET151*, a BRD4 inhibitor, has been reported to suppress cell proliferation in MCL. *BRD4* can target genes involved in the BCR signaling pathway, including *BLNK* (B cell linker), *PAX5*, and *IKZF3* (IKAROS family zinc finger 3) [51], thereby suggesting that *BRD4* may be a potential treatment target in lymphoma. The *STAG2* gene encodes a cohesion complex subunit that regulates chromatid separation, mutations of which can result in enhanced chromosomal stability and be detected in acute leukemias [52] and adult T cell lymphoma [53]. Results of this work might provide a new molecular classification method for predicting the prognosis of PCNSL, which needs to be further verified in clinical practice.

This study has several limitations, one of which lies in the lack of ideal normal controls (such as paraneoplastic normal brain tissue) and the use of only 8 paired bone marrow tissues as negative controls. To minimize the error induced by this small negative control size, the mutation sites and CNVs were filtered by GRCh37 and/or the GenomAD, ExAC, and 1000 Genomes Project databases. Moreover, FFPE samples were used in this study. The DNA extracted from these samples is problematic for mutation testing, especially for amplicon-based massively parallel sequencing, owing to DNA fragmentation and the artificial C:G > T:A SNVs caused by the deamination of cytosine to uracil. However, the FFPE samples used in this study were pretreated with uracil DNA glycosylase, which can eliminate uracil-containing DNA molecules [54]. Moreover, the PPFE samples were submitted to WGS and not to amplicon-based massively parallel sequencing, and the proportion of C > T SNVs (51%) detected in the present cohort is no greater than that detected in a previous report (58%) [55].

In conclusion, this study further describes the genomic landscape of Chinese PCNSL using WGS, including SVs, CNVs and gene mutations in PCNSL, and explores their clinical values. The *TMSB4X* mutation and a high level of CNV are significantly associated with lower OS, while the *CD79B* mutation and a high level of CNV are significantly associated with lower PFS. A prognostic model containing KPS and six mutated genes was also built, by which the 1-, 2-, and 3-year OS of PCNSL patients can be predicted accurately and conveniently. This study depicts the genomic characteristics of Chinese PCNSLs, thereby enriching the present understanding of the genetic pathogenesis of PCNSL.

Acknowledgments

We thank Shanghai Yuanqi Biomedical Technology Co. Ltd. (Shanghai, China) for the bioinformatics analysis. This study was supported by funds from the Translational Research Grant of National Clinical Research Center for Hematologic Disease (No. 2020ZKZC01), the National Natural Science Foundation of China (Nos. 81830006, 82170219, and 81800188) and the Lymphoma Research Fund of China Anti-Cancer Association.

Compliance with ethics guidelines

Xianggui Yuan, Teng Yu, Jianzhi Zhao, Huawei Jiang, Yuanyuan Hao, Wen Lei, Yun Liang, Baizhou Li, and Wenbin Qian declare that they have no conflict of interest. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the *Helsinki Declaration* of 1975, as revised in 2000. Informed consent was obtained from all patients for being included in the study.

Electronic Supplementary Material Supplementary material is available in the online version of this article at https://doi.org/10.1007/s11684-023-0994-x and is accessible for authorized users.

References

1. Grommes C, DeAngelis LM. Primary CNS lymphoma. J Clin Oncol 2017; 35(21): 2410–2418
2. Baraniskin A, Schroers R. Liquid biopsy and other non-invasive diagnostic measures in PCNSL. Cancers (Basel) 2021; 13(11): 2665

3. van der Meulen M, Dinmohamed AG, Visser O, Doorduijn JK, Bromberg JEC. Improved survival in primary central nervous system lymphoma up to age 70 only: a population-based study on incidence, primary treatment and survival in the Netherlands, 1989–2015. Leukemia 2017; 31(8): 1822–1825
4. Kim P, Omuro A. Consolidation therapy in primary central nervous system lymphoma. Curr Treat Options Oncol 2020; 21(9): 74
5. Courts C, Montesinos-Rongen M, Brunn A, Bug S, Siemer D, Hans V, Blümcke I, Klapper W, Schaller C, Wiestler OD, Küppers R, Siebert R, Deckert M. Recurrent inactivation of the *PRDM1* gene in primary central nervous system lymphoma. J Neuropathol Exp Neurol 2008; 67(7): 720–727
6. Montesinos-Rongen M, Schmitz R, Brunn A, Gesk S, Richter J, Hong K, Wiestler OD, Siebert R, Küppers R, Deckert M. Mutations of CARD11 but not TNFAIP3 may activate the NF-kB pathway in primary CNS lymphoma. Acta Neuropathol 2010; 120(4): 529–535
7. Montesinos-Rongen M, Zühlke-Jenisch R, Gesk S, Martín-Subero JI, Schaller C, Van Roost D, Wiestler OD, Deckert M, Siebert R. Interphase cytogenetic analysis of lymphoma-associated chromosomal breakpoints in primary diffuse large B-cell lymphomas of the central nervous system. J Neuropathol Exp Neurol 2002; 61(10): 926–933
8. Braggio E, Van Wier S, Ojha J, McPhail E, Asmann YW, Egan J, da Silva JA, Schiff D, Lopes MB, Decker PA, Valdez R, Tibes R, Eckloff B, Witzig TE, Stewart AK, Fonseca R, O'Neill BP. Genome-wide analysis uncovers novel recurrent alterations in primary central nervous system lymphomas. Clin Cancer Res 2015; 21(17): 3986–3994
9. Fukumura K, Kawazu M, Kojima S, Ueno T, Sai E, Soda M, Ueda H, Yasuda T, Yamaguchi H, Lee J, Shishido-Hara Y, Sasaki A, Shirahata M, Mishima K, Ichimura K, Mukasa A, Narita Y, Saito N, Aburatani H, Nishikawa R, Nagane M, Mano H. Genomic characterization of primary central nervous system lymphoma. Acta Neuropathol 2016; 131(6): 865–875
10. Zhou Y, Liu W, Xu Z, Zhu H, Xiao D, Su W, Zeng R, Feng Y, Duan Y, Zhou J, Zhong M. Analysis of genomic alteration in primary central nervous system lymphoma and the expression of some related genes. Neoplasia 2018; 20(10): 1059–1069
11. Wang PP, Liu SH, Chen CT, Lv L, Li D, Liu QY, Liu GL, Wu Y. Circulating tumor cells as a new predictive and prognostic factor in patients with small cell lung cancer. J Cancer 2020; 11(8): 2113–2122
12. Yang L, Luquette LJ, Gehlenborg N, Xi R, Haseley PS, Hsieh CH, Zhang C, Ren X, Protopopov A, Chin L, Kucherlapati R, Lee C, Park PJ. Diverse mechanisms of somatic structural variations in human cancer genomes. Cell 2013; 153(4): 919–929
13. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, Johns AL, Miller D, Nones K, Quek K, Quinn MC, Robertson AJ, Fadlullah MZ, Bruxner TJ, Christ AN, Harliwong I, Idrisoglu S, Manning S, Nourse C, Nourbakhsh E, Wani S, Wilson PJ, Markham E, Cloonan N, Anderson MJ, Fink JL, Holmes O, Kazakoff SH, Leonard C, Newell F, Poudel B, Song S, Taylor D, Waddell N, Wood S, Xu Q, Wu J, Pinese M, Cowley MJ, Lee HC, Jones MD, Nagrial AM, Humphris J, Chantrill LA, Chin V, Steinmann AM, Mawson A, Humphrey ES, Colvin EK, Chou A, Scarlett CJ, Pinho AV, Giry-Laterriere M, Rooman I, Samra JS, Kench JG, Pettitt JA, Merrett ND, Toon C, Epari K, Nguyen NQ, Barbour A, Zeps N, Jamieson NB, Graham JS, Niclou SP, Bjerkvig R, Grützmann R, Aust D, Hruban RH, Maitra A, Iacobuzio-Donahue CA, Wolfgang CL, Morgan RA, Lawlor RT, Corbo V, Bassi C, Falconi M, Zamboni G, Tortora G, Tempero MA; Australian Pancreatic Cancer Genome Initiative; Gill AJ, Eshleman JR, Pilarsky C, Scarpa A, Musgrove EA, Pearson JV, Biankin AV, Grimmond SM. Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 2015; 518(7540): 495–501
14. Wu S, Ou T, Xing N, Lu J, Wan S, Wang C, Zhang X, Yang F, Huang Y, Cai Z. Whole-genome sequencing identifies ADGRG6 enhancer mutations and FRS2 duplications as angiogenesis-related drivers in bladder cancer. Nat Commun 2019; 10(1): 720
15. Chapuy B, Roemer MGM, Stewart C, Tan Y, Abo RP, Zhang L, Dunford AJ, Meredith DM, Thorner AR, Jordanova ES, Liu G, Feuerhake F, Ducar MD, Illerhaus G, Gusenleitner D, Linden EA, Sun HH, Homer H, Aono M, Pinkus GS, Ligon AH, Ligon KL, Ferry JA, Freeman GJ, van Hummelen P, Golub TR, Getz G, Rodig SJ, de Jong D, Monti S, Shipp MA. Targetable genetic features of primary testicular and primary central nervous system lymphomas. Blood 2016; 127(7): 869–881
16. Braggio E, McPhail ER, Macon W, Lopes MB, Schiff D, Law M, Fink S, Sprau D, Giannini C, Dogan A, Fonseca R, O'Neill BP. Primary central nervous system lymphomas: a validation study of array-based comparative genomic hybridization in formalin-fixed paraffin-embedded tumor specimens. Clin Cancer Res 2011; 17(13): 4245–4253
17. Booman M, Szuhai K, Rosenwald A, Hartmann E, Kluin-Nelemans H, de Jong D, Schuuring E, Kluin P. 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 2008; 216(2): 209–217
18. Krysiak K, Gomez F, White BS, Matlock M, Miller CA, Trani L, Fronick CC, Fulton RS, Kreisel F, Cashen AF, Carson KR, Berrien-Elliott MM, Bartlett NL, Griffith M, Griffith OL, Fehniger TA. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. Blood 2017; 129(4): 473–483
19. Vater I, Montesinos-Rongen M, Schlesner M, Haake A, Purschke F, Sprute R, Mettenmeyer N, Nazzal I, Nagel I, Gutwein J, Richter J, Buchhalter I, Russell RB, Wiestler OD, Eils R, Deckert M, Siebert R. The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. Leukemia 2015; 29(3): 677–685
20. Bruno A, Boisselier B, Labreche K, Marie Y, Polivka M, Jouvet A, Adam C, Figarella-Branger D, Miquel C, Eimer S, Houillier C, Soussain C, Mokhtari K, Daveau R, Hoang-Xuan K. Mutational analysis of primary central nervous system lymphoma. Oncotarget 2014; 5(13): 5065–5075
21. Takashima Y, Sasaki Y, Hayano A, Homma J, Fukai J, Iwadate Y, Kajiwara K, Ishizawa S, Hondoh H, Tokino T, Yamanaka R. Target amplicon exome-sequencing identifies promising diagnosis and prognostic markers involved in RTK-RAS and PI3K-AKT signaling as central oncopathways in primary central nervous system lymphoma. Oncotarget 2018; 9(44): 27471–27486
22. Zhu Q, Wang J, Zhang W, Zhu W, Wu Z, Chen Y, Chen M, Zheng

L, Tang J, Zhang S, Wang D, Wang X, Chen G. Whole-genome/exome sequencing uncovers mutations and copy number variations in primary diffuse large B-cell lymphoma of the central nervous system. Front Genet 2022; 13: 878618

23. Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD, Scott DW, Ding J, Roth A, Chiu R, Corbett RD, Chan FC, Mendez-Lago M, Trinh DL, Bolger-Munro M, Taylor G, Hadj Khodabakhshi A, Ben-Neriah S, Pon J, Meissner B, Woolcock B, Farnoud N, Rogic S, Lim EL, Johnson NA, Shah S, Jones S, Steidl C, Holt R, Birol I, Moore R, Connors JM, Gascoyne RD, Marra MA. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood 2013; 122(7): 1256–1265
24. Ren W, Ye X, Su H, Li W, Liu D, Pirmoradian M, Wang X, Zhang B, Zhang Q, Chen L, Nie M, Liu Y, Meng B, Huang H, Jiang W, Zeng Y, Li W, Wu K, Hou Y, Wiman KG, Li Z, Zhang H, Peng R, Zhu S, Pan-Hammarström Q. Genetic landscape of hepatitis B virus-associated diffuse large B-cell lymphoma. Blood 2018; 131(24): 2670–2681
25. Schmitz R, Wright GW, Huang DW, Johnson CA, Phelan JD, Wang JQ, Roulland S, Kasbekar M, Young RM, Shaffer AL, Hodson DJ, Xiao W, Yu X, Yang Y, Zhao H, Xu W, Liu X, Zhou B, Du W, Chan WC, Jaffe ES, Gascoyne RD, Connors JM, Campo E, Lopez-Guillermo A, Rosenwald A, Ott G, Delabie J, Rimsza LM, Tay Kuang Wei K, Zelenetz AD, Leonard JP, Bartlett NL, Tran B, Shetty J, Zhao Y, Soppet DR, Pittaluga S, Wilson WH, Staudt LM. Genetics and pathogenesis of diffuse large B-cell lymphoma. N Engl J Med 2018; 378(15): 1396–1407
26. Xu PF, Li C, Xi SY, Chen FR, Wang J, Zhang ZQ, Liu Y, Li X, Chen ZP. Whole exome sequencing reveals the genetic heterogeneity and evolutionary history of primary gliomas and matched recurrences. Comput Struct Biotechnol J 2022; 20: 2235–2246
27. Brastianos PK, Horowitz PM, Santagata S, Jones RT, McKenna A, Getz G, Ligon KL, Palescandolo E, Van Hummelen P, Ducar MD, Raza A, Sunkavalli A, Macconaill LE, Stemmer-Rachamimov AO, Louis DN, Hahn WC, Dunn IF, Beroukhim R. Genomic sequencing of meningiomas identifies oncogenic SMO and AKT1 mutations. Nat Genet 2013; 45(3): 285–289
28. Kim J, Hwang K, Kwon HJ, Lee JE, Lee KS, Choe G, Han JH, Kim CY. Clinicopathologic characteristics of grade 2/3 meningiomas: a perspective on the role of next-generation sequencing. Front Oncol 2022; 12: 885155
29. Ferreri AJM, Blay JY, Reni M, Pasini F, Spina M, Ambrosetti A, Calderoni A, Rossi A, Vavassori V, Conconi A, Devizzi L, Berger F, Ponzoni M, Borisch B, Tinguely M, Cerati M, Milani M, Orvieto E, Sanchez J, Chevreau C, Dell'Oro S, Zucca E, Cavalli F. Prognostic scoring system for primary CNS lymphomas: the International Extranodal Lymphoma Study Group experience. J Clin Oncol 2003; 21(2): 266–272
30. Abrey LE, Ben-Porat L, Panageas KS, Yahalom J, Berkey B, Curran W, Schultz C, Leibel S, Nelson D, Mehta M, DeAngelis LM. Primary central nervous system lymphoma: the Memorial Sloan-Kettering Cancer Center prognostic model. J Clin Oncol 2006; 24(36): 5711–5715
31. Radke J, Ishaque N, Koll R, Gu Z, Schumann E, Sieverling L, Uhrig S, Hübschmann D, Toprak UH, López C, Hostench XP, Borgoni S, Juraeva D, Pritsch F, Paramasivam N, Balasubramanian GP, Schlesner M, Sahay S, Weniger M, Pehl D, Radbruch H, Osterloh A, Korfel A, Misch M, Onken J, Faust K, Vajkoczy P, Moskopp D, Wang Y, Jödicke A, Trümper L, Anagnostopoulos I, Lenze D, Küppers R, Hummel M, Schmitt CA, Wiestler OD, Wolf S, Unterberg A, Eils R, Herold-Mende C, Brors B; ICGC MMML-Seq Consortium; Siebert R, Wiemann S, Heppner FL. The genomic and transcriptional landscape of primary central nervous system lymphoma. Nat Commun 2022; 13(1): 2558
32. Gonzalez-Aguilar A, Idbaih A, Boisselier B, Habbita N, Rossetto M, Laurenge A, Bruno A, Jouvet A, Polivka M, Adam C, Figarella-Branger D, Miquel C, Vital A, Ghesquières H, Gressin R, Delwail V, Taillandier L, Chinot O, Soubeyran P, Gyan E, Choquet S, Houillier C, Soussain C, Tanguy ML, Marie Y, Mokhtari K, Hoang-Xuan K. Recurrent mutations of MYD88 and TBL1XR1 in primary central nervous system lymphomas. Clin Cancer Res 2012; 18(19): 5203–5211
33. Pecqueux C, Arslan A, Heller M, Falkenstein M, Kaczorowski A, Tolstov Y, Sultmann H, Grullich C, Herpel E, Duensing A, Kristiansen G, Hohenfellner M, Navone NM, Duensing S. FGF-2 is a driving force for chromosomal instability and a stromal factor associated with adverse clinico-pathological features in prostate cancer. Urol Oncol 2018; 36(8): 365 e315–365 e326
34. Guo K, Ma Z, Zhang Y, Han L, Shao C, Feng Y, Gao F, Di S, Zhang Z, Zhang J, Tabbò F, Ekman S, Suda K, Cappuzzo F, Han J, Li X, Yan X. HDAC7 promotes NSCLC proliferation and metastasis via stabilization by deubiquitinase USP10 and activation of β-catenin-FGF18 pathway. J Exp Clin Cancer Res 2022; 41(1): 91
35. Xu Z, Cai Y, Liu W, Kang F, He Q, Hong Q, Zhang W, Li J, Yan Y, Peng J. Downregulated exosome-associated gene *FGF9* as a novel diagnostic and prognostic target for ovarian cancer and its underlying roles in immune regulation. Aging (Albany NY) 2022; 14(4): 1822–1835
36. Chen YX, Liu XJ, Yang L, He JJ, Jiang YM, Mai J. Systematic analysis of expression profiles and prognostic significance of the *FGF* gene family in pancreatic adenocarcinoma. Oncol Lett 2022; 24(6): 435
37. Wu Y, Yi Z, Li J, Wei Y, Feng R, Liu J, Huang J, Chen Y, Wang X, Sun J, Yin X, Li Y, Wan J, Zhang L, Huang J, Du H, Wang X, Li Q, Ren G, Li H. FGFR blockade boosts T cell infiltration into triple-negative breast cancer by regulating cancer-associated fibroblasts. Theranostics 2022; 12(10): 4564–4580
38. Peng J, Sridhar S, Siefker-Radtke AO, Selvarajah S, Jiang DM. Targeting the FGFR pathway in urothelial carcinoma: the future is now. Curr Treat Options Oncol 2022; 23(9): 1269–1287
39. Krull JE, Wenzl K, Hartert KT, Manske MK, Sarangi V, Maurer MJ, Larson MC, Nowakowski GS, Ansell SM, McPhail E, Habermann TM, Link BK, King RL, Cerhan JR, Novak AJ. Somatic copy number gains in MYC, BCL2, and BCL6 identifies a subset of aggressive alternative-DH/TH DLBCL patients. Blood Cancer J 2020; 10(11): 117
40. Walker BA, Mavrommatis K, Wardell CP, Ashby TC, Bauer M, Davies FE, Rosenthal A, Wang H, Qu P, Hoering A, Samur M, Towfic F, Ortiz M, Flynt E, Yu Z, Yang Z, Rozelle D, Obenauer J, Trotter M, Auclair D, Keats J, Bolli N, Fulciniti M, Szalat R, Moreau P, Durie B, Stewart AK, Goldschmidt H, Raab MS, Einsele H, Sonneveld P, San Miguel J, Lonial S, Jackson GH, Anderson KC, Avet-Loiseau H, Munshi N, Thakurta A, Morgan

GJ. Identification of novel mutational drivers reveals oncogene dependencies in multiple myeloma. Blood 2018; 132(6): 587–597
41. An HW, Kim SY, Kwon JW, Seok SH, Woo SH, Kim DY, Park JW. *In vivo* CRISPR-Cas9 knockout screening using quantitative PCR identifies thymosin beta-4 X-linked that promotes diffuse-type gastric cancer metastasis. Mol Carcinog 2021; 60(9): 597–606
42. Chu Y, You M, Zhang J, Gao G, Han R, Luo W, Liu T, Zuo J, Wang F. Adipose-derived mesenchymal stem cells enhance ovarian cancer growth and metastasis by increasing thymosin beta 4X-linked expression. Stem Cells Int 2019; 2019: 9037197
43. Khodabakhshi AH, Morin RD, Fejes AP, Mungall AJ, Mungall KL, Bolger-Munro M, Johnson NA, Connors JM, Gascoyne RD, Marra MA, Birol I, Jones SJM. Recurrent targets of aberrant somatic hypermutation in lymphoma. Oncotarget 2012; 3(11): 1308–1319
44. Yuniati L, Scheijen B, van der Meer LT, van Leeuwen FN. Tumor suppressors BTG1 and BTG2: beyond growth control. J Cell Physiol 2019; 234(5): 5379–5389
45. Liao D. Emerging roles of the EBF family of transcription factors in tumor suppression. Mol Cancer Res 2009; 7(12): 1893–1901
46. Hodkinson BP, Schaffer M, Brody JD, Jurczak W, Carpio C, Ben-Yehuda D, Avivi I, Forslund A, Özcan M, Alvarez J, Ceulemans R, Fourneau N, Younes A, Balasubramanian S. Biomarkers of response to ibrutinib plus nivolumab in relapsed diffuse large B-cell lymphoma, follicular lymphoma, or Richter's transformation. Transl Oncol 2021; 14(1): 100977
47. Rohde M, Bonn BR, Zimmermann M, Lange J, Möricke A, Klapper W, Oschlies I, Szczepanowski M, Nagel I, Schrappe M; MMML-MYC-SYS Project; ICGC MMML-Seq Project; Loeffler M, Siebert R, Reiter A, Burkhardt B. Relevance of ID3-TCF3-CCND3 pathway mutations in pediatric aggressive B-cell lymphoma treated according to the non-Hodgkin Lymphoma Berlin-Frankfurt-Münster protocols. Haematologica 2017; 102(6): 1091–1098
48. Martín-Garcia D, Navarro A, Valdés-Mas R, Clot G, Gutiérrez-Abril J, Prieto M, Ribera-Cortada I, Woroniecka R, Rymkiewicz G, Bens S, de Leval L, Rosenwald A, Ferry JA, Hsi ED, Fu K, Delabie J, Weisenburger D, de Jong D, Climent F, O'Connor SJ, Swerdlow SH, Torrents D, Beltran S, Espinet B, González-Farré B, Veloza L, Costa D, Matutes E, Siebert R, Ott G, Quintanilla-Martinez L, Jaffe ES, López-Otín C, Salaverria I, Puente XS, Campo E, Beà S. *CCND2* and *CCND3* hijack immunoglobulin light-chain enhancers in cyclin D1$^-$ mantle cell lymphoma. Blood 2019; 133(9): 940–951
49. Donati B, Lorenzini E, Ciarrocchi A. BRD4 and cancer: going beyond transcriptional regulation. Mol Cancer 2018; 17(1): 164
50. Chapuy B, McKeown MR, Lin CY, Monti S, Roemer MG, Qi J, Rahl PB, Sun HH, Yeda KT, Doench JG, Reichert E, Kung AL, Rodig SJ, Young RA, Shipp MA, Bradner JE. Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. Cancer Cell 2013; 24(6): 777–790
51. Tsukamoto T, Nakahata S, Sato R, Kanai A, Nakano M, Chinen Y, Maegawa-Matsui S, Matsumura-Kimoto Y, Takimoto-Shimomura T, Mizuno Y, Kuwahara-Ota S, Kawaji Y, Taniwaki M, Inaba T, Tashiro K, Morishita K, Kuroda J. BRD4-regulated molecular targets in mantle cell lymphoma: insights into targeted therapeutic approach. Cancer Genomics Proteomics 2020; 17(1): 77–89
52. Chung NG, Kim MS, Yoo NJ, Lee SH. Somatic mutation of STAG2, an aneuploidy-related gene, is rare in acute leukemias. Leuk Lymphoma 2012; 53(6): 1234–1235
53. Hashemi Zonouz T, Abdulbaki R, Bandyopadhyay BC, Nava VE. Novel mutations in a lethal case of lymphomatous adult T cell lymphoma with cryptic myocardial involvement. Curr Oncol 2021; 28(1): 818–824
54. Serizawa M, Yokota T, Hosokawa A, Kusafuka K, Sugiyama T, Tsubosa Y, Yasui H, Nakajima T, Koh Y. The efficacy of uracil DNA glycosylase pretreatment in amplicon-based massively parallel sequencing with DNA extracted from archived formalin-fixed paraffin-embedded esophageal cancer tissues. Cancer Genet 2015; 208(9): 415–427
55. Fukumura K, Kawazu M, Kojima S, Ueno T, Sai E, Soda M, Ueda H, Yasuda T, Yamaguchi H, Lee J, Shishido-Hara Y, Sasaki A, Shirahata M, Mishima K, Ichimura K, Mukasa A, Narita Y, Saito N, Aburatani H, Nishikawa R, Nagane M, Mano H. Genomic characterization of primary central nervous system lymphoma. Acta Neuropathol 2016; 131(6): 865–875