Molecular and clinical diversity in primary central nervous system lymphoma

I. Hernández-Verdin, E. Kirasic, K. Wienand, K. Mokhtari, S. Eimer, H. Loiseau, A. Rousseau, J. Paillassa, G. Ahle, F. Lerintiu, et al.

ESMO GOOD SCIENCE BETTER MEDICINE BEST PRACTICE

ANNALS OF ONCOLOGY driving innovation in oncology

ORIGINAL ARTICLE

Molecular and clinical diversity in primary central nervous system lymphoma

I. Hernández-Verdin[1], E. Kirasic[1], K. Wienand[2,3,4], K. Mokhtari[1,5], S. Eimer[6], H. Loiseau[7,8], A. Rousseau[9,10], J. Paillassa[11], G. Ahle[12], F. Lerintiu[13], E. Uro-Coste[14,15,16], L. Oberic[17], D. Figarella-Branger[18,19], O. Chinot[20,21], G. Gauchotte[22,23,24,25], L. Taillandier[26], J.-P. Marolleau[27], M. Polivka[28], C. Adam[29], R. Ursu[30], A. Schmitt[31], N. Barillot[1], L. Nichelli[32], F. Lozano-Sánchez[33], M.-J. Ibañez-Juliá[34], M. Peyre[1,35], B. Mathon[1,35], Y. Abada[33], F. Charlotte[36], F. Davi[37], C. Stewart[38], A. de Reyniès[39], S. Choquet[36], C. Soussain[40], C. Houillier[33], B. Chapuy[2,3], K. Hoang-Xuan[1,33] & A. Alentorn[1,33*]

[1]Institut du Cerveau-Paris Brain Institute-ICM, Inserm, Sorbonne Université, CNRS, Paris, France; [2]Department of Hematology and Medical Oncology, University Medical Center Göttingen, Göttingen; [3]Department of Hematology, Oncology and Cancer Immunology, Campus Benjamin Franklin, Charité—Universitätsmedizin Berlin, Berlin; [4]Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany; [5]Department of Neuropathology, Groupe Hospitalier Pitié Salpêtrière, APHP, Paris; [6]Department of Pathology, CHU de Bordeaux, Hôpital Pellegrin, Bordeaux; [7]Department of Neurosurgery, Bordeaux University Hospital Center, Pellegrin Hospital, Bordeaux; [8]EA 7435—IMOTION, University of Bordeaux, Bordeaux; [9]Department of Pathology, PBH, CHU Angers, Angers; [10]CRCINA, Université de Nantes—université d'Angers, Angers; [11]Department of Hematology, CHU Angers, Angers; [12]Department of Neurology, Hôpitaux Civils de Colmar, Colmar; [13]Department of Neuropathology, Hôpitaux Civils de Colmar, Strasbourg; [14]Department of Pathology, CHU de Toulouse, IUC-Oncopole, Toulouse; [15]INSERM U1037, Cancer Research Center of Toulouse (CRCT), Toulouse; [16]Université Toulouse III Paul Sabatier, Toulouse; [17]Department of Hematology, IUC Toulouse Oncopole, Toulouse; [18]Neuropathology Department, University Hospital Timone, Aix Marseille University, Marseille; [19]Inst Neurophysiopathol, CNRS, INP, Aix-Marseille University, Marseille; [20]Department of Neuro-oncology, CHU Timone, APHM, Marseille; [21]Institute of NeuroPhysiopathology, CNRS, INP, Aix-Marseille University, Marseille; [22]Department of Biopathology, CHRU Nancy, CHRU/ICL, Bâtiment BBB, Vandoeuvre-lès-Nancy; [23]Department of Legal Medicine, CHRU Nancy, Vandoeuvre-lès-Nancy; [24]INSERM U1256, University of Lorraine, Vandoeuvre-lès-Nancy; [25]Centre de Ressources Biologiques, BB-0033-00035, CHRU, Nancy; [26]Department of Neuro-oncology, CHRU-Nancy, Université de Lorraine, Nancy; [27]Department of Hematology, CHU Amiens-Picardie, Amiens; [28]Department of Anatomopathology, Lariboisière Hospital, Assistance Publique—Hopitaux de Paris, University of Paris, Paris; [29]Pathology Department, Bicêtre University Hospital, Public Hospital Network of Paris, Le Kremlin Bicêtre; [30]Department of Neurology, Université de Paris, AP-HP, Hôpital Saint Louis, Paris; [31]Department of Hematology, Institut Bergonié Hospital, Bordeaux; [32]Department of Neuroradiology, Sorbonne Université, Assistance Publique—Hôpitaux de Paris, Groupe Hospitalier Pitié-Salpêtrière-Charles Foix, Paris; [33]Department of Neurology-2, Sorbonne Université, Assistance Publique—Hôpitaux de Paris, Groupe Hospitalier Pitié-Salpêtrière-Charles Foix, Paris; [34]Department of Neurology, Perpignan Hospital, Perpignan; [35]Department of Neurosurgery, Sorbonne Université, Assistance Publique—Hôpitaux de Paris, Groupe Hospitalier Pitié-Salpêtrière-Charles Foix, Paris; Departments of [36]Pathology; [37]Hematology, APHP, Hôpital Pitié-Salpêtrière and Sorbonne University, Paris, France; [38]Broad Institute of MIT and Harvard, Cambridge, USA; [39]INSERM UMR_S1138—Centre de Recherche des Cordeliers—Université Pierre et Marie Curie et Université Paris Descartes, Paris; [40]Hematology Unit, Institut Curie, Saint-Cloud, France

Available online 17 November 2022

Background: Primary central nervous system lymphoma (PCNSL) is a rare and distinct entity within diffuse large B-cell lymphoma presenting with variable response rates probably to underlying molecular heterogeneity.

Patients and methods: To identify and characterize PCNSL heterogeneity and facilitate clinical translation, we carried out a comprehensive multi-omic analysis [whole-exome sequencing, RNA sequencing (RNA-seq), methylation sequencing, and clinical features] in a discovery cohort of 147 fresh-frozen (FF) immunocompetent PCNSLs and a validation cohort of formalin-fixed, paraffin-embedded (FFPE) 93 PCNSLs with RNA-seq and clinico-radiological data.

Results: Consensus clustering of multi-omic data uncovered concordant classification of four robust, non-overlapping, prognostically significant clusters (CS). The CS1 and CS2 groups presented an immune-cold hypermethylated profile but a distinct clinical behavior. The 'immune-hot' CS4 group, enriched with mutations increasing the Janus kinase (JAK)—signal transducer and activator of transcription (STAT) and nuclear factor-κB activity, had the most favorable clinical outcome, while the heterogeneous-immune CS3 group had the worse prognosis probably due to its association with meningeal infiltration and enriched *HIST1H1E* mutations. CS1 was characterized by high Polycomb repressive complex 2 activity and *CDKN2A/B* loss leading to higher proliferation activity. Integrated analysis on proposed targets suggests potential use of immune checkpoint inhibitors/JAK1 inhibitors for CS4, cyclin D-Cdk4,6 plus phosphoinositide 3-kinase (PI3K) inhibitors for CS1, lenalidomide/demethylating drugs for CS2, and enhancer of

*Correspondence to: Dr Agustí Alentorn, Department of Neurology-2, Mazarin, Groupe Hospitalier Pitié Salpêtrière, 75013 Paris, France. Tel: +33-142164160
E-mail: agusti.alentorn@icm-institute.org (A. Alentorn).

zeste 2 polycomb repressive complex 2 subunit (EZH2) inhibitors for CS3. We developed an algorithm to identify the PCNSL subtypes using RNA-seq data from either FFPE or FF tissue.
Conclusions: The integration of genome-wide data from multi-omic data revealed four molecular patterns in PCNSL with a distinctive prognostic impact that provides a basis for future clinical stratification and subtype-based targeted interventions.
Key words: PCNSL, microenvironment, tumor heterogeneity, multi-omics

INTRODUCTION

Primary central nervous system lymphoma (PCNSL) is a rare subtype of extranodal non-Hodgkin's lymphoma within diffuse large B-cell lymphoma (DLBCL), but has a less favorable prognosis than its systemic counterpart and has been proved to be molecularly a different biological entity.[1-6] The standard treatment relies on high-dose methotrexate (HD-MTX) regimen with or without consolidation and is associated with treatment resistance or relapses in up to 60% of the patients.[7,8]

Biologically, initial studies have found PCNSL to be at late B-cell germinal center (GC) exit stages and to have constitutive nuclear factor-κB (NF-κB) activity driven by mutations in genes of the B-cell receptor (BCR) pathway, of the toll-like receptor (TLR) pathway (*MYD88*), and *CARD11*.[6,9-17] Recently, DLBCL has been divided into different molecular clusters and PCNSL has been related to the so-called 'MCD' (based on the co-occurrence of $MYD88^{L265P}$ and *CD79B* mutations) or cluster 5 (C5) DLBCL, both converging in the presence of frequent $MYD88^{L265P}$, *CD79B*, *PIM1*, *BTG2* mutations, immunoglobulin heavy locus & B-cell lymphoma 6 (IgH-BCL6) translocations, copy gains of 3q12.3 and 9p24.1 [programmed death-ligand 1 (PD-L1)/PD-L2], and copy losses of 6p21-22 [human leukocyte antigen (HLA) locus], 6q21, and 9p21.3 (*CDKN2A* biallelic loss).[1,2,18-21] Although the role of the tumor microenvironment (TME) in PCNSL has gained terrain recently, a comprehensive description, including its interaction with methylation/mutation data, is still lacking.[22-30] Therefore, PCNSL heterogeneity has not been properly addressed mainly due to the lack of multi-omic data integration and the limited number of patients.[19]

Here, we carried out an integrative analysis of different data layers that included mutations, copy number alterations (CNAs), fusions, gene expression, T-cell receptor (TCR)/BCR clonotypes, TME, methylation, tumor localization, and diverse clinical data to identify molecular subtypes of treatment-naive immunocompetent Epstein—Barr virus-negative (EBV^-) PCNSL with clinically distinct behaviors. Additionally, to facilitate routine clinical implementation, we developed an algorithm that uses only gene expression data from either formalin-fixed, paraffin-embedded (FFPE) or fresh-frozen (FF) tissue, to identify the PCNSL molecular subtypes associated with multi-omic features.

PATIENTS AND METHODS

Patients

A total of 147 biopsies from treatment-naive FF (discovery cohort) and 93 FFPE (validation cohort) tumor samples from immunocompetent EBV^- PCNSL were recollected from different hospitals across France (see Supplementary Table S1, available at https://doi.org/10.1016/j.annonc.2022.11.002). None of the patients received corticosteroids for 1 week before the brain biopsy. All patients had a complete systemic evaluation to rule out secondary central nervous system DLBCL. Diagnoses and section selection for sequencing were established at the reference institution by specialized (neuro)pathologists. The tumor cell content for each sample was estimated to be at least 60% based on histomorphological evaluation. FFPE tissue sections available for all data (FF and FFPE cohorts) were used for immunophenotypic characterization with antibodies directed against CD20, CD10, BCL6, CD3, Ki67, and MUM1/IRF4. EBV-encoded small RNAs *in situ* hybridization was carried out as previously described.[31] We obtained appropriate consent from relevant institutional review boards, which coordinated the consent process at each tissue-source site; written informed consent was obtained from all participants. The Pitié Salpêtirère Hospital ethics committee approved the study (Ile-de-France VI, N° DC-2009-957) and CNIL (DR-2013-279). All patients received HD-MTX regimens according to French National 'Lymphome oculo-cérébral, LOC' PCNSL network.[7] Moreover, 19/134 (14.2%; FF cohort) and 24/93 (25.8%; FFPE cohort) received intensive chemotherapy with autologous stem cell rescue (IC-ASCR). Magnetic resonance images (MRI) of PCNSL at the time of first diagnosis and recurrence in sufficient quality were available for 90 patients within the validation cohort. We carried out, on the FF cohort, exome sequencing ($n = 115$), RNA sequencing (RNA-seq, $n = 123$), and DNA methylation profiling ($n = 64$). In the FFPE cohort, only RNA-seq was carried out ($n = 93$). Study design and sample size for each sequencing approach are displayed in Figure 1A.

Multi-omic data integration for PCNSL molecular subtyping

The multi-omic cohort comprised 85 FF samples from which we obtained six data types including mutations, CNA events, gene expression, immune cell proportions, TCR/BCR clonotypes, and fusion transcripts (Figure 1A). The most appropriate cluster number was obtained by clustering prediction index (CPI) and Gap-statistics analyses. High robust clustering using these data was obtained by consensus clustering resulting from 10 different multi-omics clustering algorithms (iClusterBayes, moCluster, CIMLR, IntNMF, ConsensusClustering, COCA, NEMO, PINSPlus, SNF, and LRA).[32-42] Clustering performance was evaluated for each cluster number ($n = 2$-8) using the weighted silhouette width.[43] Genomic data were deposited at the European

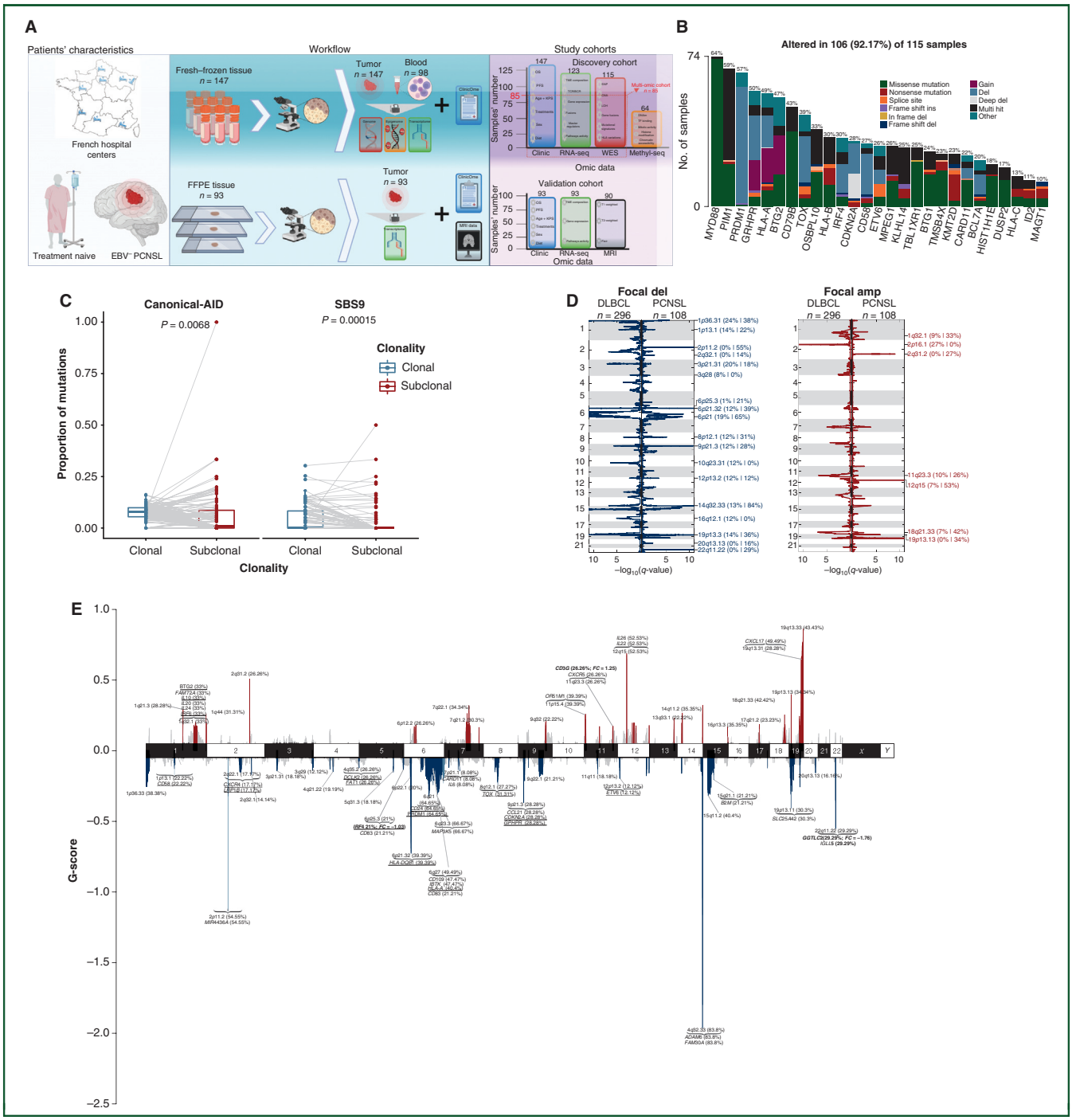

Figure 1. Study design and PCNSL mutational landscape. (A) The study design including the workflow and the data composition for each cohort is demonstrated. WES was carried out using tumor tissue and matched peripheral blood samples for 98 PCNSL patients, while for 17 patients it was tumor-only. For the methyl-seq, we also included four control peripheral blood samples. Different types of information were available for each omic data, as seen inside each bar plot. The multi-omic cohort consisted of 85 samples that had WES, RNA-seq, and clinic data within the discovery fresh-frozen cohort (red dashed box). (B) The number of affected samples within the top driver genes (identified by dNdScv algorithm) in the cohort of 115 PCNSL patients is shown. Bar plots are filled according to mutation type (missense, nonsense, splice site, frameshift, multihit, or other) or CNA events (gain, deletion, or deep deletion). The frequency of affected samples within the cohort is annotated at the top of each bar plot for each driver gene. (C) The boxplot comparison of the proportion of mutations attributed to c-AID or SBS9 mutational processes when occurring at clonal times versus subclonal times is shown. *P* values were calculated by Wilcoxon matched-pairs signed rank test for paired data (per sample). (D) The GISTIC2.0-defined recurrent copy number focal deletions (blue, left) and gains (red, right) as mirror plots in DLBCL ($n = 296$ from Chapuy et al., 2018[1]) and PCNSL ($n = 108$ from this study) are shown. Chromosome position is on the *y*-axis, and significance is on the *x*-axis. CNAs are labeled with their associated cytoband/arm followed in brackets by the frequency of the alteration (DLBCL | PCNSL). (E) The GISTIC2.0 results of significantly recurring (events with a q-value of <0.1) focal amplifications (red) and deletions (blue) are shown, where the chromosomes are plotted in the *x*-axis and the GISTIC-scores (G-score) are plotted in the *y*-axis. Genes affected for each focal event are annotated followed by the percentage of altered samples ($n = 108$). Genes are underlined if they are driver genes or in bold if they are transcriptionally affected by the focal event (related to Supplementary Figure S8 and Table S4, available at https://doi.org/10.1016/j.annonc.2022.11.002).
c-AID, canonical activation-induced cytidine deaminase; CNA, copy number alteration; DLBCL, diffuse large B-cell lymphoma; EBV⁻, Epstein—Barr virus-negative; FC, fold change; KPS, Karnofsky performance status; methyl-seq, methylation sequencing; MRI, magnetic resonance imaging; OS, overall survival; PCNSL, primary central nervous system lymphoma; PFS, progression-free survival; RNA-seq, RNA sequencing; SNP, single-nucleotide polymorphism; TME, tumor microenvironment; WES, whole-exome sequencing.

Genome-phenome Archive (EGAD00001008706; http://www.ebi.ac.uk/ega/), which is hosted by the European Bioinformatics Institute (EBI). The RBraLymP algorithm is publicly accessible at https://github.com/iS4i4S/PCNSL-RBraLymP. All other materials are available upon request from the authors.

Statistical analyses

All statistical analyses were carried out using the R statistical programming environment (v4.0; The R Project for Statistical Computing, Vienna, Austria). Differences in proportions and binary/categorical variables were calculated from two-sample Z-tests or Fisher's exact test. The Kruskal–Wallis test was used to test for a difference in distribution between three or more independent groups, and the Mann–Whitney *U* test was used for differences in distributions between two population groups unless otherwise noted. *P* values were corrected for multiple comparisons using the Benjamini–Hochberg method when applicable. Overall survival (OS) and progression-free survival (PFS) analysis were assessed using log-rank Kaplan–Meier curves and multivariate Cox proportional hazards regression modeling. See Supplementary Appendix for full details, available at https://doi.org/10.1016/j.annonc.2022.11.002.

RESULTS

PCNSL mutational landscape

We identified 32 544 mutations in the 115 PCNSL FF samples analyzed (85% tumor-normal pairs, median = 3.23 mutations/Mb; range = 0.02-85.49 mutations/Mb; Supplementary Table S2 and Figure S1, available at https://doi.org/10.1016/j.annonc.2022.11.002). We applied the dNdScv[44] algorithm to identify driver mutations finding the hallmark mutations of PCNSL like *MYD88* (64%), *PIM1* (59%), *PRDM1* (57%), *GRHPR* (50%), *HLA-A/B/C* (49%, 30%, and 13%), *BTG2* (47%), *CD79B* (43%), *CDKN2A* (28%), *ETV6* (26%), *TBL1XR1* (25%), *KLHL14* (25%), *CARD11* (22%), and *HIST1H1E* (18%) which are involved in BCR-TLR-mediated NF-kB signaling, antigen presentation, cell cycle, histone modification, and B-cell differentiation regulation[1,2,12,18,19,21,45] (Figure 1B, Supplementary Table S3 and Figure S2A, available at https://doi.org/10.1016/j.annonc.2022.11.002). Moreover, we detected canonical activation-induced cytidine deaminase (c-AID) off-target mutations and found that they represent 7.9% [95% confidence interval (CI) 6.8% to 8.5%] of single-nucleotide variant mutations and fall within driver genes like *PIM1* (47%), *CD79B* (10%), *IRF4* (9%), and *HIST1H1E* (6%) (Supplementary Table S4 and Figure S2B, available at https://doi.org/10.1016/j.annonc.2022.11.002). Next, we used the clonal allelic status of each mutation along with the mutational process that probably originated them to gain information about the relative timing of these processes. Interestingly, both c-AID and non-c-AID (COSMIC signature SBS9) mutations are significantly more active at clonal stages ($P = 0.007$ and $P = 0.018$, respectively), hence reflecting the importance of AID activity in early stages of PCNSL tumorigenesis (Figure 1C), as previously described.[1,18,21,45,46]

Regarding focal CNA, we identified significant recurrent amplifications in 18q21.33 (42%) and 19p13.13 (34%), and deletions in 6p21 (39%), 6q21 (65%), 6q27 (49%), and 9p21.3 (28%) which have a higher frequency than those observed by Chapuy et al. in systemic DLBCL (Figure 1D).[1,18] Furthermore, we found recently described amplifications in 1q32.1 (33%, *IL10*) and 11q23.3 (26%, *CD3G*), and deletions in 6p25.3 (21%, *IRF4*), 22q11.22 (29%, *GGTLC2*), and 14q32.33 (84%) that produce significant expression changes in *CD3G* [fold change (FC) = 1.25], *IRF4* (FC = −1.03), and *GGTLC2* [FC = −1.76, false discovery rate (FDR) q-value <0.1], respectively[21] (Figure 1E, Supplementary Table S5 and Figure S3, available at https://doi.org/10.1016/j.annonc.2022.11.002).

Multi-omic data integration reveals PCNSL molecular subtypes with clinical outcome implications

To address the PCNSL heterogeneity, we carried out a cluster of clusters analysis using six levels of information including messenger RNA (mRNA) expression, mutations, CNA, fusion transcripts, TCR/BCR clonotypes, and immune cell fractions available in the multi-omic cohort (Figure 2A, Supplementary Figure S4, available at https://doi.org/10.1016/j.annonc.2022.11.002). Using the optimal cluster number resulting from CPI, Gap-statistics, and weighted silhouette width, we identified four PCNSL subtypes (CS1-CS4) that display different clinical outcomes in OS (global log-rank $P < 0.001$, Figure 2B). Patients in CS4 had the longest OS (median 66.8 months; 95% CI 19.8-67.2 months) and lived significantly longer than those in both clusters CS2 (median 18 months; 95% CI 8.3-53.4 months; $P = 0.024$) and CS3 (median 13.8 months; 95% CI 6.1-16.7 months; $P = 0.003$), and slightly longer, but not significantly, to those in CS1 (median 26.2 months; 95% CI 13.3-63.9 months; $P = 0.094$). Additionally, these observations remained significant after adjusting by age and Karnofsky performance status (KPS) in Cox proportional hazards ratio multivariate models (Figure 2B). Interestingly, CS4 was independently associated with a better response when considering PFS in univariate and multivariate models (Figure 2C). Furthermore, we did not observe significant differences in the median number of predicted immunogenic neoantigens ($P = 0.44$), the fraction of c-AID mutations ($P = 0.25$), or the number of patients receiving IC-ASCR ($P = 0.32$, Supplementary Table S6, available at https://doi.org/10.1016/j.annonc.2022.11.002).

Transcriptomic data correctly assign multi-omic defined PCNSL subtypes in FF and FFPE samples

Given the difficulty of acquiring FF tissue and of analyzing and implementing multi-omic data into routine clinical practice, we sought to evaluate the use of only RNA expression, obtained from FFPE or FF tissue, to categorize patients into the four subtypes. To do so, we first identified, for each subtype, the top 100 unique up-regulated

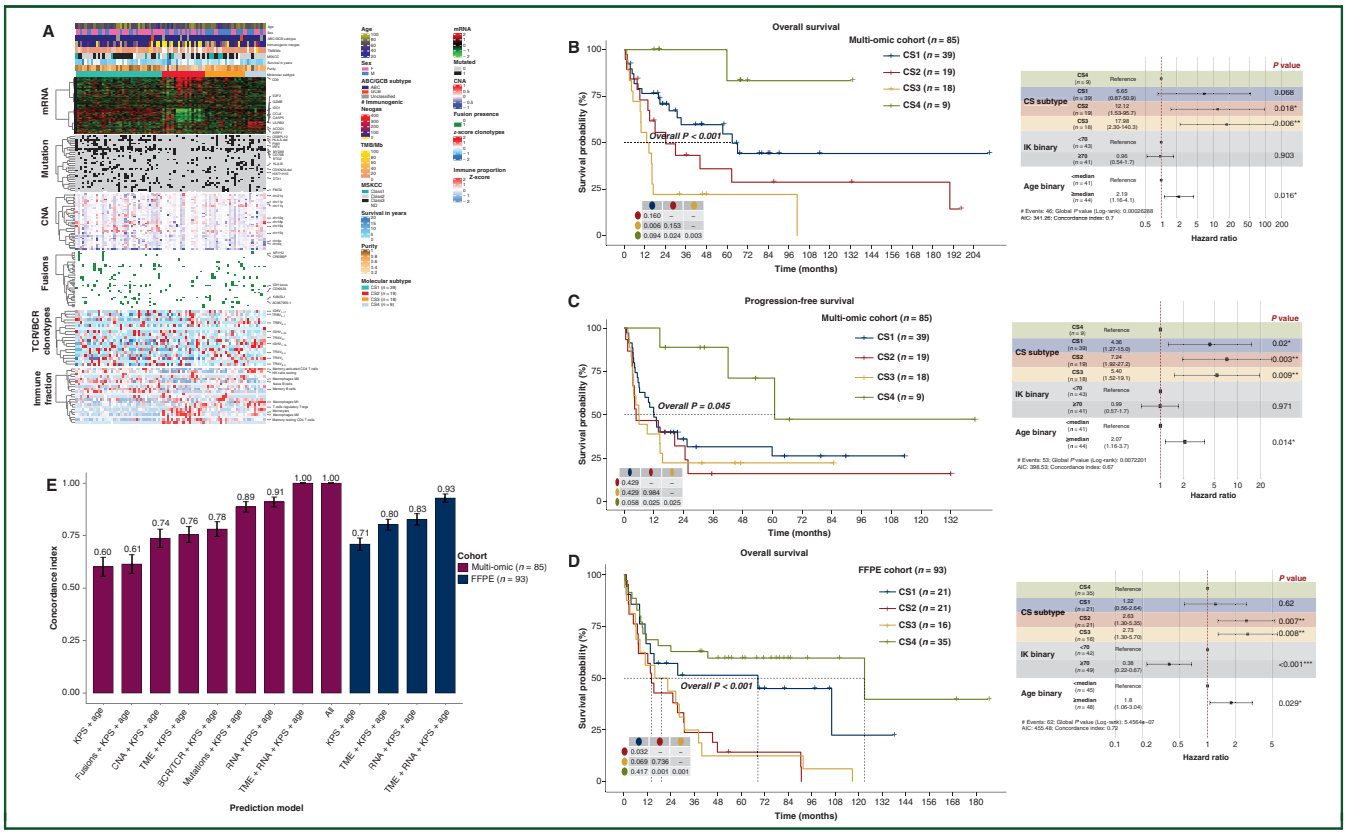

Figure 2. Multi-omic data integration reveals PCNSL molecular subtypes with clinical outcome implications. (A) The resulting consensus heatmap based on 10 integrative clustering algorithms to define the clusters (CS1-CS4) is shown, where each of the 10 algorithms uses cluster of clusters analysis to integrate six levels of omic information (*y*-axis) in the order: (i) mRNA expression (2087 variables), (ii) mutations (31 variables), (iii) CNA (40 variables), (iv) fusion transcripts (43 variables), (v) TCR/BCR clonotypes (19 variables), and (vi) immune cell fractions (22 variables). Additional genomic and clinical features are annotated at the top. (B) Kaplan–Meier (left) or hazard ratio (right) estimates of overall survival among patients belonging to each cluster that resulted from the Consensus cluster of clusters analysis are shown. (C) same as (B) but using progression-free survival. (D) Kaplan–Meier (left) or hazard ratio (right) estimates of overall survival among patients belonging to each cluster using an FFPE validation cohort ($n = 93$) are shown. (E) Harrell's concordance index (value annotated at top of each bar) obtained when evaluating the contribution of each omic-level data to outcome prediction models using Cox proportional hazards regression is shown. The prediction was overfit when using ALL omic data on the FF cohort. Observations were validated (same direction and effect) using RNA and TME data from the FFPE cohort. Error bars indicate 95% confidence intervals.

BCR, B-cell receptor; CNA, copy number alterations; CS, cluster; DLBCL, diffuse large B-cell lymphoma; FF, fresh-frozen; FFPE, formalin-fixed, paraffin-embedded; mRNA, messenger RNA; MSKCC, Memorial Sloan Kettering Cancer Center; PCNSL, primary central nervous system lymphoma; TCR, T-cell receptor; TMB, tumor mutational burden; TME, tumor microenvironment. *$P < 0.05$; **$P \leq 0.01$; ***$P \leq 0.001$.

biomarkers by carrying out differential expression analysis of one group versus the others and then extracting the non-overlapping genes across subtypes (P adjusted < 0.05; Supplementary Table S7, available at https://doi.org/10.1016/j.annonc.2022.11.002); afterward we applied different prediction methods on the mRNA data (Supplementary Figures S5-S7, available at https://doi.org/10.1016/j.annonc.2022.11.002). We obtained a Cohen's κ coefficient of 0.90 ($P < 0.001$, Supplementary Figure S8A, available at https://doi.org/10.1016/j.annonc.2022.11.002) when evaluating the accuracy of correctly assigning patients from the multi-omic cohort. Additionally, when expanding to the FF-RNA complete set ($n = 123$) or when using the FFPE cohort ($n = 93$), we observed the same behaviors regarding clinical outcome (global $P < 0.001$) across molecular subtypes in both univariate and multivariate models using OS and PFS (Figure 2D, Supplementary Figure S8, available at https://doi.org/10.1016/j.annonc.2022.11.002).

Next, we evaluated the contribution of each of the six data types to outcome prediction models by using Harrell's concordance index (C-index).[47] A C-index of 0.60 (95% CI 0.56-0.65) in FF and 0.71 (95% CI 0.68-0.74) in FFPE was observed using KPS and age, which are the clinical features currently used in the Memorial Sloan Kettering Cancer Center prognostic score for PCNSL.[48] When adding different data types to the multi-omic cohort modeling, we observed higher predictive power using mRNA expression compared to the others (C-index $= 0.91 \pm 0.02$). We further validated these observations in the FFPE cohort obtaining a C-index of 0.83 (95% CI 0.80-0.85) and 0.93 (95% CI 0.91-0.95) when adding the mRNA level or the TME + RNA levels to the model, respectively (Figure 2D). Altogether, these results show that RNA-seq data alone from FFPE or FF tissue can be used to correctly identify PCNSL subgroups.

Distinct genetic signatures within PCNSL subtypes and DLBCLs

Next, we aimed to characterize the differences in genetic alterations across groups by carrying out inside-outside cluster Fisher tests (q-value ≤ 0.1). The CS4 cluster

presents 10 enriched events that included mutations in *SOCS1,* which is a negative regulator of the Janus kinase (JAK)—signal transducer and activator of transcription 3 (STAT3) pathway, *MPEG1*, *PIM2*, and deletion of 17q25.1 involving *GRB2* that indirectly regulates the NF-κB pathway. We observed 43 events within the CS1 cluster including mutations involved in the NF-κB pathway (*RIPK1* via 6p25.3 deletion), B-cell differentiation (*IRF4* via 6p25.3 deletion, *TOX*, and *BCL6*), proliferation via interruption of cell cycle arrest (*CDKN2A/2B* fusions and *FOXC1*), and B-cell lymphomagenesis (e.g. *ETV6*, *OSBPL10*). Patients within the CS3 cluster exhibit 12 events from which *HIST1H1E* arises as the top enriched, and has been proved to enhance self-renewal properties and disrupt chromatin architecture in B-cell lymphomas.[1,2,18,45,49] The CS2 cluster did not present any genomic enriched events with respect to the other subtypes. Furthermore, most of these distinctive events arrived as early events (clonal) in tumorigenesis like *IRF4* and *BCL6* in CS1 (Figure 3A, Supplementary Table S8, available at https://doi.org/10.1016/j.annonc.2022.11.002). Of note, most of these mutations were not observed in the clusters previously defined by Chapuy et al. or were more CS-specific (e.g. 9p11.2 del; Figure 3B).[1] We also used the LymphGen algorithm to identify the correspondence of our samples to the genetic subtypes described by Wright et al.[45] We observed a similar proportion of samples that were classified as either MCD (40%, 34/85) or other (39%, 33/85), while the remaining felt within either A53 (13%, 11/85) or BN2 (8%, 7/85). The MCD was mainly composed of CS1 PCNSL subtypes (21/34 versus 13/34, $P = 0.015$) while the LymphGen's 'Other' had more CS2 samples ($P < 0.001$, Figure 3C). Furthermore, as expected, most samples (91%, 77/85) were activated B cells as per the cell-of-origin classification.

B-cell differentiation stages, pathways, and TME distinctions between PCNSL molecular subtypes

To determine whether or not these mutations could affect oncogenic pathways and B-cell differentiation programs transcriptionally, we recovered and analyzed the expression of different curated gene signatures initially proposed and used by Wright et al. and Kotlov et al.[45,50] CS1 was characterized by the up-regulation of PI3K, glycolytic activity, and cell proliferation signatures, which is consistent with the observed mutations in the cell cycle arrest; additionally, it presented hyperactivation of the Polycomb repressive complex 2 (PRC2) which has been proven to inhibit major histocompatibility complex (MHC)-I expression, through histone methylation (Figure 4A, $P < 0.05$).[51] Moreover, p53 activity was enriched in the CS2 subtype.[52,53] Interestingly, even though all subtypes presented mutations within the NF-κB pathway, it was transcriptionally active only in clusters CS3 and CS4. Additionally, mitogen-activated protein kinase (MAPK) and JAK—STAT pathways were up-regulated in those clusters, respectively (Figure 4A).

Regarding B-cell differentiation programs, CS1 expresses a mixture of GC cells, which is consistent with both the observed *MYC* induction activity, 6p25.3-19q13.12 deletions, and *BCL6* mutations (Figure 2C). On the other hand, the cluster CS4 presents an enrichment in terminally differentiated plasma cells that goes in line with *BCL6* negative regulation, the absence of *MYC* induction, and *BCL6* mutations. The most heterogeneous cluster was CS3, presenting features of both GC and mature B cells. Intriguingly, CS2 did not present any B-cell stage enrichment but instead a lymphatic endothelial cell (LEC) gene signature (Figure 4A, Supplementary Figure S9, available at https://doi.org/10.1016/j.annonc.2022.11.002). These different B-cell states, including plasmablasts/plasma cells, have been previously described in DLBCL and PCNSL using bulk/single-cell RNA-seq approaches.[54-56] Additionally, Milpied et al. reported the co-existence of different B-cell states within a patient with follicular lymphoma.[57] To further corroborate these observations, we analyzed BCR clonotypes and the immunoglobulin (Ig) heavy-chain variable (V_H) and constant region expression across clusters. We observed across clusters the presence of BCR clones' diversity and $IgV_{H4\text{-}34}$ expression, which further supports the presence of self-antigen-dependent chronic active BCR signaling across clusters (Supplementary Figure S10, available at https://doi.org/10.1016/j.annonc.2022.11.002).[45,58]

Then, we aimed to describe the TME differences between subtypes by using CIBERSORTx-derived immune deconvolution and Kotlov's B-cell lymphoma-specific TME gene signatures.[50] The CS1 cluster is immunologically 'neutral'; meanwhile the CS2 cluster, which is immunologically depleted, exhibits expression of vascular endothelial cells, memory resting $CD4^+$ T cells, monocytes, and activation of γ-aminobutyric acid synthesis, which has been recently linked to B cells that inhibit $CD8^+$ T cells' killer function and promote monocyte differentiation into anti-inflammatory macrophages.[59] The CS4 cluster has a hot-inflammatory TME due to the presence of active $CD8^+$ T cells and natural killer cells (with high cytolytic activity score), and a high expression of MHC-I/II, and immune checkpoint molecules.[60] Conversely, heterogeneity was again observed for the CS3 subtype, with only inactivated macrophages M0 being more significantly enriched (Figure 4A, Supplementary Figure S11, available at https://doi.org/10.1016/j.annonc.2022.11.002). When analyzing TCR sequences, we found significantly higher T-cell clonotype diversity in CS4 versus others ($P < 0.05$, Supplementary Figure S12, available at https://doi.org/10.1016/j.annonc.2022.11.002). Similar results were observed in the FFPE cohort (Supplementary Figure S13, available at https://doi.org/10.1016/j.annonc.2022.11.002).

Moreover, immune staining revealed higher presence of both $CD3^+$ and $CD163^+$ cells in CS4, in line with the transcriptionally observed hot TME (Figure 4B). To evaluate the LEC signature, we stained PROX1, a homeobox gene master regulator (MR) in LEC differentiation that forms part of our transcriptionally defined LEC gene signature and that was used by Louveau et al. when describing LECs in the central nervous system,[61,62] finding a higher number of $PROX1^+$ cells in the CS2 subtype compared to the others. Finally, we

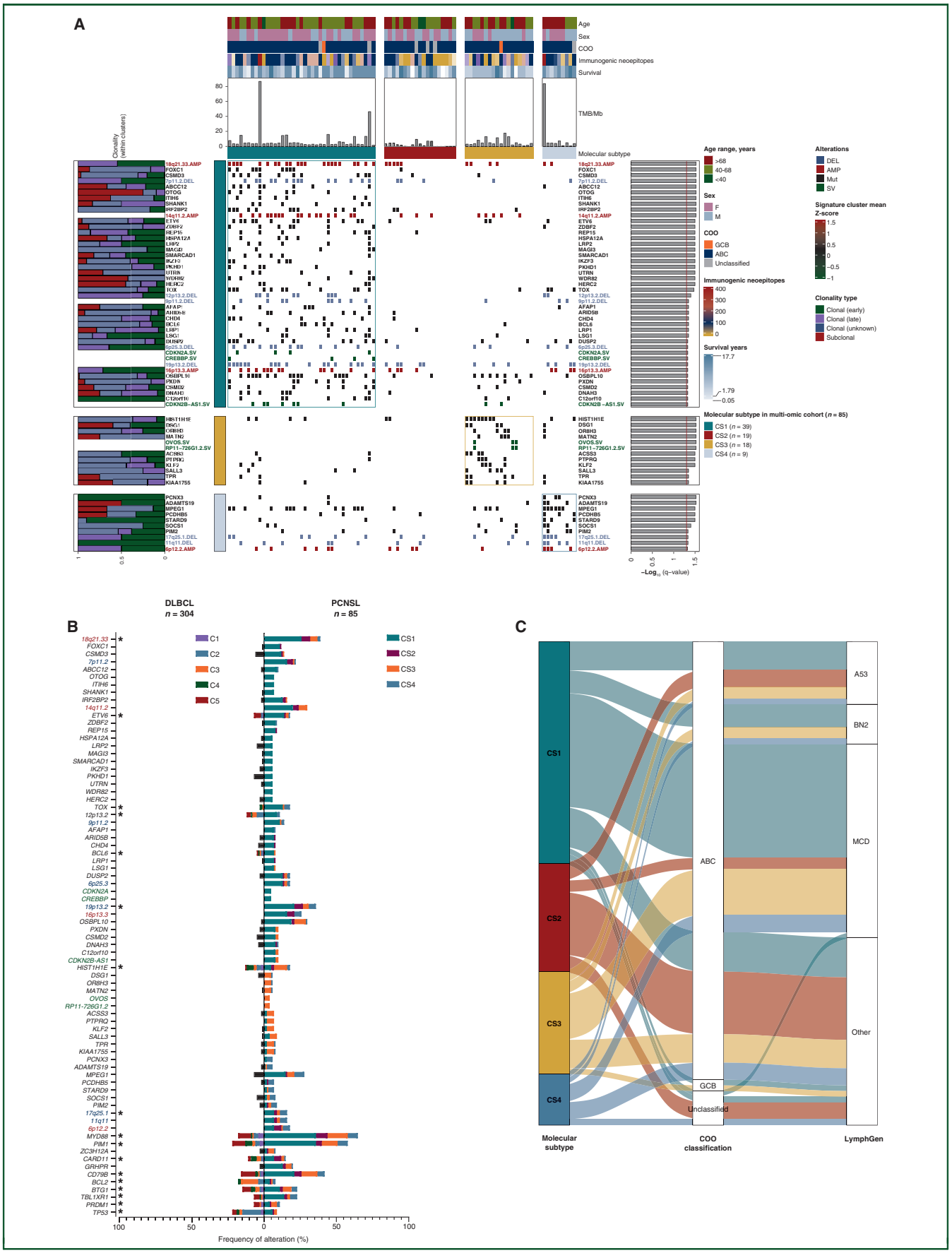

Figure 3. Distinct genetic signatures within PCNSL subtypes and systemic DLBCL. (A) The landmark genetic alterations are shown for each PCNSL subtype (boxed for each cluster) identified by a one-sided Fisher test (event within cluster versus outside cluster) and ranked by significance (FDR corrected q-value ≤0.1 selected, red

observed loss of repressive H3 lysine 23 trimethylation (H3K27me3^{+} cells) in the CS3 subtype, a phenomenon that has been previously associated with B-cell lymphomas carrying *HIST1H1B-E* mutations.[49]

Subtype-specific MRs

To elucidate subtype-specific MRs, we determined the transcription factors (TFs) of each molecular PCNSL subtype using the genes that are differentially expressed within each cluster and a curated collection of TF targets.[63] We found increased TF activity of *TEAD1* (in CS1 and CS2); *PRDM14* (CS1) which has roles as histone methyltransferase and leukemia initiator[64]; *IRF4, SPIB* (both targets of lenalidomide), *MYC*, and *PROX1* (CS2); *E2F1* and *IRF3* (CS3); and *STAT3/1* and *NFKB1* (CS4) (Figure 5A, Supplementary Tables S9 and S10 and Figure S14, available at https://doi.org/10.1016/j.annonc.2022.11.002).

Epigenetic attributes across PCNSL subtypes

Firstly, we evaluated the DNA methylation—transcriptome correlation by considering all PCNSL samples ($n = 64$; i.e. not grouping by subtypes) and a total of 27 111 genes with available data, and found that only 12.4% were correlated with their promoter methylation (P adjusted < 0.05 and $\rho < 0$). Interesting genes included *TERT*, *CD79A*, and *MGMT* ($\rho = -0.50$, -0.59, and -0.62, respectively; P adjusted $<$ 0.001, Supplementary Table S11A, available at https://doi.org/10.1016/j.annonc.2022.11.002), which goes in line with the literature.[65] When grouping by molecular subtypes, we found *CCL22*, which is a chemoattractant for primary activated T lymphocytes,[66] less expressed in the CS2 cluster and associated with gene hypermethylation (Supplementary Table S11, available at https://doi.org/10.1016/j.annonc.2022.11.002).

We proceeded to investigate epigenetic differences among subtypes. CS2 displayed higher hypermethylation globally, within promoters, and at chromosome ends ($P =$ 0.006, $P < 0.001$, and $P < 0.001$; Figure 5B); however, the CS1 subtype presented higher methylation within CpG islands ($P = 0.001$, Supplementary Figures S15 and S16, available at https://doi.org/10.1016/j.annonc.2022.11.002). Interestingly, a hypermethylator phenotype has been previously associated with a depleted TME in systemic DLBCL.[50]

Gene Ontology analyses on differentially methylated promoters revealed B-cell differentiation programs to be hypomethylated in CS1 but hypermethylated in CS2 (Figure 5C, Supplementary Table S12, available at https://doi.org/10.1016/j.annonc.2022.11.002). We observed cell adhesion hypermethylation only in CS2, which was previously reported when comparing PCNSL versus DLBCL.[67] Genomic region enrichment analysis on hypermethylated promoters identified strong enrichment of binding sites for the histone/chromatin proteins H3K27me3, and EZH2 in CS1, as expected from the high PRC2 observed transcriptionally. On the other hand, the CS2 subtype exhibited enrichment for BCL11A, NF-kB, IRF4, and BCL6 (Figure 5D, Supplementary Table S13, available at https://doi.org/10.1016/j.annonc.2022.11.002). These findings together with the observed increased LEC gene signature and PROX1 immune staining support the idea of a phenotypic shift to more LEC-like in this PCNSL subtype.

CS3 subtype is associated with meningeal infiltration to cerebrospinal fluid

Here, we investigated if brain MRI analysis ($n = 90$, FFPE cohort) could provide more insights about the molecular subtypes. CS4 tumors arose more in the isthmus of the corpus callosum (7/34 versus 0/56, $P < 0.001$). Conversely, CS2/CS3 were more frequent in the brainstem (3/19 and 4/16 versus 1/55; $P = 0.049$ and $P = 0.008$, respectively), when compared to the other clusters. Strikingly, we found no association with tumor size or multiple lesions. However, meningeal infiltration of the cerebrospinal fluid (CSF) was only found within CS3 tumors (6/16 versus 0/74, $P < 0.001$; Figure 6A, Supplementary Tables S14 and S15, available at https://doi.org/10.1016/j.annonc.2022.11.002).

Therapeutic associations of the PCNSL molecular subtypes

We next wondered if the PCNSL molecular subtypes may exhibit greater benefit from specific targeted therapies. Even though the hallmark PCNSL alterations targeting My-T-BCR protein supercomplex, CD79A/B BCR subunits, and the CBM (CARD11-BCL10-MALT1) complex were relatively constant across subgroups, the NF-κB hyperactive group (CS4) presented more *GRB2*/*LYN* deletions and absence of *PLCG2* mutations, which either represses the BCR complex or affects the CBM complex activation. Furthermore, NF-κB activity could not be explained by self-antigen-dependent chronic active BCR signaling up-regulation since $IgV_{H4\text{-}34}$ expression was similar across groups.[45,43,68] These observations suggest that CS4 and CS3 may be more sensitive to Bruton's tyrosine kinase (BTK) inhibitors. The CS4 cluster also presented high JAK—STAT activity and mutated *SOCS1* (JAK1 repressor), making it potentially responsive to JAK1 inhibitors.[69,70] Moreover, the absence of PRC2 activity and presence of MHC-I, only monoallelic deletions in *HLA-A/B2M/CD58* (Supplementary Figure S17, available at https://doi.org/10.1016/j.annonc.2022.11.002), and checkpoint molecules'

line, bar plot to the right). The left bar plot shows the relative contribution of temporal acquisition for each alteration event (only within the enriched cluster) to indicate how early or late during tumorigenesis the event might have happened. Additional genomic and clinical features are annotated at the top. (B) A mirror bar plot with the frequencies of recurrent genetic alterations in PCNSL's clusters ($n = 85$) compared to those in DLBCL's clusters ($n = 304$, Chapuy et al., 2018) is shown. Asterisks denote the known driver events in DLBCL and colors the alteration type (mutation = black; gain = red; loss = blue; structural variant = green). (C) The DLBCL cell-based COO transcriptomic signature subgroups and the Wright et al.[45] DLBCL subtypes (MCD, N1, A53, BN2, ST2, EZB, and others) within our CS PCNSL molecular subtypes are shown.

ABC, activated B cell; AMP, amplification; COO, cell of origin; CS, cluster; DEL, deletion; DLBCL, diffuse large B-cell lymphoma; F, female; FDR, false discovery rate; Fusion, fusion transcript; GCB, germinal center B cell; M, male; Mut, mutation; PCNSL, primary central nervous system lymphoma; TMB, tumor mutational burden.

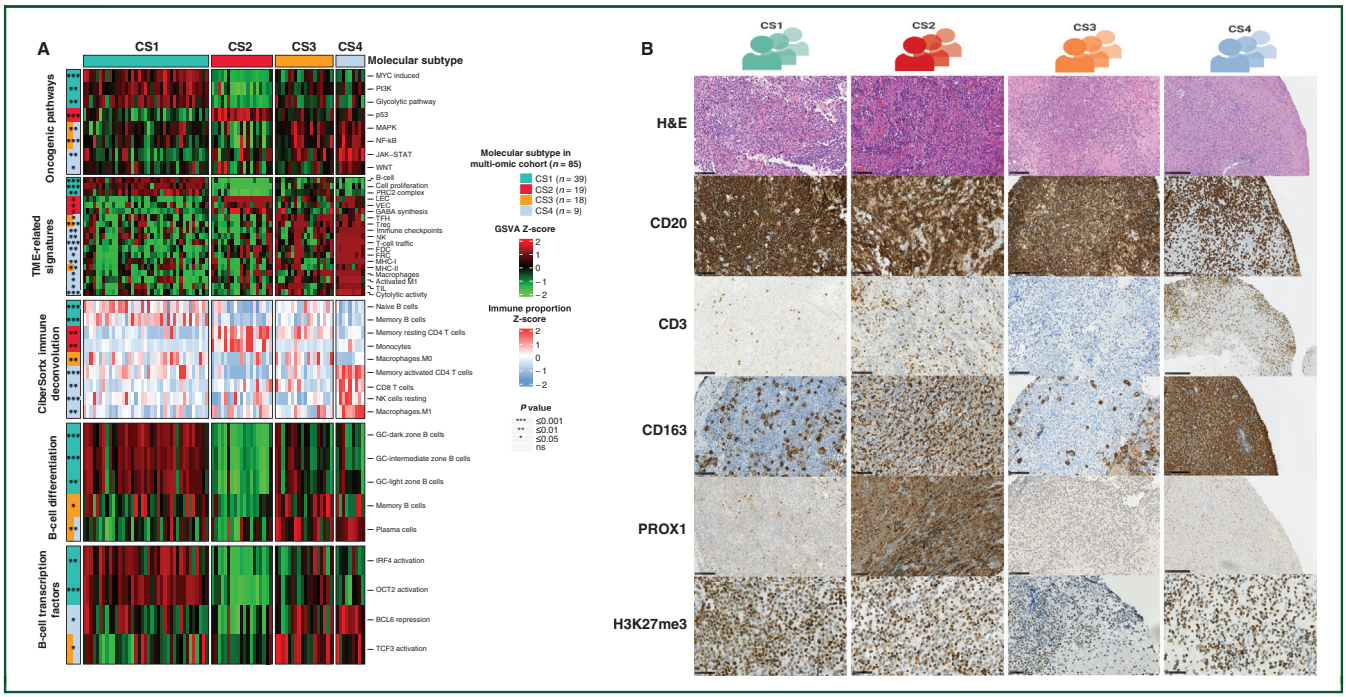

Figure 4. Phenotypic analysis of the multi-omic defined PCNSL subtypes. (A) A heatmap with either gene signature activity (measured by GSVA) or immune cell proportions (CIBERSORTx deconvoluted) across molecular subtypes is shown. *P* values indicate higher expression of the colored group when compared against the others (Wilcoxon test, left side of plot). (B) The hematoxylin—eosin staining and IHC of CD20, CD3, CD163, PROX1, and H3K27me3 in samples belonging to the different PCNSL subtypes are shown. Scale bars are indicated in each figure.

CS, cluster; FDC, follicular dendritic cells; FRC, fibroblastic reticular cells; GABA, γ-aminobutyric acid; GC, germinal center; GSVA, gene set variation analysis; H&E, hematoxylin—eosin; IHC, immunohistochemistry; JAK—STAT, Janus kinase—signal transducer and activator of transcription; LEC, lymphatic endothelial cell; MHC, major histocompatibility complex; NK, natural killer; PCNSL, primary central nervous system lymphoma; PRC2, Polycomb repressive complex 2.

expression indicate a potential use of immune checkpoint inhibitors (ICIs) for CS4. On the other hand, EZH2 inhibitors in combination with ICI could potentially increase MHC-I expression and immune detection in CS3.[71]

Additionally, we observed a higher frequency of cases with genetic alterations involved in the cell cycle for CS1 (97%, $P < 0.001$, Supplementary Figures S18 and S19, available at https://doi.org/10.1016/j.annonc.2022.11.002); hence, cyclin D-Cdk4,6 plus PI3K inhibitors could be beneficial for CS1 patients.

Despite not presenting enriched genetic signatures, the CS2 cluster may be potentially susceptible to inhibition of the TFs IRF4 (e.g. lenalidomide), SPIB, and MEIS1 (e.g. MEISi-1), demethylating drugs, and/or inhibition of GAD67.[45,59]

To further evaluate the proposed therapeutic targets, we downloaded the RNA-seq expression, the genome-wide CRISPR loss-of-function screens' scores, and the drug screening response data from 20 DLBCL cell lines.[72-74] We applied the partition around medoids method[70] using our PCNSL RNA-seq data as training expression and found DLBCL cell lines KARPAS-422, A3-KAW, and SU-DHL-4 to be transcriptionally similar to PCNSL subtypes CS1, CS2, and CS3, respectively. Next, we investigated the cumulative gene dependency on different pathways measured by CRISPR screen Chronos score since it has been demonstrated to exhibit the lowest copy number and screen quality bias.[73] We found the CS3-related cell line (SU-DHL-4) to have higher dependency on proteins involved in BCR subunits and BCR-dependent NF-κB activation (Figure 6B), while the CS1-related one benefited from *CDKN2A/B* knockouts. The PRC2 elements were equally essential for both CS1 and CS3. Interestingly, the *IRF4/SPIB* knockouts were selectively essential for the CS2-related cell line which corresponds to the MR analysis' results. When looking at the half maximal inhibitory concentration (IC_{50}) of ibrutinib in these cell lines, we found the lowest value in the CS3-related cell line ($IC_{50} = 1.71$; Figure 6C), thus supporting our hypothesis about CS3 subtype sensitivity to BTK inhibitors. Moreover, the CS1-related cell line showed the lowest IC_{50} toward AMG-319 ($IC_{50} = 10.19$), a PI3K-mammalian target of rapamycin (mTOR) inhibitor.

DISCUSSION

Identifying groups of patients with shared biologic and prognostic markers is extremely challenging mainly due to high genetic, phenotypic, and TME heterogeneity. Thanks to a collaborative nationwide effort, here, we carried out a multi-omic analysis of a large cohort of immunocompetent EBV^- PCNSL. We identified four PCNSL molecular subtypes with specific oncogenic pathways, gene expression phenotypes, methylation profiles, TME, tumor location, outcome, and potential therapeutic targets. Moreover, our study gives plausible explanations to the PCNSL response heterogeneity based on finding that many previously PCNSL characteristic

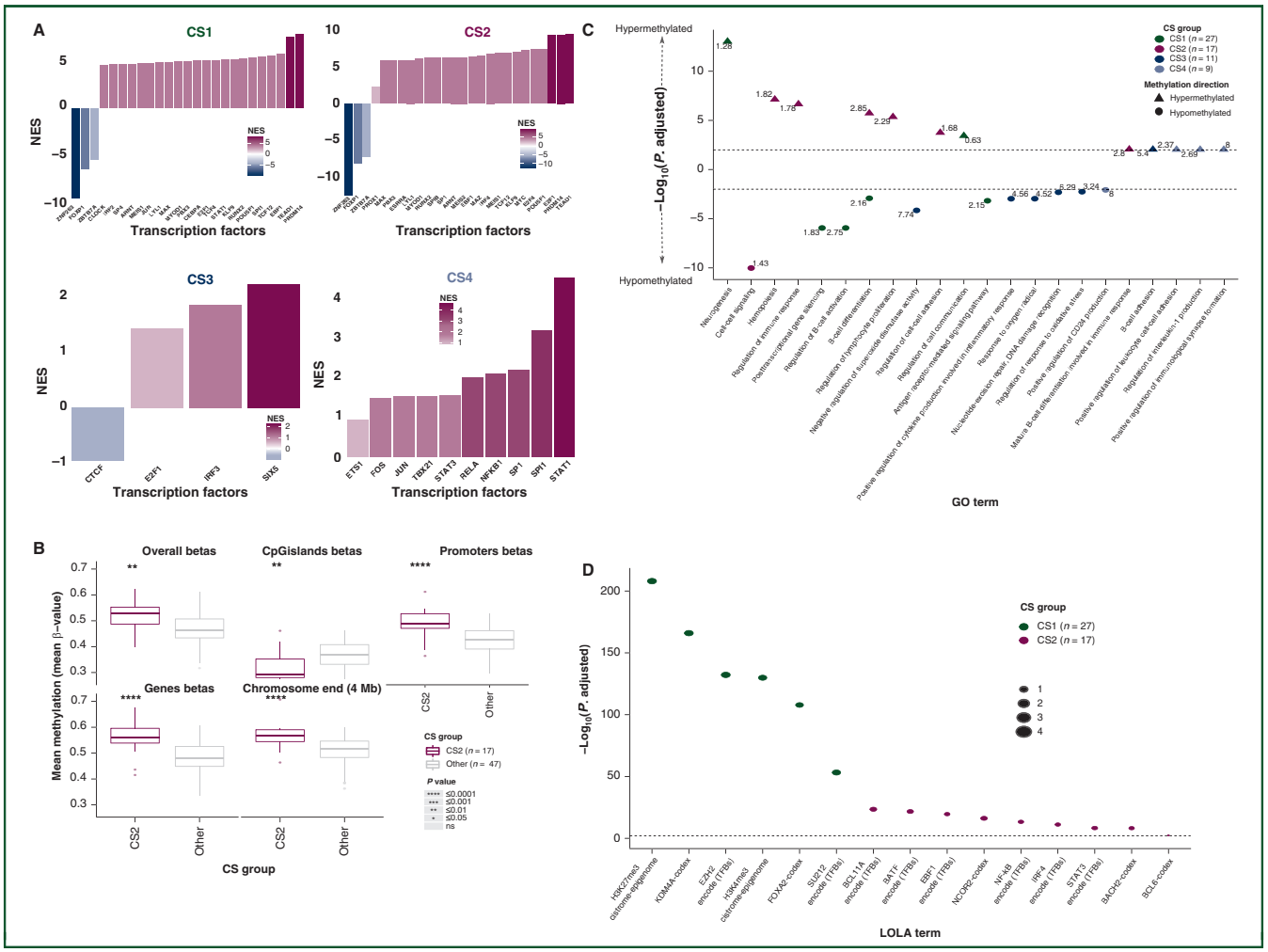

Figure 5. TFs and epigenetic attributes across PCNSL subtypes. (A) The TF regulon normalized enrichment scores (NESs) in the TFs found in the PCNSL subtypes by DoRothEA analysis are shown. (B) The boxplot comparison of CS2 versus other clusters using the mean methylation levels (mean β-values) globally, on CpG islands, on promoters, on gene body ('genes'), and at chromosome ends (4 Mb) is shown. (C) GO enrichment analysis on differentially methylated promoters (DMP) across subtypes, where the log(OddsRatio) is annotated next to its associated *P*-adjusted value, is shown. (D) The locus overlap (LOLA) region set enrichment analysis for hypermethylated promoters across the four PCNSL subtypes is shown. The *x*-axis presents the targets followed by the database. *P* values were calculated using a two-sided Fisher's exact test and then adjusted for multiple testing by the FDR method.
CS, cluster; FDR, false discovery rate; GO, Gene Ontology; PCNSL, primary central nervous system lymphoma; TF, transcription factors.

features, based on MCD or C5 DLBCL subtypes,[1,45] are cluster-specific (CS1-CS4) and not shared across all PCNSL tumors. For example, PCNSLs (MCD/C5 DLBCLs) are mainly characterized by mutations leading to constitutive NF-κB activation but this was only observed, transcriptionally, for CS4 and CS3; however, the outcome for these clusters is very different mainly due to meningeal infiltration to the CSF probability, TME, *HIST1H1E* mutations, and B-cell differentiation program differences. Hence, we propose different tailored treatments according to the pathway activation of each CS, suggesting, for example, that CS4 might be more likely to respond to ICI treatment, and BTK/JAK1 inhibitors. Our initial multi-omic analysis along with our DepMap/Genomics of Drug Sensitivity in Cancer (GDSC) analysis on DLBCL cell lines suggest the CS3 subtype to be particularly more susceptible to BTK inhibitors, while the CS2 subtype to lenalidomide, and the CS1 subtype to PI3K inhibitors (Figure 7). When analyzing the TME and epigenetic features across subtypes, we found similar results to those previously described by Kotlov et al. in DLBCLs where the hypermethylated subtypes (CS1/CS2) had a depleted TME.[50]

On top of this and given the importance of routine clinical implementation, we propose RBraLymP (RNA-based Brain Lymphoma Profiler), which uses gene expression data from either FFPE or FF tissue, to identify the PCNSL molecular subtypes associated with multi-omic features. The RBraLymP algorithm is publicly accessible at https://github.com/iS4i4S/PCNSL-RBraLymP such that existing and new therapy efforts can be directed to the most appropriate patients.

Our study presents various limitations. Identification of subtypes was achieved using relevant features obtained from different static bulk-omics data, and hence does not respond to tumor intra-heterogeneity or single-cell's cross-talk as other approaches like single-cell RNA-seq or spatial transcriptomics would potentially do. Another limitation would be the absence of a validation PCNSL cohort having

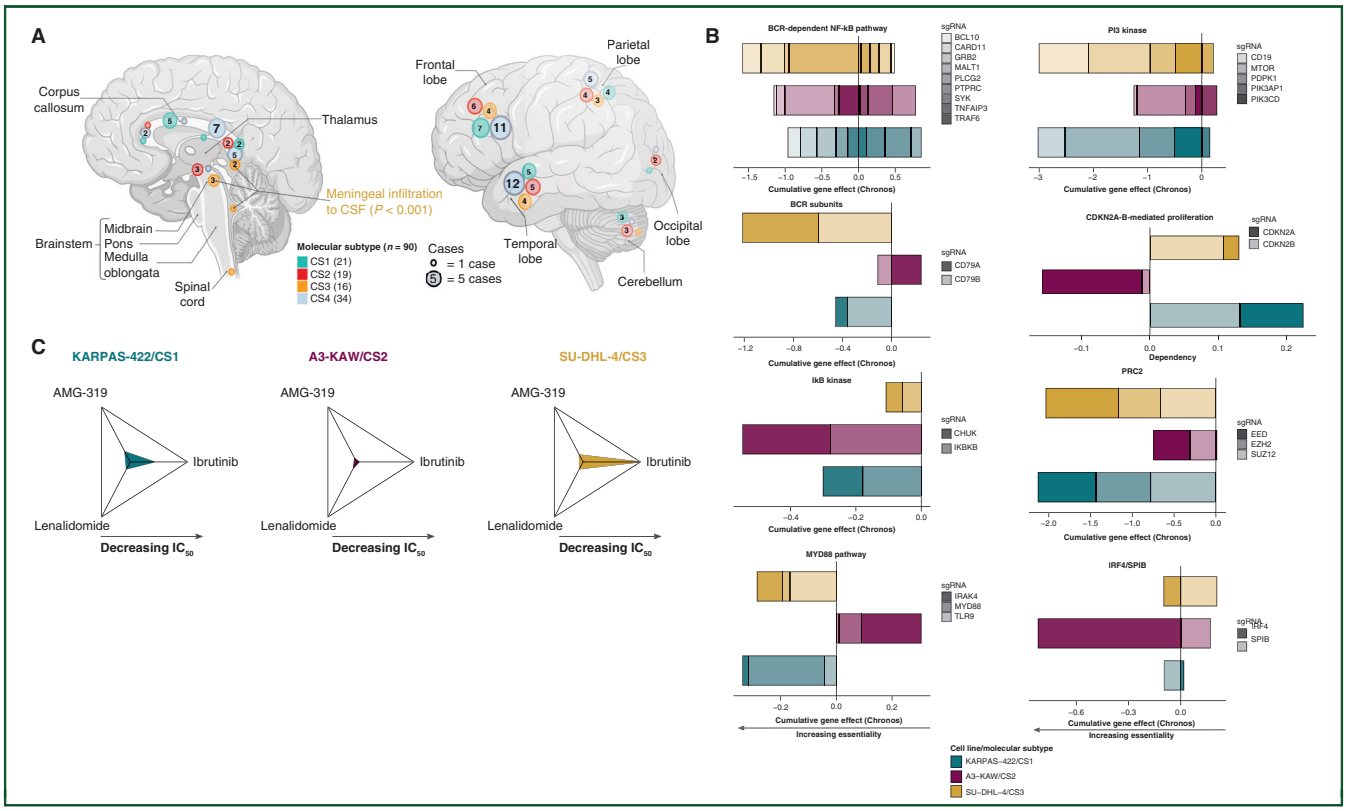

Figure 6. Tumor location distinctions and functional genomics of the PCNSL subtypes. (A) The tumor location of 90 PCNSLs (FFPE cohort) in the human central nervous system grouped by molecular subtype is shown, where the number of cases is indicated within the circles. Tumors occurring in midline locations are depicted in the sagittal view (left panel); meanwhile, tumors occurring in the cerebral and cerebellar hemispheres are depicted in the exterior view (right panel). *P* value refers to a one-sided Fisher test (event within cluster versus outside cluster). (B) The cumulative gene effect scores (Chronos CRISPR loss-of-function screens) on different lymphomagenic pathways are shown. (C) Radar plots of drug sensitivity across different CS PCNSL-related cell lines are shown. Shown values are the inverse of the IC_{50}.
BCR, B-cell receptor; CS, cluster; CSF, cerebrospinal fluid; FFPE, formalin-fixed, paraffin-embedded; IC_{50}, concentration that causes 50% inhibition of growth; NF-κB, nuclear factor-κB; PCNSL, primary central nervous system lymphoma; PRC2, Polycomb repressive complex 2.

the same data types as the discovery cohort; however, this is at the moment difficult given the rarity of the disease. Finally, our analysis on DLBCL cell lines would benefit from further validations using PCNSL-derived cell lines, organoids, or patient-derived xenograft models with drug sensitivity and RNA-seq data.

In summary, our multi-omics analysis builds on the current classification of DLBCL by the addition of the molecular heterogeneity within PCNSL that may inform on its pathogenesis. Our study discovered a link between genetic and neoplastic signaling pathways, pointing to potential treatment targets. Selecting treatment for PCNSL based on individual genetic changes is not desirable from the standpoint of precision medicine, as it is likely that combinations of genetic aberrations influence therapeutic response. The genetic subgroups we define could serve as a conceptual foundation for developing targeted therapeutic approaches for these poorly understood and with high mortality malignancies.

ACKNOWLEDGEMENTS

We thank Yannick Marie and Emeline Mundwiller (Paris Brain Institute), Lucile Armenoult, and Mira Ayadi (Carte d'Identité de Tumeurs, Ligue contre le Cancer) for the help with exome, RNA sequencing, and bisulfite DNA sequencing libraries' preparation; Silvia Duran for helpful discussions; Figure 7 was created with biorender.com and all the clinicians, pathologists, researchers, and patients associated with the French National 'Lymphome Oculo-Cérébral' included in Rare cancers of the central nervous system, RENOCLIP-LOC Network, approved by the French National Institute of Cancer (INCa).

FUNDING

This work was supported by a grant from Investissements d'avenir and by the grant INCa-DGOS-Inserm_12560 of the SiRIC CURAMUS [grant number INCa-DGOS-Inserm_12560], the program 'investissements d'avenir' [grant number ANR-10-IAIHU-06], PRT-K/INCa grant LOC-model reference 2017-1-RT-04 (no grant number), DRCI de l'APHP (no grant number), CRC2013_105_R1/projet Tri LOC, BETPSY project, overseen by the French National Research Agency, as part of the second 'Investissements d'Avenir' program [grant number ANR-18-RHUS-0012], Foundation RAM active investments, ARTC foundation (no grant number), an unrestricted grant from Bristol Myers Squibb (BMS) [grant number RDON06618], ICGex project (no grant number), and IDeATIon project with an unrestricted grant from MSD Avenir (no grant number).

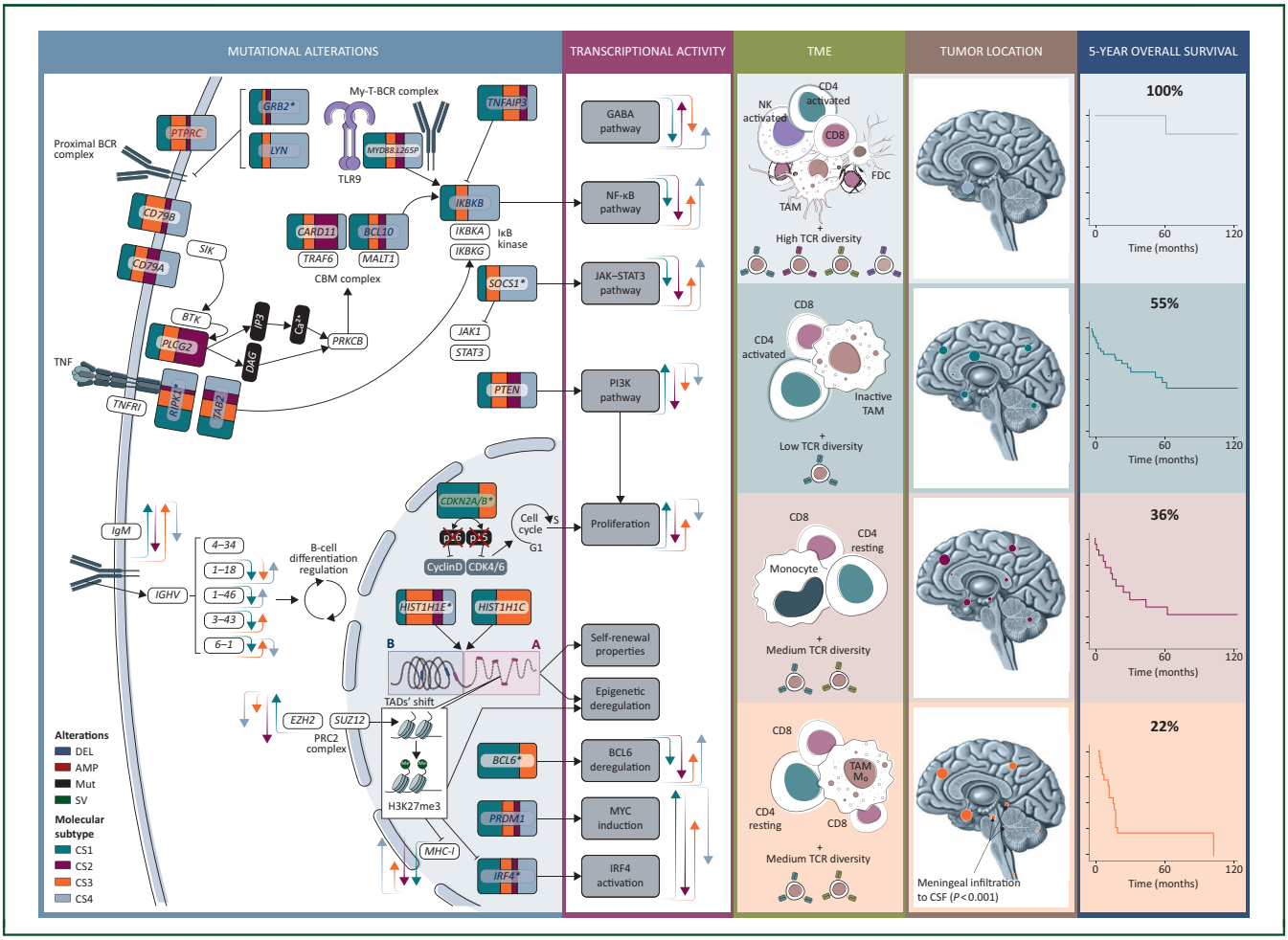

Figure 7. Pathogenetic and therapeutic implications of the PCNSL subtypes. Shown is a schematic representation summarizing the major molecular findings and proposed potential therapeutic targets. Contribution of each molecular subtype to the indicated alteration where color bar width indicates the prevalence of each subtype (Mutational alterations section). Asterisks indicate if a genetic alteration is enriched in any CS subtype (related to Figure 3A). Arrows indicate transcriptional gene signature activity where the height indicates the relative up- or negative regulation (according to Figure 4A).
AMP, amplification; CS, cluster; DEL, deletion; FDC, follicular dendritic cells; GABA, γ-aminobutyric acid; JAK—STAT3, Janus kinase—signal transducer and activator of transcription 3; Mut, mutation; NF-κB, nuclear factor-κB; NK, natural killer; PCNSL, primary central nervous system lymphoma; SV, fusion transcripts; TAM, tumor-associated macrophages; TCR, T-cell receptor.

DISCLOSURE

GA reports grants from Biogen, Novartis, Roche, Sanofi, Abbvie, Pfizer, and CSL Behring, outside the submitted work. AA reports research grant with an unrestricted grant from Bristol Myers Squibb (BMS). All other authors have declared no conflicts of interest.

REFERENCES

1. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med*. 2018;24(5):679-690.
2. Schmitz R, Wright GW, Huang DW, et al. Genetics and pathogenesis of diffuse large B-cell lymphoma. *N Engl J Med*. 2018;378(15):1396-1407.
3. Yoshida K, Shiraishi Y, Chiba K, et al. Whole-genome sequencing of primary central nervous system lymphoma and diffuse large B-cell lymphoma. *Blood*. 2016;128(22). 4112-4112.
4. Sehn LH, Salles G. Diffuse large B-cell lymphoma. *N Engl J Med*. 2021;384(9):842-858.
5. Deckert M, Batchelor TT, Ferry JA, et al. Primary diffuse large B cell lymphoma of the CNS. In: *World Health Organization Classification of Tumours of the Central Nervous System*.. 5th ed. Lyon, France: International Agency for Research on Cancer; 2021:351-355.
6. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
7. Houillier C, Soussain C, Ghesquières H, et al. Management and outcome of primary CNS lymphoma in the modern era: an LOC network study. *Neurology*. 2020;94(10):e1027-e1039.
8. Zhou Y, Liu W, Xu Z, et al. Analysis of genomic alteration in primary central nervous system lymphoma and the expression of some related genes. *Neoplasia*. 2018;20(10):1059-1069.
9. Swerdlow SH, Campo E, Harris NL, et al. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*., Vol. 2. Lyon, France: International Agency for Research on Cancer; 2008.
10. Braggio E, Van Wier S, Ojha J, et al. Genome-wide analysis uncovers novel recurrent alterations in primary central nervous system lymphomas. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2015;21(17):3986-3994.
11. Bonzheim I, Giese S, Deuter C, et al. High frequency of MYD88 mutations in vitreoretinal B-cell lymphoma: a valuable tool to improve diagnostic yield of vitreous aspirates. *Blood*. 2015;126(1):76-79.
12. Bruno A, Boisselier B, Labreche K, et al. Mutational analysis of primary central nervous system lymphoma. *Oncotarget*. 2014;5(13):5065-5075.
13. Kraan W, van Keimpema M, Horlings HM, et al. High prevalence of oncogenic MYD88 and CD79B mutations in primary testicular diffuse large B-cell lymphoma. *Leukemia*. 2014;28(3):719-720.

14. Ngo VN, Young RM, Schmitz R, et al. Oncogenically active MYD88 mutations in human lymphoma. *Nature*. 2011;470(7332):115-119.
15. Vater I, Montesinos-Rongen M, Schlesner M, et al. The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. *Leukemia*. 2015;29(3):677-685.
16. Kersten MJ, Kraan W, Doorduijn J, et al. Diffuse large B cell lymphomas relapsing in the CNS lack oncogenic MYD88 and CD79B mutations. *Blood Cancer J*. 2014;4(12). e266-e266.
17. Brunn A, Nagel I, Montesinos-Rongen M, et al. Frequent triple-hit expression of MYC, BCL2, and BCL6 in primary lymphoma of the central nervous system and absence of a favorable MYClowBCL2low subgroup may underlie the inferior prognosis as compared to systemic diffuse large B cell lymphomas. *Acta Neuropathol (Berl)*. 2013;126(4):603-605.
18. Chapuy B, Roemer MGM, Stewart C, et al. Targetable genetic features of primary testicular and primary central nervous system lymphomas. *Blood*. 2016;127(7):869-881.
19. Fukumura K, Kawazu M, Kojima S, et al. Genomic characterization of primary central nervous system lymphoma. *Acta Neuropathol (Berl)*. 2016;131(6):865-875.
20. Cady FM, O'Neill BP, Law ME, et al. Del(6)(q22) and *BCL6* rearrangements in primary CNS lymphoma are indicators of an aggressive clinical course. *J Clin Oncol*. 2008;26(29):4814-4819.
21. Radke J, Ishaque N, Koll R, et al. The genomic and transcriptional landscape of primary central nervous system lymphoma. *Nat Commun*. 2022;13(1):2558.
22. Alame M, Cornillot E, Cacheux V, et al. The immune contexture of primary central nervous system diffuse large B cell lymphoma associates with patient survival and specific cell signaling. *Theranostics*. 2021;11(8):3565-3579.
23. Marcelis L, Antoranz A, Delsupehe A-M, et al. In-depth characterization of the tumor microenvironment in central nervous system lymphoma reveals implications for immune-checkpoint therapy. *Cancer Immunol Immunother.* 2020;69(9):1751-1766.
24. Four M, Cacheux V, Tempier A, et al. PD1 and PDL1 expression in primary central nervous system diffuse large B-cell lymphoma are frequent and expression of PD1 predicts poor survival: PD1 expression in lymphoma. *Hematol Oncol*. 2017;35(4):487-496.
25. Komohara Y, Horlad H, Ohnishi K, et al. M2 macrophage/microglial cells induce activation of STAT3 in primary central nervous system lymphoma. *J Clin Exp Hematop*. 2011;51(2):93-99.
26. Kumari N, Krishnani N, Rawat A, Agarwal V, Lal P. Primary central nervous system lymphoma: prognostication as per international extranodal lymphoma study group score and reactive CD3 collar. *J Postgrad Med*. 2009;55(4):247.
27. Ponzoni M, Berger F, Chassagne-Clement C, et al. Reactive perivascular T-cell infiltrate predicts survival in primary central nervous system B-cell lymphomas. *Br J Haematol*. 2007;138(3):316-323.
28. Miyasato Y, Takashima Y, Takeya H, et al. The expression of PD-1 ligands and IDO1 by macrophage/microglia in primary central nervous system lymphoma. *J Clin Exp Hematop*. 2018;58(2):95-101.
29. Sasayama T, Tanaka K, Mizowaki T, et al. Tumor-associated macrophages associate with cerebrospinal fluid interleukin-10 and survival in primary central nervous system lymphoma (PCNSL): tumor-associated macrophages in PCNSL. *Brain Pathol*. 2016;26(4):479-487.
30. Cho H, Kim SH, Kim S-J, et al. Programmed cell death 1 expression is associated with inferior survival in patients with primary central nervous system lymphoma. *Oncotarget*. 2017;8(50):87317-87328.
31. Wilms T, Khan G, Coates PJ, et al. No evidence for the presence of Epstein-Barr virus in squamous cell carcinoma of the mobile tongue. *PLoS One*. 2017;12(9):e0184201.
32. Mo Q, Shen R, Guo C, Vannucci M, Chan KS, Hilsenbeck SG. A fully Bayesian latent variable model for integrative clustering analysis of multi-type omics data. *Biostatistics*. 2018;19(1):71-86.
33. Meng C, Helm D, Frejno M, Kuster B. moCluster: identifying joint patterns across multiple omics data sets. *J Proteome Res*. 2016;15(3):755-765.
34. Ramazzotti D, Lal A, Wang B, Batzoglou S, Sidow A. Multi-omic tumor data reveal diversity of molecular mechanisms that correlate with survival. *Nat Commun*. 2018;9(1):4453.
35. Chalise P, Fridley BL. Integrative clustering of multi-level 'omic data based on non-negative matrix factorization algorithm. *PLoS One*. 2017;12(5):e0176278.
36. Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. *Mach Learn*. 2003;52(1/2):91-118.
37. Hoadley KA, Yau C, Wolf DM, et al. Multiplatform analysis of 12 cancer types reveals molecular classification within and across tissues of origin. *Cell*. 2014;158(4):929-944.
38. Rappoport N, Shamir R. NEMO: cancer subtyping by integration of partial multi-omic data. *Bioinformatics*. 2019;35(18):3348-3356.
39. Nguyen H, Shrestha S, Draghici S, Nguyen T. PINSPlus: a tool for tumor subtype discovery in integrated genomic data. *Bioinformatics*. 2019;35(16):2843-2846.
40. Wang B, Mezlini AM, Demir F, et al. Similarity network fusion for aggregating data types on a genomic scale. *Nat Methods*. 2014;11(3):333-337.
41. Wu D, Wang D, Zhang MQ, Gu J. Fast dimension reduction and integrative clustering of multi-omics data using low-rank approximation: application to cancer molecular classification. *BMC Genomics*. 2015;16(1):1022.
42. Lu X, Meng J, Zhou Y, Jiang L, Yan F. *MOVICS* : an R package for multi-omics integration and visualization in cancer subtyping. *Bioinformatics*. 2021;36(22-23):5539-5541.
43. Phelan JD, Young RM, Webster DE, et al. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. *Nature*. 2018;560(7718):387-391.
44. Martincorena I, Raine KM, Gerstung M, et al. Universal patterns of selection in cancer and somatic tissues. *Cell*. 2017;171(5):1029-1041.e21.
45. Wright GW, Huang DW, Phelan JD, et al. A probabilistic classification tool for genetic subtypes of diffuse large B cell lymphoma with therapeutic implications. *Cancer Cell*. 2020;37(4):551-568.e14.
46. Pasqualucci L, Bhagat G, Jankovic M, et al. AID is required for germinal center—derived lymphomagenesis. *Nat Genet*. 2008;40(1):108-112.
47. Goeman JJL. $_1$ penalized estimation in the cox proportional hazards model. *Biom J*. 2010;52(1):70-84.
48. Abrey LE, Ben-Porat L, Panageas KS, et al. Primary central nervous system lymphoma: the memorial Sloan-Kettering cancer center prognostic model. *J Clin Oncol*. 2006;24(36):5711-5715.
49. Yusufova N, Kloetgen A, Teater M, et al. Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. *Nature*. 2021;589(7841):299-305.
50. Kotlov N, Bagaev A, Revuelta MV, et al. Clinical and biological subtypes of b-cell lymphoma revealed by microenvironmental signatures. *Cancer Discov.* 2021;11(6):1468-1489.
51. Fangazio M, Ladewig E, Gomez K, et al. Genetic mechanisms of HLA-I loss and immune escape in diffuse large B cell lymphoma. *Proc Natl Acad Sci*. 2021;118(22):e2104504118.
52. Schubert M, Klinger B, Klünemann M, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun*. 2018;9(1):20.
53. Holland CH, Szalai B, Saez-Rodriguez J. Transfer of regulatory knowledge from human to mouse for functional genomics analysis. *Biochim Biophys Acta Gene Regul Mech*. 2020;1863(6):194431.
54. Wei B, Liu Z, Fan Y, et al. Analysis of cellular heterogeneity in immune microenvironment of primary central nervous system lymphoma by single-cell sequencing. *Front Oncol*. 2021;11:683007.
55. Steen CB, Luca BA, Esfahani MS, et al. The landscape of tumor cell states and ecosystems in diffuse large B cell lymphoma. *Cancer Cell*. 2021;39(10):1422-1437.e10.
56. Holmes AB, Corinaldesi C, Shen Q, et al. Single-cell analysis of germinal-center B cells informs on lymphoma cell of origin and outcome. *J Exp Med*. 2020;217(10):e20200483.

57. Milpied P, Cervera-Marzal I, Mollichella M-L, et al. Human germinal center transcriptional programs are de-synchronized in B cell lymphoma. *Nat Immunol*. 2018;19(9):1013-1024.
58. Young RM, Wu T, Schmitz R, et al. Survival of human lymphoma cells requires B-cell receptor engagement by self-antigens. *Proc Natl Acad Sci*. 2015;112(44):13447-13454.
59. Zhang B, Vogelzang A, Miyajima M, et al. B cell-derived GABA elicits IL-10^{+} macrophages to limit anti-tumour immunity. *Nature*. 2021;599:471-476.
60. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell*. 2015;160(1-2):48-61.
61. Jalkanen S, Salmi M. Lymphatic endothelial cells of the lymph node. *Nat Rev Immunol*. 2020;20(9):566-578.
62. Louveau A, Smirnov I, Keyes TJ, et al. Structural and functional features of central nervous system lymphatic vessels. *Nature*. 2015;523(7560):337-341.
63. Alvarez MJ, Shen Y, Giorgi FM, et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. *Nat Genet*. 2016;48(8):838-847.
64. Dettman EJ, Simko SJ, Ayanga B, et al. Prdm14 initiates lymphoblastic leukemia after expanding a population of cells resembling common lymphoid progenitors. *Oncogene*. 2011;30(25):2859-2873.
65. Chu LC, Eberhart CG, Grossman SA, Herman JG. Epigenetic silencing of multiple genes in primary CNS lymphoma. *Int J Cancer*. 2006;119(10):2487-2491.
66. Wiedemann GM, Knott MML, Vetter VK, et al. Cancer cell-derived IL-1α induces CCL22 and the recruitment of regulatory T cells. *OncoImmunology*. 2016;5(9):e1175794.
67. Nakamura T, Yamashita S, Fukumura K, et al. Genome-wide DNA methylation profiling identifies primary central nervous system lymphoma as a distinct entity different from systemic diffuse large B-cell lymphoma. *Acta Neuropathol (Berl)*. 2017;133(2):321-324.
68. Jang IK, Cronshaw DG, Xie L-K, et al. Growth-factor receptor-bound protein-2 (Grb2) signaling in B cells controls lymphoid follicle organization and germinal center reaction. *Proc Natl Acad Sci*. 2011;108(19):7926-7931.
69. Linossi EM, Nicholson SE. Kinase inhibition, competitive binding and proteasomal degradation: resolving the molecular function of the suppressor of cytokine signaling (SOCS) proteins. *Immunol Rev*. 2015;266(1):123-133.
70. Phillips TJ, Forero-Torres A, Sher T, et al. Phase 1 study of the PI3Kδ inhibitor INCB040093 ± JAK1 inhibitor itacitinib in relapsed/refractory B-cell lymphoma. *Blood*. 2018;132(3):293-306.
71. Dersh D, Phelan JD, Gumina ME, et al. Genome-wide screens identify lineage- and tumor-specific genes modulating MHC-I- and MHC-II-restricted immunosurveillance of human lymphomas. *Immunity*. 2021;54(1):116-131.e10.
72. Tsherniak A, Vazquez F, Montgomery PG, et al. Defining a cancer dependency map. *Cell*. 2017;170(3):564-576.e16.
73. Dempster JM, Boyle I, Vazquez F, et al. Chronos: a cell population dynamics model of CRISPR experiments that improves inference of gene fitness effects. *Genome Biol*. 2021;22(1):343.
74. Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res*. 2012;41(D1):D955-D961.