Molecular and clinical diversity in primary central nervous system lymphoma: a LOC Network study

Agusti Alentorn (agusti.alentorn@icm-institute.org)
Paris Brain Institute

Isaias Hernández-Verdin
Paris Brain Institute

Eva Kirasic
Paris Brain Institute

Kirsty Wienand
Göttingen Comprehensive Cancer Center, Georg-August-University https://orcid.org/0000-0001-9548-4420

Sandrine Eimer
CHU Bordeaux

Hugues Loiseau
Centre Hospitalier Universitaire de Bordeaux

Audrey Rousseau
CHU Angers

Jérôme Paillassa
CHU Angers

Guido Ahle
CH Colmar

Felix Lerintiu
CH Colmar

Emmanuelle Uro-Coste
Department of Pathology, CHU de Toulouse, IUC-Oncopole

Lucie Oberic
Department of hematology, IUC Toulouse Oncopole

Dominique Figarella-Branger
Assistance Publique Hopitaux de Marseille

Olivier Chinot
CHU La Timone

Guillaume Gauchotte
CHRU Nancy

Luc Taillandier
CHRU Nancy

Jean-Pierre Marolleau
CHU Amiens
Marc Polivka
Lariboisière Hospital
Clovis Adam
CHU Bicêtre
Renata Ursu
Hôpital Saint Louis
Anna Schimitt
Institut Bergonié Hospital
Noemie Barillot
Paris Brain Institute
Lucia Nichelli
Groupe Hospitalier Pitié Salpêtrière
Fernando Lozano-Sánchez
Groupe Hospitalier Pitié Salpêtrière
Maria-José Ibañez-Juliá
CH Perpignan
Matthieu Peyre
Hôpital Pitié Salpêtrière
Bertrand Mathon
Groupe Hospitalier Pitié Salpêtrière
Yah-se Abada
Groupe Hospitalier Pitié Salpêtrière
Frederic Charlotte
APHP https://orcid.org/0000-0003-2803-6606
Frédéric Davi
Hematology Department, Hôpital de la Pitié-Salpêtrière, AP-HP
Chip Stewart
Broad Institute of Harvard and MIT
Aurélien de Reyniès
Centre de Recherche des Cordeliers
Sylvain Choquet
Groupe Hospitalier Pitié Salpêtrière https://orcid.org/0000-0002-7791-0470
Carole Soussain
Institut Curie
Caroline Houillier
Groupe Hospitalier Pitié Salpêtrière
Bjoern Chapuy

Universitätsmedizin Berlin

Khe Hoang-Xuan

Sorbonne Universités, UPMC Univ Paris 06, UMR S 1127

Article

Keywords:

Posted Date: March 17th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1438980/v1

Abstract

Primary central nervous system lymphoma (PCNSL) is a distinct extranodal lymphoma presenting with limited stage disease but variable response rates to treatment despite homogenous pathological presentation. The likely underlying molecular heterogeneity and its clinical impact is poorly understood.

We performed a comprehensive genome-wide analysis of 147 PCNSL from fresh-frozen tumor tissue from immunocompetent, treatment naïve PCNSL patients, employing whole-exome sequencing, assessment of somatic copy number alterations and DNA methylation, and RNA expression. These data were integrated and correlated with the clinico-radiological characteristics and outcomes of the patients. We validated our results in an independent series of 93 PCNSL formalin-fixed, paraffin-embedded (FFPE) samples.

Consensus clustering of multi-omics data identified four robust, non-overlapping, prognostically significant clusters (CS) within PCNSL. The CS1 group, characterized by high proliferation and Polycomb Repressive Complex 2 (PRC2) complex activity had an intermediate outcome between CS2/CS3 and CS4. Patients who had PCNSL with an “immune-hot” (CS4) profile had the most favorable clinical outcome. In contrast, patients with the immune-cold hypermethylated CS2 and the heterogenous-immune CS3 groups had a poor prognosis. Nearly all PCNSL patients with meningeal infiltration harbored *HIST1H1E* mutations, enriched in the CS3 group. The integrated analysis suggests that the CS4 group may be more susceptible to immunotherapy.

The integration of genome-wide data from multi-omics data revealed four molecular patterns in PCNSL with a distinctive prognostic impact that significantly improved the current clinical stratification. This molecular classification using FFPE samples facilitates routine use in clinical practice and provides potential precision-medicine strategies in PCNSL.

Introduction

Primary central nervous system lymphoma (PCNSL) is a rare subtype of extranodal non-Hodgkin lymphoma that often presents as diffuse large B cell lymphoma (DLBCL), but has a less favorable prognosis compared to its systemic counterpart and has been proved to be molecularly a different biological entity[1–4]. 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[5,6].

Biologically, initial studies revealed that PCNSL arise from late germinal center (GC) B-cell stages and has constitutive NF-κB activity driven by mutations in genes of the B-cell receptor (BCR) pathway or the toll-like receptor (TLR) pathway, including *CD79B*, *MYD88* and *CARD11*[7]. Recently, DLBCL has been divided in different molecular clusters and PCNSL have been related to the “MCD” or Cluster 5 (C5) DLBCL, both converging in the presence of frequent *MYD88*^{L265P}, *CD79B*, *PIM1*, *BTG2* mutations, copy gains of 3q12.3,

9p24.1 (PD-L1/PD-L2), and copy losses of 6p21-22 (HLA locus), 6q21, and 9p21.3 (*CDKN2A* biallelic loss)[1,2,7,8]. Currently, PCNSL heterogeneity has not been properly addressed mainly due to the lack of multi-omic data integration and the limited number of patients[8].

Here, we performed an integrative analysis of mutations, somatic copy-number alterations (SCNA), fusions, gene expression, TCR/BCR clonotypes, tumor microenvironment (TME), methylation, tumor localization, and clinical data to identify molecular subtypes of PCNSL with clinically distinct behaviors. Additionally, to facilitate routine clinical implementation, we developed an algorithm that uses 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.

Methods

Patients

A total of 147 FF (discovery cohort) and 93 FFPE (validation cohort) tumor samples from immunocompetent Epstein-Barr negative PCNSL were recollected from different French hospitals (see Table S1 in Supplementary Appendix 1) after written informed consent and ethics approval (Pitié Salpêtrière Hospital ethics committee) were obtained. All the tumors were newly diagnosed PCNSL, pathologically confirmed according to the World Health Organization classification and were treated with standard HD-MTX based chemotherapy regimen according to "*Lymphome oculo-cérébral*" (LOC) network recommendations[5].

Sequencing platforms and multi-omic data integration for PCNSL molecular subtyping

We performed, on the FF cohort, exome sequencing (n=115) to call mutations and obtain CNA events, RNA sequencing (n=123) to analyze gene expression, immune cell proportions, TCR/BCR clonotypes, and fusion transcripts, and DNA methylation profiling (n=64). High robust clustering was obtained by consensus clustering resulting from 10 different multi-omics clustering algorithms that are integrated in the R package "MOVICS"[9].

Statistical analyses

Differences in proportions and binary/categorical variables were calculated from Fisher's exact test. Kruskal-Wallis test was used to test for a difference in distribution between three or more independent groups, and Mann Whitney U test was used for differences in distributions between two population groups unless otherwise noted. Overall Survival (OS) analysis was assessed using log-rank Kaplan-Meier curves and multivariate Cox proportional hazards regression modeling. See Supplementary Appendix 1 for full details.

Results

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

We performed a cluster of clusters analysis using six levels of omic information (Figure 1A and Fig. S1-S3) to identify four PCNSL subtypes (CS1 to CS4) that display different clinical outcomes in OS (Global log-rank $p < 0.001$, Figure 1B). Patients in CS4 had the longest OS (median=66.8 months; 95% confidence interval [CI]=19.8-67.2) and lived significantly longer than those in both clusters CS2 (median=18 months; $CI_{95\%}$=8.3-53.4; p=0.024) or CS3 (median=13.8 months; $CI_{95\%}$=6.1-16.7; p=0.003), and slightly longer, but not significantly, to those in CS1 (median = 26.2 months; $CI_{95\%}$=13.3-63.9; p=0.094). Additionally, these observations remained significant after adjusting by age and Karnofsky Performance Status (KPS) in Cox proportional hazard ratio multivariate models (Fig. S4A). Interestingly, CS4 was independently associated with a better response when considering progression free survival in univariate and multivariate models (Fig. S4B-C). Finally, we did not observe significant differences in the median number of predicted immunogenic neoantigens (p=0.44, Table S2 and Fig. S5).

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 PCNSL CS (Table S3). We obtained a Cohen's kappa coefficient of 0.90 ($p<0.001$) when evaluating the accuracy of correctly assigning patients from the multi-omic cohort. Additionally, when expanding to the FF-RNA complete set or when using the FFPE cohort, we validated the prognostic importance of the molecular subtypes CS1-CS4 for clinical outcome (Global $p<0.001$) in univariate and multivariate models (Figure 1C and Fig. S6-S12).

Next, we evaluated the contribution of each omic-level data to outcome prediction models by using Harrell's concordance index (C-index)[10]. A C-index of 0.60 (0.56-0.65 at $CI_{95\%}$) in FF and 0.71 (0.68-0.74 at $CI_{95\%}$) 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[11]. When adding different omic-data to the FF cohort modeling, we observed higher predictive power using mRNA expression compared to the other omic data (C-index=0.91±0.02 at $CI_{95\%}$). We further validated these observations in the FFPE cohort obtaining a C-index of 0.83 (0.80-0.85 at $CI_{95\%}$) and 0.93 (0.91-0.95 at $CI_{95\%}$) when adding the mRNA level or the TME and RNA levels to the model, respectively (Figure 1D). Altogether, these results show that RNA-seq data from FFPE or FF tissue can be used to correctly identify PCNSL subgroups and significantly increase accuracy of outcome prediction.

Mutational landscape of PCNSL

We identified 32,544 mutations in the 115 PCNSL samples analyzed (median=3.23 mutations/Mb; range=0.02-85.49; Table S4 and Fig. S13). We applied the dNdScv[12] algorithm to identify driver mutations identifying 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%), *TBL1XR1*

(25%), *KLHL14* (25%), *CARD11* (22%), and *HIST1H1E* (18%) which are involved in BCR and TLR-mediated NF-kB signaling, antigen presentation, cell-cycle, histone modification and B-cell differentiation regulation [7,8,13] (Figure 2A, Table S5 and Fig. S14). Moreover, we detected canonical activation-induced cytidine deaminase (c-AID) off-target mutations and found they represent 7.9% (6.8-8.5% at 95% CI) of SNV mutations and fall within driver genes like *PIM1* (47%), *CD79B* (10%), *IRF4* (9%), and *HIST1H1E* (6%) (Table S6, Fig. S15-S16). Interestingly, clonal mutations are significantly more driven by both c-AID and non c-AID (Cosmic signature SBS9) mutational processes than sublonal mutations (p=0.007 and 0.018, respectively), hence reflecting the importance of AID activity in the early stages of PCNSL tumorigenesis (Fig. S17-18).

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 in primary systemic DLBCL (Figure 2B)[1,7]. Furthermore, we found additional, not previously 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* (FC= 1.25), *IRF4* (FC=-1.03) and *GGTLC2* (FC=-1.76, FDR q-value<0.1), respectively (Table S7, Fig. S19-S22).

Distinct genetic signatures within PCNSL subtypes and systemic DLBCL

Afterwards, we aimed to characterize the differences in genetic alterations across groups for each mutation, focal SCNA, and fusions. The CS4 cluster presents ten enriched events that included mutations in *SOCS1,* which is a negative regulator of the JAK-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 NF-κB pathway (*RIPK1*/6p25.3 deletion), B-cell differentiation (*IRF4*/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[1,2,7,14,15]. The CS2 cluster did not present any genomic characteristic events. Furthermore, most of these distinctive events were inferred as early events (clonal) in tumorigenesis like *IRF4* and *BCL6* in CS1 (Figure 2C, Table S8-S9). Of note, most of these mutations were not observed in the clusters previously defined in systemic DLBCL (e.g., 9p11.2 del; Figure 2D)[1].

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

We analyzed the expression of different previously curated gene signatures[14,16] (see Methods). CS1 was characterized by the upregulation of PI3K, glycolytic activity, and cell proliferation signatures; additionally, it presented hyperactivation of the Polycomb Repressive Complex 2 (PRC2) complex which reduces MHC-I expression, through histone methylation (Figure 3A, p<0.05, and Fig. S23-S25)[17]. Moreover, p53 activity was enriched in the CS2 cluster[16,18]. Interestingly, even though all clusters presented mutations within the

NF-κB pathway, it was transcriptionally active only in clusters CS3 and CS4. Additionally, MAPK and JAK-STAT pathways were upregulated in those clusters, respectively (Figure 3A).

Regarding B-cell differentiation programs, CS1 expressed a mixture of GC cells which is consistent with the 6p25.3-19q13.12 deletions, and *BCL6* mutations (Figure 2C). On the other hand, cluster CS4 presents an enrichment in terminally differentiated plasma cells that goes in line with *BCL6* downregulation, the absence of *MYC* induction, and *BCL6* mutations. The most heterogeneous cluster was the CS3, presenting features of both GC and mature B-cells (plasma cells and memory B-cells). Intriguingly, the cluster CS2 did not present any B-cell stage enrichment but instead a lymphatic endothelial cell (LEC) gene signature (Figure 3A and Fig. S26-27).

Then, we aimed to describe the TME differences between subtypes by using CIBERSORTx derived immune deconvolution and B-cell lymphoma specific TME gene signatures[19]. CS1 cluster is immunologically "neutral" meanwhile the CS2, which is immunologically depleted, exhibits expression of vascular endothelial cells (VEC), memory resting $CD4^+$ T-cells, monocytes, and activation of GABA synthesis, which has been recently linked to B-cells that inhibit $CD8^+$ T-cells' killer function and promote monocyte differentiation into anti-inflammatory macrophages[20]. The CS4 cluster has a hot-inflammatory TME due to the presence of active $CD8^+$ T-cells and NK cells (with high cytolytic activity score)[21]. Conversely, heterogeneity was again observed for the CS3 subtype, being only inactivated macrophages M0 more significantly enriched (Figure 3A, Fig. S28-S41, and Table S9-S10).

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 on the molecular subtypes. We observed no brain lobe preference between PCNSL subgroups but, in general, tumors arose less in the occipital lobe (4/90 cases versus 86/90, $p<0.001$). In addition, CS4 tumors arose more in the isthmus of the corpus callosum (7/34 cases versus 0/56, $p<0.001$). Conversely, CS2/CS3 were more frequent in the brainstem (4/16 and 3/19 cases versus 1/55, $p=0.005$), when compared to the other clusters. Strikingly, we found no association with tumor size nor multiple lesions. However, meningeal infiltration of the cerebrospinal fluid (CSF) was only found within CS3 tumors (6/16 cases versus 0/74, $p<0.001$, Figure 3B and Table S11).

Epigenetic attributes across PCNSL subtypes

We proceeded to investigate epigenetic differences among subtypes (n=64). The CS2 displayed higher hypermethylation globally ($p=0.006$, Fig. S42-S43). Interestingly, GO analyses on differentially methylated promoters revealed B-cell differentiation programs to be hypomethylated in CS1 but hypermethylated in CS2; while interleukin-1 was hypermethylated in CS4 (Figure 4A and Table S12). Genomic region enrichment analysis on hypermethylated promoters identified strong enrichment of binding sites for the histone/chromatin proteins H3K27me3 and E2H2 in CS1, and NF-κB, IRF4, and BCL6 in CS2 (Figure 4B, Fig. S44 and Table S13).

From multi-omics to potential therapeutic targets

To generate an explanatory bridge between the different multi-omic layers and ultimately potential therapeutic targets across subtypes, we integrated all multi-omic data and evaluated their contribution to specific pathways. Even though the hallmark PCNSL alterations targeting My-T-BCR protein supercomplex, *CD79A/B* BCR subunits, *TNFAIP3*, *RIPK1*, *TAB2*, and the CBM (*CARD11-BCL10-MALT1*) complex were relatively constant across subgroups, the NF-κB hyperactive group (CS4) exhibited more *GRB2/ LYN* deletions and had no *PLCG2* mutations, which represses the BCR complex and activates the CBM complex, respectively. Furthermore, NF-κB activity could not be explained by self-antigen-dependent chronic active BCR signaling upregulation since $IgV_{H4\text{-}34}$ expression was similar across groups (Figure 5) [14,22,23]. These observations suggest that CS4 and CS3 may be more sensitive to BTK inhibitors (e.g., ibrutinib). The CS4 cluster also presented high JAK-STAT activity and mutated *SOCS1* (a JAK1 repressor), making it potentially responsive to JAK1 inhibitors (e.g., INCB040093)[24,25]. Regarding antigen presentation-related genes, we observed only monoallelic deletions in *HLA-A*, *B2M,* and *CD58* but not in *HLA-B* or *HLA-C*. Moreover, the absence of PRC2 complex activity and presence of MHC-I and checkpoint molecules expression indicate a potential use of immune checkpoint inhibitors (ICI) for CS4. On the other hand, EZH2 inhibitors (e.g., tazemetostat) in combination with ICI could potentially increase MHC-I expression and immune detection in CS3[26]. Interestingly, the CS3 cluster is enriched with *HIST1H1E/C* mutations which have been recently demonstrated to confer enhanced fitness, and self-renewal properties to B-cells[15].

Additionally, we observed a higher frequency of cases with genetic alterations involved in the cell cycle for CS1 (97%, $p<0.001$, e.g., *CDKN2A/2B* fusions); hence, cyclin D-CDK4 and CDK6 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), and/or inhibition of GAD67[14,20].

Discussion

Identifying groups of patients with shared biologic and prognostic markers is extremely challenging mainly due to high genetic, phenotypic and TME heterogeneity. We identified four PCNSL molecular subtypes with specific oncogenic pathways, gene expression phenotypes, methylation profiles, TME, tumor location, outcome, and potential therapeutic targets (Fig. 5). Moreover, our study gives plausible explanations to the PCNSL response heterogeneity based on finding that many previously PCNSL characteristic features, based on MCD or C5 DLBCL subtypes [1,14], are cluster-specific (C1-C4) 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 tumor location, TME, and B-cell differentiation program differences.

Moreover, 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.

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.

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.

Declarations

Acknowledgments

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 5 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 in part supported by a grant from Investissements d'avenir and by the grant INCa-DGOS-Inserm_12560 of the SiRIC CURAMUS, the program "investissements d'avenir" ANR-10-IAIHU-06, PRT-K/INCa grant LOC-model reference 2017-1-RT-04, DRCI de l'APHP, 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 No. ANR-18-RHUS-0012), Foundation RAM active investments, ARTC foundation, an unrestricted grant from Bristol Myers Squibb (BMS): RDON06618, ICGex project and IDeATIon project with an unrestricted grant from MSD Avenir.

Declaration of interest

G.A reports grants from Biogen, Novartis, Roche, Sanofi, Abbvie, Pfizer and CSL Behring, outside the submitted work. A.A. reports research grant with an unrestricted grant from Bristol Myers Squibb (BMS). The rest of the authors *declare* no potential conflicts of interest to this study.

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. **24**(5):679–90 (2018).
2. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N Engl J Med. **378**(15):1396–407 (2018).
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 **128**(22):4112–4112 (2016).
4. Sehn LH, Salles G. Diffuse Large B-Cell Lymphoma. N Engl J Med. **384**(9):842–58 (2021).
5. Houillier C, Soussain C, Ghesquières H, et al. Management and outcome of primary CNS lymphoma in the modern era: An LOC network study. Neurology **94**(10):e1027–39 (2020).
6. 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 **20**(10):1059–69 (2018).
7. Chapuy B, Roemer MGM, Stewart C, et al. Targetable genetic features of primary testicular and primary central nervous system lymphomas. Blood **127**(7):869–81 (2016).
8. Fukumura K, Kawazu M, Kojima S, et al. Genomic characterization of primary central nervous system lymphoma. Acta Neuropathol (Berl) **131**(6):865–75 (2016).
9. Lu X, Meng J, Zhou Y, Jiang L, Yan F. *MOVICS*: an R package for multi-omics integration and visualization in cancer subtyping. Bioinformatics **36**(22–23):5539–41 (2021).
10. Goeman JJ. L_1 Penalized Estimation in the Cox Proportional Hazards Model. Biom J. **52**(1):70–84 (2010).
11. 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 **24**(36):5711–5 (2006).
12. Martincorena I, Raine KM, Gerstung M, et al. Universal Patterns of Selection in Cancer and Somatic Tissues. Cell **171**(5):1029–1041.e21 (2017).
13. Bruno A, Boisselier B, Labreche K, et al. Mutational analysis of primary central nervous system lymphoma. Oncotarget **5**(13):5065–75 (2014).
14. 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 **37**(4):551–568.e14 (2020).
15. Yusufova N, Kloetgen A, Teater M, et al. Histone H1 loss drives lymphoma by disrupting 3D chromatin architecture. Nature **589**(7841):299–305 (2021).
16. Schubert M, Klinger B, Klünemann M, et al. Perturbation-response genes reveal signaling footprints in cancer gene expression. Nat Commun **9**(1):20 (2018).
17. 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. **118**(22):e2104504118 (2021).

18. Holland CH, Szalai B, Saez-Rodriguez J. Transfer of regulatory knowledge from human to mouse for functional genomics analysis. *Biochim Biophys Acta BBA - Gene Regul Mech.* **1863**(6):194431 (2020).
19. Kotlov N, Bagaev A, Revuelta MV, et al. Clinical and Biological Subtypes of B-cell Lymphoma Revealed by Microenvironmental Signatures. Cancer Discov. **11**(6):1468–89 (2021).
20. Zhang B, Vogelzang A, Miyajima M, et al. B cell-derived GABA elicits IL–10 + macrophages to limit anti-tumour immunity. Nature **599**(7885):471–476 (2021).
21. Rooney MS, Shukla SA, Wu CJ, Getz G, Hacohen N. Molecular and Genetic Properties of Tumors Associated with Local Immune Cytolytic Activity. Cell **160**(1–2):48–61 (2015).
22. 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. **108**(19):7926–31 (2011).
23. Phelan JD, Young RM, Webster DE, et al. A multiprotein supercomplex controlling oncogenic signalling in lymphoma. Nature **560**(7718):387–91 (2018).
24. 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. **266**(1):123–33 (2015).
25. 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 **132**(3):293–306 (2018).
26. 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 **54**(1):116–131.e10 (2021).

Figures

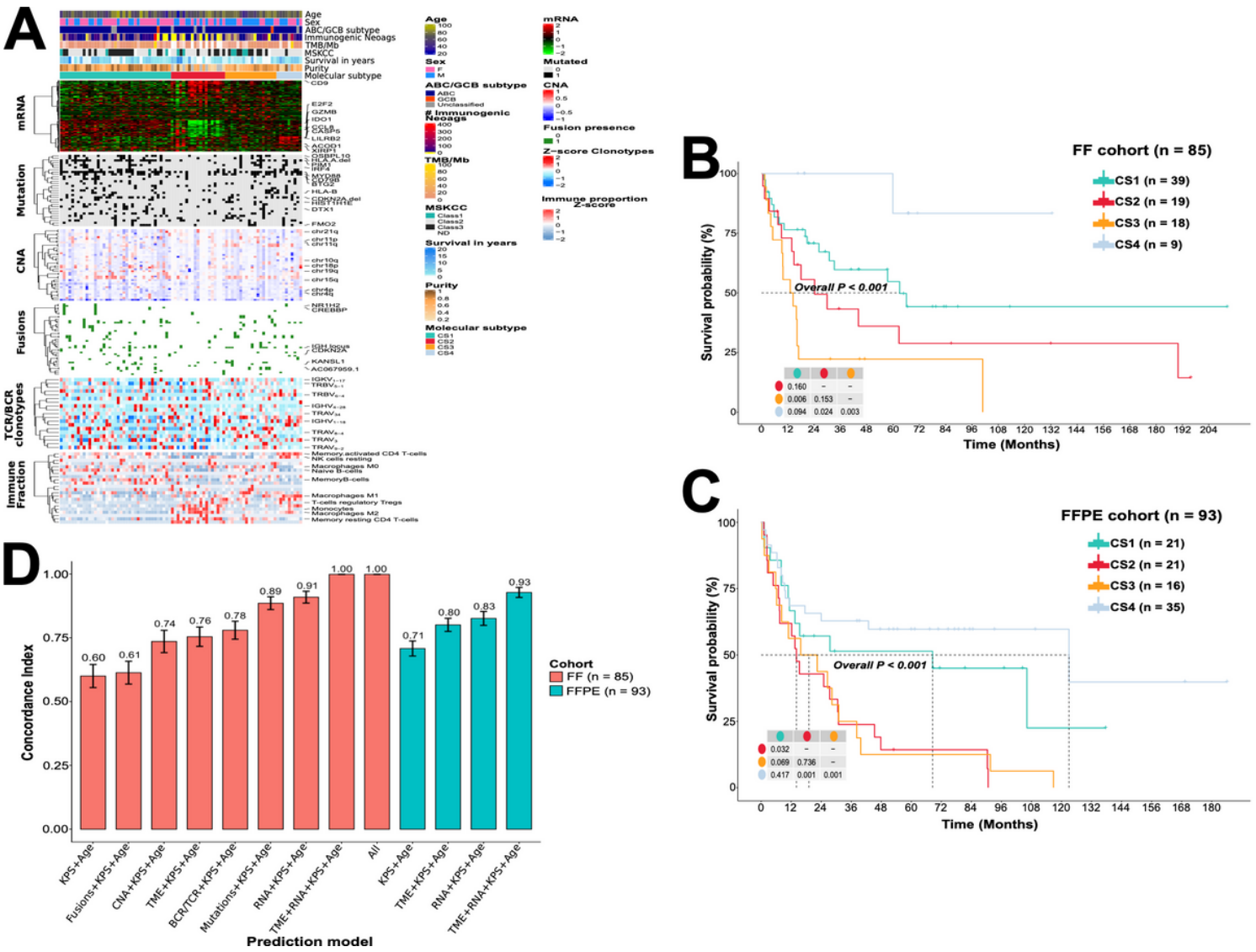

Figure 1

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

Panel A shows the resulting consensus heatmap based on 10 integrative clustering algorithms to define the clusters (CS1 to CS4) 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 (2,087 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. Panel B shows Kaplan-Meier estimates of overall survival among patients belonging to each cluster that resulted from the Consensus cluster of clusters analysis. Panel C shows Kaplan-Meier estimates of overall survival among patients belonging to each cluster using an FFPE validation cohort (n=93). Age and KPS multivariate models for both cohorts are shown in Fig. S4-S13 in Supplementary Appendix 1. Panel D shows the Harrell's concordance index (value annotated at top of each bar) obtained when evaluating each omic-level data to outcome prediction models using Cox proportional hazards regression. 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 the 95% confidence intervals.

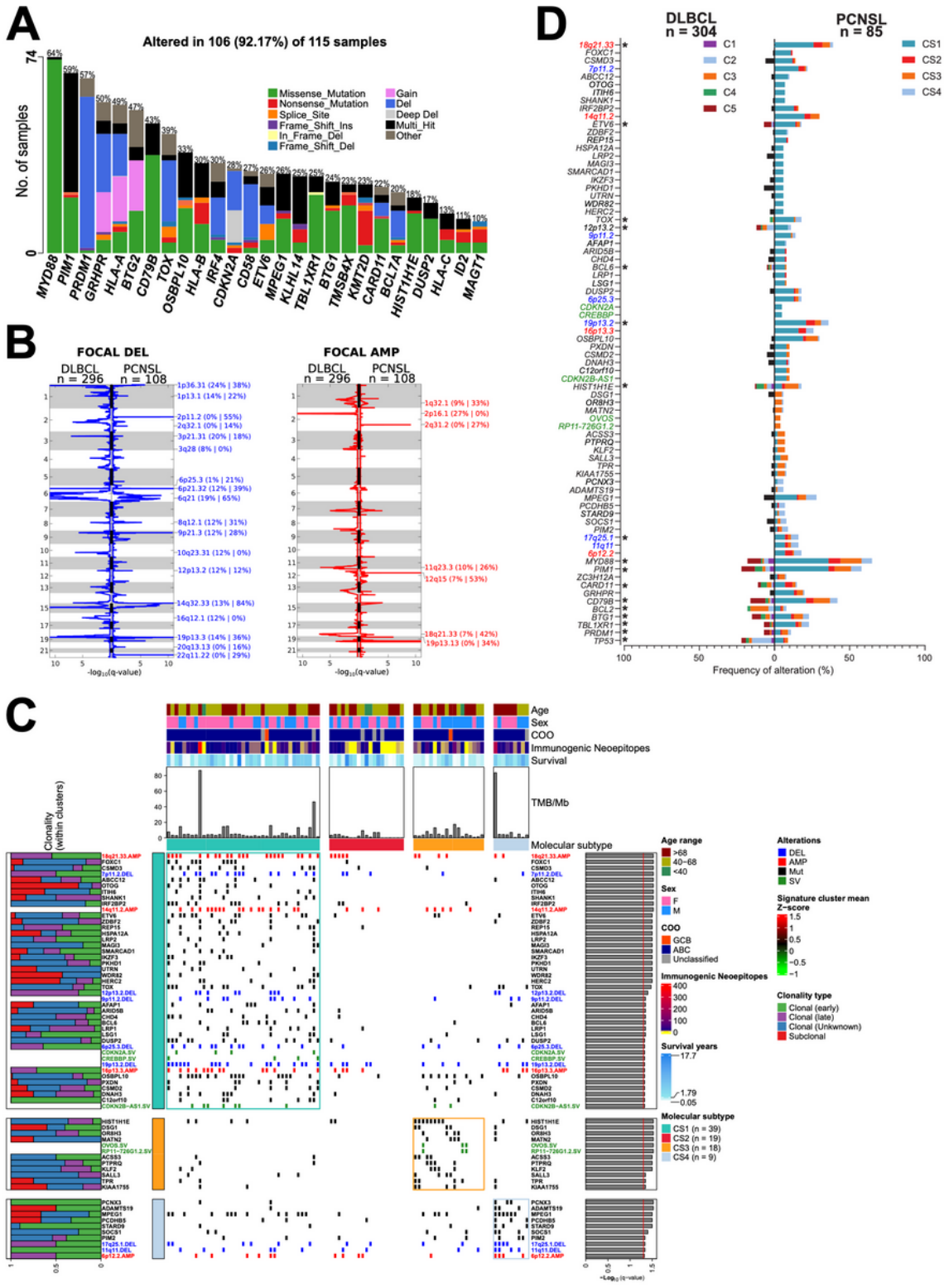

Figure 2

Distinct genetic signatures within PCNSL subtypes and systemic DLBCL

Panel A) shows the number of affected samples within the top driver genes (identified by dNdScv algorithm) in the cohort of 115 PCNSL patients. Barplots 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 barplot for each driver gene. B) shows the GISTIC2.0-defined recurrent copy number focal deletions (blue, left) and gains (red, right) as mirror plots in DLBCL (n=304 from Chapuy et al., 2018) and PCNSL (n=108 from this study). Chromosome position is on the y-axis, and significance is on the x-axis. SCNAs are labeled with their associated cytoband/arm followed in brackets by the frequency of the alteration (DLBCL | PCNSL). C) shows the landmark genetic alterations for each PCNSL subtype (boxed for each cluster) identified by a one-sided Fisher test (event within-cluster vs outside-cluster) and ranked by significance (q-value ≤0.1 selected, red 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. COO, cell of origin; F, female; M, male; DEL, deletion; AMP, amplification; Mut, mutation; Fusion, fusion transcript. Panel D shows 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). Asterisks denote the known driver events in DLBCL and colors the alteration type (mutation = black; gain = red; loss = blue; structural variant = green).

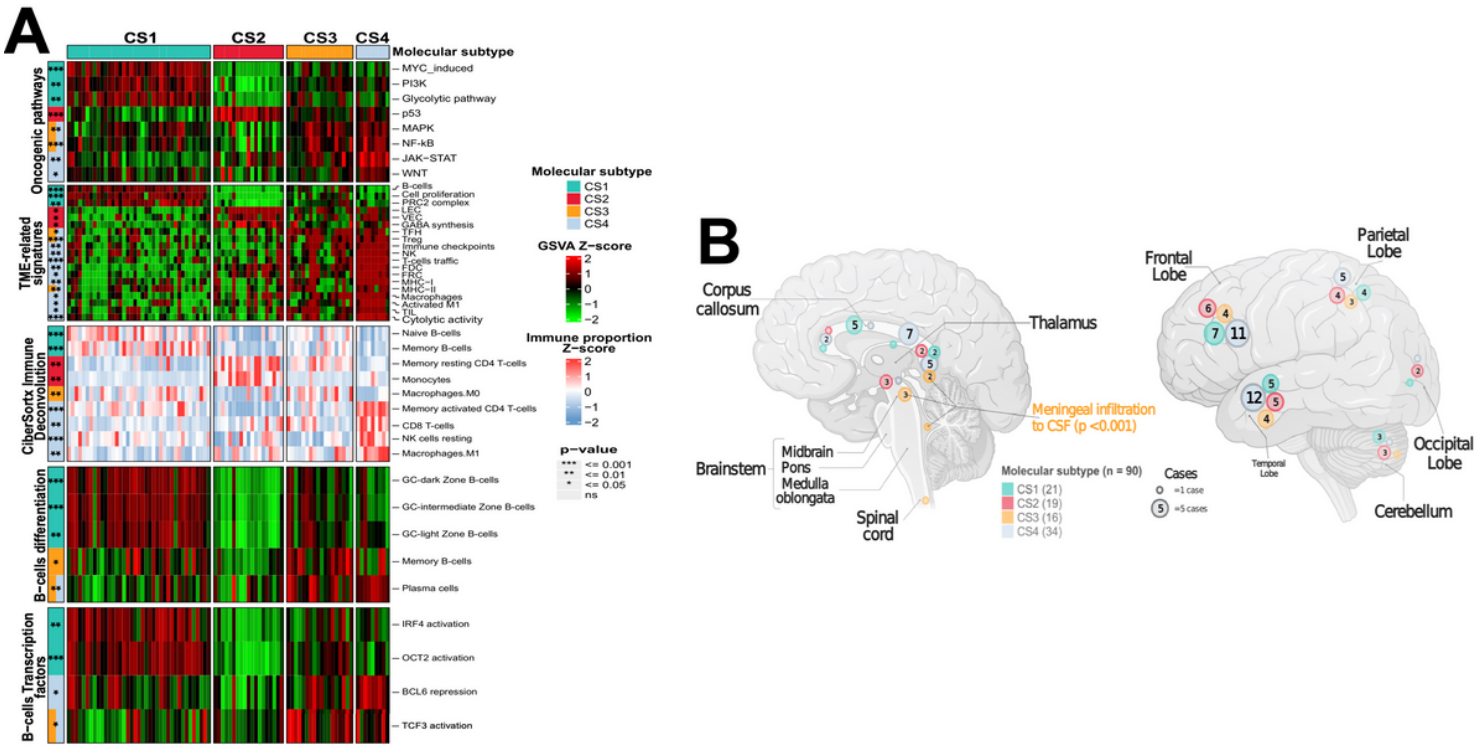

Figure 3

Phenotypic and tumor location distinctions of the multi-omic defined PCNSL subtypes

Panel A shows a heatmap with either gene signature activity (measured by GSVA) or immune cell proportions (CiberSortx deconvoluted) across molecular subtypes. P-values indicate higher expression of the colored group when compared against the others (Wilcoxon-test, left side of plot). FRC, Fibroblastic reticular cells; FDC, follicular dendritic cells. Panel B shows the tumor location of 90 PCNSLs (FFPE cohort) in the human central nervous system grouped by molecular subtype 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 vs outside-cluster).

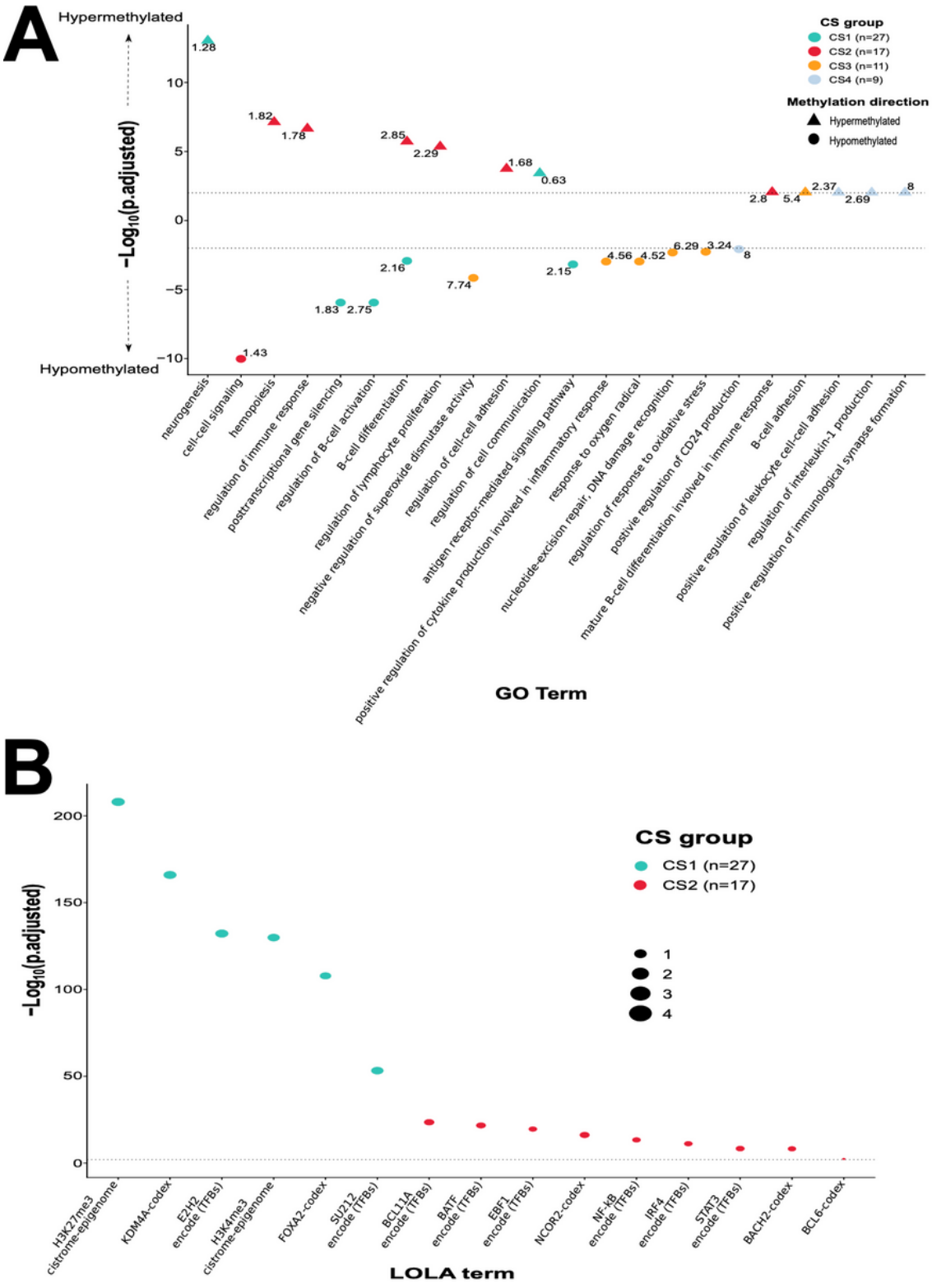

Figure 4

Epigenetic attributes across PCNSL subtypes

Panel A shows GO enrichment analysis on DMP across subtypes where the log(OddsRatio) is annotated next to its associated p-adjusted value. Panel B shows the locus overlap (LOLA) region set enrichment analysis for hypermethylated promoters across the four PCNSL subtypes. The CS3 group is characterized

by enrichment of H3K27me3 probably resulting from the high PRC2 complex transcriptomic activity, whereas the CS2 by enrichment of BCL11A, NF-κB, and IRF4 which is in line with the observed low transcriptomic activity of the related targets (Figure 3A). 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. The complete lists of GO and LOLA enrichments results are provided in Tables S12 and S13.

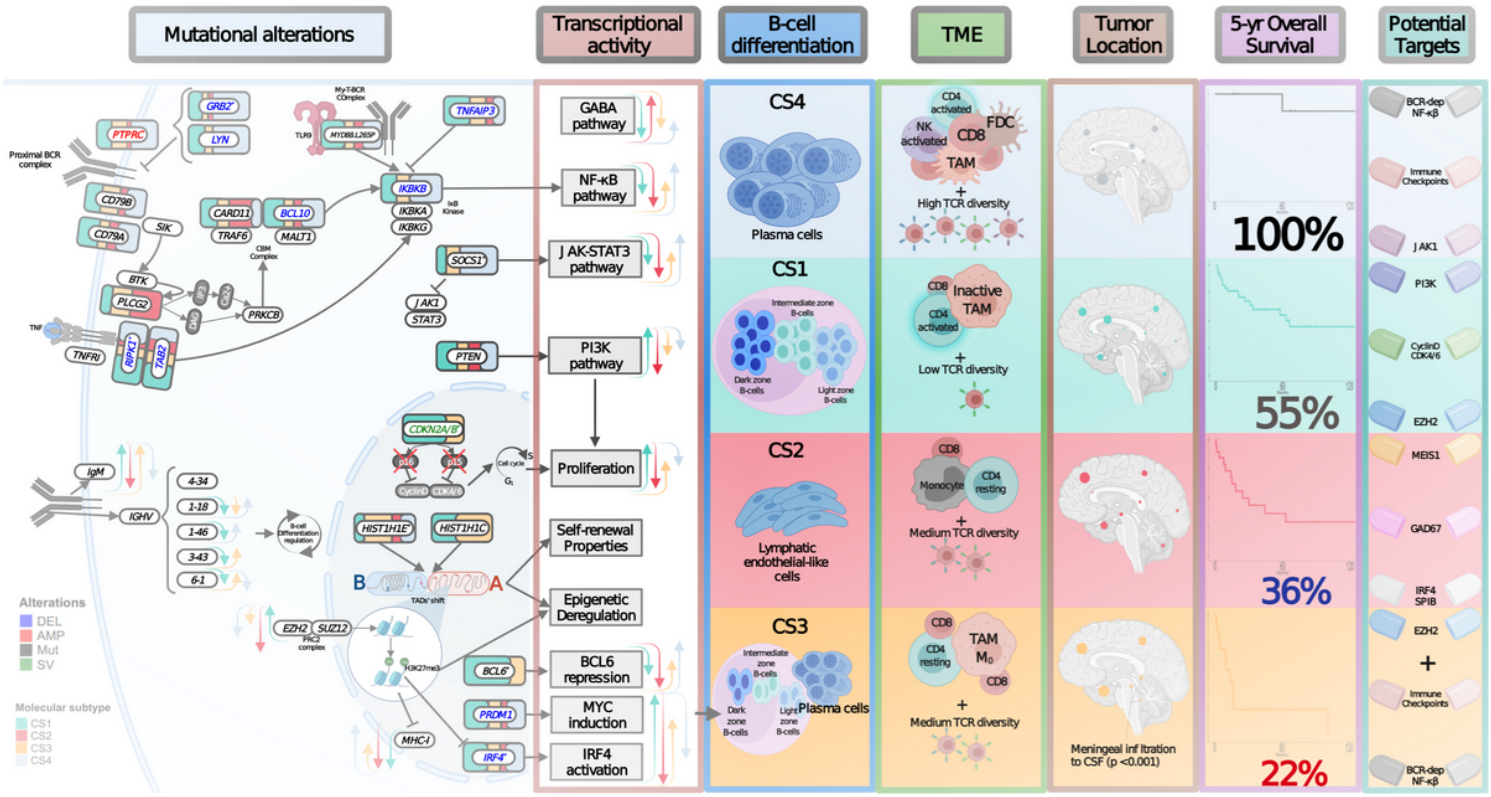

Figure 5

From multi-omics to potential therapeutic targets

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 2C). Arrows indicate transcriptional gene signature activity where the height indicates the relative up or downregulation (according to Figure 3A). DEL, deletion; AMP, amplification; Mut, mutation; SV refers to fusion transcripts. NF-κB activity could not be explained by self-antigen-dependent chronic active BCR signaling upregulation since IgV_{H4-34} expression was not significantly different across groups. The IgV_H regions more expressed in CS4 ($p < 0.05$) were the V_{1-18}, V_{1-46}, and V_{3-9} that have been associated with more differentiated B-cell stages; meanwhile, the naive-transition stage-related, V_{3-43} and V_{4-30-2} regions, were upregulated in CS3.

HIST1H1E/C mutations confer self-renewal properties to B-cells and induce shifts from compartment B to compartment A chromatin. Analyses for determining the TCR/BCR diversity, the immunoglobulin heavy-chain variable (V_H) and constant regions expression, and the master regulators (*MEIS1*, *IRF4*, and *SPIB*) listed as potential CS2 targets are provided in Supplementary Appendix 1.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryAppendixtext.docx
- Allfigures.pdf
- TableS1.xlsx
- TableS3.xlsx
- TableS4.xlsx
- TableS5.xlsx
- TableS6.xlsx
- TableS7.xlsx
- TableS8.xlsx
- TableS9.xlsx
- TableS10.xlsx
- TableS11.xlsx
- TableS12.xlsx
- TableS13.xlsx
- TableS14.xlsx
- flatAlentornepc.pdf
- flatAlentornrs.pdf