Primary Central Nervous System Lymphoma Tumor Biopsies Show Heterogeneity in Gene Expression Profiles, Genetic Subtypes, and in vitro Drug Sensitivity to Kinase Inhibitors

Yuan Xiao Zhu[1, 6], Jasper CH Wong[2], Talal Hilal[3], Alanna Maguire[1, 4], Jon Ocal[1], Katie Zellner[1], Xianfeng Chen[5], Brian K Link[7], Thomas M Habermann[8], Matthew J Maurer[9], James R Cerhan[10], Patrick B Johnston[8], Andrew L Feldman[11], David W. Scott[2], Allison Rosenthal[3] and Lisa Rimsza[1]

[1]Department of Laboratory Medicine and Pathology, Mayo Clinic Arizona, Phoenix, AZ
[2]Centre for Lymphoid Cancer, British Columbia Cancer, Vancouver, BC V5Z 1L3, Canada
[3]Department of Internal Medicine, Hematology-Oncology, Mayo Clinic Arizona, Phoenix, AZ
[4]Department of Cancer Biology, Mayo Clinic Arizona, Phoenix, AZ
[5]Department of Research Biostatistics, Mayo Clinic Arizona, Scottsdale, AZ
[6]Division of Hematology/Oncology, Mayo Clinic Arizona, Scottsdale, AZ
[7]Department of Internal Medicine, University of Iowa, Iowa city, IA
[8]Division of Hematology, Mayo Clinic Rochester, MN
[9]Division of Biostatistics and Medicine, Mayo Clinic Rochester, MN
[10]Division of Epidemiology, Mayo Clinic, Rochester, MN
[11]Department of Laboratory Medicine and Pathology, Mayo Clinic Rochester, MN

Support: National Cancer Institute P01CA229100; MER supported by P50 CA97274 and U01 CA195568

Corresponding Author:

Lisa M. Rimsza, M.D.
Professor and Chair
Department of Pathology and Laboratory Medicine
University of Arizona, College of Medicine Tucson
1501 N. Campbell Ave, #5205
Tucson, AZ 85724-1234
Office: 520-626-6097
Email: lrimsza@arizona.edu

Key Words: PCNSL; Central Nervous System; Lymphoma; DLBCL; Gene Expression Profiling, Activated B cell

ABSTRACT

Primary central nervous system lymphoma (PCNSL) is clinically challenging due to its location and small biopsy size, leading to a lack of comprehensive molecular and biologic description. We previously demonstrated that 91% of PCNSL belong to the activated B-cell-like (ABC) molecular subtype of diffuse large B-cell lymphoma (DLBCL). Here we investigated the expression of 739 cancer related genes in HIV (-) patients using NanoString digital gene expression profiling in 25 ABC-PCNSL and 43 ABC-systemic DLBCL, all tumors were EBV (-). We found that two-thirds of ABC-PCNSL samples had a transcriptional landscape distinct from ABC-systemic DLBCL samples. Of the 739 genes measured, 135 were identified as differentially expressed between these ABC-PCNSL and ABC-systemic DLBCL ($p<0.05$). Compared with ABC-systemic DLBCL, ABC-PCNSL showed higher gene expression in several cancer related gene sets including genes related to Hedgehog, DNA damage repair, Wnt and MAPK signaling. Hierarchical clustering 28 PCNSL samples (25 ABC and 3 GCB subtypes) identified two transcriptional subgroups, P1 (n=9) and P2 (n=19). P2 showed higher activities across most of the cancer related pathways and had a significantly shorter patient survival time ($p<0.01$). Whole exome sequencing showed that some distinct genetic features of PCNSL compared to DLBCL. The genetic subtypes ("LymphGen") of PCNSL consisted mainly of "MCD" and "Other" subtypes, which did not correlate with clinical survival. These data provide more information about unique characters of PCNSL, which may help to identify novel drug targets for developing therapeutic strategies.

Introduction

Primary central nervous system large B-cell lymphoma (PCNSL) is a distinct disease that presents and remains confined to the central nervous system. PCNSL is a mature B-cell lymphoma with large cell cytology, predominantly of the activated B-cell-like (ABC) subtype of diffuse large B-cell lymphomas (DLBCL) (Hilal et al., 2019). PCNSL may occur in immunocompromised patients with human immunodeficiency virus (HIV) and express Epstein-Barr virus associated mRNAs (EBER). However, most cases arise in patients that are HIV (-) and whose tumors lack EBER (Swerdlow et al., 2016). PCNSL is included in the global category of large B-cell lymphomas of immune-privileged sites, alongside cases occurring in the testis or vitreoretina, with the rationale that these locations may share common mechanisms of lymphomagenesis (Alaggio et al., 2022).

Despite a favorable initial response to chemotherapy and radiation, the prognosis of PCNSL is inferior to that of DLBCL presenting outside the CNS. The incidence of PCNSL has increased in older adults (>65-70 years) over the last four decades while improvement in survival has not (Mendez et al., 2018), likely due to the difficulty of safely administering intensive CNS-penetrating chemotherapy combinations in that age group (Hilal, 2020). In a nationwide population-based study, the 2-year and 5-year overall survival among patients >60 years of age was 44% and 28%, respectively (Houillier et al., 2020). Further analyses of potential prognostic groups and therapeutic targets are therefore justified to identify therapeutics that can be administered with acceptable toxicity to older adults.

Herein, we compared PCNSL to systemic DLBCL using gene expression profiling (GEP), mutation and copy number analysis (CNA), patient survival data, and drug sensitivity assays to identify unique biology and altered cellular pathways that may be amenable to therapeutic targeting. We focused on primary disease (cases with secondary involvement were excluded) from patients without HIV infection, EBV (-)

tumors, and ABC subtype determined by gold standard molecular methods, to create a uniform cohort of patients and comparisons.

Materials and Methods

Patient samples

Previously, 37 PCNSL and 43 systemic DLBCL diagnostic samples from untreated patients with annotated clinical outcomes were obtained from the Molecular Epidemiology Resource (MER) (Cerhan et al., 2017). For rigorous comparison, this work focused exclusively on HIV (-) patients and EBV (-) tumors that were molecularly confirmed as ABC subtype from previous work (Hilal et al., 2019). Cases with secondary involvement of the CNS were excluded. The only criteria for PCNSL sample inclusion in this follow-up biological study was adequate remaining material, leaving 25 PCNSL from the original 37. The systemic DLBCL cases were likewise ABC subtyped by Lymph2Cx assay, HIV and EBV (-). Hematoxylin and eosin (H&E) stained slides were reviewed by a hematopathologist to determine tumor content and verify the diagnosis. Cases with less than 60% tumor content were enriched by macro-dissection of the viable tissue. This study was conducted in compliance with IRB approved protocols (PR16-000507 and PR2207-02).

Cell lines

Two human PCNSL cell lines, TK and HKBML, were purchased from the JCRB cell bank and RIKEN BRC cell bank in Japan, respectively. Both cell lines have been reported and used in previous PCNSL research (Fujimoto et al., 2020; Miyoshi et al., 1982). TK cells were cultured in RPMI1640 medium supplemented with 10% fetal calf serum (FCS) and HKBML cells were maintained in HamF12 with 15% FCS. HBL-1 was gifted from Dr. Riccardo Dalla-Favera's lab (Columbia University, New York, NY) and maintained in IMDM with 10% FCS. OCI-Ly3 and TMD8 were gifted from Dr. Louis Staudt's lab at the National Cancer Institute (Bethesda, MD) and maintained in RPMI with 10% FCS.

DNA and RNA extraction

DNA and RNA were extracted with the AllPrep DNA/RNA FFPE kit (Qiagen, USA) from at least 2x 10 µm formalin-fixed paraffin-embedded (FFPE) unstained sections. RNA was quantified using a NanoDrop (Thermo-Scientific, USA). DNA was quantified using the High Sensitivity dsDNA Qubit kit (Invitrogen, USA).

NanoString and gene expression analysis

A minimum of 360ng of total RNA derived from FFPE patient samples, or 100ng of RNA from cell lines, was used to determine the Cell of Origin (COO) using the Lymph2Cx assay (Hilal et al., 2019; Scott et al., 2014). Likewise, an RNA input of 300ng per patient sample or 100ng per cell line was used for a modified PanCancer Pathways (739 genes plus 40 housekeepers) panel from NanoString (Seattle, WA), which included 606 pathway genes for 13 canonical pathways, 124 cancer driver genes, and 9 BCL2 family genes. Both the Lymph2Cx and PanCancer Pathways panels were run on the NanoString nCounter system (Seattle, WA). RNA from an immortalized DLBCL cell line (DB cells, ATCC) was included as a batch control in each run for quality control and normalization (supplemental figure 1). Synthetic RNA oligomers (5 µl of 1.8 µM) were included as positive controls in the Lymph2Cx. The collected RCC data were first reviewed and evaluated using the NanoString nSolver 4.0 Analysis software for quality control, followed by technical normalization using synthetic controls and biological normalization via housekeeping genes. The data collected from the PanCancer Pathways panel were analyzed with an Advanced Analysis software plugin (version 2.0, R-based statistical tool) for nSolver 4.0 to detect and visualize differentially expressed genes, perform unsupervised hierarchical clustering, and assess pathway score changes in different groups of samples.

Whole exome sequencing, copy number, and EBV analysis

DNA from 29 PCNSL and 5 cell lines were sequenced using the Agilent SureSelect Human All Exon v6 + UTR kit at Canada's Michael Smith Genome Sciences Center at BC Cancer, Vancouver, Canada. Libraries were sequenced on an Illumina high-throughput lane (150bp paired-ends). Reads were aligned to GRCh37 reference genome using BWA-MEM (v.0.7.17) (Li & Durbin, 2009). High-confidence simple somatic variants (SSMs) were determined using SLMS-3, a voting approach of four variant callers (Strelka2, Lofreq, Mutect2, and SAGE), wherein consensus variants were selected if three out of four variant callers identified them (Kim et al., 2018; McKenna et al., 2010; Wilm et al., 2012). SAGE can be found at https://github.com/hartwigmedical/hmftools/tree/master/sage. The full pipeline can be found on https://github.com/LCR-BCCRC/lcr-modules (slms3). A curated set of genes implicated in other B-cell lymphomas was used to subset to genes of interest. The genetic subtype of each sample was determined using the LymphGen algorithm, as previously described (Wright et al., 2020).

Potential copy number alterations (CNAs) were identified using PureCN (v.2.0.1) (Riester et al., 2016), which calculated and corrected for tumor ploidy and purity. High confidence CNAs were determined using GISTIC v2.0.23 (Mermel et al., 2011) with a q-value cut-off of 0.1. Samples showing a tumor purity of 0.4 or less were excluded from copy number analyses due to high levels of noise (Supplemental data 1). Manual review of the *CDKN2A* locus was performed using CNVkit (v.0.9.9) (Talevich et al., 2016). Observations were subsequently cross-referenced with the loci of known oncogenes and tumor suppressors implicated in DLBCL. GISTIC plots were generated using Oviz-Bio's visualization suite (Jia et al., 2020).

Mutational and copy number profiles were compared to published systemic ABC-DLBCL whole exome sequencing (WES) libraries generated by Agilent SureSelect v2 (Chapuy et al., 2018) and Agilent SureSelect v5 (Reddy et al., 2017; Schmitz et al., 2018). Mutational calls were restricted down to regions

shared by all capture kits. Publicly available datasets were processed using PureCN (v.2.0.1). Regions that contained high repeats, segmental duplications, and centromeric and telomeric regions were excluded from copy number analyses.

To detect EBV transcripts in the sequencing data and determine EBV positivity, the GATK (v.4.1.8.1) pipeline PathSeq was used to map any unmapped reads in WES libraries to a reference EBV genome with a threshold of 0.0002% reads. We leveraged other WES libraries that used the same capture kit with experimentally confirmed EBV status to threshold our cohort.

Immunoblotting assays

Western blots used 30mg of protein subjected to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) gels followed by transfer to PVDF membranes. Membranes were probed with primary antibodies overnight, washed, and incubated with horseradish peroxidase (HRP)-conjugated-secondary antibodies. Detection was performed by the Enhanced Chemical Luminescence (ECL) method. Anti-AKT, anti-pAKT (S473), anti-PERK (137F5), anti ERK(Thr202/Y204) and anti b-catenin antibodies were from Cell Signaling Technology (Danvers, MA), and mouse anti-b actin antibody was from GenScript (Piscataway, NJ).

Drug treatment and cell viability assay

PCNSL cell lines representing P1, P2, and systemic ABC-DLBCL were treated with methotrexate (MTX) and temozolomide (TMZ), 2 commonly used chemotherapy drugs. We also tested novel agents and drugs inhibiting highly expressed pathways including Bruton's tyrosine kinase (BTK) (ibrutinib), AKT (afuresertib), which is a key player in PI3K pathway, BCL2 (venetoclax), immunomodulatory drugs (pomalidomide), and fibroblastic growth factor receptor (FGFR) (BGJ398). All drugs were purchased from Selleck Chemicals LLC (Houston, TX). Cell viability and cell growth was

measured by MTS or MTT cell proliferation assay kits according to the manufacturer's instructions (Abcam, Waltham, MA). Experiments were performed in triplicate and repeated at least twice.

Correlation of transcriptional subtypes of PCNSL with clinical survival analysis

Kaplan-Meier survival analysis was performed on those two subgroups using R library ggsur rvp lot (https://cloud.r-project.org/package=survminer)**.** Correlations between LymphGen subgroups and MCD genetic subtypes with survival were performed similarly. Overall survival was defined from the time of diagnosis to death from any cause.

RESULTS

ABC-PCNSL shows higher gene expression in Hedgehog, DNA damage repair, Wnt and MAPK associated signaling compared to ABC- systemic DLBCL

In this study, 28 RNA samples from 37 PCNSL patients passed QC and yield enough RNA for expression profile. Of 28 PCNSL, 25 PCNSL were identified as ABC subtypes, and 3 were identified as GCB subtypes and subsequently excluded from this analysis. For rigorous comparison, we first compared the expression of 739 cancer-related genes in 25 cases of HIV (-) ABC-PCNSL with 43 cases of HIV (-) systemic ABC-DLBCL (both nodal and extra-nodal), which had sufficient remaining material from our clinical testing database. An overview of the heatmap of normalized expression data indicated that most of the HIV (-) ABC-PCNSL cohort (P) has a distinct transcriptional landscape from that of the HIV (-) ABC-systemic DLBCL cohort (NS) (Figure 1A). Of the 739 genes, 135 (18.3%) showed significant differential expression between these two groups ($p<0.05$, adj $p<0.05$, supplemental data 2 and Figure 1B). Analysis showed significant upregulation of KMT2D, MCL1, NTRK2 and MAPK10, and downregulation of LAMA5 and FLNA (Figure 1C and Supplemental data 2), in the ABC-PCNSL cohort compared with the ABC-systemic DLBCL cohort of nodal/systemic cases. Using pathway score analysis (provided in the analysis software), which condense each samples' pathway gene expression profiles into small sets of pathway

scores (increasing score corresponds to mostly increasing expression), we identified that ABC-PCNSLhad higher pathway scores in signaling related to hedgehog, DNA damage repair, wnt, and MAPK, and lower scores in the gene sets associated with cell cycle-apoptosis, notch and Jak-stat signaling and driver genes (Figure 1C).

We hypothesized that some of the differentially expressed genes between ABC-systemic DLBCL and ABC-PCNSL could originate from low purity related to background CNS tissues (Supplemental data 1). Therefore, we compared our differentially expressed gene list with the 474 CNS-enriched genes identified from a previous study of gene expression in normal CNS compared to malignant B cells (Montesinos-Rongen et al., 2008). We found that 5 CNS-enriched genes were on our list, including *FGFR2*, *HDAC10, JAG2*, *MAPK8IP1* and *SKP1* (Supplemental data 3). We further compared our list with another 475 highly expressed genes in the human brain from a published dataset (Human Protein Atlas),and identified 3 genes that were on our list (Supplemental data 3), including *FGFR2*, *GNG7*, and *ZIC2*. We performed immunohistochemistry staining of FGFR2 to confirm staining specifity (Supplemental Figure 2). FGFR2 was expressed in the kidney tubules, and little to no protein staining was seen in the normal brain and normal tonsil. In PCNSL, FGFR2 expression was mainly detected in the surrounding brain tissue and stroma rich areas (Supplemental Figure 2).

PCNSL separated into two groups based on gene expression data with differing survival In the GEP analysis of 28 PCNSL samples (including 25 ABC-PCNSL and 3 GCB-PCNSL), using unsupervised hierarchical clustering, we found two clusters of cases that we termed PCNSL group 1 (P1, n=9) and PCNSL group 2 (P2, n= 19) (Figure 2A). The P1 and P2 samples were from different runs and contained both ABC and GCB types (one GCB sample in P2, two in P1) confirming this was not a batch effect (Figure 2B). The P2 group had more samples with high tumor purity (based on copy number data) (Figure 2B, 44%) than P1 group (23.5%), but one subcluster in P1 group showed a similar percentage of high-purity

samples (40%) to P2. Further analysis indicated that P2 had higher pathway scores across most of the cancer related pathways (mainly *PI3K, MAPK* and *RAS*), but DNA damage repair pathways were lower (Figure 2C-D). We next correlated the two transcriptional subtypes with survival data (Supplemental data 4). We used a coxph model (Proportional Hazards Regression Model), including age as a covariate, to rule out the impact of age on survival. Patients with a P2 subtype showed a significantly shorter overall survival, p=0.003 (Figure 3A). Analysis of differentially expressed genes between P2 and P1 identified 171 genes ($p<0.05$, Supplemental data 5), with the top 10 consisting associated with gene sets of chromatin modification, Ras, cell cycle-apoptosis, *PI3K*, TGF-beta, MAPK and JAK-STAT (Figure 3B).

Genomic subtypes of ABC-PCNSL are not associated with transcriptional subtypes or survival

We performed whole exome sequencing to determine the genetic subtype classification as published for DLBCL (Wright et al., 2020). The majority of PCNSL patient samples had mutations in *MYD88*^{L265P} (18/29; 62.1%), *PIM1* (17/29; 58.6%), and *CD79B* (14/29; 48.3%), representative of the "MCD" LymphGen subclass of DLBCL as expected for the ABC subtype of DLBCL and, by inference, ABC-PCNSL (Figure 4A). Indeed, 59% of PCNSL samples were identified as "MCD", 17% as "MCD-composite", and 24% as "Other" using the LymphGen algorithm (Figure 4B). Consistent with previous descriptions of PCNSL, the most frequent gene mutations affected the NF-kB pathway including *MYD88*^{L265P} (62%), *CD79B* (48%), *CARD11* (23%), *TBL1XR1* (27%), and *NFKBIZ* 3'UTR mutations (14%). We further identified other common mutations associated with the MCD subtype, including *KLHL14* (40%), *SETD1B* (33%), *HLA-B* (33%), *BTG1* (33%), *OSBPL10* (33%), *PRDM1* (27%), *ETV6* (20%), *IRF4* (20%), and *MPEG1* (17%) (Wright et al., 2020). We also found *DUSP2* (27%) mutations, which have been associated with higher FOXP3 and PD-L1

positive cell populations in PCNSL (Zhu et al., 2022). No association of LymphGen subtypes with clinical survival was observed in our cohort, likely due to low case numbers (data not shown).

Since the WES libraries covered the non-coding space, we also examined the 3'UTR and 5'UTR areas of regions that have been described as targets of aberrant somatic hypermutation (aSHM) (Arthur et al., 2018; Dreval et al., 2023). The predominant genes that demonstrated frequent mutations are *PIM1* (75%), *BCL6* (69%), *DTX1* (66%), *TMSB4X* (66%), *IGLL5* (62%), *BTG2* (62%), *PAX5* (59%), *BTG1* (55%), and *BCL7A* (52%) (Figure 4C). Notably, mutations were centered around the transcription start site (TSS) of *MYC* (31%) and *BCL2* (24%) (Supplemental Figure 3).

When relating these mutational patterns to the transcriptional subtypes, no genetic lesions were enriched in either the P1 or P2 group (Supplemental Figure 4). Interestingly, within our initial cohort of PCNSL cases, three samples were classified as germinal center B-cell-like (GCB). When comparing the mutational profile of the GCB-PCNSL against ABC-PCNSL, as expected for what is known in DLBCL, we observed significant enrichment ($q < 0.1$, Fisher's exact test) in *GNA13* (3/3, 100%), *CARD11* (3/3, 100%), and *UBE2A* (2/3, 67%) mutations in GCB-PCNSL, while ABC-PCNSL harbored fewer *GNA13* mutations in 2/26 (8%), *CARD11* mutations in 3/26 (12%), and *UBE2A* mutations in 1/26 (4%) cases (Figure 4D). When comparing regions that have been annotated as common aSHM sites in DLBCL, we observed unique mutations in the *CD74* locus around the TSS in the GCB-PCNSL (2/3, 67%). These results need to be verified in a larger cohort of the uncommon GCB-PCNSL cases.

Since EBV (+) PCNSL tumors are biologically distinct from EBV (-) PCNSL, we confirmed the EBV (-) status by querying unmapped reads in the WES libraries that mapped to the EBV genome. All PCNSL samples showed no reads aligning to the EBV genome, except one case, which showed some read support for EBV, though it was just below the threshold for consideration (Supplemental Figure 5).

Using PureCN v2.0.1, we determined the tumor purity and correlated the results with ploidy of the PCNSL tumors from the WES libraries demonstrating 31% (9/29) displayed polyploidy. We next confined the data set to 16 samples that had a tumor purity of at least 40% for further analyses with GISTIC v.2.0.23. Arm-level amplifications affected chromosomes 7p (44%), 12p (31%), 12q (31%), and 18q (44%), while arm-level deletions affected chromosome 6q (56%). Notably, focal amplifications were observed in 31% (5/16) of tumors at 2p13.3, which contains the gene encoding the splicing regulator PCBP-1 (*PCBP1*). Amplifications at 14q24.1 (*ZFP36L1*), 3q26.32 (*TBL1XR1)*, and 1p13.1 (*PDE4B*) were also observed in 43% (7/16), 31% (5/16), and 31% (5/16) of cases, respectively. Recurrent focal deletions at 9p21.3 (*CDKN2A*) and 19p13.11 (*MEF2B*) were observed in 56% (9/16) and 31% (5/16) of the PCNSL tumors, respectively (Supplemental Figure 6).

The three ABC-systemic DLBCL lines (OCI-Ly3, TMD-8, HBL-1) were classified as "MCD" and harbored mutations representative of the "MCD" subclass (*MYD88*^{L265P}, *CD79B*, and *PRDM1)*. In comparison, both ABC-PCNSL cell lines (TK and HKBML) lacked aberrations in the canonical MCD genes and were classified as "Other" by LymphGen (Figure 4A and Supplemental Figure 7). Notably, the PCNSL TK line displayed a mutation in *CARD11,* which has been associated with increased resistance to ibrutinib (Caeser et al., 2021). OCI-Ly3 also harbored *CARD11* mutations.

PCNSL displayed distinct genetic features compared to MCD-classified systemic ABC-DLBCL

To compare the mutational landscape of PCNSL to systemic ABC-DLBCL, we gathered 395 exome data from published sources (Chapuy et al., 2018; Reddy et al., 2017; Schmitz et al., 2018). We first compared PCNSL to all systemic ABC-DLBCL. Overall, PCNSL showed a significantly ($P < 0.01$, t-test) higher tumor mutation burden (TMB) relative to systemic ABC-DLBCL with a median TMB of 2.86 mutations/MB compared to 2.31 mutations/MB in ABC-DLBCL. Since the majority of PCNSL cases were identified as the "MCD" LymphGen subgroup, we next restricted the analysis to the subset of MCD-classified systemic

ABC-DLBCL (MCD-ABC-DLBCL) as the most biologically comparable group. Compared to the 112 MCD-ABC-DLBCL, the TMB of PCNSL was similar to that of the MCD-ABC-DLBCL with a median of 2.62 mutations/MB (Figure 5A).

In comparison to systemic ABC-DLBCL, we observed significant ($q < 0.1$, Fisher's exact test) mutations in genes associated with the LymphGen subgroup "MCD", including *MYD88*^{L265P}, *CD79B*, and *BTG1* in PCNSL (Supplemental Figure 8). In the strictest comparison, restricted to just MCD-ABC-DLBCL, most differences between the groups were no longer observed (Supplemental Figure 9) except for increased mutations ($q < 0.1$, Fisher's exact test) in *SOCS1*, *HNRNPU*, *GNAI2*, *GNA13, CXCR4,* and *CIITA* in PCNSL compared with MCD-ABC-DLBCL (Figure 5B and Supplemental Figure 10). Consistent with other studies that found increased JAK/STAT pathway usage in PCNSL, SOCS1 is a negative regulator of the JAK/STAT pathway (Liau et al., 2018; Mondello et al., 2020). We also observed mutations affecting G-protein signaling (*GNA13*, *GNAI2)* and in *CIITA*, which encodes a regulator of class II major histocompatibility complex expression. Notably, aberrations in *GNA13, GNAI2,* and *CIITA* were associated with a mutational signature found in the mesenchymal-lymphoma microenvironment (Kotlov et al., 2021). Interestingly, we observed *CXCR4* mutations in 14% of the PCNSL tumors versus 1% in the MCD-ABC-DLBCL. *CXCR4* mutations are rare in DLBCL, often associated with lymphoplasmacytic lymphoma, a low-grade B cell lymphoma,, and have been implicated in increased resistance to ibrutinib (Castillo et al., 2019). Finally, we identified unique mutations in the RNA-binding protein HNRNPU in 14% of the PCNSL, similar to the truncating mutations of *HNRNPU* previously identified in Burkitt lymphoma (Thomas et al., 2023).

The copy number profiles of 89 MCD-ABC-DLBCL were compared to 16 PCNSL tumors with higher purity using GISTIC v.2.0.23 to determine recurrent copy number alterations. The overall profile of PCNSL resembled that of MCD-ABC-DLBCL, with some key differences. Notably, we observed unique arm-level

deletions affecting chromosome 10p in PCNSL overall ($q < 0.1$, Fisher's exact test) (Figure 5C). In contrast, we observed unique arm-level amplifications at chromosome 3q in the MCD-ABC-DLBCL. Chromosome 6q deletions frequently occurred in both groups, although 6q21 deletions containing *PRDM1* occurred at a higher frequency in PCNSL (88%) compared to MCD-ABC-DLBCL (47%) (Figure 5D and E). Similarly, we observed a higher incidence of focal deletions at 19p13.2 containing *MEF2B* (38%) and at 3p21.31 containing *RHOA* (38%) in PCNSL compared to MCD-ABC-DLBCL (11% and 6%, respectively). Focal amplifications at 14q24.1 (*ZFP36L1*), 11q13.1 (*MALAT1*), and 1p31.3 (*PDE4B*), were observed at a higher prevalence in PCNSL compared to MCD-ABC-DLBCL. Other copy number alterations commonly observed in MCD-ABC-DLBCL including deletions in 12p13.2 (*ETV6*), 1p13.1 (*CD58*), and 1p36.31 (*NOL9, KLHL21*), were detected at similar levels in PCNSL (Figure 5E). After manual curation of coverage profiles at the *CDKN2A* locus across all samples, we found similar levels of *CDKN2A* deletions in MCD-ABC-DLBCL 43% (38/89) compared to PCNSL 56% (9/16).

Two PCNSL cell lines show variable sensitivity to chemotherapy, BTK, FGFR, and PI3K inhibitors

To better understand the gene expression profiling data from the PCNSL samples, and to evaluate any associate between pathway score and drug response, two PCNSL cell lines (TK and HKBML) and three systemic DLBCL cell lines (TMD8, OCI-Ly3, and HBL-1), all verified as ABC subtypes by Lymph2Cx, were run using the PanCancer Pathways panel (Figure 6 A, duplicates for each cell line). Three lines, including HKBML, showed higher pathway scores across most of the cancer related pathways than the TK and OCI-LY3 lines (Figure 6A).

Next, we tested drug responses of all lines to ibrutinib, which targets BTK and has been used to treat ABC-DLBCL patients (Grommes et al., 2017). As expected, TK and OCI-Ly3both showed resistance to ibrutinib (Figure 6B). Interestingly, both TK and OCI-LY3 showed lower pathway scores than other cell lines. Consistent with the pathway score analysis, western blotting of the two PCNSL cell lines

demonstrated HKBML had more activated PI3K and ERK signaling than TK (Figure 6C). Further testing with additional drugs/compounds on those two PCNSL cell lines showed that, compared to the TK cell line, HKBML (with a higher pathway score) was less responsive to MTX, pomalidomide, venetoclax, and an FGF inhibitor (BGJ 398), and was more sensitive to TMZ and afuresertib, a PI3 kinase pathway inhibitor (Figure 6D). Overall, these findings suggest that the discovered transcriptional heterogeneity of PCNSL may impact responsiveness to therapy.

Discussion

Evaluating gene expression profiles for CNS lymphoma is more challenging than for systemic DLBCL due to small sample sizes, variable cell density, variable infiltration by immune effector cells, and background brain tissue (Rubenstein et al., 2006). We minimized this by macro-dissecting samples to reach a minimum of 60% tumor. An additional strength of this work is the focus on cases that were both EBV (-) and HIV (-) with exclusion of secondary CNS lymphomas in order to create a rigorous and clean cohort. The main limitations include the small sample size (yet relatively large for this rare disease) and thus limited ability to address multiple testing and lack of replication.

Previously, Tun et al used gene expression profiles to compare PCNSL to other subtypes of non-Hodgkin lymphoma and found a unique CNS signature in which there was significant differential expression of multiple extracellular matrix and adhesion-related pathways (Tun et al., 2008). Specifically, the secreted phosphoprotein 1 (*SPP1*) gene was highly upregulated, while the nicotinamide N-methyltransferase (*NNMT*) vascular endothelial growth factor (*VEGF*) genes were downregulated. *SPP1* reportedly has a role in osteoclast attachment to bone matrix, but also as a cytokine that upregulates expression of interferon-gamma and interleukin-12, the latter involved in both T and NK cell activation. Similar findings were also reported by Rubinstein et al in their microarray study (Rubenstein et al., 2006). Another GEP study found that PCNSL did not differ markedly from systemic DLBCL, but did show that

SPP1 is highly expressed in PCNSL B cells (Montesinos-Rongen et al., 2008). Our study showed the same finding and confirmed over-expression of *SPP1* and down-regulation of VEGFC in PCNSL (Figure 1C and Supplemental data 2). The current study additionally identified new genes and pathways that differentially expressed between ABC-PCNSL and systemic ABC-DLBCL samples, which may help in identifying new therapeutic strategies. For example, we found that MCL-1 expression is higher and BCL-2-is lower in ABC-PCNSL samples compared with ABC-systemic DLBCL, suggesting ABC-PCNSL patients may be less responsive to Venetoclax and inhibition of MCL-1 could be effective.

We observed unique mutations in PCNSL including *GNA13, GNAI2,* and *CIITA*, likely representing a lymphoma tumor microenvironment mutational signature (Kotlov et al., 2021). We also observed *CXCR4* mutations in 14% of PCNSL; unlike in lymphoplasmacytic lymphoma in which 43% of *CXCR4* mutations were S338X mutations, the majority of mutations in PCNSL occurred in the transmembrane domain of CXCR4 with consequent implications for ibrutinib resistance (Gustine et al., 2019). In MCD-ABC-PCNSL mutations in *SOCS1*, a negative regulator of the JAK-STAT pathway, were also enriched.

Chapuy et al aimed to identify genetic features of PCNSL and primary testicular lymphoma using WES. They identified several recurrent mutations involved in NF-kB signaling pathway, including *MYD88, CD79B*, and *PIM1*, identical to our own findings. Functional studies confirmed the role of *MYD88* mutations in activating NF-kB pathway and promoting survival of PCNSL cells (Chapuy et al., 2016). Another study using whole genome and RNA sequencing to characterize the genetic landscape of both primary and secondary large B-cell lymphomas, EBV (+) and EBV (-), in the CNS found significantly more single nucleotide variants (SNVs) and indels compared to systemic DLBCL, even in the intronic and intergenic regions (Radke et al., 2022). In addition, that group of investigators found recurrent mutations in JAK-STAT, NF-kB, and B-cell receptor signaling pathways, alongside other mutations also found in other ABC-DLBCL such as *MYD88*, *CD79B*, as well as deletions of *MTAP* and C*DKN2A* at 9p21

and *CDKN2D* at 19p13 in the PCNSL cases. The PCNSL cases also had more frequent deletion of the human leukocyte antigen (*HLA*) type I D locus (*HLA-D*) at 6p21, implicating immune evasion in lymphomagenesis. In fact, homozygous deletions of the *HLA* type I and II loci have long been known to be more common in PCNSL compared to systemic DLBCL (Riemersma et al., 2000). Other features of PCNSL included mutational signatures related to DNA replication and mitosis. In addition, *TERT* gene expression (but not telomere content) and IgM expression (as compared to IgG expression) were both significantly higher in PCNSL compared to the secondary cases. Finally, the EBV (+) cases (primary and secondary) lacked recurrent mutational hotspots apart from immunoglobulin and HLA-DRB loci (Radke et al., 2022).

Due to small sample sizes and limited amounts of RNA extracted from PCNSL samples, we could not perform duplicate runs of patient samples or perform further validation using RT-PCR or immunostaining. However, previous studies from our group and others have repeatedly demonstrated that the data generated from FFPE samples using nCounter technology were correlated with that generated from matched fresh frozen material or QPCR data {Scott, 2014 #12}{Chen, 2016 #205}{Ryan, 2023 #206}. After analyzing two PCNSL cell lines with different pathway scores and their responses to different drugs and inhibitors, we found consistency between pathway scores with activation of certain pathways, such as PI3 signaling. PI3K inhibitors have been evaluated in PCNSL. The first trial was of buparlisib, which closed prematurely due to limited response (Grommes et al., 2016). Another study evaluated PQR309, a novel dual PI3K/mTOR inhibitor, but the results did not seem promising, and data are yet published (NCT02669511, NCT03120000). An attractive approach is combining therapies targeting different pathways. Copanlisib, a pan-PI3K inhibitor is currently being evaluated in combination with ibrutinib in relapsed PCNSL in a phase Ib/II trial (NCT03581942). Results are expected in 2024.

Altogether, these findings add to the existing genomic and transcriptional landscape of PCNSL and suggest heterogeneity within the disease with possible targetable biology.

References:

Alaggio, R., Amador, C., Anagnostopoulos, I., Attygalle, A. D., Araujo, I. B. de O., Berti, E., Bhagat, G., Borges, A. M., Boyer, D., Calaminici, M., Chadburn, A., Chan, J. K. C., Cheuk, W., Chng, W.-J., Choi, J. K., Chuang, S.-S., Coupland, S. E., Czader, M., Dave, S. S., … Xiao, W. (2022). The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia*, *36*(7), Article 7. https://doi.org/10.1038/s41375-022-01620-2

Arthur, S. E., Jiang, A., Grande, B. M., Alcaide, M., Cojocaru, R., Rushton, C. K., Mottok, A., Hilton, L. K., Lat, P. K., Zhao, E. Y., Culibrk, L., Ennishi, D., Jessa, S., Chong, L., Thomas, N., Pararajalingam, P., Meissner, B., Boyle, M., Davidson, J., … Morin, R. D. (2018). Genome-wide discovery of somatic regulatory variants in diffuse large B-cell lymphoma. *Nat Commun*, *9*(1), 4001. https://doi.org/10.1038/s41467-018-06354-3

Caeser, R., Walker, I., Gao, J., Shah, N., Rasso-Barnett, L., Anand, S., Martin, J. E., & Hodson, D. J. (2021). Acquired CARD11 mutation promotes BCR independence in Diffuse Large B Cell Lymphoma. *JCO Precis Oncol*, *5*, 145–152. https://doi.org/10.1200/PO.20.00360

Castillo, J. J., Xu, L., Gustine, J. N., Keezer, A., Meid, K., Dubeau, T. E., Liu, X., Demos, M. G., Kofides, A., Tsakmaklis, N., Chen, J. G., Munshi, M., Guerrera, M. L., Chan, G. G., Patterson, C. J., Yang, G., Hunter, Z. R., & Treon, S. P. (2019). CXCR4 mutation subtypes impact response and survival outcomes in patients with Waldenström macroglobulinaemia treated with ibrutinib. *British Journal of Haematology*, *187*(3), 356–363. https://doi.org/10.1111/bjh.16088

Cerhan, J. R., Link, B. K., Habermann, T. M., Maurer, M. J., Feldman, A. L., Syrbu, S. I., Thompson, C. A., Farooq, U., Novak, A. J., Slager, S. L., Allmer, C., Lunde, J. J., Macon, W. R., Inwards, D. J., Johnston, P. B., Micallef, I. N. M., Nowakowski, G. S., Ansell, S. M., Kay, N. E., … Witzig, T. E. (2017). Cohort Profile: The Lymphoma Specialized Program of Research Excellence (SPORE) Molecular Epidemiology Resource (MER) Cohort Study. *Int J Epidemiol*, *46*(6), 1753–1754i. https://doi.org/10.1093/ije/dyx119

Chapuy, B., Roemer, M. G., Stewart, C., Tan, Y., Abo, R. P., Zhang, L., Dunford, A. J., Meredith, D. M., Thorner, A. R., Jordanova, E. S., Liu, G., Feuerhake, F., Ducar, M. D., Illerhaus, G., Gusenleitner, D., Linden, E. A., Sun, H. H., Homer, H., Aono, M., … Shipp, M. A. (2016). Targetable genetic features of primary testicular and primary central nervous system lymphomas. *Blood*, *127*(7), 869–881. https://doi.org/10.1182/blood-2015-10-673236

Chapuy, B., Stewart, C., Dunford, A. J., Kim, J., Kamburov, A., Redd, R. A., Lawrence, M. S., Roemer, M. G. M., Li, A. J., Ziepert, M., Staiger, A. M., Wala, J. A., Ducar, M. D., Leshchiner, I., Rheinbay, E., Taylor-Weiner, A., Coughlin, C. A., Hess, J. M., Pedamallu, C. S., … Shipp, M. A. (2018). Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. *Nat Med*, *24*(5), 679–690. https://doi.org/10.1038/s41591-018-0016-8

Dreval, K., Hilton, L. K., Cruz, M., Shaalan, H., Ben-Neriah, S., Boyle, M., Collinge, B., Coyle, K. M., Duns, G., Farinha, P., Grande, B. M., Meissner, B., Pararajalingam, P., Rushton, C. K., Slack, G. W., Wong, J., Mungall, A. J., Marra, M. A., Connors, J. M., … Morin, R. D. (2023). Genetic subdivisions

of follicular lymphoma defined by distinct coding and noncoding mutation patterns. *Blood*, *142*(6), 561–573. https://doi.org/10.1182/blood.2022018719

Ennishi, D., Jiang, A., Boyle, M., Collinge, B., Grande, B. M., Ben-Neriah, S., Rushton, C., Tang, J., Thomas, N., Slack, G. W., Farinha, P., Takata, K., Miyata-Takata, T., Craig, J., Mottok, A., Meissner, B., Saberi, S., Bashashati, A., Villa, D., … Scott, D. W. (2019). Double-Hit Gene Expression Signature Defines a Distinct Subgroup of Germinal Center B-Cell-Like Diffuse Large B-Cell Lymphoma. *J Clin Oncol*, *37*(3), 190–201. https://doi.org/10.1200/JCO.18.01583

Fujimoto, K., Shinojima, N., Hayashi, M., Nakano, T., Ichimura, K., & Mukasa, A. (2020). Histone deacetylase inhibition enhances the therapeutic effects of methotrexate on primary central nervous system lymphoma. *Neurooncol Adv*, *2*(1), vdaa084. https://doi.org/10.1093/noajnl/vdaa084

Grommes, C., Pastore, A., Palaskas, N., Tang, S. S., Campos, C., Schartz, D., Codega, P., Nichol, D., Clark, O., Hsieh, W. Y., Rohle, D., Rosenblum, M., Viale, A., Tabar, V. S., Brennan, C. W., Gavrilovic, I. T., Kaley, T. J., Nolan, C. P., Omuro, A., … Mellinghoff, I. K. (2017). Ibrutinib Unmasks Critical Role of Bruton Tyrosine Kinase in Primary CNS Lymphoma. *Cancer Discov*, *7*(9), 1018–1029. https://doi.org/10.1158/2159-8290.CD-17-0613

Grommes, C., Pentsova, E., Nolan, C., Wolfe, J., Mellinghoff, I. K., & Deangelis, L. (2016). 335P - Phase II study of single agent buparlisib in recurrent/refractory primary (PCNSL) and secondary CNS lymphoma (SCNSL). *Annals of Oncology*, *27*, vi106. https://doi.org/10.1093/annonc/mdw367.13

Gustine, J. N., Xu, L., Tsakmaklis, N., Demos, M. G., Kofides, A., Chen, J. G., Liu, X., Munshi, M., Guerrera, M. L., Chan, G. G., Patterson, C. J., Keezer, A., Meid, K., Dubeau, T., Yang, G., Hunter, Z. R., Treon, S. P., & Castillo, J. J. (2019). CXCR4(S338X) clonality is an important determinant of ibrutinib outcomes in patients with Waldenstrom macroglobulinemia. *Blood Adv*, *3*(19), 2800–2803. https://doi.org/10.1182/bloodadvances.2019000635

Hilal, T. (2020). Primary central nervous system lymphoma: Consensus, controversies, and future directions. *Adv Cell Gene Ther*, *3*(3, e82).

Hilal, T., Maguire, A., Kosiorek, H. E., Rimsza, L. M., & Rosenthal, A. C. (2019). Clinical features and cell of origin subtyping using gene expression profiling in HIV-negative patients with primary central nervous system lymphoma. *Leuk Lymphoma*, *60*(14), 3581–3583. https://doi.org/10.1080/10428194.2019.1633637

Houillier, C., Soussain, C., Ghesquieres, H., Soubeyran, P., Chinot, O., Taillandier, L., Lamy, T., Choquet, S., Ahle, G., Damaj, G., Agape, P., Molucon-Chabrot, C., Amiel, A., Delwail, V., Fabbro, M., Jardin, F., Chauchet, A., Moles-Moreau, M. P., Morschhauser, F., … Gyan, E. (2020). Management and outcome of primary CNS lymphoma in the modern era: An LOC network study. *Neurology*, *94*(10), e1027–e1039. https://doi.org/10.1212/WNL.0000000000008900

Jia, W., Li, H., Li, S., Chen, L., & Li, S. C. (2020). Oviz-Bio: A web-based platform for interactive cancer genomics data visualization. *Nucleic Acids Res*, *48*(W1), W415–W426. https://doi.org/10.1093/nar/gkaa371

Kim, S., Scheffler, K., Halpern, A. L., Bekritsky, M. A., Noh, E., Kallberg, M., Chen, X., Kim, Y., Beyter, D., Krusche, P., & Saunders, C. T. (2018). Strelka2: Fast and accurate calling of germline and somatic variants. *Nat Methods*, *15*(8), 591–594. https://doi.org/10.1038/s41592-018-0051-x

Kotlov, N., Bagaev, A., Revuelta, M. V., Phillip, J. M., Cacciapuoti, M. T., Antysheva, Z., Svekolkin, V., Tikhonova, E., Miheecheva, N., Kuzkina, N., Nos, G., Tabbo, F., Frenkel, F., Ghione, P., Tsiper, M., Almog, N., Fowler, N., Melnick, A. M., Leonard, J. P., … Cerchietti, L. (2021). Clinical and Biological Subtypes of B-cell Lymphoma Revealed by Microenvironmental Signatures. *Cancer Discov*, *11*(6), 1468–1489. https://doi.org/10.1158/2159-8290.CD-20-0839

Lassman, A. B., Sepulveda-Sanchez, J. M., Cloughesy, T. F., Gil-Gil, M. J., Puduvalli, V. K., Raizer, J. J., De Vos, F. Y. F., Wen, P. Y., Butowski, N. A., Clement, P. M. J., Groves, M. D., Belda-Iniesta, C., Giglio, P., Soifer, H. S., Rowsey, S., Xu, C., Avogadri, F., Wei, G., Moran, S., & Roth, P. (2022). Infigratinib in Patients with Recurrent Gliomas and FGFR Alterations: A Multicenter Phase II Study. *Clin Cancer Res*, *28*(11), 2270–2277. https://doi.org/10.1158/1078-0432.CCR-21-2664

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liau, N. P. D., Laktyushin, A., Lucet, I. S., Murphy, J. M., Yao, S., Whitlock, E., Callaghan, K., Nicola, N. A., Kershaw, N. J., & Babon, J. J. (2018). The molecular basis of JAK/STAT inhibition by SOCS1. *Nat Commun*, *9*(1), 1558. https://doi.org/10.1038/s41467-018-04013-1

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Mendez, J. S., Ostrom, Q. T., Gittleman, H., Kruchko, C., DeAngelis, L. M., Barnholtz-Sloan, J. S., & Grommes, C. (2018). The elderly left behind-changes in survival trends of primary central nervous system lymphoma over the past 4 decades. *Neuro Oncol*, *20*(5), 687–694. https://doi.org/10.1093/neuonc/nox187

Mermel, C. H., Schumacher, S. E., Hill, B., Meyerson, M. L., Beroukhim, R., & Getz, G. (2011). GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*, *12*(4), R41. https://doi.org/10.1186/gb-2011-12-4-r41

Miyoshi, I., Kubonishi, I., Yoshimoto, S., Hikita, T., Dabasaki, H., Tanaka, T., Kimura, I., Tabuchi, K., & Nishimoto, A. (1982). Characteristics of a brain lymphoma cell line derived from primary intracranial lymphoma. *Cancer*, *49*(3), 456–459. https://doi.org/10.1002/1097-0142(19820201)49:3<456::aid-cncr2820490311>3.0.co;2-k

Mondello, P., Cuzzocrea, S., Arrigo, C., Pitini, V., Mian, M., & Bertoni, F. (2020). STAT6 activation correlates with cerebrospinal fluid IL-4 and IL-10 and poor prognosis in primary central nervous system lymphoma. *Hematological Oncology*, *38*(1), 106–110. https://doi.org/10.1002/hon.2679

Montesinos-Rongen, M., Brunn, A., Bentink, S., Basso, K., Lim, W. K., Klapper, W., Schaller, C., Reifenberger, G., Rubenstein, J., Wiestler, O. D., Spang, R., Dalla-Favera, R., Siebert, R., & Deckert, M. (2008). Gene expression profiling suggests primary central nervous system lymphomas to be derived from a late germinal center B cell. *Leukemia*, *22*(2), 400–405. https://doi.org/10.1038/sj.leu.2405019

Radke, J., Ishaque, N., Koll, R., Gu, Z., Schumann, E., Sieverling, L., Uhrig, S., Hubschmann, D., Toprak, U. H., Lopez, C., Hostench, X. P., Borgoni, S., Juraeva, D., Pritsch, F., Paramasivam, N., Balasubramanian, G. P., Schlesner, M., Sahay, S., Weniger, M., … Heppner, F. L. (2022). The genomic and transcriptional landscape of primary central nervous system lymphoma. *Nat Commun*, *13*(1), 2558. https://doi.org/10.1038/s41467-022-30050-y

Reddy, A., Zhang, J., Davis, N. S., Moffitt, A. B., Love, C. L., Waldrop, A., Leppa, S., Pasanen, A., Meriranta, L., Karjalainen-Lindsberg, M. L., Norgaard, P., Pedersen, M., Gang, A. O., Hogdall, E., Heavican, T. B., Lone, W., Iqbal, J., Qin, Q., Li, G., … Dave, S. S. (2017). Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell*, *171*(2), 481-494 e15. https://doi.org/10.1016/j.cell.2017.09.027

Riemersma, S. A., Jordanova, E. S., Schop, R. F., Philippo, K., Looijenga, L. H., Schuuring, E., & Kluin, P. M. (2000). Extensive genetic alterations of the HLA region, including homozygous deletions of HLA class II genes in B-cell lymphomas arising in immune-privileged sites. *Blood*, *96*(10), 3569–3577.

Riester, M., Singh, A. P., Brannon, A. R., Yu, K., Campbell, C. D., Chiang, D. Y., & Morrissey, M. P. (2016). PureCN: copy number calling and SNV classification using targeted short read sequencing. *Source Code Biol Med*, *11*, 13. https://doi.org/10.1186/s13029-016-0060-z

Rubenstein, J. L., Fridlyand, J., Shen, A., Aldape, K., Ginzinger, D., Batchelor, T., Treseler, P., Berger, M., McDermott, M., Prados, M., Karch, J., Okada, C., Hyun, W., Parikh, S., Haqq, C., & Shuman, M. (2006). Gene expression and angiotropism in primary CNS lymphoma. *Blood*, *107*(9), 3716–3723. https://doi.org/10.1182/blood-2005-03-0897

Schmitz, R., Wright, G. W., Huang, D. W., Johnson, C. A., Phelan, J. D., Wang, J. Q., Roulland, S., Kasbekar, M., Young, R. M., Shaffer, A. L., Hodson, D. J., Xiao, W., Yu, X., Yang, Y., Zhao, H., Xu, W., Liu, X., Zhou, B., Du, W., … Staudt, L. M. (2018). Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. *N Engl J Med*, *378*(15), 1396–1407. https://doi.org/10.1056/NEJMoa1801445

Scott, D. W., Wright, G. W., Williams, P. M., Lih, C. J., Walsh, W., Jaffe, E. S., Rosenwald, A., Campo, E., Chan, W. C., Connors, J. M., Smeland, E. B., Mottok, A., Braziel, R. M., Ott, G., Delabie, J., Tubbs, R. R., Cook, J. R., Weisenburger, D. D., Greiner, T. C., … Rimsza, L. M. (2014). Determining cell-of-origin subtypes of diffuse large B-cell lymphoma using gene expression in formalin-fixed paraffin-embedded tissue. *Blood*, *123*(8), 1214–1217. https://doi.org/10.1182/blood-2013-11-536433

Swerdlow, S. H., Campo, E., Pileri, S. A., Harris, N. L., Stein, H., Siebert, R., Advani, R., Ghielmini, M., Salles, G. A., Zelenetz, A. D., & Jaffe, E. S. (2016). The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*, *127*(20), 2375–2390. https://doi.org/10.1182/blood-2016-01-643569

Talevich, E., Shain, A. H., Botton, T., & Bastian, B. C. (2016). CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput Biol*, *12*(4), e1004873. https://doi.org/10.1371/journal.pcbi.1004873

Thomas, N., Dreval, K., Gerhard, D. S., Hilton, L. K., Abramson, J. S., Ambinder, R. F., Barta, S., Bartlett, N. L., Bethony, J., Bhatia, K., Bowen, J., Bryan, A. C., Cesarman, E., Casper, C., Chadburn, A., Cruz, M., Dittmer, D. P., Dyer, M. A., Farinha, P., … Morin, R. D. (2023). Genetic subgroups inform on pathobiology in adult and pediatric Burkitt lymphoma. *Blood*, *141*(8), 904–916. https://doi.org/10.1182/blood.2022016534

Tun, H. W., Personett, D., Baskerville, K. A., Menke, D. M., Jaeckle, K. A., Kreinest, P., Edenfield, B., Zubair, A. C., O'Neill, B. P., Lai, W. R., Park, P. J., & McKinney, M. (2008). Pathway analysis of primary central nervous system lymphoma. *Blood*, *111*(6), 3200–3210. https://doi.org/10.1182/blood-2007-10-119099

Wilm, A., Aw, P. P., Bertrand, D., Yeo, G. H., Ong, S. H., Wong, C. H., Khor, C. C., Petric, R., Hibberd, M. L., & Nagarajan, N. (2012). LoFreq: A sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets. *Nucleic Acids Res*, *40*(22), 11189–11201. https://doi.org/10.1093/nar/gks918

Wright, G. W., Huang, D. W., Phelan, J. D., Coulibaly, Z. A., Roulland, S., Young, R. M., Wang, J. Q., Schmitz, R., Morin, R. D., Tang, J., Jiang, A., Bagaev, A., Plotnikova, O., Kotlov, N., Johnson, C. A., Wilson, W. H., Scott, D. W., & Staudt, L. M. (2020). A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. *Cancer Cell*, *37*(4), 551-568 e14. https://doi.org/10.1016/j.ccell.2020.03.015

Zhu, Q., Wang, J., Zhang, W., Zhu, W., Wu, Z., Chen, Y., Chen, M., Zheng, L., Tang, J., Zhang, S., Wang, D., Wang, X., & Chen, G. (2022). Whole-Genome/Exome Sequencing Uncovers Mutations and Copy Number Variations in Primary Diffuse Large B-Cell Lymphoma of the Central Nervous System. *Frontiers in Genetics*, *13*. https://www.frontiersin.org/articles/10.3389/fgene.2022.878618

Acknowledgement

The authors wish to acknowledge the gene expression data shared by Dr. Manuel Montesinos-Rongen as well as the Lymphoma & Leukemia Molecular Profiling Project for their critical review.

Figure legends

Figure 1**. Differentially expressed genes between ABC-PCNSL (P) and ABC-systemic DLBCL (NS) were identified**. **(A)** Heatmap overview of normalized expression data of P and NS on the PanCancer Pathways panel. The heatmap was scaled to give all genes equal variance and was generated via unsupervised clustering. Orange indicates high expression; blue indicates low expression. This plot is meant to provide a high-level exploratory view of the data. **(B)** A volcano plot displaying each gene's log10(p-value) and log2 fold change with the selected covariate. Genes with high statistical significance lie at the top of the plot above the horizontal lines, and genes with high differential expression fall to either side. Horizontal lines indicate the various False Discovery Rates (FDR) or appropriate p-value thresholds if there is no adjustment to the p-values. Genes are colored if the resulting p-value is below the given FDR or p-value threshold. The 40 most statistically significant genes are labeled in the plot. **(C)** Pathway score analysis of P and NS indicated a higher gene expression in Hedgehog, DNA damage repair, Wnt and MAPK associated signaling compared to ABC- systemic DLBCL. Covariate plots comparing the pathway scores between NS and P.

Figure 2. **Two transcriptional subgroups were identified in PCNSL samples**. **(A-B)** Heatmap of the normalized data from 28 cases of HIV negative PCNSL revealed two transcriptional clusters (P1 and P2) using unsupervised hierarchical clustering. The samples from different runs (A) and with different tumor purities (B) were marked. H: tumor purity $\geq$ 0.6; M: tumor purity between 0.3 and 0.6; L: tumor purity $\leq$ 0.3; N: no information about purity. **(C)** Heatmap of pathway scores across all PCNSL samples indicated that P2 has higher pathway scores across most of the cancer related pathways. **(D)** Covariate plots

comparing the pathway scores between P2 and P1 identified top pathway score changes, including high scores in PI3K, MAPK, RAS signaling and low scores in DNA damage signaling.

Figure 3. **The patients in P2 have a short clinical overall survival**. **(A)** Kaplan-Meier survival analysis of patient data from two transcriptional subgroups of PCNSL indicates a shorter survival time for patients from the P2 subgroup compared with the P1 subgroup (p=0.003)**. (B)** The top 25 genes with the greatest differential expression between P1 and P2 are listed.

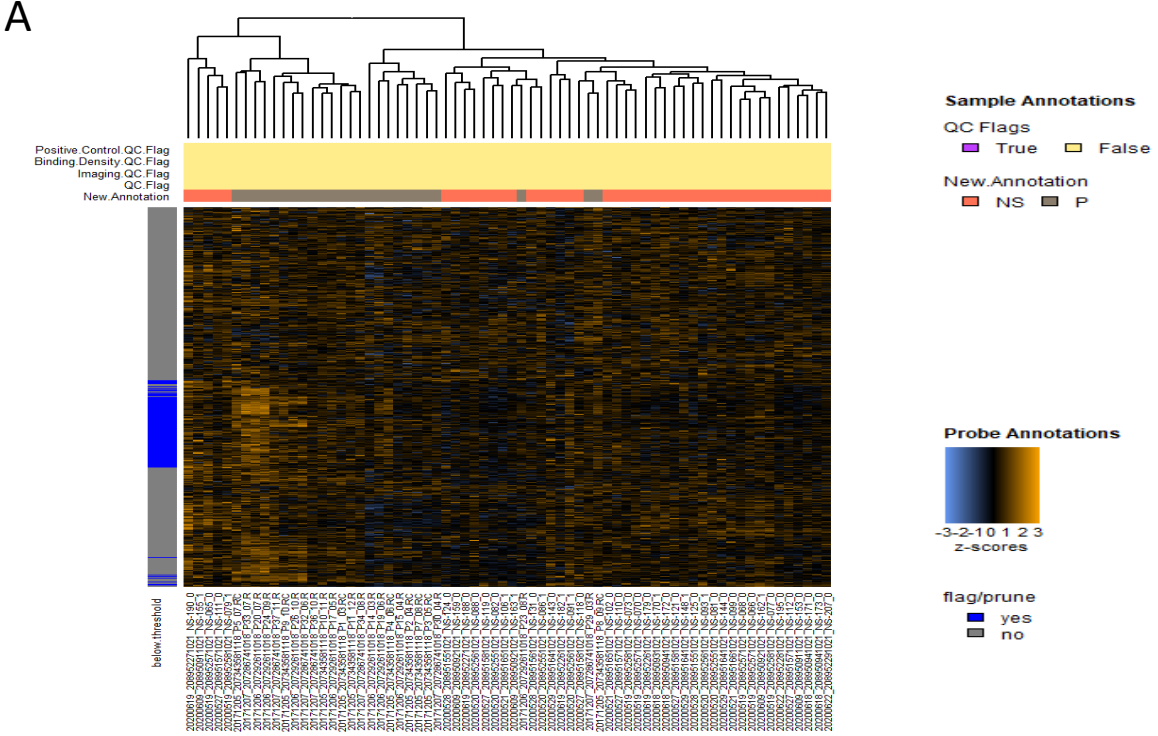
Figure 4. **WES analysis revealed genetic heterogeneity in PCNSL. (A-B)** Oncoplot showing the mutational landscape of PCNSL and cell lines derived from ABC-PCNSL and systemic ABC-DLBCL. The tracks below the Oncoplot show additional pathological features, including COO, LymphGen subgroup assignment, and PCNSL expression subtype. The majority (59%) of ABC-PCNSL patient samples were identified as "MCD", 17% as "MCD-composite", and 24% as "Other". **(C)** Bar plot showing the percentage of PCNSL with mutations in common aSHM regions described in DLBCL. **(D)** Fisher's exact tests were conducted between ABC-PCNSL (blue) and GCB-PCNSL (orange). The forest plot and corresponding bar plot reveals mutations in *GNA13, CD74, CARD11,* and *UBE2A* as significantly ($q < 0.1$) associated with GCB-PCNSL.

Figure 5 **PCNSL demonstrated distinct genetic features compared to nodal MCD-ABC-DLBCL. (A)** Boxplots comparing tumor mutation burden (TMB) of PCNSL compared to nodal ABC-DLBCL and nodal ABC-DLBCL classified as MCD by the LymphGen subgroup. PCNSL showed a significantly ($P < 0.01$, t-test) higher tumor mutation burden (TMB) relative to systemic ABC-DLBCL. However, compared to MCD-ABC-DLBCL, the TMB of PCNSL was more similar to that of the MCD-ABC-DLBCL. **(B)** Fisher's exact tests were conducted between PCNSL (pink) and MCD-ABC-DLBCL (green). The forest plot and corresponding bar

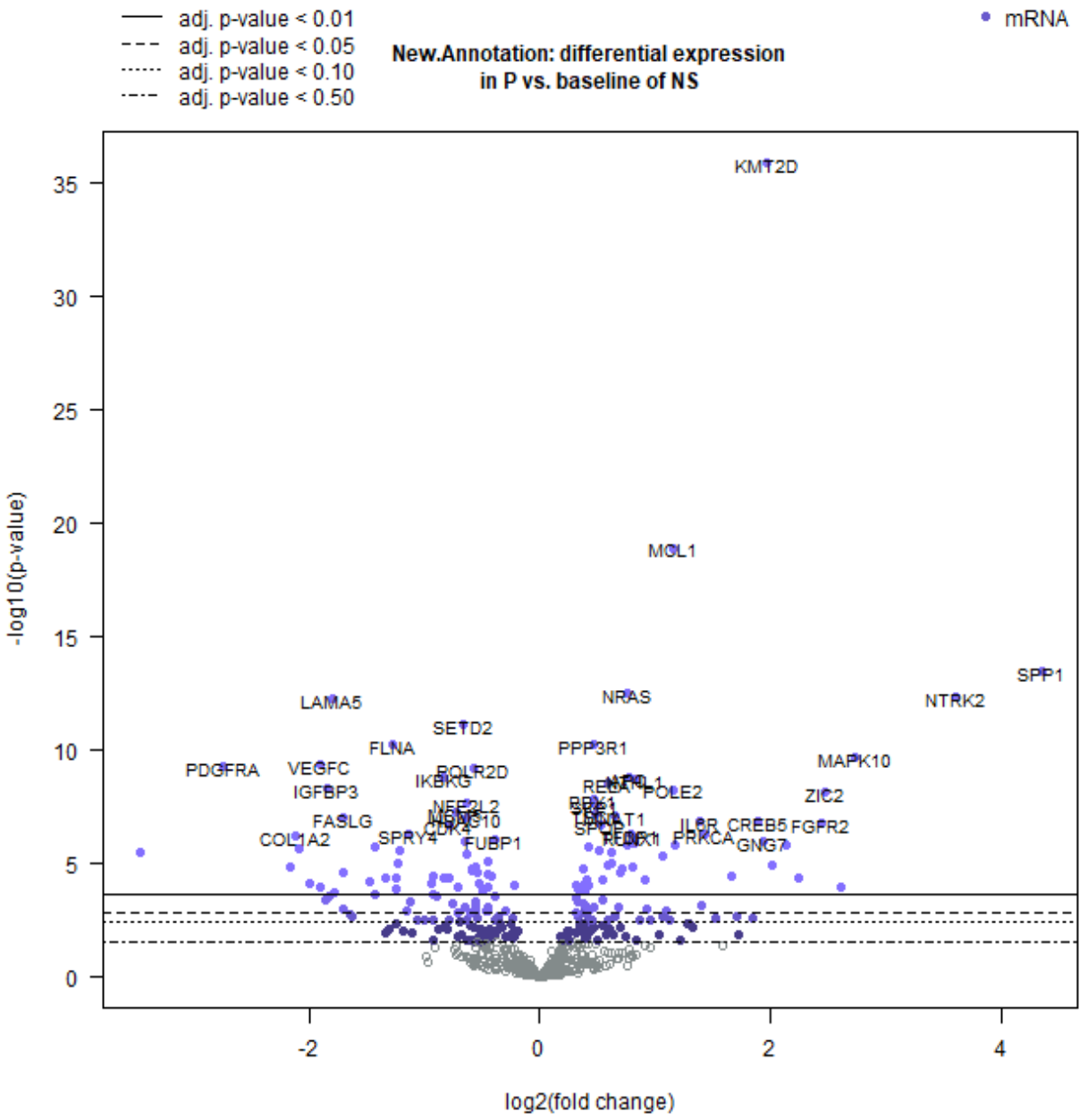
plot reveals a significant enrichment ($q < 0.1$, Fisher's exact test) in mutations in *SOCS1*, *HNRNPU*, *GNAI2*, *GNA13, CXCR4,* and *CIITA* in PCNSL compared to MCD-ABC-DLBCL. **(C)** Copy number profiles highlighting amplifications and deletions per chromosome of PCNSL (top) and MCD-ABC-DLBCL (bottom). **(D-E)** A visualization of significant amplification and deletion peaks in the aggregate of PCNSL and MCD-ABC-DLBCL, and genes associated with these peaks. Forest plot highlighting significantly enriched peaks in the aggregate of PCNSL and MCD-ABC-DLBCL as defined by GISTIC v.2.0.23. Genes of interest associated to these peaks are annotated beside the peak.

Figure 6. **Two PCNSL cell lines analogous to P1 and P2 were shown to have variable sensitivity to conventional anti-lymphoma chemotherapy, BTK, FGFR, and PI3K inhibitors. (A)** Two ABC-PCNSL cell lines (TK and HKBML) and three systemic ABC-DLBCL cell lines were tested on a PanCancer Pathways panel. Heatmap of pathway score indicated some similarity between the PCNSL cell lines and two transcriptional subgroups identified in PCNSL patient samples. **(B)** MTS assay was performed 72 hours after five DLBCL cell lines were treated with a specific BTK inhibitor, Ibrutinib. Shown is the result of a representative experiment with triplicates. **(C)** Western blotting reveals potential correlation between activated AKT and Ibrutinib sensitivities in the tested cell lines. **(D)** The PCNSL cell lines were further tested by MTT assay for responses to four clinical drugs for PCNSL and lymphoma (MTX, TMZ, pomalidomide and venetoclax) and to two inhibitors against AKT (Afuresertib) or FGFR kinase (BGJ398). **(E)** Combination of MTX with FGF inhibitor generated a likely synergistic antitumor activity on HKBML.

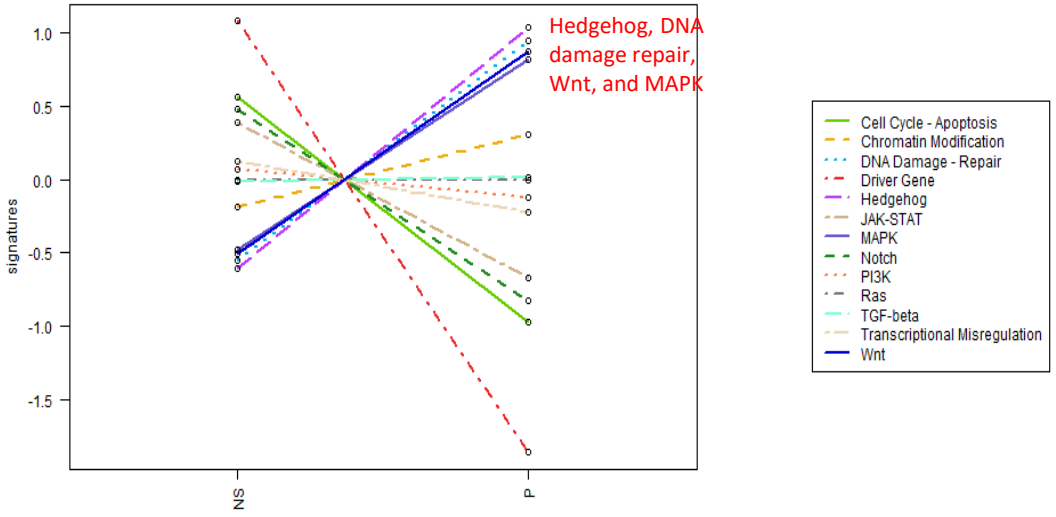
A


Positive.Control.QC.Flag
Binding.Density.QC.Flag
Imaging.QC.Flag
QC.Flag
New.Annotation
below.threshold
Sample Annotations
QC Flags
True
False
New.Annotation
NS
P
Probe Annotations
-3 -2 -1 0 1 2 3
z-scores
flag/prune
yes
no

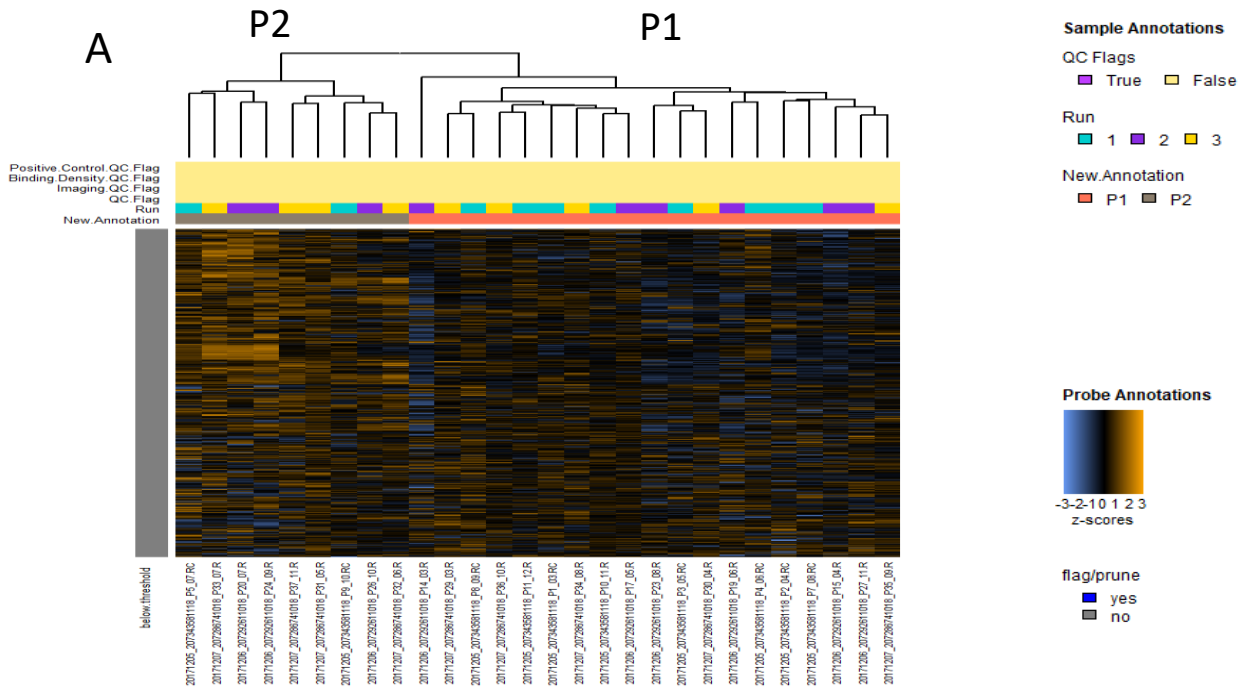
B

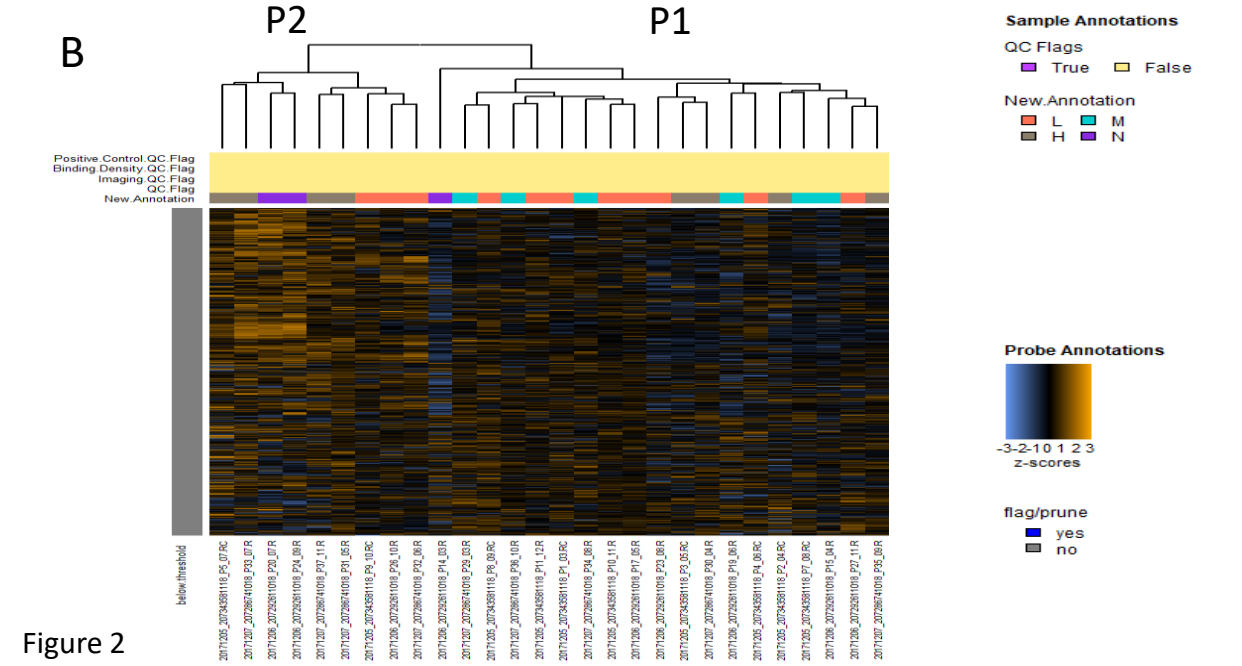

adj. p-value < 0.01
adj. p-value < 0.05
adj. p-value < 0.10
adj. p-value < 0.50
mRNA
New.Annotation: differential expression
in P vs. baseline of NS
-log10(p-value)
log2(fold change)
KMT2D
MGL1
SPP1
NTRK2
NRAS
LAMA5
SETD2
FLNA
PPP3R1
PDGFRA
VEGFC
IGFBP3
IKBKG
POLR2D
MAPK10
ZIC2
POLE2
FASLG
COL1A2
SPRY4
FUBP1
CREB5
FGFR2
GNG7
PRKCA

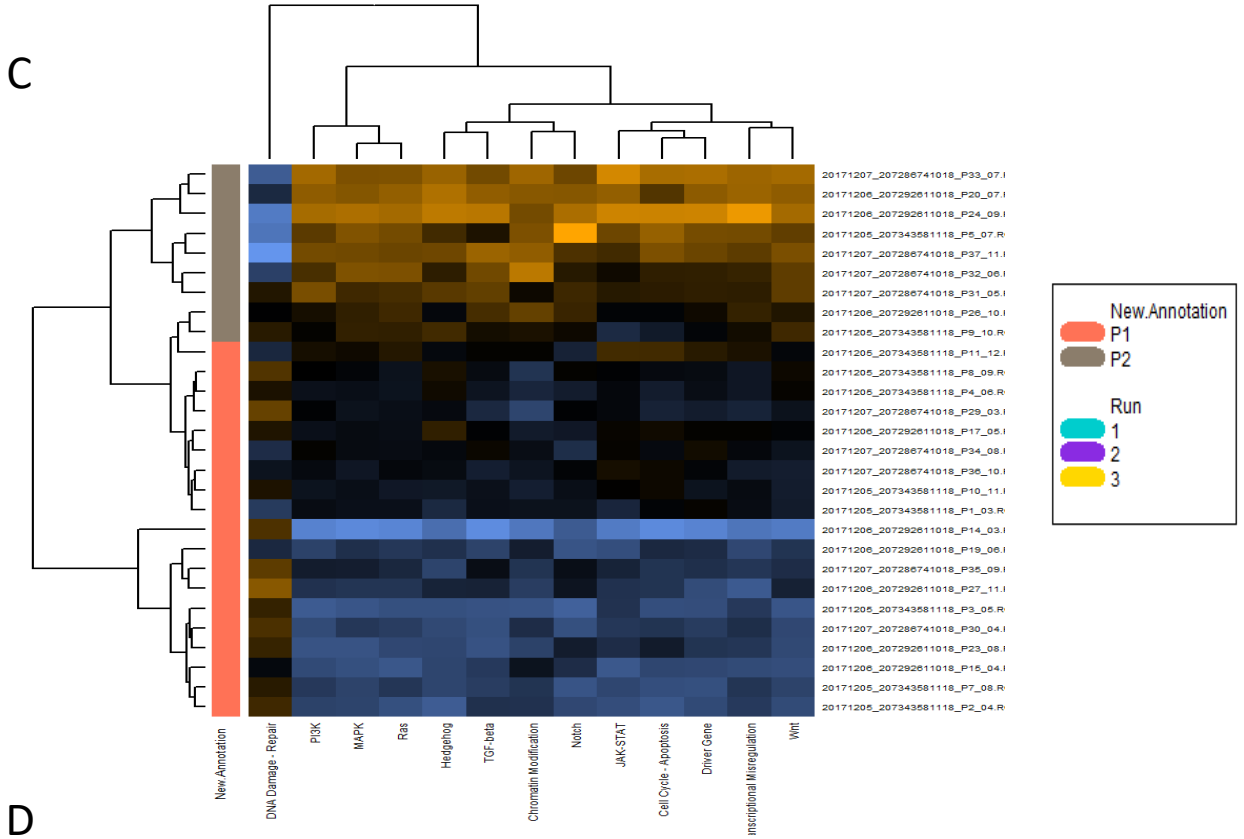
C


Hedgehog, DNA damage repair, Wnt, and MAPK
signatures
NS
P
Cell Cycle - Apoptosis
Chromatin Modification
DNA Damage - Repair
Driver Gene
Hedgehog
JAK-STAT
MAPK
Notch
PI3K
Ras
TGF-beta
Transcriptional Misregulation
Wnt

Figure


A
P2
P1
Positive.Control.QC.Flag
Binding.Density.QC.Flag
Imaging.QC.Flag
QC.Flag
Run
New.Annotation
below.threshold
Sample Annotations
QC Flags
True
False
Run
1
2
3
New.Annotation
P1
P2
Probe Annotations
-3 -2 -1 0 1 2 3
z-scores
flag/prune
yes
no


B
P2
P1
Positive.Control.QC.Flag
Binding.Density.QC.Flag
Imaging.QC.Flag
QC.Flag
New.Annotation
below.threshold
Sample Annotations
QC Flags
True
False
New.Annotation
L
M
H
N
Probe Annotations
-3 -2 -1 0 1 2 3
z-scores
flag/prune
yes
no


C
New.Annotation
DNA Damage - Repair
PI3K
MAPK
Ras
Hedgehog
TGF-beta
Chromatin Modification
Notch
JAK-STAT
Cell Cycle - Apoptosis
Driver Gene
inscriptional Misregulation
Wnt
New.Annotation
P1
P2
Run
1
2
3

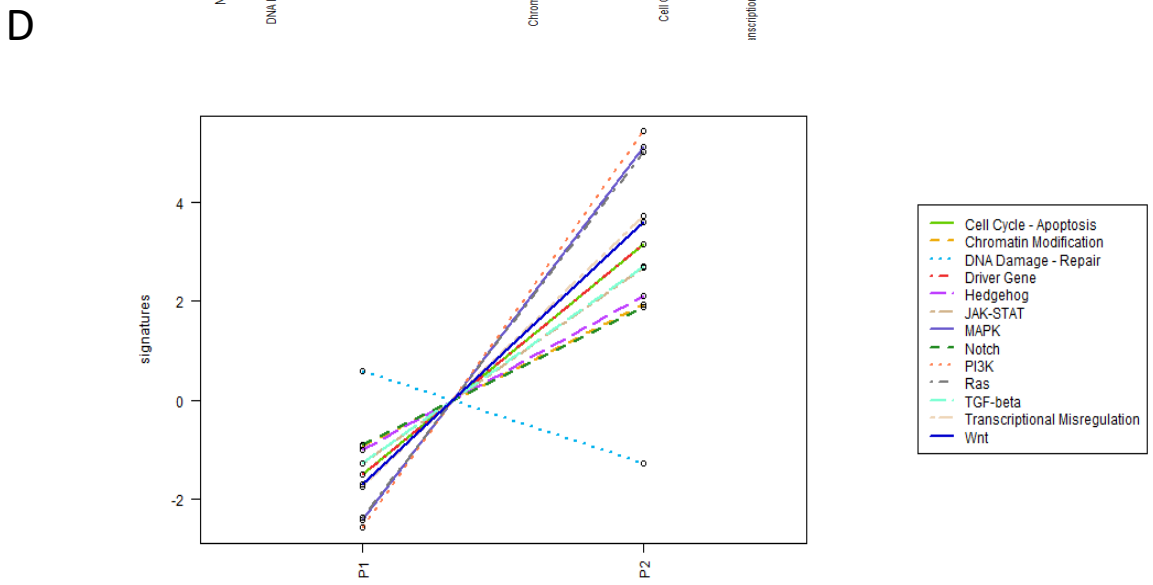
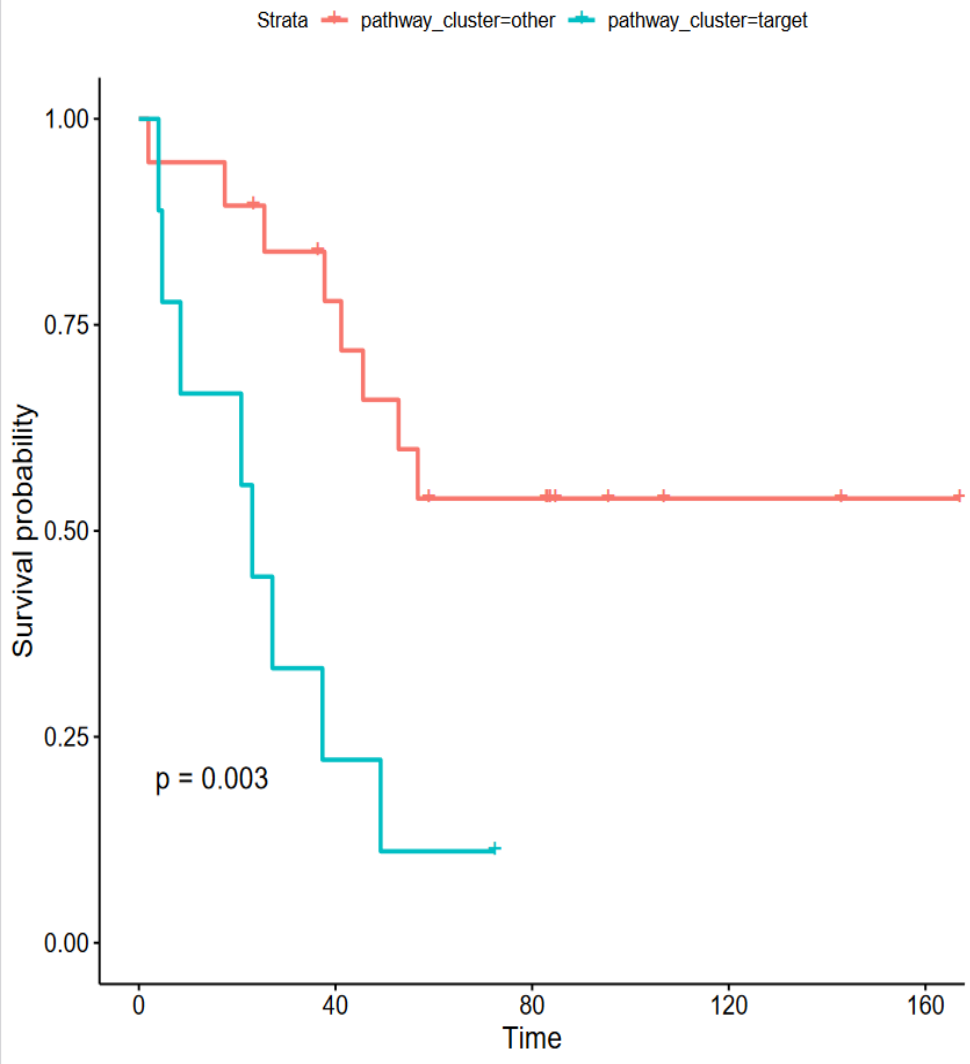

D
signatures
4
2
0
-2
P1
P2
Cell Cycle - Apoptosis
Chromatin Modification
DNA Damage - Repair
Driver Gene
Hedgehog
JAK-STAT
MAPK
Notch
PI3K
Ras
TGF-beta
Transcriptional Misregulation
Wnt

Figure 2

A



B

Top 25 differentially expressed genes (P2 vs P1)

	Log2 fold change	P-value	BY.p.value	Gene.sets
HDAC11-mRNA	1.81	2.05E-07	0.00109	Chromatin Modification
GRIN1-mRNA	4.86	4.64E-06	0.0123	Ras
IL1R1-mRNA	1.73	1.20E-05	0.0212	Cell Cycle - Apoptosis, MAPK
CDK4-mRNA	-1	1.79E-05	0.0238	Cell Cycle - Apoptosis, PI3K
MDM2-mRNA	-0.802	2.81E-05	0.0298	Cell Cycle - Apoptosis, PI3K, Transcriptional Misregulation
PPP2R2C-mRNA	2.52	5.47E-05	0.0356	PI3K
ID4-mRNA	1.65	5.49E-05	0.0356	TGF-beta
CREB5-mRNA	2.16	6.81E-05	0.0356	PI3K
NTRK2-mRNA	2.27	6.91E-05	0.0356	MAPK
SPRY2-mRNA	1.29	7.19E-05	0.0356	JAK-STAT
LIFR-mRNA	1.93	7.69E-05	0.0356	JAK-STAT
APC-mRNA	0.625	8.04E-05	0.0356	Driver Gene, Wnt
CCNA2-mRNA	-1.02	8.78E-05	0.0358	Cell Cycle - Apoptosis
MSH2-mRNA	-0.732	0.000115	0.0437	Driver Gene
GNAQ-mRNA	1.26	0.000152	0.0538	Driver Gene
CACNA2D1-mRNA	1.28	0.000221	0.0733	MAPK
TSPAN7-mRNA	2.55	0.000281	0.0841	Transcriptional Misregulation
BAD-mRNA	0.64	0.000285	0.0841	Cell Cycle - Apoptosis, PI3K, Ras
GNA11-mRNA	0.917	0.000305	0.0853	Driver Gene
MAPT-mRNA	2.19	0.000324	0.0861	MAPK
DTX4-mRNA	1.67	0.000374	0.0944	Notch
UBE2T-mRNA	-0.791	0.000459	0.111	DNA Damage - Repair
HMGA1-mRNA	-0.677	0.000495	0.114	Chromatin Modification
HSPA1A-mRNA	1.93	0.000516	0.114	MAPK
CDC6-mRNA	-0.882	0.000583	0.12	Cell Cycle - Apoptosis

Figure 3

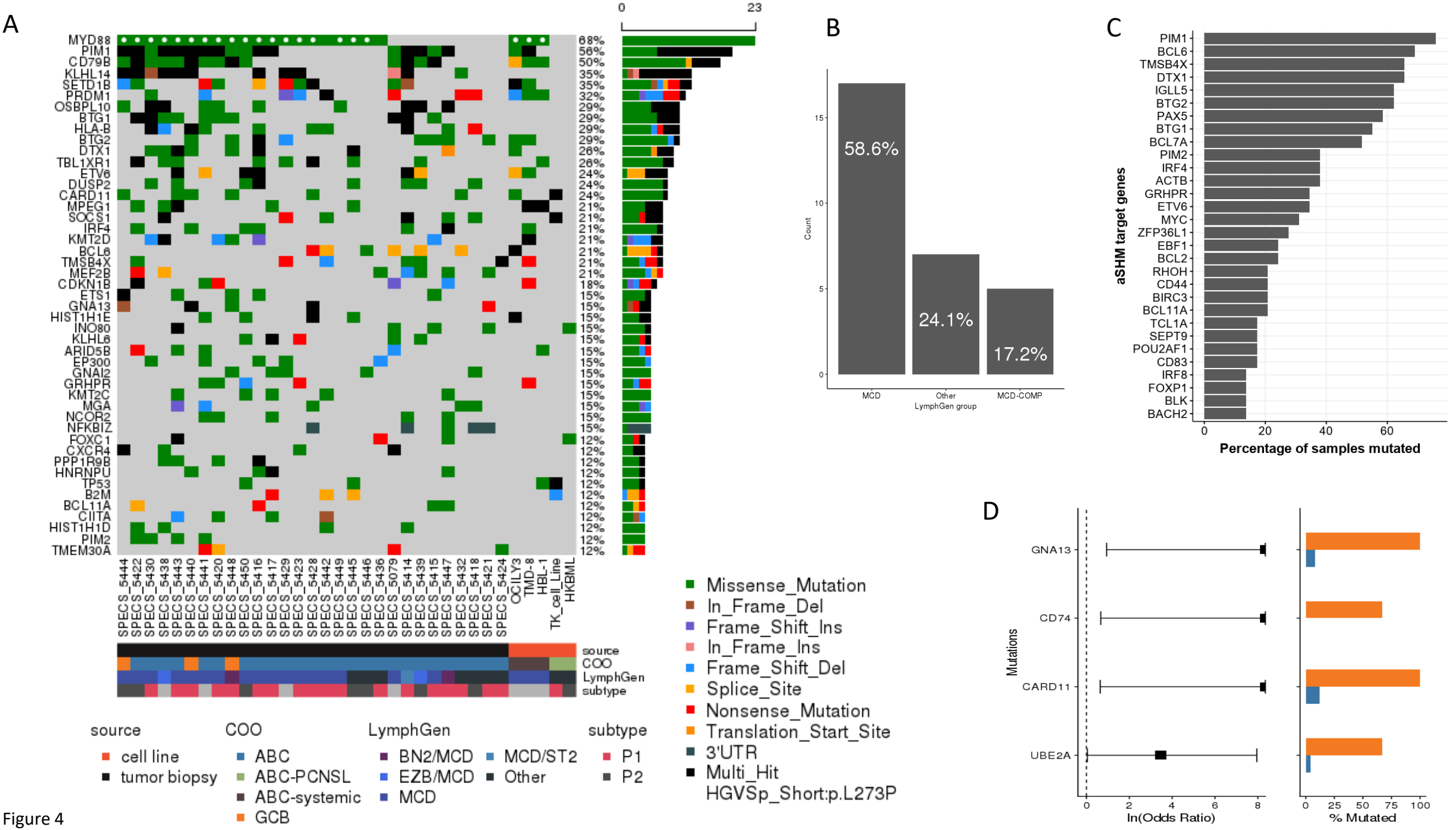
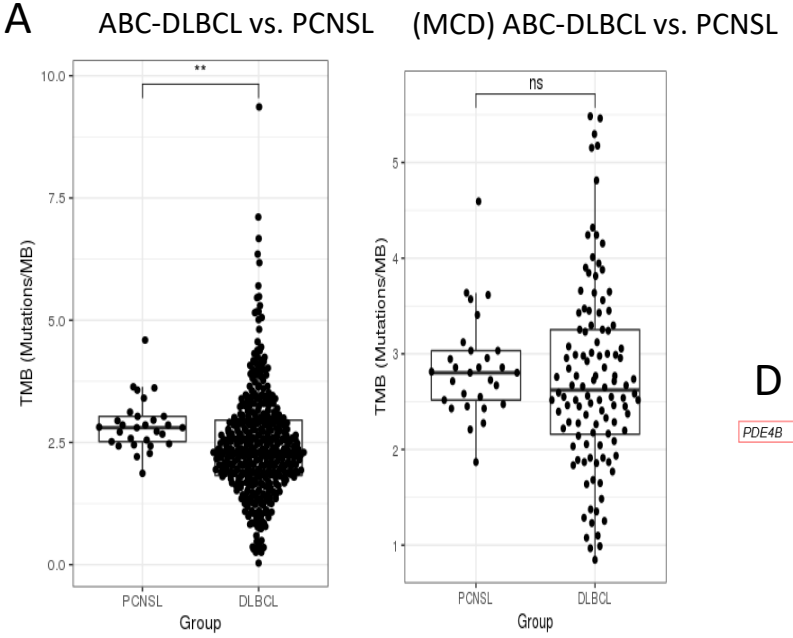
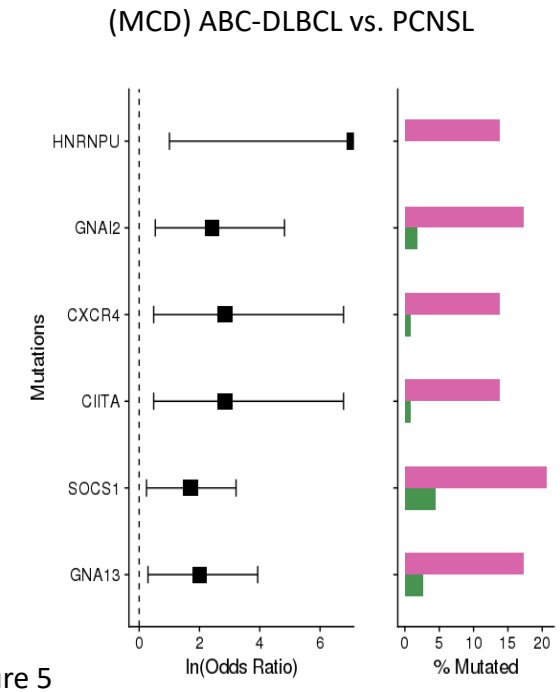

A
MYD88 68%
PIM1 56%
CD79B 50%
KLHL14 35%
SETD1B 35%
PRDM1 32%
OSBPL10 29%
BTG1 29%
HLA-B 29%
BTG2 29%
DTX1 26%
TBL1XR1 26%
ETV6 24%
DUSP2 24%
CARD11 24%
MPEG1 21%
SOCS1 21%
IRF4 21%
KMT2D 21%
BCL6 21%
TMSB4X 21%
MEF2B 21%
CDKN1B 18%
ETS1 15%
GNA13 15%
HIST1H1E 15%
INO80 15%
KLHL6 15%
ARID5B 15%
EP300 15%
GNAI2 15%
GRHPR 15%
KMT2C 15%
MGA 15%
NCOR2 15%
NFKBIZ 15%
FOXC1 12%
CXCR4 12%
PPP1R9B 12%
HNRNPU 12%
TP53 12%
B2M 12%
BCL11A 12%
CIITA 12%
HIST1H1D 12%
PIM2 12%
TMEM30A 12%
0
23
SPECS_5444
SPECS_5422
SPECS_5430
SPECS_5438
SPECS_5443
SPECS_5440
SPECS_5441
SPECS_5420
SPECS_5448
SPECS_5450
SPECS_5416
SPECS_5417
SPECS_5429
SPECS_5423
SPECS_5428
SPECS_5442
SPECS_5449
SPECS_5445
SPECS_5446
SPECS_5436
SPECS_5079
SPECS_5414
SPECS_5439
SPECS_5415
SPECS_5447
SPECS_5432
SPECS_5418
SPECS_5421
SPECS_5424
OCILY3
TMD-8
HBL-1
TK_cell_Line
HKBML
source
COO
LymphGen
subtype
source
cell line
tumor biopsy
COO
ABC
ABC-PCNSL
ABC-systemic
GCB
LymphGen
BN2/MCD
EZB/MCD
MCD
MCD/ST2
Other
subtype
P1
P2
Missense_Mutation
In_Frame_Del
Frame_Shift_Ins
In_Frame_Ins
Frame_Shift_Del
Splice_Site
Nonsense_Mutation
Translation_Start_Site
3'UTR
Multi_Hit
HGVSp_Short:p.L273P
B
Count
58.6%
24.1%
17.2%
MCD
Other
MCD-COMP
LymphGen group
C
aSHM target genes
PIM1
BCL6
TMSB4X
DTX1
IGLL5
BTG2
PAX5
BTG1
BCL7A
PIM2
IRF4
ACTB
GRHPR
ETV6
MYC
ZFP36L1
EBF1
BCL2
RHOH
CD44
BIRC3
BCL11A
TCL1A
SEPT9
POU2AF1
CD83
IRF8
FOXP1
BLK
BACH2
Percentage of samples mutated
D
Mutations
GNA13
CD74
CARD11
UBE2A
ln(Odds Ratio)
% Mutated

Figure 4

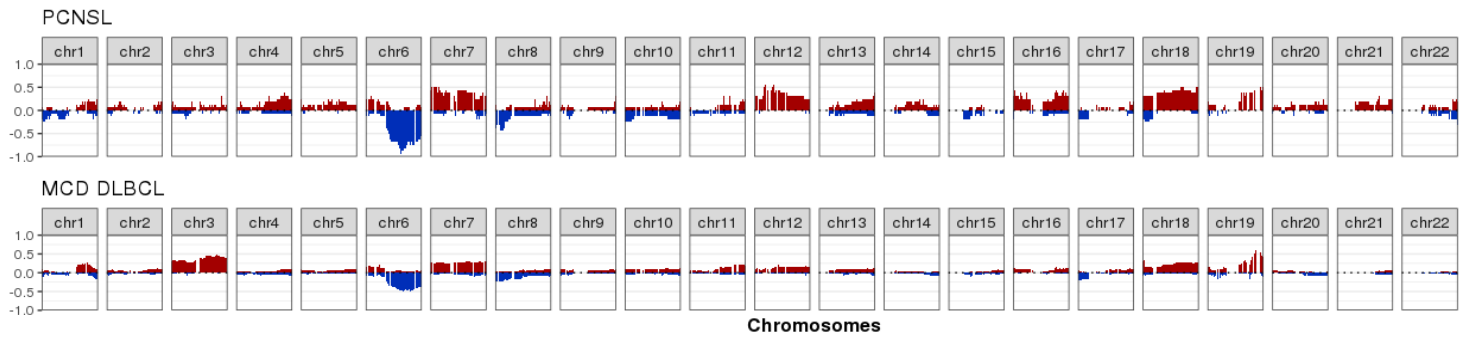
A



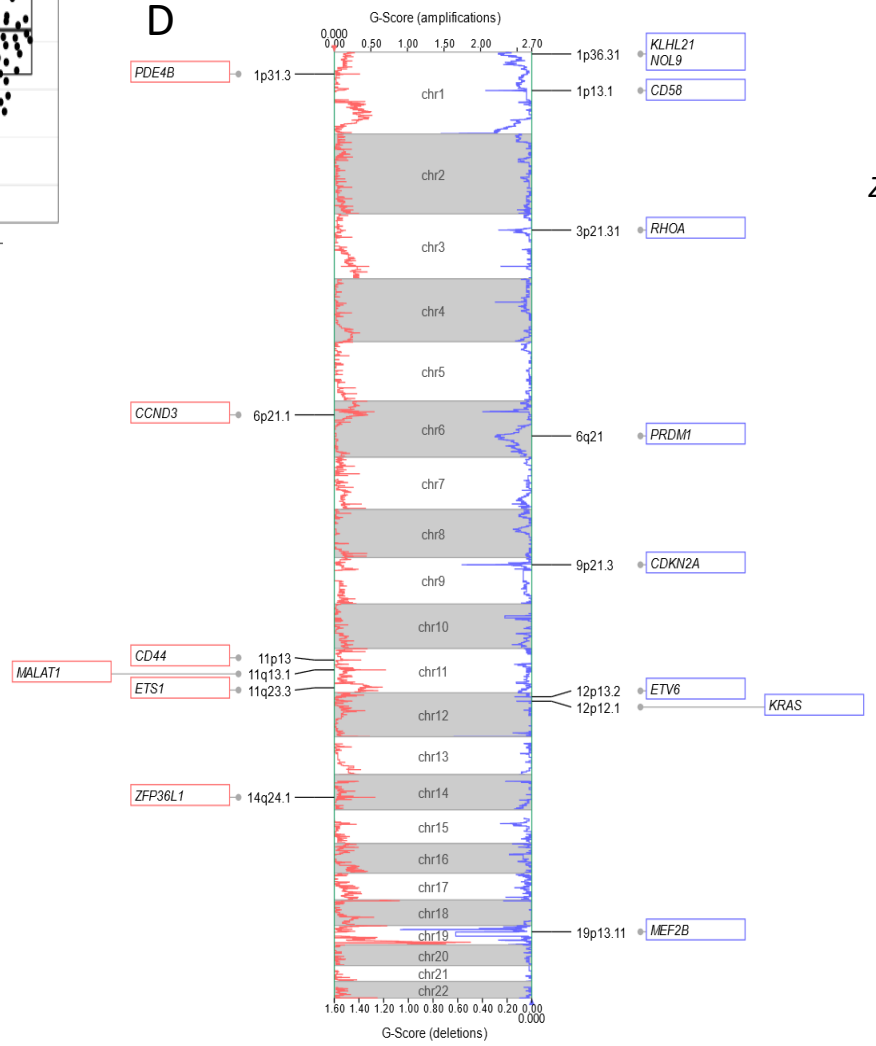
B



C



D



E

	name	Events_PCNSL	Events_DLBCL
ZFP36L1	14q24.1_Amp	11/16	2/89
	22q13.33_Amp	11/16	2/89
	16p13.3_Del	5/16	2/89
PRDM1	6q21_Del	15/16	42/89
	12q24.33_Amp	11/16	14/89
	16q24.1_Amp	9/16	10/89
RHOA	3p21.31_Del	6/16	5/89
PDE4B	1p31.3_Amp	5/16	4/89
	7p12.2_Amp	11/16	21/89
	13q32.3_Amp	7/16	10/89
	4q32.3_Amp	6/16	8/89
	16p13.13_Amp	6/16	8/89
	12p13.31_Amp	7/16	11/89
	8q24.22_Amp	6/16	10/89
MEF2B	19p13.2_Del	6/16	10/89
MALAT1	11q13.1_Amp	7/16	13/89
	7q36.3_Amp	10/16	24/89
	12q24.33_Del	0/16	30/89

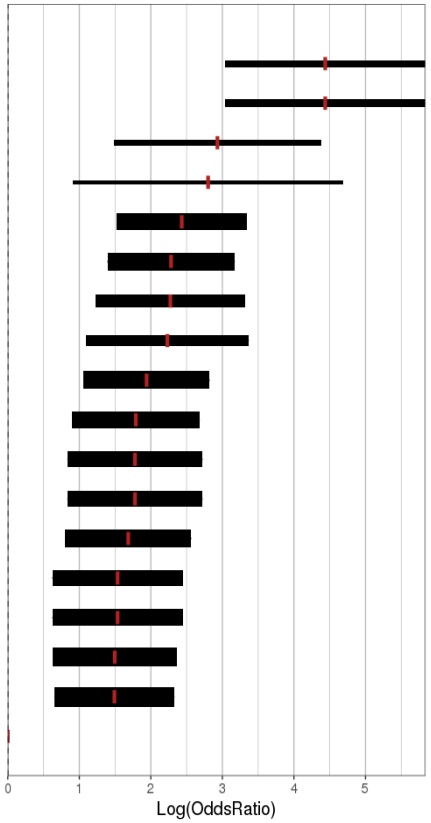
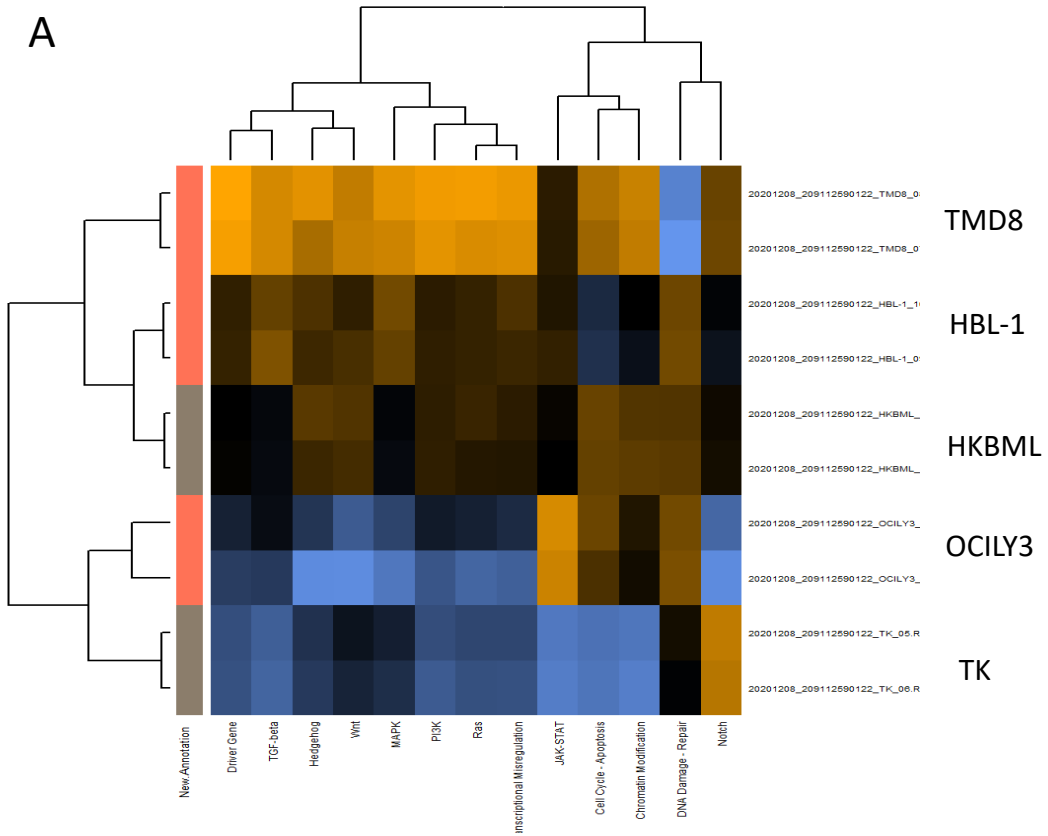
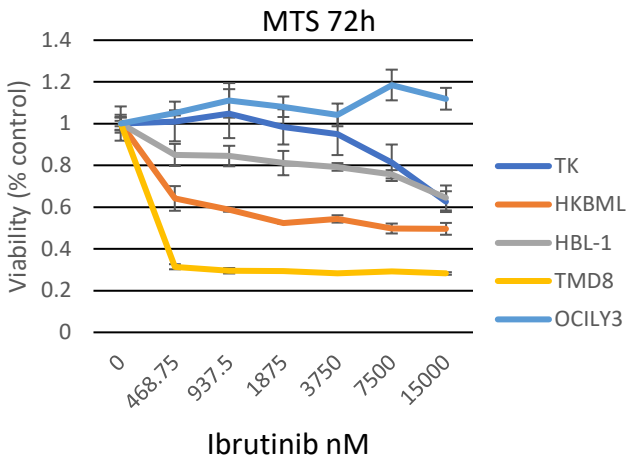
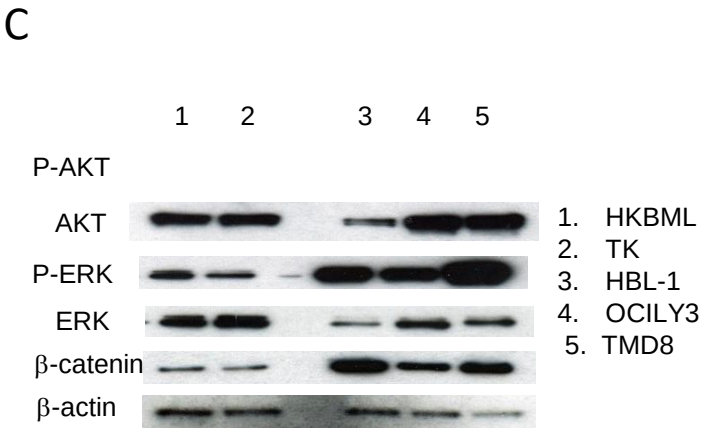
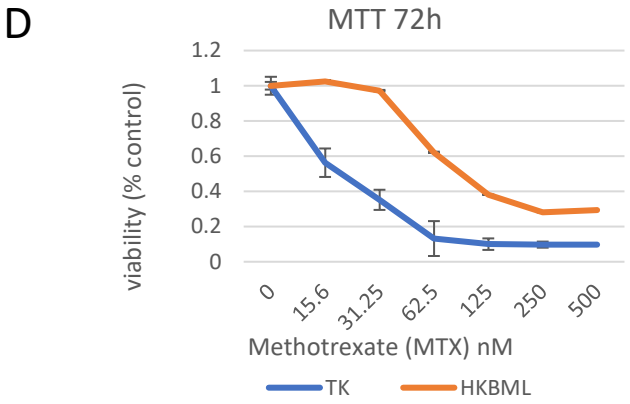


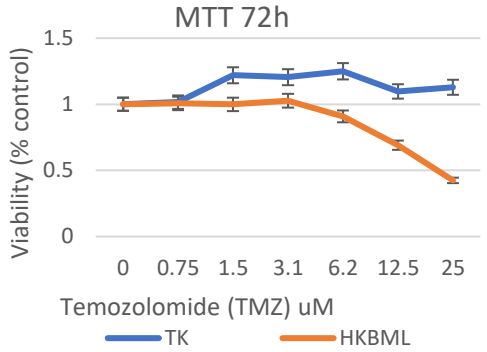
Figure 5

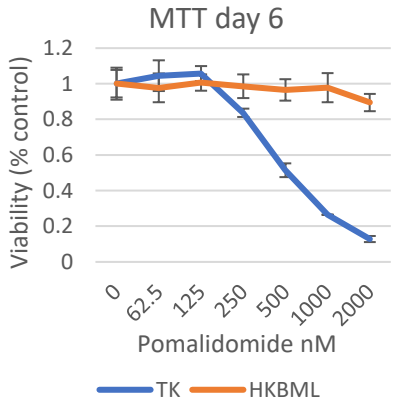

A
TMD8
HBL-1
HKBML
OCILY3
TK
New Annotation
Driver Gene
TGF-beta
Hedgehog
Wnt
MAPK
PI3K
Ras
JAK-STAT
Cell Cycle - Apoptosis
Chromatin Modification
DNA Damage - Repair
Notch

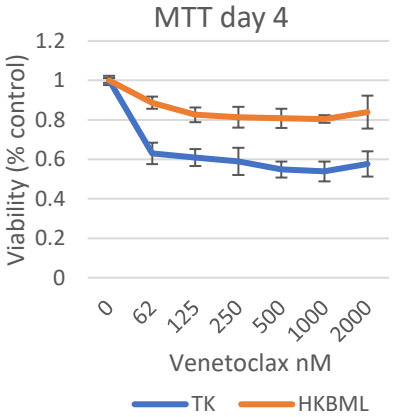

B
MTS 72h
Viability (% control)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
468.75
937.5
1875
3750
7500
15000
Ibrutinib nM
TK
HKBML
HBL-1
TMD8
OCILY3


C
1
2
3
4
5
P-AKT
AKT
P-ERK
ERK
β-catenin
β-actin
1. HKBML
2. TK
3. HBL-1
4. OCILY3
5. TMD8


D
MTT 72h
viability (% control)
1.2
1
0.8
0.6
0.4
0.2
0
0
15.6
31.25
62.5
125
250
500
Methotrexate (MTX) nM
TK
HKBML


MTT 72h
Viability (% control)
1.5
1
0.5
0
0
0.75
1.5
3.1
6.2
12.5
25
Temozolomide (TMZ) uM
TK
HKBML


MTT day 6
Viability (% control)
1.2
1
0.8
0.6
0.4
0.2
0
0
62.5
125
250
500
1000
2000
Pomalidomide nM
TK
HKBML


MTT day 4
Viability (% control)
1.2
1
0.8
0.6
0.4
0.2
0
0
62
125
250
500
1000
2000
Venetoclax nM
TK
HKBML

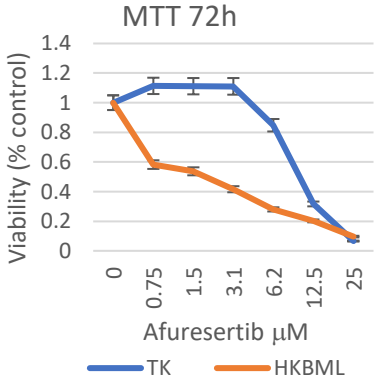

MTT 72h
Viability (% control)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.75
1.5
3.1
6.2
12.5
25
Afuresertib µM
TK
HKBML

Figure 6