

Integrative genomic and transcriptomic analysis in plasmablastic lymphoma identifies disruption of key regulatory pathways

Hanno M. Witte,^{1,2,*} Axel Künstner,^{3-5,*} Nadine Hertel,^{1,6,*} Heinz-Wolfram Bernd,⁶ Veronica Bernard,⁶ Stephanie Stölting,⁶ Hartmut Merz,⁶ Nikolas von Bubnoff,^{1,5} Hauke Busch,^{3-5,†} Alfred C. Feller,^{6,†} and Niklas Gebauer^{1,5,†}

¹Department of Hematology and Oncology, University Hospital of Schleswig-Holstein, Campus Lübeck, Lübeck, Germany; ²Department of Hematology and Oncology, Federal Armed Forces Hospital Ulm, Ulm, Germany; ³Medical Systems Biology Group, and ⁴Institute for Cardiogenetics, University of Lübeck, Lübeck, Germany; ⁵University Cancer Center Schleswig-Holstein, University Hospital of Schleswig-Holstein, Campus Lübeck, Lübeck, Germany; and ⁶Hämatopathologie Lübeck, Reference Centre for Lymph Node Pathology and Hematopathology, Lübeck, Germany

Key Points

- WES coupled with RNA-sequencing of a large PBL cohort reveals genetic drivers of oncogenesis in RTK-RAS, NF-κB, and JAK/STAT signaling.
- The mutational landscape and SCN data emphasize the distinctness of EBV⁺/EBV⁻ PBL from both DLBCL and multiple myeloma.

Plasmablastic lymphoma (PBL) represents a clinically heterogeneous subtype of aggressive B-cell non-Hodgkin lymphoma. Targeted-sequencing studies and a single-center whole-exome sequencing (WES) study in HIV-positive patients recently revealed several genes associated with PBL pathogenesis; however, the global mutational landscape and transcriptional profile of PBL remain elusive. To inform on disease-associated mutational drivers, mutational patterns, and perturbed pathways in HIV-positive and HIV-negative PBL, we performed WES and transcriptome sequencing (RNA-sequencing) of 33 PBL tumors. Integrative analysis of somatic mutations and gene expression profiles was performed to acquire insights into the divergent genotype–phenotype correlation in Epstein-Barr virus–positive (EBV⁺) and EBV⁻ PBL. We describe a significant accumulation of mutations in the JAK signal transducer and transcription activator (*OSMR*, *STAT3*, *PIM1*, and *SOCS1*), as well as receptor tyrosine-kinase RAS (*ERBB3*, *NRAS*, *PDGFRB*, and *NTRK*) pathways. We provide further evidence of frequent perturbances of NF-κB signaling (*NFKB2* and *BTK*). Induced pathways, identified by RNA-sequencing, closely resemble the mutational profile regarding alterations accentuated in interleukin-6/JAK/STAT signaling, NF-κB activity, and MYC signaling. Moreover, class I major histocompatibility complex–mediated antigen processing and cell cycle regulation were significantly affected by EBV status. An almost exclusive upregulation of phosphatidylinositol 3-kinase/AKT/mTOR signaling in EBV⁺ PBL and a significantly induced expression of *NTRK3* in concert with recurrent oncogenic mutations in EBV⁻ PBL hint at a specific therapeutically targetable mechanism in PBL subgroups. Our characterization of a mutational and transcriptomic landscape in PBL, distinct from that of diffuse large B-cell lymphoma and multiple myeloma, substantiates the pathobiological independence of PBL in the spectrum of B-cell malignancies and thereby refines the taxonomy for aggressive lymphomas.

Submitted 9 June 2021; accepted 2 October 2021; prepublished online on *Blood Advances* First Edition 29 October 2021; final version published online 21 January 2022. DOI 10.1182/bloodadvances.2021005486.

*H.M.W., A.K., and N.H. contributed equally to this study.

†H.B., A.C.F., and N.G. are joint senior authors.

Raw fastq files have been deposited in the European genome-phenome archive (EGA) under the accession number EGAD00001006795. The samples PBL_WES_005200 (EGA ID: EGAR00002570178) and PBL_WES_004923 (EGA

ID: EGAR00002570203) were retrieved from the same patient, at the same time point, but from 2 different locations and uploaded to EGA. In the current study, only PBL_WES_004923 was used for analysis.

The full-text version of this article contains a data supplement.

© 2022 by The American Society of Hematology. Licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0), permitting only noncommercial, nonderivative use with attribution. All other rights reserved.

Introduction

Plasmablastic lymphoma (PBL) is a rare and clinically aggressive large B-cell lymphoma of post-germinal center origin.¹ Both immunoblastic and/or plasmablastic morphology alongside immunohistochemical features of terminal B-cell differentiation (CD79a⁺, MUM-1⁺, CD38/CD138⁺; CD20⁻, PAX5⁻) illustrate its biological and clinical behavior intermediate between diffuse large B-cell lymphoma (DLBCL) and multiple myeloma (MM).²⁻⁴ According to chromogenic in situ hybridization, Epstein-Barr virus (EBV)-encoded RNA is detectable in ~70% of cases.

PBL exhibits a pronounced incidence in HIV-positive patients, although observations regarding the presence of an underlying HIV infection in ~50% to 60% of cases partially contradict initial reports claiming a near pathognomonic association.³ The disease can indeed also occur in patients with immunodeficiency of other cause (eg, secondary to organ transplant recipients or age-related immunosenescence) but is also occasionally encountered in immunocompetent individuals.⁵⁻⁷ The incidence of PBL accounts for ~2% of all lymphomas arising in patients with AIDS.⁷

Due to the rarity of PBL and the ongoing lack of an evidence-based standard treatment approach, resulting from a deficiency in prospectively gathered clinical data, translational research remains challenging.^{7,8} Compared with other aggressive B-cell lymphomas, PBL outcomes are dismal. Despite recent prognostic advances with intensified treatment approaches and alternative therapeutic options, including CD38-directed monoclonal antibody, proteasome inhibitors, and immunomodulatory treatment strategies, these are highly limited in their applicability in elderly and frail patients, warranting the need for alternative, if possible, specifically targeted therapeutic approaches.^{2,9,10}

The most frequently detected chromosomal aberrations are low-level amplifications and/or translocations involving *MYC*, frequently juxtaposing *MYC* to the immunoglobulin heavy chain locus, present in approximately one-half of cases.¹¹⁻¹³ Comparative genomic hybridization studies revealed a copy number variation profile, intermediate between DLBCL and MM, while favoring an association with the former.¹⁴

PBL was shown to exhibit a transcriptional profile distinct from DLBCL regarding B-cell receptor signaling and targets of the transcription factors *MYC* and *MYB*.¹⁵ Moreover, preliminary studies suggested a significant impact of EBV status on gene expression profiles affecting T-cell activation signatures and mechanisms of immune escape.^{16,17}

Recently, in a targeted sequencing study, Garcia-Reyero et al reported on an elevated frequency of *PRDM1/BLIMP1* and *STAT3* mutations together with other genetic aberrations, including mutations in *BRAF* and *NOTCH* signaling.^{18,19} According to another targeted sequencing study that focused on a panel of oncogenes implicated in large B-cell lymphomas, the overall mutational burden in PBL was markedly low. Only 60% of patients were thus shown to harbor oncogenic mutations.¹⁹ Together with its morphologic and immunophenotypical peculiarities, this finding suggests the presence of alternative drivers of lymphomagenesis in PBL that are distinct from those associated with other aggressive B-cell lymphomas. Of note, several recent studies on the mutational landscape of PBL

hinted at pervasive mutations in *JAK-STAT3* and *RAS-MAPK* as well as noncanonical *NF-κB* signaling. However, these studies, including 15 cases studied by using whole-exome sequencing (WES), were exclusively investigating HIV-positive patients or PBL as a posttransplantation lymphoproliferative disease.²⁰⁻²² Most recently, in an amplicon panel-based and array-based study, Ramis-Zaldivar et al²³ provided a first report, investigating divergences between HIV-positive and HIV-negative cases alongside the impact of an underlying EBV infection. They observed that EBV-negative PBL cases harbored a higher mutational and copy number load and more frequent *TP53*, *CARD11*, and *MYC* mutations, whereas EBV-positive PBL tended to have more mutations affecting the *JAK-STAT* pathway.

Although previous studies have hinted at a limited number of potential genetic drivers in PBL lymphomagenesis, a comprehensive characterization of the genomic and transcriptomic landscape in a representative cohort of patients is lacking thus far. The current study combined WES and RNA sequencing (RNA-seq) of 33 PBL samples. Through the integrated analysis of mutational and gene expression data, we outline oncogenic drivers, copy number alterations, and pathway perturbations. We thus refine the molecular taxonomy of terminally differentiated B-cell malignancies and substantiate the independent role of PBL, distinct from DLBCL and MM.

Materials and methods

Case selection and clinicopathologic characteristics

In a retrospective approach, we reviewed our institutional database to identify patients with PBL whose primary diagnostic biopsy specimen had been referred to the reference center for Haematopathology University Hospital Schleswig Holstein Campus Lübeck and Hämatopathologie Lübeck for centralized histopathologic panel evaluation between January 1998 and December 2019. Further clinicopathologic characteristics are presented in the supplemental Materials and Methods.

This retrospective study was approved by the ethics committee of the University of Lübeck (reference no. 18-311) and conducted in accordance with the Declaration of Helsinki. Patients at the Reference Center for Hematopathology have provided written informed consent regarding routine diagnostic and academic assessment, including genomic studies of their biopsy specimen in addition to transfer of their clinical data.

Extraction of nucleic acids, WES, and RNA-seq

Genomic tumor DNA and tumor RNA were each extracted from 3 formalin-fixed, paraffin-embedded (FFPE) tissue sections of 5 μm thickness using a Maxwell RSC DNA FFPE kit and Maxwell RSC RNA FFPE kit (both Promega). WES was performed by using a hybrid capture approach with the SureSelect Human All Exon V6 library preparation kit (Agilent Technologies) followed by massively parallel sequencing of enriched exonic sequences on a NovaSeq platform (Illumina). All samples with successful library preparation (33 of 41) were taken forward to sequencing.

PBL samples were sequenced to an average depth of 201× (±50×; median, 203×). Sequencing was performed at Novogene (UK) Co., Ltd.

Raw fastq files have been deposited in the European genome-phenome archive (EGA) under the accession number EGAD00001006795. The transcriptional profile of PBL was investigated by RNA-seq, when exome sequencing had previously been conducted successfully. A Ribo-Zero Magnetic Kit (Human, Mouse, Rat) (Illumina) was used for ribosomal RNA depletion, and library preparation was performed by using the NEBNext UltraT Directional RNA Library Prep Kit (New England BioLabs). Sequencing in cases with successful library preparation (20 of 33) was performed at Novogene (UK) Co., Ltd., and all genomic studies were guided by the principles of the European Medicines Agency International Council for Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use guideline E18 on genomic sampling and management of genomic data.

Variant calling

As described in Gebauer et al,²⁴ raw fastq reads trimmed (adapter and quality values) using FASTP²⁵ (version 0.20.0; minimum length, 50 bp; maximum unqualified bases, 30%; trim tail set to 1) and trimmed reads were mapped to GRCh37/hg19 by using BWA MEM (version 0.7.15).²⁶ Resulting SAM files were cleaned and sorted and converted into BAM format using PICARD TOOLS (version 2.18.4). Next, mate-pair information was fixed, duplicates were removed, and base quality recalibration was performed by using PICARD TOOLS²⁷ and dbSNP version 138. Single nucleotide variants (SNVs) and short insertions as well as deletions (indels) were identified following the best practices for somatic mutation calling provided by GATK.²⁸ Briefly, the GATK MUTECT2²⁹ (version 4.1.5.0) algorithm was applied to all BAM files with GnomAD variants as germline resource and the b37 exome panel data as a panel of normal, capturing the expected repertoire of germline variants to be expected in a Central European study population. Afterward, FFPE read orientation artifacts were identified and removed according to GATK guidelines. Filtered variants were annotated by using ANNOVAR³⁰ (24 October 2019 version). Coverage for reference and alternative alleles for each variant was extracted by using VCFQUERY (VCFTOOLS version 0.1.13³¹). The top 20 frequently mutated genes (FLAGS³²) were removed from further analysis. Somatic variants were filtered as follows: minimum coverage of 40, minimum variant allele frequency of 10%, and only variants with a frequency <0.001 in 1000 genomes, GnomAD, or ExAC were considered for subsequent downstream analysis. To identify genes that are more often mutated than expected, MutSigCV (version 1.41)³³ was applied, and potential driver genes were identified using $P < .05$. Selected mutations in *PRDM1*, *SOCS1*, *STAT3*, and *TP53* in samples from which sufficient tissue was available for confirmatory investigations were subjected to either Sanger sequencing or amplicon-based next-generation sequencing (supplemental Figure 1; supplemental Table 2).

Details on data analysis regarding copy number aberrations, network diffusion analysis, gene expression profiling, and fusion gene detection, as well as subsequent statistical methods, are provided in the supplemental Materials and Methods.

Results

Clinical characteristics of the study group

We collected 41 cases of PBL at diagnosis with sufficient FFPE tissue samples for molecular studies (age range, 29-86 years); 33

cases were included in the final analysis, following 8 dropouts due to insufficient yield at library preparation for WES. An underlying HIV infection was detectable in 14 (42.4%) of 33 patients, and PBL tumor cells were shown to express EBV-encoded RNA in 20 (60.6%) of 33 cases. HIV-positive patients were significantly younger (mean age, 44 vs 70 years; $P = .0001$) and had a significantly lower international prognostic index (IPI; $P = .006$) and a trend toward lower-stage disease at diagnosis ($P = .076$). Whenever evaluable, human herpesvirus-8 was absent (immunohistochemistry, $n = 6$ [18.2%]), and serologic test results were available in 9 (27.3%) cases. The majority of patients in our study were male (23 of 33 [69.7%]) and presented with advanced-stage disease (19 of 33 stage III/IV [57.6%]) and an adverse prognostic constellation (18 of 33 IPI >2). Most patients received an intensive CHOP-like (cyclophosphamide, doxorubicin, vincristine, and prednisolone) therapeutic front-line regimen (21 of 33 [63.6%]), resulting in an overall response rate of 72.7% (24 of 33). Median progression-free and overall survival was 9 and 22 months, respectively. Two elderly, frail patients refused any type of anticancer treatment and rapidly died due to PBL progression.

Baseline characteristics of PBL cases included in the current study are briefly summarized in Table 1. Treatment outcome in our cohort was dismal overall, although slightly more favorable in the HIV-positive subgroup (supplemental Figure 2) as could be expected from previous studies (median overall survival, 19 vs 45 months) as significantly more HIV-negative patients received an intensive CHOP-like first-line regimen ($P = .001$).^{7,34} Most samples (except case 01) included in the current study were previously investigated for the impact of *MYC* gene rearrangements and amplifications on clinical outcome.³⁴ Beyond these observations, we detected no significant association between an underlying HIV or EBV infection and any of the other reported clinicopathologic features.

Mutational landscape of PBL identified by WES

To characterize the mutational landscape in an extensive cohort of PBL cases, WES was successfully performed in 33 patient-derived tumor biopsy specimens. Matched germline DNA was not available for comparative analysis. We therefore applied the analytical framework outlined earlier to analyze WES data in the absence of paired germline DNA. After the primary identification of SNVs and indels in individual samples and subsequent filtering to correct for FFPE-derived artifacts and spurious mutations, we applied the MutSigCV algorithm and thereby initially identified 110 potentially significant candidate driver genes ($P < .05$) (supplemental Table 3).³³

At an overall low to intermediate tumor mutational burden (median, 3.273; range, 1.377-9.870 mutations/Mbase), we describe 8450 presumably harmful somatic mutations (cutoffs are discussed in the Materials and methods) involving 4505 genes after variant filtering. Of these, SNVs and indels represented 66% of the mutations (5219 SNVs for 4505 genes). Among them, missense mutations were the most frequent alterations (91%), followed by nonsense (4%) and splice-site (2%) mutations; indels posed 2% of somatic mutations (Figure 1). An oncoplot showing significantly mutated genes with alterations in ≥ 3 cases is provided in Figure 2A. Confirmatory Sanger sequencing or amplicon-based next-generation sequencing identified variants, previously detected by WES on all investigated occasions (supplemental Figure 1; supplemental Table 2).

Table 1. Baseline clinicopathologic characteristics in patients with PBL

Characteristic	PBL group (N = 33)
Age, median (range), y	61.5 (29-86)
Sex	
Female	10 (33.3%)
Male	23 (69.7%)
HIV status	
HIV-positive	14 (42.4%)
HIV-negative	19 (57.6%)
EBV status	
EBV-positive	20 (60.6%)
EBV-negative	13 (39.4%)
HIV- and EBV-positive	13 (39.4%)
R-IPI	
0	4 (12.1%)
1-2	11 (33.3%)
>2	18 (54.5%)
Stage (Ann Arbor)	
I	6 (18.2%)
II	8 (24.2%)
III	8 (24.2%)
IV	11 (33.3%)
B symptoms	
Yes	15 (45.5%)
No	18 (54.5%)
Chromosomal aberrations	
<i>MYC</i> overall	26 (78.8%)
<i>MYC</i> amplification	12 (36.4%)
<i>MYC</i> split (+/- amplification)	14 (42.4%)
CD30 by immunohistochemistry	
Positive	7 (21.2%)
Negative	26 (78.8%)
Extranodal sites	
0	4 (12.1%)
1-2	27 (81.8%)
>2	2 (6.1%)
ECOG performance status	
0-2	25 (75.8%)
>2	8 (24.2%)
Lactate dehydrogenase	
Normal	10 (30.3%)
Elevated	23 (69.7%)
CNS involvement at diagnosis	
Yes	0 (0.0%)
No	33 (100.0%)
Front-line therapy regimen	
CHOP-like	21 (63.6%)

Table 1. (continued)

Characteristic	PBL group (N = 33)
R-based	5 (15.2%)
Others*	8 (24.2%)
Refusal of treatment	2 (6.1%)
Front-line therapy response rates (n = 32)	
CR	10 (30.3%)
VGPR	4 (12.1%)
PR	5 (15.2%)
SD	5 (15.2%)
PD	7 (21.2%)

CHOP, cyclophosphamide, doxorubicin, vincristine, prednisolone; CNS, central nervous system; CR, complete response; ECOG, Eastern Cooperative Oncology Group; GCB, germinal-center B-cell like; PD, progressive disease; PR, partial remission; R, rituximab; R-IPI, revised-international prognostic index; SD, stable disease; VGPR, very good partial remission.

*Indicates other regimen (eg, bendamustine) or palliative cytoreductive treatment.

We observed no evidence of microsatellite instability (MSI)-associated hypermutations. This finding appears similar to results obtained in DLBCL (0.3%) but unlike and significantly lower than other aggressive lymphomas (eg, primary mediastinal B-cell lymphoma).³⁵

All PBL cases harbored mutations in genes implicated in oncogenesis according to our bioinformatics annotations. Upon integrated analysis of molecular and clinical data, we investigated the 17 most frequent genomic alterations identified in the current study (mutations with a frequency $\geq 15\%$ in the study cohort) for their impact on overall survival and progression-free survival. We identified a significant impact ($P < .05$) of *CSMD3* mutational status (supplemental Figure 3) on both overall survival and progression-free survival; upon correction for multiple testing according to the Bonferroni method, an additional effect regarding *ZNF462* mutation status failed to reach statistical significance, and neither gene was significantly mutated according to the subsequent MutSigCV analysis. Finally, within a Cox proportional hazards model, *CSMD3* mutation status failed to display prognostic independence from established IPI factors, which prompts the need for additional, preferably prospective clinical correlations of mutational data in PBL, in the search for molecular prognosticators.

Recurrent copy number alterations in PBL

We investigated our PBL cohort for somatic copy number variants (SCNVs) using the CONTROL-FREEC (v11.4)³⁶ algorithm in tumor-only mode followed by GISTIC2.0 (version 2.0.23)³⁷ analysis, while excluding chromosomes X and Y as well as common CNVs according to the Broad Institute's panel of normal (supplemental Materials). Subsequently, these observations were cross-referenced with the loci of known oncogenes, tumor suppressors, and elements of potential oncogenic pathways.³⁸ We found recurrent copy number gains in oncogenes (eg, *MEF2B* and *CSF1R*) previously implicated in lymphomagenesis and previously reported to be incidentally amplified in a variety of human cancers, as well as copy number losses in tumor suppressors such as *NPM1* (Figure 2B-C); common CNVs, as

defined by the aforementioned referenced panel of normal, were identified in the expected frequencies.³⁹⁻⁴¹ Additional information on minimal common regions alongside all affected genes within a given SCNv is provided in supplemental Table 4.

Significantly mutated candidate driver genes and mutational signatures

The list of significant candidate driver genes included several genes previously implicated in PBL pathogenesis, such as *STAT3* and *PRDM1* alongside mutational drivers in other B-cell lymphoproliferative disorders such as DLBCL and MM (*NRAS*, *BRAF*, and *TP53*) (Figure 2). We observed no subtype-specific predominance regarding *MYC*, EBV, and/or HIV status of statistical significance in our limited cohort. However, *STAT3* mutations were found to be significantly enriched in HIV-positive PBL (Fisher's exact test, $P = .0463$), whereas they were rare in HIV-negative patients, hinting at partially distinct molecular paths of pathogenesis. Similar trends bordering on statistical significance of enrichment were observed in *LNP1* for HIV-positive patients as well as *TP53*, *PRDM1*, and *IRS4* in HIV-negative patients. Significantly mutated, putative driver genes in PBL identified according to our pipeline exhibited a high degree of estimated pathogenicity with a median combined annotation dependent depletion (CADD) score of 25.45.

To further characterize the mutational profiles driving PBL pathogenesis, we performed mutational signature analysis as recently described in DLBCL and other aggressive B-cell lymphomas.^{35,42} Here, we identified the presence of 2 signatures, SBS2 (indicating impaired activity of the AID/APOBEC family of cytidine deaminases) and SBS6 (associated with defective DNA mismatch repair), with an overall predominance of the latter (supplemental Table 5). This phenomenon was significantly accentuated in EBV-negative PBL cases (supplemental Figure 4). Unlike our previous work on PBL and *MYC* in a more comprehensive, clinically annotated cohort, no correlation was observed between *MYC* status and overall survival ($P = .62$) or progression-free survival ($P = .91$), both when analyzing amplifications and translocations together or separately.³⁴ The number of mutations was similar between *MYC* wild-type and *MYC* altered samples (again, translocations and amplifications; Wilcoxon test, $P = .44$). *MYC* rearrangements were not correlated with EBV status (χ^2

test, $P = .1493$). Regarding genes and *MYC* aberrations, *OGFOD2*, *CASC4*, and *CTDSPL2* were found to be mutated more often in *MYC* wild-type samples (4 of 7 vs 2 of 25; $\text{fdr} = 0.0122$, $\text{fdr} = 0.0258$, $\text{fdr} = 0.0258$). Furthermore, the expected association between gene expression of *MYC* and cytogenetic *MYC* status failed to reach statistical significance within our limited sample set (Student t test, $P = .095$); splitting *MYC* status into amplification and translocation also revealed no statistically significant differences (Student t test, $P > .05$).

Deregulation of JAK/STAT, RTK-RAS, NF-κB, and phosphatidylinositol 3-kinase/AKT/mTOR signaling in PBL by distribution of mutations and network diffusion analysis

The predicted functional impact of all reported mutations is shown in supplemental Table 6 (CADD score) as well as protein domain annotations for SNVs and indels. Selected MutSigCV genes showed an average CADD score of 22.2 (standard deviation, 8.1; median, 23.4; minimum, 1.21; maximum, 37.0). We further investigated the putative functional relevance of mutations proven to be significant upon MutSigCV analysis. In this context, we matched our results from MutSigCV analysis with the pfam database to describe the impact on the protein function in more detail. Supplemental Table 6 summarizes the impact of SCNVs on post-transcriptomic changes. Here, we identified 4754 (64.8%) missense mutations, 231 (3.1%) nonsense mutations, 5 (0.1%) nonstop mutations, 121 splice site mutation, and 108 frame shift mutation harboring the potential to affect protein folding and the sequence of amino acids. Focusing on the most frequently mutated genes in our PBL cohort, by way of example, we were hereby able to confirm the relevance of *STAT3*, *HIST1H1E*, *PRDM1*, *TP53*, *NRAS*, and *PIM1* mutations on protein functions.

The *STAT3* protein plays a key role in cell growth and apoptosis.⁴³ Our results indicate relevant disruption of the *STAT3* binding domain (pfam 02864) and the SH2 domain in *STAT3* (cd10374).⁴⁴ The proto-oncogene *NRAS* encodes a protein with intrinsic guanine triphosphatase activity, which is involved in the regulation of cell growth and division. Thus, *NRAS* mutations promote carcinogenesis.⁴⁵ Here, we identified frequent *NRAS* missense mutations

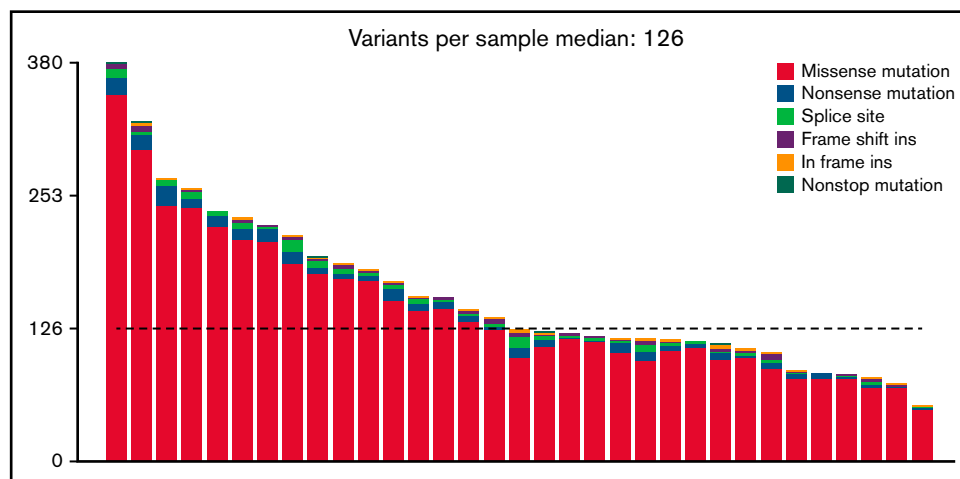


Figure 1. Number and type of mutations per sample.

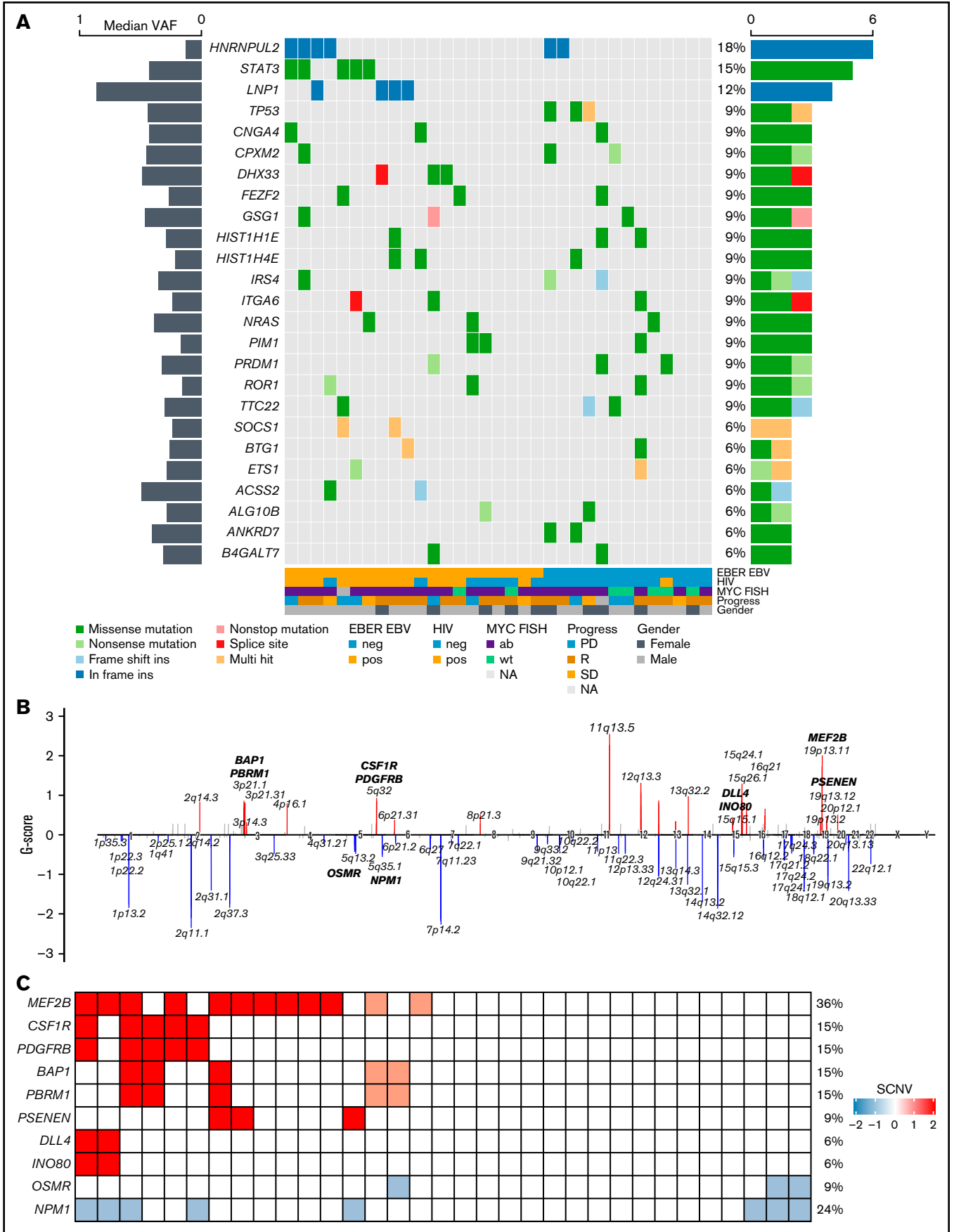


Figure 2.

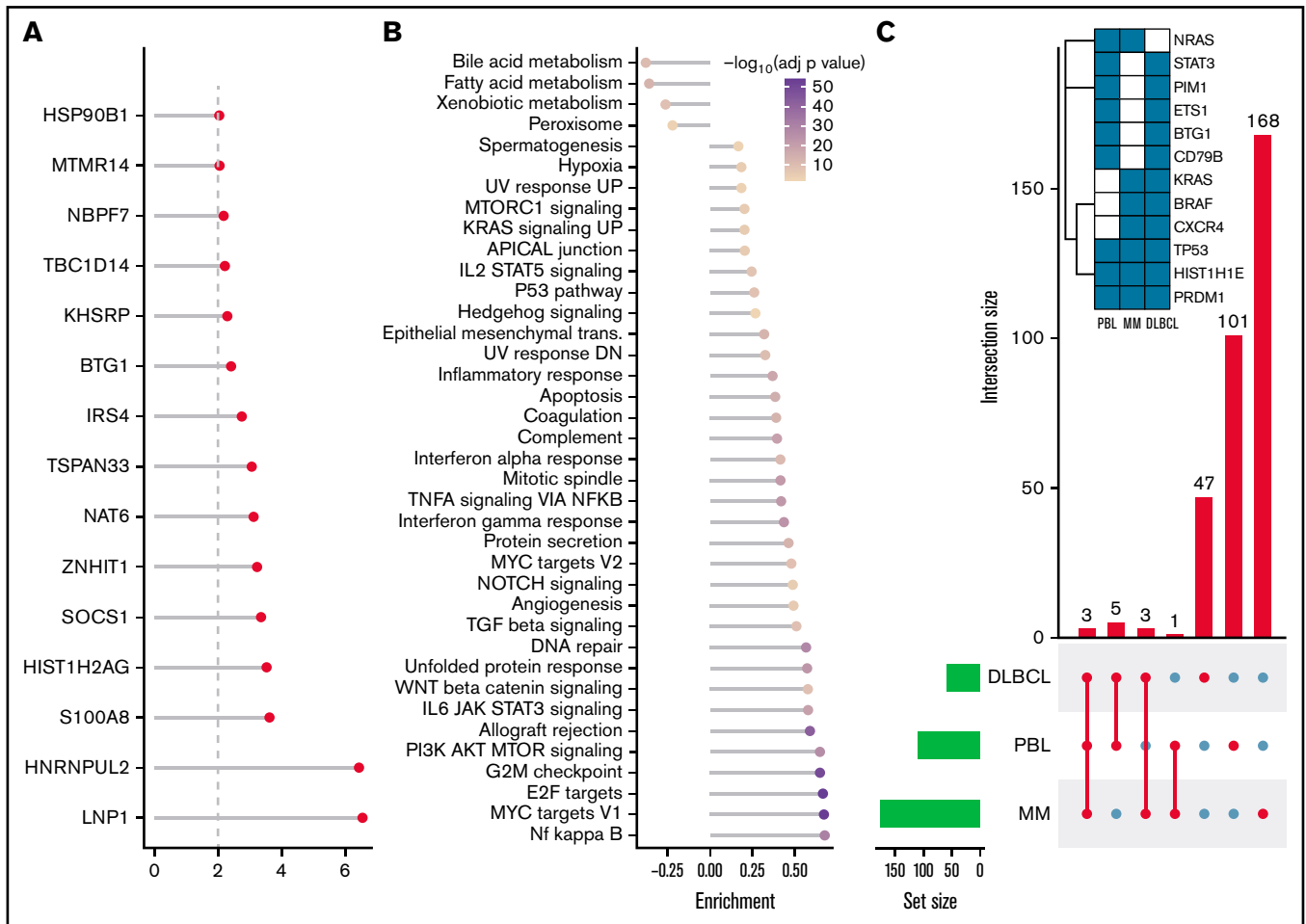


Figure 3. Mutational analysis for significantly mutated genes, network propagation and mutational overlap with related entities. Significance levels for all PBL MutSigCV genes ($P < .01$) (A) and HALLMARK gene sets and Nfkb pathway enrichment (B) for network diffusion analysis of significant MutSigCV genes (MutSigCV, $P < .01$). (C) UpSet plot showing the overlap of MutSigCV genes using PBL, DLBCL (1144 cases retrieved from 5 cohorts encompassing various subtypes of DLBCL via cBioPortal),^{42,58–61} and MM (205 cases from Lohr et al⁶² retrieved via cBioPortal) data; inlet shows the overlap between the 3 tumor entities.

affecting the H-N-K Ras-like domain (cd04138). Mutations in this domain are known to lead to the dysregulation of cell growth.⁴⁶ We also found frequent *TP53* missense mutations affecting the p53 DNA-binding domain (cd08367). The relevance of this mutation on carcinogenesis has been reported plurally.⁴⁷

Cumulatively, we detected genetic lesions, putatively resulting in gain of function in JAK/STAT signaling, in 64% of PBL cases. Expanding on previously reported *STAT3* mutations in PBL, we further identified recurrent mutations and/or SCNVs affecting *OSMR*, *PIM1*, and *ITGA4*. Although JAK/STAT signaling was the predominant target of somatic mutation in PBL, the overall distribution regarding the mutational signature within the pathway was heterogeneous (33 mutations in 24 genes in 21 patients) (supplemental Figures 5 and 6). Several genes, including *OSMR*, were previously

implicated in alternative signaling pathways, including the enhancement of Ras/Raf/MAPK signaling, rendering the gene a rather promiscuous target of pathway allocation. Most of these genetic alterations have been previously reported to constitute gain-of-function mutations resulting in constitutive JAK/STAT pathway activation in various types of T cells as well as primary or central nervous system/or central nervous system B-cell lymphoma and MM.^{35,48-50}

Merging mutational and SCNV data, we observed alterations that putatively resulted in gain of function in RTK-RAS signaling in 76% of patients (supplemental Figures 5, 6, and 8). Activating mutations affecting *ERBB3* and *NRAS* were most frequent in our series and have been previously described in a number of other tumor types, including MM, whereas an amplification of *PDGFRB* was found to be the overall dominating genetic lesion

Figure 2. Genomic and mutational landscape in plasmablastic lymphoma. (A) An oncoplot of genes mutated in at least 3 PBL samples after PBL MutSigCV analysis ($P < .05$) with allele frequencies added (left bar) and samples affected (right bar). Different types of mutations are color coded. In addition, covariates are shown below the plot for each sample. EBV-encoded RNA [EBER] EBV/HIV: neg, negative; pos, positive. MYC fluorescence in situ hybridization (FISH): ab, aberrated; NA, not available; wt, wild type. Progress: PD, progression disease; R, remission; SD, stable disease. (B) Location of SCNVs along the genome (red bars denote gains; blue bars denote losses; gene names refer to affected oncogenes and genes related to oncogenetic pathways). (C) Selected genes (oncogenes and oncogenetic pathway genes) affected by SCNVs.

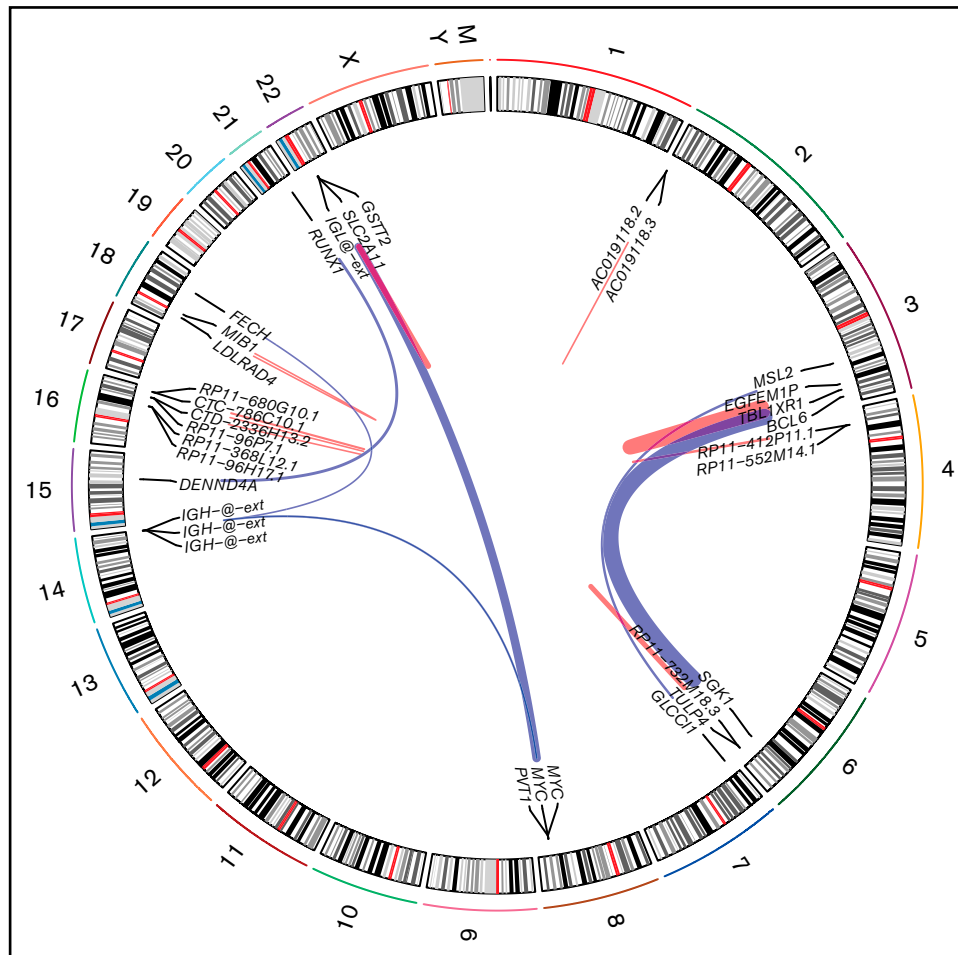


Figure 4. Fusion identified in PBL samples with respect to their genomic location. Red links indicate intrachromosomal fusions, and blue links indicate interchromosomal fusions.

in this pathway.^{51,52} A widespread mutational pattern was observed, affecting a substantial fraction of the pathway (27 of 85 genes [32%]), including *ALK*, *BRAF*, *CBLC*, and *ERBB2*, similar to other B-cell malignancies as well as other solid tumors.^{53,54} Intriguingly, we also identified 2 cases harboring a presumably harmful *NTRK3* mutation (CADD score 29.8). This finding is of relevance in light of the recent tissue-agnostic approval by the US Food and Drug Administration of the pan-TRK inhibitors larotrectinib and entrectinib alongside the recent identification of *NTRK* point mutations in hematologic malignancies, known to cause similar biological effects as *NTRK* fusions.⁵⁵

In accordance with previous studies, we found mutations and SCNVs affecting candidate driver genes of lymphomagenesis directly affecting NF- κ B signaling in 58% of cases (supplemental Figure 7A).²⁰ Although this was among the predominant pathways, identified through our network diffusion approach, mutations affecting the pathway were broadly distributed, with only 2 genes showing mutations in >1 patient (*NFKB2*, $n = 4$; *ATM*, $n = 2$), both previously implicated in the pathogenesis of lymphoplasmacytic lymphoma, intriguingly another type of terminally differentiated B-cell lymphoma.⁵⁶ Upon extension of the analysis onto the NF- κ B Pathway Interactome Map, thus explicitly broadening the scope beyond

the most compelling representatives of a given pathway, several cases were found to exhibit mutations and SCNVs in *CDX1*, *TWF2*, and *ABCB9* as well as the previously described genetic lesions in *PRDM1* (supplemental Figure 7B).^{18,19,57} Intriguingly and in keeping with recent observations by Ramis-Zaldivar et al,²³ we observed an accumulation of mutations in genes implicated in epigenetic and chromatin modifiers (11 of 33 cases).

After SNV and indel evaluation with MutSigCV, both a gene set enrichment as well as a network diffusion approach (Figures 3A-B) on genes with $P < .01$ were used to assess the impact of significant genetic events on neighboring genes. This evaluation highlighted the perturbing role of several recurrent alterations in interleukin-6 (IL-6)/JAK/STAT (*OSMR*, *STAT3*, *PIM1*, and *SOCS1*) as well as NF- κ B (*NFKB2* and *BTK*) and to a lesser extent phosphatidylinositol 3-kinase/AKT/mTOR (PI3K/AKT/mTOR) (*ACACA* and *HSP90B1*) signaling pathways (supplemental Figure 5) as the mutations detected within the latter pathway do not concede with the most compelling drivers of PI3K signaling. The former 2 of these pathways were previously described as being significantly disrupted in PBL.^{18,19} Furthermore, through a network diffusion approach, we found cellular metabolism to be significantly preserved, whereas NF- κ B, MYC target genes, cell cycle regulators, and immunologic

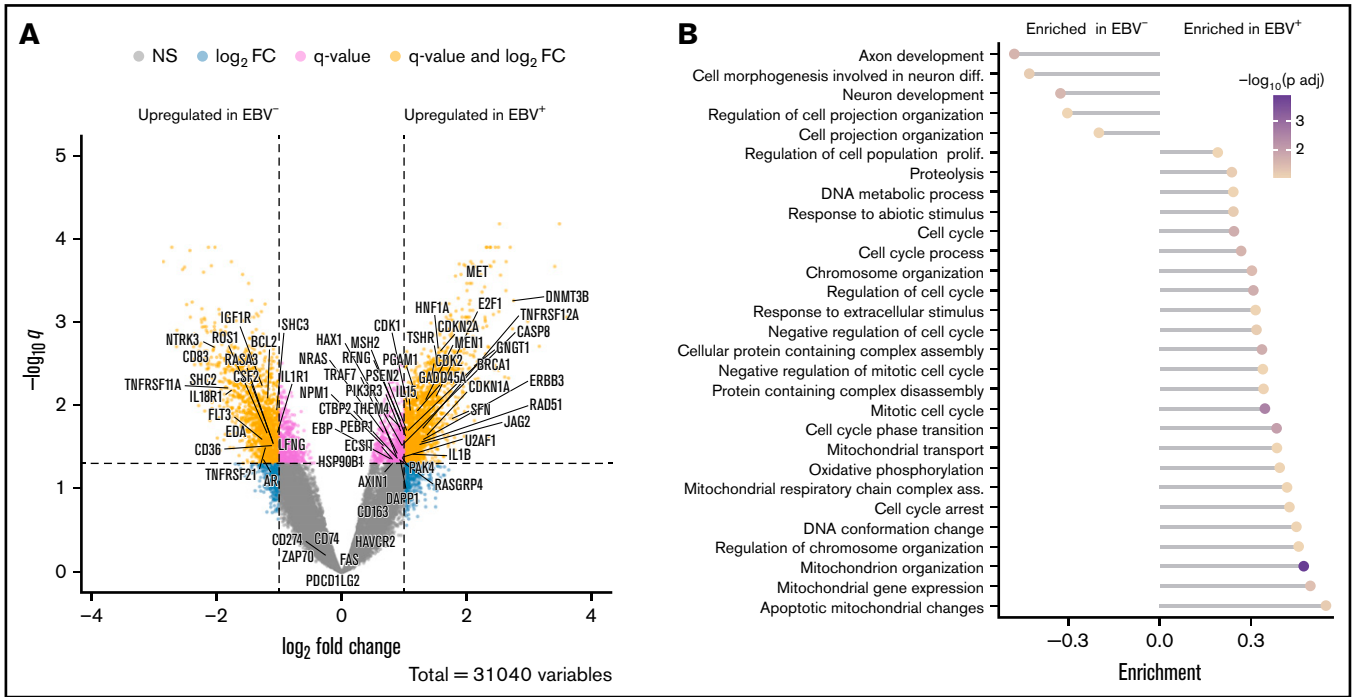


Figure 5. Distinctive expression pattern between EBV⁺ and EBV⁻ PBL. (A) Volcano plot shows log₂ fold-change (FC) and q-values of expression analysis (limma: EBV positive [EBV⁺] vs EBV negative [EBV⁻]; positive log₂ FC refer to upregulated genes in EBV⁺, negative values to upregulated genes in EBV⁻). Genes related to NF- κ B, PI3K AKT TOR, IL6 JAK STAT3, RTK_RAS, and NOTCH pathways are shown. Genes from Dunmire et al⁶⁶ are also included. (B) Gene set enrichment analysis against Gene Ontology gene sets on differentially expressed genes (log₂ fold-change > 1, p_{adj} < 0.01). Results were manually curated (full list is shown in supplemental Figure 9); enrichment scores (length of bar) and log scaled q-values (color-coded dots) are displayed.

processes (including allograft rejection and IL6/JAK/STAT3 signaling) were significantly mutated (Figure 3B).

Comparative analysis of PBL mutational landscape and related entities

To refine the genomic taxonomy of B-cell lymphomas and to investigate the mutational commonalities and differences between PBL and other related pathological entities chosen, due to similar features of (terminal) B-cell differentiation (DLBCL and MM), comparative analysis of candidate mutational drivers in PBL as described earlier and The Cancer Genome Atlas cohorts of DLBCL ($n = 1,151^{42,58-60}$) and MM ($n = 211^{62}$) was performed. Interestingly, we identified three overlapping candidate drivers common to all entities (*HIST1H1E*, *PRDM1*, and *TP53*) as well as 5 genes exclusively mutated in both PBL and DLBCL (*PIM1*, *BTG1*, *CD79B*, *ETS1*, and *STAT3*), whereas only *NRAS* was found to be an exclusively shared driver between PBL and MM. The vast majority of novel candidate genes reported in our study were, however, found to be PBL exclusive (Figure 3C).

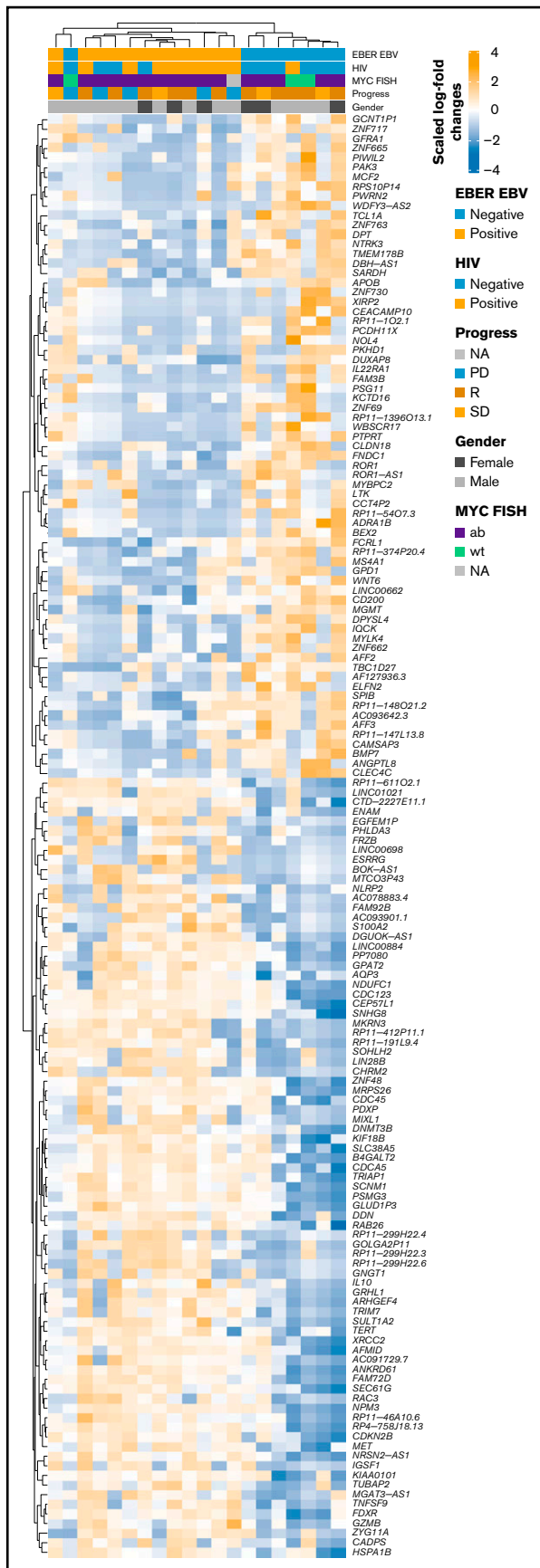
Fusion genes

Using STAR-FUSION (version 1.9.0)⁶³ running in de novo reconstruction mode and hg19 as reference genome, we were able to identify fusion genes from RNA-seq data (Figure 4). Ten samples were shown to carry 22 fusions (supplemental Table 7). In total, 16 unique fusion events were detected. We did not detect any novel, recurrent fusion events beyond the expectedly high frequency of *MYC* aberrations. To verify the fusions identified using STAR-FUSION,

we ran FUSIONCATCHER and FUSIONINSPECTOR to validate our findings. FUSIONSCATCHER verified 11 of 22 fusions; 3 of 4 *MYC* fusions were verified by FUSIONCATCHER. FUSIONINSPECTOR verified 18 of 22 fusions and 3 of 4 *MYC* fusions. All of the latter were subsequently validated by fluorescence in situ hybridization. In total, 19 of 22 fusions were verified.

RNA-seq

Using STAR ALIGNER (version 2.7.2b)⁶⁴ and MIXNORM,⁶⁵ we performed a gene set enrichment analysis employing GSVA and HALLMARK gene sets (supplemental Figure 8). The RNA-seq expression profiles of all 20 patients, whose data sets met our predefined standards for quality and quantity (>10 Gb/sample at Q30% > 90%) for subsequent analysis, were subjected to hierarchical clustering of both gene sets and cases. We hereby uncovered 2 dominant clusters of gene sets alongside 4 patient clusters. Induced pathways were found to reflect the genomic profile with regard to the heterogeneous distribution of alterations accentuated in IL-6/JAK/STAT signaling, inflammatory response genes, NF- κ B activity, and *MYC* signaling. Integrated analysis of WES and RNA-seq regarding mutational data in an exemplary subset of genes (*STAT3*, *TP53*, *DHX33*, *NRAS*, and *HIST1H4E*) yielded a concordance rate of 43.8%, which was, however, attributable to low coverage/expression within the RNA-seq data. Furthermore, the aspect of allele-specific expression needs to be considered. Successful expression analysis of genes, affected by SCNAs according to our WES data, was hampered by the low number of cases with availability of both types of data.



Comparative gene expression profiling between EBV-positive and EBV-negative PBL

RNA-seq profiles were further investigated for genes showing significant differential expression between EBV-positive PBL and EBV-negative PBL identified through the linear model fit with P value correction ($p_{\text{adj}} < 0.01$ and \log_2 fold-change > 1). Comparative analysis yielded 1099 genes with significant differential expression between the 2 groups (Figure 5A); a tabular delineation of gene expression profiles in our cohort is provided in supplemental Table 8. Of these genes, 623 were upregulated in EBV-positive PBL. Unsupervised hierarchical clustering of differentially expressed genes (\log_2 fold-change > 2) is represented on the heatmap shown in Figure 6. Subsequently, functional annotation of this gene expression profile was performed by gene set enrichment analysis, using the Gene Ontology "Biological Process." Hereby, we uncovered a predominant induction of metabolic (eg, mitochondrial) processes in EBV-positive PBL (Figure 5B; supplemental Figure 10). Intriguingly, class I major histocompatibility complex-mediated antigen processing and presentation were also significantly affected by the presence of an underlying EBV infection, as was cell cycle regulation. These findings thus partially confirm previous results, whereas only *IL15* expression was induced in EBV-positive PBL, unlike the other genes within the IEGS33.^{16,67} Differential expression was also prominent in several known EBV targets, including *CD83*, *RAD51*, *PGAM1*, *CDK2*, and *DNMT3B*, a known molecular link between B-cell state and the EBV latency program.⁶⁸

Upon restriction of our comparative gene expression analysis on genes with established functions as tumor suppressors or oncogenes, we identified several significantly upregulated genes in EBV-positive PBL (including *MET*, *BCL2*, and *NRAS*). EBV-negative PBL was significantly enriched for downregulated tumor suppressors (including *FLT3*, *MEN1*, *MSH2*, and *NPM1*) (Figure 5A). Intriguingly, we identified significant EBV-dependent deregulations within all major oncogenic pathways previously found to be affected by oncogenic mutations. Although most pathways displayed a homogeneous distribution of deregulated genes between both subgroups, key elements of PI3K/AKT/mTOR (including *CDK1*, *CDK2*, and *HSP90B1*) were almost exclusively upregulated in EBV-positive PBL (Figure 5A; supplemental Figure 9; supplemental Table 9). This is in keeping with the known induction of PI3K signaling in EBV-infected cells as a potent collaborator of *MYC*.^{69,70} Specifically, in PBL, Mine et al⁷¹ described an IL-6 dependency of PBL-1, an EBV-driven PBL cell line, for proliferation and survival. Furthermore, mTOR inhibitors effectively induced cell death in this model, suggesting its potential therapeutic role in a subset of patients with PBL.

Further analysis on a case-by-case level, screening signaling pathways (predominantly affected by genetic lesions in PBL) revealed a highly induced expression of *NTRK3* in EBV-negative PBL. We hereby establish a significant, EBV-dependent transcriptional

Figure 6. Heatmap of scaled \log_2 fold-change of significantly expressed genes ($P < .01$, \log_2 fold-change > 2) EBV-positive vs EBV-negative samples. Clustering was performed by using Spearman correlations, and covariates are shown within the figure. ab, aberrated; FISH, fluorescence in situ hybridization; NA, not available; PD, progression disease; R, remission; SD, stable disease; wt, wild type.

divergence in contrast to the widely similar mutational landscape shared by both PBL subtypes.

A comprehensive tabular summary of the entire processed data generated in the current study, using the pipeline outlined here, is provided as supplemental Table 9. Lollipop plots are provided for selected genes with special emphasis on known cancer-related genes and MutSigCV top genes within our cohort in supplemental Figure 11.

Discussion

In this comprehensive genomic and transcriptomic study of PBL, we report on WES and RNA-seq data obtained in a cohort of PBL tumors. To the best of our knowledge, we present the largest cohort and most comprehensive data set for this entity, as previous reports on PBL were limited by lower case numbers, targeted sequencing approaches, and lack of transcriptional data, whereas we systematically defined recurrent mutations, CNAs, and gene fusions alongside predominant mutation and transcriptional signatures. The current study made 2 novel and essential observations.

First, a significant enrichment for mutations affecting both NF- κ B as well as IL-6/JAK/STAT signaling pathways was identified, as could be expected from previous studies reporting on targeted sequencing approaches.^{18,19} However, adding to previous and complementing most recent observations (the former primarily derived from HIV-positive patients), we unveil several novel and recurrent, activating genetic alterations involving the classic RTK-RAS pathway in the vast majority of cases (23 of 33 [69.7%]) as a predominant driver of PBL lymphomagenesis, regardless of age and/or HIV status.²³ These observations are, however, in line with recent observations in a targeted sequencing study on 11 patients with PBL-type posttransplant lymphoproliferative disorders.²² Intriguingly, compared with previous studies in highly selected HIV-positive PBL cohorts, we find that *STAT3* mutations, considered to be essential drivers in PBL lymphomagenesis, as well as *LNP1* mutations, occur less frequently in HIV-negative cases, whereas *TP53*, *PRDM1*, and *IRS4* mutations and RTK-RAS perturbations seem to be more essential in this subgroup.^{20,21} Upon comparison with the most recent amplicon-panel derived study by Ramis-Zaldivar et al,²³ we observed vast similarities with the current study, embedding it firmly within the landscape of published data on the molecular pathogenesis of PBL. Similarities included trends toward more mutations in *STAT3* in EBV/HIV-positive cases and enrichment in *TP53* mutations in EBV/HIV-negative cases, along with enrichment in mutations affecting epigenetic and chromatin modifiers. However, by contrast, these observations failed to reach statistical significance in terms of affected pathways, which may be attributable to the more concise methodologic approach of Ramis-Zaldivar et al as well as the small sample sizes and strong heterogeneity with regard to the composition of both study groups. By using MutSigCV, we describe a set of 110 candidate driver genes ($P < .05$; 15 genes, $P < .01$) (Figure 3) with substantial evidence of somatic and functionally impairing mutation status in the vast majority of studied tumors.³³ Additional q-value analysis would have profusely limited this set of candidate driver genes in our relatively small study cohort, precluding known drivers of oncogenesis in aggressive lymphomas (eg, *TP53*, *PRDM1*, and others) from subsequent analysis. Using a widely established cutoff regarding the q-value (eg, < 0.1) would have resulted in a reduction in size of the candidate driver gene sets ($n = 2$ genes [*LNP1* and *HNRNPUL2*]) brought on by the limitations in

cohort size. Mutations in *HNRNPUL2* have not been implicated in lymphoma pathogenesis but in glioblastoma and colorectal carcinoma.^{72,73} Although there was a high grade of confidence in our filtering pipeline, this suspicious finding requires functional validation in further studies. To retain pathobiologically meaningful data, q-value filtering was therefore omitted in our exploratory study as our data set disallows further screening for significantly mutated genes beyond the provided data (q-values are reported in supplemental Table 3). Subsequent studies on larger cohorts are required to further refine the definitive gene set of significant drivers in PBL and its subtypes. Our observations regarding *NTRK3* in EBV-negative PBL, in concert with the previously identified oncogenic *NTRK3* mutations, hint at a specific, therapeutically targetable mechanism in a clinically relevant subset of PBL cases.

Of potential therapeutic relevance, we found 2 mutations affecting *BTK* at diagnosis, potentially disrupting B-cell receptor signaling, resistant to specific inhibition (eg, ibrutinib). Moreover, the identification of oncogenic mutations in *NTRK3* alongside its pronounced expression in EBV-positive PBL holds potential therapeutic promise for a subset of patients with PBL given the tissue-agnostic approval by the US Food and Drug Administration of the pan-TRK inhibitors larotrectinib and entrectinib.⁵⁵ There remains, however, a lack of functional validations of potential novel therapeutic agents in preclinical models of PBL beyond preliminary studies of bortezomib, which has already found its way into clinical PBL practice.⁷⁴ This poses a promising subject for successive studies. Upon comparison of mutational patterns and frequencies in PBL and related B-cell malignancies (DLBCL and MM), we provide evidence of pathogenetic independence of PBL from DLBCL and MM based on its genomic landscape, thereby supplementing previous studies on differential gene expression signatures.^{15,16} Our observations regarding MSI-associated hypermutations are clinically relevant as MSI has been associated with increased sensitivity to programmed cell death protein 1 blockade across a variety of malignancies, and checkpoint inhibitor therapies are currently being explored as salvage treatment options in a subset of relapsed/refractory aggressive B-cell lymphomas.⁷⁵⁻⁷⁷ Of note, copy number variation analysis derived from WES data revealed aberration affecting several genes, previously implicated in lymphomagenesis, such as *CSF1R* and *MEF2B*.^{39,40} In our case series, however, we were unable to reproduce the findings by Liu et al,²¹ who described recurrent copy number gains of *CD44*. This observation may well be attributed to the pathogenetically more heterogeneous study group in the current article.

Second, our RNA-seq data mirrored the genomic profile with regard to the heterogeneous distribution of altered pathways, and we further identified several transcriptional clusters, enriched for IL-6/JAK/STAT signaling, inflammatory response genes, NF- κ B activity, and MYC signaling. Of special interest, however, comparative analysis of gene expression between EBV-positive PBL and EBV-negative PBL identified a marked impact on class I major histocompatibility complex-mediated antigen processing and presentation, as well as cell cycle regulation, by the presence of an underlying EBV infection. Moreover, key elements of PI3K/AKT/mTOR signaling (including *CDK1*, *CDK2*, and *HSP90B1*) were almost exclusively upregulated in EBV-positive PBL, potentially advising more specific and phenotypically tailored treatment approaches to be explored in PBL, taking into account the tumor's respective EBV status. The proposed IL-6/JAK/STAT dependency reflected by our WES and

RNA-seq data appears particularly to be in keeping with a viral IL-6–driven xenograft model of PBL in which viral IL-6 collaborated with *MYC* in the formation of spontaneous plasmablastic tumors in a mouse model.⁷⁸ Intriguingly, RNA-seq revealed downregulation of tumor suppressors (including *FLT3*, *MEN1*, *MSH2*, and *NPM1*) to be a recurrent feature significantly accentuated in EBV-negative PBL. In contrast to the widely similar mutational landscapes shared by both PBL subtypes, our observations uncovered a fundamental impact of EBV status on PBL gene expression. Our characterization of a distinct mutational and transcriptomic landscape in PBL, distinct from both DLBCL and MM, with pronounced mutational impairment of IL-6/JAK/STAT, RTK-RAS, and NF- κ B signaling alongside transcriptional deregulation of the PI3K/AKT/mTOR pathway predominantly encountered in EBV-positive PBL, substantiate the pathobiological independence of PBL in the spectrum of terminally differentiated B-cell malignancies and thereby refine the taxonomy for aggressive lymphomas.

Acknowledgments

The authors thank Tanja Oeltermann and Biggi Steinfeldt for their skilled technical assistance. A.K. and H.B. acknowledge

computational support from the OMICS compute cluster at the University of Lübeck.

H.B. acknowledges funding by the Deutsche Forschungsgemeinschaft (German Research Foundation) under Germany's Excellence Strategy (EXC 22167-390884018).

Authorship

Contribution: N.G. and A.C.F. were responsible for study concept; N.G., A.K., H.M.W., N.H., V.B., H.M., H.-W.B., and S.S. performed data collection; A.K., N.G., H.M.W., H.-W.B., V.B., H.B., and N.v.B. contributed to data analysis and creation of figures and tables; N.G. drafted the initial version of the manuscript; and all authors contributed to critical revision and approval of the final version.

Conflict-of-interest disclosure: The authors declare no competing financial interests.

ORCID profiles: A.K., 0000-0003-0692-2105; N.v.B., 0000-0001-9593-8947; H.B., 0000-0003-4763-4521.

Correspondence: Niklas Gebauer, Department of Hematology and Oncology, UKSH Campus Lübeck, Ratzenburger Allee 160, 23538 Lübeck, Germany; e-mail: Niklas.Gebauer@uksh.de.

References

1. Swerdlow SH, Campo E, Pileri SA, et al. The 2016 revision of the World Health Organization classification of lymphoid neoplasms. *Blood*. 2016;127(20):2375-2390.
2. Loghavi S, Alayed K, Aladily TN, et al. Stage, age, and EBV status impact outcomes of plasmablastic lymphoma patients: a clinicopathologic analysis of 61 patients. *J Hematol Oncol*. 2015;8(1):65.
3. Delecluse HJ, Anagnostopoulos I, Dallenbach F, et al. Plasmablastic lymphomas of the oral cavity: a new entity associated with the human immunodeficiency virus infection. *Blood*. 1997;89(4):1413-1420.
4. Vega F, Chang CC, Medeiros LJ, et al. Plasmablastic lymphomas and plasmablastic plasma cell myelomas have nearly identical immunophenotypic profiles [published correction appears in *Mod Pathol*. 2005;18(6):873]. *Mod Pathol*. 2005;18(6):806-815.
5. Sanguedolce F, Zanelli M, Zizzo M, et al. Clinical, pathological and molecular features of plasmablastic lymphoma arising in the gastrointestinal tract: a review and reappraisal [published correction appears in *Pathol Res Pract*. 2020;216(11):153132]. *Pathol Res Pract*. 2020;216(6):152973.
6. Arora N, Eule C, Gupta A, Li HC, Sadeghi N. Clinicopathologic features, management, and outcomes of plasmablastic lymphoma: a 10-year experience. *Am J Hematol*. 2019;94(5):E127-E129.
7. Tchernonog E, Faurie P, Coppo P, et al. Clinical characteristics and prognostic factors of plasmablastic lymphoma patients: analysis of 135 patients from the LYSA group. *Ann Oncol*. 2017;28(4):843-848.
8. Teruya-Feldstein J, Chiao E, Filippa DA, et al. CD20-negative large-cell lymphoma with plasmablastic features: a clinically heterogenous spectrum in both HIV-positive and -negative patients. *Ann Oncol*. 2004;15(11):1673-1679.
9. Castillo JJ, Bibas M, Miranda RN. The biology and treatment of plasmablastic lymphoma. *Blood*. 2015;125(15):2323-2330.
10. Castillo JJ, Guerrero-Garcia T, Baldini F, et al. Bortezomib plus EPOCH is effective as frontline treatment in patients with plasmablastic lymphoma. *Br J Haematol*. 2019;184(4):679-682.
11. Miao L, Guo N, Feng Y, et al. High incidence of *MYC* rearrangement in human immunodeficiency virus-positive plasmablastic lymphoma. *Histopathology*. 2020;76(2):201-211.
12. Valera A, Balagué O, Colomo L, et al. IG/*MYC* rearrangements are the main cytogenetic alteration in plasmablastic lymphomas. *Am J Surg Pathol*. 2010;34(11):1686-1694.
13. Taddesse-Heath L, Meloni-Ehrig A, Scheerle J, Kelly JC, Jaffe ES. Plasmablastic lymphoma with *MYC* translocation: evidence for a common pathway in the generation of plasmablastic features. *Mod Pathol*. 2010;23(7):991-999.
14. Chang CC, Zhou X, Taylor JJ, et al. Genomic profiling of plasmablastic lymphoma using array comparative genomic hybridization (aCGH): revealing significant overlapping genomic lesions with diffuse large B-cell lymphoma. *J Hematol Oncol*. 2009;2(1):47.
15. Chapman J, Gentles AJ, Sujoy V, et al. Gene expression analysis of plasmablastic lymphoma identifies downregulation of B-cell receptor signaling and additional unique transcriptional programs. *Leukemia*. 2015;29(11):2270-2273.

16. Gravelle P, Péricart S, Tosolini M, et al. EBV infection determines the immune hallmarks of plasmablastic lymphoma. *OncImmunology*. 2018; 7(10):e1486950.
17. Laurent C, Fabiani B, Do C, et al. Immune-checkpoint expression in Epstein-Barr virus positive and negative plasmablastic lymphoma: a clinical and pathological study in 82 patients. *Haematologica*. 2016;101(8):976-984.
18. Montes-Moreno S, Martínez-Magunacelaya N, Zecchini-Barrese T, et al. Plasmablastic lymphoma phenotype is determined by genetic alterations in MYC and PRDM1. *Mod Pathol*. 2017;30(1):85-94.
19. Garcia-Reyero J, Martínez Magunacelaya N, Gonzalez de Villambrosia S, et al. Genetic lesions in MYC and STAT3 drive oncogenic transcription factor overexpression in plasmablastic lymphoma. *Haematologica*. 2021;106(4):1120-1128.
20. Cinar M, Rong HR, Chineke I, et al. Genetic analysis of plasmablastic lymphomas in HIV (+) patients reveals novel driver regulators of the noncanonical NF- κ B pathway. Paper presented at 60th American Society of Hematology (ASH) Annual Meeting. 29 November 2018. Vol. 132 (suppl 1):1565. San Diego, CA: American Society of Hematology.
21. Liu Z, Filip I, Gomez K, et al. Genomic characterization of HIV-associated plasmablastic lymphoma identifies pervasive mutations in the JAK-STAT pathway. *Blood Cancer Discov*. 2020;1(1):112-125.
22. Leeman-Neill RJ, Soderquist CR, Montanari F, et al. Phenogenomic heterogeneity of post-transplant plasmablastic lymphomas [published online ahead of print 3 December 2020]. *Haematologica*. doi: 10.3324/haematol.2020.267294.
23. Ramis-Zaldivar JE, Gonzalez-Farre B, Nicolae A, et al. MAPK and JAK-STAT pathways dysregulation in plasmablastic lymphoma. *Haematologica*. 2021;106(10):2682-2693.
24. Gebauer N, Künstner A, Ketzler J, et al. Genomic insights into the pathogenesis of Epstein-Barr virus-associated diffuse large B-cell lymphoma by whole-genome and targeted amplicon sequencing. *Blood Cancer J*. 2021;11(5):102.
25. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884-i890.
26. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Available at: <https://arxiv.org/abs/1303.3997v2>. Accessed 30 November 2020
27. Institute B. Picard Toolkit. GitHub Repository; 2019. Available at: <https://broadinstitute.github.io/picard/>. Accessed 30 November 2020
28. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297-1303.
29. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol*. 2013;31(3):213-219.
30. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.
31. Danecek P, Auton A, Abecasis G, et al; 1000 Genomes Project Analysis Group. The variant call format and VCFtools. *Bioinformatics*. 2011; 27(15):2156-2158.
32. Shyr C, Tarailo-Graovac M, Gottlieb M, Lee JJ, van Karnebeek C, Wasserman WW. FLAGS, frequently mutated genes in public exomes [published correction appears in *BMC Med Genomics*. 2017;10(1):69]. *BMC Med Genomics*. 2014;7(1):64.
33. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013; 499(7457):214-218.
34. Witte HM, Hertel N, Merz H, et al. Clinicopathological characteristics and MYC status determine treatment outcome in plasmablastic lymphoma: a multi-center study of 76 consecutive patients. *Blood Cancer J*. 2020;10(5):63.
35. Chapuy B, Stewart C, Dunford AJ, et al. Genomic analyses of PMBL reveal new drivers and mechanisms of sensitivity to PD-1 blockade. *Blood*. 2019;134(26):2369-2382.
36. Boeva V, Popova T, Bleakley K, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. *Bioinformatics*. 2012;28(3):423-425.
37. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41.
38. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. *Science*. 2013;339(6127):1546-1558.
39. Lamprecht B, Walter K, Kreher S, et al. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat Med*. 2010;16(5):571-579, 1p following 579.
40. Pon JR, Wong J, Saberi S, et al. MEF2B mutations in non-Hodgkin lymphoma dysregulate cell migration by decreasing MEF2B target gene activation. *Nat Commun*. 2015;6(1):7953.
41. AACR Project GENIE Consortium. AACR Project GENIE: powering precision medicine through an international consortium. *Cancer Discov*. 2017; 7(8):818-831.
42. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes [published corrections appear in *Nat Med*. 2018;24(8):1292 and *Nat Med*. 2018;24(8):1290-1291]. *Nat Med*. 2018;24(5):679-690.
43. Wingelhofer B, Neubauer HA, Valent P, et al. Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. *Leukemia*. 2018;32(8):1713-1726.

44. de Araujo ED, Orlova A, Neubauer HA, et al. Structural implications of STAT3 and STAT5 SH2 domain mutations. *Cancers (Basel)*. 2019;11(11):E1757.
45. Qian L, Chen K, Wang C, Chen Z, Meng Z, Wang P. Targeting NRAS-mutant cancers with the selective STK19 kinase inhibitor chelidonine. *Clin Cancer Res*. 2020;26(13):3408-3419.
46. Quan Y, Liu G, Yu W, et al. Expression, purification, and characterization of Ras protein (BmRas1) from *Bombyx mori*. *Comp Funct Genomics*. 2012;2012:747539.
47. Kato S, Han SY, Liu W, et al. Understanding the function-structure and function-mutation relationships of p53 tumor suppressor protein by high-resolution missense mutation analysis. *Proc Natl Acad Sci U S A*. 2003;100(14):8424-8429.
48. Nicolae A, Xi L, Pham TH, et al. Mutations in the JAK/STAT and RAS signaling pathways are common in intestinal T-cell lymphomas. *Leukemia*. 2016;30(11):2245-2247.
49. Vater I, Montesinos-Rongen M, Schlesner M, et al. The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing. *Leukemia*. 2015;29(3):677-685.
50. Liu Z, Liu H, He J, Lin P, Tong Q, Yang J. Myeloma cells shift osteoblastogenesis to adipogenesis by inhibiting the ubiquitin ligase MURF1 in mesenchymal stem cells. *Sci Signal*. 2020;13(633):eaay8203.
51. Leich E, Weißbach S, Klein HU, et al. Multiple myeloma is affected by multiple and heterogeneous somatic mutations in adhesion- and receptor tyrosine kinase signaling molecules. *Blood Cancer J*. 2013;3(2):e102.
52. Rossi A, Voigtlaender M, Janjetovic S, et al. Mutational landscape reflects the biological continuum of plasma cell dyscrasias. *Blood Cancer J*. 2017;7(2):e537.
53. Pek M, Yatim SMJM, Chen Y, et al. Oncogenic KRAS-associated gene signature defines co-targeting of CDK4/6 and MEK as a viable therapeutic strategy in colorectal cancer. *Oncogene*. 2017;36(35):4975-4986.
54. Gebauer N, Biersack H, Czerwinska AC, et al. Favorable prognostic impact of RAS mutation status in multiple myeloma treated with high-dose melphalan and autologous stem cell support in the era of novel agents: a single center perspective. *Leuk Lymphoma*. 2016;57(1):226-229.
55. Joshi SK, Qian K, Bisson WH, et al. Discovery and characterization of targetable NTRK point mutations in hematologic neoplasms. *Blood*. 2020;135(24):2159-2170.
56. Hunter ZR, Xu L, Tsakmaklis N, et al. Insights into the genomic landscape of MYD88 wild-type Waldenström macroglobulinemia. *Blood Adv*. 2018;2(21):2937-2946.
57. Tieri P, Termanini A, Bellavista E, Salvioli S, Capri M, Franceschi C. Charting the NF-κB pathway interactome map. *PLoS One*. 2012;7(3):e32678.
58. Lohr JG, Stojanov P, Lawrence MS, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci U S A*. 2012;109(10):3879-3884.
59. Morin RD, Mungall K, Pleasance E, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood*. 2013;122(7):1256-1265.
60. Reddy A, Zhang J, Davis NS, et al. Genetic and functional drivers of diffuse large B cell lymphoma. *Cell*. 2017;171(2):481-494.e15.
61. Hoadley KA, Yau C, Hinoue T, et al; Cancer Genome Atlas Network. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. *Cell*. 2018;173(2):291-304.e6.
62. Lohr JG, Stojanov P, Carter SL, et al; Multiple Myeloma Research Consortium. Widespread genetic heterogeneity in multiple myeloma: implications for targeted therapy. *Cancer Cell*. 2014;25(1):91-101.
63. Haas BJ, Dobin A, Li B, Stransky N, Pochet N, Regev A. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol*. 2019;20(1):213.
64. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15-21.
65. Yin S, Wang X, Jia G, Xie Y. MIXnorm: normalizing RNA-seq data from formalin-fixed paraffin-embedded samples. *Bioinformatics*. 2020;36(11):3401-3408.
66. Dunmire SK, Odumade OA, Porter JL, et al. Primary EBV infection induces an expression profile distinct from other viruses but similar to hemophagocytic syndromes. *PLoS One*. 2014;9(1):e85422.
67. Kumar P, Murakami M, Kaul R, Saha A, Cai Q, Robertson ES. Deregulation of the cell cycle machinery by Epstein-Barr virus nuclear antigen 3C. *Future Virol*. 2009;4(1):79-91.
68. Guo R, Zhang Y, Teng M, et al. DNA methylation enzymes and PRC1 restrict B-cell Epstein-Barr virus oncoprotein expression. *Nat Microbiol*. 2020;5(8):1051-1063.
69. Barel M, Balbo M, Le Romancer M, Frade R. Activation of Epstein-Barr virus/C3d receptor (gp140, CR2, CD21) on human cell surface triggers pp60src and Akt-GSK3 activities upstream and downstream to PI 3-kinase, respectively. *Eur J Immunol*. 2003;33(9):2557-2566.
70. Sander S, Calado DP, Srinivasan L, et al. Synergy between PI3K signaling and MYC in Burkitt lymphomagenesis. *Cancer Cell*. 2012;22(2):167-179.
71. Mine S, Hishima T, Suganuma A, et al. Interleukin-6-dependent growth in a newly established plasmablastic lymphoma cell line and its therapeutic targets. *Sci Rep*. 2017;7(1):10188.
72. Jiang H, Wang Y, Ai M, et al. Long noncoding RNA CRNDE stabilized by hnRNPUL2 accelerates cell proliferation and migration in colorectal carcinoma via activating Ras/MAPK signaling pathways. *Cell Death Dis*. 2017;8(6):e2862.

73. Weller M, Weber RG, Willscher E, et al. Molecular classification of diffuse cerebral WHO grade II/III gliomas using genome- and transcriptome-wide profiling improves stratification of prognostically distinct patient groups. *Acta Neuropathol.* 2015;129(5):679-693.
74. Matsuki E, Miyakawa Y, Asakawa S, et al. Identification of loss of p16 expression and upregulation of MDR-1 as genetic events resulting from two novel chromosomal translocations found in a plasmablastic lymphoma of the uterus. *Clin Cancer Res.* 2011;17(8):2101-2109.
75. Overman MJ, McDermott R, Leach JL, et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. *Lancet Oncol.* 2017;18(9):1182-1191.
76. Salem ME, Bodor JN, Puccini A, et al. Relationship between MLH1, PMS2, MSH2 and MSH6 gene-specific alterations and tumor mutational burden in 1057 microsatellite instability-high solid tumors. *Int J Cancer.* 2020;147(10):2948-2956.
77. Kline J, Godfrey J, Ansell SM. The immune landscape and response to immune checkpoint blockade therapy in lymphoma. *Blood.* 2020;135(8):523-533.
78. Rosean TR, Holman CJ, Tompkins VS, et al. KSHV-encoded vIL-6 collaborates with deregulated c-Myc to drive plasmablastic neoplasms in mice. *Blood Cancer J.* 2016;6(2):e398.