Title: Mutational landscape of high-grade B-cell lymphoma with *MYC*-, *BCL2* and/or *BCL6* rearrangements characterized by whole-exome sequencing

3

Authors: Axel Küstner^{1,2,3*}, Hanno M. Witte^{4,5*}, Jörg Riedl^{4,6*}, Veronica Bernard⁶,
Stephanie Stölting⁶, Hartmut Merz⁶, Vito Olschewski⁴, Wolfgang Peter^{7,8}, Julius Ketzer⁹,
Yannik Busch⁷, Peter Trojok⁷, Nikolas von Bubnoff^{3,4}, Hauke Busch^{1,2,3**}, Alfred C. Feller^{6**},
Niklas Gebauer^{3,4**}

8

Affiliations:

¹ Medical Systems Biology Group, University of Lübeck, Ratzeburger Allee 160, 23538
Lübeck, Germany

² Institute for Cardiogenetics, University of Lübeck, Ratzeburger Allee 160, 23538 Lübeck,
Germany

³ University Cancer Center Schleswig-Holstein, University Hospital of Schleswig- Holstein,
Campus Lübeck, 23538 Lübeck, Germany

⁴ Department of Hematology and Oncology, University Hospital of Schleswig-Holstein,
Campus Lübeck, Ratzeburger Allee 160, 23538 Lübeck, Germany

⁵ Department of Hematology and Oncology, Federal Armed Forces Hospital Ulm, Oberer
Eselsberg 40, 89081 Ulm

⁶ Hämatopathologie Lübeck, Reference Centre for Lymph Node Pathology and
Hematopathology, Lübeck, Germany

⁷ HLA Typing Laboratory of the Stefan-Morsch-Foundation, 557565 Birkenfeld, Germany

⁸ Institut für Transfusionsmedizin, Universitätsklinikum Köln. Kerpenerstr. 62, 50937 Köln

⁹ Department of Paediatrics, University Hospital of Schleswig-Holstein, Campus Luebeck,
23538 Luebeck, Germany

26

27 **Scientific category: Lymphoid Neoplasia**

28 **Running Title:** HGBL-DH/TH mutational landscape

29 **Key Words:** High-grade B-cell lymphoma, mutational landscape, whole-exome sequencing

30 **Word Count (Text): 3.972**

31 **Word Count (Abstract): 250**

32 **References: 43**

33 **Tables: 1**

34 **Figures: 5**

35 **Supplementary Tables: 7**

36 **Supplementary Figures: 7**

37 **Competing Interests:** The authors declare no conflicts of interest.

38

39 *These authors contributed equally to this manuscript

40 **Shared senior authorship

41 **Date of submission: July 14th2021**

42 **Corresponding Author:**

43 PD Dr. med. Niklas Gebauer

44 Department of Hematology and Oncology

45 UKSH Campus Luebeck

46 Ratzenburger Allee 160

47 23538 Luebeck

48 Email: Niklas.Gebauer@uksh.de

49

50 Abstract

51 High-grade B-cell lymphoma accompanied with *MYC* and *BCL2* and/or *BCL6* rearrangements
52 (HGBL-DH/TH) poses a cytogenetically-defined provisional entity among aggressive B-cell
53 lymphomas that is traditionally associated with unfavorable prognosis.

54 To better understand the mutational and molecular landscape of HGBL-DH/TH we here
55 performed whole-exome sequencing and deep panel next-generation-sequencing (NGS) of 47
56 clinically annotated cases. Oncogenic drivers, mutational signatures and perturbed pathways
57 were compared with data from follicular lymphoma (FL), diffuse large-B-cell lymphoma
58 (DLBCL) and Burkitt lymphoma (BL).

59 We find an accumulation of oncogenic mutations in NOTCH, IL6/JAK/STAT and NFκB
60 signaling pathways and delineate the mutational relationship within the continuum between
61 FL/DLBCL, HGBL-DH/TH and BL. Further, we provide evidence of a molecular divergence
62 between *BCL2* and *BCL6* rearranged HGBL-DH. Beyond a significant congruency with the
63 C3/EZB DLBCL cluster in *BCL2* rearranged cases on an exome-wide level, we observe an
64 enrichment of the SBS6 mutation signature in *BCL6* rearranged cases. Differential gene set
65 enrichment and subsequent network propagation analysis according to cytogenetically defined
66 subgroups revealed an impairment of TP53 and MYC pathway signaling in *BCL2* rearranged
67 cases, whereas *BCL6* rearranged cases lacked this enrichment, but instead exhibited showed
68 impairment of E2F targets. Intriguingly, HGBL-TH displayed intermediate mutational
69 features in all three aspects.

70 This study elucidates a recurrent pattern of mutational events driving FL into *MYC*-driven
71 *BCL2* rearranged HGBL, unveiling the mutational pathogenesis of this provisional entity.
72 Through this refinement of the molecular taxonomy for aggressive, germinal-center derived
73 B-cell lymphomas, this calls into question the current WHO classification system, especially
74 regarding the status of *MYC/BCL6* rearranged HGBL.

75

76 **Introduction**

77 High-grade B-cell lymphoma with *MYC*-, *BCL2* and/or *BCL6* (HGBL) rearrangements poses
 78 a novel, yet provisional, cytogenetically-defined entity within the current WHO classification
 79 of lymphoid tumors. It is presumably allocated in the pathobiological continuum between
 80 diffuse large B-cell (DLBCL) and Burkitt lymphoma (BL) (1). The t(8;14)(q24;q32)
 81 IgH/*MYC* rearrangement constitutes the molecular hallmark of BL. This or further derivative
 82 chromosomal rearrangements that juxtapose *MYC* to a genomic enhancer, occur in
 83 approximately 10% of DLBCL and have been shown to correlate with inferior clinical
 84 outcome (2). The rearrangement is a driver of oncogenesis that is in approx. 50% of cases
 85 accompanied by additional rearrangements involving *BCL2* and/or *BCL6* and referred to as
 86 double-hit (DH) or triple-hit (TH) lymphomas (3-11).

87 While the clinical outcome in HGBL-DH/TH patients is generally poor, recent studies have
 88 emphasized the significant impact of *MYC* translocation partners and defined *MYC*/Ig-
 89 rearrangements to be the most reliable predictors of adverse outcome (2, 7).

90 In a prior study, we discovered an elevated frequency of *TP53* impairment in *MYC*-driven
 91 DH/TH, whose presence was subsequently demonstrated for a subset of patients with a
 92 single-hit *MYC* translocation as well, indicating inferior outcome (12, 13).

93 By conventional cytogenetics HGBL-DH/TH were shown to recurrently harbor a complex
 94 karyotype (4). Data on the genetic basis of this entity, however, remains elusive. Several
 95 preliminary studies, predominantly focusing on HGBL-DH/TH with DLBCL morphology,
 96 have employed a panel-based next generation sequencing (NGS) approach (14-16). The
 97 insights from these studies were all restricted by gene-panel design and the associated
 98 clinicopathological data, yet their central assertions included a significant enrichment in
 99 mutations affecting *CREBBP*, *BCL2* and *KMT2D* alongside an overall reflection of the
 100 phenotypical gray-zone between DLBCL and BL. Most recently, Cucco *et al.* elucidated
 101 significant aspects of the molecular signature of HGBL in a panel-based sequencing and gene

expression study, employing a 70-gene HaloPlex panel and an array-based gene expression approach. The authors restricted their study to samples with DLBCL morphology that stemmed from a clinical trial and the UK's population-based Haematological Malignancy Research Network (17).

A comprehensive, exome-wide assessment of oncogenic driver mutations in HGBL-DH/TH, including cases with BL-like morphology is, however, still warranted and of vital importance to the refinement of the pathogenetic understanding of this clinically challenging entity ultimately enabling targeted therapeutic approaches.

We therefore conducted a whole-exome sequencing (WES) study on a large cohort of HGBL-DH/TH, validated by panel-based NGS and supplemented these data with a comprehensive clinicopathological assessment of the study group. Here we report on oncogenic drivers, somatic copy number alterations (SCNAs) and putative pathway perturbations, thus refining the molecular taxonomy of *MYC*-driven germinal-center-derived aggressive lymphomas.

Materials and methods

Case selection and clinicopathological characteristics

In a retrospective approach, we reviewed our institutional database to identify HGBL patients whose primary diagnostic biopsy specimen had been referred to the Reference center for Hematopathology University Hospital Schleswig Holstein Campus Lübeck and Hämatopathologie Lübeck for centralized histopathological panel evaluation between January 2007 and December 2019. For additional Information on clinicopathological work-up, please see **Supplementary materials and methods** and **Supplementary Table 1**.

This retrospective study was approved by the ethics committee of the University of Lübeck (reference-no 18-356) and conducted in accordance with the declaration of Helsinki. Patients had given written informed consent regarding routine diagnostic and academic assessment of

their biopsy specimen including molecular studies at the Reference center for Hematopathology and transfer of their clinical data.

Whole exome and targeted amplicon-based sequencing

Whole exome sequencing of $n = 47$ HGBL-DH/TH samples was performed by a hybrid capture approach with the Agilent SureSelect Human All Exon V6 library preparation kit (Agilent Technologies) followed by Illumina short read sequencing on a NovaSeq platform (Illumina) to an average depth of 303x (standard deviation $\pm 152x$; median 237x) by Novogene (UK) Co., Ltd. Seeking to validate the initial delineation of the exome sequencing-derived mutational landscape in HGBL we employed our in-house custom AmpliSeq panel (Thermo Fisher Scientific, Waltham, Massachusetts, USA) for targeted amplicon sequencing (tNGS), encompassing all coding exons of 43 genes (see **Supplementary Table 2**) in 21 cases. Raw paired-end data (*fastq* format) was trimmed and quality filtered using FASTP (v0.20.0; minimum length 50bp, max. unqualified bases 30%, trim tail set to 1) (18) and trimmed reads were mapped to GRCh37/hg19 using BWA MEM (v0.7.15) (19). Resulting alignment files in *SAM* format were cleaned, sorted, and converted into *BAM* format using PICARD TOOLS (v2.18.4). Single nucleotide variants (SNVs), as well as short insertions and deletions (InDels) were identified following the best practices for somatic mutations calling provided by GATK (20). Somatic copy number aberrations (SCNAs) were identified by CONTROL-FREEC (v11.4) (21). For further details on nucleic acid extraction, panel sequencing, single nucleotide and copy number variant calling please see **Supplementary methods**.

Mutational deleteriousness and significance, network propagation, gene set enrichment and mutational cluster analysis

The MUTSIGCV algorithm was employed on WES data to delineate significantly mutated genes within the study cohort while deleteriousness was assessed via the CADD v1.3. The acquired genomic data were then processed through the LymphGen algorithm and underwent manual screening for an enrichment in overlapping aberrations with the molecular clusters proposed by Chapuy *et al.* followed by validation through a logistic regression framework (22, 23). Cytogenetically defined subgroups (HGBL with *MYC* and *BCL2* aberrations, HGBL with *MYC* and *BCL6* aberrations, and HGBL-TH) underwent differential downstream analysis by a network propagation approach simulating a protein-protein interaction network. Subsequently a gene set variation analysis was performed against HALLMARK gene sets. For details including statistical approaches correlating molecular and clinicopathological findings please see **Supplementary methods**.

Data availability

Sequencing data in bam format from WES and panel sequencing have been deposited in the European genome-phenome archive (EGA) under the accession number EGAS00001005420.

Results

Clinicopathological characteristics of the study group

We collected 47 cases of HGBL-DH/TH at diagnosis with sufficient FFPE tissue samples for molecular studies (median age 71; range 35 – 89 years) all of which were included in the final analysis, following successful library preparation for WES. There was insufficient clinical follow-up in 9/47 (19%) cases. An underlying HIV infection was clinically excluded in all cases. The majority of patients in our study were male (25/47; 53%) and presented with advanced stage disease (24/38 stage III/IV; 63%) and an adverse prognostic constellation (24/38 (63%) R-IPI > 2). Most patients received an intensive CHOP-like therapeutic frontline

approach (25/38; 66%). The overall response rate after first line (immuno-) chemotherapy was 76% resembling a general therapeutic response in 29/38 cases.

Table 1 summarizes the baseline characteristics of all HGBL-DH/TH cases included in the current study. Histologically, the predominant morphology was that of DLBCL (NOS) (32/47), however, Burkitt-like morphology and immunophenotype was present in 15/47 patients. Cytogenetically 21/47 cases presented with *MYC/BCL2*, 17/47 with *MYC/BCL6* double-hit constellation and 9 triple-hit lymphomas were included. *MYC* translocation partner revealed *MYC*-Ig rearrangement in 8/17 cases. The treatment outcome in our cohort was unfavorable yet in keeping with previous data reported by Rosenwald and colleagues (2). For confirmatory purposes, we included four cases, which were assessed for *TP53* mutation status in a previous study and were able to validate both cytogenetic as well as molecular observations (12).

The mutational landscape of HGBL-DH/TH identified by WES

To characterize the mutational landscape in an extensive cohort of HGBL-DH/TH cases, we successfully performed WES in 47 patient-derived tumor biopsies and matched constitutional DNA in seven cases. We further applied the analytical framework outlined above to analyze WES data in the absence of paired germline DNA in the majority of cases. Following the primary identification of SNVs and indels in individual samples and subsequent filtering to correct for FFPE-derived artefacts and spurious mutations, we applied the MutSig2CV algorithm and thereby identified 22 significant candidate driver genes ($p < 0.001$; 13 genes with $q < 0.1$; **Supplementary Table 3**) (24).

All HGBL-DH/TH cases carried mutations in genes of oncogenic relevance according to our bioinformatic annotations. In total, we described 10,092 presumably harmful somatic mutations (cut-off see materials and methods) involving 5,521 genes after variant filtering. Of these, SNVs and InDels represented 74.1% of the mutations (7,479 SNVs). Among them,

missense mutations were the most frequent alterations (85.2%), followed by nonsense (5.7%) and InDels (5.6%), while splice-site mutations posed 3.3% of somatic mutations (**Figure 1A**). Displaying an overall intermediate tumor mutational burden (median 3.974; range 1.065 – 18.234 mutations/Mbase; **Figure 1B**), HGBL-DH/TH revealed no evidence of MSI-related hypermutations, which is in keeping with observations in DLBCL (0.3%), yet differs from other aggressive lymphomas (e.g., primary mediastinal B-cell lymphoma) (25).

Upon comparative analysis of WES and targeted resequencing data, we were able to demonstrate a concordance rate of 92.0% (46/50 in 18 matched samples) of mutational calls, prompting high confidence in mutational calls derived from WES, even in non-germline-paired cases. A comprehensive description of all variants described by WES as well as panel based NGS is provided in **Supplementary Tables 4 and 5**. Nevertheless, we observed a significant enrichment of non-germline matched samples in non-synonymous SNVs, which prompted us to include significantly mutated genes according to the MUTSIGCV analysis, only. As an exception to this rule, we also included *MYC* mutations below the statistical significance level due to their previously established clinical and functional relevance.

Recurrent copy number alterations in HGBL-DH/TH

We investigated our HGBL-DH/TH cohort for SCNVs employing the CONTROL-FREEC (21) algorithm in tumor-normal and tumor-only mode, respectively, followed by GISTIC2.0 (26) analysis. The analysis excluded chromosomes X and Y as well as common benign copy number variants defined by the Broad Institute's panel of normal. Upon cross-referencing our findings with genomic loci of known oncogenes, tumor-suppressors and elements of significant signaling pathways we identified recurrent copy number gains in oncogenes such as *MEF2B* and *CSF1R*, which have previously been implicated in the pathogenesis of malignant lymphomas (27, 28). Further, copy number losses in tumor suppressors like *NPM1*

were recurrently identified (**Figure 2A, B**). No significant differences were detected for genes affected by copy number alterations between the three cytogenetically defined subgroups (Fisher exact test $p > 0.05$ after Bonferroni correction for multiple testing). Common CNVs, as defined by the above referenced panel of normal were encountered at the expected frequencies (29, 30).

Significantly mutated candidate driver genes and mutational signatures

Putative candidate driver genes comprised several genes previously implicated in HGBL-DH/TH pathogenesis, such as *KMT2D*, *CREBBP* and *TP53* alongside several further mutated genes such as *CDKN2A*, *LNPI* or *SI* (14). Established mutational drivers known from other B-cell lymphoproliferative disorders (e.g., FL, DLBCL, BL) were recurrently encountered (e.g., *CCND3*, *ARID1A*) (**Figures 2C; 3A**) (31, 32). In our limited cohort, distinctions regarding subtype-specific mutational signatures were found to be marginal among *BCL2/BCL6* status or Burkitt-like vs non-Burkitt like morphology. However, we found *CCND3* mutations, previously reported as driver mutations in BL pathogenesis, to be significantly enriched in HGBL-DH/TH patients with Burkitt-like morphology (9/15 vs 3/28). This observation hints at partially similar molecular paths of pathogenesis between BL and HGBL with Burkitt-like morphology. Further, an enrichment of mutations affecting *CREBBP* in HGBL-DH/TH patients with *BCL2* rearrangement was observed, which is well in keeping with its proposed fundamental role in FL pathogenesis. Distribution of mutations within selected, significantly mutated genes is depicted in **Supplementary Figure 1**. Additional profiling of mutational signatures driving HGBL-DH/TH revealed a predominance of the SBS5 signature alongside the emphasized occurrence of the SBS6 signature (implicated in defective DNA mismatch repair) in patients with *BCL6* rearrangements (**Supplementary Figure 2, Supplementary Table 6**).

Comparative analysis of mutational landscape in HGBL-DH/TH, related entities and molecular clusters in DLBCL

Next, we sought to refine the genomic taxonomy of aggressive germinal-center-derived B-cell lymphomas and to investigate the mutational commonalities and differences between HGBL-DT/TH and other related pathological entities. These were subsequently selected for their similar features of B-cell differentiation (ABC-type DLBCL, FL, GCB-type DLBCL and BL) and a comparative analysis of candidate mutational drivers in HGBL-DH/TH (as described earlier) and cBioPortal cohorts of ABC-type DLBCL (n = 67), GCB-type DLBCL (n = 45 (22)), BL (n = 108(33)) and FL (n = 199(33)) was conducted. Interestingly, we identified one overlapping candidate driver common to all entities (*KMT2D*). Additionally, *CREBBP* was found in all entities except ABC-type DLBCL. Mutations affecting *EZH2*, *IRF8* and *TNFRSF14* were, however, specifically occurred in HGBL-DT/TH and FL/GCB-type DLBCL, while *CCND3* mutations appeared to be a pathogenetic feature shared between BL and HGBL-DT/TH. Additionally, *TP53* mutations posed a predominant feature of aggressive lymphomas present in all HGBL-DH/TH, GCB-type DLBCL and BL types and therefore most likely acquired during high-grade transformation (**Figure 3G, H**). In basic accordance with previous studies, our data suggest a common origin especially for *BCL2* rearranged HGBL-DH/TH and FL/GCB-type DLBCL (14, 17).

Upon comparative investigation of our current data and mutational clusters, previously described in DLBCL, we observed a striking predominance of C3/EZB cluster cases in the *BCL2* rearranged subgroup according to the integrative molecular classification proposed by Chapuy *et al.* and Wright *et al.*, respectively. This is in keeping with a significant enrichment of these cases with DLBCL morphology in terms of *MYC* rearrangement status (**Figure 4, Supplementary Figure 3**) (22, 23). Complementary to our analysis, employing the LymphGen algorithm (cf. **Supplementary Table 7**), a logistic regression indicated a significantly different number of mutated genes in C3 between the HGBL subtypes. HGBL

harboring only *BCL2* were shown to exhibit the highest number of mutated C3 genes, while HGBL with *BCL6* alterations had the lowest number of mutated C3 genes (*BCL2/6* cases: (p = 6.059×10^{-5} , adj R² = 0.3108). In contrast to triple hit cases, HGBL with *MYC* and an isolated additional *BCL6* rearrangement showed a significant decrease in the number of mutated C3 genes (p = 2.00×10^{-5} , estimate: -1.7589) (**Supplementary Figure 4, Supplementary Table 8**). Within the subgroup of *BCL6* rearranged cases, the BN2 cluster was more prominent than the EZB cluster. In keeping with their strong affinity towards the C3/EZB cluster, *BCL2* rearranged cases exhibited an enrichment for mutations in *CREBBP* and *KMT2D*, while *BCL6* rearranged cases were, in contrast, enriched for mutations in *ARID1A*. The vast majority of triple-hit cases was also classified within the EZB cluster.

Mutational impairment of NOTCH, RTK-RAS and TP53 signaling in HGBL-DH/TH

Cumulatively, we detected genetic lesions, putatively impairing NOTCH signaling in 74% of HGBL-DH/TH patients. Expanding on previously reported *CREBBP*, *EP300* and *DTX1* mutations in HGBL we further identified recurrent mutations affecting *NCOR1* and others (**Supplementary Figure 5**) (14, 17). NOTCH signaling was thereby the predominant target of somatic mutation in HGBL-DH/TH, albeit with a quite heterogeneous mutational pattern affecting 35/47 patients with lesions in 28/71 genes (**Supplemental Figure 6A**). Most of these genomic aberrations had been previously reported to be gain-of-function mutations putatively resulting in constitutive NOTCH pathway activation in various types of predominantly germinal-center derived B-cell lymphoma. Several of these mutational hits including *NCOR1* and *DTX1* have been shown to herald adverse clinical outcome (34, 35).

This remained the case when undertaking a differential downstream analysis within the cytogenetically defined subgroups (HGBL with *MYC* and *BCL2* aberrations, HGBL with *MYC* and *BCL6* aberrations and HGBL-TH), which was prompted by their significantly divergent distribution onto molecular clusters. Through this analysis a mutational signature

became apparent that is additionally dominated by impairment of TP53 and MYC signaling in *BCL2* rearranged cases. *BCL6* rearranged cases lacked this enrichment, while HGBL-TH cases revealed intermediate mutational features. As another remarkable feature throughout all subgroups we observed alterations, putatively resulting in gain-of-function in IL6/JAK/STAT signaling in 74% of patients (**Supplementary Figure 5**). Activating mutations affecting *PIMI* and *SOCS1* were most frequently encountered in our case series and have been previously implicated in HGBL-DH/TH pathogenesis (17). These candidate driver genes are supplemented by mutations in *LTB* and *STAT3* (both previously identified in HIV-associated plasmablastic lymphoma) among others (36). Beyond this, we observed a relatively dispersed mutational pattern with putative driver events affecting 28 genes within the NOTCH-pathway (**Supplemental Figure 6A**).

In accordance with previous studies, we found mutations directly impacting NF-κB signaling in 62% of cases (**Supplementary Figure 5**) (37). While this was among the predominant pathways identified through our network propagation approach, mutations affecting the pathway were narrowly detectable with only *BCL2* harboring mutations in more than three patients (34%) followed by recurrent SNVs and indels in *PARP1* (6%) and *BIRC3* (4%) and *CARD11* (4%).

Following SNV and InDel evaluation with MUTSIGCV, a network propagation approach (**Figures 3B, C**) was employed on significantly mutated genes to delineate the functional implications of significant genetic events on neighboring genes. This further underscored the mutational impairment of the aforementioned pathways alongside WNT and PI3K signaling. These observations are in accordance with preliminary impressions derived from targeted sequencing studies, employing panel-based approaches (14, 17). In addition to the divergent results from our mutational pathway analysis, we identified an enrichment of E2F targets impacted by significantly mutated genes in both double- and triple-hit cases affected by *BCL6* rearrangements (**Figures 3D, E, F**).

Survival analysis

Upon integrated analysis of molecular and clinical data we investigated genomic alterations present in > 15% of patients for their impact on overall survival (OS) and progression-free survival (PFS). Hereby we identified *ARID1A* mutations to be predictive of worse clinical outcome in our cohort (OS: $p = 0.0049$; PFS: 0.025). Subsequent Bonferroni correction for multiple testing was performed. Thus, we identified a significant impact of mutations affecting *ARID1A* which was maintained regarding OS when correction for multiple testing was applied while its primarily significant effect on PFS was reduced to a trend of borderline statistical significance (**Figure 5**). A Cox proportional hazard model revealed this effect to be independent of the established clinical IPI prognosticators (age, LDH, extra nodal manifestations, stage, and performance status; OS: $p < 0.001$; HR: 13.989; 95%CI: 3.362 – 58.205; PFS: $p = 0.001$; HR: 6.648; 95%CI: 2.098 – 21.061). Within the cytogenetically defined subgroups, we identified no alterations with independent impact on clinical outcome. However, a trend of borderline statistical significance towards inferior outcome in *MYC/BCL2* rearranged cases harboring *FOXO1* mutations was observed (**Supplementary Figure 7**).

Discussion

Here we report on whole-exome sequencing data from an extensive cohort of HGBL-DH/TH tumors, which is to the best of our knowledge the hitherto largest cohort and most extensive molecular data set for this entity. Previous reports on HGBL-DH/TH were limited by low sample numbers and/or targeted sequencing approaches. Contrary to this, WES here allowed to systematically define recurrent mutations, predominant mutational signatures and SCNVs in their respective clinicopathological context from which we report three central observations.

Firstly, being the first exome-wide mutational investigation for this rare subtype of lymphoma, we identify a significant overlap of mutational drivers between HGBL-DH/TH and FL as well as GCB-type DLBCL (e.g., *TNFRSF14*, *EZH2* and *IRF4*) as its high-grade counterpart. Aggressive transformation was associated with the acquisition of mutations in *TP53*. Moreover, shared features, including *CCND3* and *CDKN2A* mutations underscore a close molecular relation between HGBL-DH/TH and BL (38, 39). This is additionally reflected in the enrichment of HGBL-DH/TH patients with Burkitt-like morphology for *CCND3* mutations. Further, we identify a number of significant mutational drivers not captured by previous, panel-based sequencing studies. Most frequently among these, we find *SI* mutations that have been previously implicated in CLL progression as well as mutations in *POU2AF1*, which has been recently found to be an augmented target of mutations during aggressive transformation of FL to DLBCL (40, 41). Although *MYC* did not meet the predefined MUTSIGCV significance level in our study, we still observed mutations in 19% of cohort samples (**Supplementary Figure 8**), which is in agreement with previous panel-based studies (17).

Secondly, upon screening the mutational landscape in HGBL-DH/TH in comparison to the molecular clusters of DLBCL, proposed by Chapuy *et al.* and the LymphGen algorithm proposed by Wright *et al.*, we unveil a striking overlap of *BCL2* rearranged cases with the C3/EZB cluster, which was previously shown to be enriched for *MYC* rearrangements and oncogenic drivers implicated in FL pathogenesis (22, 23). We argue that a predominant subset of HGBL-DH/TH most likely corresponds to these transformed FL. This offers a potential explanation for the inferior clinical outcome of C3 DLBCL patients, despite their GCB-phenotype, through an enrichment for *MYC* rearranged HGBL-DH/TH cases. Of note, we find the predominant impairment of *TP53* and to a lesser extent *MYC* signaling in *BCL2* rearranged cases to be in keeping with a previous study on an independent set of HGBL-DH/TH, in which we found *TP53* mutations to be a recurrent feature of HGBL-DH with

BCL2, but not *BCL6* rearrangements (12). Intriguingly, we further observed two molecular subtypes in *MYC/BCL6* only rearranged cases. While selected cases were categorized within the EZB cluster, several cases revealed an association with the BN2 cluster, potentially hinting at a *MYC*-driven high-grade transformation of a precursor lesion with an origin within the marginal zone, as previously described (23). In addition to these observations, we found triple-hit cases to reflect EZB lymphomas in the vast majority of cases, potentially hinting at *BCL6* rearrangements as late and non-defining events in HGBL-TH lymphomagenesis. Supporting this assumption, Pedrosa *et al.* have shown DLBCL with *BCL2* and *BCL6*, but without *MYC* rearrangements to be exclusively associated with the EZB cluster (42). Considering the significantly mutated genes, our observations underscore previous assumptions regarding a molecular divergence between *BCL2* and *BCL6* rearranged HGBL-DH (14, 17, 43). The predominant mutational distinction between these groups was the presumably FL-derived enrichment for *CREBBP* mutations in the *BCL2* rearranged subgroup. On an exome-wide level we observed an enrichment of the SBS6 signature (implicated in defective DNA mismatch repair) and a significantly diminished congruency with the C3/EZB DLBCL cluster in the *BCL6* rearranged subgroup. Of note, these findings fundamentally dispute the combined characterization in the current WHO classification, despite several shared clinical aspects common to all subtypes of HGBL-DH/TH (1, 2). Beyond *de novo* DLBCL with *BCL6* rearrangement, potential alternative explanations for this phenomenon include both clonal evolution and subsequent aggressive transformation from rare cases of *BCL6* rearranged marginal zone lymphomas alongside *BCL2* non-rearranged/*BCL6* rearranged FL, which were previously shown to be characterized by a heterogenous mutational landscape (44, 45). From our data, we further deduce an intermediate role for HGBL-TH, which may indicate two divergent paths of clonal evolution originating from a *BCL2* or a *BCL6* driven disease with subsequent acquisition of the alternative rearrangement. Lastly, we describe a pronounced mutational impairment of NOTCH, IL6/JAK/STAT and

NFκB signaling pathways and recurrent oncogenetically relevant genes affected by SCNVs (including *MEF2B*, which was previously shown to be enriched in mutations/aberrations within the C3 DLBCL cluster) thereby systematically characterize the oncogenetic footprint of this subgroup of lymphoma. This is further combined with the identification of novel putative mutational drivers (e.g., *NCOR1*, *DTX1*, *LTB* and *STAT3*) alongside several previously established mutational hotspots in HGBL-DH/TH. Moreover, among these significantly mutated genes we describe *ARID1A* which emerges as a potential prognosticator of treatment response and outcome from our correlative assessment of clinical and molecular features of our present cohort, which was found to be independent from previously established clinical prognostic factors.

We acknowledge the shortcomings inherent to the retrospective design of the study alongside the limited availability of germline DNA for matched pair analysis. The latter aspect is reflected in a significantly elevated number of mutations in non-matched samples. This prompted us to limit our subsequent analysis to significantly mutated genes (except for *MYC* and *BCL2* mutations, which were additionally included based on their proven relevance in prior studies (14, 17, 22, 46) and thereby equalizing the abovementioned effect. Pairing of our WES-results with RNA-seq data, preferably in an extended, clinically annotated cohort, which was beyond the scope of the present study, would further deepen our molecular understanding of HGBL-DH/TH, especially regarding cases with prominent Burkitt or Burkitt-like morphology.

In summary, our identification of distinct mutational landscapes among HGBL-DH/TH, derived from an exome-wide sequencing approach shows both overlapping and distinctive features compared with germinal center derived lymphomas such as GCB-type DLBCL and low-grade FL as well as BL. Our work further underscores the developing notion of a recurrent pattern of mutational events driving a potentially unidentified preexisting FL into *MYC*-driven HGBL-DH/TH, offering insight into the molecular pathogenesis of this

provisional entity. By refining the molecular taxonomy for aggressive, germinal-center derived B-cell lymphomas, these results call into question the current WHO classification system, especially regarding the status of *MYC/BCL6* rearranged HGBL.

Acknowledgments

A.K. and H.B. acknowledge computational support from the OMICS compute cluster at the University of Lübeck. The research was supported by a grant to N.G. by the Stefan-Morsch-Foundation alongside infrastructural support. H.B. acknowledges funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy— EXC 22167-390884018).

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri A, Stein H, Thiele J, Vardiman JW. WHO Classification of Tumors of Haematopoietic and Lymphoid Tissues. Vol. 2. Lyon: IARC; 2017. 586 p. ISBN: 9789283244943.
2. Rosenwald A, Bens S, Advani R, Barrans S, Copie-Bergman C, Elsensohn MH, Natkunam Y, Calaminici M, Sander B, Baia M, Smith A, Painter D, Pham L, Zhao S, Ziepert M, Jordanova ES, Molina TJ, Kersten MJ, Kimby E, Klapper W, Raemaekers J, Schmitz N, Jardin F, Stevens WBC, Hoster E, Hagenbeek A, Gribben JG, Siebert R, Gascoyne RD, Scott DW, Gaulard P, Salles G, Burton C, de Jong D, Sehn LH, Maucort-Boulch D. Prognostic Significance of MYC Rearrangement and Translocation Partner in Diffuse Large B-Cell Lymphoma: A Study by the Lunenburg Lymphoma Biomarker Consortium. J Clin Oncol. 2019 Dec 10;37(35):3359-3368. Epub 2019/09/10. doi:10.1200/JCO.19.00743. Cited in: Pubmed; PMID 31498031.
3. Aukema SM, Kreuz M, Kohler CW, Rosolowski M, Hasenclever D, Hummel M, Kuppers R, Lenze D, Ott G, Pott C, Richter J, Rosenwald A, Szczepanowski M, Schwaenen C, Stein H, Trautmann H, Wessendorf S, Trumper L, Loeffler M, Spang R, Kluin PM, Klapper W, Siebert R, Molecular Mechanisms in Malignant Lymphomas Network P. Biological

characterization of adult MYC-translocation-positive mature B-cell lymphomas other than
molecular Burkitt lymphoma. *Haematologica*. 2014 Apr;99(4):726-35.
doi:10.3324/haematol.2013.091827. Cited in: Pubmed; PMID 24179151.

4. Aukema SM, Siebert R, Schuurin E, van Imhoff GW, Kluin-Nelemans HC, Boerma EJ,
Kluin PM. Double-hit B-cell lymphomas. *Blood*. 2011 Feb 24;117(8):2319-31.
doi:10.1182/blood-2010-09-297879. Cited in: Pubmed; PMID 21119107.

5. Copie-Bergman C, Cuilliere-Dartigues P, Baia M, Briere J, Delarue R, Canioni D, Salles
G, Parrens M, Belhadj K, Fabiani B, Recher C, Petrella T, Ketterer N, Peyrade F, Haioun C,
Nagel I, Siebert R, Jardin F, Leroy K, Jais JP, Tilly H, Molina TJ, Gaulard P. MYC-IG
rearrangements are negative predictors of survival in DLBCL patients treated with
immunochemotherapy: a GELA/LYSA study. *Blood*. 2015 Nov 26;126(22):2466-74.
doi:10.1182/blood-2015-05-647602. Cited in: Pubmed; PMID 26373676.

6. Oki Y, Noorani M, Lin P, Davis RE, Neelapu SS, Ma L, Ahmed M, Rodriguez MA,
Hagemeister FB, Fowler N, Wang M, Fanale MA, Nastoupil L, Samaniego F, Lee HJ, Dabaja
BS, Pinnix CC, Medeiros LJ, Nieto Y, Khouri I, Kwak LW, Turturro F, Romaguera JE, Fayad
LE, Westin JR. Double hit lymphoma: the MD Anderson Cancer Center clinical experience.
Br J Haematol. 2014 Sep;166(6):891-901. doi:10.1111/bjh.12982. Cited in: Pubmed; PMID
24943107.

7. Pedersen MO, Gang AO, Poulsen TS, Knudsen H, Lauritzen AF, Nielsen SL, Klausen TW,
Norgaard P. MYC translocation partner gene determines survival of patients with large B-cell
lymphoma with MYC- or double-hit MYC/BCL2 translocations. *Eur J Haematol*. 2014
Jan;92(1):42-8. doi:10.1111/ejh.12212. Cited in: Pubmed; PMID 24118498.

8. Petrich AM, Nabhan C, Smith SM. MYC-associated and double-hit lymphomas: a review
of pathobiology, prognosis, and therapeutic approaches. *Cancer*. 2014 Dec 15;120(24):3884-
95. doi:10.1002/cncr.28899. Cited in: Pubmed; PMID 25060588.

9. Pillai RK, Sathanoori M, Van Oss SB, Swerdlow SH. Double-hit B-cell lymphomas with
BCL6 and MYC translocations are aggressive, frequently extranodal lymphomas distinct

from BCL2 double-hit B-cell lymphomas. *Am J Surg Pathol*. 2013 Mar;37(3):323-32.
doi:10.1097/PAS.0b013e31826cebad. Cited in: Pubmed; PMID 23348205.

10. Sarkozy C, Traverse-Glehen A, Coiffier B. Double-hit and double-protein-expression lymphomas: aggressive and refractory lymphomas. *Lancet Oncol*. 2015 Nov;16(15):e555-67.
doi:10.1016/S1470-2045(15)00005-4. Cited in: Pubmed; PMID 26545844.

11. Wang XJ, Medeiros LJ, Lin P, Yin CC, Hu S, Thompson MA, Li S. MYC cytogenetic status correlates with expression and has prognostic significance in patients with MYC/BCL2 protein double-positive diffuse large B-cell lymphoma. *Am J Surg Pathol*. 2015 Sep;39(9):1250-8. doi:10.1097/PAS.0000000000000433. Cited in: Pubmed; PMID 25828389.

12. Gebauer N, Bernard V, Gebauer W, Thorns C, Feller AC, Merz H. TP53 mutations are frequent events in double-hit B-cell lymphomas with MYC and BCL2 but not MYC and BCL6 translocations. *Leuk Lymphoma*. 2015 Jan;56(1):179-85. eng. Epub 2014/04/01. doi:10.3109/10428194.2014.907896. Cited in: Pubmed; PMID 24679006.

13. Schiefer AI, Kornauth C, Simonitsch-Klupp I, Skrabbs C, Masel EK, Streubel B, Vanura K, Walter K, Migschitz B, Stoiber D, Sexl V, Raderer M, Chott A, da Silva MG, Cabecadas J, Mullauer L, Jager U, Porpaczy E. Impact of Single or Combined Genomic Alterations of TP53, MYC, and BCL2 on Survival of Patients With Diffuse Large B-Cell Lymphomas: A Retrospective Cohort Study. *Medicine (Baltimore)*. 2015 Dec;94(52):e2388. Epub 2015/12/31. doi:10.1097/MD.00000000000002388. Cited in: Pubmed; PMID 26717387.

14. Evrard SM, Pericart S, Grand D, Amara N, Escudie F, Gilhodes J, Bories P, Traverse-Glehen A, Dubois R, Brousset P, Parrens M, Laurent C. Targeted next generation sequencing reveals high mutation frequency of CREBBP, BCL2 and KMT2D in high-grade B-cell lymphoma with MYC and BCL2 and/or BCL6 rearrangements. *Haematologica*. 2019 Apr;104(4):e154-e157. Epub 2018/10/13. doi:10.3324/haematol.2018.198572. Cited in: Pubmed; PMID 30309852.

15. Stengel A, Kern W, Meggendorfer M, Haferlach T, Haferlach C. Detailed molecular analysis and evaluation of prognosis in cases with high grade B-cell lymphoma with MYC and BCL2 and/or BCL6 rearrangements. *Br J Haematol*. 2019 Jun;185(5):951-954. Epub 2018/11/22. doi:10.1111/bjh.15653. Cited in: Pubmed; PMID 30460680.

533

534 16. Momose S, Weissbach S, Pischmarov J, Nedeva T, Bach E, Rudelius M, Geissinger E,
535 Staiger AM, Ott G, Rosenwald A. The diagnostic gray zone between Burkitt lymphoma and
536 diffuse large B-cell lymphoma is also a gray zone of the mutational spectrum. *Leukemia*.
537 2015 Aug;29(8):1789-91. doi:10.1038/leu.2015.34. Cited in: Pubmed; PMID 25673238.

538

539 17. Cucco F, Barrans S, Sha C, Clipson A, Crouch S, Dobson R, Chen Z, Thompson JS, Care
540 MA, Cummin T, Caddy J, Liu H, Robinson A, Schuh A, Fitzgibbon J, Painter D, Smith A,
541 Roman E, Tooze R, Burton C, Davies AJ, Westhead DR, Johnson PWM, Du MQ. Distinct
542 genetic changes reveal evolutionary history and heterogeneous molecular grade of DLBCL
543 with MYC/BCL2 double-hit. *Leukemia*. 2020 May;34(5):1329-1341. Epub 2019/12/18.
544 doi:10.1038/s41375-019-0691-6. Cited in: Pubmed; PMID 31844144.

545

546 18. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor.
547 *Bioinformatics*. 2018 Sep 1;34(17):i884-i890. Epub 2018/11/14.
548 doi:10.1093/bioinformatics/bty560. Cited in: Pubmed; PMID 30423086.

549 19. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
550 2013. [cited 2020 07.06.2020]. Available from: <https://arxiv.org/abs/1303.3997>.

551

552 20. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K,
553 Altshuler D, Gabriel S, Daly M, DePristo MA. The Genome Analysis Toolkit: a MapReduce
554 framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010
555 Sep;20(9):1297-303. Epub 2010/07/21. doi:10.1101/gr.107524.110. Cited in: Pubmed; PMID
556 20644199.

557

558 21. Boeva V, Popova T, Bleakley K, Chiche P, Cappel J, Schleiermacher G, Janoueix-Lerosey
559 I, Delattre O, Barillot E. Control-FREEC: a tool for assessing copy number and allelic content
560 using next-generation sequencing data. *Bioinformatics*. 2012 Feb 1;28(3):423-5. Epub
561 2011/12/14. doi:10.1093/bioinformatics/btr670. Cited in: Pubmed; PMID 22155870.

562

563 22. Chapuy B, Stewart C, Dunford AJ, Kim J, Kamburov A, Redd RA, Lawrence MS,
564 Roemer MGM, Li AJ, Ziepert M, Staiger AM, Wala JA, Ducar MD, Leshchiner I, Rheinbay
565 E, Taylor-Weiner A, Coughlin CA, Hess JM, Peadarallu CS, Livitz D, Rosebrock D,
566 Rosenberg M, Tracy AA, Horn H, van Hummelen P, Feldman AL, Link BK, Novak AJ,

- 567 Cerhan JR, Habermann TM, Siebert R, Rosenwald A, Thorner AR, Meyerson ML, Golub TR,
568 Beroukhim R, Wulf GG, Ott G, Rodig SJ, Monti S, Neuberg DS, Loeffler M, Pfreundschuh
569 M, Trumper L, Getz G, Shipp MA. Molecular subtypes of diffuse large B cell lymphoma are
570 associated with distinct pathogenic mechanisms and outcomes. *Nat Med*. 2018
571 May;24(5):679-690. Epub 2018/05/02. doi:10.1038/s41591-018-0016-8. Cited in: Pubmed;
572 PMID 29713087.
573
- 574 23. Wright GW, Huang DW, Phelan JD, Coulibaly ZA, Roulland S, Young RM, Wang JQ,
575 Schmitz R, Morin RD, Tang J, Jiang A, Bagaev A, Plotnikova O, Kotlov N, Johnson CA,
576 Wilson WH, Scott DW, Staudt LM. A Probabilistic Classification Tool for Genetic Subtypes
577 of Diffuse Large B Cell Lymphoma with Therapeutic Implications. *Cancer Cell*. 2020 Apr
578 13;37(4):551-568 e14. Epub 2020/04/15. doi:10.1016/j.ccell.2020.03.015. Cited in: Pubmed;
579 PMID 32289277.
580
- 581 24. Lawrence MS, Stojanov P, Polak P, Kryukov GV, Cibulskis K, Sivachenko A, Carter SL,
582 Stewart C, Mermel CH, Roberts SA, Kiezun A, Hammerman PS, McKenna A, Drier Y, Zou
583 L, Ramos AH, Pugh TJ, Stransky N, Helman E, Kim J, Sougnez C, Ambrogio L, Nickerson
584 E, Shefler E, Cortes ML, Auclair D, Saksena G, Voet D, Noble M, DiCara D, Lin P,
585 Lichtenstein L, Heiman DI, Fennell T, Imielinski M, Hernandez B, Hodis E, Baca S, Dulak
586 AM, Lohr J, Landau DA, Wu CJ, Melendez-Zajgla J, Hidalgo-Miranda A, Koren A,
587 McCarroll SA, Mora J, Crompton B, Onofrio R, Parkin M, Winckler W, Ardlie K, Gabriel
588 SB, Roberts CWM, Biegel JA, Stegmaier K, Bass AJ, Garraway LA, Meyerson M, Golub
589 TR, Gordenin DA, Sunyaev S, Lander ES, Getz G. Mutational heterogeneity in cancer and the
590 search for new cancer-associated genes. *Nature*. 2013 Jul 11;499(7457):214-218. Epub
591 2013/06/19. doi:10.1038/nature12213. Cited in: Pubmed; PMID 23770567.
592
- 593 25. Chapuy B, Stewart C, Dunford AJ, Kim J, Wienand K, Kamburov A, Griffin GK, Chen
594 PH, Lako A, Redd RA, Cote CM, Ducar MD, Thorner AR, Rodig SJ, Getz G, Shipp MA.
595 Genomic analyses of PMBL reveal new drivers and mechanisms of sensitivity to PD-1
596 blockade. *Blood*. 2019 Dec 26;134(26):2369-2382. Epub 2019/11/08.
597 doi:10.1182/blood.2019002067. Cited in: Pubmed; PMID 31697821.
598
- 599 26. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R, Getz G. GISTIC2.0
600 facilitates sensitive and confident localization of the targets of focal somatic copy-number

alteration in human cancers. *Genome Biol.* 2011;12(4):R41. Epub 2011/04/30.
doi:10.1186/gb-2011-12-4-r41. Cited in: Pubmed; PMID 21527027.

27. Brescia P, Schneider C, Holmes AB, Shen Q, Hussein S, Pasqualucci L, Basso K, Dalla-Favera R. MEF2B Instructs Germinal Center Development and Acts as an Oncogene in B Cell Lymphomagenesis. *Cancer Cell.* 2018 Sep 10;34(3):453-465 e9. Epub 2018/09/12.
doi:10.1016/j.ccell.2018.08.006. Cited in: Pubmed; PMID 30205047.

28. Edginton-White B, Cauchy P, Assi SA, Hartmann S, Riggs AG, Mathas S, Cockerill PN, Bonifer C. Global long terminal repeat activation participates in establishing the unique gene expression programme of classical Hodgkin lymphoma. *Leukemia.* 2019 Jun;33(6):1463-1474. Epub 2018/12/14. doi:10.1038/s41375-018-0311-x. Cited in: Pubmed; PMID 30546079.

29. Lamprecht B, Walter K, Kreher S, Kumar R, Hummel M, Lenze D, Kochert K, Bouhlei MA, Richter J, Soler E, Stadhouders R, Johrens K, Wurster KD, Callen DF, Harte MF, Giefing M, Barlow R, Stein H, Anagnostopoulos I, Janz M, Cockerill PN, Siebert R, Dorken B, Bonifer C, Mathas S. Derepression of an endogenous long terminal repeat activates the CSF1R proto-oncogene in human lymphoma. *Nat Med.* 2010 May;16(5):571-9, 1p following 579. Epub 2010/05/04. doi:10.1038/nm.2129. Cited in: Pubmed; PMID 20436485.

30. Pon JR, Wong J, Saberi S, Alder O, Moksa M, Grace Cheng SW, Morin GB, Hoodless PA, Hirst M, Marra MA. MEF2B mutations in non-Hodgkin lymphoma dysregulate cell migration by decreasing MEF2B target gene activation. *Nat Commun.* 2015 Aug 6;6:7953. Epub 2015/08/08. doi:10.1038/ncomms8953. Cited in: Pubmed; PMID 26245647.

31. Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, Richards KL, Dunphy CH, Choi WW, Srivastava G, Lugar PL, Rizzieri DA, Lagoo AS, Bernal-Mizrachi L, Mann KP, Flowers CR, Naresh KN, Evens AM, Chadburn A, Gordon LI, Czader MB, Gill JJ, Hsi ED, Greenough A, Moffitt AB, McKinney M, Banerjee A, Grubor V, Levy S, Dunson DB, Dave SS. The genetic landscape of mutations in Burkitt lymphoma. *Nat Genet.* 2012 Dec;44(12):1321-5. Epub 2012/11/13. doi:10.1038/ng.2468. Cited in: Pubmed; PMID 23143597.

32. Dunleavy K, Little RF, Wilson WH. Update on Burkitt Lymphoma. *Hematol Oncol Clin North Am.* 2016 Dec;30(6):1333-1343. Epub 2016/11/28. doi:10.1016/j.hoc.2016.07.009. Cited in: Pubmed; PMID 27888884.
33. Ma MCJ, Tadros S, Bouska A, Heavican T, Yang H, Deng Q, Moore D, Akhter A, Hartert K, Jain N, Showell J, Ghosh S, Street L, Davidson M, Carey C, Tobin J, Perumal D, Vose JM, Lunning MA, Sohani AR, Chen BJ, Buckley S, Nastoupil LJ, Davis RE, Westin JR, Fowler NH, Parekh S, Gandhi M, Neelapu S, Stewart D, Bhalla K, Iqbal J, Greiner T, Rodig SJ, Mansoor A, Green MR. Subtype-specific and co-occurring genetic alterations in B-cell non-Hodgkin lymphoma. *Haematologica.* 2021 Apr 1. Epub 2021/04/02. doi:10.3324/haematol.2020.274258. Cited in: Pubmed; PMID 33792219.
34. Reddy A, Zhang J, Davis NS, Moffitt AB, Love CL, Waldrop A, Leppa S, Pasanen A, Meriranta L, Karjalainen-Lindsberg ML, Norgaard P, Pedersen M, Gang AO, Hogdall E, Heavican TB, Lone W, Iqbal J, Qin Q, Li G, Kim SY, Healy J, Richards KL, Fedoriw Y, Bernal-Mizrachi L, Koff JL, Staton AD, Flowers CR, Paltiel O, Goldschmidt N, Calaminici M, Clear A, Gribben J, Nguyen E, Czader MB, Ondrejka SL, Collie A, Hsi ED, Tse E, Au-Yeung RKH, Kwong YL, Srivastava G, Choi WWL, Evens AM, Pilichowska M, Sengar M, Reddy N, Li S, Chadburn A, Gordon LI, Jaffe ES, Levy S, Rempel R, Tzeng T, Happ LE, Dave T, Rajagopalan D, Datta J, Dunson DB, Dave SS. Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma. *Cell.* 2017 Oct 5;171(2):481-494 e15. Epub 2017/10/07. doi:10.1016/j.cell.2017.09.027. Cited in: Pubmed; PMID 28985567.
35. Meriranta L, Pasanen A, Louhimo R, Cervera A, Alkods A, Autio M, Taskinen M, Rantanen V, Karjalainen-Lindsberg ML, Holte H, Delabie J, Lehtonen R, Hautaniemi S, Leppa S. Deltex-1 mutations predict poor survival in diffuse large B-cell lymphoma. *Haematologica.* 2017 May;102(5):e195-e198. Epub 2017/02/12. doi:10.3324/haematol.2016.157495. Cited in: Pubmed; PMID 28183850.
36. Liu Z, Filip I, Gomez K, Engelbrecht D, Meer S, Lalloo PN, Patel P, Perner Y, Zhao J, Wang J, Pasqualucci L, Rabadan R, Willem P. Genomic characterization of HIV-associated plasmablastic lymphoma identifies pervasive mutations in the JAK-STAT pathway. *Blood Cancer Discov.* 2020 Jul;1(1):112-125. Epub 2020/11/24. doi:10.1158/2643-3249.bcd-20-0051. Cited in: Pubmed; PMID 33225311.

37. Cinar M, Rong HR, Chineke I, Louissaint A, Chapman-Fredricks JR, Vega F, Gunthel CF, Love C, Mosunjac M, Flowers CR, Dave S, Lossos IS, Bernal-Mizrachi L. Genetic Analysis of Plasmablastic Lymphomas in HIV (+) Patients Reveals Novel Driver Regulators of the Noncanonical NF- κ B Pathway. In: 621. LYMPHOMA—GENETIC/EPIGENETIC BIOLOGY. 60th American Society of Hematology (ASH) Annual Meeting; 2018 NOVEMBER 29, 2018; San Diego. American Society of Hematology; 2018. Available from: <https://ashpublications.org/blood/article/132/Supplement%201/1565/273130/Genetic-Analysis-of-Plasmablastic-Lymphomas-in-HIV>.
38. Ma MCJ, Tadros S, Bouska A, Heavican T, Yang H, Deng Q, Moore D, Akhter A, Hartert K, Jain N, Showell J, Ghosh S, Street L, Davidson M, Carey C, Tobin J, Perumal D, Vose JM, Lunning MA, Sohani AR, Chen BJ, Buckley S, Nastoupil LJ, Davis RE, Westin JR, Fowler NH, Parekh S, Gandhi M, Neelapu S, Stewart D, Bhalla K, Iqbal J, Greiner T, Rodig SJ, Mansoor A, Green MR. Subtype-specific and co-occurring genetic alterations in B-cell non-Hodgkin lymphoma. bioRxiv. 2021:674259. doi:10.1101/674259.
39. Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, Wright G, Shaffer AL, Hodson DJ, Buras E, Liu X, Powell J, Yang Y, Xu W, Zhao H, Kohlhammer H, Rosenwald A, Kluin P, Muller-Hermelink HK, Ott G, Gascoyne RD, Connors JM, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Olgwang MD, Reynolds SJ, Fisher RI, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Pittaluga S, Wilson W, Waldmann TA, Rowe M, Mbulaiteye SM, Rickinson AB, Staudt LM. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature*. 2012 Oct 4;490(7418):116-20. Epub 2012/08/14. doi:10.1038/nature11378. Cited in: Pubmed; PMID 22885699.
40. Landau DA, Tausch E, Taylor-Weiner AN, Stewart C, Reiter JG, Bahlo J, Kluth S, Bozic I, Lawrence M, Bottcher S, Carter SL, Cibulskis K, Mertens D, Sougnez CL, Rosenberg M, Hess JM, Edelmann J, Kless S, Kneba M, Ritgen M, Fink A, Fischer K, Gabriel S, Lander ES, Nowak MA, Dohner H, Hallek M, Neuberg D, Getz G, Stilgenbauer S, Wu CJ. Mutations driving CLL and their evolution in progression and relapse. *Nature*. 2015 Oct 22;526(7574):525-30. Epub 2015/10/16. doi:10.1038/nature15395. Cited in: Pubmed; PMID 26466571.

702

703 41. Gonzalez-Rincon J, Mendez M, Gomez S, Garcia JF, Martin P, Bellas C, Pedrosa L,
704 Rodriguez-Pinilla SM, Camacho FI, Quero C, Perez-Callejo D, Rueda A, Llanos M, Gomez-
705 Codina J, Piris MA, Montes-Moreno S, Barcena C, Rodriguez-Abreu D, Menarguez J, de la
706 Cruz-Merino L, Monsalvo S, Parejo C, Royuela A, Kwee I, Cascione L, Arribas A, Bertoni F,
707 Mollejo M, Provencio M, Sanchez-Beato M. Unraveling transformation of follicular
708 lymphoma to diffuse large B-cell lymphoma. PLoS One. 2019;14(2):e0212813. Epub
709 2019/02/26. doi:10.1371/journal.pone.0212813. Cited in: Pubmed; PMID 30802265.

710

711 42. Pedrosa L, Fernandez-Miranda I, Perez-Callejo D, Quero C, Rodriguez M, Martin-Acosta
712 P, Gomez S, Gonzalez-Rincon J, Santos A, Tarin C, Garcia JF, Garcia-Arroyo FR, Rueda A,
713 Camacho FI, Garcia-Cosio M, Heredero A, Llanos M, Mollejo M, Piris-Villaespesa M,
714 Gomez-Codina J, Yanguas-Casas N, Sanchez A, Piris MA, Provencio M, Sanchez-Beato M.
715 Proposal and validation of a method to classify genetic subtypes of diffuse large B cell
716 lymphoma. Sci Rep. 2021 Jan 21;11(1):1886. Epub 2021/01/23. doi:10.1038/s41598-020-
717 80376-0. Cited in: Pubmed; PMID 33479306.

718

719 43. Clipson A, Barrans S, Zeng N, Crouch S, Grigoropoulos NF, Liu H, Kocialkowski S,
720 Wang M, Huang Y, Worrillow L, Goodlad J, Buxton J, Neat M, Fields P, Wilkins B, Grant
721 JW, Wright P, Ei-Daly H, Follows GA, Roman E, Watkins AJ, Johnson PW, Jack A, Du MQ.
722 The prognosis of MYC translocation positive diffuse large B-cell lymphoma depends on the
723 second hit. J Pathol Clin Res. 2015 Jul;1(3):125-133. Epub 2016/06/28. doi:10.1002/cjp2.10.
724 Cited in: Pubmed; PMID 27347428.

725

726 44. Ye H, Remstein ED, Bacon CM, Nicholson AG, Dogan A, Du MQ. Chromosomal
727 translocations involving BCL6 in MALT lymphoma. Haematologica. 2008 Jan;93(1):145-6.
728 eng. Epub 2008/01/02. doi:10.3324/haematol.11927. Cited in: Pubmed; PMID 18166802.

729

730 45. Nann D, Ramis-Zaldivar JE, Müller I, Gonzalez-Farre B, Schmidt J, Egan C, Salmeron-
731 Villalobos J, Clot G, Mattern S, Otto F, Mankel B, Colomer D, Balagué O, Szablewski V,
732 Lome-Maldonado C, Leoncini L, Dojcinov S, Chott A, Copie-Bergman C, Bonzheim I, Fend
733 F, Jaffe ES, Campo E, Salaverria I, Quintanilla-Martinez L. Follicular lymphoma t(14;18)-
734 negative is genetically a heterogeneous disease. Blood Adv. 2020 Nov 24;4(22):5652-5665.

eng. Epub 2020/11/20. doi:10.1182/bloodadvances.2020002944. Cited in: Pubmed; PMID 33211828.

46. Kridel R, Chan FC, Mottok A, Boyle M, Farinha P, Tan K, Meissner B, Bashashati A, McPherson A, Roth A, Shumansky K, Yap D, Ben-Neriah S, Rosner J, Smith MA, Nielsen C, Giné E, Telenius A, Ennishi D, Mungall A, Moore R, Morin RD, Johnson NA, Sehn LH, Tousseyn T, Dogan A, Connors JM, Scott DW, Steidl C, Marra MA, Gascoyne RD, Shah SP. Histological Transformation and Progression in Follicular Lymphoma: A Clonal Evolution Study. PLoS Med. 2016 Dec;13(12):e1002197. eng. Epub 2016/12/14. doi:10.1371/journal.pmed.1002197. Cited in: Pubmed; PMID 27959929.

Table 1. Clinical characteristics of the study group.

Characteristics	HGBL (n = 47)	DHL- <i>BCL2</i> (n = 21)	DHL- <i>BCL6</i> (n = 17)	THL (n = 9)
Insufficient FU	9 (19.1%)	4 (19.0%)	2 (11.8%)	3 (33.3%)
Age (yrs.; median + range)	71 (35 – 89)	73 (35 – 88)	72 (35 – 89)	66 (42 – 76)
Sex				
Female	22 (46.8%)	8 (38.1%)	11 (64.7%)	3 (33.3%)

Male	25 (53.2%)	13 (61.9%)	6 (35.3%)	6 (66.7%)
R-IPI				
0	1 (2.6%)	1 (5.9%)	-	-
1-2	13 (34.2%)	5 (29.4%)	6 (40.0%)	2 (33.3%)
>2	24 (63.2%)	11 (64.7%)	9 (60.0%)	4 (66.7%)
Stage (Ann Arbor)				
I	5 (13.2%)	2 (11.8%)	3 (20.0%)	-
II	9 (23.7%)	4 (23.5%)	3 (20.0%)	2 (33.3%)
III	5 (13.2%)	1 (5.9%)	2 (13.3%)	2 (33.3%)
IV	19 (50.0%)	10 (58.8%)	7 (46.7%)	2 (33.3%)
B-Symptoms				
Yes	20 (52.6%)	10 (58.8%)	7 (46.7%)	3 (50.0%)
No	18 (47.4%)	7 (41.2%)	8 (53.3%)	3 (50.0%)
Extranodal sites				
0	9 (23.7%)	3 (17.6%)	4 (26.7%)	2 (33.3%)
1-2	28 (73.7%)	14 (82.4%)	10 (66.7%)	4 (66.7%)
>2	1 (2.6%)	-	1 (6.7%)	-
ECOG PS				
0-1	20 (52.6%)	8 (47.1%)	8 (53.3%)	4 (66.7%)
≥2	18 (47.4%)	9 (52.9%)	7 (46.7%)	2 (33.3%)
LDH				
Normal	7 (18.4%)	3 (17.6%)	3 (20.0%)	1 (16.7%)
Elevated	31 (81.6%)	14 (82.4%)	12 (80.0%)	5 (83.3%)
CNS involvement at diagnosis				
Yes	2 (5.3%)	-	2 (13.3%)	-
No	36 (94.7%)	17 (100.0%)	13 (86.7%)	6 (100.0%)
Morphology				
DLBCL-like	32 (68.1%)	17 (80.9%)	12 (70.6%)	3 (33.3%)
Burkitt-like	15 (31.9%)	4 (19.0%)	5 (29.4%)	6 (66.7%)
Frontline therapy regimen				
CHOP-like	25 (65.8%)	11 (64.7%)	10 (66.7%)	4 (66.7%)
R-based	31 (81.6%)	14 (82.4%)	12 (80.0%)	5 (83.3%)
Intensified*	13 (34.2%)	7 (41.1%)	3 (20.0%)	3 (50.0%)
Less intensive**	9 (23.7%)	4 (23.5%)	4 (26.7%)	1 (16.7%)
Refusal	1 (2.6%)	-	1 (6.7%)	-
CHOP, cyclophosphamide/hydroxycarbonyl/vincristine/prednisone; CNS, central nervous system; DHL, double-hit lymphoma; DLBCL, diffuse large B-Cell Lymphoma; ECOG, Eastern cooperative oncology group; FU, follow-up; HGBL, high grade B-cell lymphoma; LDH, Lactate dehydrogenase; PS, performance status; R, rituximab; R-IPI, revised International Prognostic Index; THL, triple-hit lymphoma; Yrs., years.				
*Intensified regimens: B-ALL, GMALL, CHOEP (additional etoposide), EPOCH				
**Less intensive regimens: Bendamustin, mini-CHOP (50% dose reduction), rituximab mono				

Figure Legends:

Figure 1. Panel (A) shows the number of variants stratified by variant classification while panel (B) delineates the number of mutations per sample with a median of 153 mutations per sample.

764

765 **Figure 2.** Co-oncoplot for genes identified as significant driver genes by MUTSIGCV ($p <$
766 0.001 ; $n = 22$) in our cohort stratified by cytogenetical subtypes is shown in Panel (A);
767 different types of mutations are colour coded. Additionally, covariates are shown below the
768 plot for each sample. (B) Location of SCNVs along the genome (red bars denote gains; blue
769 bars denote losses; gene names refer to affected oncogenes and genes related to oncogenetic
770 pathways). Panel (C) displays selected genes (oncogenes and genes implicated in oncogenetic
771 pathways) affected by SCNVs.

772

773 **Figure 3.** (A) Significance levels for all HGBL-DH/TH MUTSIGCV genes ($p < 0.001$; gene
774 names in orange indicate $q < 0.1$) and (B) HALLMARK gene sets and NF κ B pathway
775 enrichment for network propagation analysis of significant MUTSIGCV genes (MUTSIGCV p
776 < 0.001). (C) Significant levels for *MYC/BCL2* subgroup (MUTSIGCV $p < 0.001$) and (D)
777 results for enrichment for network propagation analysis against HALLMARK gene sets
778 including NF κ B pathway. Panels (E) and (F) show enrichment results for cytogenetical
779 subgroups *MYC/BCL6* and triple hit, respectively (*MYC/BCL6* genes included: CDKN2A,
780 CD78B, LNP1, KRTAP13-1, UBE2A, CCND3; triple hit genes included: CDKN2A, LNP1,
781 CCND3).

782 UpSet plot (G) showing the overlap of MUTSIGCV genes using our HGBL-DH/TH (D/THL)
783 cohort, the cytogenetical subgroups (BCL2, BCL6, THL), as well as cohorts of ABC-type
784 DLBCL ($n = 67(33)$), GCB-type DLBCL ($n = 45(22)$), BL ($n = 108(38)$) and FL ($n =$
785 $199(38)$) (all retrieved via cBioPortal); (H) shows the overlap between the five lymphoma
786 subtypes and the three cytogenetical subtypes of HGBL-DH/TH.

787

788 **Figure 4.** Allocation of HGBL-DH/TH samples unto the molecular subgroups/clusters of
789 DLBCL, according to LymphGen based on their mutational signature. Additionally,

covariates are shown above the plot for each sample. Row names refer to chromosomal alterations, genes and fusions and are background coloured by their specific subtype (blue = BN2, orange = EZB).

Figure 5. Overall (A) and progression-free survival (B) according to *ARID1A* mutational status. Numbers at risk alongside hazard ratios and p-values according to log-rank testing are provided. Subsequent Bonferroni correction for multiple testing (all genes with a mutational frequency > 15% were investigated) identified a significant impact of *ARID1A* mutations regarding OS while its primarily significant effect on PFS was reduced to a trend of borderline statistical significance.

Figure 1A **Figure 1B**

Figure 2A

Figure 2B

Figure 2C

Figure 3A

Figure 3B

Figure 3C

Figure 3D

Figure 3E

Figure 3F

Figure 3G

Figure 3H

MYC/BCL2/BCL6

LymphGen Subtype

Prior Fol. Lymphoma

Morphology

Gender

Matched normal tissue

Number Mutations

Number Clones

MYC/BCL2/BCL6

- MYC/BCL2
- MYC/BCL6
- Triple hit

LymphGen Subtype

- BN2
- EZB
- Other

Prior Fol. Lymphoma

- na
- no
- yes

Morphology

- Burkitt-like
- DLBCL

Gender

- female
- male

Matched normal tissue

- no
- yes

Number Mutations

- 800
- 600
- 400
- 200
- 0

Number Clones

- 6
- 4
- 2
- 0

Alterations

- Mutation
- Truncation
- Amplification
- Translocation

BCL2 Fusion

BCL6 Fusion

BCL2

EZH2

TNFRSF14

KMT2D

CREBBP

REL

FAS

IRF8

EP300

Chrom. 12p

MEF2B

CIITA

ARID1A

GNA13

STAT6

PTEN

Chrom. 21

EBF1

GNAI2

C10orf12

BCL7A

HLA-DMB

S1PR2

MAP2K1

FBXO11

MIR17HG

BCL6

NOTCH2

TNFAIP3

DTX1

CD70

BCL10

UBE2A

TMEM30A

KLF2

SPEN

CCND3

NOL9

TP63

ETS1

HIST1H1D

PRKCB

HIST1H2BK

TRIP12

KLHL21

TRRAP

PABPC1

Figure 5A

Figure 5B