haematologica
Journal of The Ferrata Storti Foundation

Focal structural variants revealed by whole genome sequencing disrupt the histone demethylase KDM4C in B cell lymphomas

by Cristina López, Nikolai Schleussner, Stephan H. Bernhart, Kortine Kleinheinz, Stephanie Sungalee, Henrike L. Sczabiel, Helene Kretzmer, Umut H. Toprak, Selina Glaser, Rabea Wagener, Ole Ammerpohl, Susanne Bens, Maciej Giefing, Juan C. González Sánchez, Gordana Apic, Daniel Hübschmann, Martin Janz, Markus Kreuz, Anja Mottok, Judit M. Müller, Julian Seufert, Steve Hoffmann, Jan O. Korbel, Robert B. Russell , Roland Schüle, Lorenz Trümper, Wolfram Klapper, Bernhard Radwimmer, Peter Lichter, ICGC MMML-Seq Consortium, Ralf Küppers, Matthias Schlesner, Stephan Mathas, and Reiner Siebert

Received: September 13, 2021.
Accepted: March 18, 2022.

Citation: Cristina López, Nikolai Schleussner, Stephan H. Bernhart, Kortine Kleinheinz, Stephanie Sungalee, Henrike L. Sczabiel, Helene Kretzmer, Umut H. Toprak, Selina Glaser, Rabea Wagener, Ole Ammerpohl, Susanne Bens, Maciej Giefing, Juan C. González Sánchez, Gordana Apic, Daniel Hübschmann, Martin Janz, Markus Kreuz, Anja Mottok, Judit M. Müller, Julian Seufert, Steve Hoffmann, Jan O. Korbel, Robert B. Russell , Roland Schüle, Lorenz Trümper, Wolfram Klapper, Bernhard Radwimmer, Peter Lichter, ICGC MMML-Seq Consortium, Ralf Küppers, Matthias Schlesner, Stephan Mathas, and Reiner Siebert. Focal structural variants revealed by whole genome sequencing disrupt the histone demethylase KDM4C in B cell lymphomas. Haematologica. 2022 Apr 28. doi: 10.3324/haematol.2021.280005. [Epub ahead of print]

Focal structural variants revealed by whole genome sequencing disrupt the histone demethylase *KDM4C* in B cell lymphomas

Cristina López[1,2*], Nikolai Schleussner[3,4,*], Stephan H. Bernhart[5,6,7], Kortine Kleinheinz[8], Stephanie Sungalee[9], Henrike L. Sczakiel[3,4], Helene Kretzmer[5,6,7,10], Umut H. Toprak[11,12,13], Selina Glaser[1], Rabea Wagener[1,2], Ole Ammerpohl[1,2], Susanne Bens[1,2], Maciej Giefing[2,14], Juan C. González Sánchez[15], Gordana Apic[15], Daniel Hübschmann[16,17,18], Martin Janz[3,4], Markus Kreuz[19], Anja Mottok[1], Judith M. Müller[20], Julian Seufert[11], Steve Hoffmann[5,6,7,21], Jan O. Korbel[9], Robert B. Russell[15], Roland Schüle[20,22], Lorenz Trümper[23], Wolfram Klapper[24], Bernhard Radlwimmer[25], Peter Lichter[25], ICGC MMML-Seq Consortium, Ralf Küppers[26], Matthias Schlesner[11,27], Stephan Mathas[3,4#], Reiner Siebert[1,2#]

[1]Institute of Human Genetics, Ulm University and Ulm University Medical Center, Ulm, 89081, Germany;
[2]Institute of Human Genetics, Christian-Albrechts-University, Kiel, 24105, Germany;
[3]Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, 13125, Germany,
[4]Hematology, Oncology and Tumor Immunology, Charité - Universitätsmedizin Berlin, Berlin, 12200, Germany, and Experimental and Clinical Research Center, a joint cooperation between the MDC and the Charité, Berlin, 13125, Germany;
[5]Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, 04107, Germany;
[6]Bioinformatics Group, Department of Computer, University of Leipzig, Leipzig, 04107, Germany;
[7]Transcriptome Bioinformatics, LIFE Research Center for Civilization Diseases, University of Leipzig, Leipzig, 04107, Germany;
[8]Department for Bioinformatics and Functional Genomics, Institute of Pharmacy and Molecular Biotechnology and Bioquant, University of Heidelberg, Heidelberg, 69120, Germany;
[9]EMBL Heidelberg, Genome Biology Unit, Heidelberg, 69117, Germany;
[10]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany;
[11]Bioinformatics and Omics Data Analytics (B240), German Cancer Research Center (DKFZ), Heidelberg, 69120, Germany;
[12]Faculty of Biosciences, Heidelberg University, Heidelberg, 69120, Germany;
[13]Hopp-Children's Cancer Center at the NCT Heidelberg (KiTZ), Division of Neuroblastoma Genomics (B087, German Cancer Research Center (DKFZ), Heidelberg, 69120, Germany;
[14]Institute of Human Genetics, Polish Academy of Sciences, Poznan, 60-479, Poland;
[15]BioQuant and Biochemie Zentrum Heidelberg (BZH), Heidelberg University, Heidelberg, 69120, Germany;
[16]German Cancer Consortium (DKTK), Heidelberg, 69120, Germany
[17]Heidelberg Institute of Stem Cell Technology and Experimental Medicine (HI-STEM), Heidelberg, 69120, Germany
[18]Computational Oncology, Molecular Precision Oncology Program, National Center for Tumor Diseases (NCT), German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), Heidelberg, 69120, Germany
[19]Institute for Medical Informatics Statistics and Epidemiology, Leipzig, 04107, Germany;

[20]Klinik für Urologie und Zentrale Klinische Forschung, Klinikum der Albert-Ludwigs-Universität Freiburg, Freiburg, 79104, Germany;
[21]Leibniz Institute on Ageing-Fritz Lipmann Institute (FLI), Computational Biology, Jena, 07745, Germany;
[22]BIOSS Centre of Biological Signalling Studies, Albert-Ludwigs-University Freiburg, Freiburg, 79104, Germany;
[23]Department of Hematology and Oncology, Georg-August-University of Göttingen, Göttingen, 37075, Germany;
[24]Hematopathology Section, Christian-Albrechts-University, Kiel, 24105, Germany;
[25]Division of Molecular Genetics, German Cancer Research Center (DKFZ), 69120, Heidelberg, Germany;
[26]Institute of Cell Biology (Cancer Research), University of Duisburg-Essen, Essen, 45147, Germany, and German Cancer Consortium (DKTK)
[27]Biomedical Informatics, Data Mining and Data Analytics, Augsburg University, Augsburg, 86159, Germany
*contributed equally as co-first author;
# contributed equally as co-senior author.

Correspondence: R.Siebert: reiner.siebert@uni-ulm.de; S.M.: stephan.mathas@charite.de

Data sharing: Sequencing data are available from the European Genome-phenome archive (EGA) (accession number EGAS00001002199)

Acknowledgments

This study has been supported by the German Ministry of Science and Education (BMBF) in the framework of the ICGC MMML-Seq (01KU1002A-J), ICGC DE-Mining (01KU1505G and 01KU1505E). This work was also supported by the BMBF-funded Heidelberg Center for Human Bioinformatics (HD-HuB) within the German Network for Bioinformatics Infrastructure (de.NBI) (#031A537A, #031A537C). We thank the High Throughput Sequencing unit of the DKFZ Genomics and Proteomics Core Facility for providing whole-genome sequencing services. We thank Prof. Elias Campo and Blanca González-Farré for support on IHC experiments, and Prof. Georg Bornkamm for providing pRTS-1. Former grant support of MMML by the Deutsche Krebshilfe (2003-2011) is gratefully acknowledged. C.L. was supported by an Alexander von Humboldt Foundation post-doctoral fellowship. R.Siebert and M.G. received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952304. Former support to M.G. in form of the FEBS long-term fellowship and the “Support for International Mobility of Scientists” fellowship of the Polish Ministry of Science and Higher Education is gratefully acknowledged. The support of the technical staff of the Institutes of Human Genetics in Kiel and Ulm, as well as, the former Master student Ivan Potreba are gratefully acknowledged.

Author contributions

C.L. performed PCR and Sanger sequencing validations. C.L. S.B. and R.Siebert performed FISH validation analyses. N.S., M.J. and S.M. realized the functional analyses. S.S., U.H.T., J.S., K.K., S.H.B., D.H., M.K., and M.S. performed analysis of next-generation sequencing data. J.M.M. and R.Schüle developed the KDM4C antibody. M.G. provided the genomic data of cHL cell lines. M.S, S.H. and J.O.K. supervised next-generation sequencing analysis and interpreted data. LR.B.R. and J.C.G. generated the protein modelling. A.M. and W.K. performed IHC analysis. H.E., S.G, O.A, B.R, and P.L., analysed the WGBS and methylation arrays. H.L.S. and M.J. provided KDM4C-knockout data of cHL cell lines. N.S. and H.L.S. cloned KDM4C expression vectors. R.K. provided normal B-cell samples. C.L., R.Siebert and S.M. interpreted data and wrote the manuscript. R.Siebert and S.M. designed the study. R.W. and C.L. supported coordination of the project. L.T. performed the clinical coordination of the project. R.Siebert coordinated the ICGC MMML-Seq network. All authors read and approved the final manuscript.

Disclosure of Conflicts of Interest

The authors declare no conflict of interest.

Abstract

Histone methylation-modifiers, like *EZH2* and *KMT2D*, are recurrently altered in B-cell lymphomas. To comprehensively describe the landscape of alterations affecting genes encoding histone methylation-modifiers in lymphomagenesis we investigated whole genome and transcriptome data of 186 mature B-cell lymphomas sequenced in the ICGC MMML-Seq project. Besides confirming common alterations of *KMT2D* (47% of cases), *EZH2* (17%), *SETD1B* (5%), *PRDM9* (4%), *KMT2C* (4%), and *SETD2* (4%) also identified by prior exome or RNAseq studies, we here unravel *KDM4C* in chromosome 9p24, encoding a histone demethylase, to be recurrently altered. Focal structural variation was the main mechanism of *KDM4C* alterations, which was independent from 9p24 amplification. We identified *KDM4C* alterations also in lymphoma cell lines including a focal homozygous deletion in a classical Hodgkin lymphoma cell line. By integrating RNAseq and genome sequencing data we predict *KDM4C* structural variants to result in loss-of-function. By functional reconstitution studies in cell lines, we provide evidence that KDM4C can act as tumor suppressor. Thus, we show that identification of structural variants in whole genome sequencing data adds to the comprehensive description of the mutational landscape of lymphomas and, moreover, establish *KDM4C* as putative tumor suppressive gene recurrently altered in subsets of B-cell derived lymphomas.

Introduction

The majority of mature B-cell malignancies originates from the germinal center (GC) B-cell or the post-GC stage[1]. Historically, Hodgkin lymphoma (HL) and non-Hodgkin lymphomas (NHL) are distinguished. The neoplastic Hodgkin-/Reed-Sternberg cells in classical HL (cHL) are supposed to be derived from pre-apoptotic GC B-cells[1]. The most prevalent types of GC-derived NHL are diffuse large B-cell lymphoma (DLBCL) and follicular lymphoma (FL)[2]. DLBCL is an aggressive disease, composed of various subtypes including the gene expression-based subgroups GC B-cell like (GCB) and activated B-cell like (ABC) types[2,3], or recently genomically defined groups[4–10]. FL is a more indolent disease which occasionally transforms into DLBCL[11].

The immanent genomic instability of B-cells during the GC reaction, which is required for the formation of antibody diversity, is assumed to be causative for malignant transformation of GC or post-GC B-cells[12]. The GC reaction requires a tightly controlled balance between proliferation and growth arrest, and for full cellular activation or a rather resting stage. Epigenetic modifiers are key regulators of these "on-off" stages[12]. In line, they are common targets of genomic alterations in GC-derived B-cell lymphomas[11], including the genes encoding the histone methyltransferases KMT2D, KDMT2C and EZH2, the histone acetyltransferases CREBBP and EP300[6,13–15], or the chromatin remodelers including members of the SWI/SNF complex[16].

A series of recent studies investigated genomic alteration frequencies in oncogenic drivers including epigenetic modifiers in huge series of patients with DLBCL, but also FL. These studies, investigating in total more than 1,800 DLBCL (Chapuy *et al*., n=304[5]; Schmitz *et al*., n=574[6]; Reddy *et al*., n= 1,001[14]), reported aberrations in *KMT2D* (24-31%), *CREBBP* (11-17%), and *EP300* (6-8%) [5,6,14]. Consequently, alterations in these histone modifiers contribute to the definition of genetic subgroups of DLBCL, like cluster 3 in Chapuy *et al*.[5], or the EZB group in Schmitz *et al*.[6] and Wright *et al*.[4], the BCL2 group in Lacy et al.[8] and the EZH2 group in Hübschmann et al[7]. Most of these studies characterized the genomic alteration landscape predominantly by exome or otherwise targeted sequencing, in part combined with transcriptome sequencing[10]. Methodically, however, although this approach detects single nucleotide variants (SNVs), small insertions and deletions (indels) and gross imbalances, it has an inherent weakness in the detection of some structural variants (SVs) like intragenic deletions, chromosomal translocations, and inversions[17].

To overcome this shortcoming, we here mined data from whole genome sequencing (WGS) of 186 GC-derived B-cell lymphomas (Supplementary Information) and 183 corresponding

germline DNA samples generated by us in the framework of the International Cancer Genome Consortium (ICGC MMML-Seq, https://dcc.icgc.org)[7,18–20]. Given the increasing pathogenetic, diagnostic and therapeutic importance of altered histone methylation in B-cell lymphomas[11,13,21], we focused here on 79 genes encoding histone methylation modifiers.

Methods

Whole genome and transcriptome sequencing data

We mined WGS data and available RNAseq data of 186 GC-derived B-cell lymphomas from the ICGC MMML-Seq network[7] (Supplemental Methods). The ICGC MMML-Seq study has been approved by the Institutional Review Board of the Medical Faculty of the University of Kiel (A150/10) and Ulm (349/11), and of the recruiting centers. Methods and procedures used by the ICGC MMML-Seq have been detailed in various publications[7,18,20,22,23] of the network. Sequencing data are available from the European Genome-phenome archive (EGA) (accession number EGAS00001002199). The genomic status of 79 genes encoding histone methylation modifiers (selected based on, http://crdd.osdd.net/raghava/dbem/index.php, accessed 01/03/2018; Online Supplementary Table S1) was investigated.

Cell lines and cell line data

Nineteen B- and T-cell NHL and 4 cHL-derived cell lines were used in the study (Online Supplementary Table S2). The identity of the cell lines used was confirmed by STR analysis using the StemElite ID System (Promega). Copy number data[24] and/or exome data[25] from previously published studies from the 4 cHL cell lines herein analyzed and 2 additional cHL cell lines were included as well as previously reported WGS data from the cHL cell line L1236[26].

Bioinformatics analyses

The computational approaches for the analysis of WGS and transcriptome data were recently described[20] (Online Supplementary Material and Methods). Briefly, WGS data was analyzed using the DKFZ core variant calling workflows of the ICGC PCAWG project (Allele-specific copy-number alterations were analyzed using ACE-seq[20]and SVs were called using the SOPHIA algorithm[20] and DELLY v0.5.9[27,28]. To determine the incidence of SVs in Pan-cancer Analysis of Whole Genomes (PCWAG), filtered structural variant calls were generated by SOPHIA for the

PCAWG cohorts processed with the same tools and settings as with the lymphoma cohort used in this study.

Transcriptome data were mapped with segemehl 0.2.0[29]. Gene expression values were counted using RNAcounter 1.5.2, using the "--nh" option and counting only exonic reads (-t exon).

Mechismo (http://mechismo.russelllab.org/) was used to predict the potential effect of SVs and SNVs detected by WGS.

Verification of KDM4C alterations by PCR-based Sanger sequencing and FISH

Verification analyses on DNA level included PCR-amplification with subsequent Sanger sequencing and fluorescent *in situ* hybridization (FISH) (Online Supplementary Materials and Methods). Alternative *KDM4C* fusion transcripts were validated using specific primers to amplify breakpoint fusion sequences from tumor RNA derived cDNA (Online Supplementary Materials and Methods) and Sanger sequencing. Verifications using FISH were done using two home-made FISH-probes, i.e. using locus-specific and break-apart *KDM4C* probes (Online Supplementary Material and Methods). Digital image acquisition, processing, and evaluation of FISH assays were performed using ISIS digital image analysis version 5.0 (MetaSystems, Altussheim, Germany). The same FISH approaches were used to evaluate the genomic status of *KDM4C* in lymphoma cell lines.

Functional analyses

For functional analyses, the respective cells were transfected with doxycycline (Dox)-inducible KDM4C expression constructs or transduced with KDM4C-encoding lentiviruses (Online Supplementary Materials and Methods. For generation of KDM4C-inducible cells, cells were electroporated in OPTI-MEM I using Gene-Pulser II (Bio-Rad). Twenty-four hours after transfection, Hygromycin B (Sigma-Aldrich, Taufkirchen, Germany) was added. After 21 – 28 days of culture in the presence of Hygromycin B, cells were suitable for functional assays. Where indicated, GFP^+ cells were enriched 72 hours after Dox-induction using a FACS Aria. Production of lentiviruses and lentiviral transduction of cells was performed as described[31] (Online Supplementary Materials and Methods). KDM4C protein expression by Western blot

and immunohistochemistry was assessed using a rabbit polyclonal anti-KDM4C antibody raised against amino acids 1007-1056.

Results

Aberrations in genes encoding histone methylation modifiers in the IGCG MMML-Seq cohort determined by WGS

We analyzed WGS data of 186 GC-derived B-cell lymphomas[7] for somatic aberrations potentially perturbing gene function, including SNVs, indels, SVs and focal copy number aberrations (CNAs) affecting at least two cases in 79 genes encoding histone methylation modifiers. The genes most commonly affected were *KMT2D* (47% of cases), *EZH2* (17%), *SETD1B* (5%), *KDM4C* (4%), *PRDM9* (4%), *KMT2C* (4%), *and SETD2* (4%) (Figure 1A).

The genes identified and their frequency of alteration in our cohort are grossly in line with previous analyses of DLBCL cohorts[5,6,14]. However, alterations of *KDM4C* have not been emphasized as recurrent finding in previous whole exome studies of GC-B-cell lymphomas. *KDM4C* (also called *JMJD2C*) encodes a member of the Jumonji family of demethylases, which activate genes by removing methyl groups from histones H3K9 and H3K36[32,33].

Genomic alterations affecting the KDM4C gene locus

We detected aberrations of *KDM4C* in 7/186 (3.7%) of all cases, and in 4/75 (5.3%) of bona fide DLBCL. In detail, we detected a total of 7 focal aberrations in *KDM4C*, including 4 heterozygous deletions, 1 duplication, 1 translocation and 1 nonsynonymous SNV each, affecting 7 patients (Figure 1B; Online Supplementary Table S2). The minimum read number supporting *KDM4C* alterations was 4 reads. The cancer cell fraction or the proportion of cancer cells with a *KDM4C* alteration among all cancer cells suggest that these alterations are clonal in GC-B-cell lymphomas (Online Supplementary Table S2). We verified all *KDM4C* aberrations by FISH using homemade FISH assays and/or PCR and Sanger sequencing (Online Supplementary Table S2).

Pathologic and genetic features of the lymphomas with KDM4C gene alterations

The cases displaying *KDM4C* alterations included 4/75 DLBCL (5.3%), 1/17 FL-DLBCL (5.9%), 1/4 (25%) large B-cell lymphoma (LBCL) with *IRF4*-rearrangement, and 1/1 primary mediastinal B cell lymphoma (PMBL) (Table 1). We explored whether cases with *KDM4C*-aberrations diagnosed as DLBCL (n=4) cluster in a specific genomic subgroup according to NMF clustering described previously by our group[7], which resembles features of genetic DLBCL subgroups previously also described by others[4,5] (Table 1). Using the described 4-cluster classifier limited to DLBCL cases, we identified 2 MYD88-like and 2 TP53-like cases. Using the 9-subclusters classifier established for the whole ICGC MMML-Seq FL and DLBCL cohort, we assigned 2 cases to the PIM1-like, 2 cases to the PAX5-like and one case to the MYD88-like cluster (2 cases other than DLBCL or FL not included in the previous NMF analyses, i.e. one PMBL and one *IRF4* rearrangement-positive LCL). Furthermore, we explored the distribution of *KDM4C*-altered cases diagnosed as DLBCL (n=4) among transcriptional subgroups based on the cell-of-origin signature. Two cases were classified as GCB-like and two cases as ABC-like lymphomas. Thereafter, we extended the analysis to the whole cohort and we observed that 4/7 cases with *KDM4C* aberrations displayed a GCB signature and 2/7 cases an ABC signature (1 case assigned to Type III/unclassified). Overall, *KDM4C* aberrations were not significantly enriched in any specific genetic nor transcriptional subtype of the analyzed lymphomas.

In addition, we explored if *KDM4C* aberrations were associated with 9p24.1 amplification as described previously[34]. In 5 out of 7 cases with *KDM4C* focal aberrations, we did not detect 9p24.1 gain (Online Supplementary Figure S1 and Online Supplementary Figure S2). This indicates that the vast majority of focal *KDM4C* aberrations occurred independently from 9p24 amplifications. We also explored aneuploidy of chromosome 9 and the general genomic ploidy of the *KDM4C*-altered cases. We did not observe a significant increase of chromosome 9 gains (Online Supplementary Figure S2) or polyploid genome contents (Online Supplementary Figure S3) in cases with *KDM4C* aberration as compared to cases without such a change. Furthermore, we investigated if the *KDM4C* aberrations occurred preferentially in lymphomas with highly rearranged genomes that prompted to carry these events by chance. We identified a mean of 73.85 SVs per case (range 34-144) in lymphomas with *KDM4C* aberration compared to a mean of 83.24 SVs per case (range 5-1696) in lymphomas lacking *KDM4C* aberrations. The mean number of SVs was not significantly different between both groups of patients (Wilcoxon test, *P*=0.125, Online Supplementary Figure S4).

KDM4C alterations as potential tumor drivers

KDM4C is located in a genomic region early replicating in B-cells (at the border to a late replicating region) based on the ENCODE Repli-seq data from different lymphoblastoid cell lines[35] (Online Supplementary Figure S5). This fact suggests that mutation of the gene is not a passenger effect caused by late replication. Furthermore, we examined if the breakpoints were located in the RGYW/WRCY motif typically associated to the enzyme activation-induced cytidine deaminase (AID)[36] that is active in GC B-cells. However, none of the breakpoints directly hit these motifs. These facts suggest that mutation of the gene in lymphoma is not a bystander effect caused by late replication or aberrant somatic hypermutation.

Next we analyzed the incidence of SVs affecting the *KDM4C* locus in the 45 tumor datasets with WGS available included in the Pan-Cancer Analysis of Whole Genomes (PCWAG)[37]. We identified 14 cohorts with equal or higher incidence of *KDM4C* alterations as reported herein (Online Supplementary Figure S6). Importantly, the cohort with the highest incidence of SV breakpoints affecting *KDM4C* was the DLBCL-US cohort, corroborating our findings and highlighting the relevance of *KDM4C* alterations in this subgroup of B-cell lymphoma. Other cancer types harboring SVs in *KDM4C* were renal, head and neck, ovarian, hepatocarcinoma, esophageal, bladder, prostate, osteosarcoma, and gastric cancer, suggesting that *KDM4C* alterations might not be restricted to lymphomas. A role of *KDM4C* in these tumors has been discussed previously[38–45].

Molecular consequence of KDM4C aberrations

The KDM4C protein consists of an N-terminal catalytic domain (JmjN and JmjC) followed by three zinc-fingers (PHD or C2H2) and two C-terminal Tudor domains[46]. By integrating RNAseq and WGS data in the 6 cases with focal SVs and available transcriptome data (4 focal heterozygous deletions, 1 translocation, and 1 duplication), we detected alternative *KDM4C* transcripts in five of them, which were verified by RT-PCR and sequencing (Online Supplementary Table S2). *In silico* analyses using mechismo (http://mechismo.russelllab.org/) predicted these alternative transcripts to result in altered proteins lacking the catalytic or recognition (epigenetic readers) domains (Figure 2A). *In silico* modelling of the nonsynonymous somatic SNV (c.G80A, p.R27Q) detected in case 4177842, which lies in the JmJN domain, suggests a possible functional consequence on the protein function (Online Supplementary Figure S1).

Lack of epigenetic alterations at the KDM4C locus due to KDM4C mutation

As to the recently proposed role of a circular RNAs (circRNAs) derived from the *KDM4C* locus (circKDM4C) in repression of proliferation and metastasis in breast cancer[47], we also explored expression of circKDM4C in the RNAseq data of our cohort. We identified circKDM4C expression in 3 cases (3/180, 1.7%), but none of them showed an alteration of the *KDM4C* locus (data not shown). In addition, we investigated the epigenetic architecture at the *KDM4C* locus by mining previously published whole genome bisulfite sequencing (WGBS) and array-based DNA methylation, as well as, chromatin state data[22]. The promoter and transcription start site (TSS) region showed strong hypomethylation in all subtypes of lymphomas and B-cell controls (Online Supplementary Figure S7). We did not observe any differential DNA methylation with regard to expression or alteration of the *KDM4C* locus with the notable exception of CpG cg13880654 associated with an intronic enhancer site which was hypermethylated in the majority of lymphoma cell lines other than Burkitt lymphoma cell lines. Together, the pattern of alterations strongly suggests KDM4C protein loss-of-function as common principle of the *KDM4C* gene aberrations (Figure 2A and Online Supplementary Figure S1), indicative of a tumour suppressor function of KDM4C protein.

Transcriptional analyses

Using the RNAseq data, we investigated the KDM4C transcript expression in the seven cases with *KDM4C* alterations compared to 173 GC-B-cell lymphomas without *KDM4C* aberrations. To reduce confounders, we performed these analyses for each lymphoma subtype separately. No statistically significant difference in *KDM4C* transcript expression was observed between samples with altered as compared to wildtype *KDM4C* in FL-DLBCL (P= 0.4706), in DLBCL (P= 0.2095), or on LBCL with IRF4 break (P= 1), though this analysis was clearly limited by the low number of *KDM4C* altered cases per group (Online Supplementary Figure S8).

Next, we performed differential expression analyses of RNAseq data comparing cases with and without *KDM4C* alterations. There were only two morphologic subgroups where a sufficient number of cases with and without *KDM4C* alterations with RNAseq data were available, namely DLBCL and FL-DLBCL. We performed the differential expression analyses in both of these groups separately. None of the previously described target genes of KDM4C (Online Supplementary Table S3) were among the 107 differentially expressed genes between *KDM4C* mutated and wildtype cases in the DLBCL group or the 19 differentially expressed genes in the FL-DLBCL group (Online Supplementary Table S4). Notably, there appears to be a small but

significant difference in gene expression between the *KDM4C* mutated and wildtype cases in the DLBCL group. In contrast, the results in the FL-DLBCL group are in line with common fluctuations that may or may not be due to the *KDM4C* status, as it is only a one against all other comparison, and comparing a random FL-DLBCL case against all others often shows even bigger differences. We analyzed the genes differentially expressed between mutated and unmutated *KDM4C* in DLBCL using string-db.org[48]. While there are more interactions than randomly expected between the 64 proteins known to string-db (14 vs 7, $P<0001$), the only enrichment found was in signal peptide domain from UniProt keywords (25 of 64, FDR 0.0135). Next, we intersected the differential expressed genes based on *KDM4C* mutation status in the DLBCL group with genes differentially expressed between subtypes of DLBCL (ABC, GCB and Type III). However, no significant enrichment in the number of overlapping genes was detected. Finally, we examined the expression levels of H3 in the cases harboring *KDM4C* aberrations as compared to the cases lacking *KDM4C* alterations, but no significantly differential expression on H3 was detected ($P > 0.1$).

Functional analysis of KDM4C

We explored public data and screened a total of 23 lymphoma cell lines for the presence of inactivating *KDM4C* alterations using the same approach employed to validate the SVs described above and combined this with published genomic data from two additional cHL cell lines (Online Supplementary Table S2). We detected *KDM4C* deletions in the mycosis fungoides-derived cell line My-La, and the cHL-derived cell line L1236 (Online Supplementary Table S2). More specifically, in WGS data of L1236[26] we identified a *KDM4C* deletion of approximately 60 kb (chr9: 6,775,810-6,836,328 bp (hg19), comprising exons 2, 3 and 4) at one allele, with an approximately 32 kb internal deletion of the second allele (chr9: 6,795,791-6,828,233 bp (hg19), affecting exons 3 and 4), resulting in homozygous loss of exons 3 and 4 of *KDM4C* (Figure 1B). Whereas the larger heterozygous deletion ablates the canonical translation initiation codon, the 32 kb deletion is predicted to encode an (if translated) non-functional protein in L1236 lacking the N-terminus with the catalytic JmjN domain (Figure 2A). The *KDM4C* heterozygous deletion in My-La cells is in agreement with the conventional cytogenetic analysis describing a del(9)(p21)[49]. In addition, SUP-HD1 and KARPAS-422 cell lines carry SNVs c.G1713A, p.W571*, and c.C2498T, p.P833L, respectively (Figure 2A and Online Supplementary Table S2) reported in the COSMIC cell lines database (cancer.sanger.ac.uk/cell_lines), which we validated by PCR and Sanger sequencing and which

are predicted as pathogenic variants by FATHMM. Evolutionary and protein structure predictions also suggest that p.P833L is a loss-of-function variant (Online Supplementary Figure S2). Thus, remarkably, with L1236 and SUP-HD1 two out of 6 bona fide cHL cell lines show potential inactivating changes in the *KDM4C* gene suggesting a tumor suppressive role in both, GC-B-cell lymphomas and cHL.

In agreement with the genomic data, KDM4C immunoblotting of the various cell lines revealed a complete loss of KDM4C protein expression in L1236 cells, and a strong reduction in My-La, but also in the cell lines Se-Ax and SU-DHL-1 using a homemade rabbit polyclonal anti-KDM4C (against aminoacids 1007-1056) (Figure 2B). To address the functional consequences of *KDM4C* deletions, we constructed an episomally replicating vector for doxycycline (Dox)-inducible KDM4C re-expression in L1236 cells as well as KDM4C lentiviruses for re-expression in SU-DHL-1 cells (Figure 2C). In both cell lines, KDM4C re-expression resulted in a loss of KDM4C-expressing cells over time, in agreement with a tumor suppressor function in these cells. Neither in KARPAS-422 (p.P833L) nor in NAMALWA (no *KDM4C* aberration), such an effect was observed (Figure 2D).

In addition, we aimed to investigate the expression of KDM4C by immunohistochemistry (IHC) in primary tissues as well as cell lines using the anti-KDM4C antibody used for the immunoblotting technique. We were able to prove the lack of expression of KDM4C protein in L1236 and to specifically detect ectopically expressed KDM4C in formalin fixed and paraffin-embedded transfected HEK293 cells (data not shown). However, this as well as several commercially available anti-KDM4C antibodies failed in our hands to reliably quantify KDM4C protein expression in primary tissues (data not shown).

Discussion

Herein, we analyzed WGS data of 186 GC-B-cell lymphomas to investigate somatic aberrations, including SNVs, indels, SVs, and focal copy number aberrations (CNAs) in 79 genes encoding histone methylation regulators, to identity epigenetic modifiers potentially involved in GC-B-cell lymphomagenesis[6,11–16]. We identified *KDM4C*, encoding a histone demethylase, as recurrently altered in B-cell lymphomas (7/186, 4%).

Integrating RNAseq and WGS data in the 6 cases with focal SVs, we detected alternative *KDM4C* transcripts in five of them. *In silico* analyses using mechismo predicted these alternative transcripts to result in altered proteins lacking the catalytic or recognition

(epigenetic readers) domains suggesting a loss of function. In contrast, *KDM4C* has been previously described as an oncogene in lymphomas, which is activated by large chromosome 9p gains. These gains mostly derive from co-amplification with *JAK2*, *CD274*, and *PDCD1LG2*, recurrently found in cHL and PMBL[34]. Nonetheless, recent studies in these lymphoma subtypes refined the minimally gained region and point to *CD274* and *PDCD1LG2* as main targets, whereas *KDM4C* is not consistently gained (Online Supplementary Figure S9). We here detected *KDM4C* to be altered due to focal SVs (6/7) rather than gross imbalances. Moreover, we observed that the vast majority of focal *KDM4C* aberrations occurred independently from 9p24 amplifications. Overall, *KDM4C* might belong to the increasing list of genes with oncogenic and tumor suppressive function depending on the cellular context and the type of aberration. Examples from hematologic neoplasms include *EZH2* or the CEBP gene family[50–52].

In addition to genomic mutations, we explored different other layers that could contribute to dysregulated *KDM4C* in GC-B-cell lymphomas. We used previously published data derived from WGBS, array-based DNA methylation, and chromatin state data[22] to investigate epigenetic alterations at the *KDM4C* locus. However, our findings do not provide evidence for epigenetic inactivation of the *KDM4C* locus. Furthermore, we investigated the role of expression of cirKDM4C, a circular RNA from the KDM4C locus recently described to be of pathogenic relevance in solid tumors[47,53]. Although we detected circKDM4C expression in 17% of the cases, none of them showed an alteration of the *KDM4C* locus. Taken together, these results suggest that genomic alterations are the main mechanism involved in *KDM4C* gene dysregulation in GC-B-cell lymphomas. The fact that these genomic alterations are mostly focal SV and that SNVs of *KDM4C* (1/7) seem to be rare in GC-B-cell lymphomas likely explains why alterations of this gene have been underestimated in previous whole exome analyses[5,6].

KDM4C alterations are not exclusive to GC-B-cell NHL. Exploring a set of cell lines we observed that L1236[26] and SUP-HD1, and thus two out of 6 (33.3%) bona fide cHL cell lines, show potentially inactivating changes in the *KDM4C* gene suggesting a tumor suppressive role in both, GC-B-cell NHL and cHL. We could not detect changes in KDM4C transcript expression between cases with altered and wildtype KDM4C using RNAseq data of the GCB-derived lymphomas in the ICGC cohort. In contrast, using KDM4C immunoblotting in the cell lines we observed a complete loss of KDM4C protein expression in L1236 cells, and a strong reduction in My-La, which is in line with the genomic analysis. Subsequent functional analyses using KDM4C re-expression in cell lines with reduced or lack of KDM4C expression support the hypothesis of a tumor suppressor function of KDM4C at least in a subset of lymphomas.

In conclusion, our work not only adds *KDM4C* to the list of histone methylation modifiers recurrently altered in B-cell lymphomas, but it also supports a function of *KDM4C* as tumor suppressor at least in a subset of lymphoma types. Moreover, our data demonstrate that focal SVs contribute to the mutational burden of distinct genes, which might be missed by pure exome and/or RNAseq approaches. This has to be considered if mutational landscapes are defined for classification schemes like those proposed for lymphomas.

References

1. Swerdlow SH, Campo E, Harris NL, Jaffe ES, Pileri SA, Stein H TJ. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues. 4th ed.; 2017.

2. Rosenwald A, Wright G, Chan WC, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med. 2002;346(25):1937-1947.

3. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med. 2006;354(23):2419-2430.

4. Wright GW, Huang DW, Phelan JD, et al. A Probabilistic Classification Tool for Genetic Subtypes of Diffuse Large B Cell Lymphoma with Therapeutic Implications. Cancer Cell. 2020;37(4):551-568.e14.

5. Chapuy B, Stewart C, Dunford AJ, et al. Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes. Nat Med. 2018;24(5):679-690.

6. Schmitz R, Wright GW, Huang DW, et al. Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma. N Engl J Med. 2018;378(15):1396-1407.

7. Hübschmann D, Kleinheinz K, Wagener R, et al. Mutational mechanisms shaping the coding and noncoding genome of germinal center derived B-cell lymphomas. Leukemia. 2021;35(7):2002-2016.

8. Lacy SE, Barrans SL, Beer PA, et al. Targeted sequencing in DLBCL, molecular subtypes, and outcomes: a Haematological Malignancy Research Network report. Blood. 2020;135(20):1759-1771.

9. Runge HFP, Lacy S, Barrans S, et al. Application of the LymphGen classification tool to 928 clinically and genetically-characterised cases of diffuse large B cell lymphoma (DLBCL). Br J Haematol. 2021;192(1):216-220.

10. Morin RD, Arthur SE, Hodson DJ. Molecular profiling in diffuse large B-cell lymphoma: why so many types of subtypes? Br J Haematol. 2022;196(4):814-829.

11. Pasqualucci L, Dalla-Favera R. Genetics of diffuse large B-cell lymphoma. Blood. 2018;131(21):2307-2319.

12. De Silva NS, Klein U. Dynamics of B cells in germinal centres. Nat Rev Immunol.

2015;15(3):137-148.

13. Morin RD, Johnson NA, Severson TM, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet. 2010;42(2):181-185.

14. 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.

15. Hashwah H, Schmid CA, Kasser S, et al. Inactivation of CREBBP expands the germinal center B cell compartment, down-regulates MHCII expression and promotes DLBCL growth. Proc Natl Acad Sci U S A. 2017;114(36):9701-9706.

16. Krysiak K, Gomez F, White BS, et al. Recurrent somatic mutations affecting B-cell receptor signaling pathway genes in follicular lymphoma. Blood. 2017;129(4):473-483.

17. Singleton AB. Exome sequencing: a transformative technology. Lancet Neurol. 2011;10(10):942-946.

18. Richter J, Schlesner M, Hoffmann S, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat Genet. 2012;44(12):1316-1320.

19. ICGC/TCGA Pan-Canger Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. Nature. 2020;578(7793):82-93.

20. López C, Kleinheinz K, Aukema SM, et al. Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. Nat Commun. 2019;10(1):1459.

21. Li H, Kaminski MS, Li Y, et al. Mutations in linker histone genes HIST1H1 B, C, D, and E; OCT2 (POU2F2); IRF8; and ARID1A underlying the pathogenesis of follicular lymphoma. Blood. 2014;123(10):1487-1498.

22. Kretzmer H, Bernhart SH, Wang W, et al. DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. Nat Genet. 2015;47(11):1316-1325.

23. Doose G, Haake A, Bernhart SH, et al. MINCR is a MYC-induced lncRNA able to modulate MYC's transcriptional network in Burkitt lymphoma cells. Proc Natl Acad Sci U S A. 2015;112(38):E5261-5270.

24. Otto C, Giefing M, Massow A, et al. Genetic lesions of the TRAF3 and MAP3K14 genes in classical Hodgkin lymphoma. Br J Haematol. 2012;157(6):702-708.

25. Liu Y, Abdul Razak FR, Terpstra M, et al. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. Leukemia. 2014;28(11):2248-2251.

26. Schneider M, Schneider S, Zühlke-Jenisch R, et al. Alterations of the CD58 gene in classical Hodgkin lymphoma. Genes Chromosomes Cancer. 2015;54(10):638-645.

27. Rausch T, Zichner T, Schlattl A, Stütz AM, Benes V, Korbel JO. DELLY: structural variant discovery by integrated paired-end and split-read analysis. Bioinformatics. 2012;28(18):i333-i339.

28. Northcott PA, Buchhalter I, Morrissy AS, et al. The whole-genome landscape of medulloblastoma subtypes. Nature. 2017;547(7663):311-317.

29. Hoffmann S, Otto C, Kurtz S, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. PLoS Comput Biol. 2009;5(9):e1000502.

30. Betts MJ, Lu Q, Jiang Y, et al. Mechismo: predicting the mechanistic impact of mutations and modifications on molecular interactions. Nucleic Acids Res. 2015;43(2):e10.

31. Schleussner N, Merkel O, Costanza M, et al. The AP-1-BATF and -BATF3 module is essential for growth, survival and TH17/ILC3 skewing of anaplastic large cell lymphoma. Leukemia. 2018;32(9):1994-2007.

32. Cloos PAC, Christensen J, Agger K, et al. The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3. Nature. 2006;442(7100):307-311.

33. Loh Y-H, Zhang W, Chen X, George J, Ng H-H. Jmjd1a and Jmjd2c histone H3 Lys 9 demethylases regulate self-renewal in embryonic stem cells. Genes Dev. 2007;21(20):2545-2557.

34. Rui L, Emre NCT, Kruhlak MJ, et al. Cooperative epigenetic modulation by cancer amplicon genes. Cancer Cell. 2010;18(6):590-605.

35. Hansen RS, Thomas S, Sandstrom R, et al. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. Proc Natl Acad Sci U S A. 2010;107(1):139-144.

36. Yu K, Huang FT, Lieber MR. DNA Substrate Length and Surrounding Sequence Affect the

Activation-induced Deaminase Activity at Cytidine *. J Biol Chem. 2004;279(8):6496-6500.

37. Campbell PJ, Getz G, Korbel JO, et al. Pan-cancer analysis of whole genomes. Nature. 2020;578(7793):82-93.

38. Krill-Burger JM, Lyons MA, Kelly LA, et al. Renal cell neoplasms contain shared tumor type-specific copy number variations. Am J Pathol. 2012;180(6):2427-2439.

39. Lee DH, Kim GW, Jeon YH, Yoo J, Lee SW, Kwon SH. Advances in histone demethylase KDM4 as cancer therapeutic targets. FASEB J. 2020;34(3):3461-3484.

40. Shao N, Cheng J, Huang H, et al. GASC1 promotes hepatocellular carcinoma progression by inhibiting the degradation of ROCK2. Cell Death Dis. 2021;12(3):253.

41. Ma X, Ying Y, Sun J, et al. circKDM4C enhances bladder cancer invasion and metastasis through miR-200bc-3p/ZEB1 axis. Cell death Discov. 2021;7(1):365.

42. Chen GQ, Ye P, Ling RS, et al. Histone Demethylase KDM4C Is Required for Ovarian Cancer Stem Cell Maintenance. Stem Cells Int. 2020;2020:8860185.

43. Kleszcz R, Skalski M, Krajka-Kuźniak V, Paluszczak J. The inhibitors of KDM4 and KDM6 histone lysine demethylases enhance the anti-growth effects of erlotinib and HS-173 in head and neck cancer cells. Eur J Pharm Sci. 2021;166:105961.

44. Yang D, Xu T, Fan L, Liu K, Li G. microRNA-216b enhances cisplatin-induced apoptosis in osteosarcoma MG63 and SaOS-2 cells by binding to JMJD2C and regulating the HIF1α/HES1 signaling axis. J Exp Clin Cancer Res. 2020;39(1):201.

45. Lang T, Xu J, Zhou L, et al. Disruption of KDM4C-ALDH1A3 feed-forward loop inhibits stemness, tumorigenesis and chemoresistance of gastric cancer stem cells. Signal Transduct Target Ther. 2021;6(1):336.

46. Huang Y, Fang J, Bedford MT, Zhang Y, Xu R-M. Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science. 2006;312(5774):748-751.

47. Liang Y, Song X, Li Y, et al. circKDM4C suppresses tumor progression and attenuates doxorubicin resistance by regulating miR-548p/PBLD axis in breast cancer. Oncogene. 2019;38(42):6850-6866.

48. Jensen LJ, Kuhn M, Stark M, et al. STRING 8--a global view on proteins and their functional interactions in 630 organisms. Nucleic Acids Res. 2009;37(Database

issue):D412-416.

49. Netchiporouk E, Gantchev J, Tsang M, et al. Analysis of CTCL cell lines reveals important differences between mycosis fungoides/Sézary syndrome vs. HTLV-1(+) leukemic cell lines. Oncotarget. 2017;8(56):95981-95998.

50. Gan L, Yang Y, Li Q, Feng Y, Liu T, Guo W. Epigenetic regulation of cancer progression by EZH2: from biological insights to therapeutic potential. Biomark Res. 2018;6(1):10.

51. Tolomeo M, Grimaudo S. The "Janus" Role of C/EBPs Family Members in Cancer Progression. Int J Mol Sci. 2020;21(12):1-18.

52. Shaulian E. AP-1 - The Jun proteins: Oncogenes or tumor suppressors in disguise? Cell Signal. 2010;22(6):894-899.

53. Ma SD, Xu X, Jones R, et al. PD-1/CTLA-4 Blockade Inhibits Epstein-Barr Virus-Induced Lymphoma Growth in a Cord Blood Humanized-Mouse Model. PLoS Pathog. 2016;12(5):e1005642.

Table 1. Characteristics and biological features of the cohort. Overview on age, sex, and lymphoma subtype distribution of the cohort and of *KDM4C* affected cases

		Whole cohort (n=186)	*KDM4C* focal aberrations (n=7)
Median age at diagnosis, years (range)		62.5 (16-89)	57 (16-75)
Male; n (%)		84 (45)	3 (42.9)
Diagnostics	**DLBCL**	75 (40.3)	4 (57)
	FL	86 (46.2)	0
	FL-DLBCL	17 (9.1)	1 (14.3)
	LBCL-with *IRF4* breaks	4 (2.2)	1 (14.3)
	B-NOS	1 (0.55)	0
	DH-BL	2 (1.1)	0
	PMBL	1 (0.55)	1 (14.3)
COO	**GCB**	124/182 (68.1)	4 (57.1)
	ABC	29/182 (15.9)	2 (28.6)
	Type III	29/182 (15.9)	1 (14.3)
Hallmark events	***MYC* breaks; n (%)**	18/185 (9.7)	0
	***BCL2* breaks; n (%)**	106/185 (5.7)	2 (28.6)
	***BCL6* breaks; n (%)**	51 (27.4)	2 (28.6)
NMF Subclusters (4 categories, only DLBCL)[7]	**BCL2-like; n (%)**	10/72 (13.9)	0/4
	BCL6-like; n (%)	17/72 (23.6)	0/4
	MYD88-like; n (%)	26/72 (36.1)	2/4 (50)
	TP53-like; n (%)	19/72 (26.4)	2/4 (50)
NMF Subclusters (9 categories)[7]	**B2M like; n (%)**	13/179 (7.3)	0/5
	BCL2-like; n (%)	35/179 (19.5)	0/5
	BCL6-like; n (%)	23/179 (12.8)	0/5
	MYD88-like; n (%)	17/179 (9.5)	1/5 (20)
	PAX5-like; n (%)	19/179 (10.6)	2/5 (40)
	PIM1-like; n (%)	18/179 (10)	2/5 (40)
	TP53-like; n (%)	19/179 (10.6)	0/5
	CSMD1-like; n (%)	20/179 (11.2)	0/5
	SOCS1-like; n (%)	15/179 (8.4)	0/5

n, number; DLBCL, diffuse large B-cell lymphoma; FL, follicular lymphoma; LBCL, large B-cell lymphoma; B-NOS, B-cell lymphoma not otherwise specified; PMBL, primary mediastinal B-cell lymphoma; COO, cell of origin based on gene expression classified; GCB, germinal center B-cell like; ABC, activated B-cell like.

Figure legends

Figure 1. Genomic aberrations of genes encoding histone methylation modifiers in GC-derived B-cell lymphomas (A) Genes encoding histone methylation modifiers affected by SNVs, SVs, focal CNAs, and indels in at least two of the studied GC-derived B-cell lymphomas. The events are shown independently of the transcript isoform affected. Genes encoding histone methyltransferases are named in grey, genes encoding demethylases in blue. Categorization according to diagnosis and cell of origin (COO)-classification is shown. The respective type of variant is indicated by color. Only focal aberrations are displayed. (B) UCSC genome browser track (accessed 03.02.2021) displaying the *KDM4C* genomic aberrations detected in primary lymphomas and lymphoma cell lines. The type of variant is indicated by color and each focal aberration in a separate line. Note, the size of the squares showing the SNVs and translocations is not representative of the real size of the genomic aberration to improve the visualization. The heterozygous deletions and the homozygous deletion are labeled in yellow and red, respectively. Duplication is indicated with brown color and the translocation in purple. The breakpoints of the genomic aberrations are annotated based on hg19 and the *KDM4C* transcripts displayed as provided by UCSC Genes track. Note, that the L1236 cell line shows homozygously and heterozygously deleted regions.

Figure 2. Potential tumor suppressive role of the KDM4C protein (A) Predicted KDM4C protein alterations associated with genomic and transcriptomic changes. Top panel: structure of KDM4C protein (Uniprot: Q9H3R0) with highlighted positions of the two N-terminal JmjN and JmjC catalytic domains, the PHD zinc finger domains, and the two C-terminal Tudor recognition domains. Boxes highlight the predicted residual protein structure (if translated) in lymphoma cases and cell lines with somatic KDM4C variants. In case 4177842, the SNV affects the JmjN domain. In cases 4191799 and 4110120 the regulatory/recognition region, in cases 41074893, 4128852 and 4115001 the catalytic domains are affected. The predicted protein in KARPAS-422 affects the regulatory region; in L1236 the predicted protein abrogates the catalytic domain JmjC, and the regulatory/recognition region, and consequently generates a non-functional protein; in SUP-HD1 the regulatory/recognition domains are predicted to be lost by the truncating SNV. (B) Immunoblotting of KDM4C protein. Left column, upper three panels, KDM4C protein expression in whole cell extracts of various human leukemia/lymphoma-derived cell lines, as indicated. β-actin or α-tubulin are shown as controls. Note, that L1236 cells completely lack KDM4C protein expression, and KDM4C expression in

SU-DHL-1, My-La and Se-Ax cells is strongly reduced. Upper right panel, KDM4C immunoblotting of L1236 cells before and after transfection with doxycyclin (Dox)-inducible KDM4C-expression vector or respective control (Mock) construct, before (left) and after (right) enrichent of Dox-induced, GFP^+ cells. β-actin is shown as control. Right column, center and bottom, immunoblotting of KDM4C of KARPAS-422 and NAMALWA after transfection with Dox-inducible KDM4C or control (Mock) construct, without and after addition of Dox. β-actin is shown as control. Left column bottom, immunoblotting of KDM4C of SU-DHL-1 cells after lentiviral transduction with KDM4C expression or respective control (Mock) construct. Note that L428, L1236, REH and NAMALWA extracts were included as positive and negative controls. (C, D) Reconstitution of KDM4C results in loss of cells re-expressing KDM4C in L1236 with absent or SU-DHL-1 with reduced endogenous expression of the protein, but not in those with robust endogenous KDM4C expression (KARPAS-422, NAMALWA). (C) Left, L1236 were transfected with Dox-inducible KDM4C-expression or respective control (Mock) construct. GFP-expression is driven in parallel to KDM4C by a bidirectional promoter. GFP-positive cells were enriched by flow cytometry 48 hours after addition of Dox, and, thereafter, the percentage of GFP-positive cells was monitored over time by flow cytometry. The percentage of KDM4C- or Mock-transfected GFP-positive cells normalized to the percentage of GFP^+ cells at day 2 after Dox-induction and enrichment of GFP^+ cells is shown. Right, SU-DHL-1 cells were lentivirally transduced with KDM4C, and the percentage of KDM4C, $EBFP^+$ cells over time was determined as described for L1236 cells. (D) KARPAS-422 (left) and NAMALWA (right) cells were treated and monitored as described in (C) for L1236, with the difference that due to high transfection rates for both cell lines no enrichment of GFP^+ cells was required.

Figure 1

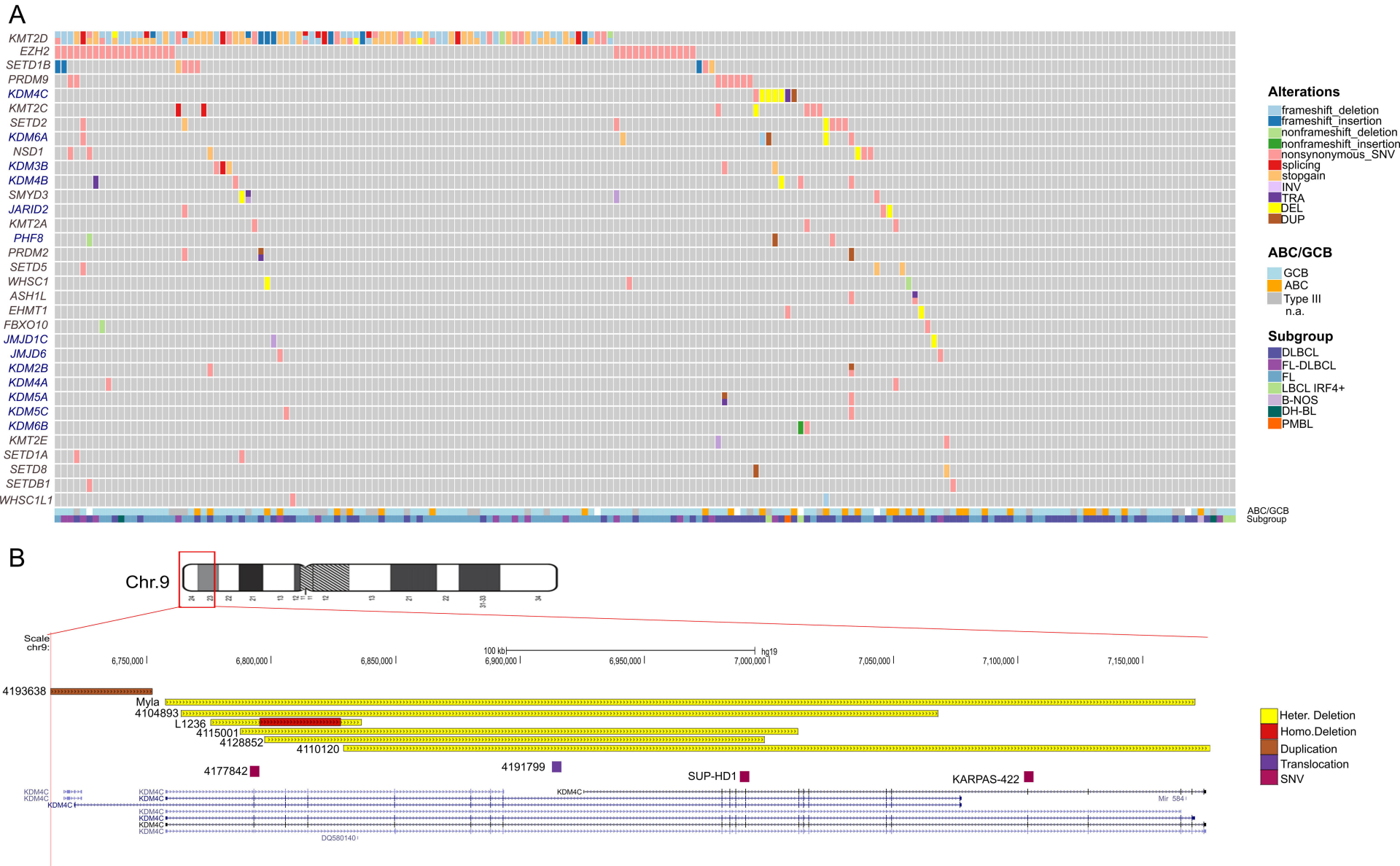
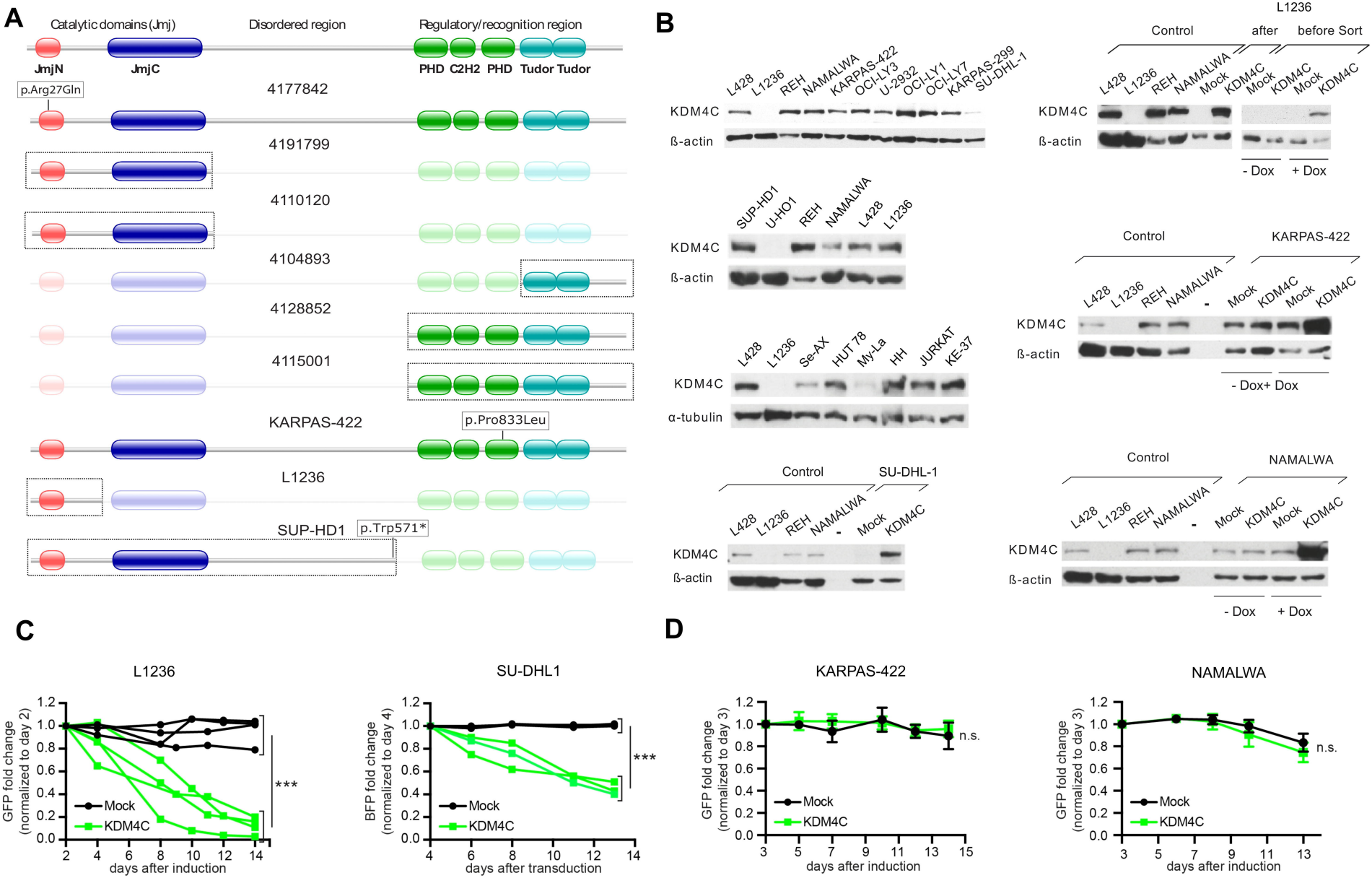

A
KMT2D
EZH2
SETD1B
PRDM9
KDM4C
KMT2C
SETD2
KDM6A
NSD1
KDM3B
KDM4B
SMYD3
JARID2
KMT2A
PHF8
PRDM2
SETD5
WHSC1
ASH1L
EHMT1
FBXO10
JMJD1C
JMJD6
KDM2B
KDM4A
KDM5A
KDM5C
KDM6B
KMT2E
SETD1A
SETD8
SETDB1
WHSC1L1
ABC/GCB
Subgroup
Alterations
frameshift_deletion
frameshift_insertion
nonframeshift_deletion
nonframeshift_insertion
nonsynonymous_SNV
splicing
stopgain
INV
TRA
DEL
DUP
ABC/GCB
GCB
ABC
Type III
n.a.
Subgroup
DLBCL
FL-DLBCL
FL
LBCL IRF4+
B-NOS
DH-BL
PMBL
B
Chr.9
Scale
chr9:
100 kb
hg19
6,750,000
6,800,000
6,850,000
6,900,000
6,950,000
7,000,000
7,050,000
7,100,000
7,150,000
4193638
Myla
4104893
L1236
4115001
4128852
4110120
4177842
4191799
SUP-HD1
KARPAS-422
KDM4C
Mir 584
DQ580140
Heter. Deletion
Homo.Deletion
Duplication
Translocation
SNV

Figure 2


A
Catalytic domains (Jmj)
Disordered region
Regulatory/recognition region
JmjN
JmjC
PHD C2H2 PHD Tudor Tudor
p.Arg27Gln
4177842
4191799
4110120
4104893
4128852
4115001
KARPAS-422
p.Pro833Leu
L1236
SUP-HD1
p.Trp571*
B
L428 L1236 REH NAMALWA KARPAS-422 OCI-LY3 U-2932 OCI-LY1 OCI-LY7 KARPAS-299 SU-DHL-1
KDM4C
ß-actin
L1236
Control
after
before Sort
L428 L1236 REH NAMALWA Mock KDM4C Mock KDM4C Mock KDM4C
- Dox
+ Dox
SUP-HD1 U-HO1 REH NAMALWA L428 L1236
Control
KARPAS-422
L428 L1236 REH NAMALWA - Mock KDM4C Mock KDM4C
- Dox+ Dox
L428 L1236 Se-AX HUT 78 My-La HH JURKAT KE-37
α-tubulin
Control
SU-DHL-1
L428 L1236 REH NAMALWA - Mock KDM4C
Control
NAMALWA
L428 L1236 REH NAMALWA - Mock KDM4C Mock KDM4C
C
L1236
SU-DHL1
GFP fold change (normalized to day 2)
BFP fold change (normalized to day 4)
Mock
KDM4C
***
days after induction
days after transduction
D
KARPAS-422
NAMALWA
GFP fold change (normalized to day 3)
n.s.
days after induction

Supplementary data

Focal structural variants revealed by whole genome sequencing disrupt the histone demethylase *KDM4C* in B cell lymphomas

Cristina López[1,2*], Nikolai Schleussner[3,4,*], Stephan H. Bernhart[5,6,7], Kortine Kleinheinz[8], Stephanie Sungalee[9], Henrike L. Sczakiel[3,4], Helene Kretzmer[5,6,7,10], Umut H. Toprak[11,12,13], Selina Glaser[1], Rabea Wagener[1,2], Ole Ammerpohl[1,2], Susanne Bens[1,2], Maciej Giefing[2,14], Juan C. González Sánchez[15], Gordana Apic[15], Daniel Hübschmann[16,17,18], Martin Janz[3,4], Markus Kreuz[19], Anja Mottok[1], Judith M. Müller[20], Julian Seufert[11], Steve Hoffmann[5,6,7,21], Jan O. Korbel[9], Robert B. Russell[15], Roland Schüle[20,22], Lorenz Trümper[23], Wolfram Klapper[24], Bernhard Radlwimmer[25], Peter Lichter[25], ICGC MMML-Seq Consortium, Ralf Küppers[26], Matthias Schlesner[11,27], Stephan Mathas[3,4#], Reiner Siebert[1,2#]

[1]Institute of Human Genetics, Ulm University and Ulm University Medical Center, Ulm, 89081, Germany;
[2]Institute of Human Genetics, Christian-Albrechts-University, Kiel, 24105, Germany;
[3]Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association (MDC), Berlin, 13125, Germany,
[4]Hematology, Oncology and Tumor Immunology, Charité - Universitätsmedizin Berlin, Berlin, 12200, Germany, and Experimental and Clinical Research Center, a joint cooperation between the MDC and the Charité, Berlin, 13125, Germany;
[5]Interdisciplinary Center for Bioinformatics, University of Leipzig, Leipzig, 04107, Germany;
[6]Bioinformatics Group, Department of Computer, University of Leipzig, Leipzig, 04107, Germany;
[7]Transcriptome Bioinformatics, LIFE Research Center for Civilization Diseases, University of Leipzig, Leipzig, 04107, Germany;
[8]Department for Bioinformatics and Functional Genomics, Institute of Pharmacy and Molecular Biotechnology and Bioquant, University of Heidelberg, Heidelberg, 69120, Germany;
[9]EMBL Heidelberg, Genome Biology Unit, Heidelberg, 69117, Germany;
[10]Department of Genome Regulation, Max Planck Institute for Molecular Genetics, Berlin, Germany;
[11]Bioinformatics and Omics Data Analytics (B240), German Cancer Research Center (DKFZ), Heidelberg, 69120, Germany;
[12]Faculty of Biosciences, Heidelberg University, Heidelberg, 69120, Germany;
[13]Hopp-Children's Cancer Center at the NCT Heidelberg (KiTZ), Division of Neuroblastoma Genomics (B087, German Cancer Research Center (DKFZ), Heidelberg, 69120, Germany;
[14]Institute of Human Genetics, Polish Academy of Sciences, Poznan, 60-479, Poland;
[15]BioQuant and Biochemie Zentrum Heidelberg (BZH), Heidelberg University, Heidelberg, 69120, Germany;
[16]German Cancer Consortium (DKTK), Heidelberg, 69120, Germany
[17]Heidelberg Institute of Stem Cell Technology and Experimental Medicine (HI-STEM), Heidelberg, 69120, Germany

[18]Computational Oncology, Molecular Precision Oncology Program, National Center for Tumor Diseases (NCT), German Cancer Research Center (DKFZ) and German Cancer Consortium (DKTK), Heidelberg, 69120, Germany
[19]Institute for Medical Informatics Statistics and Epidemiology, Leipzig, 04107, Germany;
[20]Klinik für Urologie und Zentrale Klinische Forschung, Klinikum der Albert-Ludwigs-Universität Freiburg, Freiburg, 79104, Germany;
[21]Leibniz Institute on Ageing-Fritz Lipmann Institute (FLI), Computational Biology, Jena, 07745, Germany;
[22]BIOSS Centre of Biological Signalling Studies, Albert-Ludwigs-University Freiburg, Freiburg, 79104, Germany;
[23]Department of Hematology and Oncology, Georg-August-University of Göttingen, Göttingen, 37075, Germany;
[24]Hematopathology Section, Christian-Albrechts-University, Kiel, 24105, Germany;
[25]Division of Molecular Genetics, German Cancer Research Center (DKFZ), 69120, Heidelberg, Germany;
[26]Institute of Cell Biology (Cancer Research), University of Duisburg-Essen, Essen, 45147, Germany, and German Cancer Consortium (DKTK)
[27]Biomedical Informatics, Data Mining and Data Analytics, Augsburg University, Augsburg, 86159, Germany
*contributed equally as co-first author;
# contributed equally as co-senior author.

Correspondence: R.Siebert: reiner.siebert@uni-ulm.de; S.M.: stephan.mathas@charite.de

Supplementary Material and Methods

ICGC MMML-seq cohort

The ICGC MMML-Seq cohort comprises untreated tumor tissue and corresponding germline material (cells from peripheral blood or buffy coats without IGHV clonal rearrangement) obtained with informed consent of the respective patients and/or in minors their legal guardian. The ICGC MMML-Seq study has been approved by the Institutional Review Board of the Medical Faculty of the University of Kiel (A150/10) and Ulm (349/11), and of the recruiting centers. Tumor samples were reviewed by expert hematopathologists and classified according to the WHO 2008 guidelines.[6]

In the framework of the ICGC MMML-Seq network (http://dcc.icgc.org), we mined in the present study data of a total of 186 GC-derived B-cell lymphomas, including 86 follicular Lymphoma (FL), 75 diffuse large B cell lymphoma (DLBCL), 1 primary mediastinal B-cell lymphoma (PMBL), 17 transformed DLBCL (from FL, namely: FL-DLBCL), 1 B-cell lymphoma not otherwise specified (B-NOS), 4 large B-cell lymphoma (LBCL) with *IRF4*-rearrangement, and 2 double hit lymphomas with molecular Burkitt lymphoma (mBL) signature. With the exception of 5 cases (4 LBCL with *IRF4*-rearrangement and 1 PMBL), genomic and transcriptomic data of the cohort have been previously published with regard to other features.[5]

Cell lines

Nineteen B and T-cell non-Hodgkin and 4 classical Hodgkin lymphoma (cHL) derived cell lines were used in the study (Online Supplementary Table S2). Cell lines were obtained from DSMZ (German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany) or from Prof. Stephan Mathas. Cell lines were tested negative for mycoplasma contamination, and their authenticity was confirmed by STR analysis using the StemElite ID System (Promega). In addition, copy number data[7] and/or exome data[8] from previously published studies from the 4 cHL herein studied and 2 additional cHL cell lines were included as well as previously reported WGS data from L1236.[9]

METHOD DETAILS

Methods and procedures used in the ICGC MMML-Seq have been described in previous publications of the network and are summarized in the subsequent section.

Sample Processing

The study was performed in accordance with the ICGC guidelines (www.icgc.org). DNA and RNA extraction, the detection and sequencing of immunoglobulin rearrangements, and whole-genome and transcriptome sequencing of the patients have been published previously.[1] The analysis of WGS was performed as previously described[1,3]. Somatic SNVs and indels in matched tumor normal pairs were identified using the DKFZ core variant calling workflows of the ICGC PCAWG project (https://dockstore.org/containers/quay.io/pancancer/pcawg-dkfz-workflow). Initial candidate variants for SNVs in the tumor were generated by samtools and bcftools (version 0.1.19), followed by a lookup of the corresponding positions in the control. The work flow to identify putative somatic variants and indels was performed as published recently.[3] SNVs and indels were annotated using ANNOVAR[10] according to GENCODE gene annotation (version 19) and overlapped with variants from dbSNP (build 141) and the 1000 Genomes Project database. SNVs classified as splicing, non-synonymous changes, stop-gains, and stop-losses were predicted to affect protein function.

Detection of copy number alterations and genomic structural variants

Allele-specific copy-number alterations were analyzed using ACE-seq (allele-specific copy-number estimation from whole genome sequencing; unpublished data). The detailed method was previously published[3]. ACE-seq was used to determine absolute allele-specific copy numbers as well as tumor ploidy and tumor cell content based on coverage ratios of tumor and control as well as the B-allele frequency (BAF) of heterozygous single nucleotide polymorphisms (SNPs). Moreover, ploidies were manually checked and compared with FISH results.

Structural variants (SV) were called using the SOPHIA algorithm (unpublished data) and DELLY v0.5.9 as previously described[3]. The source code of SOPHIA is available at https://bitbucket.org/utoprak/sophia/ and the method detail has been reported previously.[3] We used DELLY algorithm to call simple and complex SVs. A high confident set of somatic SVs of size >1kb, supported by at least four read pairs, and filtered for absence in the paired normal control tissue was derived. Moreover, we removed SVs detected either in ≥1% of a set of 1105 germline samples from healthy samples constructed from the DELLY's consensus germline SVs called in PCAWG normal tissue sample.

To determine the incidence of SVs in Pan-cancer Analysis of Whole Genomes (PCWAG), filtered structural variant calls were generated by SOPHIA for the PCAWG cohorts processed with the same tools and settings as with the lymphoma cohort used in this study. For each patient existence of a *KDM4C* altering event was determined, using a list of SVs with two breakpoints each, as follows: i) If one breakpoint is directly on *KDM4C* (chr9:6682372-7175648, hg19) and the second is off the designated region on any chromosome the event is a *KDM4C* hit; ii) If both breakpoints are on *KDM4C* and they fall on different exons or introns, the event is a KDM4C hit; iii) If both breakpoints are on *KDM4C* and they fall on the same exon, the event is a *KDM4C* hit; iv) If both breakpoints are on the same intron the event is not a *KDM4C* hit.

Integration of different genomic variant types

SNVs, indels, SVs and CNAs were integrated to account for all variant types in the recurrence analysis. Whilst all genes with SNVs or indels in coding regions (nonsynonymous, splicing, frameshift event) and ncRNA were included, SVs and CNAs were handled differently. Any genes between the breakpoints of focal SVs (<1 Mbp) were considered. However, duplications and deletions called by SOPHIA in the range of 10 kbp and 1 Mbp had to be verified by ACEseq, discarding subclonal events with less than 0.7 copy number deviation from the average ploidy. For larger SVs only genes that were directly hit by a breakpoint were considered. Only focal CNA events (<1 Mbp) were taken into account for variant integration, as these are more likely to target specific genes within the affected region than large events such as whole chromosome arm events. To capture the precise target, focal SVs and CNAs were combined and local maxima of overlapping regions with more than one event were identified.

Clonality analysis of *KDM4C* alterations

The cancer cell fraction (CCF; the proportion of cancer cells with a KDM4C alteration among all cancer cells) was calculated as follows:

$$CCF = \frac{VAF * (TCN * TCC + 2 * (1 - TCC))}{TCC * A}$$

TCN: total copy number at locus (i.e. copy number of reference allele plus copy number of alternative allele)

TCC: tumor cell content

A: copy number of alternative allele

VAF: variant allele fraction, calculated as variant reads divided by all reads

- $VAF = \frac{v}{n+v}$

For deletions, variant reads were calculated as the sum of all split reads at both breakpoints, while normal reads were the average of non-variant supporting reads overlapping the breakpoints.

Case 4191799 has an unbalanced translocation affecting the KDM4C locus. It harbours two intact copies of chromosome 9 and one intact copy of chromosome 8. An additional copy of the telomeric part of 9p is translocated to the (8p-truncated) second copy of chromosome 8. As the total copy number is different at the two breakpoints, the CCF calculation has been performed independently for breakpoints *a* and *b*. Variant reads were taken as the sum of all split reads at both breakpoints, while normal reads were the number of non-variant supporting read at each breakpoint. The variant allele fractions were calculated as

- $VAF_a = \frac{v}{n_a+v}, VAF_b = \frac{v}{n_b+v}$

The alternative allele copy number was 1 at both breakpoints, while the reference allele copy number was 1 at breakpoint *a*, and 2 at breakpoint *b*.

For case 4193638 no cancer cell fraction has been determined as the increased copy number at this locus allows for different possible combinations of reference and alternative allele copy numbers.

To evaluate if the observed variant read counts are compatible with the variant being clonal in the cancer cells, 95% confidence intervals for the binomial probability for the observed number of variant reads and total reads at the breakpoint loci were calculated using the binconf function from the R package Hmisc using the score test-based method after Wilson. The number of observations was the sum of the normal reads at breakpoint *a*, the normal reads at breakpoint *b*, and the variant reads. The number of "successes" was the number of variant reads. Next, the expected proportion of variant reads assuming that the variants were clonal was calculated as follows:

- $p = \frac{TCC*A}{TCC*(A+R_{aT}+R_{bT})+(1-TCC)*(R_{aN}+R_{bN})}$

  - R_{aT}, R_{bT}: copy number of reference allele at breakpoint *a* and *b* in tumor
  - R_{aN}, R_{bN}: copy number of reference allele at breakpoint *a* and *b* in normal

Finally, it was evaluated if *p* is within the 95% confidence interval.

Whole-transcriptome analysis

Transcriptome data were mapped with segemehl 0.2.0[11] allowing for spliced alignments and using a minimum accuracy of 90%. Gene expression values were counted using RNAcounter 1.5.2, using the "--nh" option and counting only exonic reads (-t exon). We looked for backsplicing reads from the *KDM4C* locus in the segemehl output.

Gene expression-based GC-B-cell lymphomas classification

The gene expression classifier was performed as previously described[5]. In brief, a set of 23 differentially expressed genes to distinguish between DLBCL of type ABC and GCB has been established previously.[12] We used these genes to classify our cohort based on RNA-Seq data. We applied the thresholds of <0.25 for GCB-DLBCL and >=0.66 for ABC-DLBCL. Samples in between were labeled as unclassified lymphoma. The classifier was applied to all samples with available RNA-Seq data included in the analysis (n=180).

PCR and Sanger sequencing

SNVs and SVs involving *KDM4C* locus were amplified by polymerase chain reaction from the genomic DNA using specific primers. Primer sequences and PCR conditions are available upon request. Amplicons were purified (MinElute 96 UF PCR Purification Kit, Qiagen, Hilden, Germany) and cycle-sequenced using fluorescent dye-termination (Big Dye Terminator V1.1 Cycle Sequencing Kit, Applied Biosystems, Darmstadt, Germany) and an ABI 3100 or ABI 310 automatic capillary genetic analyser. Alternative *KDM4C* fusion transcripts were validated using specific primers to amplify breakpoint fusion sequences. The RNAs from the tumor samples were treated with DNaseI (RNAase-Free, Ambion, Thermofisher Scientific, Darmstadt, Germany). Complementary DNA (cDNA) was synthesized from 1 µg of total RNA using Quanti Tect Reverse Transcription kit (Qiagen, Hilden, Germany) according to manufacturer´s instruction. The cDNAs were amplified by polymerase chain reaction using specific primers and conditions and the amplicons were purified and sequenced as described above.

***KDM4C* fluorescent in situ hybridization (FISH)**

FISH on interphase nuclei was performed on frozen tissue sections and on methanol: acetic acid fixed cells from cell lines. Chromosomal aberrations affecting the *KDM4C* locus were analysed by FISH using two-FISH-probe designs, i.e. using locus-specific and break-apart *KDM4C* probes. The locus-

specific *KDM4C* assay comprises BAC clone RP11-940A10 spanning *KDM4C* labelled in spectrum green which was combined with a control probe for *NOTCH1* gene locus (9q34.3), which including BACs RP11-251M1 and RP11-83N9 both labelled in spectrum orange. The *KDM4C* break-apart assay contains two differentially labelled BAC clones RP11-356C23 (spectrum orange) and RP11-996O17 (spectrum green) flanking the *KDM4C* locus. BAC clones were obtained from Life Technologies, (Darmstadt, Germany). Digital image acquisition, processing, and evaluation were performed using ISIS digital image analysis version 5.0 (MetaSystems, Altussheim, Germany).

DNA constructs

For generation of the pcDNA3-*KDM4C* expression construct, *KDM4C* was amplified by use of primers *KDM4C*_FLAG_BamHI s 5´- GCGGATCCGCCTGCGAGGTGGCCGAGGTGG and *KDM4C*_EcoRIintern as 5´-GCGAATTCTCCAGCCTCCTGGGTTATC for amplification of the 5´-part of the cDNA (referred to as "part A"), and by use of primers *KDM4C*_EcoRIintern s 5´- GCGAATTCATGATCACTTTCCCATATG and *KDM4C*_Stop_XhoI as 5´- GCCTCGAGGCCTACTGTCTCTTCTGGCAC for amplification of the 3´-part of the cDNA. Thereafter, parts A and B were cloned consecutively *via* BamHI and EcoRI as well as EcoRI and XhoI into a modified pcDNA3 construct containing an N-terminal FLAG tag. For the pRTS-1[13] based inducible *KDM4C* expression vector, *KDM4C* was amplified by use of primers *KDM4C*_5´_ATG_XbaI s 5´- GCTCTAGAGCCACCATGGAGGTGGCCGAGGTGGAAAG and *KDM4C*_3´STOP_XbaI as 5´- GCTCTAGACTACTGTCTCTTCTGGCAC using pcDNA3-*KDM4C* as template. The amplified *KDM4C*-product was ligated *via* XbaI into pUC19-Sfi, and mobilized by SfiI digestion for cloning into pRTS-1. For generation of the lentiviral *KDM4C* expression construct, human *KDM4C* was amplified from the pRTS1-*KDM4C* plasmid by use of primers *KDM4C*_AgeI s 5´-CAACCGGTGCCACCATGGAGGTGGCCGAGGTGG and *KDM4C*_XbaI as 5´-CATCTAGACTGTCTCTTCTGGCACTT and cloned into lentiCRISPRv2-EBFP[14] *via* AgeI and XbaI. All DNA constructs were verified by Sanger sequencing.

Cell lines, culture conditions, transfections and transduction, functional analyses

Cell lines were cultured as previously described.[15] Where indicated, 1 μg/ml doxycycline (Dox; Sigma Aldrich, Taufkirchen, Germany) was added. For generation of KDM4C-inducible cells, cells were electroporated in OPTI-MEM I using Gene-Pulser II (Bio-Rad); L1236 with 960 μF and 160 V, KARPAS-422 with 500 μF and 300 V, NAMALWA with 500 μF and 300 V, HEK293 with 960 μF and 180 V. Twenty-four hours after transfection, Hygromycin B (L1236 110 μg/ml; Karpas-422 50 μg/ml; Namalwa 250 μg/ml; Sigma-Aldrich, Taufkirchen, Germany) was added. After 21 – 28 days of culture in the presence of Hygromycin B, cells were suitable for functional assays. Where indicated, GFP+

cells were enriched 72 hours after dox-induction using a FACS Aria. Production of lentiviruses and lentiviral transduction of cells was performed as described[14]. In brief, $5x10^5$ HEK293T/17 cells were seeded in DMEM (10% FCS; 1% Penicillin/Streptomycin; 1% sodium pyruvate; 1% Glutamax) the day before transfection. For transfection, 10 μg of the lentiviral *KDM4C*-plasmid, 5 μg of the packaging plasmid psPAX2 (Addgene plasmid #12260, a kind gift of D. Tron) and 5 μg of the packaging plasmid pCMV-VSV-G (Addgene plasmid #8454, a kind gift of B. Weinberg) were transfected using calcium-phosphate-precipitation. 48 hours after transfection, the viral supernatant was harvested. For viral transduction, $8x10^5$ SU-DHL-1 cells were seeded in 2 ml, and 2 – 3 ml of viral supernatant was added. Centrifugation was carried out with 2,000 g for 90 min at 32 °C. The day after transduction, cells were washed three times with 1 x PBS. The percentage of viable GFP-positive cells was determined by propidium iodide (PI)-staining, and FACS analyses were performed using a FACS Canto II over time at the indicated times. The percentage of PI-negative, GFP-positive cells at each time point was normalized to the percentage of PI-negative, GFP-positive cells at the starting day of the respective assay, as indicated.

Immunoblotting

Protein preparation and immunblotting were performed as described[15]. The following primary antibodies were used: rabbit polyclonal anti-KDM4C (raised against AA1007-1056 of human JMJD2C; generated in-house, R. Schüle laboratory), mouse monoclonal anti-β-actin (A5316; Sigma Aldrich, Taufkirchen, Germany), mouse monoclonal anti-α-tubulin (MCA78A; Serotec, Puchheim, Germany). Filters were incubated with horseradish peroxidase-conjugated secondary antibodies. Bands were visualized with the enhanced chemiluminescence system (Amersham, Freiburg, Germany).

Immunohistochemisty

FFPE slides from L1236 and HEK293 cells, transfected with pcDNA3 as a control or with pcDNA3-KDM4C, were subjected to immunohistochemical staining using homemade anti-KDM4C antibody (R. Schüle laboratory; see immunoblotting section). The FFPE slides were deparaffinized using roti-histol and hydrated in a degree series of ethanol (absolute, 80% and 70%). Thereafter the antigen retrieval was performed using a sodium citrate solution at pressure cooker 10 minutes. FFPE slides were rinsed in TBS and endogenous peroxidase blocked using Triton X-100. Thereafter, the FFPE slides were incubated with anti-KDM4C antibody (1:100) at 4°C overnight and incubated with mouse and rabbit -HRP conjugated detection IHC kit (Abcam) at room temperature following the manufacturer's instructions. To investigate the possible quantification of KDM4C protein expression several variations of IHC were tested without success.

Protein Modelling

Mechismo (http://mechismo.russelllab.org/)[16] was used to predict the potential effect of SVs and SNVs detected by whole genome sequencing.

Online Supplementary Tables on Excel file

Supplementary Table S1. Histone methylation genes list analyzed (provided as excel format)
The table includes the histone methylation genes analyzed in the study, showing the function (histone demethylase or methyltransferase), the gene name, the Uniprot ID and in which pathway the protein is involved.

Supplementary Table S2. Genomic aberrations affecting *KDM4C* (provided as excel format).

The table displays the focal structural variants (deletion, duplication, translocation) and single nucleotide variants (SNVs) within the *KDM4C* locus detected in the ICGC cohort and the cell lines analyzed in this study. Moreover, an overview of validated genomic somatic SVs and SNVs in analyzed cohort and alternative KDM4C transcript are displayed in the table. The diagnosis and the cell of origin of the ICGC cases and cell lines explored herein are included in the table. *KDM4C* genomic status reported in COSMIC cell line database (https:// cancer.sanger.ac.uk/cell_linesThe KDM4C) and the copy number status of *KDM4C* gene in previously published studies in classical Hodgkin (cHL) cell lines using SNP 6.0[7] are highlighted by asterisk (*) and hashtag ($^{\#}$), respectively. *KDM4C* deletions described in L1236 (labelled with dolar ($^{\$}$) were retrieved from whole genome sequencing data previously published[9]. The *KDM4C* validations using Fluorescent in situ hybridization (FISH), and PCR and Sanger sequencing are depicted in the table. *KDM4C* assay comprises BAC clone RP11-940A10 spanning *KDM4C* labelled in spectrum green that was combined with a control probe for *NOTCH1* gene locus (9q34.3) (including BACs RP11-251M1 and RP11-83N9 both labelled in spectrum orange). The *KDM4C* break-apart assay contains the two differentially labelled BAC clones RP11-356C23 (spectrum orange) and RP11-996O17 (spectrum green) flanking the *KDM4C* locus. Note, the exome sequencing data previously published[8] has not reported the homozygous deletion in L1236 nor the truncating mutation in SUP-HD1.

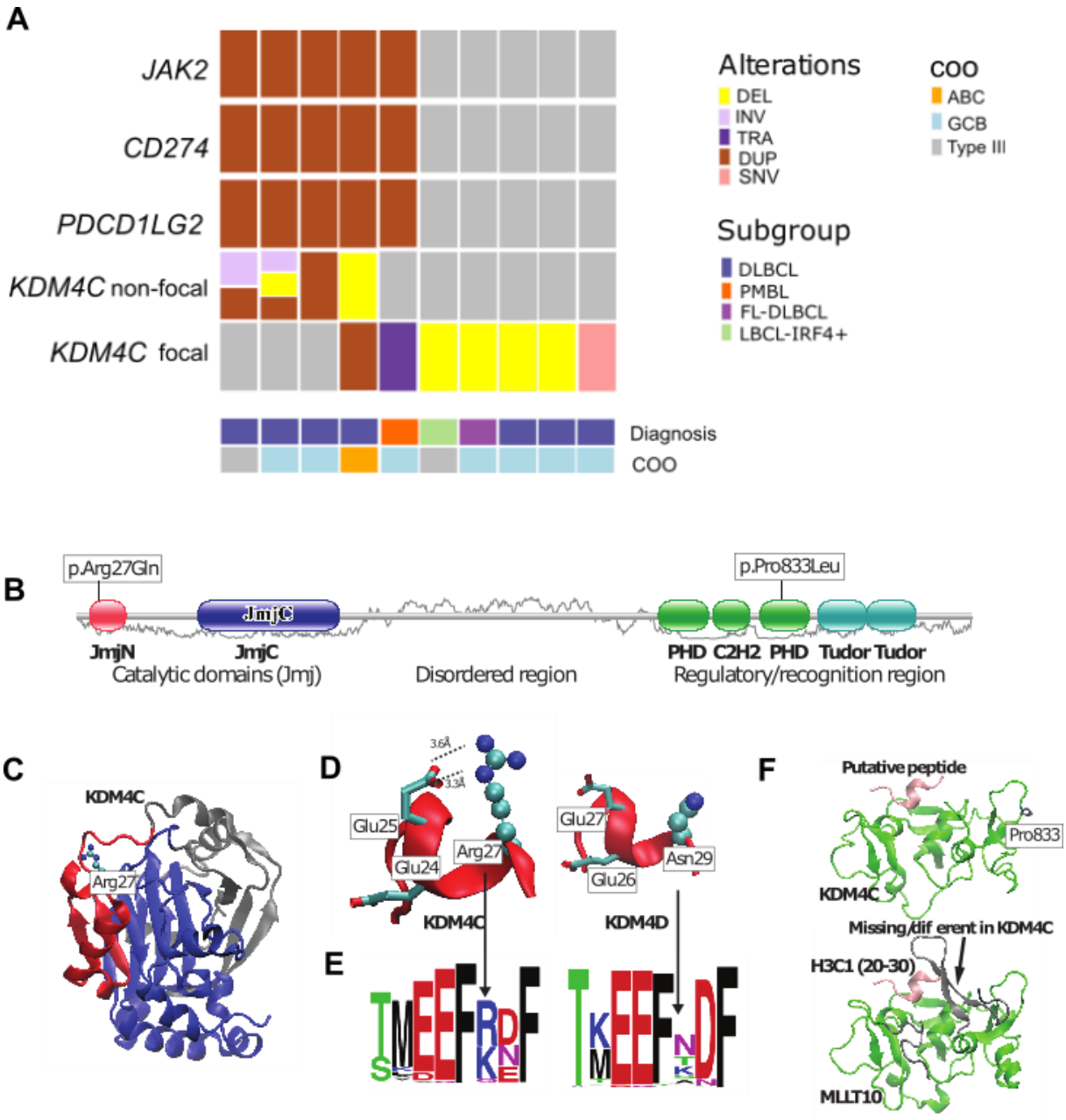
Supplementary Table S3. Targets of KDM4C previously described (provided as excel format).

Supplementary Table S4. Differential expressed genes between *KDM4C* altered cases and *KDM4C* wildtype in FL-DLBCL subgroup (provided as excel format).

Supplementary Figures

Extension of the Legend Figure 2A. Modelling protein of KDM4C aberrations

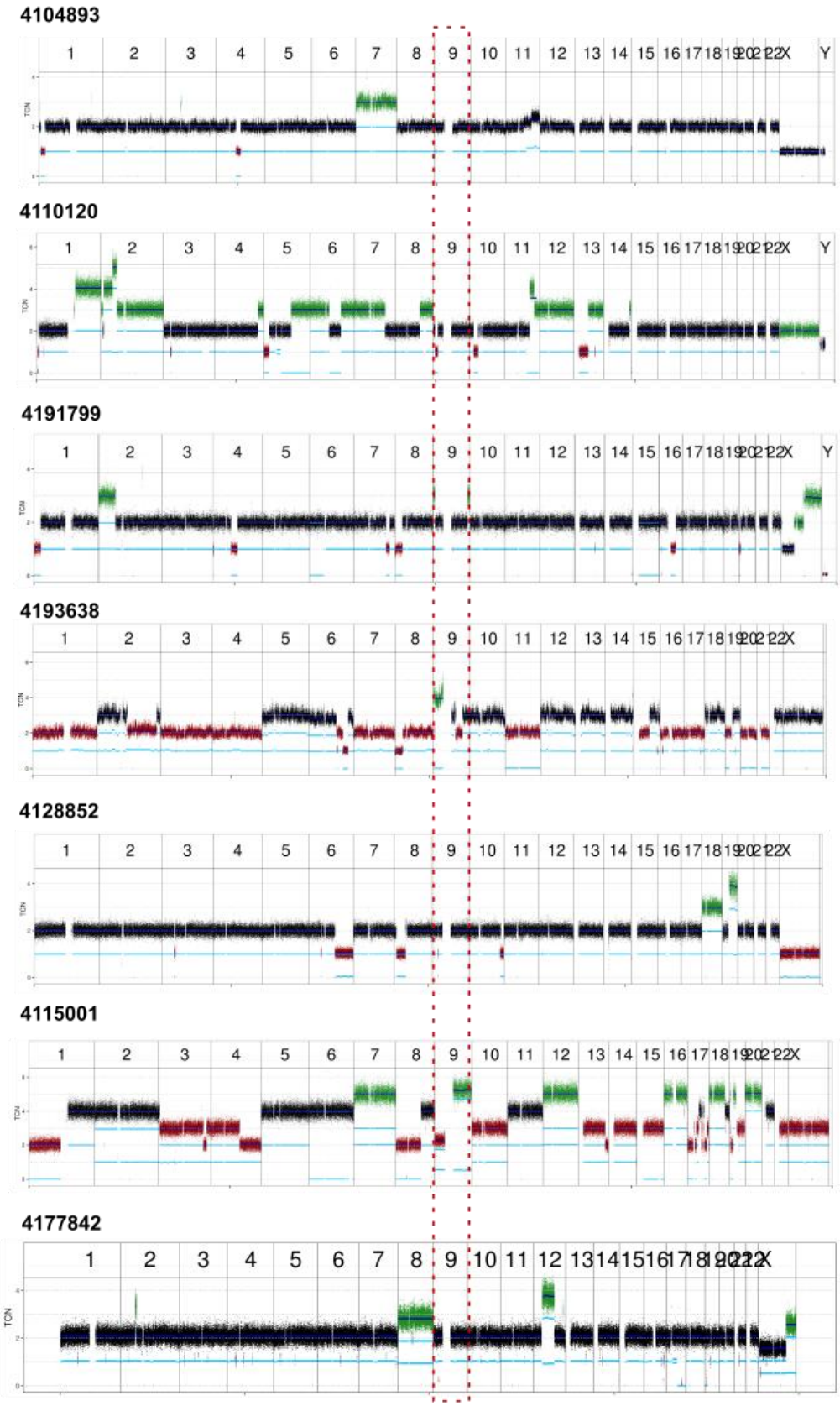
The absence of a catalytic domain in cases 4128852, 41074893, 4115001 can lead to a protein that competes for binding of ligands without catalysis, and argues that this essentially leads to a loss of activity by jamming function of remaining wt-KDM4C proteins. The absence of the recognition domains in case 4191799 will lead to an enzyme that won't necessarily reach its right targets, and could end up acting elsewhere, though it would not hit its intended target as often as the full-length protein. KDM4C in case 4110120 seems unlikely to be functional, due to loss of one catalytic domain and the recognition domains. In addition, case 4177842 harbours a nonsynonymous SNV (c,G80A, p,R27Q) affecting also the JmJH catalytic domain. The predicting impact of this SNV is a pathogenic mutation with consequence at protein level. All predictions were performed using mechismo[16].



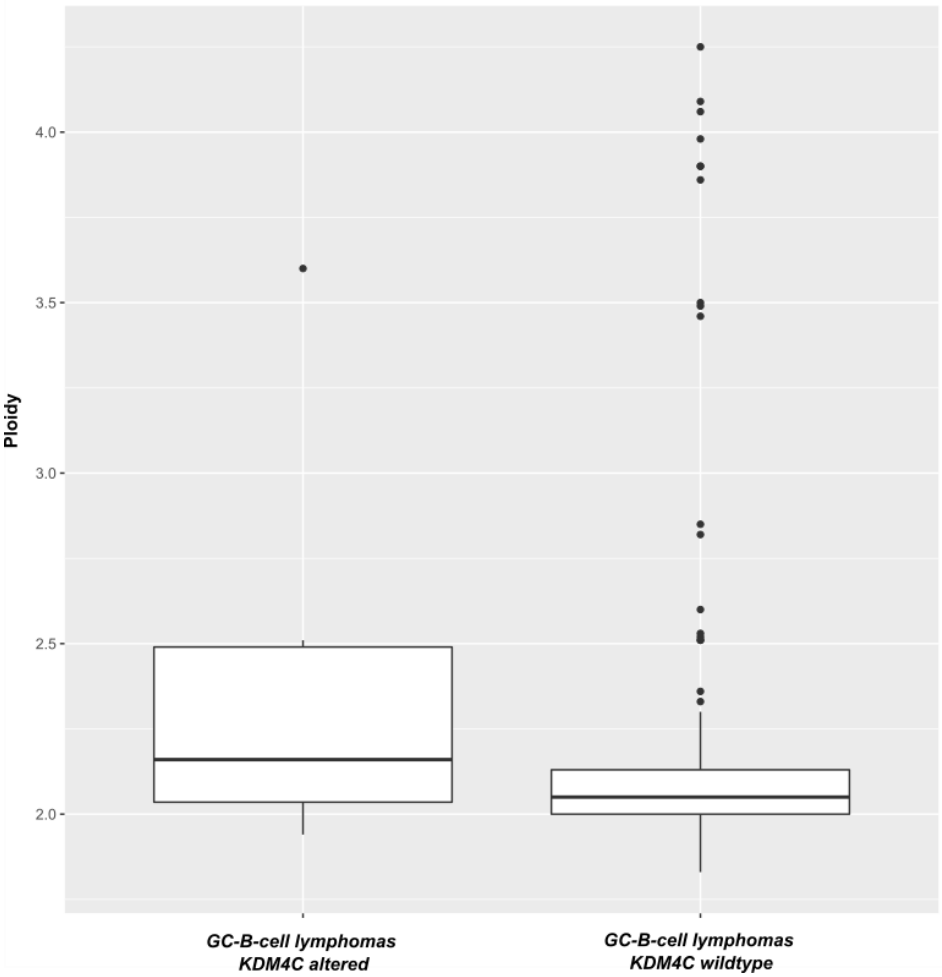
Supplementary Figure S1. Copy number status of genomic cluster in 9p24.1 (*JAK2*, *CD274* and *PDCD1LG2* genes) versus alterations affecting *KDM4C and* KDM4C domain structure and analysis of Arg27 (comparison with the equivalent position in KDM4D) and Pro833.

(A) Oncoprint displaying the cases with copy number gains (focal and non-focal, in contrast to Fig. 1 which only included focal aberrations) involving the *JAK2*, *CD274* and/or *PCD1LG2* genes in GC-derived B-cell lymphomas compared to the genomic status of *KDM4C* in each case. The analyzed genes are ordered from telomere to centromere of chromosome 9. The cases are classified based on diagnosis, and also annotated according to the cell of origin (COO) classification. (B) Domain structure of KDM4C with domains labelled (after Uniprot/Pfam) and showing the position of p.Arg27Gln detected in KDM4C in 4177842 and p.Pro833Leu. The graph behind the domain is a

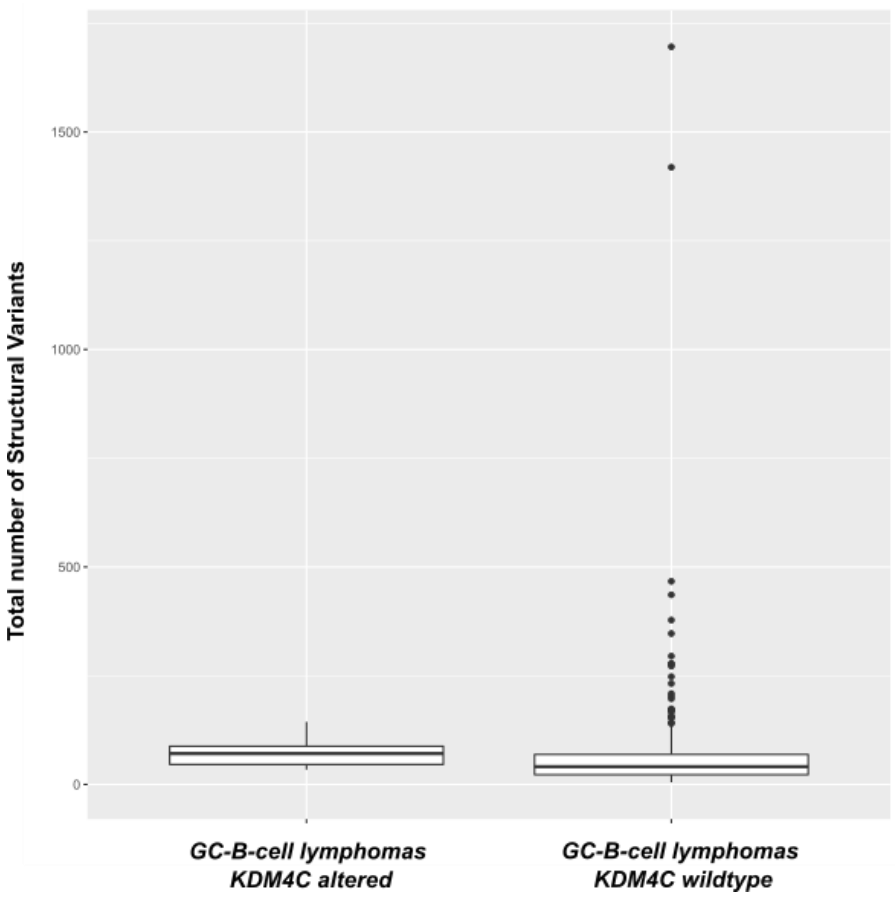
prediction of disordered sequence from IUPred[17]. Both SNVs are reported in the COSMIC cell lines database (cancer.sanger.ac.uk/cell_lines), and were validated by PCR and Sanger sequencing and are predicted as pathogenic variants by FATHMM. (C) KDM4C structure (PDB code 4xdo) showing the JmJN (red) and JmjC (blue) domains together with grey colour denoting the intervening sequence that is also part of the catalytic core. (D) Zoom on the region of Arg27 (spheres) in KDM4C showing Glu24 and Glu25 (sticks), with the putative salt-bridge between Glu25 and Arg27 shown as dashed lines (distances in Angstroms). The equivalent region in KDM4D (PDB code 3dxt) is also shown for comparison. While there are no known interactions here[16] at Arg27, it is adjacent to a pair of Glu residues E24 & E25 which are conserved in the subfamily member KDM4D, where they have been shown to be PARsylated, a modification negatively regulating its activity[18]. Curiously, KDM4D lacks an Arg in the position equivalent to R27 and has an Asn instead, an amino acid similar to Gln. In the KDM4C structure, R27 is hydrogen bonded with one of these Glu residues (PDB:4xdo). It is thus possible that p.R27Q could render KDM4C more like KDM4D in this region, preventing this hydrogen bond and leading to a similar PARsylation and resulting loss-of-function. (E) Sequence motive logos for KDM4C (left) and KDM4D (right) around the positions shown in c. The height of each amino acid symbol denotes its prevalence across orthologs at each position. (F) Model of KDM4C PHD finger domain showing the location of Pro833 (top) in KDM4C detected in the KARPAS-422 cell line and the template structure (MLLT10 bound to H3C1 peptide; PDB:5dah; bottom). The PHD finger is shown in green and the peptides in pink. The grey region in the bottom structure shows a C-terminal, binding-site region that is absent/different in KDM4C and which is contacted by the region around Pro-833. Predictors of variant impact (PolyPhen2, MutationAssessor, SIFT) give conflicting views (from low-impact to damaging) likely because different sequence alignments are used. Position 833 shows a conservation of a small amino acid sequence (Pro, Ser, Ala) across orthologs of KDM4C (EggNOG database), but shows poorer conservation across wider homologs/paralogs (UniProt), where generally a polar amino acid is preferred. This argues for a specific functional role in KDM4C proteins that is different from wider members of this domain family. Inspection of the structure (modelled on PDB:5dah using MODELER; Supplemental Fig. 2) suggests that Pro-833 is involved in orienting the region C-terminal of the PHD finger to form the peptide binding site. Taken together, evolutionary and protein structure predictions also suggest that p.P833L is a loss-of-function variant.



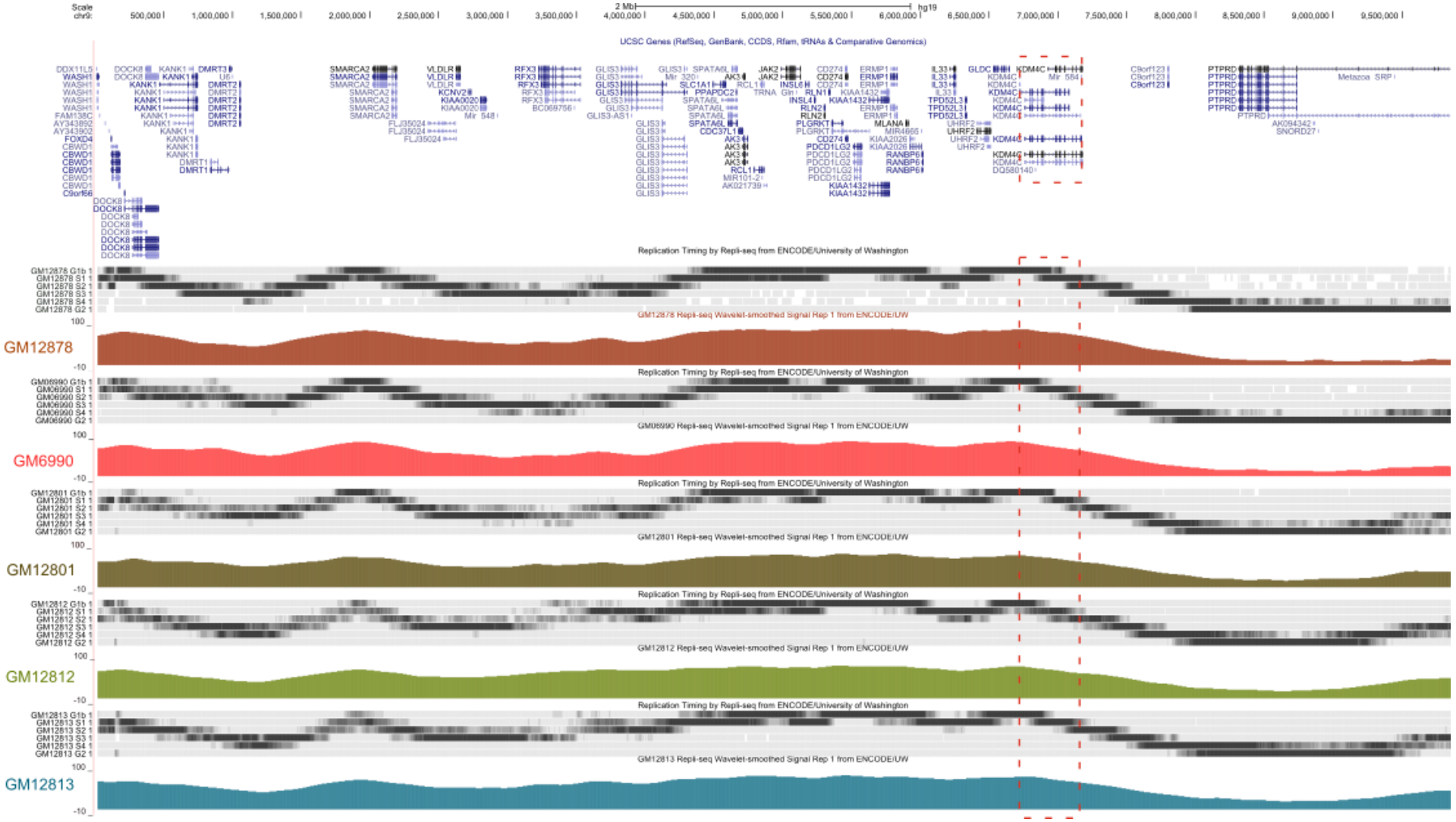
Supplementary Figure S2. Imbalance profiles of germinal center derived B cell lymphomas harboring *KDM4C* aberrations. Losses are indicated in red and gains in green. Chromosome 9, where *KDM4C* locus is located, is highlighted by a red rectangle.



Supplementary Figure S3. Box-plot of overall ploidy in GC-B-cell lymphomas cases with and without *KDM4C* alterations.

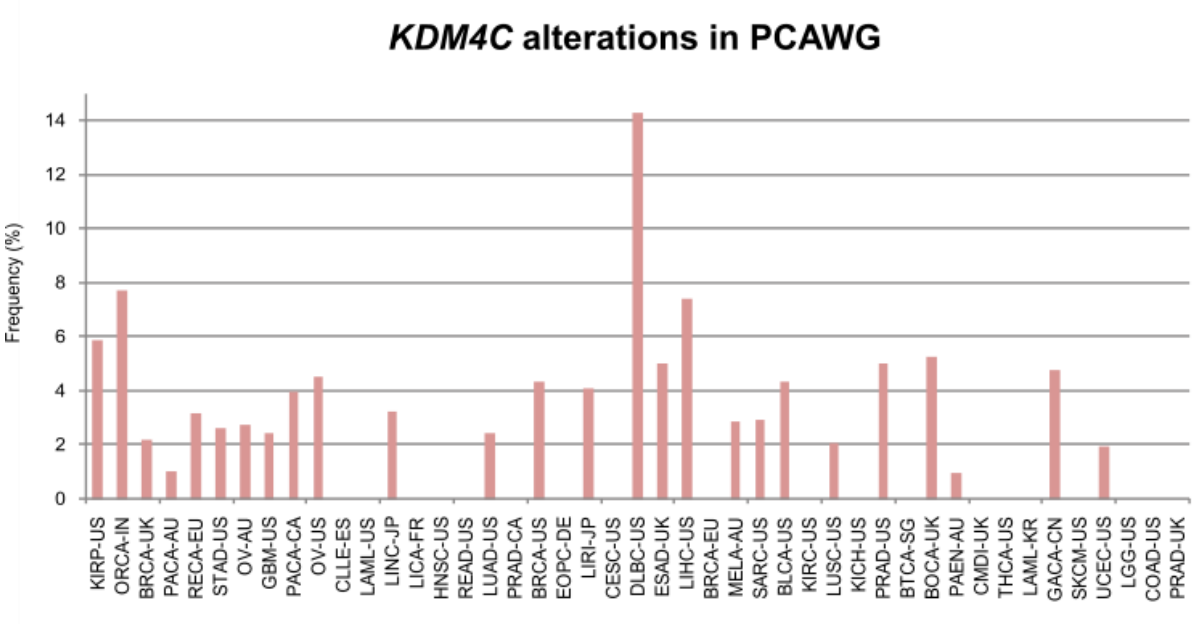


Supplementary Figure S4. Box-plot of the total number of structural variants detected by WGS in lymphomas with and without *KDM4C* alterations.

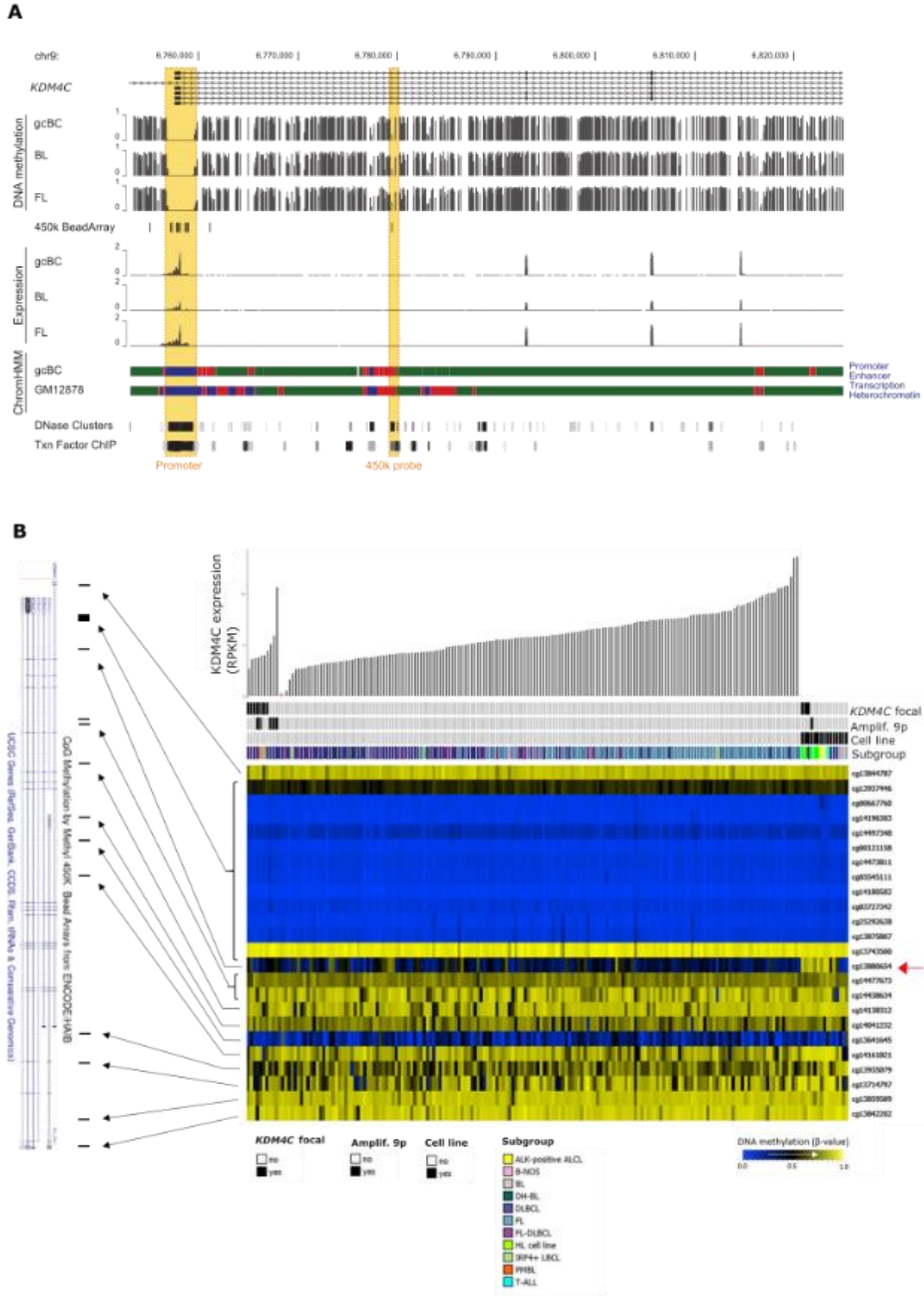


Supplementary Figure S5. Replication timing of the 9p region.

The figure shows the replication timing by Repli-seq from the ENCODE track in the Genome browser (https://genome.ucsc.edu, accessed 10/12/2021) of different lymphoblastoid cell lines[19]. Each line contains the replication timing for the six cell cycle fractions: G1/G1b, S1, S2, S3, S4, and G2. Replication patterns are visualized as a continuous function based on sequencing tag density (*Percentage-normalized Signal*) and as a wavelet-smoothed transform of the six-fraction profile (*Wavelet-smoothed Signal*)[19]. The *KDM4C* locus is highlighted by red rectangle.

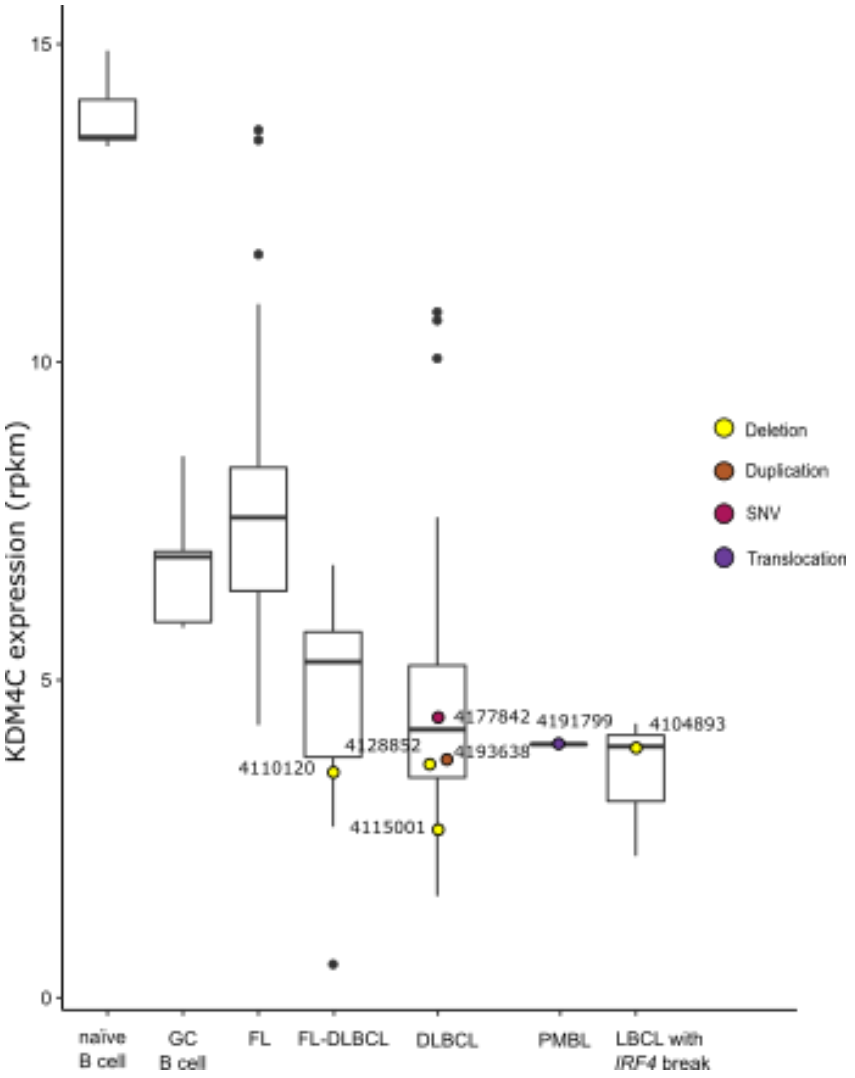


Supplementary Figure S6. Graphic displaying the frequency of structural variants on *KMD4C* in 45 tumor cohorts included PCAWG dataset[20].


A
chr9:
6,760,000
6,770,000
6,780,000
6,790,000
6,800,000
6,810,000
6,820,000
KDM4C
DNA methylation
gcBC
BL
FL
450k BeadArray
Expression
gcBC
BL
FL
ChromHMM
gcBC
GM12878
Promoter
Enhancer
Transcription
Heterochromatin
DNase Clusters
Txn Factor ChIP
Promoter
450k probe
B
KDM4C expression (RPKM)
KDM4C focal
Amplif. 9p
Cell line
Subgroup
UCSC Genes (RefSeq, GenBank, CCDS, Rfam, tRNAs & Comparative Genomics)
CpG Methylation by Methyl 450K Bead Arrays from ENCODE/HAIB
KDM4C focal
no
yes
Amplif. 9p
no
yes
Cell line
no
yes
Subgroup
ALK-positive ALCL
B-NOS
BL
DH-BL
DLBCL
FL
FL-DLBCL
HL cell line
IRF4+ LBCL
PMBL
T-ALL
DNA methylation (β-value)
0.0
0.5
1.0

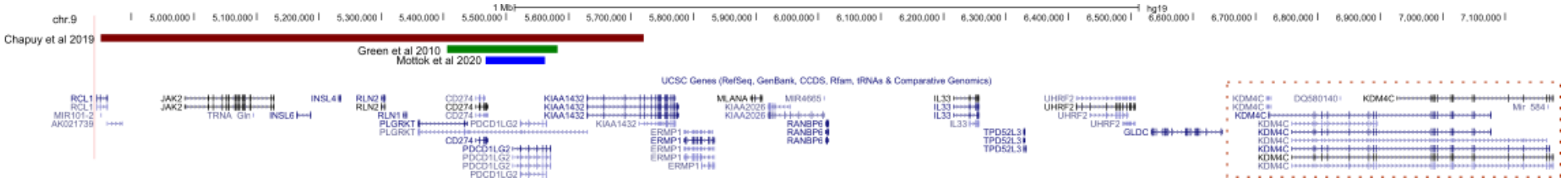
Supplementary Figure S7. Analysis of DNA methylation and chromatin states at the *KDM4C* locus in the ICGC MMML-Seq cohort by mining data from Kretzmer et al., 2015[2].

(A) UCSC genome browser visualisation of the *KDM4C* gene locus showing average DNA methylation and gene expression in germinal center B cells (gcBC), BL and FL as well as chromatin states. The promoter region (left highlight) is enriched in DNase clusters, transcription factor binding sites and is depleted in methylation independent of KDM4C expression levels. The right highlight indicates a differentially methylated 450k Bead Chip methylation array probe cg13880654 hypermethylated in various lymphoma cell lines (see panel b), which overlaps with DNase clusters and transcription factor binding sites and is annotated as an enhancer by ChromHMM. (B) Heatmap displaying the DNA methylation of the CpGs associated with the *KDM4C* gene based on Illumina 450K BeadChip analysis of the ICGC MMML-seq cohort. Twenty-four CpGs within the gene *KDM4C* were identified and displayed for 176 lymphoma samples (10 cases were excluded, not passing the quality control for the analysis) and 15 cell lines. The lymphoma samples and cell lines, respectively, are ordered according to *KDM4C* focal aberration, and/or amplification in 9p24.1 (including *JAK2*, *CD274* and *PDCD1LG2)* status, followed by *KDM4C* gene expression level (is a bar plot at the top of figure S7B), RPKM values, two cases without RNAseq data highlighted with the red asterisk). The 24 CpGs are ordered according to genomic location and mapped to the *KDM4C* locus (left). The CpG 13880654 hypermethylated in lymphoma cell lines except Burkitt lymphoma cell lines is highlighted by an arrow.



Supplementary Figure S8. KDM4C expression in GC-derived B-cell lymphomas.

Mean expression of KDM4C in normal naïve B cell and germinal center derived B cells (GCB cells) from 5 healthy controls and GC-derived B-cell lymphoma, including follicular lymphoma (FL), diffuse large B cell lymphoma (DLBCL), FL-DLBCL, primary mediastinal diffuse large B cell lymphoma (PMBCL), large B-cell lymphoma (LBCL) with *IRF4*-rearrangement. The GC-B-cell lymphomas cases harboring aberrations in *KDM4C* are labeled with the corresponding ICGC MMML-Seq ID and the dot color assigned indicates the type of genomic *KDM4C* aberration (deletion in yellow, duplication in brown, single nucleotide variant (SNV) in dark pink, and translocation in purple).



Supplementary Figure S9. Minimal region of copy number gain on 9p24.1 based on high-resolution studies of PMBL and HL.

Chapuy et al 2019[21] (red color) and Mottok et al., 2019[22] (green color) inferred the copy number aberrations from whole exome sequencing data and used GISTIC algorithm to defined critical region of gain in PMBL cohorts. Green et al 2010[23] analyzed a cohort of HL, DLBCL and PMBL determining the copy number by Affymetrix SNP6.0 microarrays. *KDM4C* locus is highlighted with the red dotted rectangle at the far right. Screenshot based on the USCS browser (https://genome.ucsc.edu, accessed on 09.12.2021).

Remaining members of the ICGC MMML-Seq Consortium

Coordination (C1): Susanne Wagner[2], Gesine Richter[2]
Data Center (C2): Jürgen Eils[20], Jules Kerssemakers[20], Christina Jaeger-Schmidt[20], Ingrid Scholz[20]
Clinical Centers (WP1): Christoph Borst[21], Friederike Braulke[22], Birgit Burkhardt[23], Alexander Claviez[24], Martin Dreyling[25], Sonja Eberth[25], Hermann Einsele[26], Norbert Frickhofen[27], Siegfried Haas[21], Dennis Karsch[28], Nicole Klepl[22], Michael Kneba[28], Jasmin Lisfeld[23], Luisa Mantovani-Löffler[29], German Ott[30], Marius Rohde[31], Christina Stadler[22], Peter Staib[32], Stephan Stilgenbauer[33], Lorenz Trümper[22], Thorsten Zenz[34],
Normal Cells (WPN): Martin-Leo Hansmann[35], Dieter Kube[22], Ralf Küppers[36], Marc A. Weniger[36]
Pathology and Analyte Preparation (WP2-3): Siegfried Haas[37], Michael Hummel[38], Wolfram Klapper[39], Ulrike Kostezka[40], Dido Lenze[38], Peter Möller[41], Andreas Rosenwald[42], Monika Szczepanowski[39],
Sequencing and genomics (WP4-7): Sietse Aukema[2], Vera Binder[43], Arndt Borkhardt[43], Andrea Haake[2], Jessica Hoell[43], Ellen Leich[42], Inga Nagel[2], Jordan Pischimariov[42], Stefan Schreiber[44], Inga Vater[2]
Bioinformatics (WP8-9): Hans Binder[6,7], Benedikt Brors[45], Gero Doose[6,7,8], Jürgen Eils[20], Roland Eils[12,20], Lydia Hopp[6], Helene Kretzmer[6,7,8], David Langenberger[6,7,8], Markus Loeffler[14], Maciej Rosolowski[14], Peter F. Stadler[7].

[20] Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), Heidelberg, 69120, Germany;
[21]Department of Internal Medicine/Hematology, Friedrich-Ebert-Hospital, Neumünster;
[22]Department of Hematology and Oncology, Georg-August-University of Göttingen, 37075 Göttingen, Germany;
[23]University Hospital Münster-Pediatric Hematology and Oncology, 48149, Münster, Germany;
[24]Department of Pediatrics, University Hospital Schleswig-Holstein, Campus Kiel, 24105, Kiel, Germany;
[25]Department of Medicine III - Campus Grosshadern, University Hospital Munich, Munich, Germany;
[26]University Hospital Würzburg, Department of Medicine and Poliklinik II, University of Würzburg, Würzburg;
[27]Department of Medicine III, Hematology and Oncology, Dr. Horst-Schmidt-Kliniken of Wiesbaden, Wiesbaden;
[28]Department of Internal Medicine II: Hematology and Oncology, University Medical Centre, Campus Kiel, Kiel;
[29]Hospital of Internal Medicine II, Hematology and Oncology, St-Georg Hospital Leipzig, Leipzig, Germany;
[30]Department of Pathology, Robert-Bosch-Hospital, Stuttgart, Germany;
[31]Pediatric Hematology and Oncology, University Hospital Giessen, 35392, Giessen, Germany;
[32]Clinic for Hematology and Oncology, St.-Antonius-Hospital, Eschweiler;
[33]Department for Internal Medicine III, Ulm University, 89081, Ulm, Germany;
[34]National Centre for Tumor Disease, Heidelberg, Germany;
[35]Senckenerg Institute of Pathology, University of Frankfurt Medical School, Frankfurt am Main, 60590, Germany;
[36]Institute of Cell Biology (Cancer Research), Medical School, University of Duisburg-Essen, 45147, Essen, Germany;
[37]Friedrich-Ebert Hospital Neumünster, Clinics for Hematology, Oncology and Nephrology, Neumünster, Germany;
[38]Institute of Pathology, Charité-University Medicine Berlin, 10117, Germany;
[39]Hematopathology Section, Christian-Albrechts-University, 24105, Kiel, Germany;
[40]Comprehensive Cancer Center Ulm (CCCU), University Hospital Ulm, 89081, Ulm, Germany;
[41]Institute of Pathology, University of Ulm and University Hospital of Ulm, 89081, Ulm, Germany;
[42]Institute of Pathology, Comprehensive Cancer Center Mainfranken, University of Würzburg, 97080, Würzburg, Germany;

[43]Medical Faculty, Department of Pediatric Oncology, Hematology and Clinical Immunology, Heinrich-Heine-University, 40225, Düsseldorf, Germany;
[44]Department of General Internal Medicine, University Kiel, Kiel, Germany;
[45]Division of Applied Bioinformatics (G200), German Cancer Research Center (DKZF), 69120, Heidelberg, Germany;

References

1. Richter J, Schlesner M, Hoffmann S, et al. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. *Nat Genet*. 2012;44(12):1316-1320.

2. Kretzmer H, Bernhart SH, Wang W, et al. DNA methylome analysis in Burkitt and follicular lymphomas identifies differentially methylated regions linked to somatic mutation and transcriptional control. *Nat Genet*. 2015;47(11):1316-1325.

3. López C, Kleinheinz K, Aukema SM, et al. Genomic and transcriptomic changes complement each other in the pathogenesis of sporadic Burkitt lymphoma. *Nat Commun*. 2019;10(1):1459.

4. Doose G, Haake A, Bernhart SH, et al. MINCR is a MYC-induced lncRNA able to modulate MYC's transcriptional network in Burkitt lymphoma cells. *Proc Natl Acad Sci U S A*. 2015;112(38):E5261-70.

5. Hübschmann D, Kleinheinz K, Wagener R, et al. Mutational mechanisms shaping the coding and noncoding genome of germinal center derived B-cell lymphomas. *Leukemia*. 2021;35(7):2002-2016.

6. Swerdlow, S.H., Campo, E., Harris, N.L., Jaffe, E.S., Pileri, S.A., Stein, H., Thiele, J., Vardiman J. *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues*. IARC; 2008.

7. Otto C, Giefing M, Massow A, et al. Genetic lesions of the TRAF3 and MAP3K14 genes in classical Hodgkin lymphoma. *Br J Haematol*. 2012;157(6):702-708.

8. Liu Y, Abdul Razak FR, Terpstra M, et al. The mutational landscape of Hodgkin lymphoma cell lines determined by whole-exome sequencing. *Leukemia*. 2014;28(11):2248-2251.

9. Schneider M, Schneider S, Zühlke-Jenisch R, et al. Alterations of the CD58 gene in classical Hodgkin lymphoma. *Genes Chromosomes Cancer*. 2015;54(10):638-645.

10. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2010;38(16):e164.

11. Hoffmann S, Otto C, Kurtz S, et al. Fast mapping of short sequences with mismatches, insertions and deletions using index structures. *PLoS Comput Biol*. 2009;5(9):e1000502.

12. Hummel M, Bentink S, Berger H, et al. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. *N Engl J Med*. 2006;354(23):2419-2430.

13. Bornkamm GW, Berens C, Kuklik-Roos C, et al. Stringent doxycycline-dependent control of gene activities using an episomal one-vector system. *Nucleic Acids Res*. 2005;33(16):1-11.

14. Schleussner N, Merkel O, Costanza M, et al. The AP-1-BATF and -BATF3 module is essential for growth, survival and TH17/ILC3 skewing of anaplastic large cell lymphoma. *Leukemia*. 2018;32(9):1994-2007.

15. 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.

16. Betts MJ, Lu Q, Jiang Y, et al. Mechismo: predicting the mechanistic impact of mutations and modifications on molecular interactions. *Nucleic Acids Res*. 2015;43(2):e10.

17. Mészáros B, Erdös G, Dosztányi Z. IUPred2A: Context-dependent prediction of protein

disorder as a function of redox state and protein binding. *Nucleic Acids Res*. 2018;46(W1):W329-W337.

18. Le May N, Iltis I, Amé J-C, et al. Poly (ADP-ribose) glycohydrolase regulates retinoic acid receptor-mediated gene expression. *Mol Cell*. 2012;48(5):785-798.

19. Hansen RS, Thomas S, Sandstrom R, et al. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc Natl Acad Sci U S A*. 2010;107(1):139-144.

20. Campbell PJ, Getz G, Korbel JO, et al. Pan-cancer analysis of whole genomes. *Nature*. 2020;578(7793):82-93.

21. 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.

22. Mottok A, Hung SS, Chavez EA, et al. Integrative genomic analysis identifies key pathogenic mechanisms in primary mediastinal large B-cell lymphoma. *Blood*. 2019;134(10):802-813.

23. Green MR, Monti S, Rodig SJ, et al. Integrative analysis reveals selective 9p24.1 amplification, increased PD-1 ligand expression, and further induction via JAK2 in nodular sclerosing Hodgkin lymphoma and primary mediastinal large B-cell lymphoma. *Blood*. 2010;116(17):3268-3277.