

The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing

I. Vater, M. Montesinos-Rongen, Matthias Schlesner, A. Haake, F. Purschke, R. Sprute, N. Mettenmeyer, I. Nazzari, I. Nagel, J. Gutwein, J. Richter, I. Buchhalter, R. B. Russell, O. D. Wiestler, R. Eils, M. Deckert, R. Siebert

Angaben zur Veröffentlichung / Publication details:

Vater, I., M. Montesinos-Rongen, Matthias Schlesner, A. Haake, F. Purschke, R. Sprute, N. Mettenmeyer, et al. 2015. "The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing." *Leukemia* 29 (3): 677–85.
<https://doi.org/10.1038/leu.2014.264>.

Nutzungsbedingungen / Terms of use:

licgercopyright

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under these conditions:

Deutsches Urheberrecht

Weitere Informationen finden Sie unter: / For more information see:

<https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/>



The mutational pattern of primary lymphoma of the central nervous system determined by whole-exome sequencing

I Vater^{1,8}, M Montesinos-Rongen^{2,8}, M Schlesner^{3,8}, A Haake¹, F Purschke², R Sprute², N Mettenmeyer², I Nazzari², I Nagel¹, J Gutwein¹, J Richter¹, I Buchhalter³, RB Russell^{4,5}, OD Wiestler⁶, R Eils^{3,5,7}, M Deckert^{2,9} and R Siebert^{1,9}

INTRODUCTION

Primary CNS lymphoma (PCNSL), defined as diffuse large B-cell lymphoma (DLBCL) confined to the CNS, differs from extracerebral DLBCL by its less favorable prognosis.^{1–3} In phenotypic and molecular terms, PCNSL exhibits some features similar to extracerebral DLBCL, but differs in others.³ It is tempting to speculate that molecular characteristics considerably contribute to the poor prognosis of PCNSL.

Various genetic and epigenetic alterations are involved in PCNSL pathogenesis.² They contribute to the late germinal center (GC) exit B-cell phenotype (IGM+CD20+BCL6+MUM1/IRF4+) with an arrest in terminal B-cell differentiation.^{2,4} This can be explained at least in part by continued BCL6 activity, which in a subset of PCNSL may be caused by chromosomal translocations affecting the *BCL6* locus (17–47%),^{3,5–7} by impaired immunoglobulin (*IG*) class-switch recombination due to *su* deletions,⁸ and by deleterious *PRDM1* mutations (19%).⁹ Furthermore, the process of somatic hypermutation (SHM) that physiologically targets rearranged *IG* genes does extensively modify *IG* genes in PCNSL.^{10,11} This process also aberrantly targets the proto-oncogenes *PIM1*, *MYC*, *RHOH/TTF*, and *PAX5* in PCNSL at high frequency (90%).¹² Molecular changes converge on pathways crucially regulating B-cell activation, proliferation, and apoptosis. Genes of the B-cell receptor (BCR) pathway (*SHIP*, *CD79B*, *CBL*, and *BLNK*),¹³ of the toll-like receptor (TLR) pathway (*MYD88*),¹⁴ and of the BCL10-CARD11-

MALT1 (BCM) complex (*CARD11*), that is, the BCR pathway target, harbors functionally relevant mutations in PCNSL¹⁵ or are amplified (*MALT1*).^{7,16} Altogether, these molecular alterations stimulate the NF- κ B pathway and sustain high-BCR and NF- κ B activity.^{3,17}

Our current knowledge on the molecular biology of PCNSL on the one hand relies on array-based profiling studies of genomic and transcriptional alterations.^{4,16,18–21} These have indicated deregulated pathways as well as candidate genes located in minimal regions of imbalance, like HLA class II genes in the region of homozygous loss in 6p21.¹⁹ On the other hand, the main pathogenetic insights are derived from locus-specific approaches using FISH or sequencing of candidate genes that identified recurrently translocated genes, like *BCL6*^{6,7,12} or recurrently mutated genes, like *MYD88* and *CD79B*,^{13,14} respectively. Although these studies provided potentially important insights into PCNSL development, they are, inevitably, restricted to candidates of interest. Thus, mutations in the genes not targeted by recurrent imbalances or otherwise identified as candidates may have been systematically missed. Recent advantages in sequencing technology allow genome-wide and, thus, unbiased mutation analysis. In the recent past, genome-wide sequencing studies using whole genome, exome, or transcriptome sequencing have been performed in a number of tumors, including various B-cell lymphomas.^{22–27} In addition to the detection of pathogenetically relevant driver genes, these techniques provide a unique

¹Institute of Human Genetics, Christian-Albrechts-University Kiel & University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany; ²Institute of Neuropathology, University of Cologne, Cologne, Germany; ³Division of Theoretical Bioinformatics (B080), German Cancer Research Center (DKFZ), Heidelberg, Germany; ⁴Cell Networks, University of Heidelberg, Heidelberg, Germany; ⁵BioQuant, Heidelberg University, Heidelberg, Germany; ⁶German Cancer Research Center (DKFZ), Heidelberg, Germany and ⁷Department for Bioinformatics and Functional Genomics, Institute for Pharmacy and Molecular Biotechnology (IPMB), Heidelberg University, Heidelberg, Germany. Correspondence: Dr M Deckert, Institute of Neuropathology, University Hospital of Cologne, Kerpener Straße 62, Köln D-50924, Germany.

E-mail: martina.deckert@uni-koeln.de

⁸These authors contributed equally to this work.

⁹These authors contributed equally to this work.

advantage of the depiction of mutational patterns of tumors.²⁸ The latter might be used to identify active or passive mutational mechanisms in cancer, like aberrant SHM (aSHM) in B-cell lymphomas. Although whole genome or exome data of more than 100 extracerebral DLBCL have been published, the number of genomes published from PCNSL is still limited. This discrepancy may be attributed to the fact that usually biopsy size is very restricted in this rare disease, owing to the fact that the use of stereotactic biopsies is the golden diagnostic standard. One study performing exome sequencing in 4 PCNSL identified mutations of *MYD88*,¹⁸ which were already identified previously in a candidate gene approach,¹⁴ as well as mutations of *TBL1XR1*.¹⁸ Nevertheless, the entire landscape of mutations has not been published in this study.

To get a deeper insight into the mutational pattern of mutations as well as into the mutational processes active in PCNSL, we have performed whole-exome sequencing (WES) in a series of 9 PCNSL. We show that aSHM significantly contributes to the mutational load of PCNSL. The spectrum of coding mutations shows a considerable overlap with that of DLBCL particularly affecting the pathways described above. In addition, our data suggest that alterations in genes involved in CNS development might contribute to the specific manifestation of PCNSL in the brain.

MATERIALS AND METHODS

Patients and diagnoses

A retrospective collection of a total of 9 PCNSL was studied, in which sufficient amount of cryopreserved tissue was available. Four and five samples were derived from stereotactic and open neurosurgical approaches, respectively (Table 1). These cases have not been included in previous studies. Diagnoses were established according to the WHO classification.²⁹ All samples had a tumor load of at least 80%. Systemic lymphoma manifestation was excluded by extensive staging. All patients were HIV-negative. The study was approved by local Ethics Commissions in Cologne and Kiel (06–187, 07–109, B208/12) and performed in accordance with the Declaration of Helsinki. In line with the protocols approved by the Ethics Commissions, all samples were derived from deceased patients and, therefore, corresponding germline material was not available for the studies.

DNA extraction

High-molecular weight DNA was extracted from frozen tissue with the Gentra Puregene DNA Purification Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. WES and PCR-based verification were performed using the same DNA sample of the respective cases.

PCR-based clonality analyses

The clonal rearrangements of the IG heavy chain (*IGH*) locus were amplified with VH gene segment family-specific primers as published.¹⁰

Whole-exome sequencing

WES was performed as custom service (GATC Biotech, Konstanz, Germany). Fragmentation was done with the E210 (Covaris, Woburn, MA, USA), followed by target enrichment with the SureSelectXT Human All Exon V4 Kit (Agilent, Santa Clara, CA, USA). Quantification and QC was determined using the Agilent 2100 Bioanalyzer (Agilent). Paired-end sequencing (2*100 bp) was performed on the HiSeq 2000 (Illumina, San Diego, CA, USA) using the TruSeq SBS Kit v3 (Illumina) according to the standard protocol.

Bioinformatic mutation analyses

Bioinformatic analyses of next generation sequencing data were performed according to Jones *et al.*,³⁰ with slight modifications. Sequencing reads were mapped and aligned to the hs37d5 reference assembly (hg19+decoy sequences) using Burrows–Wheeler Aligner (BWA) (version 0.6.1),³¹ and were processed with SAMtools (version 0.1.17)³² and Picard tools (version 1.61; <http://picard.sourceforge.net>).

An in-house analysis pipeline based on SAMtools mpileup and bcftools was used to detect single nucleotide variants (SNVs). In addition to previously described filters to remove artifacts, we excluded variants located in regions of low mappability or overlapping with the hiSeq-DepthTop10Pct, Encode DAC Blacklisted Regions, and Duke Excluded Regions tracks from the UCSC Genome Browser. High-confidence somatic SNVs were not allowed to overlap with any two of the following features: tandem repeats, simple repeats, low-complexity, satellite repeats, and segmental duplications. The following heuristic criteria were required: (i) at least five tumor reads at the position; (ii) more than one variant read per strand or at least five variant reads in total and variant allele fraction > 0.1.

Indels were called with Platypus³³ on the tumor BAM file and an unrelated BAM file (sequenced blood from an unrelated individual) as control to subtract common artifacts.

Since matched germline material was unavailable, somatic variants were identified by comparison with genetic variation databases. Variants present in the 1000 genomes database with an MAF > 1% were removed. Variants detected in an in-house control database derived from deep whole-genome sequencing data (>30× coverage) from the blood of 80 unrelated individuals were filtered out.

SNVs and Indels were annotated using Annovar³⁴ and in-house-developed scripts.

To correlate regional SNV density with replication timing, we binned the somatic SNVs from the 9 PCNSL cases into 1 MB windows. We retrieved the length of target regions overlapping with each 1 MB window and calculated the SNV count per kb of the target region. RepSeq replication-timing data for the lymphoblastoid cell line GM12878³⁵ was downloaded from the ENCODE project (<http://genome.ucsc.edu/ENCODE>), averaged for each 1 MB window and correlated with the SNV densities.

SNV clusters were detected by merging the somatic SNV calls from the 9 PCNSL cases. Each genomic region containing five or more SNVs within 1 kb was considered as seed cluster. Overlapping seed clusters and seed clusters with a distance of less than 1 kb were merged to retrieve the final SNV clusters.

Verification and validation of mutations by Sanger sequencing

Selected SNVs detected by WES were verified by Sanger sequencing. PCR was performed using standard procedures (for primers, Biomers, Ulm,

Table 1. Clinicopathologic characteristics of the PCNSL studied

Case	Sex	Age (years)	Location	Biopsy	IgH	IGHV mut. freq.	BCL6	IHC	MUM1	Ki67
#1	m	79	White-matter, parieto-occipital	Open resection	IGHV4–34; IGHD2–21; IGHH4	24.9	+	+	+	+
#2	m	71	Insular region	Stereotactic biopsy	IGHV4–34; IGHD1–26; IGHH5	2.1	+	+	+	+
#3	m	71	Corpus callosum	Stereotactic biopsy	IGHV4–34; IGHD3–10; IGHH4	0.0	+	+	+	+
#4	m	65	Corpus callosum	Open resection	IGHV3–7; IGHD3–16; IGHH4	10.5	+	+	+	+
#5	f	67	Cerebellum	Open resection	IGHV3–21; IGHD6–19; IGHH5	1.2	(+)	+	+	+
#6	m	63	White-matter, bifrontal	Open resection	IGHV4–34; IGHD5–18; IGHH6	6.3	(+)	+	+	+
#7	m	37	Cerebellum	Open resection	IGHV2–5; IGHD3–22; IGHH4	1.7	(+)	+	+	+
#8	m	58	Temporal lobe	Stereotactic biopsy	IGHV3–7; no IGHD; IGHH5	11.1	+	(+)	(+)	(+)
#9	f	65	White-matter, temporo-parietal	Stereotactic biopsy	IGHV4–34; IGHD3–10; IGHH6	10.6	n.a.	n.a.	n.a.	n.a.

Abbreviations: IgH, Immunoglobulin heavy chain locus; IHC, Immunohistochemistry; n.a., not applicable. Grading score: BCL6: + high, (+) low, grading according to Brunn *et al.*³⁸; MUM1: + 90–100%, (+) 80%; Ki67: + 90–100%, (+) 60%.

Germany, see Supplementary Table 1), followed by direct sequencing using an ABI 3130 or 310 Genetic Analyzer and the Big Dye Terminator v1.1 or v3.1 Cycle Sequencing Kit (Life Technologies, Carlsbad, CA, USA). For validation, an independent cohort of 22 stereotactic biopsy samples of PCNSL (Supplementary Table 5) was analysed for mutations in the genes *MPEG1*, *CCND3*, and *SUSD2* following the same procedure.

Mining of gene expression data

Gene expression profiling (GEP) data obtained from analysis of 21 PCNSL as well as 20 DLBCL and 10 normal CNS samples published previously served as the basis for further analyses.⁴ Raw data were normalized by Affymetrix software (Affymetrix, Santa Clara, CA, USA) and further analyzed in geWorkbench 2.5.0.³⁶ In brief, all values below 128 were replaced by 128 and log-2 transformed thereafter. Supervised analysis of PCNSL with DLBCL was performed for all the genes, which were mutated in at least 2 PCNSL. Significantly differentially expressed tags were calculated based on Student's *t*-test. Only tags that were at least two-fold significantly differentially expressed are shown. Samples of normal brain tissue were included for comparison.

Immunohistochemistry

20 paraffin-embedded PCNSL were studied using rabbit anti-PTPRD polyclonal antibody (clone bs-11471R, Bioss, Woburn, MA, USA) with 3,3'-diaminobenzidine (Sigma, Deisenhofen, Germany) as chromogene and H₂O₂ as co-substrate on an automated immunostainer (Bond, Leica, Wetzlar, Germany).

RESULTS AND DISCUSSION

Histopathologic and molecular characterization of the study population

To detect genetic drivers involved in the pathogenesis of PCNSL, we performed WES of 9 PCNSL from immunocompetent patients. In accordance with previous observations,³⁷ all tumors showed the typical clinical, morphological, immunohistochemical, and molecular features of PCNSL with a CD20+CD79B+BCL6+MUM1/IRF+Ki67^{high} phenotype (Table 1). Histopathologically, the tumor-cell content was at least 80% in all samples. Moreover, all 9 PCNSL displayed a clonal *IGH* rearrangement and 8/9 PCNSL had introduced somatic mutations into the rearranged VH gene fragment (Table 1). All sequences were in-frame and did not contain a stop codon; thus, they were considered as being potentially functional.

Whole-exome sequencing and verification of sequence variants

Coverage of the WES data for the 9 PCNSL is provided in Supplementary Table 2. The median coverage of high-quality sequences on target was 103.35 × (range: 76.22–114.49 ×). As matched normal DNA was unavailable, we identified private mutations that were present in the PCNSL but absent from public (dbSNP, 1000genome) and in-house databases, and considered these as candidate somatic mutations in analogy to Fontebasso *et al.*³⁸ This filtering approach left a median of 723 (range: 520–953) potentially somatic SNVs and 51 (range: 35–60) small indels per PCNSL (Figure 1; Supplementary Table 3). A total of 1800 of these SNVs (median per PCNSL: 202, range: 139–251) and 122 indels (median per PCNSL: 14, range: 7–22) showed potentially protein-changing features (Figure 1).

To verify our technical approach of exome sequencing, we subjected a total of 85 mutations affecting 16 genes called by WES to PCR-based Sanger sequencing. In this regard, we decided to focus on the genes that likely have a functional role in the pathogenesis of PCNSL and for which (some) evidence has been obtained in PCNSL or other B-cell neoplasms.^{3,24,26} Of these 85 mutations, 77 could be validated demonstrating a high concordance rate of 90.6% between WES and Sanger sequencing (Supplementary Figure 1). Only eight mutations detected by WES could not be verified by Sanger sequencing. Remarkably, these

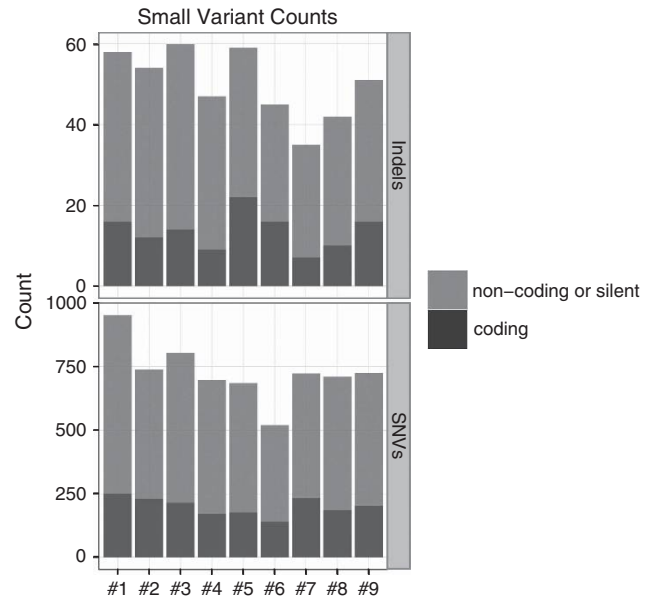


Figure 1. Distribution of mutations. Number of potentially somatic indels (upper panel) and SNVs (lower panel) per PCNSL case. The grey bar indicates noncoding variants, the black bar coding variants (for indels: all indels overlapping exonic regions or splice sites; for SNVs: nonsynonymous, stopgain, stoploss, and splice-site).

mutations all affected the *PIM1* gene in PCNSL #1 and #3. *PIM1* is a target of aSHM in PCNSL¹² as well as in systemic DLBCL.³⁹ In line with the earlier observations made independently in another series of 10 PCNSL,¹² the pattern of mutations in *PIM1* detected in the PCNSL analysed herein clearly shows evidence for aSHM. Moreover, the data of this study do not only provide evidence for aSHM also targeting *PIM1* but also indicate that part of the mutations have escaped detection or were hard to detect by Sanger sequencing likely because of their subclonal nature (Figure 2).

Validation of mutations in selected genes in an independent cohort of PCNSL

An independent cohort of 22 stereotactic biopsy samples of PCNSL was studied for the presence of mutations affecting the *MPEG1*, *CCND3*, and *SUSD2* genes. Sanger sequencing detected 7/22 (32%) in *MPEG1*. This prevalence is similar to WES data. Also for *CCND3*, mutation prevalence did not vary significantly between WES and validation cohorts being 3/9 (33%) in the prior and 1/22 (5%) in the latter. Sanger sequencing failed to detect *SUSD2* mutations in the cohort used for validation (0/22, 0%). This discrepancy may be because of the increased sensitivity of WES to detect mutations as compared with Sanger sequencing. In addition, a sampling bias may also have a role.

PCNSL show a homogenous genome-wide distribution of mutations

Having shown the validity of the WES data we next investigated the genome-wide pattern of mutations. The spectrum of somatic point mutations was conserved between the cases, with C>T/G>A transitions being the most frequent events (Figure 3a). Similarly, the trinucleotide context affected by mutations was quite conserved between the PCNSL again showing the predominance of the C>T/G>A transitions (Figure 3b). This predominance of C>T/G>A mutations is in line with the fact that this is a generally common mutation in the genome linked to deamination of unmethylated cytosines.²⁸

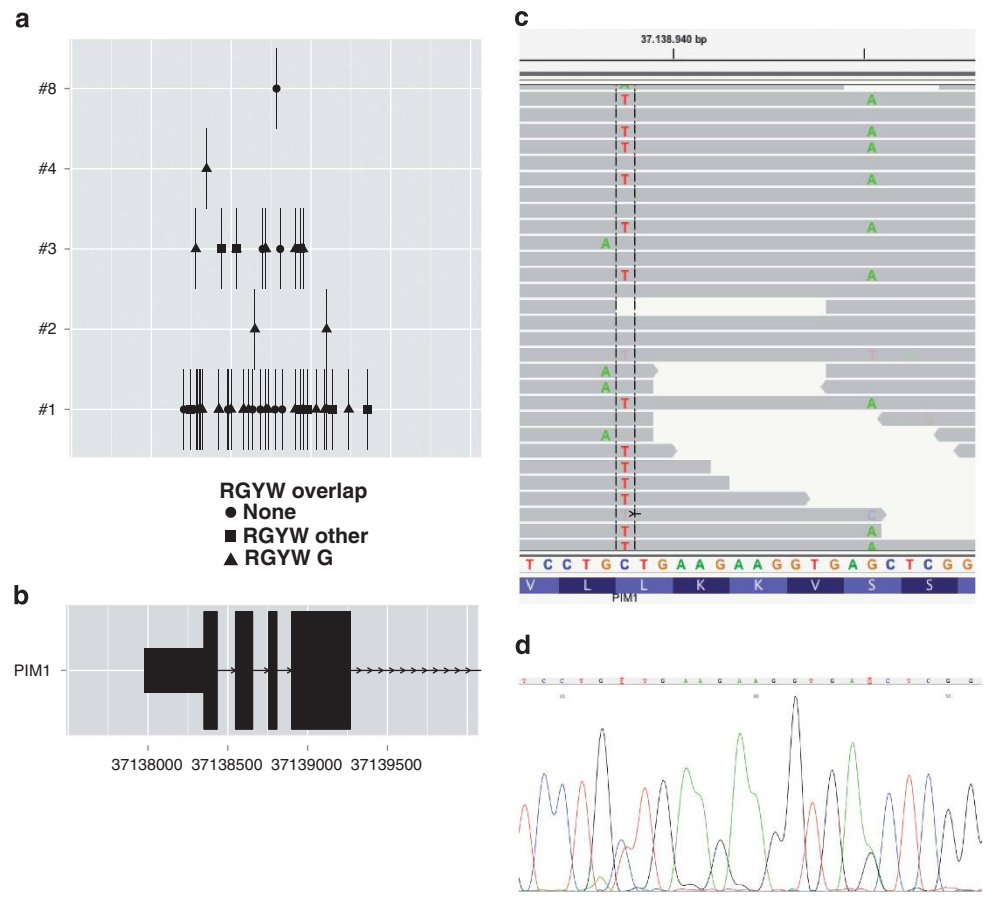


Figure 2. Verification of mutations. (a) The *PIM1* gene is affected by clusters of mutations. Each symbol indicates a somatic SNV. The shape of the symbol shows the relation of the SNV to the RGYW sequence motif (R = A or G; Y = C or T; W = A or T), which is preferably targeted by SHM. Triangle: the 'G' of RGYW is affected (the hotspot of the motif); square: another position of RGYW is affected; circle: no overlap with RGYW. (b) Exon-intron structure of *PIM1*. Arrows indicate introns, thick bars exons, and thin bars UTR's. (c and d) Comparison of the *PIM1* exon 4 in PCNSL #3 by WES (c) and Sanger sequencing (d). Based only on Sanger sequencing data, the first mutation would be missed and only the second and third mutations in (d) would be assessed. Together with the WES data, the first mutation in (d) could be verified. Nevertheless, (c) and (d) demonstrates the limited power of mutation detection merely based on Sanger sequencing.

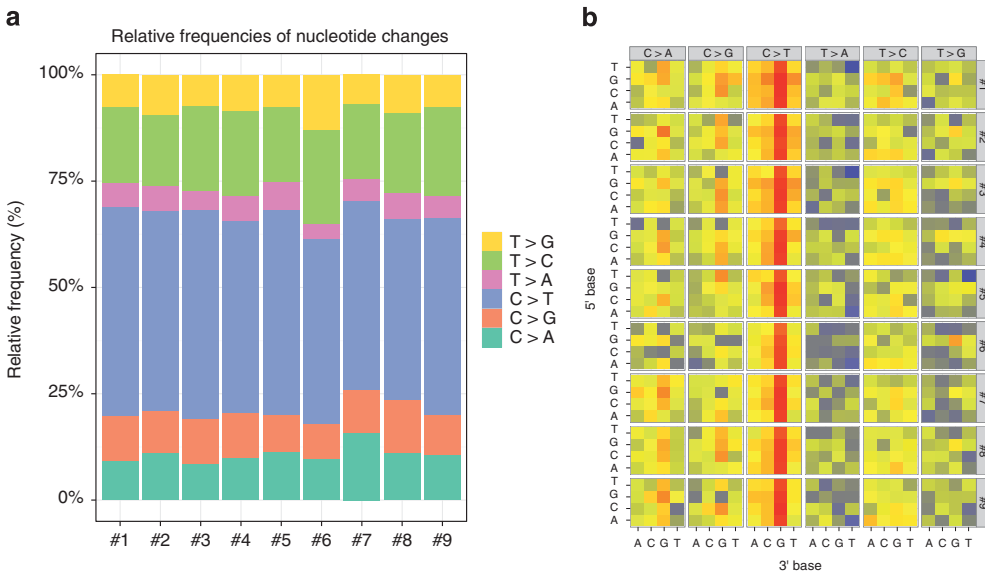


Figure 3. Somatic SNV profiles. (a) Spectra of base substitutions for the PCNSL cases. (b) Heatmap indicating the count of each mutation type in the context of the neighbouring bases. The counts are normalized for the frequency of the respective trinucleotide in the reference genome, and log-transformed values of the normalized ratios plotted. The y axis shows the SNV's 5' base, the x axis the 3' base.

Clusters of mutations in PCNSL point to aberrant activity of the somatic hypermutation machinery

To identify potential clusters of mutations in the genome, we next plotted the number of somatic point mutations per kb against the replication time in B lymphocytes (Figure 4a). Although clusters of somatic mutations (also called kataegis) are known to accumulate in late replicating regions, we here identified a series of gene loci showing high mutational rates in early replicating regions.⁴⁰ These almost exclusively represent genes well known to undergo SHM/aSHM in PCNSL and DLBCL, such as *IGH*, *IGK*, *IGL*, *MYC*, *PAX5*, *BCL2*, and *PIM1*.^{10–12,39} As SHM preferentially targets RGYW motifs, in particular the G of this motif, we investigated the prevalence of this motif among all mutations. As shown in Figure 4b, a median of 22.2% (range: 20.0–24.7%) of somatic SNVs in the 9 PCNSL overlap with the RGYW motif targeted by SHM, and a median of 7.9% (range: 6.2–12.6%) affect its hotspot position. Thus, our data strongly indicate that the process of aSHM significantly contributes to the mutational landscape of PCNSL.

Interestingly, PCNSL #2, 3, 5, and 7, which show low/no *IGH* mutation frequencies, show evidence for aSHM. These data may either indicate that certain SHM confer different changes in fitness of the tumor cell or that once a tumor cell has optimized its BCR for specific antigen recognition, a switch of the SHM machinery to

target other, potentially oncogenic genes, may provide a growth advantage.

Next, we asked whether hitherto unrecognized targets of aSHM can be identified in our dataset. First, we performed a genome-wide search for SNV clusters containing more than five SNVs per kb in the 9 PCNSL. Considering the frequency of mutations in the RGYW motif as well as the fact that the SHM machinery usually affects active genes within 2–2.5 kb from transcription start we then searched for clusters of mutations fulfilling these criteria. These analyses suggested *KLHL14*, *OSBPL10*, and *SUSD2* as potential hitherto unknown targets of aSHM in PCNSL (Figures 4c and d). These data provide the basis for future analysis of these genes in the setting of clinical trials to address their potential role as biomarker.

Genes recurrently affected by protein-coding changes in PCNSL

A total of 150 genes were affected by protein-coding small mutations in at least two PCNSL cases (Supplementary Table 4) with a total of 27 genes affected in at least 3 PCNSL (Figure 5). The genes displaying the highest frequency of protein-coding small mutations in our cohort were *CD79B*, *PIM1*, and *ODZ4* each affected in 4/9 PCNSL and *MUC16* affected in 5/9 PCNSL. *MUC16* is recurrently identified as highly mutated gene in cancer genomic

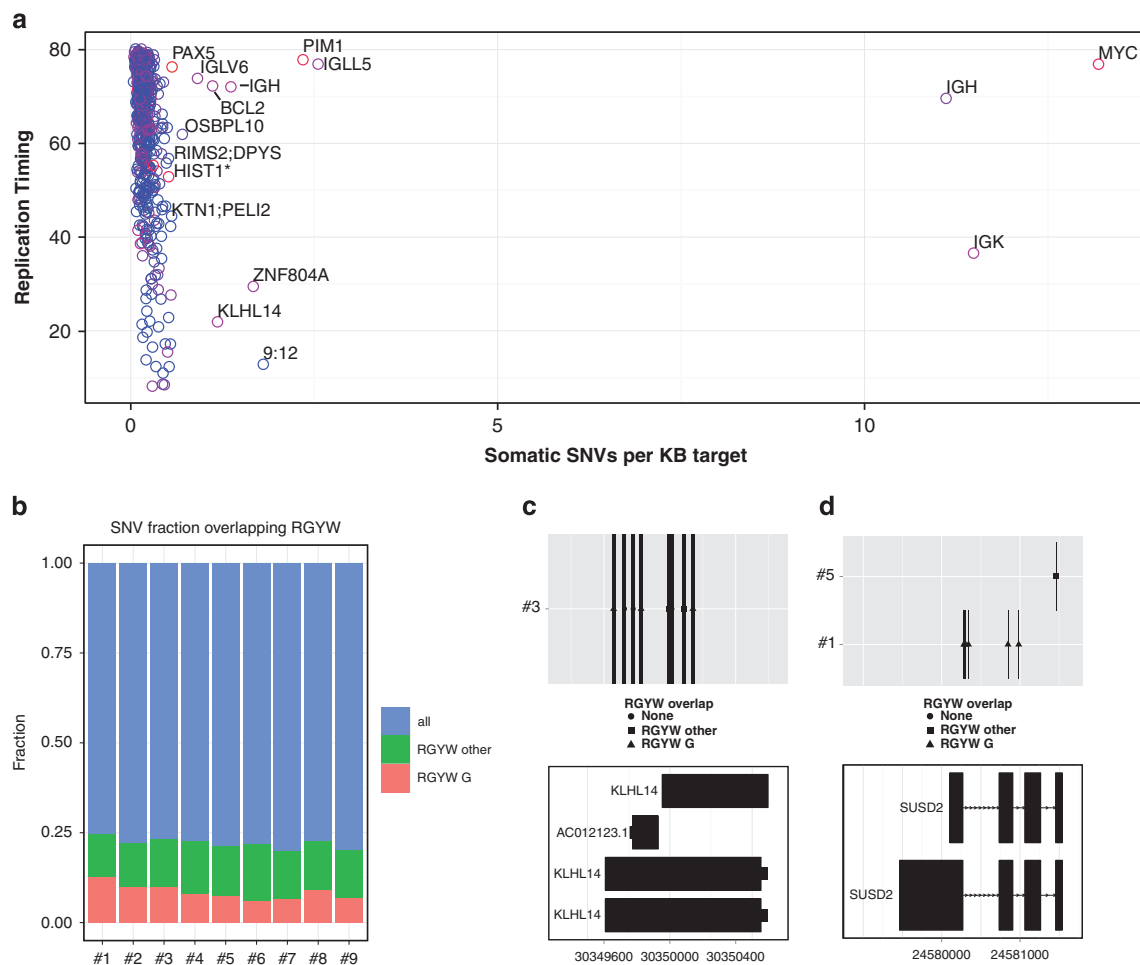


Figure 4. aSHM in PCNSL. **(a)** Correlation of regional SNV density with replication timing. SNVs from all PCNSL cases were merged, and the SNV density (number of SNVs per kb target region) for each 1 MB window of the reference genome determined. This SNV density is correlated with the average replication timing (determined for the lymphoblastoid cell line GM12878³⁵) of the 1 MB window. The color indicates the fraction of SNVs in the window which affect the 'G' of RGYW. **(b)** Fraction of SNVs overlapping with the RGYW motif (green) or targeting its hotspot position 'G' (red). **(c)** and **(d)** Clusters of hypermutation affecting the *KLHL14* and *SUSD2* genes. See Figure 2 for details.

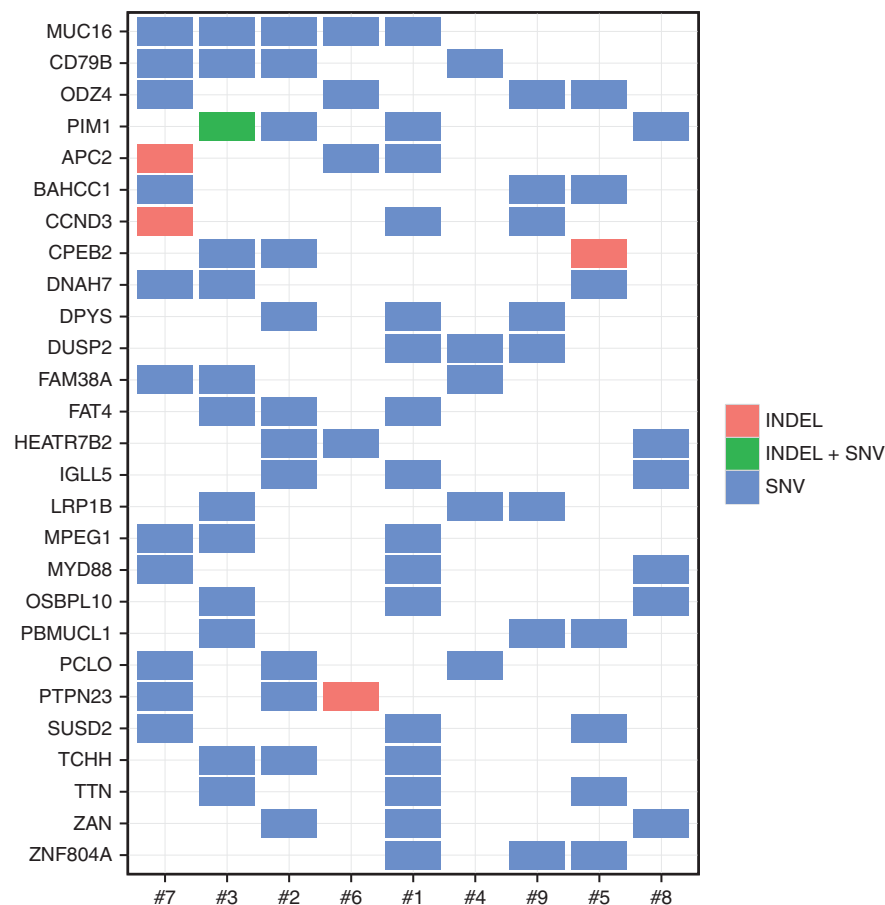


Figure 5. Genes affected by mutation in at least three PCNSL.

studies of various tumors and has been reported by Morin *et al.*²⁶ to be present in 7/40 systemic DLBCL analysed. In contrast, the *PIM1* oncogene is a well-known target of the SHM machinery in PCNSL as well as in systemic DLBCL.^{3,39} These data extend and confirm observations suggesting increased mutation frequencies of genes targeted by SHM as well as aSHM in PCNSL as compared with systemic DLBCL.^{12,39} With regard to *CD79B* we have previously shown this (mutations in 5/20 PCNSL) as well as other genes of the BCR pathway to be mutated in 44% of PCNSL.¹³ Similarly, *MYD88* mutations identified in 3/9 PCNSL in the present study have been described previously by others and us to be present in 38 and 50% of PCNSL, respectively.^{14,18} These data demonstrate that despite the limited number of cases and the lack of germline material our approach is able to pick up pathogenetically relevant mutations in PCNSL. This is further confirmed by a study of also 9 PCNSL published during the review process of our manuscript, which also identified *MYD88*, *CD79B*, *IGLL5*, *PIM1*, *MUC16*, *OSBPL10*, and *MPEG1* among the most frequently mutated genes.⁴¹

A subset of genes recurrently affected by protein-coding changes maps to regions of recurrent chromosomal imbalances in PCNSL. Chromosomal imbalances can be mechanisms alternative to mutation leading to deregulation of genes involved in the pathogenesis of cancer. As, thus, the location of such chromosomal imbalances might point to pathogenetically relevant genes, we asked whether recurrently mutated genes identified by the exome sequencing approach overlapped with minimal regions of recurrent imbalances detected in recent SNP array analyses of

PCNSL.^{16,18,42,43} Of the genes with at least two mutations in our data set, a total of 35 map to regions of recurrent genomic imbalance or partial uniparental disomies (pUPD) in PCNSL, including the histone encoding genes *HIST1H3H* and *HIST1H2AE* in the pUPD region in 6p, the *PTPRD* gene in the pUPD region in 9p24 and the *NCOR2* and *POU2F2* in the gained regions in 12q24.31 and 19q13.2, respectively. A total of seven genes affected with at least three mutations in our exome study were described to be involved in recurrent genomic imbalances. These genes comprise *PCLO*, *ZAN*, *FAM38A*, *BAHCC1*, and *APC2* located in recurrently gained chromosomal regions, as well as *PBMUCL1* and *SUSD2*, located in regions of recurrent loss.

Differential expression of genes recurrently affected by protein-coding changes in PCNSL

Considering that activation or inactivation of cancer genes can also occur from altered expression mediated, for example, by epigenetic changes, we next investigated the expression of the recurrently mutated genes mining previously published data on GEP of PCNSL.⁴ Comparison of PCNSL and systemic DLBCL GEP data with regard to the identified mutated genes revealed five of them to exhibit a differential mRNA expression. Among these genes, again were *NCOR2* and *POU2F2* located in the gained regions in 12q24.31 and 19q13.2 that are overexpressed in PCNSL as compared with DLBCL (Supplementary Figure 2). The *POU2F2* protein is prominently expressed by the tumor cells of PCNSL (this study, data not shown).

The mutational pattern of PCNSL partially overlaps with that of systemic DLBCL.

As already outlined above, our study is in line with previous investigations that show clear evidence for ongoing aSHM in PCNSL.¹² We here provide further evidence that pathways and cellular functions altered in other B-cell lymphomas are also targeted by protein-coding mutations in PCNSL. These include members of the BCR, TLR, and NF- κ B pathways (such as *CARD14*, *CD79B*, *CD79A*, *TLR2*, *TLR6*, and *TLR10*) as well as proteins involved in cell-cycle regulation (*CCND3*, *CDK18*, and *CDK20*), immune recognition (*HLA-B*), or chromatin structure and modifications (*CREBBP*, *MLL2*, *ARID1A*, *ARID1B*, *SMARCA4*, *SMARCA5*, and *SMARCC1*). Thus, it is not surprising that genes affected by mutations in our PCNSL exome study widely resemble those recently described to be recurrently mutated in DLBCL.^{24,25} The mutated genes include known targets of aSHM⁴⁴ like *PIM1*, *MYC*, *BCL2*, *IRF4*, and *SOC51* which showed at least three protein-changing mutations per affected PCNSL sample. In this regard, it is also interesting to note that the *MYC*^{high}*BCL2*^{high} phenotype is significantly increased in PCNSL as compared with systemic DLBCL accounting for 82% of PCNSL with the *MYC*^{neg}*BCL2*^{neg} phenotype being absent.^{45–47} Moreover, we identified recurrent mutations (≥ 2 PCNSL samples affected) in *BTG2*, *CCND3*, *CD79B*, *DUSP2*, *LRR1Q3*, *MUC16*, *MYD88*, *PCLO*, *PIM1*, and *POU2F2*, all of which are well known or have recently been suggested to be functionally relevant in DLBCL.^{24,25,27,48} These molecular and immunological findings may indicate that the lymphoma entity of PCNSL is more homogeneous than systemic DLBCL, at least at the phenotypic level. Alternatively, these mutations may have been ‘undercalled’ in DLBCL probably due to lower tumor-cell content and/or the existence of multiple subclones.

The *HIST1* gene cluster is recurrently targeted by somatic mutations in PCNSL.

Deletions including homozygous deletions in 6p21 in PCNSL have been proposed to target the HLA class II locus constituting a mechanism of immune escape.^{16,19,49–51} Although this might be true it does not explain why many PCNSL show pUPD terminally of the HLA locus in 6p. Although our exome data also provide evidence for deregulated immune recognition, we here also identify recurrent somatic mutations of the *HIST1* gene cluster located in 6p22 within the region of pUPD in PCNSL. Our findings strongly indicate that the enrichment of mutations in this cluster is due to aSHM. First, 7 out of 18 SNVs in this region affect the G of the SHM hotspot motif RGYW. Second, we have observed clusters of SNVs in this region in PCNSL #1, which is the tumor with the strongest hypermutation signature in the aSHM target genes identified in our cohort. Strikingly, here the putative aSHM extends over multiple genes spanning a region of roughly 2 MB without focal SNV hotspots. These findings suggest a role of deregulated chromatin structure in the pathogenesis of PCNSL, which in turn could contribute to the considerable chromosomal instability present in these tumors.^{7,16}

Frequent somatic mutations in genes involved in CNS development in PCNSL

ODZ4 was among the top four genes most frequently targeted by protein-coding mutations being affected in 4/9 PCNSL. Morin *et al.*²⁶ reported nonsynonymous mutations of this gene in 2/40 DLBCL. Thus, the frequency of changes in PCNSL seems to be significantly higher ($P=0.0074$, Fisher’s Exact Test) than in systemic DLBCL. Remarkably, so far, *ODZ4* mutations have yet been mutually exclusive with *PIM1* mutations. *ODZ4* encodes the teneurin transmembrane protein 4 (TENM4), a type II transmembrane protein highly expressed in the CNS in oligodendrocytes and neurons.⁵² It has been suggested that teneurin may act as

both an axonal and an oligodendroglial factor crucial for the myelination of small-diameter axons in the CNS and that it regulates oligodendrocyte differentiation.⁵³ *Fur/fur* mice that lack teneurin 4 expression show hypomyelination in the CNS involving the corpus callosum,⁵³ a region frequently affected by PCNSL.²⁹ Furthermore, a susceptibility locus near *ODZ4* was identified to be associated with bipolar disorder.⁵⁴ Most (3/4) mutations we identified herein in PCNSL in *ODZ4* lie in the extracellular region: K609R is predicted to lie on the surface of an EGF domain and could, based on other EGF structures, interact with a ligand. G1763R and S1888T lie between the unique, cysteine-rich, YD repeats that are involved in binding to carbohydrates, such as heparins (e.g. Minet *et al.*). S142F is intracellular and predicted to lie in a GSK3, CK1 or PKA phosphorylation site. Thus, all changes are candidates for disrupting *ODZ4*/teneurin 4 function (see Supplementary References).

In addition to TENM4, protein tyrosine phosphatase receptors (PTPR) are involved in neuriteogenesis and axonal growth,^{55,56} and defects of *ODZ4* and *PTPRD* have been detected in aggressive high-stage neuroblastoma.^{57,58} In our study, we identified protein changing mutations in *PTPRD*, *CSMD2*, and *CSMD3* in 3/9, 4/9, and 4/9 PCNSL, respectively (Supplementary Figure 1).

Moreover, to show that *PTPRD* is indeed expressed by the tumor cells of PCNSL, immunohistochemistry was performed in an independent series of 20 PCNSL studied previously.³⁷ In all of these PCNSL, *PTPRD* was expressed by the tumor cells (Figure 6). The *PTPRD* mutation K1031M lies on the surface of an extracellular FN3 domain and positioned, modelled on other FN3 domains, such that this positive-to-hydrophobic change would affect intra-domain or inter-protein interactions. In contrast, R1348K affects an intracellular site likely on the surface of the phosphatase catalytic domain as modelled on the structure of the human phosphatase PTP Sigma (see Supplementary References).

CSMD3, located on 8q22.3–q24.1 to which benign adult familial myoclonic epilepsy type I has been mapped, is expressed mainly in the brain.⁵⁹ *CSMD2* and *CSMD3* have been suggested as tumor suppressors in several tumors including oligodendrogliomas.⁵⁹ The mutations detected in *CSMD2* and *CSMD3* in PCNSL all affect the extracellular part of the protein and lie in or between CUB or Sushi domains, which are known to play roles in oligomerization and interactions with other proteins. *CSMD2* R2033C is predicted

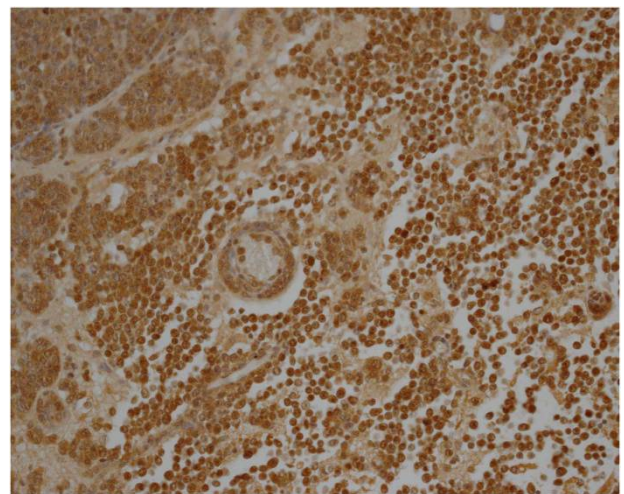


Figure 6. Immunohistochemistry for PTPRD expression in PCNSL. All tumors of a series of 20 PCNSL expressed PTPRD. 80% (16/20) showed an expression of $>80\%$ of the tumor cells and an expression of $>50\text{--}79\%$ of the tumor cells in the remainder. The photomicrograph shows PTPRD expression by $>90\%$ of the tumor cells in a representative PCNSL. Anti-PTPRD immunostaining, slight counterstaining with hemalum, original magnification $\times 200$.

to lie at a predicted interface with another CUB domain. G2280E similarly lies on the surface of a CUB domain (though not at a known interface). R687* would delete several CUB and Sushi domains at the C-terminus of the protein, although possibly leaving a viable protein (as the mutation would not necessarily disrupt folding of any one domain). The CSMD3 mutation E861G lies within a Sushi domain and is more buried than those in CSMD2 and modelling this on the structure of complement factor B in complex with complement C3B and Factor D predicts this residue to lie adjacent to a cysteine involved in a disulphide bond. The modification could thus disrupt the overall structure of the domain and thus interactions with other proteins (see Supplementary References).

Taken together, our data reveal alterations in genes that are not only physiologically expressed in the CNS but also involved in tumorigenesis. It is tempting to speculate that ODZ4, a homologue of ODZ2, which is involved in a chromosomal translocation in MALT lymphomas,⁶⁰ as well as CSMD2/CSMD3 and PTPRD may play a similar role in PCNSL influencing both tumorigenesis and differentiation.

CONCLUSIONS

In summary, our WES study of PCNSL provides novel insights into the pathogenesis of PCNSL. On the one hand, we provide evidence for the genome-wide activity and pathogenetic relevance of aSHM in PCNSL. On the other hand, focusing on coding mutations we show a considerable overlap of deregulated pathways between PCNSL and systemic DLBCL. In addition to genes affected also in DLBCL, the transformational process in PCNSL may take advantage of mutations in genes facilitating tumor manifestation in the CNS, for example, mutations in genes involved in CNS differentiation. Whether the changes in these genes are indeed drivers with functional relevance for the pathogenesis of PCNSL or rather reflect bystander alterations needs to be addressed in future studies.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

We thank all our technicians. This study has been supported by the Deutsche Krebsstiftung (grant nos. 109471, 109472), ICGC MML-Seq project/German Ministry for Education and Science (BMBF Grant no. 01KU1002 A-J), and the Jackstädt foundation.

REFERENCES

- Deckert M, Engert A, Brück W, Ferreri AJ, Finke J, Illerhaus G *et al.* Modern concepts in the biology, diagnosis, differential diagnosis and treatment of primary central nervous system lymphoma. *Leukemia* 2011; **25**: 1797–1807.
- Deckert M, Montesinos-Rongen M, Brunn A, Siebert R. Systems biology of primary CNS lymphoma: from genetic aberrations to modeling in mice. *Acta Neuropathol* 2014; **127**: 175–188.
- Montesinos-Rongen M, Siebert R, Deckert M. Primary lymphoma of the central nervous system: just DLBCL or not? *Blood* 2009; **113**: 7–10.
- Montesinos-Rongen M, Brunn A, Bentink S, Basso K, Lim WK, Klapper W *et al.* Gene expression profiling suggests primary central nervous system lymphomas to be derived from a late germinal center B cell. *Leukemia* 2008; **22**: 400–405.
- Cady FM, O'Neill BP, Law ME, Decker PA, Kurtz DM, Giannini C *et al.* Del(6)(q22) and BCL6 rearrangements in primary CNS lymphoma are indicators of an aggressive clinical course. *J Clin Oncol* 2008; **26**: 4814–4819.
- Montesinos-Rongen M, Akasaka T, Zühlke-Jenisch R, Schaller C, Van Roost D, Wiestler OD *et al.* Molecular characterization of BCL6 breakpoints in primary diffuse large B-cell lymphomas of the central nervous system identifies GAPD as novel translocation partner. *Brain Pathol* 2003; **13**: 534–538.
- Montesinos-Rongen M, Zühlke-Jenisch R, Gesk S, Martin-Subero JL, Schaller C, Van Roost D *et al.* Interphase cytogenetic analysis of lymphoma-associated chromosomal breakpoints in primary diffuse large B-cell lymphomas of the central nervous system. *J Neuropathol Exp Neurol* 2002; **61**: 926–933.
- Montesinos-Rongen M, Schmitz R, Courts C, Stenzel W, Bechtel D, Niedobitek G *et al.* Absence of immunoglobulin class switch in primary lymphomas of the central nervous system. *Am J Pathol* 2005; **166**: 1773–1779.
- Courts C, Montesinos-Rongen M, Brunn A, Bug S, Siemer D, Hans V *et al.* Recurrent inactivation of the PRDM1 gene in primary central nervous system lymphoma. *J Neuropathol Exp Neurol* 2008; **67**: 720–727.
- Montesinos-Rongen M, Küppers R, Schlüter D, Spieker T, Van Roost D, Schaller C *et al.* Primary central nervous system lymphomas are derived from germinal-center B cells and show a preferential usage of the V4-34 gene segment. *Am J Pathol* 1999; **155**: 2077–2086.
- Thompson AR, Ellison DW, Stevenson FK, Zhu D. V(H) gene sequences from primary central nervous system lymphomas indicate derivation from highly mutated germinal center B cells with ongoing mutational activity. *Blood* 1999; **94**: 1738–1746.
- Montesinos-Rongen M, Van Roost D, Schaller C, Wiestler OD, Deckert M. Primary diffuse large B-cell lymphomas of the central nervous system are targeted by aberrant somatic hypermutation. *Blood* 2004; **103**: 1869–1875.
- Montesinos-Rongen M, Schäfer E, Siebert R, Deckert M. Genes regulating the B cell receptor pathway are recurrently mutated in primary central nervous system lymphoma. *Acta Neuropathol* 2012; **124**: 905–906.
- Montesinos-Rongen M, Godlewska E, Brunn A, Wiestler OD, Siebert R, Deckert M. Activating L265P mutations of the MYD88 gene are common in primary central nervous system lymphoma. *Acta Neuropathol* 2011; **122**: 791–792.
- Montesinos-Rongen M, Schmitz R, Brunn A, Gesk S, Richter J, Hong K *et al.* Mutations of CARD11 but not TNFAIP3 may activate the NF-kappaB pathway in primary CNS lymphoma. *Acta Neuropathol* 2010; **120**: 529–535.
- Schwindt H, Vater I, Kreuz M, Montesinos-Rongen M, Brunn A, Richter J *et al.* Chromosomal imbalances and partial uniparental disomies in primary central nervous system lymphoma. *Leukemia* 2009; **23**: 1875–1884.
- Courts C, Montesinos-Rongen M, Martin-Subero JL, Brunn A, Siemer D, Zühlke-Jenisch R *et al.* Transcriptional profiling of the nuclear factor-kappaB pathway identifies a subgroup of primary lymphoma of the central nervous system with low BCL10 expression. *J Neuropathol Exp Neurol* 2007; **66**: 230–237.
- Gonzalez-Aguilar A, Idhah A, Boisselier B, Habbita N, Rossetto M, Laurence A *et al.* Recurrent mutations of MYD88 and TBL1XR1 in primary central nervous system lymphomas. *Clin Cancer Res* 2012; **18**: 5203–5211.
- Riemersma SA, Jordanova ES, Schop RF, Philippo K, Looijenga LH, Schuurin E *et al.* Extensive genetic alterations of the HLA region, including homozygous deletions of HLA class II genes in B-cell lymphomas arising in immune-privileged sites. *Blood* 2000; **96**: 3569–3577.
- Rubenstein JL, Fridlyand J, Shen A, Aldape K, Ginzinger D, Batchelor T *et al.* Gene expression and angiogenesis in primary CNS lymphoma. *Blood* 2006; **107**: 3716–3723.
- Tun HW, Personett D, Baskerville KA, Menke DM, Jaecle KA, Kreinest P *et al.* Pathway analysis of primary central nervous system lymphoma. *Blood* 2008; **111**: 3200–3210.
- Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A *et al.* Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; **403**: 503–511.
- Jima DD, Zhang J, Jacobs C, Richards KL, Dunphy CH, Choi WW *et al.* Deep sequencing of the small RNA transcriptome of normal and malignant human B cells identifies hundreds of novel microRNAs. *Blood* 2010; **116**: e118–e127.
- Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C *et al.* Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. *Proc Natl Acad Sci USA* 2012; **109**: 3879–3884.
- Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD *et al.* Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. *Nature* 2011; **476**: 298–303.
- Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD *et al.* Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. *Blood* 2013; **122**: 1256–1265.
- Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A *et al.* Analysis of the coding genome of diffuse large B-cell lymphoma. *Nat Genet* 2011; **43**: 830–837.
- Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV *et al.* Signatures of mutational processes in human cancer. *Nature* 2013; **500**: 415–421.
- Deckert M, Paulus W. Malignant Lymphomas. In: Louis DN, Ohgaki H, Wiestler OD, Cavenee WK (eds) *WHO Classification of Tumors Pathology & Genetics of Tumors of the Nervous System* 4th edn. IARC: Lyon, 2007; pp 188–192.
- Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ *et al.* Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat Genet* 2013; **45**: 927–932.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 2009; **25**: 1754–1760.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009; **25**: 2078–2079.

- 33 Rimmer A, Mathieson I, Lunter G, McVean G. Platypus: An Integrated Variant Caller 2012 [cited 2014 January 25, 2014]; Available from www.well.ox.ac.uk/platypus.
- 34 Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 2010; **38**: e164.
- 35 Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M *et al*. Sequencing newly replicated DNA reveals widespread plasticity in human replication timing. *Proc Natl Acad Sci USA* 2010; **107**: 139–144.
- 36 Floratos A, Smith K, Ji Z, Watkinson J, Califano A. geWorkbench: an open source platform for integrative genomics. *Bioinformatics* 2010; **26**: 1779–1780.
- 37 Brunn A, Nagel I, Montesinos-Rongen M, Klapper W, Vater I, Paulus W *et al*. Frequent triple-hit expression of MYC, BCL2, and BCL6 in primary lymphoma of the central nervous system and absence of a favorable MYC(low)BCL2 (low) subgroup may underlie the inferior prognosis as compared to systemic diffuse large B cell lymphomas. *Acta Neuropathol* 2013; **126**: 603–605.
- 38 Fontebasso AM, Schwartzentruber J, Khuong-Quang DA, Liu XY, Sturm D, Korshunov A *et al*. Mutations in SETD2 and genes affecting histone H3K36 methylation target hemispheric high-grade gliomas. *Acta Neuropathol* 2013; **125**: 659–669.
- 39 Pasqualucci L, Neumeister P, Goossens T, Nanjangud G, Chaganti RS, Kuppers R *et al*. Hypermutation of multiple proto-oncogenes in B-cell diffuse large-cell lymphomas. *Nature* 2001; **412**: 341–346.
- 40 Sima J, Gilbert DM. Complex correlations: replication timing and mutational landscapes during cancer and genome evolution. *Curr Opin Genet Dev* 2014; **25C**: 93–100.
- 41 Bruno A, Boisselier B, Labreche K, Marie Y, Polivka M, Jouvet A *et al*. Mutational analysis of primary central nervous system lymphoma. *Oncotarget* 2014; **5**: 5065–5075.
- 42 Braggio E, McPhail ER, Macon W, Lopes MB, Schiff D, Law M *et al*. Primary central nervous system lymphomas: a validation study of array-based comparative genomic hybridization in formalin-fixed paraffin-embedded tumor specimens. *Clin Cancer Res* 2011; **17**: 4245–4253.
- 43 Sung CO, Kim SC, 3569- S, Karube K, Shin HJ, Nam DH *et al*. Genomic profiling combined with gene expression profiling in primary central nervous system lymphoma. *Blood* 2011; **117**: 1291–1300.
- 44 Mottok A, Renne C, Seifert M, Oppermann E, Bechstein W, Hansmann ML *et al*. Inactivating SOCS1 mutations are caused by aberrant somatic hypermutation and restricted to a subset of B-cell lymphoma entities. *Blood* 2009; **114**: 4503–4506.
- 45 Brunn A, Utermöhlen O, Mihelcic M, Sanchez-Ruiz M, Carstov M, Blau T *et al*. Differential effects of CXCR4-CXCL12- and CXCR7-CXCL12-mediated immune reactions on murine P0106-125 -induced experimental autoimmune neuritis. *Neuropathol Appl Neurobiol* 2013; **39**: 772–787.
- 46 Horn H, Ziepert M, Becher C, Barth TFE, Bernd H-W, Feller AC *et al*. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. *Blood* 2013; **121**: 2253–2263.
- 47 Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S *et al*. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. *J Clin Oncol* 2012; **30**: 3452–3459.
- 48 Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB *et al*. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. *Nature* 2010; **463**: 88–92.
- 49 Booman M, Douwes J, Glas AM, Riemersma SA, Jordanova ES, Kok K *et al*. Mechanisms and effects of loss of human leukocyte antigen class II expression in immune-privileged site-associated B-cell lymphoma. *Clin Cancer Res* 2006; **12**: 2698–2705.
- 50 Booman M, Szuhai K, Rosenwald A, Hartmann E, Kluin-Nelemans H, de Jong D *et al*. Genomic alterations and gene expression in primary diffuse large B-cell lymphomas of immune-privileged sites: the importance of apoptosis and immunomodulatory pathways. *J Pathol* 2008; **216**: 209–217.
- 51 Jordanova ES, Philippo K, Giphart MJ, Schuurin E, Kluin PM. Mutations in the HLA class II genes leading to loss of expression of HLA-DR and HLA-DQ in diffuse large B-cell lymphoma. *Immunogenetics* 2003; **55**: 203–209.
- 52 Kenzelmann-Broz D, Tucker RP, Leachman NT, Chiquet-Ehrismann R. The expression of teneurin-4 in the avian embryo: potential roles in patterning of the limb and nervous system. *Int J Dev Biol* 2010; **54**: 1509–1516.
- 53 Suzuki N, Fukushi M, Kosaki K, Doyle AD, de Vega S, Yoshizaki K *et al*. Teneurin-4 is a novel regulator of oligodendrocyte differentiation and myelination of small-diameter axons in the CNS. *J Neurosci* 2012; **32**: 11586–11599.
- 54 Psychiatric GCBWDG. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet* 2011; **43**: 977–983.
- 55 Arregui CO, Balsamo J, Lilien J. Regulation of signaling by protein-tyrosine phosphatases: potential roles in the nervous system. *Neurochem Res* 2000; **25**: 95–105.
- 56 Sun QL, Wang J, Bookman RJ, Bixby JL. Growth cone steering by receptor tyrosine phosphatase delta defines a distinct class of guidance cue. *Mol Cell Neurosci* 2000; **16**: 686–695.
- 57 Boeva V, Jouannet S, Daveau R, Combaret V, Pierre-Eugène C, Cazes A *et al*. Breakpoint features of genomic rearrangements in neuroblastoma with unbalanced translocations and chromothripsis. *PLoS One* 2013; **8**: e72182.
- 58 Molenaar JJ, Koster J, Zwiijnenburg DA, van Sluis P, Valentijn LJ, van der Ploeg I *et al*. Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* 2012; **483**: 589–593.
- 59 Shimizu A, Asakawa S, Sasaki T, Yamazaki S, Yamagata H, Kudoh J *et al*. A novel giant gene CSMD3 encoding a protein with CUB and sushi multiple domains: a candidate gene for benign adult familial myoclonic epilepsy on human chromosome 8q23.3-q24.1. *Biochem Biophys Res Commun* 2003; **309**: 143–154.
- 60 Vinatzer U, Gollinger M, Mullauer L, Raderer M, Chott A, Streubel B. Mucosa-associated lymphoid tissue lymphoma: novel translocations including rearrangements of ODZ2, JMJD2C, and CNN3. *Clin Cancer Res* 2008; **14**: 6426–6431.