The *PCBP1* Gene Encoding Poly(rC) Binding Protein I Is Recurrently Mutated in Burkitt Lymphoma

Rabea Wagener,[1†] Sietse M. Aukema,[1†] Matthias Schlesner,[2] Andrea Haake,[1] Birgit Burkhardt,[3] Alexander Claviez,[4] Hans G. Drexler,[5] Michael Hummel,[6] Markus Kreuz,[7] Markus Loeffler,[7] Maciej Rosolowski,[7] Cristina López,[1] Peter Möller,[8] Julia Richter,[1] Marius Rohde,[9] Matthew J. Betts,[10] Robert B. Russell,[10] Stephan H. Bernhart,[11] Steve Hoffmann,[11] Philip Rosenstiel,[12] Markus Schilhabel,[12] Monika Szczepanowski,[13] Lorenz Trümper,[14] Wolfram Klapper,[13] and Reiner Siebert,[1]* on behalf of the ICGC MMML-Seq-Project[‡] and the "Molecular Mechanisms in Malignant Lymphomas" Network Project of the Deutsche Krebshilfe[‡]

[1]Institute of Human Genetics, Christian-Albrechts-University Kiel and University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany
[2]Deutsches Krebsforschungszentrum Heidelberg (DKFZ), Division Theoretical Bioinformatics, Heidelberg, Germany
[3]Non-Hodgkin Lymphoma Berlin-Frankfurt-Münster Group Study Center, Department of Pediatric Hematology and Oncology, University Children's Hospital, Münster, Germany
[4]Department of Pediatrics, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University, Kiel, Germany
[5]Leibniz-Institute DSMZ- German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany
[6]Institute of Pathology, Campus Benjamin Franklin, Charité–Universitätsmedizin, Berlin, Germany
[7]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Germany
[8]Institute of Pathology, Universitätsklinikum Ulm, Ulm, Germany
[9]Department of Pediatric Hematology and Oncology, Justus Liebig University, Giessen, Germany
[10]Cell Networks, Bioquant, University of Heidelberg, Heidelberg, Germany
[11]Transcriptome Bioinformatics, LIFE Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany
[12]Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University Kiel, Kiel, Germany
[13]Institute of Hematopathology, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University Kiel, Germany
[14]Department of Hematology and Oncology, Georg-August University of Göttingen, Germany

The genetic hallmark of Burkitt lymphoma is the translocation t(8;14)(q24;q32), or one of its light chain variants, resulting in *IG-MYC* juxtaposition. However, these translocations alone are insufficient to drive lymphomagenesis, which requires additional genetic changes for malignant transformation. Recent studies of Burkitt lymphoma using next generation sequencing approaches have identified various recurrently mutated genes including *ID3, TCF3, CCND3*, and *TP53*. Here, by using similar approaches, we show that *PCBP1* is a recurrently mutated gene in Burkitt lymphoma. By whole-genome sequencing, we identified somatic mutations in *PCBP1* in 3/17 (18%) Burkitt lymphomas. We confirmed the recurrence of *PCBP1* mutations by Sanger sequencing in an independent validation cohort, finding mutations in 3/28 (11%) Burkitt lymphomas and in 6/16 (38%) Burkitt lymphoma cell lines. *PCBP1* is an intron-less gene encoding the 356 amino acid poly(rC) binding protein 1, which contains three K-Homology (KH) domains and two nuclear localization signals. The mutations predominantly (10/12, 83%) affect the KH III domain, either by complete domain loss or amino acid changes. Thus, these changes are predicted to alter the various functions of PCBP1, including nuclear trafficking and pre-mRNA splicing.

Additional Supporting Information may be found in the online version of this article.

†Rabea Wagener and Sietse M. Aukema contributed equally to this work.

‡Complete list of members of the German ICGC MMML-Seq Project and MMML-Network Project appear in the Supporting Information.

Supported by: German Ministry for Education and Science (BMBF; framework of the ICGC MMML-Seq), Grant numbers: 01KU1002A to 01KU1002J; E:bio Network MMML-MYC-SYS, Grant number: 0316166; Deutsche Krebshilfe (the Network project "Molecular Mechanisms in Malignant Lymphomas"), Grant number: 70-3173-Tr3; Christoph-Schubert-Award of the KinderKrebsInitiative Buchholz/Holm-Seppensen (to R.W.); JSM-UMCG MD-PhD program, recipient of the "Nijbakker-Morra" and "Hippocrates" Foundations awards, "Foundation de Drie Lichten", the "René Vogels' Foundation" and in the framework of a "JSM-Ubbo Emmius Foundation Talent Grant" (to S.M.A.); Dr. Werner Jackstädt Foundation in the framework of a Junior Excellence Research Group on "Mechanisms of B-cell lymphomagenesis in the Senium as basic principle for the development of age adjusted therapy regimes" Grant number: S134 - 10.100 (to J.R.); Alexander von Humboldt Foundation (to C.L.); NHL-BFM study center (Gießen) from the "Forschungshilfe Station Peiper."

*Correspondence to: Prof. Dr. Reiner Siebert, Institute of Human Genetics, Christian-Albrechts-University Kiel & University Hospital Schleswig Holsteins, Campus Kiel, Schwanenweg 24, D-24105 Kiel, Germany. E-mail: rsiebert@medgen.unikiel.de

Remarkably, all six primary Burkitt lymphomas with a *PCBP1* mutation expressed MUM1/IRF4, which is otherwise detected in around 20–40% of Burkitt lymphomas. We conclude that *PCBP1* mutations are recurrent in Burkitt lymphomas and might contribute, in cooperation with other mutations, to its pathogenesis. © 2015 Wiley Periodicals, Inc.

INTRODUCTION

Burkitt lymphoma is the most common B-cell lymphoma in children and also accounts for 1–5% of lymphomas in the adult population (Boerma et al., 2004; Leoncini et al., 2008; Burkhardt et al., 2011; Linch, 2012; Miles et al., 2012). Its genetic hallmark is the translocation t(8;14)(q24;q32) involving the *MYC* gene and the immunoglobulin heavy chain (*IGH*) gene or its corresponding light chain variants involving the *IGK* or *IGL* locus. These translocations result in deregulated *MYC* expression by bringing the *MYC* oncogene under the influence of enhancers of the immunoglobulin genes (Dalla-Favera et al., 1982; Bertrand et al., 2007; Aukema et al., 2014). Burkitt lymphoma is characterized by an overall low genomic complexity (Boerma et al., 2009; Scholtysik et al., 2010; Lundin et al., 2013; Aukema et al., 2014). In up to 40% of the tumors the *IG-MYC* translocation is the sole karyotype abnormality (Berger et al., 1989; Johansson et al., 1995; Boerma et al., 2009). However, the *IG-MYC* translocation alone is insufficient to induce Burkitt lymphoma and for the full transformation to overt lymphoma additional genetic hits are required (Sander et al., 2012; Lundin et al., 2013; Bauer et al., 2014). Indeed, using whole-genome and exome sequencing, recurrent mutations in for example *ID3*, *TCF3* or *CCND3* have recently been identified in Burkitt lymphoma (Love et al., 2012; Richter et al., 2012; Schmitz et al., 2012). Here, we show that *PCBP1*, encoding the poly(rC) binding protein 1, is also recurrently mutated in Burkitt lymphoma.

MATERIALS AND METHODS

ICGC MMML-Seq Exploration Cohort

We obtained pretreatment tumor tissue and germline material (peripheral blood, buffy coats) from 17 patients with Burkitt lymphoma/leukemia (hereafter, Burkitt lymphoma). Written informed consent, either from patients or parents/caretakers, was obtained according to the International Cancer Genome Consortium (ICGC) guidelines (www.icgc.org). We showed the 17 Burkitt lymphomas to be *IG-MYC* positive and *BCL2* and *BCL6* rearrangement negative by fluorescence in situ hybridization (FISH).The ICGC MMML-Seq study was approved by the Institutional Review Boards of the Medical Faculty of the University of Kiel (A150/10) and of the recruiting centers. Experimental protocols for whole-genome and transcriptome sequencing as well as for the bioinformatical analyses including SNV calling followed those previously published by Richter et al. (2012). Only nonsynonymous mutations without an SNP annotation (dbSNP build 141) were included in the analyses. Pathology review was performed by expert reference hematopathologists (Hummel et al., 2006; Klapper et al., 2012).

MMML-Validation Cohort

For validation, we analyzed the *PCBP1* gene in an independent cohort of 28 *IG-MYC* positive, *BCL2*, and *BCL6* rearrangement negative, molecularly defined Burkitt lymphomas (Hummel et al., 2006; Richter et al., 2012), which have previously been characterized in the framework of the Molecular Mechanisms in Malignant Lymphoma (MMML) network project. The MMML protocols have been approved centrally by the Institutional Review Board of the coordination center in Göttingen, Germany, and local centers (Hummel et al., 2006; Klapper et al., 2012; Aukema et al., 2014). All lymphomas were diagnosed by a panel of expert hematopathologists. Molecular cytogenetic (FISH) and immunohistochemical analyses were performed as previously published (Hummel et al., 2006; Salaverria et al., 2011; Klapper et al., 2012; Aukema et al., 2014). For molecular classification, lymphomas with a molecular index score of $\geq$0.95 were assigned to the group of molecular Burkitt lymphomas (mBL) (Hummel et al., 2006).

Cell Lines

We Sanger-sequenced 16 Burkitt lymphoma cell lines (BALM-16, BALM-18, BL-2, BL-41, BL-30, BL-70, BLUE-1, CA-46, DAUDI, DG-75, EB-1, EB-3, NAMALWA, RAJI, RAMOS, U-698-M) to detect mutations within the *PCBP1* gene. The majority of cell lines were obtained via the *DSMZ*—German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. For an overview of the cell lines including the BALM-16, BALM-18, and EB-3 cell lines contributed by

TABLE 1. Mutations Identified in *PCBP1* in Burkitt Lymphoma

Cohort	PID	Age[a]	Sex	*MYC*-status	Genomic position (hg19) in bp	Mutation (cDNA)	Mutation (protein)	Affected domain
ICGC MMML-Seq[b]	1	17	M	*IGK-MYC*	Chr2:70314913-70314914	c.41_42delTC	p.L14fs*22	Truncating in/after KHI
	2	16	F	*IGH-MYC*	Chr2:70315476	c.601_602insT	p.C201Lfs*12	Truncating after KHII
	3	8	M	*IGH-MYC*	Chr2:70315425	c.550C>T	p.Q184*	Truncating after KHII
MMML	4	40	M	*IGH-MYC*	Chr2:70315753	c.878G>A	p.C293Y	KHIII
	5	5	M	*IGH-MYC*	Chr2:70315759	c.884T>C	p.I295T	KHIII
	6	76	M	*IGH-MYC*	Chr2:70315848	c.973A>G	p.R325G	KHIII/ NLSII
Cell lines	BL-30[c]	19	M	*IGH-MYC*[d]	Chr2:70315779	c.907A>G	p.N303D	KHIII
	BL-41	8	M	*IGH-MYC*[d]	Chr2:70315674	c.799G>A	p.V267M	NLSI
	BL-70[c]	16	M	*IGH-MYC*[d]	Chr2:70315753	c.878G>A	p.C293Y	KHIII
	DAUDI [c]	16	M	*IGH-MYC*[d]	Chr2:70315312	c.437T>G	p.I146S	KHII
	EB-3[c]	3	M	*IGH-MYC*[e]	Chr2:70315424	c.549C>G	p.Y183*	Truncating after KHII
	NAMALWA[c]	3	F	*IGH-MYC*[d]	Chr2:70315236	c.361A>T	p.K121*	Truncating after KHII

PID, patient ID; Sex, M for male and F for female; Genomic position, affected nucleotide position based on GRCh37/hg19; Mutation (cDNA), affected nucleotide in cDNA based on transcript NM_006196. KH, K-Homology domain; NLS, nuclear localization signal. All mutations were confirmed as being somatic mutations (ICGC MMML-Seq cohort) or as being likely somatic due to their absence from the dbSNP and ESP6500 databases.
[a]Age refers to age at (primary) diagnosis in years.
[b]The somatic origin of these mutations were proven by whole-genome sequencing of the germline tissue and verified by Sanger sequencing.
[c]Previously also reported in literature (Love et al., 2012; Schmitz et al., 2012).
[d]Karyotypes and *MYC*-status from Murga-Penas et al. (Murga Penas et al., 2014).
[e]*MYC*-status from Berger and Bernheim, (Berger and Bernheim, 1985).

Prof. Dr. H. G. Drexler we refer to Matsuo et al., 1997; Matsuo et al., 2002; Drexler, 2010. The BL-30 cell line was kindly provided by Dr. R. Schmitz, National Cancer Institute, NIH, Bethesda, MD. The identity of all cell lines has been validated by STR-profiling using the Stem Elite ID Kit (Promega, Madison, WI).

Extraction of DNA and Mutational Analyses

We extracted DNA from tumor tissue, blood, and buffy coat of the ICGC MMML-Seq exploration cohort using Genomic-tip 100/G (Qiagen, Hilden, Germany). DNA from tumor tissue of the MMML-validation cohort was extracted using QIAamp DNA Mini Prep kit (Qiagen). DNA from the cell lines was extracted using Gentra Puregene Cell kit (Qiagen). The complete *PCBP1* gene consisting of one exon was amplified by PCR using AccuPrime Polymerase System (Life Technologies, Carlsband). To validate the mutations in *HNRNPD*, *SF3B3*, and *SNRPB2* identified by whole-genome sequencing, PCR assays were designed to cover the site of mutation. PCR primers and PCR conditions for Sanger sequencing of *PCBP1*, as well as the sites of mutations in *HNRNPD*, *SF3B3*, and *SNRPB2* are listed in the Supporting Information Table 1. PCR products were subjected to Sanger sequencing using the Big Dye Terminator v1.1 Cycle Sequencing Kit (Life Technologies). Sequence analysis was performed using an ABI PRISM 3130 Genetic Analyzer. Identified mutations were compared with the dbSNP database (build 141) and the Exome Variant Server, NHLBI Exome Sequencing Project (ESP) (data release ESP6500SI-V2) for identification of known single nucleotide polymorphisms. We used only nonsynonymous mutations for the follow up analyses.

Statistical and Bioinformatical Analysis

We analyzed differences in categorical variables between groups using Fisher's exact test. Comparison of gene expression between *PCBP1*-mutated ($n = 3$) and *PCBP1*-wild-type ($n = 25$) Burkitt lymphomas within the MMML cohort was computed using available gene expression data from the Affymetrix GeneChip U133A microarray (Hummel et al., 2006). We used the R package *limma* (Smyth, 2004) to compute differentially expressed genes and the nominal *P*-values were adjusted for multiple testing with the Benjamini–Hochberg method (Benjamini and Hochberg, 1995).

Protein Structure Analysis

We used Mechismo (Betts et al., 2015) to predict the effects of missense mutations on PCBP1 interactions, which identified the structure of a single strand of poly(C) DNA bound to the human

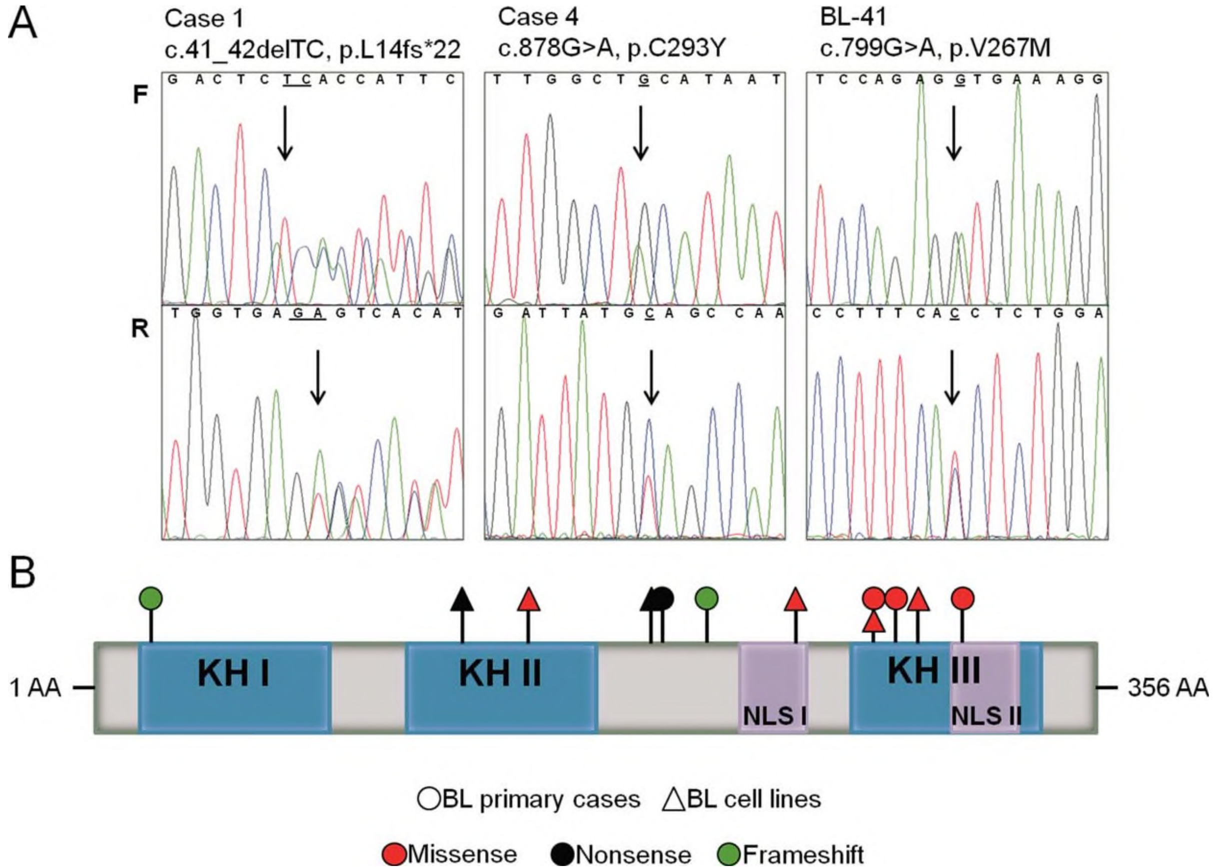

Figure 1. Overview on *PCBP1* mutations. (A) Representative electropherograms of Sanger sequences of sites of *PCBP1* mutation in primary Burkitt lymphomas (Case 2, Case 4) and a Burkitt lymphoma cell line (BL-41). F indicates sequence derived from forward primer and R from reverse primer. Arrows, as well as the underlined letters mark the site of mutation. (B) Localization of the *PCBP1* mutations identified in the analyzed Burkitt lymphoma cohorts and Burkitt lymphoma cell lines with regard to the PCBP1 protein subdomains. KH indicates K-homology domain; NLS, nuclear localization signal. The numbers indicate the amino acids (AA).

paralogue "PCBP2" (>90% sequence identity to PCBP1) as a template (Fenn et al., 2007).

RESULTS

Identification and Verification of *PCBP1* Mutations within the ICGC MMML-Seq Cohort

By exploring whole-genome sequences of four previously published prototypic pediatric *IG-MYC*-positive molecularly defined Burkitt lymphomas (Richter et al., 2012) and 13 additional *IG-MYC*-positive Burkitt lymphomas included in the ICGC-MMML-Seq study, we identified somatic mutations in the *PCBP1* gene in three of 17 (18%) (Fig. 1). All three *PCBP1* mutations, including their somatic origin as determined by their absence in the corresponding germ-line samples, were confirmed using Sanger sequencing (Fig. 1). The mutated allele frequencies (MAFs) ranged from 0.35 to 0.63 and, considering a tumor cell content of ≥80%, are mostly in line with a heterozygous mutation being present in all or the vast majority of the neoplastic cells (Supporting Information Table 2). Moreover, using the whole-genome sequencing data and SNP-array data (Scholtysik et al., 2010, COSMIC cell line database, http://cancer.sanger.ac.uk/cell_lines/conan/search, accessed 20.03.2015) deletions of the remaining *PCBP1* allele could be excluded in the Burkitt lymphomas as well in all Burkitt lymphoma cell lines with respective data. Screening of the available whole-genome sequencing data of 28 non-Burkitt lymphomas in the ICGC data portal (https://dcc.icgc.org/, accessed 06.10.2014) showed that nonsynonymous *PCBP1* mutations were not present.

Identification of *PCBP1* Mutations within the MMML Validation Cohort and Burkitt Lymphoma Cell Lines

To identify additional cases with mutations and assess the frequency of *PCBP1* mutations in an

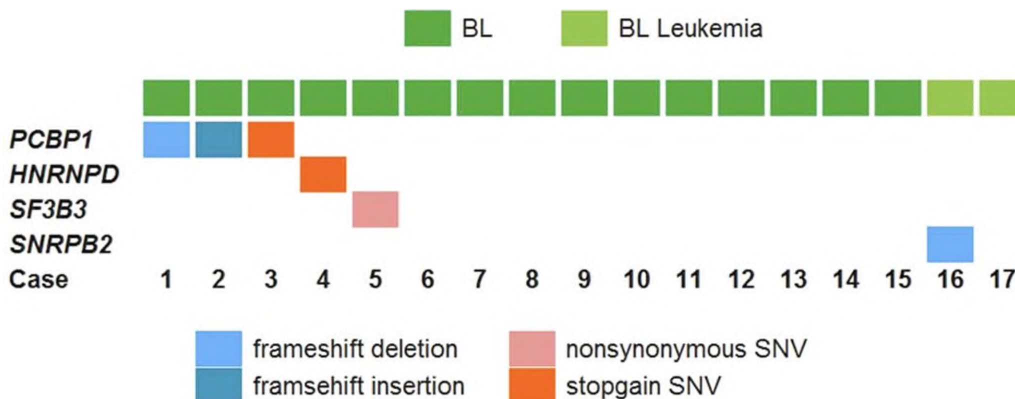

Figure 2. Heatmap showing the interplay of *PCBP1* mutations with *HNRNPD*, *SF3B3*, and *SNRPB2* mutations in 17 primary Burkitt lymphoma analyzed by whole-genome sequencing. Note that case numbers 1–17 refer to cases analyzed by whole-genome sequencing.

independent series, we screened a previously characterized validation cohort of the MMML project consisting of 28 *IG-MYC*-positive molecularly defined Burkitt lymphomas. In this cohort, three Burkitt lymphomas carried mutations in *PCBP1* (3/28, 11%). These mutations were absent from the dbSNP and ESP6500 databases, suggesting their somatic origin despite the lack of corresponding germ-line material. Thus, the overall incidence of *PCBP1* mutations in the discovery and validation cohorts combined was six out of 45 (13%) Burkitt lymphomas (Table 1). Additionally, six of the 16 (38%) investigated Burkitt lymphoma cell lines carried mutations within *PCBP1* which were also absent from the dbSNP and ESP6500 databases.

Identification and Verification of Mutations in Genes from the PCBP1 Network

To investigate the interplay of *PCBP1* with other genes, we explored the whole-genome sequencing data for protein-changing somatic mutations in genes encoding proteins described to interact with PCBP1 (Supporting Information Fig. 1). Three Burkitt lymphomas, all lacking a *PCBP1* mutation, showed somatic mutations in either *HNRNPD*, *SF3B3*, or *SNRPB2* (Fig. 2).

Investigation of Functional Changes at the Protein Level

Seven of the twelve identified mutations were missense mutations. Interestingly an identical mutation, c.878G>A/p.C293Y, was detected in Case 4 and in the Burkitt lymphoma cell line BL-70. Three of the 12 mutations were nonsense mutations; two of them affected the adjacent amino acids p.Y183 and p.Q184. Two cases carried frameshift mutations due to a deletion or insertion.

The PCBP1 protein consists of three K-Homology (KH) domains and two nuclear localization signals (NLS). Ten out of twelve mutations (83%) affected the KHIII domain either due to complete loss of the domain (5/10) or an amino acid exchange (5/10). Furthermore, 7 out of 12 mutations (58%) affected the NLSI and/or NLSII either due to loss of both (5/12) or a missense mutation (2/12) (Fig. 1B).

Of the six missense mutations, one (p.I146S) was in KHII and the remaining five were either in or next to KHIII. We predict all but p.V267M to have an effect on DNA/RNA binding. Hence, p.R325G was predicted to have a strongly disabling effect on binding to oligonucleotides (by the loss of the positive charge), with the others (p.C293Y, p.I295T, and p.N303D) having weaker predicted deleterious effects (Fig. 3).

Correlation of Mutations and Expression

To determine if the mutations of *PCBP1* had an impact on the expression of the gene, we analyzed available transcriptome data from the ICGC MMML-Seq cohort. Transcriptome analysis of Cases 2 and 3 showed a balanced expression of the wild-type and the mutated allele (relative expression of the mutated allele Case 2: 0.47; Case 3: 0.55; Supporting Information Table 2). Case 1 on the other hand, carrying a mutation leading to a truncation in/after KH I domain (p.L14fs*22), showed a more than fivefold lower relative expression of the mutated allele compared to the wild-type allele, likely due to nonsense-mediated decay

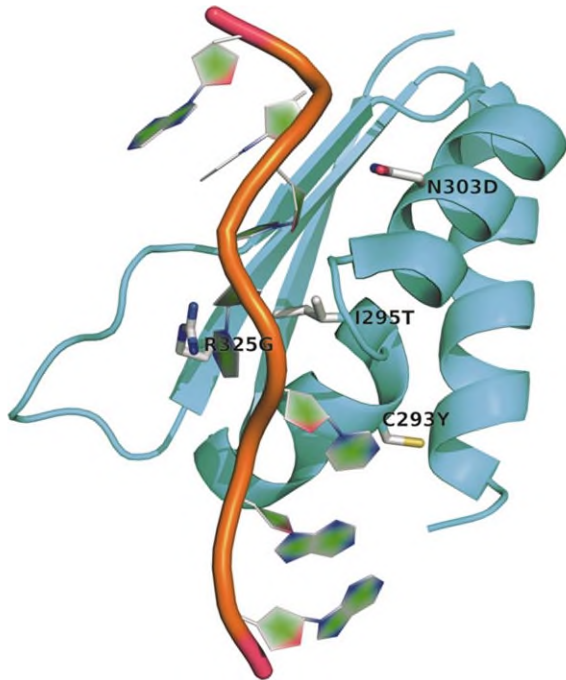

Figure 3. *PCBP1* mutations (sticks coloured by atom type) mapped to the structure of the paralogue PCBP2 (cyan ribbons) bound to a single-strand of poly(C) DNA (orange tube with nucleotide bases)(Fenn et al., 2007). P.R325G is predicted to have strongly disabling effect on oligonucleotides binding while the others (p.C293Y, p.I295T, and p.N303D) were predicted to have weaker effects. PCBP2 and PCBP1 are identical at these positions.

(number of reads for wild-type allele: 1899, for mutated allele: 291; Supporting Information Table 2).

Comparison of gene expression between Burkitt lymphomas with *PCBP1* mutation and *PCBP1* wild-type of the ICGC cohort using the available transcriptome data, as well of the MMML cohort using the gene expression data of the U133A microarray (Hummel et al., 2006) revealed no significantly differentially expressed genes (data not shown). However, due to the small sample size of mutated cases, the power is limited and genes with moderate differential expression might not be identified.

Clinico-Pathologic Characteristics

Patients with Burkitt lymphomas carrying mutations in *PCBP1* were in the majority male (5/6, 83%), as expected by the distribution in the overall cohort. All *PCBP1*-mutated cases had a CD10 and MUM1/IRF4 positive and CD5 negative immunophenotype. Five out of six cases (83%) expressed BCL6. Except for a higher frequency of MUM1 expression (6/6, 100% in mutated cases versus 12/32, 39% in wild-type cases, $P = 0.008$, Supporting Information Table 3) no significant differences were seen in the distribution of the other clinico-pathologic and molecular cytogenetic characteristics between *PCBP1* mutated and wild-type cases.

DISCUSSION

Although Burkitt lymphoma is a highly aggressive tumor it has an overall low genomic complexity with up to 40% lacking any secondary cytogenetic changes in addition to the *IG-MYC* translocation (Berger and Bernheim, 1985; Johansson et al., 1995; Boerma et al., 2009; Scholtysik et al., 2010; Aukema et al., 2014). Moreover, the *IG-MYC* translocation has also been detected in the peripheral blood of healthy individuals and in murine Peyer's patches, suggesting that secondary events as somatic mutations might be necessary for full blown lymphomagenesis (Müller et al., 1995, 1997a, 1997b; Kovalchuk et al., 2000; Bauer et al., 2014). Indeed, several studies identified recurrent secondary genetic changes in Burkitt lymphoma, as, for example, mutations in *ID3* and/or *TCF3* which were present in up to 70% of Burkitt lymphomas (Love et al., 2012; Richter et al., 2012; Schmitz et al., 2012). We here show, in two independent cohorts of sporadic Burkitt lymphoma, that the *PCBP1* gene located in chromosomal band 2p13.3 is recurrently mutated in *IG-MYC*-positive Burkitt lymphoma affecting a total of 13% (6/45) of primary cases and 38% (6/16) of Burkitt lymphoma cell lines. Although mutations of *PCBP1* have been observed at a lower frequency in two other publications (Love et al., 2012; Schmitz et al., 2012), *PCBP1* has not been mentioned as a recurrently mutated gene in Burkitt lymphoma. The apparent but nevertheless not significant differences in incidence of detected mutations in Schmitz et al. (2/28, 7%) and Love et al. (2/51, 4%) might be explained by different sequencing methods used as well as composition of the cohorts (Love et al., 2012; Schmitz et al., 2012). Both the ICGC MMML-Seq exploration cohort and the MMML validation cohort consisted of solely *IG-MYC* positive, *BCL2*, and *BCL6* rearrangement negative Burkitt lymphomas. As far as applicable, all Burkitt lymphoma moreover expressed the mBL signature. By these selection criteria, we excluded lymphomas with morphological features of Burkitt lymphomas but not representing Burkitt lymphoma at the molecular level as, for example, some double-hit lymphomas

(Aukema et al., 2011, 2014; Salaverria and Siebert, 2011).

TCF3 mutations have been reported in 11% of Burkitt lymphomas (Schmitz et al., 2012) and the frequency of *PCBP1* mutations is comparable to that of *TCF3*. The *TCF3* mutations occurring at lower frequency have been described to interact with the more commonly mutated *ID3* in disruption of the ID3-TCF3-PI3K signaling pathway (Schmitz et al., 2012). Therefore it is conceivable that *PCBP1* (mutations) might also contribute to the lymphomagenesis of Burkitt lymphoma by interplay with other mutations. Indeed, by exploring the whole-genome sequencing data to investigate the interplay of *PCBP1* with other genes, we show three *PCBP1* wild-type cases to carry somatic mutations in *HNRNPD*, *SF3B3*, and *SNRPB2*. Thus, in total, 6/17 (35%) Burkitt lymphomas showed mutations in any of these four genes. Moreover, a very recent meta-analysis of pan-cancer exome data showed that most cancer related genes are mutated at frequencies between 2% and 20%. Remarkably, this analysis also identified *PCBP1* as a novel cancer related gene (Lawrence et al., 2014). However, in five of the six *PCBP1* mutated cases we identified mutations in at least one of the four genes *ID3*, *CCND3*, *TP53*, and *MYC*, which have been recently described as potential drivers in Burkitt lymphoma (Love et al., 2012; Richter et al., 2012; Schmitz et al., 2012)(Supporting Information Fig. 2). *PCBP1* mutations might contribute to Burkitt lymphoma pathogenesis in the context of this mutational landscape by affecting the splicing pathway.

PCBP1 encodes a member of the family of poly(C)-binding proteins which are characterized by their high and specific affinity for C-rich oligonucleotide sequences. They carry three KH domains that can act as independent nucleic acid-binding units resulting in a high complexity and specificity of RNA/DNA interaction possibility (Silvera et al., 1999; Thisted et al., 2001). Additionally, PCBP1 contains two nuclear localization signals (NLS) I-II (Chkheidze and Liebhaber, 2003). Whereas NLSI lies within the variable region between the KHII and KHIII domain, NLSII is located within the KHIII domain. The PCBP1 protein has yet been described to be involved in three major functions: transcriptional regulation of gene expression, translational control, and modulation of pre-mRNA splicing (Meng et al., 2007; Nishinakamura et al., 2007; Wang et al., 2010; Ren et al., 2014). However, the major determinant of PCBP1 function may be its subcellular localization. PCBP1 is ubiquitously expressed within the cell, which reflects its manifold function as transcriptional regulator within the nucleus and as translational regulator in the cytoplasm. Additionally it has been described that PCBP1 expression is concentrated in nuclear speckles (Chkheidze and Liebhaber, 2003; Berry et al., 2006), which represent sites at which splicing factors are concentrated (Misteli et al., 1997) reflecting its function in mRNA splicing.

We here identified six *PCBP1* mutations within six Burkitt lymphomas. Additionally, we describe six *PCBP1* mutations in Burkitt lymphoma cell lines including one not yet reported mutation in the BL-41 cell line. In addition, four *PCBP1* mutations have previously been described and in part been validated in Burkitt lymphoma by Schmitz et al. and Love et al. (Love et al., 2012; Schmitz et al., 2012)(Supporting Information Fig. 3). Interestingly, the same c.878G>A/p.C293Y mutation (affecting the KHIII domain by amino acid change) observed in one of our patients and in the cell line BL-70 was also detected in one patient in each paper (Love et al., 2012; Schmitz et al., 2012). Hence, the p.C293Y mutation accounts for 4/17 (24%) of *PCBP1* mutations in Burkitt lymphoma and Burkitt lymphoma cell lines detected in the present study and described in the literature (Supporting Information Fig. 3).

Ten of those mutations affect the KHIII domain either due to complete loss of the domain (5/10) or an amino acid exchange (5/10). Furthermore, 7 out of 12 mutations (58%) affect the NLSI and/or NLSII either due to loss of both (5/12) or an amino acid change in NLSI or NLSII (2/12). The functional consequence and severity of these mutations might differ between the mutation types as truncating mutations are generally more damaging than missense mutations. Despite the nonsense-mediated decay shown in Case 1, the mutated protein might be synthesized but have impaired nuclear trafficking owing to the loss of the NLS. In contrast, the two cases with missense mutations within either the NLSI or NLSII might lead solely to a decreased PCBP1 expression within the nucleus. As shown previously by Chkheidze and Liebhaber (2003) the depletion of only one NLS signal leads to the redistribution of some of the protein to the cytoplasm whereby the depletion of both signals leads to a loss of nuclear trafficking.

Our structural analysis of the mutations using the closely related PCBP2—single-strand DNA structure (Fenn et al., 2007) predicts that p.R325G

has a strongly disabling effect on binding to oligonucleotides while the others (p.C293Y, p.I295T, and p.N303D) have weaker effects.

Comparing clinicopathologic and molecular cytogenetic characteristics of *PCBP1* mutated and wild-type Burkitt lymphomas, it was noted that all mutated cases expressed MUM1/IRF4. In contrast, *PCBP1* wild-type cases expressed MUM1/IRF4 in only 39% of the cases which is in line with various reports describing MUM1 to be expressed in 20–40% of Burkitt lymphomas (Chuang et al., 2007; Gualco et al., 2009; Nasr et al., 2010; Lu et al., 2011; Ye et al., 2011). All *PCBP1* mutated Burkitt lymphomas expressed CD10 and five of six cases expressed BCL6. Certain mutations in DLBCL patients have been associated with particular immunophenotypes (Ngo et al., 2011; Bohers et al., 2014; Vaque et al., 2014), but the implications of our observation for *PCBP1* mutations and its putative association with MUM1 expression in Burkitt lymphoma are unclear at present and await confirmation in independent cohorts. Unfortunately, due to the limited number of cases and short follow-up, it was not possible to perform meaningful survival analyses.

In conclusion, we identified *PCBP1* as an additional recurrently mutated gene in Burkitt lymphoma affecting 13% of primary cases. Mutations seem to primarily affect the KHIII domain and/or NLS, whereby the consequences of the mutations point to a reduced or even loss-of-function.

ACKNOWLEDGMENTS

The authors thank the technical staff of the Institutes of Human Genetics Kiel, and Hematopathology/Lymph Node Registry, Kiel, Germany, for their excellent technical support.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Aukema SM, Siebert R, Schuuring E, van Imhoff GW, Kluin-Nelemans HC, Boerma E, Kluin PM. 2011. Double-hit B-cell lymphomas. Blood 117:2319–2331.

Aukema SM, Kreuz M, Kohler CW, Rosolowski M, Hasenclever D, Hummel M, Küppers R, Lenze D, Ott G, Pott C, Richter J, Rosenwald A, Szczepanowski M, Schwaenen C, Stein H, Trautmann H, Wessendorf S, Trümper L, Loeffler M, Spang R, Kluin PM, Klapper W, Siebert R. 2014. Biologic characterization of adult MYC-translocation positive mature B-cell lymphomas other than molecular Burkitt lymphoma. Haematologica 99:726–735.

Bauer B, Siebert R, Traulsen A. 2014. Cancer initiation with epistatic interactions between driver and passenger mutations. J Theor Biol 358:52–60.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Series B Methodol 57:289–300.

Berger R. and Bernheim A. 1985. Cytogenetics of Burkitt's lymphoma-leukemia: A review. IARC Sci Publ 60:65–80.

Berger R, Le Coniat M, Derre J, Vecchione D. 1989. Secondary nonrandom chromosomal abnormalities of band 13q34 in Burkitt lymphoma-leukemia. Genes Chromosomes Cancer 1:115–118.

Berry AM, Flock KE, Loh HH, Ko JL. 2006. Molecular basis of cellular localization of poly C binding protein 1 in neuronal cells. Biochem Biophys Res Commun 349:1378–1386.

Bertrand P, Bastard C, Maingonnat C, Jardin F, Maisonneuve C, Courel MN, Ruminy P, Picquenot JM, Tilly H. 2007. Mapping of MYC breakpoints in 8q24 rearrangements involving non-immunoglobulin partners in B-cell lymphomas. Leukemia 21: 515–523.

Betts MJ, Lu Q, Jiang Y, Drusko A, Wichmann O, Utz M, Valtierra-Gutiérrez IA, Schlesner M, Jaeger N, Jones DT, Pfister S, Lichter P, Eils R, Siebert R, Bork P, Apic G, Gavin A, Russell RB. 2015. Mechismo: Predicting the mechanistic impact of mutations and modifications on molecular interactions. Nucleic Acids Res 43:e10.

Boerma EG, van Imhoff GW, Appel IM, Veeger NJGM, Kluin PM, Kluin-Nelemans JC. 2004. Gender and age-related differences in Burkitt lymphoma—Epidemiological and clinical data from The Netherlands. Eur J Cancer 40:2781–2787.

Boerma EG, Siebert R, Kluin PM, Baudis M. 2009. Translocations involving 8q24 in Burkitt lymphoma and other malignant lymphomas: A historical review of cytogenetics in the light of todays knowledge. Leukemia 23:225–234.

Bohers E, Mareschal S, Bouzelfen A, Marchand V, Ruminy P, Maingonnat C, Ménard A, Etancelin P, Bertrand P, Dubois S, Alcantara M, Bastard C, Tilly H, Jardin F. 2014. Targetable activating mutations are very frequent in GCB and ABC diffuse large B-cell lymphoma. Genes Chromosomes Cancer 53:144–153.

Burkhardt B, Oschlies I, Klapper W, Zimmermann M, Woessmann W, Meinhardt A, Landmann E, Attarbaschi A, Niggli F, Schrappe M, Reiter A. 2011. Non-Hodgkin's lymphoma in adolescents: Experiences in 378 adolescent NHL patients treated according to pediatric NHL-BFM protocols. Leukemia 25:153–160.

Chkheidze AN, Liebhaber SA. 2003. A novel set of nuclear localization signals determine distributions of the aCP RNA-binding proteins. Mol Cell Biol 23:8405–8415.

Chuang SS, Ye H, Du MQ, Lu CL, Dogan A, Hsieh PP, Huang WT, Jung YC. 2007. Histopathology and immunohistochemistry in distinguishing Burkitt lymphoma from diffuse large B-cell lymphoma with very high proliferation index and with or without a starry-sky pattern: A comparative study with EBER and FISH. Am J Clin Pathol 128:558–564.

Dalla-Favera R, Bregni M, Erikson J, Patterson D, Gallo RC, Croce CM. 1982. Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells. Proc Natl Acad Sci USA 79:7824–7827.

Drexler H.G. 2010. Guide to Leukemia-Lymphoma Cell Lines, 2nd ed. Braunschweig.

Fenn S, Du Z, Lee JK, Tjhen R, Stroud RM, James TL. 2007. Crystal structure of the third KH domain of human poly(C)-binding protein-2 in complex with a C-rich strand of human telomeric DNA at 1.6 Å resolution. Nucleic Acids Res 35:2651–2660.

Gualco G, Queiroga EM, Weiss LM, Klumb CEN Harrington, Jr., Bacchi WJCE. 2009. Frequent expression of multiple myeloma 1/interferon regulatory factor 4 in Burkitt lymphoma. Hum Pathol 40:565–571.

Hummel M, Bentink S, Berger H, Klapper W, Wessendorf S, Barth TF, Bernd HW, Cogliatti SB, Dierlamm J, Feller AC, Hansmann ML, Haralambieva E, Harder L, Hasenclever D, Kuhn M, Lenze D, Lichter P, Martin-Subero JI, Moller P, Muller-Hermelink HK, Ott G, Parwaresch RM, Pott C, Rosenwald A, Rosolowski M, Schwaenen C, Sturzenhofecker B, Szczepanowski M, Trautmann H, Wacker HH, Spang R, Loeffler M, Trumper L, Stein H, Siebert R. 2006. A biologic definition of Burkitt's lymphoma from transcriptional and genomic profiling. N Engl J Med 354:2419–2430.

Johansson B, Mertens F, Mitelman F. 1995. Cytogenetic evolution patterns in non-Hodgkin's lymphoma. Blood 86:3905–3914.

Klapper W, Kreuz M, Kohler CW, Burkhardt B, Szczepanowski M, Salaverria I, Hummel M, Loeffler M, Pellissery S,

Woessmann W, Schwänen C, Trümper L, Wessendorf S, Spang R, Hasenclever D, Siebert R. 2012. Patient age at diagnosis is associated with the molecular characteristics of diffuse large B-cell lymphoma. Blood 119:1882–1887.

Kovalchuk AL, Kishimoto T, Janz S. 2000. Lymph nodes and Peyer's patches of IL-6 transgenic BALB/c mice harbor T(12; 15) translocated plasma cells that contain illegitimate exchanges between the immunoglobulin heavy-chain mu locus and c-myc. Leukemia 14:1127–1135.

Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, Meyerson M, Gabriel SB, Lander ES, Getz G. 2014. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 505:495–501.

Leoncini L, Raphael M, Stein H, Harris N, Jaffe E, Kluin P. 2008. Burkitt lymphoma. In: Swerdlow S, Campo E, Harris N, Jaffe E, Pileri S, Stein H, Thiele J, Vardiman J, editors. WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues, 4th ed. Lyon: IARC, pp. 262–264.

Linch DC. 2012. Burkitt lymphoma in adults. Br J Haematol 156: 693–703.

Love C, Sun Z, Jima D, Li G, Zhang J, Miles R, Richards KL, Dunphy CH, Choi WW, Srivastava G, Lugar PL, Rizzieri DA, Lagoo AS, Bernal-Mizrachi L, Mann KP, Flowers CR, Naresh KN, Evens AM, Chadburn A, Gordon LI, Czader MB, Gill JI, Hsi ED, Greenough A, Moffitt AB, McKinney M, Banerjee A, Grubor V, Levy S, Dunson DB, Dave SS. 2012. The genetic landscape of mutations in Burkitt lymphoma. Nat Genet 44: 1321–1325.

Lu B, Zhou C, Yang W, Huang H, Gao Z, He Y, Liu H, Zhou X, Gong L. 2011. Morphological, immunophenotypic and molecular characterization of mature aggressive B-cell lymphomas in Chinese pediatric patients. Leuk Lymphoma 52:2356–2364.

Lundin C, Hjorth L, Behrendtz M, Ehinger M, Biloglav A, Johansson B. 2013. Submicroscopic genomic imbalances in burkitt lymphomas/leukemias: Association with age and further evidence that 8q24/MYC translocations are not sufficient for leukemogenesis. Genes Chromosomes Cancer 52:370–377.

Matsuo Y, MacLeod RA, Kojima K, Kuwahara K, Sakata A, Drexler HG, Nishizaki C, Fukuda S, Inoue Y, Sezaki T, Sakaguchi N, Orita K. 1997. A novel ALL-L3 cell line, BALM-16, lacking expression of immunoglobulin chains derived from a patient with hypercalcemia. Leukemia 11:2168–2174.

Matsuo Y, Drexler HG, Takeuchi M, Orita K. 2002. A novel L3-acute lymphoblastic leukemia cell line (BALM-27) carrying cytoplasmic immunoglobulin δ chain but lacking expression of cell surface immunoglobulin chains. Leuk Lymphoma 43:1441–1446.

Meng Q, Rayala SK, Gururaj AE, Talukder AH, O'Malley BW, Kumar R. 2007. Signaling-dependent and coordinated regulation of transcription, splicing, and translation resides in a single coregulator, PCBP1. Proc Natl Acad Sci USA 104:5866–5871.

Miles RR, Arnold S, Cairo MS. 2012. Risk factors and treatment of childhood and adolescent Burkitt lymphoma/leukaemia. Br J Haematol 156:730–743.

Misteli T, Caceres JF, Spector DL. 1997. The dynamics of a pre-mRNA splicing factor in living cells. Nature 387:523–527.

Müller JR, Janz S, Goedert JJ, Potter M, Rabkin CS. 1995. Persistence of immunoglobulin heavy chain/c-myc recombination-positive lymphocyte clones in the blood of human immunodeficiency virus-infected homosexual men. Proc Natl Acad Sci USA 92:6577–6581.

Müller JR, Jones GM, Janz S, Potter M. 1997a. Migration of cells with immunoglobulin/c-myc recombinations in lymphoid tissues of mice. Blood 89:291–296.

Müller JR, Mushinski EB, Williams JA, Hausner PF. 1997b. Immunoglobulin/Myc recombinations in Murine Peyer's patch follicles. Genes Chromosomes Cancer 20:1–8.

Murga Penas E, Schilling G, Behrmann P, Klokow M, Vettorazzi E, Bokemeyer C, Dierlamm J. 2014. Comprehensive cytogenetic and molecular cytogenetic analysis of 44 Burkitt lymphoma cell lines: Secondary chromosomal changes characterization, karyotypic evolution, and comparison with primary samples. Genes Chromosomes Cancer 53:497–515.

Nasr MR, Rosenthal N, Syrbu S. 2010. Expression profiling of transcription factors in B- or T-acute lymphoblastic leukemia/lymphoma and burkitt lymphoma: Usefulness of PAX5 immunostaining as pan-Pre-B-cell marker. Am J Clin Pathol 133:41–48.

Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim K, Kohlhammer H, Xu W, Yang Y, Zhao H, Shaffer AL, Romesser P, Wright G, Powell J, Rosenwald A, Muller-Hermelink H, Ott G, Gascoyne RD, Connors JM, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Fisher RI, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Staudt LM. 2011. Oncogenically active MYD88 mutations in human lymphoma. Nature 470:115–119.

Nishinakamura H, Minoda Y, Saeki K, Koga K, Takaesu G, Onodera M, Yoshimura A, Kobayashi T. 2007. An RNA-binding protein αCP-1 is involved in the STAT3-mediated suppression of NF-κB transcriptional activity. Int Immunol 19:609–619.

Ren C, Cho SJ, Jung YS, Chen X. 2014. DNA polymerase eta is regulated by poly(rC)-binding protein 1 via mRNA stability. Biochem J 464:377–386

Richter J, Schlesner M, Hoffmann S, Kreuz M, Leich E, Burkhardt B, Rosolowski M, Ammerpohl O, Wagener R, Bernhart SH, Lenze D, Szczepanowski M, Paulsen M, Lipinski S, Russell RB, Adam-Klages S, Apic G, Claviez A, Hasenclever D, Hovestadt V, Hornig N, Korbel JO, Kube D, Langenberger D, Lawerenz C, Lisfeld J, Meyer K, Picelli S, Pischimarov J, Radlwimmer B, Rausch T, Rohde M, Schilhabel M, Scholtysik R, Spang R, Trautmann H, Zenz T, Borkhardt A, Drexler HG, Moller P, Macleod RA, Pott C, Schreiber S, Trumper L, Loeffler M, Stadler PF, Lichter P, Eils R, Kuppers R, Hummel M, Klapper W, Rosenstiel P, Rosenwald A, Brors B, Siebert R. 2012. Recurrent mutation of the ID3 gene in Burkitt lymphoma identified by integrated genome, exome and transcriptome sequencing. Nat Genet 44:1316–1320.

Salaverria I, Philipp C, Oschlies I, Kohler CW, Kreuz M, Szczepanowski M, Burkhardt B, Trautmann H, Gesk S, Andrusiewicz M, Berger H, Fey M, Harder L, Hasenclever D, Hummel M, Loeffler M, Mahn F, Martin-Guerrero I, Pellissery S, Pott C, Pfreundschuh M, Reiter A, Richter J, Rosolowski M, Schwaenen C, Stein H, Trumper L, Wessendorf S, Spang R, Kuppers R, Klapper W, Siebert R, for the Molecular Mechanisms in Malignant Lymphomas Network Project of the Deutsche Krebshilfe, the German High-Grade Lymphoma Study Group, the Berlin-Frankfurt-Munster-NHL trial group. 2011. Translocations activating IRF4 identify a subtype of germinal center-derived B-cell lymphoma affecting predominantly children and young adults. Blood 118:139–147.

Salaverria I, Siebert R. 2011. The gray zone between Burkitt's lymphoma and diffuse large B-cell lymphoma from a genetics perspective. J Clin Oncol 29:1835–1843.

Sander S, Calado D, Srinivasan L, Köchert K, Zhang B, Rosolowski M, Rodig S, Holzmann K, Stilgenbauer S, Siebert R, Bullinger L, Rajewsky K. 2012. Synergy between PI3K signaling and MYC in Burkitt lymphomagenesis. Cancer Cell 22: 167–179.

Schmitz R, Young RM, Ceribelli M, Jhavar S, Xiao W, Zhang M, Wright G, Shaffer AL, Hodson DJ, Buras E, Liu X, Powell J, Yang Y, Xu W, Zhao H, Kohlhammer H, Rosenwald A, Kluin P, Muller-Hermelink HK, Ott G, Gascoyne RD, Connors JM, Rimsza LM, Campo E, Jaffe ES, Delabie J, Smeland EB, Ogwang MD, Reynolds SJ, Fisher RI, Braziel RM, Tubbs RR, Cook JR, Weisenburger DD, Chan WC, Pittaluga S, Wilson W, Waldmann TA, Rowe M, Mbulaiteye SM, Rickinson AB, Staudt LM. 2012. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. Nature 490:116–120.

Scholtysik R, Kreuz M, Klapper W, Burkhardt B, Feller AC, Hummel M, Loeffler M, Rosolowski M, Schwaenen C, Spang R, Stein H, Thorns C, Trumper L, Vater I, Wessendorf S, Zenz T, Siebert R, Kuppers R, for t. "Molecular Mechanisms in Malignant Lymphomas" Network Project of the Deutsche Krebshilfe,. 2010. Detection of genomic aberrations in molecularly defined Burkitt's lymphoma by array-based, high resolution, single nucleotide polymorphism analysis. Haematologica 95:2047–2055.

Silvera D, Gamarnik AV, Andino R. 1999. The N-terminal K homology domain of the poly(rC)-binding protein is a major determinant for binding to the poliovirus 5'-untranslated region and acts as an inhibitor of viral translation. J Biol Chem 274: 38163–38170.

Smyth GK. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 3:Article 3.

Thisted T, Lyakhov DL, Liebhaber SA. 2001. Optimized RNA targets of two closely related triple KH domain proteins, heterogeneous nuclear ribonucleoprotein K and aCP-2KL, suggest distinct modes of RNA recognition. J Biol Chem 276:17484–17496.

Vaque JP, Martinez N, Batlle-Lopez A, Perez C, Montes-Moreno S, Sanchez-Beato M, Piris MA. 2014. B-cell lymphoma mutations: Improving diagnostics and enabling targeted therapies. Haematologica 99:222–231.

Wang H, Vardy LA, Tan CP, Loo JM, Guo K, Li J, Lim SG, Zhou J, Chng WJ, Ng SB, Li HX, Zeng Q. 2010. PCBP1 suppresses the translation of metastasis-associated PRL-3 phosphatase. Cancer Cell 18:52–62.

Ye Z, Xiao Y, Shi H, Ke Z, Liu Y, Liang Y, Han A. 2011. Sporadic Burkitt lymphoma in southern China: 12 years' experience in a single institution in Guangzhou. J Clin Pathol 64: 1132–1135.