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

Rabea Wagener,† Sietse M. Aukema,† Matthias Schlesner,‡ Andrea Haake,§ Birgit Burkhardt,∥ Alexander Claviez,‡ Hans G. Drexler,‡ Michael Hummel,¶ Markus Kreuz,¶ Markus Loeffer,¶ Maciej Rosolowski,¶ Cristina Lopez,¶ Peter Moller,¶ Julia Richter,¶ Marius Rohde,¶ Matthew J. Betts,¶ Robert B. Russell,∥ Stephan H. Bernhart,∥ Steve Hoffmann,∥ Philip Rosenstiel,∥ Markus Schihabel,∥ Monika Szczepanowski,∥ Lorenz Trumper,∥ Wolfram Klapper,∥ and Reiner Siebert,∥ on behalf of the ICGC MMML-Seq-Project* and the “Molecular Mechanisms in Malignant Lymphomas” Network Project of the Deutsche Krebshilfe†

1Institute of Human Genetics, Christian-Albrechts-University Kiel and University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany
2Deutsches Krebsforschungszentrum Heidelberg (DKFZ), Division Theoretical Bioinformatics, Heidelberg, Germany
3Non-Hodgkin Lymphoma Berlin-Frankfurt-Münster Group Study Center, Department of Pediatric Hematology and Oncology, University Children’s Hospital, Münster, Germany
4Department of Pediatrics, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University, Kiel, Germany
5Leibniz-Institute DSMZ – German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany
6Institute of Pathology, Campus Benjamin Franklin, Charité – Universitätsmedizin, Berlin, Germany
7Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Germany
8Institute of Pathology, Universitätshôpital Ulm, Ulm, Germany
9Department of Pediatric Hematology and Oncology, Justus Liebig University, Giessen, Germany
10Cell Networks, Bioquant, University of Heidelberg, Heidelberg, Germany
11Transcriptome Bioinformatics, LIF e Research Center for Civilization Diseases, University of Leipzig, Leipzig, Germany
12Institute of Clinical Molecular Biology, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University, Kiel, Germany
13Institute of Hematopathology, University Hospital Schleswig-Holstein Campus Kiel/Christian-Albrechts University Kiel, Germany
14Department 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. We found mutations 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; Echo Network MMML-MYCG-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 Vogel’s 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 Senin 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 Holstein, Campus Kiel, Schwansenweg 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.

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 ≥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

<table>
<thead>
<tr>
<th>Cohort</th>
<th>PID</th>
<th>Age</th>
<th>Sex</th>
<th>MYC-status</th>
<th>Genomic position (hg19) in bp</th>
<th>Mutation (cDNA)</th>
<th>Mutation (protein)</th>
<th>Affected domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICGC</td>
<td>1</td>
<td>17</td>
<td>M</td>
<td>IGK-MYC</td>
<td>Chr2:70314913-70314914</td>
<td>c.41_42delTC</td>
<td>p.L14fs*22</td>
<td>Truncating in/after KH1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>16</td>
<td>F</td>
<td>IGK-MYC</td>
<td>Chr2:703151476</td>
<td>c.601_602insT</td>
<td>p.C201Lfs*12</td>
<td>Truncating after KH1</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8</td>
<td>M</td>
<td>IGK-MYC</td>
<td>Chr2:70315425</td>
<td>c.550C&gt;T</td>
<td>p.Q184*</td>
<td>Truncating after KH1</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>40</td>
<td>M</td>
<td>IGK-MYC</td>
<td>Chr2:70315753</td>
<td>c.878G&gt;A</td>
<td>p.C293Y</td>
<td>KHIII</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>M</td>
<td>IGK-MYC</td>
<td>Chr2:70315759</td>
<td>c.884T&gt;C</td>
<td>p.I295T</td>
<td>KHIII</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>76</td>
<td>M</td>
<td>IGK-MYC</td>
<td>Chr2:70315848</td>
<td>c.973A&gt;G</td>
<td>p.R325G</td>
<td>KHIII/ NLSII</td>
</tr>
</tbody>
</table>

| Cell lines | 1  | 36   | F   | IGK-MYC    | Chr2:703151779                | c.907T>G        | p.N303D          | KHIII           |
|            | 2  | 16   | M   | IGK-MYC    | Chr2:70315674                 | c.799G>A        | p.V267M          | NLSI            |
|            | 3  | 16   | M   | IGK-MYC    | Chr2:70315753                 | c.878G>A        | p.C293Y          | KHIII           |
|            | 4  | 29   | M   | IGK-MYC    | Chr2:70315759                 | c.884T>C        | p.I295T          | KHIII           |
|            | 5  | 21   | M   | IGK-MYC    | Chr2:70315848                 | c.973A>G        | p.R325G          | KHIII/ NLSII    |
|            | 6  | 76   | M   | IGK-MYC    | Chr2:70315848                 | c.973A>G        | p.R325G          | KHIII/ NLSII    |

|         | 7  | 3   | M   | IGK-MYC    | Chr2:703151779                | c.907T>G        | p.N303D          | KHIII           |
|         | 8  | 16   | M   | IGK-MYC    | Chr2:70315674                 | c.799G>A        | p.V267M          | NLSI            |

<table>
<thead>
<tr>
<th>Mutation</th>
<th>Affected domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.41_42delTC</td>
<td>Truncating in/after KH1</td>
</tr>
<tr>
<td>c.601_602insT</td>
<td>Truncating after KH1</td>
</tr>
<tr>
<td>c.550C&gt;T</td>
<td>Truncating after KH1</td>
</tr>
<tr>
<td>c.878G&gt;A</td>
<td>KHIII</td>
</tr>
<tr>
<td>c.884T&gt;C</td>
<td>KHIII</td>
</tr>
<tr>
<td>c.973A&gt;G</td>
<td>KHIII/ NLSII</td>
</tr>
<tr>
<td>c.907T&gt;G</td>
<td>KHIII</td>
</tr>
<tr>
<td>c.799G&gt;A</td>
<td>NLSI</td>
</tr>
<tr>
<td>c.878G&gt;A</td>
<td>KHIII</td>
</tr>
<tr>
<td>c.884T&gt;C</td>
<td>KHIII</td>
</tr>
<tr>
<td>c.973A&gt;G</td>
<td>KHIII/ NLSII</td>
</tr>
</tbody>
</table>

**Notes:**
- 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.
- Somatic origin of these mutations were proven by whole-genome sequencing of the germline tissue and verified by Sanger sequencing.
- Mutations were identified 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). 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 protein...
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 \( \geq 80\% \), 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
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.L295T, 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
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.

![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)](image)
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.C295Y, 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 15% of primary cases. Mutations seem to primarily affect the KHI11 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


Klapper W, Kreuz M, Kohler CW, Burkhardt B, Szczepanowski M, Salaverria I, Hummel M, Loeffler M, Pellosky S,
Kovalchuk AL, Kishimoto T, Janz S. 2000. Lymph nodes and peptide'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.


