Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

 OPEN ACCESS | Check for updates

Analysis of RBP expression and binding sites identifies PTBP1 as a regulator of CD19 expression in B-ALL

Nicole Ziegler[a,b], Mariela Cortés-López[c], Francesca Alt[a,b], Maximilian Sprang[d], Arsenij Ustjanzew[e], Nadine Lehmann[a,b], Khalifa El Malki[a,b], Arthur Wingerter[a,b], Alexandra Russo[a,b], Olaf Beck[a,b], Sebastian Attig[f], Lea Roth[a,b], Julian König[c], Claudia Paret[a,b,g]#, and Jörg Faber[a,b,g]#

aCenter for Pediatric and Adolescent Medicine, Department of Pediatric Hematology/Oncology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; bUniversity Cancer Center (UCT), University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; cInstitute of Molecular Biology (IMB), Mainz, Germany; dFaculty of Biology, Johannes Gutenberg University Mainz, Biozentrum I, Mainz, Germany; eInstitute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; fDepartment of Translational Oncology and Immunology at the Institute of Immunology, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany; gGerman Cancer Consortium (DKTK), Site Frankfurt/Mainz, Germany, German Cancer Research Center (DKFZ), Heidelberg, Germany

ABSTRACT

Despite massive improvements in the treatment of B-ALL through CART-19 immunotherapy, a large number of patients suffer a relapse due to loss of the targeted epitope. Mutations in the *CD19* locus and aberrant splicing events are known to account for the absence of surface antigen. However, early molecular determinants suggesting therapy resistance as well as the time point when first signs of epitope loss appear to be detectable are not enlightened so far. By deep sequencing of the *CD19* locus, we identified a blast-specific 2-nucleotide deletion in intron 2 that exists in 35% of B-ALL samples at initial diagnosis. This deletion overlaps with the binding site of RNA binding proteins (RBPs) including PTBP1 and might thereby affect CD19 splicing. Moreover, we could identify a number of other RBPs that are predicted to bind to the CD19 locus being deregulated in leukemic blasts, including NONO. Their expression is highly heterogeneous across B-ALL molecular subtypes as shown by analyzing 706 B-ALL samples accessed via the St. Jude Cloud. Mechanistically, we show that downregulation of PTBP1, but not of NONO, in 697 cells reduces CD19 total protein by increasing intron 2 retention. Isoform analysis in patient samples revealed that blasts, at diagnosis, express increased amounts of CD19 intron 2 retention compared to normal B cells. Our data suggest that loss of RBP functionality by mutations altering their binding motifs or by deregulated expression might harbor the potential for the disease-associated accumulation of therapy-resistant CD19 isoforms.

Introduction

Despite tremendous improvements in the treatment of B-ALL during the last years, the prognosis for those patients suffering a relapse is rather poor. Immunotherapy with chimeric antigen receptor (CAR)-T cells targeting CD19 appeared to be a game changer, leading to impressive remission rates in pediatric patients with relapsed or refractory ALL.[1,2] However, up to 50% of B-ALL patients receiving CAR-T cells develop disease relapse, 30–60% of them being characterized by target antigen loss.[3–7] As a consequence, other B-lineage markers, such as CD20 and CD22, are under investigation as alternative or additive targets for mono- and bivalent CARs in the treatment of B cell malignancies.[8–11]

Recent research already identified genetic alterations in the *CD19* locus that are attributed to epitope-negative protein variants.[12] Furthermore, a number of alternative splicing events, such as exon 2 skipping, deletion of exons 5–6 and intron 2 retention could be related to CD19-negative relapse.[13–16] Interestingly, some of those isoforms already exist at diagnosis.[17] Coexistence and correlations between those mechanisms can be expected. Along this line, Cortés-López et al. just provided data underscoring the complex coherence between somatic mutations within the *CD19* gene, changes in splicing and the appearance of CD19 isoforms with the potential of being therapy-resistant.[18]

Splicing events are mediated by RNA binding proteins (RBPs) that coordinate the incorporation of exons into the mature mRNA. In respect to CD19, mechanistic studies could already associate single factors, among them serine/arginine-rich splicing factor 3 (SRSF3), to alternative splicing events leading to loss of the target antigen.[13,18]

However, despite considerable improvements in understanding the causes and consequences of CD19 mis-

splicing, the molecular prerequisites determining the onset of fatal deregulations that finally lead to the dominance of therapy-resistant CD19 variants is not fully understood so far. Following the idea that certain molecular signs predicting CD19 modulations might appear early during the course of the disease, we investigated *CD19* mutations and the expression profile of RBPs in blasts at initial diagnosis. Normal B cells of healthy donors were used as a reference.

Our data suggest that disease-associated genetic mutations in RBP binding motifs as well as the deregulation of splicing modulators may lay the foundation for the prevalence of therapy-resistant CD19 isoforms by intervening in the functional network of RBPs.

Material and methods

Sample cohort

All pediatric B-ALL patients were treated according to the COALL 08–09 study protocol (v. 01.10.2010). The total number of B-ALL patients analyzed was 36, 69% of them being diagnosed with common-ALL and 28% with pre-B-ALL, one patient with pro-B-ALL. The majority of patients (~31%) featured the hyperdiploidy subtype, ~22% were of the molecular subtype ETV6-RUNX1 (Table S1). Patients of the control group were hospitalized due to a non-hematologic malignancy. Bone marrow or peripheral blood was obtained as surplus material during standard diagnostic procedures. Subsequent analysis was performed with the consent of the patients or patient's parents in agreement with the ethics committee of Rhineland-Palatinate (no. 2018–13713). Samples were handled in accordance with the current (2013) version of the Declaration of Helsinki. Remission was defined as <5% blast cells.

DNA sequencing

Exon 1 to 4 including introns of the *CD19* locus (chr16:-28942047–28944969) was amplified with the Expand long template Polymerase system (Roche). Paired-end libraries were created following the Nextera XT protocol (Illumina) which uses transposome to fragment and immediately tag the DNA with adapter sequences in a single step. Quality of the libraries was proofed using an Agilent Bioanalyzer System (Agilent Technologies). Libraries were subjected to deep sequencing on an Illumina MiSeq sequencer using 151 cycles (paired-end). Minimum depth of coverage was 1000x. Data were processed using BWA Enrichment v1.0 for generation of BAM files and the somatic variant caller of Illumina, which allows to detect low-frequency mutations (below 5%). Analysis of variants was performed with the VariantStudio software (Illumina). Variants with a population frequency less than 5% were analyzed further. Reads were visualized using the IGV software.

Prediction of RBP binding sites

RBP binding motifs were predicted using the website service of the AtTRACT database.[19] As input, we used the sequence of the *CD19* locus (exons 1 to 3) and selected motifs of at least 4 nucleotides in length. We collapsed overlapping motifs per RBP using custom R scripts based in the GenomicRanges package.

PTBP1 iCLIP2

The PTBP1 iCLIP2 Genome Browser view was generated from PTBP1 iCLIP2 experiments generated for NALM-6 cells[18]; GEO accession numbers GSM5542617- GSM5542620). The genome browser view was generated with IGV[20] with the hg38 human genome version.

Targeted RNA-Seq

RNA was extracted with the miRNeasy kit (QIAGEN). Only samples with a RIN >8 were used for library preparations using the Illumina® TruSeq® Targeted RNA Expression kits to perform multiplexed gene expression profiling. Quality of the libraries was proofed with an Agilent Bioanalyzer System (Agilent Technologies). Libraries were subjected to sequencing on an Illumina MiSeq sequencer using 1×51 cycles (single read). Data were processed in the Illumina® BaseSpace Sequence Hub to extract the raw counts. Counts were analyzed with DESeq2 to determine differentially expressed genes. Two *HPRT1* targets were used as controls to estimate the size factors before differential analysis and normalization. Log-transformed raw counts were used for heatmap visualization.

RNA-Seq data analysis and visualization of publicly available data

RNA-Seq derived HTSeq count data of 706 B-ALL samples of 14 different subtypes were obtained from St. Jude Cloud (https://stjude.cloud).[21] Transcripts of the count matrix with less than 10 counts in sum of all samples were excluded. Read counts were normalized by the median of ratios method using the DESeq2 R package (version 1.34.0).[22] One pseudocount was summed to the normalized count values. The batch effect caused by different library preparation protocols was removed using the negative binomial regression by the R function ComBat_seq from the sva package (version 3.42.0) (https://bioconductor.org). Subtypes were used as "group" parameter of the function to preserve the biological condition of interest. After batch correction, boxplots showing the decimal logarithmic transformed expression values of transcripts were visualized with ggplot2 package (version 3.3.6).

Flow cytometric immunophenotyping

Immunophenotyping was performed with bone marrow aspirates or peripheral blood at initial diagnosis following standard diagnostic procedures as described previously.[17] In brief, screening was performed by using different markers, including: CD45, CD19, CD34, IgM, kappa, lambda, CD10, CD22, CD65, CD20, CD24, CD79a, CD15, CD3, TdT, HLA-DR (all Beckman Coulter). Antibody solutions were utilized as recommended by the manufacturer. Following red blood cell lysis and two subsequent washing steps, cells were resuspended in

PBS containing 1% BSA. Cells were analyzed using a Navios Flow Cytometer (Beckman Coulter) and data was analyzed using Navios software version 1.3. Samples containing >80% leukemic blasts were chosen for cell sorting.

Cell lines

697 cells were obtained from DSMZ and cultured in RPMI medium (Gibco™) with 10% fetal bovine serum (Gibco™), 1% L-glutamine (Sigma-Aldrich) and 1% penicillin-streptomycin solution (Sigma-Aldrich). Cells were cultivated at 37°C in a humidified incubator at 5% CO_2 and subcultured every 3–4 days.

Isolation of PBMCs

For isolation of peripheral blood mononuclear cells (PBMCs), bone marrow was diluted with PBS + 2 mM EDTA and separated by density gradient centrifugation (800xg, 30 min) using Histopaque®-1077 (Sigma-Aldrich). PBMCs were washed with D-PBS (Sigma-Aldrich) and immediately frozen in FBS containing 10% DMSO (Sigma-Aldrich).

Fluorescence-activated Cell Sorting (FACS)

PBMCs isolated from bone marrow were stained at room temperature for 15 min in the dark. 7-aminoactinomycon D (7-AAD) and anti-CD45 antibody were used for cells from leukemia patients, 7-AAD and anti-CD19 (all Beckman Coulter) for cells from healthy donors. Sorting was performed in MACS buffer (PBS, 2 mM EDTA, 0.1% BSA) using a FACS Aria (Becton Dickinson). Normal B cells and leukemic blasts were defined as 7-ADD⁻/CD19⁺ and 7-ADD⁻/CD45^low, respectively.

RNA extraction and cDNA synthesis

Total RNA from sorted cells or from 697 cells was purified using the ReliaPrep™ RNA Cell Miniprep System (Promega) or the RNeasy Mini Kit (QIAGEN), respectively, following the manufacturer's protocol. Reverse transcription was performed with the PrimeScript™ RT Reagent Kit with gDNA Eraser (TaKaRa).

Quantitative RT-PCR (qRT-PCR)

qRT-PCR was performed using the PerfeCTa® SYBR® Green Fast Mix® (Quantabio) in a LightCycler 480 instrument (Roche). For sequences of primers for RBPs and CD19 isoforms see Table S7. Raw values were normalized to *HPRT*.

Semi-quantitative RT-PCR

Semi-quantitative RT-PCR was performed using cDNA of FACS-sorted B cells and leukemic blasts. To amplify CD19 isoforms, primers spanning exon 1 to exon 4 (Table S7) and Taq-DNA polymerase I (Axon Labortechnik) were used. PCR conditions were as follows: 94°C 5 min, 35 cycles of 94°C 30s, 60°C 30s, 72°C 1 min, 72°C 10 min. Fragments were visualized using a QIAxcel® DNA High Resolution Cartridge on a QIAxcel

Advanced instrument (QIAGEN) running the standard protocol.

Knockout experiments in 697 cells

For CRISPR/Cas9 experiments in 697 cells, 2 gRNAs for each target were used, applying a Dual-Guide approach. crRNA:tracrRNA duplex formation and RNP assembly were performed for each gRNA separately following a protocol for electroporation of human B cell lines provided by the manufacturer (IDT). In brief, equimolar amounts of crRNA and tracrRNA were used for hybridization at 95°C for 5 min. RNP complex assembly was performed by mixing guide RNA and Cas9 enzyme (Alt-R® S.p. Cas9 Nuclease V3; IDT) at a 1:1.2 molar ratio in PBS with subsequent incubation at RT for 20 min. Mixture was kept on ice until electroporation. 1×10^6 cells were resuspended in Electroporation Buffer (Bio-Rad), RNPs as well as Electroporation Enhancer (IDT) in a final concentration of 4.8 μM were added and electroporation was performed in a GenePulser® Cuvette with 0.2 cm electrode gap (Bio-Rad) using a GenePulser Xcell device (Bio-Rad) with the following settings: square wave, 1 pulse, 250 V, 2 ms. Immediately after electroporation, cells were kept in pre-warmed medium containing 20% FCS without antibiotics. 72 h after nucleofection, the KO cell pool was collected for downstream experiments. Only KO cell pools with KO efficiencies >50% were considered for further analysis.

Predesigned crRNAs and accessory reagents were purchased from IDT. Following crRNAs were used: Hs.Cas9.PTBP1.1.AA, Hs.Cas9.PTBP1.1.AB, Hs.Cas9.PTBP2.1.AB, Hs.Cas9.PTBP2.1.AD, Hs.Cas9.NONO.1.AA, Hs.Cas9.NONO.1.AB. Alt-R® Cas9 Negative Control crRNA #1 was used as non-targeting control (indicated as WT in the figures).

Quantification of cell surface proteins

10^5 cells were resuspended in 200 μl PBS and incubated for 15 min with 7-AAD Viability Dye (Beckman Coulter) and the following antibodies: CD45 (V500-c), CD19 (APC-H7), CD20 (BV605) (all BD). Cells were washed with 2 ml PBS, resuspended in 400 μl PBS containing 1% BSA and measured using a BD Lyric Flow Cytometer equipped with the BD FACSuite™ software version v1.5.0.925. Singlets were determined by FSC-A vs. FSC-H gating and only living cells were quantified for the expression of surface proteins using FlowJo (software version 10.8.1).

Western blotting

Cell pellets were lysed in RIPA buffer containing protease inhibitor cocktail (Roche). SDS-Page and Western blotting were performed following standard procedures. Primary antibodies were as follows: PTBP1 (1:1000; Cell Signaling), PTBP2 (1:1000; abcam), CD19 (1:1000; Cell Signaling), GAPDH (1:5000; Cell Signaling). Signal detection and quantification was performed using a Fusion Pulse imaging system (Vilber).

Statistical analysis

Data are shown as mean ± SD. Statistical significance was analyzed by two-tailed Student's *t* test (GraphPad Prism software version 9.0.1). *p* values <.05 were considered significant.

For RNAseq data analysis, pairwise differential gene expression (DGE) analysis among subtypes was performed using the DESeq2 package (version 1.34.0) by fitting the negative binomial generalized linear model for each gene and using the Wald test for significance testing. The design parameter was set to "~Library_selection_protocol + Group". Benjamini & Hochberg correction as well as "ashr" log2 fold shrinkage method were used to obtain p-adjusted values.[23] Table S8 provides results of the DGE subsetted by genes and contain log2 fold changes, log2 fold change standard errors, and the p and p-adjusted values.

Targeted RNAseq samples were handled similarly, but the two HPRT1 targets were used as controls to estimate the size factors and the fitType was set to "mean" to control for the low number of surveyed genes in a targeted setting. DEseq2′s default shrinkage was used.

Results

B-ALL patients feature disease-specific mutations in the CD19 locus at initial diagnosis

The cellular mechanisms that potentially account for epitope loss are diverse, whereas the occurrence of mutations in the gene encoding for the targeted antigen itself and alternative splicing events are gaining more and more attention. Particularly, most CART-19 therapy-resistant protein variants identified so far result from exon 2 deletions or inaccurate excision of adjacent introns.

So far, only exon mutations in the *CD19* genetic locus were investigated in patient samples. Minigene-based assays, however, suggest that also intronic mutations can give rise to therapy-resistant alternative CD19 isoforms.[18,] Thus, we performed deep sequencing of the *CD19* locus comprising exon 1 to 4 including introns in 3 controls, 20 pediatric B-ALL patients at initial diagnosis and, out of those, 15 samples in remission (Table S1). Strikingly, we identified a small deletion (NM_001178098.1:c.356–95_356-94delCT, derived from TTC>TTC/T at position 28944127) with an allele frequency of ~1% located in intron 2 (Tables S2, S3, Figure 1a) in 35% of samples at diagnosis, both in common and pre-B-ALL, but not in the control group. Remission samples of the same patients did not harbor this genetic variant, indicating its specificity to leukemic blasts. Interestingly, 80% (four out of five) patients of the ETV6-RUNX1 and none of the hyperdiploid molecular subtype carried the mutation. In one sample from initial diagnosis, another intronic mutation (NM_001178098.1:c.356–111A>G) was detected. It was located next to the NM_001178098.1:c.356–95 locus and featured an allele frequency of 50% in our sample while showing a particularly low allele frequency (0.12%) in the cohort of the 1000 Genomes Project (https://www.internationalgenome.org). However, as material of the same patient in remission was not available, we cannot judge whether this mutation is particularly associated to leukemic blasts.

Interestingly, our analysis did not reveal any blast-specific mutation affecting the coding region of CD19.

Thus, our results indicate that disease-specific subclonal mutations in intron regions appear in B-ALL patients. Moreover, those point mutations already exist at diagnosis, suggesting a potential role in the disease process including the possibility of developing therapy resistance.

Disease-specific mutations in CD19 affect RBP binding sites

In order to assess potential consequences of this mutation, we investigated whether the small deletion in *CD19* intron 2 might overlap with recognition motifs for splicing factors and thereby alter the landscape of cis-regulatory elements. The ATtRACT database predicted polypyrimidine tract binding protein (PTBP) 1, PTBP2 and zinc finger protein 36 (ZFP36) to bind to this locus (Figure 1b). Interestingly, a high number of recognition motifs for PTBP1 can be detected in this region, several of them disappearing upon the insertion of the 2-nucleotide deletion (Figure 1c).

Although we found a high frequency of PTBP1 binding sites throughout the analyzed sequence, the abundance of long and thereby more specific motifs accumulate in intron regions and, remarkably, most prominently in intron 2 (Fig. S1). iCLIP2 experiments in NALM-6 cells experimentally confirmed that PTBP1 most preferentially binds to CD19 intron 2 while highest frequencies of cross-link events could be found in the second half of the intron, comprising the mutation site (Figure 1a). Thus, we assume an outstanding function of this dedicated locus. The optimal binding site for PTBP1 is the core sequence TCTT(C) embedded in a longer pyrimidine tract.[24,25] Exactly this motif is affected by the NM_001178098.1:c.356–95_356-94delCT deletion (Figure 1b). Although the core sequence is maintained despite the 2-nucleotide loss, the pyrimidine tract downstream of the consensus sequence and thereby the distance to adjacent PTBP1 recognition motifs is shortened. It is known that sequences surrounding high affinity binding sites of RBPs equally affect splice site recognition and binding efficiency.[24,26–28] Consistently, a number of potential PTBP1 binding sites are lost in case the mutation is present (Figure 1c). In contrast, the one motif for ZFP36 binding remains. Given also the fact that ZFP36 most preferentially binds AU-rich elements and one high affinity binding motif is located 25nt downstream of this position,[29] we conclude that the blast-specific mutation has the most considerable impact on PTBP1 binding.

In order to determine whether the expression of PTBP1 and ZFP36 might be related to the disease state of B cell leukemia, qRT-PCR analysis was performed in sorted blasts and B cells (Figure 1d). Although not reaching significance due to high patient-to-patient variability, *PTBP1* showed lower mRNA abundance in blasts of 77% of patients compared to the average expression in B cells. Other than *PTBP1*, transcription of *ZFP36* was rather low in all our samples. Yet, it was significantly less expressed in blasts than in B cells.
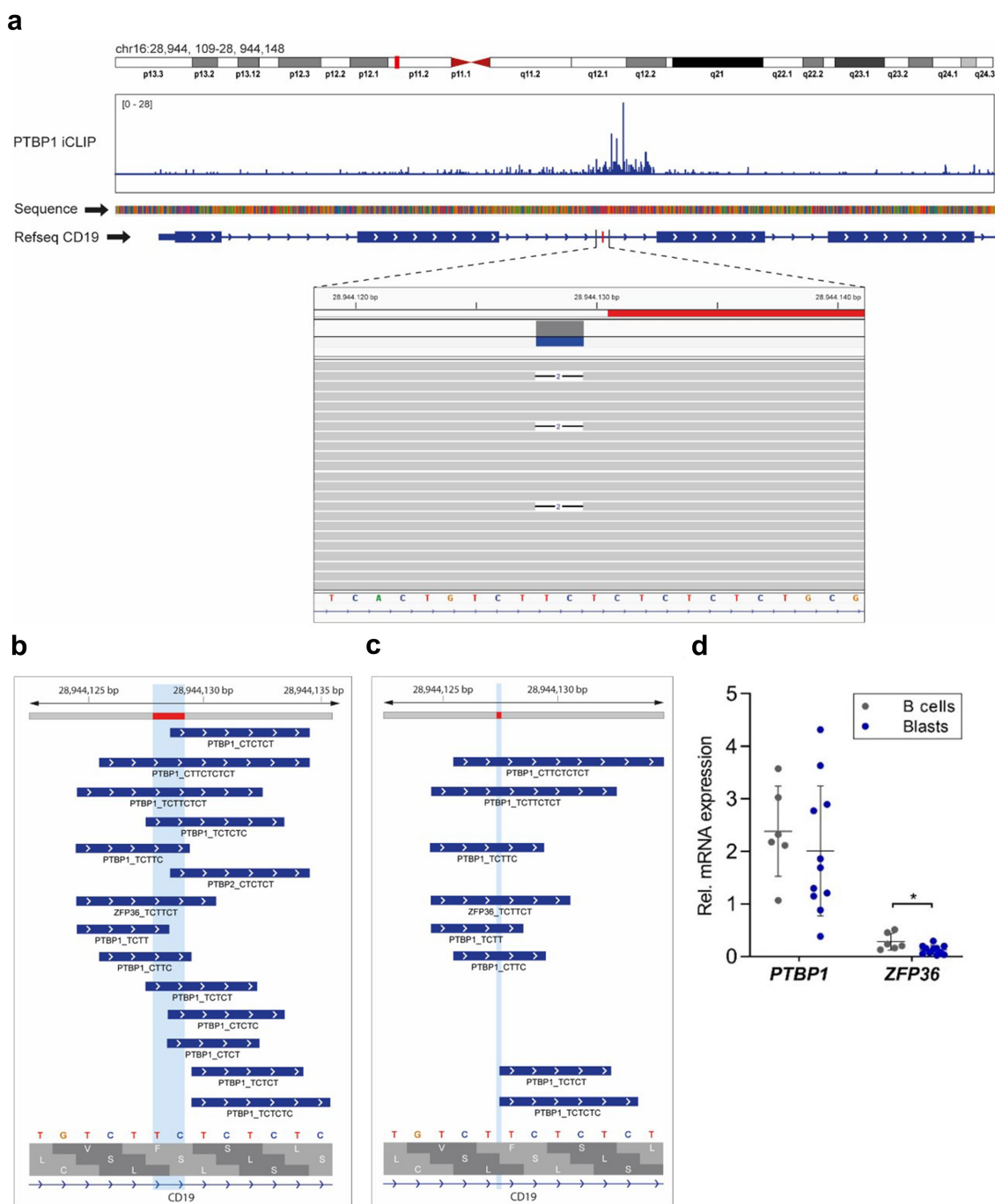
Figure 1. A blast-specific mutation in CD19 intron 2 affects RBP binding domains(a) Localization of CD19 on chr.16 p11.2 (upper panel, red square). Bar diagram showing PTBP1 iCLIP2 crosslink events on each nucleotide of endogenous CD19 exon 1–4. Lower panel shows and overview of CD19 exon 1–4, highlighting the position of the mutation in intron 2 in red. Detailed view of the DNA segment harboring the mutation site and single reads carrying the TTC>T mutation (box). Alignments are visualized with Integrative Genomics Viewer (IGV). (b, c) RBPs and binding motifs in the WT (A) and mutated sequence (b) overlapping the respective DNA locus. Nucleotides being affected by the deletion are highlighted in light blue. All binding motifs ≥4 nucleotides suggested by ATtRACT are displayed. (d) qRT-PCR analysis of *PTBP1* and *ZFP36* in sorted B cells and leukemic blasts of pediatric healthy donors and B-ALL patients (n = 6 patients in control, n = 11 patients in diseased group; *, p < .05).

Expression of RBPs is deregulated in B-ALL patients

The appearance of CART-19 therapy-resistant CD19 isoforms and, more specifically, CD19 exon 2 processing have been shown to be dependent on the presence and function of dedicated splicing factors.[13,18] To get an overview of those RBPs binding the relevant genomic region within *CD19* and might thereby impact protein processing, we extended the search for RBP binding sites to the sequence spanning exon 1 to exon 3. We thereby considered only those motifs ranging between 4 and 10 nucleotides, resulting in a list of 54 RBPs (Fig. S1). Specificity of their expression profile for the disease state was investigated by targeted RNA-Seq in nine patient samples of initial B-ALL diagnosis, 16 in remission and of 2 healthy donors (Figure 2a). Obviously, as only two B cell specimen were available for RNA-
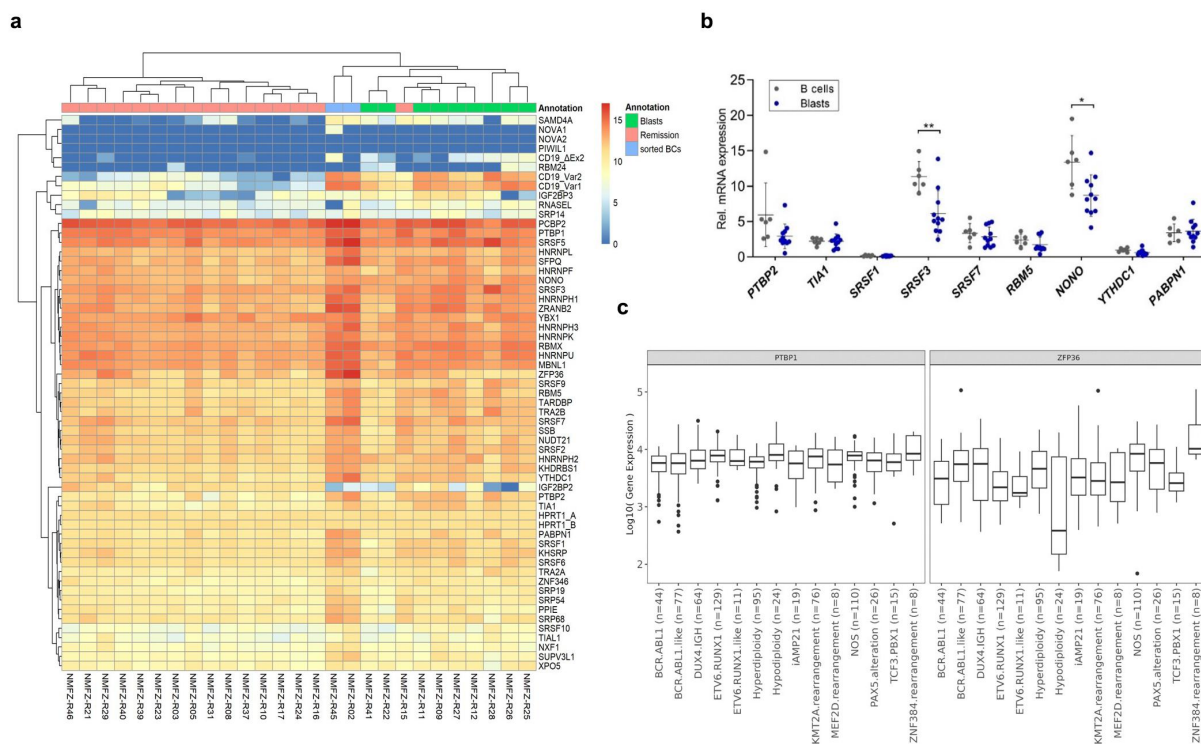
Figure 2. B-ALL patients exhibit disease-specific RBP expression profile(a) Heatmap visualization of RNA-Seq data from patients at diagnosis, in remission and sorted B-cells from healthy donors. RBP expression is shown as log-transformed data normalized to *HPRT*. (b) qRT-PCR analysis of selected RBPs in sorted B cells and leukemic blasts of pediatric healthy donors and B-ALL patients (n = 6 patients in control, n = 11 patients in diseased group; *, p < .05; **, p < .01). (c) Boxplots of log10-transformed, gene expression values of PTBP1 and ZFP36 grouped by B-ALL subtypes. Boxes range from first to third quantile, line indicates the median, whiskers show the highest and lowest values no further than 1.5*IQR from the hinge. Dots represent outliers. Values are normalized by the median of ratios method.

Seq, the results of this sample group need to be interpreted with reservation and was considered as preliminary data. Pairwise comparisons revealed 20 differentially expressed genes (DEGs) in blasts relative to B cells, while 10 and 40 DEGs were found comparing remission samples to blasts or B cells, respectively (Tables S4-S6). Overall, the majority of RBPs were in tendency less expressed in patient samples compared to B cells, which, at least partly, might result from the fact that B cell samples were sorted, while the others were not. However, these data indicate that expression levels of a high number of RBPs correlate with the disease state of B-ALL.

To analyze the disease-associated expression pattern of selected RBPs predicted by RNA-Seq data in a representative cohort of patient samples, we performed qRT-PCR analysis in blasts of 11 pediatric B-ALL patients at diagnosis and in B cells of 6 healthy donors. To allow for cell type-related conclusions, cells were isolated by FACS (Fig. S2). Gene expression of most of the splicing factors that we analyzed here, namely TIA1 cytotoxic granule associated RNA binding protein (*TIA1), SRSF1, SRSF7*, RNA binding motif protein 5 (*RBM5)*, YTH domain containing (*YTHDC1)* and poly(A) binding protein nuclear 1 (*PABPN1*) did not significantly differ between blasts and B. However, *PTBP2*, a well-described paralog of PTBP1, was in tendency less expressed in isolated leukemic blasts than in B cells (Figure 2b). Expression of *SRSF3*, which is described to regulate CD19 splicing in B-ALL patients, was significantly lower in blast cells. Moreover, we could identify non-POU domain containing octamer binding (*NONO*) being significantly decreased in leukemic blasts. Although it is described

to be associated with tumorigenesis in many types of cancer, NONO could not be related to alternative splicing events in leukemia yet.

A comparison between expression of the RBPs obtained by the two different methods is rather difficult as patients were not the same for both types of analyses. Thus, patient-specific features, such as the B-ALL molecular subtype, were unequally distributed between groups. Therefore, we re-analyzed RNA-Seq data of 706 B-ALL samples of known subtype to further assess the expression of the selected RBPs in B-ALL subtypes. As expected from our previous results, ZFP36 expression was generally low with exception of the ZNF384 subtype, while PTBP1 expression was high but heterogeneous across subtypes (Figure 2c). High heterogeneity in the expression was observed across all RBPs (Fig. S3, Table S8).

Taken together, our data reveal heterogenous expression of RBPs that is also dependent on the B-ALL subtype.

Given their ability to bind in the genomic sequence of CD19 that enables therapy-relevant splicing events, those differences in RBP abundance imply a correlation with the occurrence of particular CD19 isoforms.

Thus, beyond the appearance of intron-specific mutations that may affect the binding of certain RBPs, we show that the expression of such RBPs is generally deregulated in B-ALL.

PTBP1 is indispensable for regular CD19 splicing and surface expression

Given the high abundance of PTBP1 recognition motifs in intron 2 and the loss of several binding sites by the 2nt deletion,

PTBP1 appeared to be the one RBP being considerably affected by this mutation. Furthermore, the low expression of ZFP36 in blasts, that equally holds true in leukemic cell lines such as 697 (data not shown), further supported our assumption that PTBP1 likely is of higher importance for exon 2 splicing than ZFP36. We analyzed the interdependency of *PTBP1* and total *CD19* mRNA expression in blasts at diagnosis, revealing a positive correlation (Figure 3a). In order to investigate whether this effect might result from alternative CD19 splicing mediated by PTBP1, we performed CRISPR/Cas9-mediated knockout (KO) in the leukemic cell line 697 (Fig. S4A, C). As PTBP2 has similar function as PTBP1 and both factors can compensate for each other, we also performed KO of *PTBP2* (Fig. S4B, C). KO efficiencies were approximately 70% and 90%, respectively, and we refrained from selecting single KO cell clones but rather used the cell pools for further analysis. qRT-PCR revealed that downregulation of *PTBP1* induced an increase in *PTBP2* expression, confirming the functional relevance of the reduction of *PTBP1* levels (Fig. S4C). It is known that PTBP1 regulates PTBP2 levels by alternative splicing mediating nonsense-mediated decay, which potentially holds true also in our cellular model.[30] Consistent with the positive correlation seen in our patient cohort, CD19 surface expression was significantly reduced after *PTBP1* KO (Figure 3b). Western blot confirmed this data, showing that levels of CD19 total protein were approximately halved (Figure 3c). KO of *PTBP2*, in contrast, did not significantly reduce CD19 surface

expression or total protein abundance, which is equally mirrored in our patient cohort (Figure 3d).

In order to figure out whether downregulation of total CD19 protein is due to an altered isoform composition mediated by PTBP1, the abundance of exon 2-related variants was investigated by qRT-PCR (Figure 3e-f). Indeed, we could show that isoform distribution changes upon PTBP1 KO. While the exon 2 WT variant was less abundant, intron 2 retention (In2Ret) was significantly upregulated (figure 3f). Simultaneously with us, similar observations were made by Córtes-López et al.[18] On top of that, we determined significantly decreased expression of the isoform harboring exon 2 partial deletion. As expected, KO of *PTBP2* did not significantly affect isoform distribution. These data clearly suggest that deregulation of PTBP1, caused either by expression changes or alteration in binding capabilities, imply an accumulation of the epitope-negative splicing variant In2Ret that finally result in decreased levels of CD19 protein. CD20, that is used as an alternative B cell marker, was not affected by PTBP1 KO (Fig. S4D, E).

Intron 2 retention is increased in blasts compared to normal B cells

In order to investigate whether differences in CD19 isoform expression generally exist in leukemic blasts in comparison to normal B cells, we analyzed the occurrence of CD19 variants in
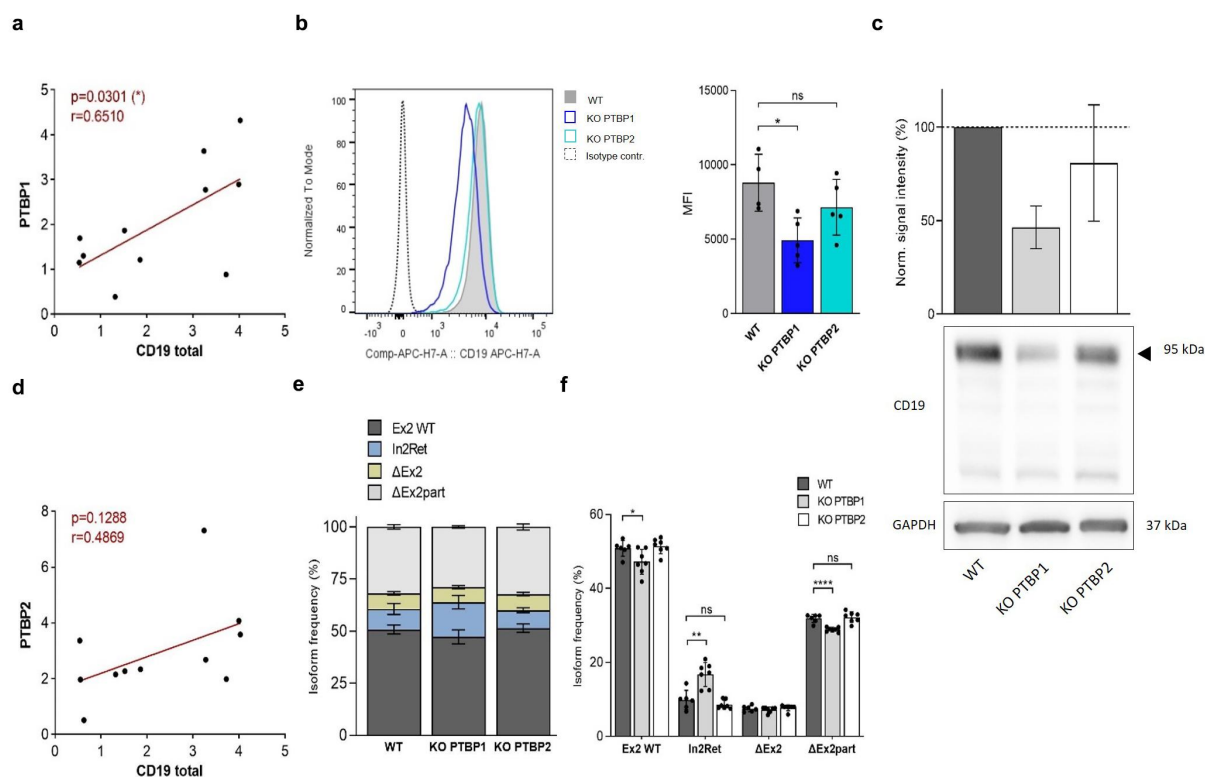

Figure 3. PTBP1 regulates CD19 protein expression by modulating alternative splicing of intron 2(a) Correlation analysis of *CD19* and *PTBP1* mRNA expression levels in sorted blasts of patients at initial diagnosis (n = 11). (b) Flow cytometric analysis of CD19 surface expression in WT (non-targeting control), *PTBP1* KO and *PTBP2* KO cells 72 h after nucleofection. Histograms are shown normalized to mode, mean fluorescent intensities were used for quantification (n = 4 independent experiments). (c) Western blot for CD19 and quantification shown as signal intensity relative to WT cells (100%) (n = 4 independent experiments). GAPDH served as loading control. (d) Correlation analysis of *CD19* and *PTBP2* mRNA expression in blasts of patients at diagnosis (n = 11). (e, f) qRT-PCR analysis of *Ex2WT, In2Ret, ΔEx2* and *ΔEx2part* in WT, *PTBP1* as well as *PTBP2* KO cells. Normalized expression values were calculated as percentage of all Ex2-related CD19 isoforms (n = 7 independent experiments). Data is shown as stacked graph (e) and bar graph for better visualization of statistical differences (f) (*, p < .05; **, p < .01; ***, p < .001). Simple linear regression and fit lines for correlation analysis were calculated using GraphPad Prism software.

blasts from patients at diagnosis and in B cells of healthy donors. Focusing on exon 2 processing, we observed that both exon2-deleted CD19 variants (*ΔEx2* and *ΔEx2part*) were already present in blasts at diagnosis, consistent with previous results (Figure 4a).[17] Additionally, we detected intron 2 retention. Importantly, all three mis-spliced CD19 isoforms were expressed also in B cells (Figure 4b), suggesting that aberrant splicing does equally occur in healthy people. Consequently, the mere presence of the *CD19* variants analyzed here is not per se predictive for the disease. However, the accumulation of dedicated isoforms might make a big difference by shifting their ratio and disbalance their regular equilibrium. To investigate this issue in our patient cohort, we precisely quantified the expression of *CD19* variants by qRT-PCR using isoform-specific primers. For better interpretation, the percentage of each exon 2-related variant was calculated (Figure 4c, Table S7). Although not reaching significance as its mean, considering the individual patient data revealed that the majority of blast samples feature lower levels of the regularly spliced *Ex2 WT* isoform compared to B cells. At the same time, abundance of *In2Ret* was significantly elevated. This suggests that the shift toward the mis-spliced variant happens at the expense of the regular one. The abundance of *ΔEx2* and *ΔEx2part* did not show apparent differences between groups. Further analysis corroborated our finding, revealing negative correlation of the percentage of the exon 2 WT isoform related to intron 2

retention (Figure 4d). Although the increase in *In2Ret* could not be decidedly attributed to the expression levels of *PTBP1* in our patient samples (Figure 4e), this does not preclude a prominent role for PTBP1 in CD19 exon 2 splicing but rather is a display of high patient-to-patient variabilities that appear to be given in this type of correlation.[18] Consequently, an extremely high number of patients is needed for statistical evaluation. Furthermore, given the general deregulation of the CD19 splicing machinery that we observe in our patient cohort, it is likely that the effects of single RBPs are at least partially masked.

NONO is a regulator of CD20 cell surface expression

As besides PTBP1 and SRSF3, NONO was significantly less expressed in our blast cohort compared to normal B cells, we considered it being an additional RBP with potential impact on CD19 splicing. Thus, we also performed KO for NONO in 697 cells, reaching an average KO efficiency of 66% (Fig. S4F). Interestingly, neither the expression of the different isoforms (Figure 5a) nor CD19 surface abundance (Figure 5b) was affected. However, we observed a higher proportion of CD20-positive cells by FACS (Figure 5c), going along with a significant increase in MFI in NONO KO cells compared to WT (Figure 5d). Taken together, this data suggests that deregulation of different RBPs as occurring in B-ALL considerably
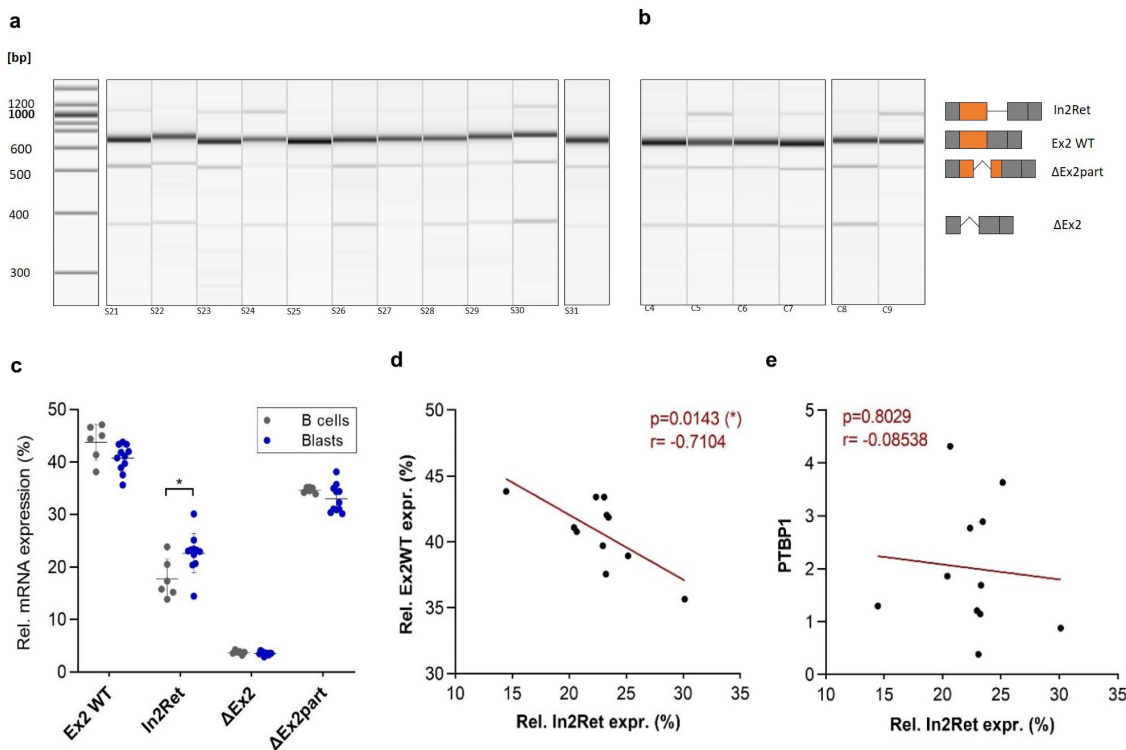

Figure 4. Intron 2 retention is not disease-specific but increased in leukemic blasts(a, b) Capillary gel electrophoresis of semi-quantitative RT-PCR visualizing CD19 exon 2 isoforms in leukemic blasts (a) and normal B cells (b). Primers spanning exon 1–4 were used. (c) Percentage of CD19 exon 2 isoforms in sorted B cells and leukemic blasts of pediatric healthy donors and B-ALL patients. Calculations were performed using raw values obtained by qRT-PCR (n = 6 patients in control, n = 11 patients in diseased group). (d) Correlation analysis of the percentage of *In2Ret* and *Ex2 WT* expression in leukemic blasts (n = 11). (e) Correlation analysis of the percentage of *In2Ret* and *PTBP1* expression in blast cells (n = 11). Simple linear regression and fit lines were calculated using GraphPad Prism software. *, p < .05.
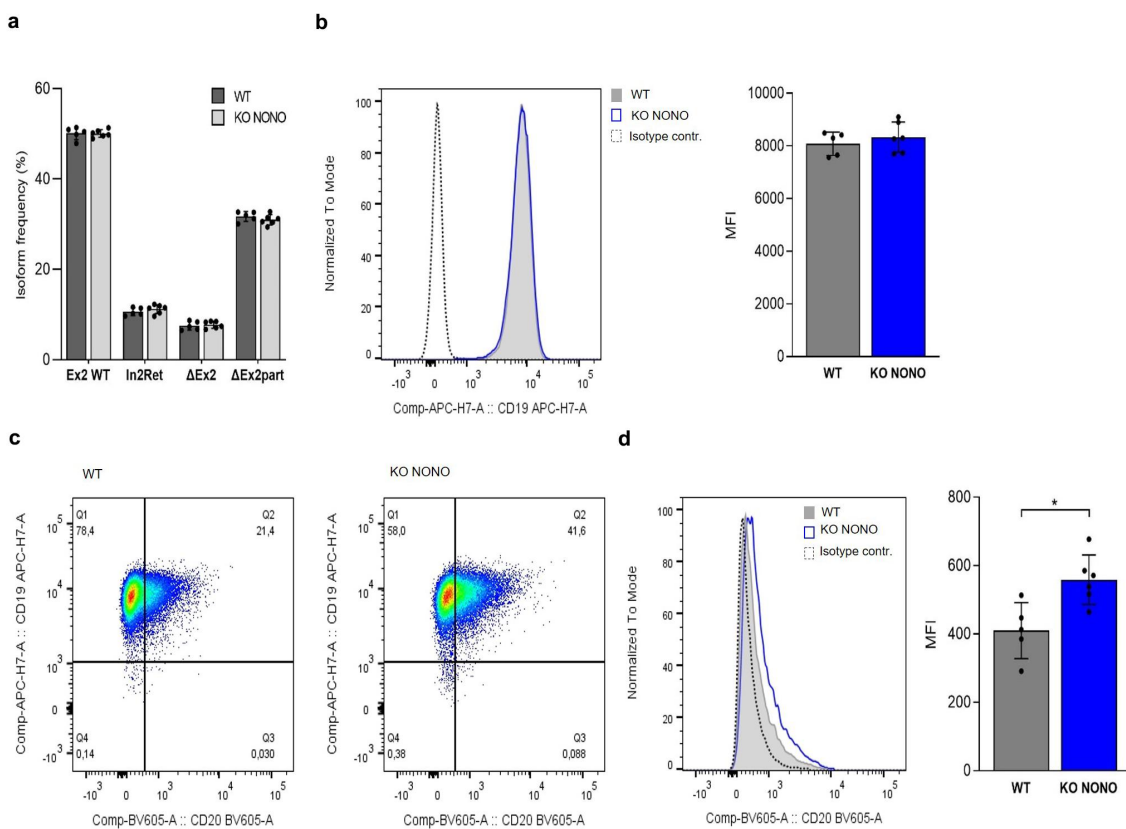
Figure 5. NONO is dispensable for CD19 splicing but regulates cell surface abundance of CD20(a) qRT-PCR analysis of *Ex2WT, In2Ret, ΔEx2* and *ΔEx2part* in WT (non-targeting control) and *NONO* KO cells. Normalized expression values were calculated as percentage of all Ex2-related CD19 isoforms. (b) Flow cytometric analysis of CD19 expression on the surface of WT and *NONO* KO cells. Histograms are shown normalized to mode, mean fluorescent intensities were used for quantification. (c) Density plots displaying CD19 and CD20 expression in WT and *NONO* KO cells. (d) Histogram view for CD20 surface expression and quantification of MFI values. *, p < .05. 5 independent experiments were analyzed.

impacts the abundance of B cell markers that serve as targets for immunotherapeutic approaches.

Discussion

The approval of CD19-directed immunotherapies dramatically improved the prognosis of patients with relapsed or refractory B-ALL and B cell lymphomas. However, about 30–50% of the patients suffer a relapse, up to 60% of them being CD19-negative.[1,3,6]

It has been shown that both CD19-specific frameshift mutations and splicing aberrations lead to epitope loss and can coexist to jointly contribute to a disease-relevant expression pattern of CD19 variants.[12–15] Alternative splicing of CD19 can thereby derive from mutations in the binding motifs as well as deregulation of RBP expression.[18] Mechanistically, it is well known that somatic mutations that disturb regulatory sequences or alter the expression of RBPs are a common cause of cancer-specific alternative splicing.[31–34] Besides directly altering the binding sites for RBPs, changes in the surrounding sequence region that is functionally relevant for binding affinity and spliceosome assembly, might impact the translation process. Functionally, isoform switches can influence mRNA stability, protein interactions and metabolic processes that finally translate into a selective advantage for tumor cells.[35,36]

The blast-specific mutation identified in our work is localized in the consensus motif of two RBPs, PTBP1 and ZFP36. The prerequisites of PTBP binding to target RNA molecules appear to be rather complex. However, there is agreement as to splicing regulation by PTBPs mostly rely on the presence of several binding sites lying in close proximity.[24,30,37] Thus, the high abundance of recognition motifs overlapping and flanking the mutated sequence rather supports than mitigates the relevance of this dedicated binding site. A multimer of the core sequence TCTT(C) embedded in a longer pyrimidine tract thereby represents a high-affinity binding site for PTB, while single guanosine nucleotides seem to be tolerated.[27] Although the TCCTC motif that is mutated in some of our leukemia patients is maintained upon the deletion, the pyrimidine tract and thereby the distance to the adjacent core motif shortens from 15 to 13 nucleotides. Furthermore, several potential binding sites are eliminated by the mutation. So far, we cannot judge whether this indeed affects PTBP1-mediated splicing, but different studies prompt that an exact distance between PTBP1 core sequences and surrounding motifs is critical for proper splicing regulation.[24,28] Specifically, pyrimidine deletions upstream or downstream of PTB binding motifs appear to be able to critically impair splicing function of PTB.[28]

Our in vitro studies revealed that PTBP1 mediates CD19 protein abundance by controlling exon 2 splicing as simultaneously suggested by another study.[18] This mechanism aligns with patient data showing that decreased *PTBP1* expression in patients at diagnosis compared to controls goes along with an increase in

intron 2 retention. Significance was not reached in our cohort due to limitations in sample size, but could be shown in the TARGET cohort.[18] Moreover, intron 2 retention was shown to increase in samples at relapse under CAR T cell therapy.[18,] Unfortunately, in our study, the same patients could not be screened by deep sequencing, so that we cannot evaluate the potential impact of CD19 mutations on these correlations. Future studies, however, will address this issue in more detail.

Due to the binding preference of ZFP36 for AU-rich elements and its general low expression across most B-ALL subtypes, we assumed a minor role of ZFP36 for CD19 exon2 processing and prioritized PTBP1 for further investigations. Additionally, while the deletion eliminates a number of PTBP1 binding motifs, the one of ZFP36 remains. However, we do not preclude a potential impact on disease-specific splicing aberrations, which might relate to its highly subtype-dependent expression profile. ZFP proteins regulate cell quiescence and proliferation of B cells and promote VDJ recombination, consequently influencing B cell development and identity.[38–40] Both PTBP1 and ZFP36 are present throughout B cell development whereby PTBP1 and CD19 underlie expression changes depending on the developmental stage.[41–43] Consequently, the effects of single RBPs and protein-protein interactions on CD19 splicing might be influenced by the B cell stage or rather the leukemia phenotype.

We expanded the expression analysis of RBPs in blasts of pediatric leukemia patients at initial diagnose and in normal B cells by further splicing factors that were selected based on our screening data and literature search.[13,43–45]

Besides *SRSF3*, which is known to be required for exon 2 inclusion,[13] and *ZFP36, NONO* was significantly less expressed in blasts compared to B cells. So far, it could not be associated with leukemia progression or CD19 splicing. Our data suggest that NONO does not impact CD19 splicing but rather regulates cell surface expression of CD20. So far, we cannot judge whether this regulation is due to a direct or an indirect mechanism, so that further investigations will be needed to evaluate its potential impact on disease-associated splicing events in B-ALL. However, NONO needs to be considered as one additional RBP defining B cell-specific and therapy-relevant marker proteins.

Different RBPs including PTBP1 are embedded in so-called mRNA regulons that coordinately regulate numerous cellular processes including immune response mechanisms.[46–49] YTHDC1, for example, is able to promote exon inclusions by recruitment of SRSF3.[44] Moreover, it was shown that PTBP2 can compensate for PTBP1 in B cells, both regulating SRSF3 activity in cancer cells. In turn, SRSF3 and other RBPs modulate PTB protein expression, illustrating that regulatory feedback mechanisms even increase the complexity of splicing events.[41,50–52] Furthermore, RBPs are generally able to modulate splicing of one another. Along this line, predominant isoforms of Heterogenous nuclear ribonucleoprotein A1 (HNRNPA1) vary between B-ALL and normal pro-B cells, which is accompanied by differences in mRNA stability.[43] Similar mechanisms certainly hold true for other RBPs that play a role in disease progression of B-ALL, and likely also for CD19 splicing. Moreover, the expression of RBPs in B-ALL samples was highly heterogeneous if analyzed across a large cohort. Hence, there is a high diversity of individual factors and regulatory circuits that together decide on the abundance of dedicated protein isoforms. However, it is conceivable that disease as well as patient-specific features define the isoform distribution of certain proteins including CD19. The disease-associated and subtype-dependent deregulation of several RBPs can serve as an evidence. Co-existence of mutations such as identified in some of our patients might underscore the idea that several cellular and molecular prerequisites combine to lay the foundation of fatal conditions misleading proper protein processing.

Given the low allele frequency, the 2-nucleotide deletion that we identified here is obviously subclonal and its relevance for the etiology of B-ALL remains to be elucidated. Yet, we cannot judge whether this mutation might accumulate in CD19-negative samples at relapse after CART-19 therapy, as such samples were only analyzed by whole exome sequencing so far and intron mutations were not investigated.[12]

Interestingly, the mutation appeared to be accumulated in B-ALL patients harboring a ETV6-RUNX1 gene fusion, as 80% of samples of this molecular subtype carried the 2-nucleotide deletion. However, a higher number of samples will be required to corroborate such correlation. Strikingly, the deletion could already be detected at diagnosis, suggesting that leukemic blasts harbor the potential to evolve into CAR-T-resistant clones directly from the beginning. With this, we pursue the idea of Rabilloud et al. who claimed the existence of CD19-negative B-ALL cells prior to CAR-T treatment.[16] Interestingly, we did not find any blast-specific mutations in the coding sequence. Thus, frameshift mutations may occur later during disease development or rather under therapy and a potential correlation with the treatment regimen before CART-19 therapy should be considered.

Screening of a bigger patient cohort to substantiate our current results as well as inclusion of samples at relapse under CART-19 therapy to determine whether the mutation might be selected during disease progression or under therapy pressure would be preferable and is subject of our current efforts.

In conclusion, we show that blast-specific mutations in intron regions with potential regulatory relevance exist at diagnosis and that blasts express a complex network of deregulated RBPs that intervene in CD19 splicing. Our data further demonstrate that the entire course of treatment needs to be considered to follow up on disease-specific features that emerge at dedicated time points and establish a CD19 epitope-negative cell population. Consecutive sampling will be critical to identify predictive markers for the likelihood of the emergence of an epitope-negative cell population upon therapy. Finally, our data can serve as a source for potential RBP candidates being involved in leukemia-specific and disease-relevant splicing modulation in defined molecular subtypes.

Disclosure statement

Funding

Author contributions

NZ, MCL, MS, AU, CP were responsible for conducting and evaluating the experiments and preparing the figures. NZ, MS and AU performed statistical analysis. NZ and CP wrote the manuscript. LR provided technical assistance for sample preparation and flow cytometric analysis. NL assisted in the evaluation of the flow cytometry data. SA performed cell sorting. CP, JK and NZ designed the study. FA, KEM, AW, AR, OB provided the patient population. FA and KEM provided guidance for patients' selection. FA and JF obtained the ethical approval for the study. CP and JF were responsible for project administration and supervision. CP, JF and JK acquired the funding. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The authors confirm that the data supporting the findings of the present study are available within the article and its supplementary material. Raw data is available on request from the corresponding author [CP].

Following St. Jude Cloud datasets were used for RNAseq analysis: Pediatric Cancer Genome Project (PCGP): This study makes use of data generated by the St. Jude Children's Research Hospital – Washington University Pediatric Cancer Genome Project and/or Childhood Solid Tumor Network.[53] Genomes for Kids (G4K): This study makes use of data generated by the St. Jude Children's Research Hospital Genomes for Kids Study.[54] Real-Time Clinical Genomics (RTCG): This study makes use of data generated by St. Jude Children's Research Hospital.[55] Pan-Acute Lymphoblastic Leukemia (PanALL): This study makes use of data generated by the Pan-Acute Lymphoblastic Leukemia Data Set of St. Jude Children's Research Hospital.[56–59]

References

1. Maude SL, Frey N, Shaw PA, Aplenc R, Barrett DM, Bunin NJ, Chew A, Gonzalez VE, Zheng Z, Lacey SF, et al. Chimeric antigen receptor T cells for sustained remissions in leukemia. N Engl J Med. 2014;371(16):1507–1517. doi:10.1056/NEJMoa1407222.
2. Park JH, Riviere I, Gonen M, Wang X, Senechal B, Curran KJ, Sauter C, Wang Y, Santomasso B, Mead E, et al. Long-term follow-up of CD19 CAR therapy in Acute Lymphoblastic Leukemia. N Engl J Med. 2018;378(5):449–459. doi:10.1056/NEJMoa1709919.
3. Sterner RC, Sterner RM. CAR-T cell therapy: current limitations and potential strategies. Blood Cancer J. 2021;11(4):69 (2021). doi:10.1038/s41408-021-00459-7.
4. Shah NN, Fry TJ. Mechanisms of resistance to CAR T cell therapy. Nat Rev Clin Oncol. 2019;16(6):372–385. doi:10.1038/s41571-019-0184-6.
5. Maude SL, Laetsch TW, Buechner J, Rives S, Boyer M, Bittencourt H, Bader P, Verneris MR, Stefanski HE, Myers GD, et al. Tisagenlecleucel in children and young adults with B-Cell Lymphoblastic Leukemia. N Engl J Med. 2018;378(5):439–448. doi:10.1056/NEJMoa1709866.
6. Xu X, Sun Q, Liang X, Chen Z, Zhang X, Zhou X, Li M, Tu H, Liu Y, Tu S et al. Mechanisms of relapse after CD19 CAR T-cell therapy for acute Lymphoblastic Leukemia and its prevention and treatment Strategies. Front Immunol. 2019;10:2664. doi:10.3389/fimmu.2019.02664.
7. Guo Z, Tu S, Yu S, Wu L, Pan W, Chang N, Zhou X, Song C, Li Y, He Y, et al. Preclinical and clinical advances in dual-target chimeric antigen receptor therapy for hematological malignancies. Cancer Sci 2021;01(9ed2021):1357–1368. doi:10.1111/cas.14799.
8. Fry TJ, Shah NN, Orentas RJ, Stetler-Stevenson M, Yuan CM, Ramakrishna S, Wolters P, Martin S, Delbrook C, Yates B, et al. CD22-targeted CAR T cells induce remission in B-ALL that is naive or resistant to CD19-targeted CAR immunotherapy. Nat Med. 2018;24(1):20–28. doi:10.1038/nm.4441.
9. Spiegel JY, Patel S, Muffly L, Hossain NM, Oak J, Baird JH, Frank MJ, Shiraz P, Sahaf B, Craig J, et al. CAR T cells with dual targeting of CD19 and CD22 in adult patients with recurrent or refractory B cell malignancies: a phase 1 trial. Nat Med. 2021;27 (8):1419–1431. doi:10.1038/s41591-021-01436-0.
10. Shah NN, Johnson BD, Schneider D, Zhu F, Szabo A, Keever-Taylor CA, Krueger W, Worden AA, Kadan MJ, Yim S, et al. Bispecific anti-CD20, anti-CD19 CAR T cells for relapsed B cell malignancies: a phase 1 dose escalation and expansion trial. Nat Med. 2020;26 (10):1569–1575. doi:10.1038/s41591-020-1081-3.
11. Larson SMW, Ji CM, Ghafouri B, Naparstek SN, Trent, J J. CD19/CD20 bispecific chimeric antigen receptor (CAR) in naïve/memory T cells for the treatment of relapsed or refractory non-hodgkin lymphoma. Cancer Discov. 2022;13:1–18. doi:10.1158/2159-8290.CD-22-0964.
12. Orlando EJ, Han X, Tribouley C, Wood PA, Leary RJ, Riester M, Levine JE, Qayed M, Grupp SA, Boyer M, et al. Genetic mechanisms of target antigen loss in CAR19 therapy of acute lymphoblastic leukemia. Nat Med. 2018;24(10):1504–1506. doi:10.1038/s41591-018-0146-z.
13. Sotillo E, Barrett DM, Black KL, Bagashev A, Oldridge D, Wu G, Sussman R, Lanauze C, Ruella M, Gazzara MR, et al. Convergence of acquired mutations and alternative splicing of CD19 enables resistance to CART-19 immunotherapy. Cancer Discov. 2015;5 (12):1282–1295. doi:10.1158/2159-8290.CD-15-1020.
14. Asnani M, Hayer KE, Naqvi AS, Zheng S, Yang SY, Oldridge D, Ibrahim F, Maragkakis M, Gazzara MR, Black KL, et al. Retention of CD19 intron 2 contributes to CART-19 resistance in leukemias with subclonal frameshift mutations in CD19. Leukemia. 2020;34 (4):1202–1207. doi:10.1038/s41375-019-0580-z.
15. Bagashev A, Sotillo E, Tang CH, Black KL, Perazzelli J, Seeholzer SH, Argon Y, Barrett DM, Grupp SA, Hu CC, et al. CD19 alterations emerging after CD19-directed immunotherapy cause retention of the Misfolded protein in the endoplasmic reticulum. Mol Cell Biol. 2018;38(21):21. doi:10.1128/MCB.00383-18.
16. Rabilloud T, Potier D, Pankaew S, Nozais M, Loosveld M, Payet-Bornet D. Single-cell profiling identifies pre-existing CD19-negative subclones in a B-ALL patient with CD19-negative relapse after CAR-T therapy. Nat Commun. 2021;12:865.
17. Fischer J, Paret C, El Malki K, Alt F, Wingerter A, Neu MA, Kron B, Russo A, Lehmann N, Roth L, et al. CD19 Isoforms enabling Resistance to CART-19 immunotherapy are expressed in B-ALL patients at initial diagnosis. J Immunother. 2017;40(5):187–195. doi:10.1097/CJI.0000000000000169.
18. Cortés-López M, Schulz L, Enculescu M, Paret C, Spiekermann B, Quesnel-Vallières M, Torres-Diz M, Unic S, Busch A, Orekhova A, et al. High-throughput mutagenesis identifies mutations and RNA-binding proteins controlling CD19 splicing and CART-19 therapy resistance. Nat Commun. 2022;13(1). doi:10.1038/s41467-022-31818-y.
19. Giudice G, Sanchez-Cabo F, Torroja C, Lara-Pezzi E. ATtRACT-a database of RNA-binding proteins and associated motifs. Database (Oxford). 2016;2016:baw035. doi:10.1093/database/baw035.
20. Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. Nat Biotechnol. 2011;29(1):24–26. doi:10.1038/nbt.1754.
21. McLeod C, Gout AM, Zhou X, Thrasher A, Rahbarinia D, Brady SW, Macias M, Birch K, Finkelstein D, Sunny J, et al. St. jude cloud: a pediatric cancer genomic data-sharing ecosystem. Cancer Discov. 2021;11(5):1082–1099. doi:10.1158/2159-8290.CD-20-1230.
22. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15(12):550. doi:10.1186/s13059-014-0550-8.
23. Stephens M. False discovery rates: a new deal. Biostatistics. 2017;18 (2):275–294. doi:10.1093/biostatistics/kxw041.
24. Perez I, Lin CH, McAfee JG, Patton JG. Mutation of PTB binding sites causes misregulation of alternative 3' splice site selection in vivo. RNA. 1997;3:764–778.

25. Feng H, Bao S, Rahman MA, Weyn-Vanhentenryck SM, Khan A, Wong J, Shah A, Flynn ED, Krainer AR, Zhang C, et al. Modeling RNA-binding protein specificity in vivo by precisely registering protein-RNA crosslink sites. Mol Cell. 2019;74(6):1189–204 e6. doi:10.1016/j.molcel.2019.02.002.

26. Dominguez D, Freese P, Alexis MS, Su A, Hochman M, Palden T, Bazile C, Lambert NJ, Van Nostrand EL, Pratt GA, et al. Sequence, structure, and context preferences of human RNA binding proteins. Mol Cell. 2018;06(9ed2018):854–67 e9. doi:10.1016/j.molcel.2018.05.001.

27. Han A, Stoilov P, Linares AJ, Zhou Y, Fu XD, Black DL, Guigo R. De novo prediction of PTBP1 binding and splicing targets reveals unexpected features of its RNA recognition and function. PLoS Comput Biol. 2014;10(1):e1003442. doi:10.1371/journal.pcbi.1003442.

28. Amir-Ahmady B, Boutz PL, Markovtsov V, Phillips ML, Black DL. Exon repression by polypyrimidine tract binding protein. RNA. 2005;11(5):699–716. doi:10.1261/rna.2250405.

29. Mukherjee N, Jacobs NC, Hafner M, Kennington EA, Nusbaum JD, Tuschl T, Blackshear PJ, Ohler U. Global target mRNA specification and regulation by the RNA-binding protein ZFP36. Genome Biol. 2014;15(1):R12. doi:10.1186/gb-2014-15-1-r12.

30. Spellman R, Llorian M, Smith CW. Crossregulation and functional redundancy between the splicing regulator PTB and its paralogs nPTB and ROD1. Mol Cell. 2007;27(3):420–434. doi:10.1016/j.molcel.2007.06.016.

31. Jung H, Lee D, Lee J, Park D, Kim YJ, Park WY, Hong D, Park PJ, Lee E. Intron retention is a widespread mechanism of tumor-suppressor inactivation. Nat Genet. 2015;47(11):1242–1248. doi:10.1038/ng.3414.

32. Supek F, Minana B, Valcarcel J, Gabaldon T, Lehner B. Synonymous mutations frequently act as driver mutations in human cancers. Cell. 2014;156(6):1324–1335. doi:10.1016/j.cell.2014.01.051.

33. Darman RB, Seiler M, Agrawal AA, Lim KH, Peng S, Aird D, Bailey SL, Bhavsar EB, Chan B, Colla S, et al. Cancer-associated SF3B1 hotspot mutations induce cryptic 3' splice site selection through use of a different branch point. Cell Rep. 2015;13(5):1033–1045. doi:10.1016/j.celrep.2015.09.053.

34. Quesada V, Conde L, Villamor N, Ordonez GR, Jares P, Bassaganyas L, Ramsay AJ, Beà S, Pinyol M, Martínez-Trillos A, et al. Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia. Nat Genet. 2011;44(1):47–52. doi:10.1038/ng.1032.

35. Climente-Gonzalez H, Porta-Pardo E, Godzik A, Eyras E. The functional impact of alternative splicing in cancer. Cell Rep. 2017;20(9):2215–2226. doi:10.1016/j.celrep.2017.08.012.

36. Peng Q, Zhou Y, Oyang L, Wu N, Tang Y, Su M, Luo X, Wang Y, Sheng X, Ma J, et al. Impacts and mechanisms of alternative mRNA splicing in cancer metabolism, immune response, and therapeutics. Mol Ther. 2021;11:19ed2021.

37. Oberstrass FC, Auweter SD, Erat M, Hargous Y, Henning A, Wenter P, Reymond L, Amir-Ahmady B, Pitsch S, Black DL, et al. Structure of PTB bound to RNA: specific binding and implications for splicing regulation. Science. 2005;309(5743):2054–2057. doi:10.1126/science.1114066.

38. Galloway A, Saveliev A, Lukasiak S, Hodson DJ, Bolland D, Balmanno K, Ahlfors H, Monzón-Casanova E, Mannurita SC, Bell LS, et al. RNA-binding proteins ZFP36L1 and ZFP36L2 promote cell quiescence. Science. 2016;352(6284):453–459. doi:10.1126/science.aad5978.

39. Zekavati A, Nasir A, Alcaraz A, Aldrovandi M, Marsh P, Norton JD, Murphy JJ. Post-transcriptional regulation of BCL2 mRNA by the RNA-binding protein ZFP36L1 in malignant B cells. PLoS One. 2014;9(7):e102625. doi:10.1371/journal.pone.0102625.

40. Newman R, Ahlfors H, Saveliev A, Galloway A, Hodson DJ, Williams R, Besra GS, Cook CN, Cunningham AF, Bell SE, et al. Maintenance of the marginal-zone B cell compartment specifically requires the RNA-binding protein ZFP36L1. Nat Immunol. 2017;18(6):683–693. doi:10.1038/ni.3724.

41. Monzon-Casanova E, Matheson LS, Tabbada K, Zarnack K, Smith CW, Turner M. Polypyrimidine tract-binding proteins are essential for B cell development. Elife. 2020;9. doi:10.7554/eLife.53557.

42. Monzon-Casanova E, Screen M, Diaz-Munoz MD, Coulson RMR, Bell SE, Lamers G, Solimena M, Smith CWJ, Turner M. The RNA-binding protein PTBP1 is necessary for B cell selection in germinal centers. Nat Immunol. 2018;19(3):267–278. doi:10.1038/s41590-017-0035-5.

43. Black KL, Naqvi AS, Asnani M, Hayer KE, Yang SY, Gillespie E, Bagashev A, Pillai V, Tasian SK, Gazzara MR, et al. Aberrant splicing in B-cell acute lymphoblastic leukemia. Nucleic Acids Res. 2018;46(21):11357–11369. doi:10.1093/nar/gky946.

44. Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, Sun H-Y, Li A, Ping X-L, Lai W-Y, et al. Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell. 2016;61(4):507–519. doi:10.1016/j.molcel.2016.01.012.

45. Katsuyama T, Moulton VR. Splicing factor SRSF1 is indispensable for regulatory T cell homeostasis and function. Cell Rep. 2021;36(1):109339. doi:10.1016/j.celrep.2021.109339.

46. Diaz-Munoz MD, Osma-Garcia IC. The RNA regulatory programs that govern lymphocyte development and function. Wiley Interdiscip Rev RNA. 2021;13(1):e1683. doi:10.1002/wrna.1683.

47. Diaz-Munoz MD, Turner M. Uncovering the role of RNA-binding proteins in gene expression in the immune system. Front Immunol. 2018;9:1094. doi:10.3389/fimmu.2018.01094.

48. Turner M, Diaz-Munoz MD. RNA-binding proteins control gene expression and cell fate in the immune system. Nat Immunol. 2018;01(20ed2018):120–129. doi:10.1038/s41590-017-0028-4.

49. Keene JD. RNA regulons: coordination of post-transcriptional events. Nat Rev Genet. 2007;8(7):533–543. doi:10.1038/nrg2111.

50. Guo J, Jia J, Jia R. PTBP1 and PTBP2 impaired autoregulation of SRSF3 in cancer cells. Sci Rep. 2015;5(1):14548. doi:10.1038/srep14548.

51. Zhou Z, Gong Q, Lin Z, Wang Y, Li M, Wang L, Ding H, Li P. Emerging Roles of SRSF3 as a Therapeutic Target for Cancer. Front Oncol. 2020;10:577636. doi:10.3389/fonc.2020.577636.

52. Izquierdo JM, Majos N, Bonnal S, Martinez C, Castelo R, Guigo R, Bilbao D, Valcárcel J. Regulation of Fas alternative splicing by antagonistic effects of TIA-1 and PTB on exon definition. Mol Cell. 2005;19(4):475–484. doi:10.1016/j.molcel.2005.06.015.

53. Downing JR, Wilson RK, Zhang J, Mardis ER, CH P, Ding L, Ley TJ, Evans WE. The Pediatric Cancer Genome Project. Nat Genet. 2012;44(6):619–622. doi:10.1038/ng.2287.

54. Newman S, Nakitandwe J, Kesserwan CA, Azzato EM, Wheeler DA, Rusch M, Shurtleff S, Hedges DJ, Hamilton KV, Foy SG, et al. Genomes for kids: the scope of pathogenic mutations in pediatric cancer revealed by comprehensive DNA and RNA sequencing. Cancer Discov. 2021;11(12):3008–3027. doi:10.1158/2159-8290.CD-20-1631.

55. Rusch M, Nakitandwe J, Shurtleff S, Newman S, Zhang Z, Edmonson MN, Parker M, Jiao Y, Ma X, Liu Y, et al. Clinical cancer genomic profiling by three-platform sequencing of whole genome, whole exome and transcriptome. Nat Commun. 2018;9(1):3962. doi:10.1038/s41467-018-06485-7.

56. Alexander TB, Gu Z, Iacobucci I, Dickerson K, Choi JK, Xu B, Payne-Turner D, Yoshihara H, Loh ML, Horan J, et al. The genetic basis and cell of origin of mixed phenotype acute leukaemia. Nature. 2018;562(7727):373–379. doi:10.1038/s41586-018-0436-0.

57. Andersson AK, Ma J, Wang J, Chen X, Gedman AL, Dang J, Nakitandwe J, Holmfeldt L, Parker M, Easton J, et al. The landscape of somatic mutations in infant MLL-rearranged acute lymphoblastic leukemias. Nat Genet. 2015;47(4):330–337. doi:10.1038/ng.3230.

58. Gu Z, Churchman ML, Roberts KG, Moore I, Zhou X, Nakitandwe J, Hagiwara K, Pelletier S, Gingras S, Berns H, et al. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. Nat Genet. 2019;51(2):296–307. doi:10.1038/s41588-018-0315-5.

59. Roberts KG, Li Y, Payne-Turner D, Harvey RC, Yang YL, Pei D, McCastlain K, Ding L, Lu C, Song G, et al. Targetable kinase-activating lesions in Ph-like acute lymphoblastic leukemia. N Engl J Med. 2014;371(11):1005–1015. doi:10.1056/NEJMoa1403088.