BCL11b interacts with RNA and proteins involved in RNA processing and developmental diseases

Haitham Sobhy[1]*, Marco De Rovere[1]*, Amina Ait-Ammar[1,2]*, Muhammad Kashif[2], Clementine Wallet[3], Fadoua Daouad[1], Thomas Loustau[3], Carine Van Lint[2 #], Christian Schwartz[3 #] and Olivier Rohr[3 #]

[1] University of Strasbourg, UR 7292, DHPI, IUT Louis Pasteur, Schiltigheim, France

[2] Université Libre de Bruxelles, ULB, Gosselies, Belgium

[3] University of Strasbourg, UPR CNRS 9002, ARN, IUT Louis Pasteur, Schiltigheim, France

* HS, MDR and AAA can be considered as equal contributors

# CVL, CS and OR can be considered as equal contributors

Correspondence should be addressed to Haitham Sobhy (haithamsobhy@gmail.com) and Olivier ROHR (olivier.rohr@unistra.fr)

Conflict of interest: Not known

Author contribution: HS conceived and designed bioinformatics analysis; HS, MDR, AAA and MK conceived and performed experiments; CW, TL, and FD provided scientific and logistic supports; CVL, CS and OR conceived the experiments, supervised the work and supported with resources. HS drafted the manuscript, all authors agreed for the final version of the manuscript.

Data deposited on SRA database with ID PRJNA661202

Keywords: Transcription factor BCL, zinc finger family, RNA processing, RNA splicing, gene expression regulation, leukemia, cell development, tumor suppressor,

ABSTRACT

BCL11b (B-cell lymphoma/leukemia 11B, CTIP2) is a C2H2 zinc-finger transcription regulator and tumor suppressor. BCL11b is involved in lymphomagenesis and in fetal, central nervous system (CNS) and immune system developments. Therefore, it may contribute in congenital disorders and cancers (e.g. leukemia). BCL11b favors persistence of HIV latency in microglia, CNS macrophages. BCL11b contributes to control cell cycle, differentiation and apoptosis in multiple organisms and cell models; however exact mechanisms are unknown. Although BCL11b recruits non-coding RNA and epigenetic enzymes to regulate gene expression, BCL11b-associated ribonucleoprotein complexes are unknown. Here, using immunoprecipitation of BCL11b-binding RNA and proteins (CLIP-seq and quantitative LC-MS/MS mass spectrometry) complemented with systems biology validations, we show that BCL11b interacts with RNA splicing and nonsense-mediated decay proteins, including FUS, SMN1, UPF1 and Drosha, which may contribute in isoform selection of protein-coding RNA isoforms from noncoding-RNAs isoforms (retained introns or nonsense mediated RNA). Interestingly, BCL11b binds to RNA transcripts and proteins encoded by the same genes (FUS, ESWR1, CHD and Tubulin). Our study highlights that BCL11b targets RNA processing and splicing proteins, and RNAs that implicate cell cycle, development, neurodegenerative, and cancer pathways. These findings will help future mechanistic understanding of developmental disorders.

IMPORTANCE

First genome-wide BCL11b-protein and RNA interactome

BCL11b interacts with RNA processing and splicing proteins

BCL11b interacts with neurodegenerative genes and sarcoma genes

BCL11b targets during cell proliferation and disease pathways

INTRODUCTION

The B-cell lymphoma/leukemia 11B protein (Bcl11b, also known as COUP-TF-interacting protein 2, CTIP2) plays crucial roles in the epigenetic regulation of gene transcription and specifically in the control of the elongation process by interacting with an inactive form of the p-TEFb complex (7SK RNA, CDK9, cyclin T1, and HEXIM1) (Cherrier et al., 2013). BCL11b, which harbors six zinc finger domains (ZnFs) of C2H2-type, shares >60% homology with the human BCL11a, as well as the mouse, chicken, and Xenopus homologues. BCL11 proteins were isolated from T-cell lines derived from patients with T-cell leukemia, suggesting a role for BCL11 proteins in blood cell development and lymphomagenesis (Chiara et al., 2021; Fu et al., 2017; Helm et al., 2023; Montefiori et al., 2021; Sidwell and Rothenberg, 2021). Moreover, BCL11b was deleted in γ-ray induced mouse lymphomas, suggesting a role as a tumor suppressor and modulator of radiation-induced DNA damage, reviewed in (Kominami, 2012).

Studying the function of BCL11b by knocking out the gene is challenging (Longabaugh et al., 2017). Mice that lack the two alleles of BCL11b die shortly after birth, exhibiting defects in multiple tissues, including the immune system, central nervous system (CNS), skin, thymus, among other organs (Eto et al., 2022; Kominami, 2012; Przybylski et al., 2022; Simon et al., 2020), which suggests the importance of BCL11b during development. For example, in brain, BCL11b serves as a marker of layer V subcerebral projection neurons and striatal interneurons (Lennon et al., 2017, 2016; Simon et al., 2020). Several reports showed that BCL11b can regulate cell / tissue development, cell cycle, cell differentiation, and apoptosis in human, mouse, chicken and Drosophila, and multiple cell models, including cancer cells, blood cells, neuronal cells and mammary duct cells (Abe et al., 2023; Cai et al., 2017; Cherrier et al., 2009; Go et al., 2012; Grabarczyk et al., 2018; Liao et al., 2016; Miller et al., 2018; Parsa et al., 2023; Yang et al., 2020; Zhang et al., 2021; Zhao et al., 2017; Du et al., 2022; Holmes et al., 2021; Liau et al., 2023; Fox et al., 2022; Zhou et al., 2022). Although few interactions are known, the

76 mechanisms by which BCL11b regulate cell proliferation and development and the target of
77 BCL11b are largely unknown (Koumoundourou et al., 2024).

78 The absence of knockout BCL11b challenges the study of BCL11b functions as well as the
79 interactions with protein and RNA in the cells. To our knowledge, BCL11b can form complex
80 with 7sk ncRNA and p-TEFb complex (Cherrier et al., 2013), or with DNA-binding proteins,
81 such as HP1 proteins (Rohr et al., 2003), NuRD (Cismasiu et al., 2008, 2005), HDACs (Marban
82 et al., 2007), HMGA1 (Eilebrecht et al., 2014), DCAF1 (Forouzanfar et al., 2019), and KU
83 proteins (Shadrina et al., 2020). The complexes including BCL11b and p-TEFb, HDACs,
84 histone methylases, or transcription factors are essential to regulate HIV latency (Cherrier et
85 al., 2013; Marban et al., 2007).

86 Although BCL11b has six C2H2 ZnF domains, to our knowledge, potential RNA-binding
87 functions have not been well characterized for any of the BCL11 proteins nor the closely related
88 proteins. In this study, we aim to identify BCL11b-binding complexes including those triggered
89 by RNA, such as p-TEFb complex. Therefore, we used immunoprecipitation of proteins
90 followed by mass spectrometry (IP-MS) as well as immunoprecipitation of cross-linked RNA
91 (CLIP-seq), to identify the proteins and RNAs bound to BCL11b. Our results reveal that
92 BCL11b can target RNA and proteins involved in neuronal development, cell cycle control,
93 RNA splicing, cancers and neurodegenerative diseases. Our study offers new insights on role
94 of BCL11b in diseases and open the route for future studies to understand the mechanisms of
95 developmental disorders.

96

97 **MATERIAL AND METHODS**

98 **Cell culture, plasmid, cell transfection and antibodies**

99 The human microglial and HEK293T cell lines were maintained in Dulbecco's modified Eagle
100 medium (DMEM). The culture media were complemented with 10% (vol/vol) heat-inactivated
101 fetal bovine serum (FBS) and 100 U/mL penicillin/streptomycin. The construct pFLAG-
102 BCL11b and control pFLAG-pcDNA3 were previously used in (Cherrier et al., 2013).
103 HEK293T cell cultured in 150-mm diameter dishes were transfected using a calcium phosphate
104 co-precipitation method with the indicated plasmids. The whole cell lysates were prepared 48
105 hours post-transfection as previously described in (Cherrier et al., 2013). The following
106 antibodies were used; BCL11b (A300–383A, Bethyl Laboratories), Goat anti-Rabbit IgG
107 (HRP; 32460, Thermo Fisher), Drosha (ChIP Grade ab12286, Abcam), Mouse anti-IgG
108 (CS200621, Millipore), Rabbit ani-IgG (PP64B, Upstate), Anti-flag M2 Affinity Gel (A2220,
109 SIGMA), and FUS, UPF1 and m-IgGκ BP-HRP (sc-47711, sc-393594 and sc-516102, Santa
110 Cruz).

111

112 **Protein extraction and co-immunoprecipitation (co-IP) for mass spectrometry**

113 The cells seeded at a density of 3×10^6 cells per dish and transfected with the indicated vectors
114 (30 µg) using the method of co-precipitation with calcium phosphate. After 48 hours of
115 transfection, the cells were lysed for 10 minutes on ice in a buffer containing: 10 mM HEPES,
116 1.5 mM MgCl2, 10 mM KCl, 0.5 mM DTT. After one minute centrifugation at 13,000 g, the
117 nuclear pellet was lysed for 30 minutes on ice in another buffer containing: 20 mM HEPES,
118 25% glycerol, 1.5 mM MgCl2, 420 mM NaCl, 0.2 mM EDTA, 0.5 mM DTT. After 2 minutes
119 centrifugation at 13000 g, the supernatant was collected and stored for further analysis by
120 immunoprecipitation and western blot. Anti-proteases were systematically added to the lysis
121 buffers. Immunoprecipitation was carried out on 500 µg of nuclear protein extracts by using the
122 M2 anti-flag gel from SIGMA. The immunoprecipitated protein complexes were subjected to

123 standard western blot analyzes. Proteins were detected using antibodies and visualized by a
124 chemiluminescence detection system (Thermo Scientific Pierce ECL Western Blotting
125 Substrate 32106).

126

127 **Mass spectrometry**

128 HEK293T cells were transfected with the flag-BCL11b vector. The nuclear proteins were
129 extracted and subjected to anti-flag immunoprecipitation. Immunoprecipitated proteins were
130 eluted and then digested as described in this previous study (Turriziani et al., 2014). The
131 peptides were separated and analyzed by reverse phase liquid chromatography coupled to a
132 high-resolution mass spectrometer (Q-Exactive connected to an Ultimate Ultra3000
133 chromatography system (Thermo Scientific, Germany)). Mass spectra were analyzed using
134 MaxQuant software. The results are representative of three independent experimental
135 duplicates. The results were obtained in the form of protein intensities calculated by
136 summarizing the ion intensities of the peptides which uniquely correspond to the protein
137 sequence identified in all the samples (Turriziani et al., 2014). The proteins interacting with
138 BCL11b were determined after normalization with the trypsin used for peptide digestion.
139 Normalization with a control condition represented by cells transfected with the empty vector
140 pcDNA3 has also been performed.

141

142 **Cross-Linking Immunoprecipitation Sequencing (CLIP-Seq):**

143 The method has been described before (Maurin et al., 2015; Tabet et al., 2016). For short, cells
144 were cultured and transfected with the pFLAG-BCL11b and pCDNA3 as control. The day prior
145 the cell lysis the magnetic beads were prepared and incubated with lysis buffer only, with
146 unspecific IgG, with anti-flag or with anti-BCL11b depending of the cell lysate to be analysed.
147 Prior to lysis , cells were covered by ice cold PBS and exposed to UV light without for the
148 crosslinking step by using the Stratalinker for 10 minutes at 4000J on ice. The cells were then

149 lysed with lysis buffer (Tris HCl pH 7.4 50mM, $CaCl_2$ 0.08mM, SDS 0.1%, KCl 100 mM,
150 $MgCl_2$ 2mM, NP40 1%, Deoxycholate 5g/L) complemented with RNasin (Promega) at 40U/ml
151 and Proteinase Inhibitor Cocktail (Roche). 10 Units of TURBO-DNase (Ambion) was added
152 for 5 minutes at 37 °C and the lysates were centrifuged for 15 minutes 13000 rpm 4°C. In the
153 meantime, the tubes with the beads were washed with lysis buffer. Lysates were than subjected
154 to two preclearing steps with beads alone and with IgG-beads to eliminate any unspecific
155 binding before an overnight incubations with anti-flag, anti-BCL11b or unspecific IgG at 4°C.
156 The day after the beads were washed three times with a high salt washing buffer (Tris HCl pH
157 7.4 50mM, EDTA 1mM, SDS 0.1%, KCl 1M, NP40 1%, Deoxycholate 5g/L) followed up by
158 a Proteinase K treatment, 200µl per tube (100mM Tris HCl pH 7.4, 10mM EDTA, 40mM NaCl,
159 Proteinase K 2mg/ml (Roche) for 20 minutes at 37° Celsius. The proteinase K activity was then
160 stopped by adding 200µl PK Urea buffer (100mM Tris HCl pH 7.4, 10mM EDTA, 40mM NaCl,
161 7M Urea) at 37° C. The RNA extraction step followed up by adding 500µl of phenol/chloroform
162 incubated 5 minutes at 37° C, 1100 rpm, the tubes were then centrifuged at RT at 13000rpm
163 and the aqueous phase transferred to a new tube. A solution of 3M sodium acetate pH5.5 and
164 1µl Glycoblue was added with 1ml of 1:1 Ethanol:Isopropanol, the tubes were mixed and the
165 RNA was left to precipitate overnight at -20° C. The tubes were then centrifuged at 13000rpm
166 at 4 ° C for 10 minutes, the RNA pellet was washed twice with 500µl 80% ethanol and left to
167 dry for 2 minutes with the lid open. The RNA was then dissolved in 10µl DPEC water and sent
168 to sequence.

169

170 **RNA-Immunoprecipitation (RIP):**

171 Microglial cells were cultured in 10cm dish, at 80%-90% confluence. The cells were lyzed and
172 total RNA was extracted using NucleoBond Xtra Maxi (740414, Macherey-Nagel) according
173 to the manufacturer's recommendations. The lysate was separated from the cellular debris by
174 centrifugation at 10000 rpm (10000g) at 4°C. The lysates were washed and pre-cleared for 1h
175 at 4°C with 25µl Dynabeads protein G (Invitrogen) with unspecific IgG and incubated overnight

176 at 4°C with 30µl Dynabeads previously coated with 5µg of the antibody of interest or unspecific
177 IgG as negative control for overnight at 4°C. The RNP complexes were disrupted by performing
178 an RNA extraction directly from the beads by adding 100µl chloroform, centrifuging at 13000
179 rpm (13000g) for 15 minutes. After this step, the liquid phase containing the RNA was
180 separated, mixed with isopropanol and Glycoblue (Invitrogen) to facilitate the formation of the
181 RNA pellet during 10 minutes incubation on ice and followed by centrifugation at 13000 rpm
182 at 4°C. The pellet was washed with ethanol 80% and re-suspended in 16µl of DEPC water.
183 RNAs were reverse transcribed to cDNA using iScript Reverse Transcription Supermix for RT-
184 qPCR (BIO-RAD) in a final volume of 20ul. cDNA was then diluted 5 times and 2µL are used
185 for the qPCR analysis with SsoAdvanced Universal SYBR Green Supermix (BIO-RAD).

186

187 **Protein extraction and co-immunoprecipitation for western-blotting**

188 HEK293T cells were transfected with Flag-BCL11b and control plasmids with JetPei
189 transfection reagent. After 48 hours, proteins were extracted with a RIPA buffer (150 mM 186
190 NaCl, 1% Nonidet P-40, 0.1% SDS (10%), 1.5mM MgCl2, 50 mM Tris-HCl, pH 8). 4 µg of anti-
191 flag or anti-IgG control antibodies were incubated with Dynabeads protein G in low salt IPLS
192 buffer (50 mM Tris HCL PH 7.5, 120 mM NaCL, 0.5 mM EDTA, 0.5 % Nonidet P-40, 10 %
193 Glycerol) and washed before addition of 1mg of protein extracts and incubated overnight at
194 4°C. The IP samples were washed twice with 1 ml of IPLS buffer at 4°C, then twice with high
195 salt concentration buffer IPHS (50 mM Tris HCL PH 7.5, 0.5 M NaCL, 0.5 mM EDTA, 0.5 %
196 Nonidet P-40, 10 % Glycerol) and finally twice with IPLS. EDTA-free Protease Inhibitor Cocktail
197 (Roche) was added to the IPLS and IPHS buffers used for the washing in order to avoid protein
198 degradation. SDS PAGE gel electrophoresis was performed and the proteins were transferred
199 to a nitrocellulose membrane using the Trans-Blot Turbo Transfer system from Bio-Rad. After
200 the transfer of proteins, the membranes were incubated for 5 minutes in the blocking buffer
201 from Bio-Rad (Every-blot blocking buffer, Ref#12010020). After the blocking step, the

202 membranes were incubated with primary and secondary antibodies mentioned in Annex A
203 (The antibodies dilutions were made in the Every-blot blocking buffer). After immunoblotting,
204 the detection was carried out using the Super Signal Chemiluminescence Detection System
205 (Thermo Fisher), and were visualized by Fusion imaging system (Vilber Lourm).

206 **PCR and RNA quantification**

207 The total RNA was extracted using NucleoBond Xtra Maxi (740414, Macherey-Nagel)
208 according to the manufacturer's recommendations. RNA was reverse transcribed into cDNA
209 using iScript Reverse Transcriptase Supermix (BIO-RAD) and oligo 12–18 Primers (Thermo
210 Fisher). The quantitative PCR was performed on the diluted cDNA (in DNase-free water) using
211 SsoAdvanced Universal SYBR Green Supermix (BIO-RAD). The PCR products were then run
212 into gel to confirm the size. The primers pairs used in the study are listed in table S4. Additional
213 PCR test were performed to confirm non-genomic DNA contaminations.

214

215 **Mass spectrometry bioinformatics analysis**

216 Proteins with three or more MS signals (MS/MS Count) were selected. The proteins should
217 have four or more peptides and at least one unique peptide to be selected. In case of two or three
218 peptides with at least two razor + unique peptides were selected, on condition that the MS/MS
219 count > 3. The unique sequence coverage was taken into consideration (over 1%). In all cases,
220 the p-values between the condition and control were calculated using student t-test. Proteins
221 with p-values < 0.05, and two-fold increase (1 log2 fold increase) were selected. For
222 downstream GO analysis, we used combination of methods, including the GO construction
223 (http://geneontology.org/) and DAVID database (https://david.ncifcrf.gov/). Additionally, we
224 downloaded OBO gene annotation files for proteins and ncRNA from
225 (http://geneontology.org/docs/download-ontology/) in 2018. For protein complexes, we
226 obtained the data from CORUM database (http://mips.helmholtz-muenchen.de/corum/).

227

228 **CLIP sequencing and bioinformatics analysis**

229 After CLIP-seq experiments, the RNA library was sequenced on Illumina Hiseq 4000 sequencer
230 as Single-Read 50 base reads following Illumina's instructions. FastQC
231 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was run to produce base quality.
232 After adapters were cut, the short reads less than 10 bases were removed. Reads were mapped
233 onto the *Homo sapiens* genome (hg19 assembly) using Bowtie2 (http://bowtie-
234 bio.sourceforge.net/bowtie2/), Samtools (http://samtools.sourceforge.net/), or STAR aligner
235 (https://github.com/alexdobin/STAR). The number of reads per gene, transcript or exon
236 counted using featureCounts, Subread package (http://subread.sourceforge.net/). The genome
237 assembly file was downloaded from GENCODE (https://www.gencodegenes.org/). The
238 protein-coding genes, exons, TSS and transcript annotation file was downloaded from UCSC
239 (http://hgdownload.cse.ucsc.edu/) or Ensembl (https://grch37.ensembl.org/), RNA annotation
240 from (https://rnacentral.org/), and OBO gene from (http://geneontology.org/) in 2018. The
241 CLIP-seq reads are visualized using Integrative Genomics Viewer (IGV) software
242 (https://igv.org/).

243 The three control conditions and three replicates in both HEK cells and microglial were used
244 for the analyses. The uniquely mapped reads (MAPQ=255) were selected. This leads to on
245 average 1.1 or 2.3 for million unique hits from control and replicate HEK experiments,
246 respectively, and 1.1 or 2.1 million unique hits from control and replicate of microglial
247 experiments, respectively. Then average of RNA CLIP tags (reads) in control conditions and
248 replicates were calculated for whole human genome. In case of microglial, the top binding genes
249 should pass three cut-offs, RNA CLIP tags over 4-fold ($\log_2$=2) to the control, the number of
250 tags of experiments should be more than 1000 those in control cases (average of number of tags
251 of experiments – average of control ≥1000), and the p-values ≤ 0.05 cut-offs. For HEK cells,
252 two-fold the control ($\log_2$ fold change =1), the average of number of tags of experiments –
253 average of control ≥200, and p-values ≤ 0.1. We selected the genes that passed the cut-offs in

254 both microglial and HEK cells datasets at the same time. This leads to a final dataset of 740
255 genes, which encode RNA transcripts bind to BCL11b as shown by CLIP-seq method. The
256 dataset then used for further GO analysis and other bioinformatics analyses.

257 To quantify the amount of RNA CLIP reads per transcripts, exons, and introns features, subread
258 package was used. For this step un-filtered BAM files were used, i.e., the reads can be mapped
259 to multiple features. Number of exons and introns were obtained from Ensembl BioMart or
260 UCSC refGene file. To get insights into the differential binding between BCL11b and the
261 different transcripts encoded from the same gene, we chose 23 genes. These genes are enriched
262 with significant CLIP reads and they characterize by the encoding both protein-coding and non-
263 coding transcripts. Boxplots were plotted using R BoxPlotR tool
264 (http://shiny.chemgrid.org/boxplotr/). Genes were visualized by UCSC browser
265 (https://genome.ucsc.edu/) or IGV (http://software.broadinstitute.org/software/igv/).

266 For over-expressed clusters, the significant cluster is two-fold the control. As a second method,
267 we counted the number of RNAs per 10kb-long bins of human genome. The clusters that pass
268 the threshold of 5-fold to the control were selected, leading to 646 cluster. The default
269 parameters were selected. We then selected the clusters that experiments are enriched over
270 control. This leads to results of 1466 and 1583 clusters in HEK293T and microglial,
271 respectively. In all cases, the significant clusters were selected for gene ontology and disease
272 ontology using DAVID, GO construction and GREAT tool (http://great.stanford.edu/). The R
273 packages "ClusterProfiler, EnrichGO, and Enrichplot" are used to determine the enrichment
274 scores. The "Pathview" R package is used to construct KEGG pathways. Volcano plots are
275 constructed using "EnhancedVolcano" R package.

276

277 **RESULTS**

278

279 **BCL11b is associated with cellular RNA**

280 Although BCL11b has six C2H2 ZnF domains, to our knowledge, potential RNA-binding
281 functions have not been well characterized for any of the BCL11 proteins; except BCL11b can
282 bind the 7SK RNA to regulate transcription elongation (Cherrier et al., 2013; Eilebrecht et al.,
283 2014). To examine the RNA-binding activity and obtain a transcriptome-wide landscape of
284 BCL11b-bound RNAs, we cross-linked BCL11b to cellular RNAs using UV radiation,
285 followed by immunoprecipitation and sequencing (CLIP-seq). These experiments were
286 performed by targeting endogenous BCL11b in microglial cells, or Flag-BCL11b in HEK293
287 cells. Microglia were previously used as model cells to study roles of BCL11b on gene silencing
288 and HIV latency (Cherrier et al., 2013; Desplats et al., 2013; Eilebrecht et al., 2014). The top
289 genes that passed the cut-offs (see material and methods) in each cell line were only selected
290 which led to a dataset of 740 different genes that have high CLIP-seq reads in both HEK and
291 microglial cells (Figure S1, S2, Suppl. data S1), we call this dataset as BCL11b-CLIP-gene-set.
292 Interestingly, we found that over 90% of RNAs in this BCL11b-CLIP-gene-set are mapped to
293 protein-coding genes (Figure S1, Table S1, and Suppl. data S1). GO terms of the BCL11b-
294 CLIP-gene-set include genes belong to development of the brain, neuron, hemopoiesis, blood
295 vessels, and hair follicles (details are found in Figure S3, Suppl. data S2). Additionally, BCL11b
296 binds to RNA transcribed from genes involved in RNA processing, for example, BCL11b binds
297 to protein-coding RNAs and non-coding RNAs (ncRNAs), involved in biogenesis of the major
298 and the minor spliceosome (Figure S1, Table S2, Suppl. data 2).

299 The pathway analysis of BCL11b-CLIP-gene-set reveals regulation of RNA transcription of
300 multiple immune response pathways including Wnt signaling and MAPK signaling (Figure 1A,
301 Figure S2, Suppl. data 2). BCL11b can regulate the same genes in Wnt signaling in different
302 cell lines, including microglia and HEK cells (Figure 1B, S3, S4, Suppl. data 1). Noteworthy,
303 we collected list of genes that harbor the top BCL11b CLIP-seq reads and compared with

304 previously published differential expression transcriptomics data (by microarray and RNA-seq
305 analyses), we observe that our study identifies novel genes and pathways that have not been
306 previously described (Figure 1B, S4) (Cherrier et al., 2013; Drashansky et al., 2021, 2019;
307 Eilebrecht et al., 2014; Hasan et al., 2019; Hosokawa et al., 2018; Longabaugh et al., 2017;
308 Tang et al., 2011). Since BCL11b functions on gene transcription have been previously
309 reported, these results suggest that BCL11b may also bind some RNA to regulate their
310 processing. Surprisingly, we found CTIP2 bound to RNA coding for regulators of the RNA
311 processing and RNA splicing pathways, such as EWSR1, FUS, SFPQ, SRRM2, and DNMT1,
312 in addition to ncRNA (Figure 1C, and figures S1), specifically to exon and exon-intron junction
313 regions (Figure 1D, E, Figure S5-S6). Noteworthy, we could not observe significant changes in
314 the expression of CTIP2-bound protein coding RNA upon over-expression of BCL11b (Figure
315 1E).

316 To confirm these results, we divided the whole genome into 10kb regions (bins) and counted
317 the number of reads corresponding to each bin (see material and methods section). For
318 additional confirmation, we performed peak calling to cluster the RNAs into bins. We then
319 selected the regions that were enriched over the control replicates of the two cell lines. The
320 results showed that over 40% of CLIP reads mapped within 0-50kb down-stream of the
321 transcription start sites (TSS), and about one third up-stream to TSS (Figure S7-S8, Suppl. data
322 S3). Gene ontology analysis for the genes localized in these regions suggested their
323 involvement in angiogenesis, response to stress, and inflammation pathways, whereas disease
324 ontology shows involvement in cancers, Alzheimer, and Huntington diseases (Figure S7-S8,
325 Suppl. data S3).

326

327 Taken together, our results demonstrate that half of the BCL11b-binding RNAs are mapped to
328 gene body regions, suggesting that BCL11b may regulate co-transcriptional RNA processing
329 and/or mRNA biogenesis.

330

331 **BCL11b binds to multiple RNA isoforms of the same genes**

332 To further examine the association of BCL11b with RNA transcripts (isoforms), we counted
333 the number of reads per exon for CLIP reads that mapped to genes. We found that about 75%
334 of the CLIP reads mapped to exons (Figure 2A). Analysis of the intronic reads revealed that
335 about 25% of the reads corresponded to isoforms that will not yield stable protein products (we
336 refer to them here as isoforms that does not code for proteins, INCP). For example, number of
337 BCL11b CLIP-seq reads are mapped to intronic regions within EWSR1, FUS, and SRRM2
338 genes, which code for retained introns (IR), nonsense mediated decay (NMD), non-stop decay
339 or processed transcript, (Figure 1C, Suppl. Figure S1). This result is consistent with our
340 proteomics data (shown in the further section) demonstrating that BCL11b interacts with
341 proteins involved in RNA splicing and NMD pathways, including FUS, SMN1, RNPs, Drosha,
342 and UPF1.

343 To further test this finding, we collected the transcripts (or RNA isoforms) of 23 genes that
344 encode isoforms coding for proteins (ICPs) and isoforms that do not code for proteins (INCPs),
345 the latter would be indicative of NMD or IR transcripts / isoforms. We then counted the number
346 of reads that correspond to each isoform of the gene, including the protein-coding isoforms
347 (ICPs) or those not coding for proteins (INCPs). We found that BCL11b binds to different
348 isoforms of the same gene in differential manner (Figure 2B, Suppl. data S4). For example, the
349 TOP3A gene encodes 21 different RNA isoforms, similarly ESWR1, Drosha, SIRT7 and
350 RAD52 genes encode 20-21 RNA isoforms. The fold increase to control of CLIP-seq for
351 TOP3A is about 1.2-folds, whereas the binding to RNA of RAD52 is up to 7-folds increase
352 (Figure 2B). The same concept applies to SMN1 (10 isoforms), COL27A1 (8 isoforms) and
353 RAD51 (11 isoforms). We counted the CLIP-seq reads per each transcript (or isoform) and then
354 normalized to the length of the transcripts. BCL11b favors binding to long isoforms coding for
355 protein (ICPs) than protein non coding (INCPs), Spearman correlation rank (R) between the
356 CLIP reads and transcripts length for ICPs is 0.51 and 0.44 for INCP, which suggest that
357 transcript length could be a contributing but not a major factor (Figure 2C). Together these

358 results suggest that BCL11b can distinguish between different RNA isoforms of the same gene,
359 which is crucial function for a protein involved in RNA metabolism or cell development.

360 We still noticed some CLIP reads that correspond to intronic regions within ICPs, but the CLIP
361 reads are mapped to the coordinates of exons in other INCP isoforms. Noteworthy, NMD or IR
362 transcripts introduce additional exons instead of introns. To explain this concept, we
363 highlighted the genomics introns in pink colors, see (Figure 1C and S1). Although these intron
364 regions are usually removed in case of ICP transcripts (thick dark bars in the panel e of Figure
365 1C and S1), we found that part of these regions corresponded to exons in INCPs (light blue bars
366 in e panel in Figure 1C and S1). Computationally, we counted number of reads mapped to exons
367 and introns, we normalized them to the number of exons or introns as shown in Figure 2D and
368 Suppl. data S4. If BCL11b binds to exons regions of RNA, we expect to find that most of the
369 CLIP reads are biased to exons coordinates. However, in Figure 2D, we observed that most of
370 the reads are mapped to introns. In case of INCPs, the number of reads were almost equally
371 distributed throughout the exons and introns. Statistically, we calculated the Spearman
372 correlation rank (R) between number of reads and number of exons, which show that number
373 of exons correlates with number of CLIP reads, in case of ICPs, but not in case of INCPs (R of
374 INCPs=0.2, and R of ICPs=0.46). This result suggests that INCPs with few numbers of exons
375 have multiple CLIP reads. The DNA (cytosine-5-)-methyltransferase 1 (DNMT1) transcripts
376 are good examples that support this observation. Number of reads mapped to retained introns
377 (IR) IR transcripts are 2-3-fold higher than the number of reads mapped to INCPs transcripts.
378 Indeed, the 5 exons IR transcripts (ENST00000591764, with 589 reads) have 3-folds more
379 reads than the 9 exons protein coding transcripts (ENST00000588952, with 196 reads) (Figure
380 2E, Suppl. data S4).

381 As mentioned above, comparing top genes that harbor highest CLIP-reads with publicly
382 available transcriptomics data showed that most of the genes are not differentially expressed
383 after BCL11b loss or gain of function (Cherrier et al., 2013; Drashansky et al., 2021, 2019;

384 Eilebrecht et al., 2014; Hasan et al., 2019; Hosokawa et al., 2018; Longabaugh et al., 2017;
385 Tang et al., 2011).

386 Taken together, our CLIP-seq results show that BCL11b can recognize the different type of
387 RNA isoforms (either protein coding or not coding), and binds to the isoform in RNA type-
388 specific manner. This characteristic is very important for the proteins that play roles in selecting
389 the correct isoforms to be translated (i.e. isoform selection). On other words, this finding
390 suggests that BCL11b may bind to specific RNA isoforms to regulate RNA processing or
391 metabolism at post-transcriptional levels. However, experimental confirmations will be needed
392 to confirm this hypothesis.

393

394 **Quantitative protein interactomics of BCL11b**

395 To characterize the protein interactants and pathways triggered by BCL11b, we used first HEK
396 cells expressing Flag-BCL11b. After isolation of BCL11b and associated proteins using anti-
397 flag antibodies, we performed quantitative LC-MS/MS. After statistical analysis, we identified
398 629 BCL11b protein partners (>two-fold change and p-value <0.05; Figure S9, supplementary
399 dataset S5).

400

401 ***BCL11b interacts with DNA damage repair proteins and transcription regulators (e.g. HIV-***
402 ***1 transcription)***

403 Our proteomics analysis shows that BCL11b binds to chromatin organization, and DNA
404 recombination proteins, including HP1 and DCAF1 proteins, and components of NuRD and
405 SNW1 complexes. Additionally, BCL11b binds to protein regulators of DNA damage repair,
406 such as H2AX, BRCA2, MCMs, PARP1, and SMCs proteins. BCL11b regulate the
407 transcription through interactions with multiple epigenetics and transcription regulators
408 including CCNT1, CDK9, SMARCAs and SMARCCs transcription regulators, as well as
409 histone methyltransferases and acetyltransferases (KMT2A and HDACs). In agreement with

410 our results, previous publications from our lab and colleagues reported that over-expression of
411 BCL11b in HEK cells and other cell lines shows similar interactions, including HDACs, HP1,
412 HMGA1, DCAF1, KU proteins, NuRD complex, and P-TEFb complex (CDK9 and CycT1)
413 (Cherrier et al., 2013; Cismasiu et al., 2008, 2005; Eilebrecht et al., 2014; Forouzanfar et al.,
414 2019; Marban et al., 2007; Rohr et al., 2003; Shadrina et al., 2020; Vickridge et al., 2024)
415 (Figure 3A, S10 and suppl. data S5). The interactions between BCL11b and the transcription
416 regulators were shown to be essential during HIV-1 latency, or epigenetics modifications and
417 DNA damage repair mechanisms. Noteworthy, some of these interactions can be dissociated
418 after RNase treatment, such as BCL11b-P-TEFb interaction, but not BCL11b-HDAC, which
419 means that some of these complexes are RNA-dependent (Cherrier et al., 2013). Noting that we
420 detected a hit of ribonucleases in our MS results, the role of BCL11b in these ribonucleoprotein
421 complexes deserves to be studied by future independent studies.

422

423 ***BCL11b interacts with proteins dedicated to cell cycle and epigenetic regulations.***

424 We observed that BCL11b interacts with proteins and complexes that are involved in DNA
425 replication and cell cycle (Figure 3A, S10, S11 and supplementary data S5), including CDK
426 proteins. Cyclin-dependent kinases (CDK) are group of proteins that regulate cell cycle and
427 transcription by phosphorylating their substrates. CDK1, 2, 4 and 6 among other were found to
428 bind to BCL11b (Figure S11). In agreement with previous reports that showed BCL11b can
429 regulate cell cycle, cell differentiation, and apoptosis by interacting with long non-coding
430 RNAs and impairs p21, p16, cyclin D1 and CDK2 pathways (Abe et al., 2023; Cai et al., 2017;
431 Cherrier et al., 2009; Go et al., 2012; Liao et al., 2016; Parsa et al., 2023; Zhang et al., 2021;
432 Zhao et al., 2017). However, the exact mechanism by which BCL11b impairs cell proliferation
433 is not known. Our data highlight the interactions between BCL11b with multiple cyclin
434 dependent kinases (e.g. CDK1, 2, and 4/6), cell division control protein (e.g. CDC7, CDC20,
435 and CDC45), and minichromosome maintenance complex (MCM) proteins, in addition to
436 SKP1 a member of the SCF ubiquitin ligase protein complex.

437

438 ***BCL11b interacts with proteins involved in translation and RNA splicing.***

439 Among the novel findings of our study is the interactions between BCL11b with multiple
440 proteins involved in RNA processing and translation, such as 40S and 60S ribosomal subunits
441 (Figure 3A-D, S10, S12, supplementary data S5). Moreover, BCL11b interacts with many
442 RNA-binding proteins, including FUS, spliceosomal subunits, small nuclear RNP (snRNPs),
443 heterogeneous nuclear RNPs (hnRNP), serine/arginine (SR)-rich proteins and the survival
444 motor neuron (SMN) protein complex. Overall, BCL11b interacts with 22 different hnRNPs or
445 snRNPs, and about 17% of proteins that constitute the spliceosome (Table S2, S3, Figure S10,
446 S12, suppl data S5).

447 FUS is another crucial factor that is involved in RNA splicing, RNA transport, and regulation
448 of gene expression in response to DNA damage (Chen et al., 2019). FUS can accumulate near
449 the transcription start sites (TSS) and binds to transcription initiation factors to inhibit
450 phosphorylation of the RNA polymerase 2 C-terminal domain (POLR2-CTD) (Chen et al.,
451 2019). We performed co-IP and immunoblot experiments to confirm BCL11b interactions with
452 SMN1 and FUS (Figure 4A, 4B). Taken together, our IP-MS and IP-immunoblot experiments
453 establish that BCL11b interacts with SMN1 and FUS proteins, suggesting that BCL11b may
454 contribute to RNA splicing and/or translation. However, further investigations will be needed
455 to confirm BCL11b functions on RNA processing.

456

457 ***BCL11b interacts with the Drosha and the DGCR8 protein complexes.***

458 FUS is a member of the Drosha and the DGCR8 complexes, which are able to regulate gene
459 expression at transcriptional and post-transcriptional levels (Pong and Gullerova, 2018). On the
460 other hand, Ago2 is the key protein in the RNA interference (RNAi) pathway. Ago2 can bind
461 to Dicer or TRBP proteins, and Drosha or TRBP complexes can regulate mRNA stability or
462 translation (Pong and Gullerova, 2018). Our IP-MS data revealed that BCL11b is associated

463 with 15 of the 20 Drosha complex proteins and 10 of the 11 DGCR8 complex proteins (Figure
464 3D, S10, S12). In addition, 20 proteins of TRBP complex, including Ago2, were observed
465 associated with BCL11b (supplementary dataset S5). To confirm these results, we performed
466 co-IP experiments, which verified the association of BCL11b with FUS and Drosha (Figure
467 1A, 1B and 1C), further suggesting that BCL11b may coordinate with FUS and Drosha to
468 regulate gene expression at transcriptional and post-transcriptional levels, such as mRNA
469 translation and stability.

470

471 ***BCL11b binds to protein from the nonsense-mediated mRNA decay (NMD) pathway.***

472 The NMD pathway degrades RNA transcripts with premature stop codons (Kurosaki et al.,
473 2019). UPF1 is a key regulator of the NMD pathway. Among the significant IP-MS hits, we
474 observed interactions between BCL11b with proteins involved in the NMD pathway. Out of
475 120 proteins identified in this specific GO molecular pathway, 61 were identified in the
476 BCL11b data set (Figure 3B, 3C, S10, S12), among these proteins two core NMD proteins
477 (SMG7 and UPF1), along with one member of the exon junction complex (RBM8A), as well
478 as multiple ribosomal proteins. It is worth to mention that UPF1 can be recruited during mRNA
479 nuclear export, ribosome stalling, or after mRNA translation termination. Co-IP experiments
480 confirmed that BCL11b associates with UPF1, linking BCL11b with proteins that control NMD
481 (Figure 4D).

482

483 **BCL11b interacts with RNA transcripts and proteins from the same gene**

484 We found BCL11b associated with 137 RNA transcripts and proteins encoded by the same gene
485 (Figure 5A, and Suppl. data S6), including EWSR1, FUS, SPFQ, UPF1, tubulins, CHDs and
486 hnRNPs. Interestingly, as we described earlier, BCL11b binds to multiple proteins and ncRNAs
487 constituents of spliceosomes (Table S2, S3, Suppl. data S6). GO analysis revealed that these
488 proteins are involved in RNA processing, NMD, and translation (Figure 5B, Suppl. data S6).

489 To our knowledge, the only known case of a protein that binds to its RNA is a tubulin. For
490 short, tubulin was reported to use a unique and unknown mechanism to auto-regulate its
491 expression level within the cell to ensure a stable production of the protein in cells, and over-
492 expression of tubulin gene does not lead to increase its protein expression level. It is thought
493 that tubulins α and β could polymerize through an unknown mediator, which then binds to the
494 nascent tubulin peptide, ribosomes and/or ribonucleases to terminate the translation (Gasic and
495 Mitchison, 2019). We found that BCL11b interacts with multiple isoforms of tubulin proteins
496 and tubulin RNAs, as well as ribonucleases and ribosomes, i.e. multiple members of the same
497 protein family (Figure 5C). Noteworthy, RNPs and ribosomal proteins are assembled in nucleus
498 before translocalized to cytoplasm, whereas tubulins are required for spindle formation during
499 cell cycle phases. Again, these results support the involvement of BCL11b in the regulation of
500 protein expressions at post-transcriptional levels, particularly the genes involved in RNA
501 splicing and cell cycle.

502

503 **Involvement of BCL11b- associated factors in diseases**

504 Our MS and CLIP-seq results show that BCL11b interacts with proteins that involved in cancer,
505 Ewing Sarcoma, or neurodegenerative diseases (Figure 6, S13-S14 Table S2-S3), such as
506 EWSR1, SMN1, FUS, TDP-43 and TAF15, SRRM2, as well as ncRNAs, such as 7SK,
507 MALAT1, NEAT1. We then searched bibliography and public databases to confirm our finding
508 related to the roles of BCL11b in disease. We found that BCL11b is modulated in lymphoma
509 and Leukemia (Fang et al., 2021; Fu et al., 2017; Kominami, 2012; Montefiori et al., 2021),
510 Ewing Sarcoma (Orth et al., 2020; Wiles et al., 2013), as well as in neuron diseases, such as
511 motor neuron amyotrophic lateral sclerosis (ALS), spinal muscular atrophy (SMA), Parkinson's
512 disease (PD), and Huntington's disease (Ahmed et al., 2015; Doktor et al., 2016; Harb et al.,
513 2022; Lennon et al., 2017, 2016; Maeda et al., 2014; Scherzer et al., 2007; Srakočić et al., 2023),
514 summarized in Figure 6A. However, the role of BCL11b in these diseases was not clear,

515 because of the lack or the difficulties of the functional studies related to protein and RNA
516 interactions of BCL11b.

517

518 **DISCUSSION AND CONCLUSION**

519 This study describes for the first time BCL11b-associated protein and RNA interactome (Figure
520 7). To our knowledge this is the first time to report a protein to bind to RNA transcript and
521 protein encoded from the same gene, such as FUS. We showed that BCL11b binds to RNA and
522 proteins encoded from genes regulating development, cell cycle, transcription, and translation.
523 In addition to its function as a transcription regulator, BCL11b binds to ribonucleic complexes
524 such as ribosomes and RNPs. We confirmed that BCL11b binds to proteins involved in RNA
525 processing, such as RNA splicing and NMD pathway, in addition to the ncRNA that are
526 involved in RNA splicing. First, we demonstrate that BCL11b binds to RNA processing
527 proteins (such as FUS, Drosha, Ago2, UPF1, and RNA splicing factors) suggesting a role in
528 the regulation of gene expression in response to DNA damage. Indeed, BCL11b overexpression
529 was shown to repress DNA damages events (Cherrier et al., 2009). Then, we highlighted the
530 potential role of BCL11b in autoregulation of some specific gene translation such as tubulin.

531 We found that BCL11b preferentially binds to RNA isoforms coding for proteins, in addition
532 to BCL11b is associated with isoforms that do not code for proteins (INCPs), such as NMD and
533 IR. Generally, translation is regulated by selecting the desired mRNA for being translated,
534 whereas, the RNA decay pathways, such as Drosha, UPF1 (NMD pathway), and Ago (RNA
535 interference) are able to decay or silence the undesirable RNA isoforms (Tuck et al., 2020;
536 Wang and Aifantis, 2020). Our results suggest that BCL11b may contribute to the selection of
537 RNA isoforms. Selecting the isoform (so-called isoform switching or selection) is a crucial
538 process through which cells select a specific RNA isoform during the development or immune
539 response. Some isoforms might be needed at specific steps during development, cell cycle or
540 response to stimuli. Therefore, malfunction of RNA splicing or processing has a great impact
541 on cell proliferation and cancer prognosis (Wang and Aifantis, 2020).

542 In conclusion, our analysis of BCL11b-protein and BCL11b-RNA interactions provide the first
543 global overview of BCL11b interactants. The results suggest that BCL11b could bind to RNA
544 processing proteins to regulate translation and/or select the desirable RNA isoform. One can
545 speculate that depletion of BCL11b could lead to isoform switching and translation of wrong
546 isoform of the protein (misfolded or short) but this needs to be formally demonstrated. Although
547 the exact role of BCL11b in diseases deserve to be studied by future research, our results offer
548 potential genome-wide RNA and proteins datasets, which will help future efforts to understand
549 the molecular basis of cancer and neurodegenerative diseases.

550

551 **Conflict of interest**: no conflict of interest is known

552

553 **Human or animal experiments**: no human or animal work have been performed

554

555 **Funding**: This work was supported by the French agency for research on AIDS and viral
556 hepatitis (ANRS); The ANRS RHIVIERA program (supported by ViiV Health care); the
557 European Union's Horizon 2020 research and innovation programme under grant agreement
558 No 691119-EU4HIVCURE-H2020-MSCA-RISE-2015; The Belgian Fund for Scientific
559 Research (FRS-FNRS, Belgium), the "Fondation Roi Baudouin", the Walloon Region (Fonds
560 de Maturation) and the university of Brussels (Action de Recherche Concertée (ARC) grant).
561 The laboratory of CVL is part of the ULB-Cancer Research Centre (U-CRC). AAA was a fellow
562 of the Wallonie-Bruxelles International program and of the Marie Skłodowska Curie COFUND
563 action. HS was a fellow of ANRS. MDR was a fellow of University of Strasbourg. CVL is
564 Directeur de Recherches of the F.R.S-FNRS, Belgium.

565

566

567 **References**

568 Abe H, Kamimura K, Okuda S, Watanabe Y, Inoue J, Aoyagi Y, et al. BCL11B expression in
569 hepatocellular carcinoma relates to chemosensitivity and clinical prognosis. Cancer Medicine
570 2023;12:15650–63. https://doi.org/10.1002/cam4.6167.

571 Ahmed I, Sbodio JI, Harraz MM, Tyagi R, Grima JC, Albacarys LK, et al. Huntington's disease:
572 Neural dysfunction linked to inositol polyphosphate multikinase. Proceedings of the National
573 Academy of Sciences 2015;112:9751–6. https://doi.org/10.1073/pnas.1511810112.

574 Cai S, Kalisky T, Sahoo D, Dalerba P, Feng W, Lin Y, et al. A Quiescent Bcl11b High Stem Cell
575 Population Is Required for Maintenance of the Mammary Gland. Cell Stem Cell 2017;20:247-
576 260.e5. https://doi.org/10.1016/j.stem.2016.11.007.

577 Chen C, Ding X, Akram N, Xue S, Luo S-Z. Fused in Sarcoma: Properties, Self-Assembly and
578 Correlation with Neurodegenerative Diseases. Molecules 2019;24:1622.

579 Cherrier T, Le Douce V, Eilebrecht S, Riclet R, Marban C, Dequiedt F, et al. CTIP2 is a negative
580 regulator of P-TEFb. Proc Natl Acad Sci U S A 2013;110:12655–60.
581 https://doi.org/10.1073/pnas.1220136110.

582 Cherrier T, Suzanne S, Redel L, Calao M, Marban C, Samah B, et al. p21WAF1 gene promoter
583 is epigenetically silenced by CTIP2 and SUV39H1. Oncogene 2009;28:3380–9.
584 https://doi.org/10.1038/onc.2009.193.

585 Chiara VD, Daxinger L, Staal FJT. The Route of Early T Cell Development: Crosstalk between
586 Epigenetic and Transcription Factors. Cells 2021;10:1074.
587 https://doi.org/10.3390/cells10051074.

588 Cismasiu VB, Adamo K, Gecewicz J, Duque J, Lin Q, Avram D. BCL11B functionally associates
589 with the NuRD complex in T lymphocytes to repress targeted promoter. Oncogene
590 2005;24:6753–64. https://doi.org/10.1038/sj.onc.1208904.

591 Cismasiu VB, Paskaleva E, Suman Daya S, Canki M, Duus K, Avram D. BCL11B is a general
592 transcriptional repressor of the HIV-1 long terminal repeat in T lymphocytes through
593 recruitment of the NuRD complex. Virology 2008;380:173–81.
594 https://doi.org/10.1016/j.virol.2008.07.035.

595 Desplats P, Dumaop W, Smith D, Adame A, Everall I, Letendre S, et al. Molecular and pathologic
596 insights from latent HIV-1 infection in the human brain. Neurology 2013;80:1415–23.
597 https://doi.org/10.1212/WNL.0b013e31828c2e9e.

598 Doktor TK, Hua Y, Andersen HS, Brøner S, Liu YH, Wieckowska A, et al. RNA-sequencing of a
599 mouse-model of spinal muscular atrophy reveals tissue-wide changes in splicing of U12-
600 dependent introns. Nucleic Acids Research 2016;45:395–416.
601 https://doi.org/10.1093/nar/gkw731.

602 Drashansky TT, Helm E, Huo Z, Curkovic N, Kumar P, Luo X, et al. Bcl11b prevents fatal
603 autoimmunity by promoting Treg cell program and constraining innate lineages in Treg cells.
604 Sci Adv 2019;5:eaaw0480. https://doi.org/10.1126/sciadv.aaw0480.

605 Drashansky TT, Helm EY, Curkovic N, Cooper J, Cheng P, Chen X, et al. BCL11B is positioned
606 upstream of PLZF and RORgammat to control thymic development of mucosal-associated

607 invariant T cells and MAIT17 program. iScience 2021;24:102307.
608 https://doi.org/10.1016/j.isci.2021.102307.

609 Du H, Wang Z, Guo R, Yang L, Liu G, Zhang Z, et al. Transcription factors Bcl11a and Bcl11b are
610 required for the production and differentiation of cortical projection neurons. Cerebral Cortex
611 2022;32:3611–32. https://doi.org/10.1093/cercor/bhab437.

612 Eilebrecht S, Le Douce V, Riclet R, Targat B, Hallay H, Van Driessche B, et al. HMGA1 recruits
613 CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. Nucleic Acids Res
614 2014;42:4962–71. https://doi.org/10.1093/nar/gku168.

615 Eto K, Machida O, Yanagishita T, Shimojima Yamamoto K, Chiba K, Aihara Y, et al. Novel BCL11B
616 truncation variant in a patient with developmental delay, distinctive features, and early
617 craniosynostosis. Hum Genome Var 2022;9:1–5. https://doi.org/10.1038/s41439-022-00220-
618 x.

619 Fang H, Wang W, El Hussein S, Morita K, Beird HC, Mitra A, et al. B-cell lymphoma/leukaemia
620 11B (BCL11B) expression status helps distinguish early T-cell precursor acute lymphoblastic
621 leukaemia/lymphoma (ETP-ALL/LBL) from other subtypes of T-cell ALL/LBL. British Journal of
622 Haematology 2021;194:1034–8. https://doi.org/10.1111/bjh.17681.

623 Forouzanfar F, Ali S, Wallet C, De Rovere M, Ducloy C, El Mekdad H, et al. HIV-1 Vpr mediates
624 the depletion of the cellular repressor CTIP2 to counteract viral gene silencing. Sci Rep
625 2019;9:13154. https://doi.org/10.1038/s41598-019-48689-x.

626 Fox PM, Tang JLY, Brand AH. The Drosophila homologue of CTIP1 (Bcl11a) and CTIP2 (Bcl11b)
627 regulates neural stem cell temporal patterning. Development 2022;149:dev200677.
628 https://doi.org/10.1242/dev.200677.

629 Fu W, Yi S, Qiu L, Sun J, Tu P, Wang Y. BCL11B-Mediated Epigenetic Repression Is a Crucial
630 Target for Histone Deacetylase Inhibitors in Cutaneous T-Cell Lymphoma. The Journal of
631 Investigative Dermatology 2017;137:1523–32. https://doi.org/10.1016/j.jid.2017.02.980.

632 Gasic I, Mitchison TJ. Autoregulation and repair in microtubule homeostasis. Current Opinion
633 in Cell Biology 2019;56:80–7. https://doi.org/10.1016/j.ceb.2018.10.003.

634 Go R, Takizawa K, Hirose S, Katsuragi Y, Aoyagi Y, Mishima Y, et al. Impairment in
635 differentiation and cell cycle of thymocytes by loss of a Bcl11b tumor suppressor allele that
636 contributes to leukemogenesis. Leukemia Research 2012;36:1035–40.
637 https://doi.org/10.1016/j.leukres.2012.04.028.

638 Grabarczyk P, Winkler P, Delin M, Sappa PK, Bekeschus S, Hildebrandt P, et al. The N-Terminal
639 CCHC Zinc Finger Motif Mediates Homodimerization of Transcription Factor BCL11B.
640 Molecular and Cellular Biology 2018;38:e00368-17. https://doi.org/10.1128/MCB.00368-17.

641 Harb K, Richter M, Neelagandan N, Magrinelli E, Harfoush H, Kuechler K, et al. Pum2 and TDP-
642 43 refine area-specific cytoarchitecture post-mitotically and modulate translation of Sox5,
643 Bcl11b, and Rorb mRNAs in developing mouse neocortex. eLife 2022;11:e55199.
644 https://doi.org/10.7554/eLife.55199.

645 Hasan SN, Sharma A, Ghosh S, Hong SW, Roy-Chowdhuri S, Im SH, et al. Bcl11b prevents
646 catastrophic autoimmunity by controlling multiple aspects of a regulatory T cell gene
647 expression program. Sci Adv 2019;5:eaaw0706. https://doi.org/10.1126/sciadv.aaw0706.

648 Helm EY, Zelenka T, Cismasiu VB, Islam S, Silvane L, Zitti B, et al. Bcl11b sustains multipotency
649 and restricts effector programs of intestinal-resident memory CD8+ T cells. Science
650 Immunology 2023;8:eabn0484. https://doi.org/10.1126/sciimmunol.abn0484.

651 Holmes TD, Pandey RV, Helm EY, Schlums H, Han H, Campbell TM, et al. The transcription
652 factor Bcl11b promotes both canonical and adaptive NK cell differentiation. Science
653 Immunology 2021;6:eabc9801. https://doi.org/10.1126/sciimmunol.abc9801.

654 Hosokawa H, Romero-Wolf M, Yui MA, Ungerback J, Quiloan MLG, Matsumoto M, et al. Bcl11b
655 sets pro-T cell fate by site-specific cofactor recruitment and by repressing Id2 and Zbtb16. Nat
656 Immunol 2018;19:1427–40. https://doi.org/10.1038/s41590-018-0238-4.

657 Kominami R. Role of the transcription factor Bcl11b in development and lymphomagenesis.
658 Proc Jpn Acad Ser B Phys Biol Sci 2012;88:72–87. https://doi.org/10.2183/pjab.88.72.

659 Koumoundourou A, Rannap M, De Bruyckere E, Nestel S, Reissner C, Egorov AV, et al.
660 Regulation of hippocampal mossy fiber-CA3 synapse function by a Bcl11b/C1ql2/Nrxn3(25b+)
661 pathway. eLife 2024;12:RP89854. https://doi.org/10.7554/eLife.89854.

662 Kurosaki T, Popp MW, Maquat LE. Quality and quantity control of gene expression by
663 nonsense-mediated mRNA decay. Nature Reviews Molecular Cell Biology 2019;20:406–20.
664 https://doi.org/10.1038/s41580-019-0126-2.

665 Lennon MJ, Jones SP, Lovelace MD, Guillemin GJ, Brew BJ. Bcl11b—A Critical
666 Neurodevelopmental Transcription Factor—Roles in Health and Disease. Frontiers in Cellular
667 Neuroscience 2017;11. https://doi.org/10.3389/fncel.2017.00089.

668 Lennon MJ, Jones SP, Lovelace MD, Guillemin GJ, Brew BJ. Bcl11b: A New Piece to the Complex
669 Puzzle of Amyotrophic Lateral Sclerosis Neuropathogenesis? Neurotoxicity Research
670 2016;29:201–7. https://doi.org/10.1007/s12640-015-9573-5.

671 Liao C-K, Fang K-M, Chai K, Wu C-H, Ho C-H, Yang C-S, et al. Depletion of B cell CLL/Lymphoma
672 11B Gene Expression Represses Glioma Cell Growth. Molecular Neurobiology 2016;53:3528–
673 39. https://doi.org/10.1007/s12035-015-9231-1.

674 Liau ES, Jin S, Chen Y-C, Liu W-S, Calon M, Nedelec S, et al. Single-cell transcriptomic analysis
675 reveals diversity within mammalian spinal motor neurons. Nat Commun 2023;14:46.
676 https://doi.org/10.1038/s41467-022-35574-x.

677 Longabaugh WJR, Zeng W, Zhang JA, Hosokawa H, Jansen CS, Li L, et al. Bcl11b and
678 combinatorial resolution of cell fate in the T-cell gene regulatory network. Proceedings of the
679 National Academy of Sciences 2017;114:5800–7. https://doi.org/10.1073/pnas.1610617114.

680 Maeda M, Harris AW, Kingham BF, Lumpkin CJ, Opdenaker LM, McCahan SM, et al.
681 Transcriptome profiling of spinal muscular atrophy motor neurons derived from mouse
682 embryonic stem cells. PLOS ONE 2014;9:e106818.
683 https://doi.org/10.1371/journal.pone.0106818.

684 Marban C, Suzanne S, Dequiedt F, de Walque S, Redel L, Van Lint C, et al. Recruitment of
685 chromatin-modifying enzymes by CTIP2 promotes HIV-1 transcriptional silencing. EMBO J
686 2007;26:412–23. https://doi.org/10.1038/sj.emboj.7601516.

687 Maurin T, Melko M, Abekhoukh S, Khalfallah O, Davidovic L, Jarjat M, et al. The FMRP/GRK4
688 mRNA interaction uncovers a new mode of binding of the Fragile X mental retardation protein
689 in cerebellum. Nucleic Acids Res 2015;43:8540–50. https://doi.org/10.1093/nar/gkv801.

690 Miller DH, Jin DX, Sokol ES, Cabrera JR, Superville DA, Gorelov RA, et al. BCL11B Drives Human
691 Mammary Stem Cell Self-Renewal In Vitro by Inhibiting Basal Differentiation. Stem Cell
692 Reports 2018;10:1131–45. https://doi.org/10.1016/j.stemcr.2018.01.036.

693 Montefiori LE, Bendig S, Gu Z, Chen X, Pölönen P, Ma X, et al. Enhancer Hijacking Drives
694 Oncogenic BCL11B Expression in Lineage-Ambiguous Stem Cell Leukemia. Cancer Discovery
695 2021;11:2846–67. https://doi.org/10.1158/2159-8290.CD-21-0145.

696 Orth MF, Hölting TLB, Dallmayer M, Wehweck FS, Paul T, Musa J, et al. High Specificity of
697 BCL11B and GLG1 for EWSR1-FLI1 and EWSR1-ERG Positive Ewing Sarcoma. Cancers
698 2020;12:644.

699 Parsa S, Soheili Z-S, Alizadeh Otaghvar H, Behrangi E, Mansouri P, Lubberts E, et al. BCL11B Is
700 Involved in Stress-Induced Differentiation of Keratinocytes and Has A Potential Role in
701 Psoriasis Pathogenesis. Cell Journal (Yakhteh) 2023;25:300–6.
702 https://doi.org/10.22074/cellj.2023.558003.1090.

703 Pong SK, Gullerova M. Noncanonical functions of microRNA pathway enzymes – Drosha,
704 DGCR8, Dicer and Ago proteins. FEBS Letters 2018;592:2973–86.
705 https://doi.org/10.1002/1873-3468.13196.

706 Przybylski GK, Korsak D, Iżykowska K, Nowicka K, Zalewski T, Tubacka M, et al. Generation of
707 Inducible BCL11B Knockout in TAL1/LMO1 Transgenic Mouse T Cell Leukemia/Lymphoma
708 Model. International Journal of Molecular Sciences 2022;23:4932.
709 https://doi.org/10.3390/ijms23094932.

710 Rohr O, Lecestre D, Chasserot-Golaz S, Marban C, Avram D, Aunis D, et al. Recruitment of Tat
711 to heterochromatin protein HP1 via interaction with CTIP2 inhibits human immunodeficiency
712 virus type 1 replication in microglial cells. J Virol 2003;77:5415–27.
713 https://doi.org/10.1128/jvi.77.9.5415-5427.2003.

714 Scherzer CR, Eklund AC, Morse LJ, Liao Z, Locascio JJ, Fefer D, et al. Molecular markers of early
715 Parkinson's disease based on gene expression in blood. Proceedings of the National Academy
716 of Sciences 2007;104:955–60. https://doi.org/10.1073/pnas.0610204104.

717 Shadrina O, Garanina I, Korolev S, Zatsepin T, Van Assche J, Daouad F, et al. Analysis of RNA
718 binding properties of human Ku protein reveals its interactions with 7SK snRNA and protein
719 components of 7SK snRNP complex. Biochimie 2020;171–172:110–23.
720 https://doi.org/10.1016/j.biochi.2020.02.016.

721 Sidwell T, Rothenberg EV. Epigenetic Dynamics in the Function of T-Lineage Regulatory Factor
722 Bcl11b. Frontiers in Immunology 2021;12.

723 Simon R, Wiegreffe C, Britsch S. Bcl11 Transcription Factors Regulate Cortical Development
724 and Function. Frontiers in Molecular Neuroscience 2020;13.

725 Srakočić S, Gorup D, Kutlić D, Petrović A, Tarabykin V, Gajović S. Reactivation of corticogenesis-
726 related transcriptional factors BCL11B and SATB2 after ischemic lesion of the adult mouse
727 brain. Sci Rep 2023;13:8539. https://doi.org/10.1038/s41598-023-35515-8.

728 Tabet R, Moutin E, Becker JA, Heintz D, Fouillen L, Flatter E, et al. Fragile X Mental Retardation
729 Protein (FMRP) controls diacylglycerol kinase activity in neurons. Proc Natl Acad Sci U S A
730 2016;113:E3619-28. https://doi.org/10.1073/pnas.1522631113.

731 Tang B, Di Lena P, Schaffer L, Head SR, Baldi P, Thomas EA. Genome-wide identification of
732 Bcl11b gene targets reveals role in brain-derived neurotrophic factor signaling. PLOS ONE
733 2011;6:e23691. https://doi.org/10.1371/journal.pone.0023691.

734 Tuck AC, Rankova A, Arpat AB, Liechti LA, Hess D, Iesmantavicius V, et al. Mammalian RNA
735 Decay Pathways Are Highly Specialized and Widely Linked to Translation. Molecular Cell
736 2020;77:1222-1236.e13. https://doi.org/10.1016/j.molcel.2020.01.007.

737 Turriziani B, Garcia-Munoz A, Pilkington R, Raso C, Kolch W, von Kriegsheim A. On-beads
738 digestion in conjunction with data-dependent mass spectrometry: a shortcut to quantitative
739 and dynamic interaction proteomics. Biology (Basel) 2014;3:320–32.
740 https://doi.org/10.3390/biology3020320.

741 Vickridge E, Faraco CCF, Lo F, Rahimian H, Liu ZY, Tehrani PS, et al. The function of BCL11B in
742 base excision repair contributes to its dual role as an oncogene and a haplo-insufficient tumor
743 suppressor gene. Nucleic Acids Research 2024;52:223–42.
744 https://doi.org/10.1093/nar/gkad1037.

745 Wang E, Aifantis I. RNA Splicing and Cancer. Trends in Cancer 2020.
746 https://doi.org/10.1016/j.trecan.2020.04.011.

747 Wiles ET, Lui-Sargent B, Bell R, Lessnick SL. BCL11B Is Up-Regulated by EWS/FLI and
748 Contributes to the Transformed Phenotype in Ewing Sarcoma. PLOS ONE 2013;8:e59369.
749 https://doi.org/10.1371/journal.pone.0059369.

750 Yang W-J, Sun Y-F, Jin A-L, Lv L-H, Zhu J, Wang B-L, et al. BCL11B suppresses tumor progression
751 and stem cell traits in hepatocellular carcinoma by restoring p53 signaling activity. Cell Death
752 & Disease 2020;11:895. https://doi.org/10.1038/s41419-020-03115-3.

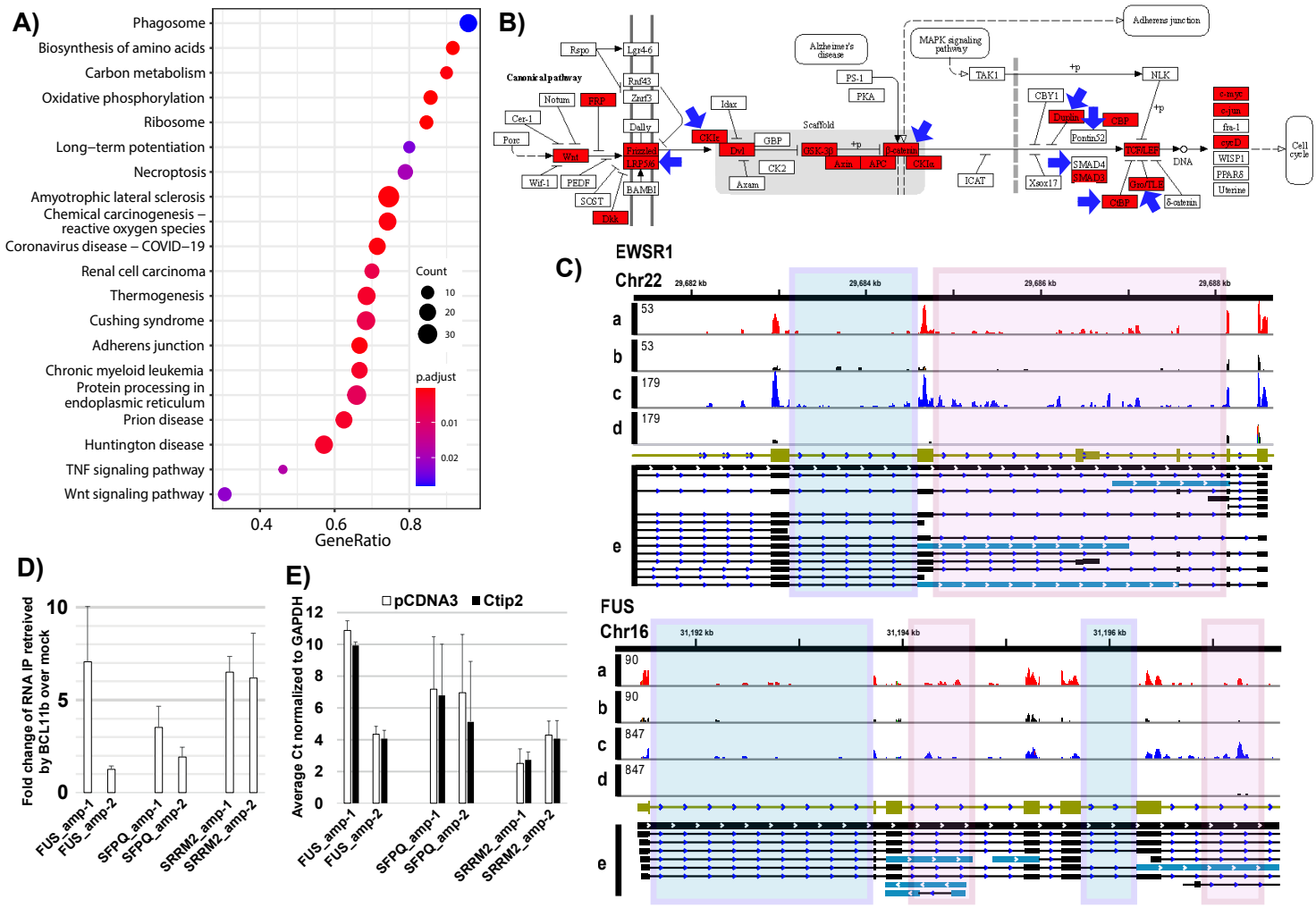
753 Zhang L, Yu L, Liu Y, Wang S, Hou Z, Zhou J. miR-21-5p promotes cell proliferation by targeting
754 BCL11B in Thp-1 cells. Oncol Lett 2021;21:119. https://doi.org/10.3892/ol.2020.12380.

755 Zhao C, Li X, Han B, You Z, Qu L, Liu C, et al. Gga-miR-219b targeting BCL11B suppresses
756 proliferation, migration and invasion of Marek's disease tumor cell MSB1. Scientific Reports
757 2017;7:4247. https://doi.org/10.1038/s41598-017-04434-w.

758 Zhou W, Gao F, Romero-Wolf M, Jo S, Rothenberg EV. Single-cell deletion analyses show
759 control of pro–T cell developmental speed and pathways by Tcf7, Spi1, Gata3, Bcl11a, Erg, and
760 Bcl11b. Science Immunology 2022;7:eabm1920.
761 https://doi.org/10.1126/sciimmunol.abm1920.

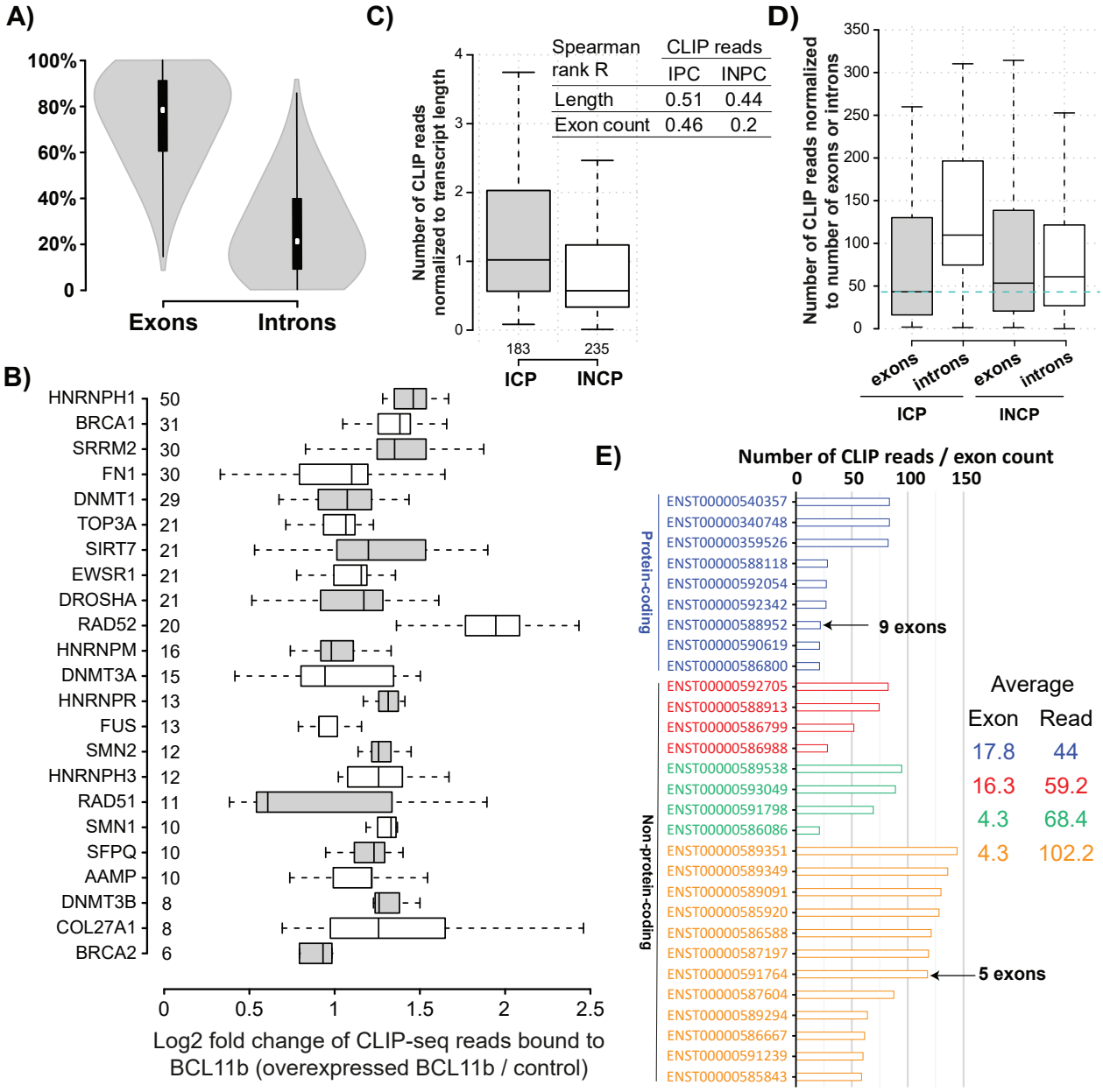
762

FIGURE 1



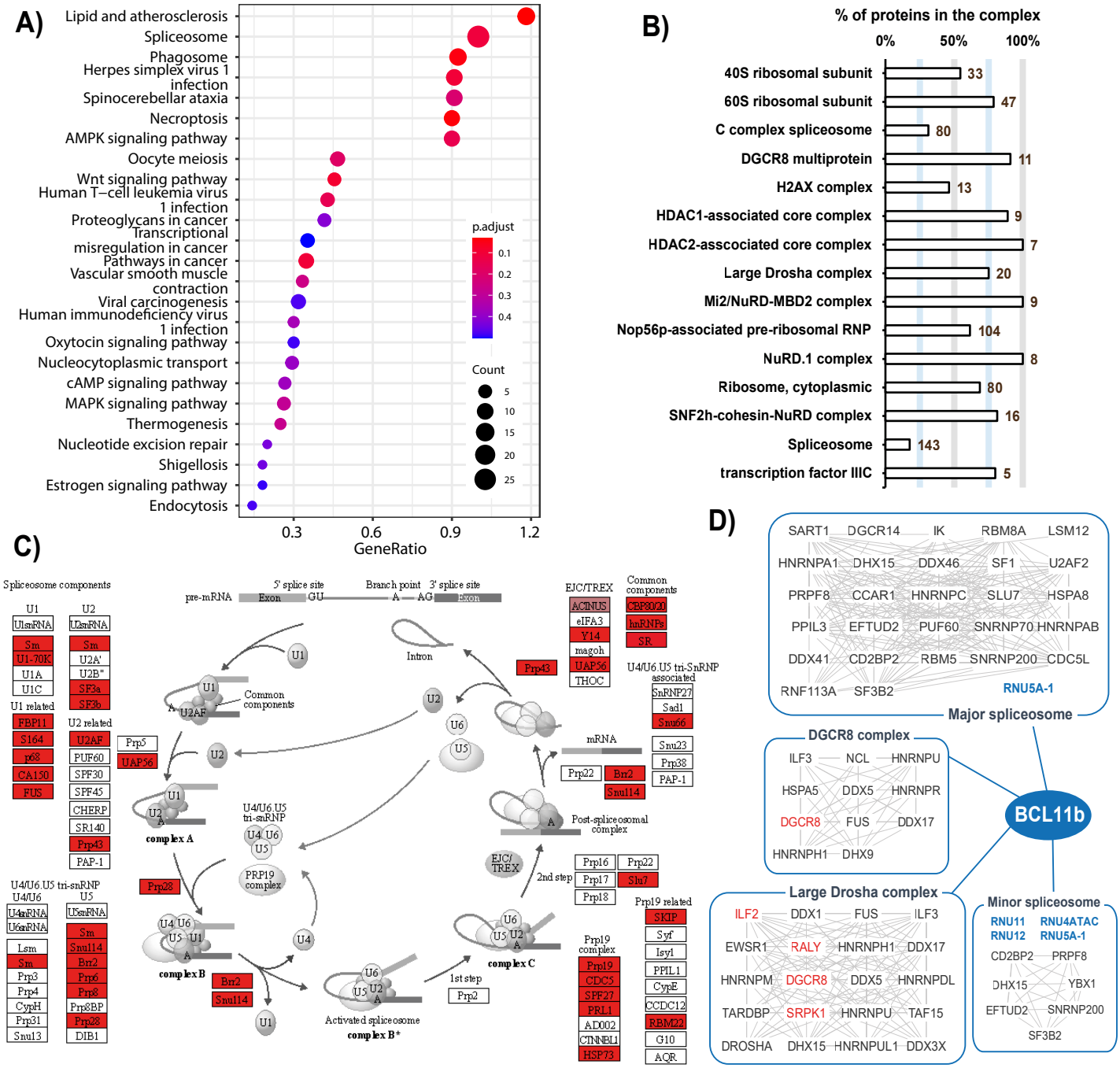
BCL11b-RNA interactions as deciphered by CLIP-seq. A) KEGG pathways of genes encode RNA bound to BCL11b, Figure S1. B) Wnt signaling is among pathways that are regulated by BCL11b-RNA interactions, the red box means the CLIP reads increased by at least 4 log2 fold change compared to control in microglia, blue arrow means CLIP reads increased at least 2 log2 fold after over-expression in HEK cells, for full pathway Figure S2. The GSEA and KEGG pathway were constructed using "ClusterProfiler, EnrichGO, Enrichplot, and Pathview" R packages. C) Representative CLIP-seq results of three protein-coding genes (EWSR1 and FUS), visualized using Integrative Genomics Viewer (IGV) tool. The olive-green bars refer to the NCBI RefSeq genes. X-axis represents the coordinates of the gene and the chromosome number. The Y-axis represents the number of CLIP reads in each condition. a) Over-expressed BCL11b in HEK cells, b) endogenous BCL11b in microglial cells, whereas, c and d are their control experiments, respectively. e) The black bars represent the GENCODE transcripts of the stated genes, the sky-blue transcripts refer to the protein non-coding transcripts that harbor significant CLIP reads and their coordinates correspond to intron of RefSeq gene, these intronic regions are highlighted with light rose colored boxes. The light blue horizontal boxes refer to introns without significant BCL11b CLIP reads. For additional figures and CLIP-seq reads alignments, see Figure S1. D) Results of RNA-IP-qPCR (RIP-PCR) of three genes, the fold change of RNA bound to endogenous BCL11b to the input control, and E) gene expression of the amplicons used in D using overexpressed BCL11b plasmid and pCDNA3 vector. For PCR primers and raw Ct values, Table S4-S5, and locations of the amplicons and gel amplification, Figure S4, S5.

FIGURE 2



BCL11b-bound RNA at transcript (isoform) level. A) Violin shape represents the number of CLIP reads mapped to exons and introns of the significant genes. B) BCL11b binds differentially to the selected 23 genes. Number next to the names are number of transcripts. C) Numbers of CLIP reads normalized to number of exons for isoforms coding for proteins (for short, ICPs) or isoforms do not code for proteins (for short, INCPs). The table shows the Spearman's rank correlation coefficient (R) between numbers of CLIP reads in case of ICPs or INCPs from one side and length or number of exons from the other side. D) Number of CLIP reads mapped to exons and those mapped to introns in case of ICPs or INCPs. E) Number of CLIP reads mapped to exons of DNMT1 gene are counted and normalized to number of exons. The table shows the average number of CLIP reads and number of exons within protein-coding transcripts and non-protein-coding transcripts. The blue colors refer to ICPs, whereas red, green and orange colors refer to nonsense mediated decay, processed transcript, and retained intron transcripts, respectively. Figures 3B-3E were constructed from Suppl. data S4.

FIGURE 3

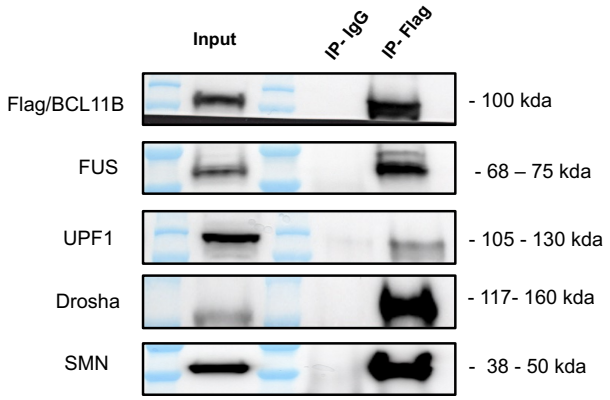


BCL11b-interacting proteins (for short BIPs) as deciphered by IP-MS. A) KEGG pathways of BCL11b-interaction proteins, additional details on GO terms and GSEA analysis can be found in Figure S9, SI data S5. B) The complexes that constitute of BIPs, X-axis show percent of BIPs to the total number of proteins in the complex; the total number of proteins in the complex is written in dark blue outside the bars. Complexes were downloaded from CORUM database, SI data S5. C) KEGG spliceosome pathway (KEGG ID: hsa03040), the red boxes represent BIPs, for additional KEGG pathways, Figure S10-S11, SI data S5. D) Four representative complexes with the name of proteins. The BIPs are found in rectangles, BCL11b-interacting ncRNAs are in blue, whereas the non-BIPs are in red colors. The arrows refer to protein-protein interactions. The interactions network were downloaded from STRING database and constructed using Cytoscape software. For additional complexes and interactions, Table, S2-S3, SI data S5.

813 **FIGURE 4**

814

815



816

817

818

819

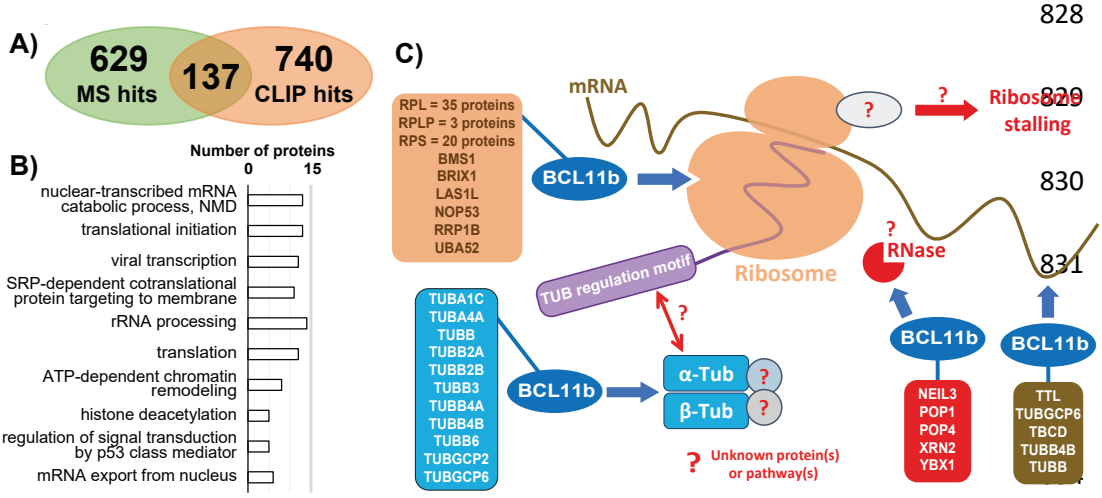
820

821

822 BCL11B interacts with FUS, Drosha, SMN1 and UPF1. Coimmunoprecipitations targeting Flag-
823 BCL11B were performed, and the indicated partners were detected by Western blot with the
824 corresponding antibodies.

825

FIGURE 5



BCL11b binds to RNA and proteins encoded from the same genes. A) Number of BCL11b-interacting proteins, from IP-MS experiment, that passed our cut-off (so-called MS hits); number of genes, which have CLIP reads passed our cut-off (so-called CLIP hits); 137 genes can bind to BCL11b in form of RNA or in form of protein. B) GO terms of 137 shared genes (BCL11b binds to RNA and protein product). C) Schematic diagram of the possible tubulin autoregulation mechanism. BCL11b binds to multiple tubulin RNA, in brown box, in addition BCL11b binds to multiple ribosomes (oranges box), tubulins (blue box) and ribonucleases (red box).

847

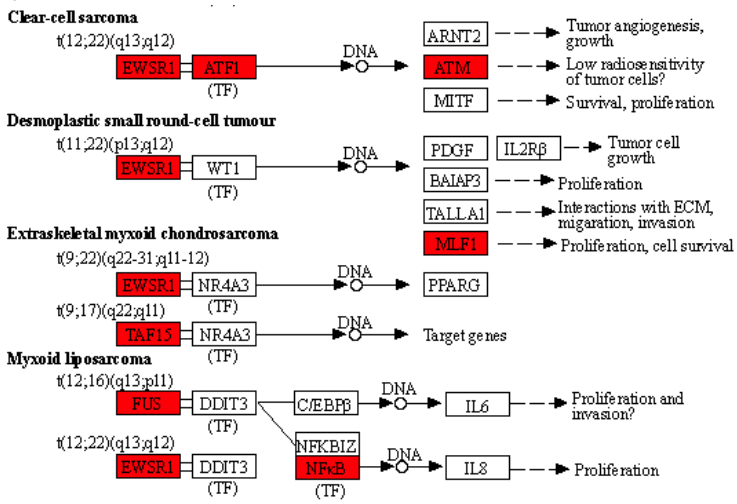
848 **FIGURE 6**

849

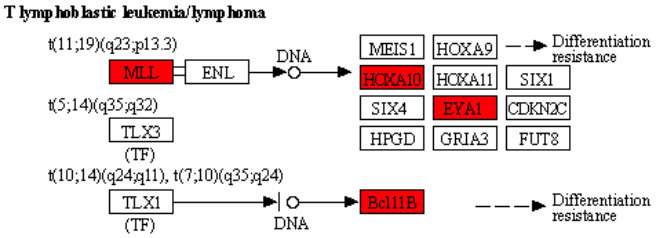
A)

	BCL11b interacts with	PD	SMA	ALS	Ewing's sarcoma
BCL11b		Down-regulated	Down-regulated		Up-regulated
SRRM2	RNA	Up-regulated	-	-	-
SMN1	Protein	-	Down-regulated	-	-
NEAT1	RNA	-	Up-regulated	-	-
EWSR1	RNA and protein	-	-	Contribute in	Contribute in
FUS	RNA and protein	-	-	Mutated	-
RNPs	RNA, protein & ncRNA	-	Defects	Defects	-
TDP43	RNA and protein	-	-	Mutated	-
TAF15	RNA and protein	-	-	Contribute in	-
VCP	RNA	-	-	Contribute in	-

B)



C)



850

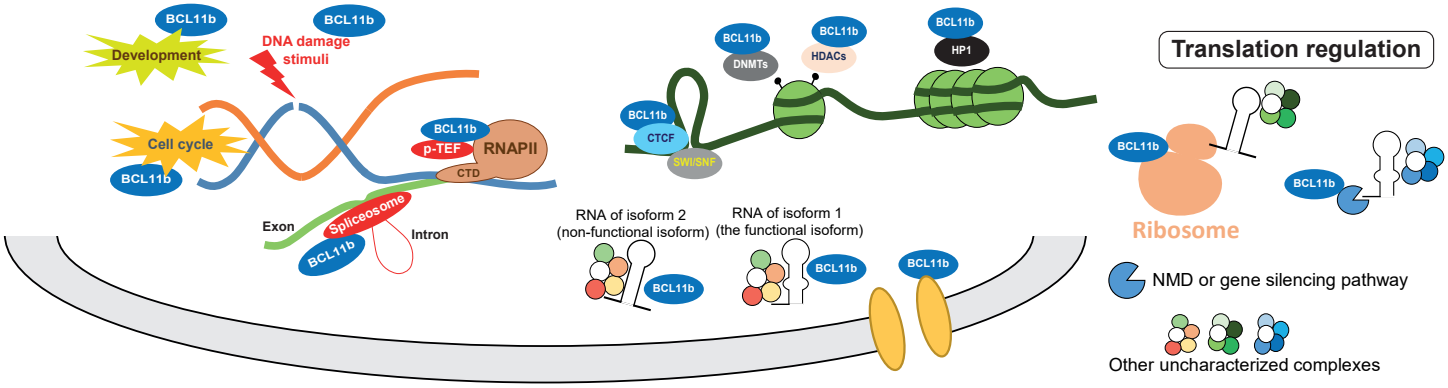
851

852 BCL11b is involved in multiple diseases. A) The table shows the contribution of BCL11b and
853 other genes in the Parkinson's disease (PD), SMA, ALS and Ewing sarcoma diseases. B and C)
854 are examples of BCL11b-interactions proteins contributing to transcriptional mis-regulation in
855 cancer, sarcoma, and lymphoma pathways (KEGG ID: hsa05202), for additional KEGG disease
856 pathways, Figures S12-S13.

857

858 **Figure 7**.

859



860

861 The graphical abstract shows the major proteins pathways, complexes and based on GO terms obtained
862 after IP-MS and CLIP-seq experiments.

863

864

865

866