

1

2

3 **Exon junction complex-associated multi-adapter RNPS1**
4 **nucleates splicing regulatory complexes to maintain**
5 **transcriptome surveillance**

6 Lena P. Schlautmann^{1,2}, Volker Boehm^{1,2}, Jan-Wilm Lackmann³, Janine Altmüller^{4,5}, Christoph

7 Dieterich^{6,7}, Niels H. Gehring^{1,2,*}

8

9 ¹ Institute for Genetics, University of Cologne, 50674 Cologne, Germany

10 ² Center for Molecular Medicine Cologne (CMMC), University of Cologne, 50937 Cologne,
11 Germany

12 ³ CECAD Research Center, University of Cologne, Joseph-Stelzmann-Str. 26, 50931
13 Cologne, Germany

14 ⁴ Cologne Center for Genomics (CCG), University of Cologne, 50931 Cologne, Germany

15 ⁵ Present address: Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Core
16 Facility Genomics, Charitéplatz 1, 10117 Berlin, Germany and Max Delbrück Center for
17 Molecular Medicine in the Helmholtz Association (MDC), Berlin, Germany

18 ⁶ Section of Bioinformatics and Systems Cardiology, Department of Internal Medicine III and
19 Klaus Tschira Institute for Integrative Computational Cardiology, Heidelberg University
20 Hospital, 69120 Heidelberg, Germany

21 ⁷ DZHK (German Centre for Cardiovascular Research), Partner site Heidelberg/Mannheim,
22 69120 Heidelberg, Germany

23

24 *Contact: Niels H. Gehring, ngehring@uni-koeln.de

25

26

27

28

29

30

31

32 **Abstract**

33

34 The exon junction complex (EJC) is an RNA-binding multi-protein complex with critical
35 functions in post-transcriptional gene regulation. It is deposited on the mRNA during splicing
36 and regulates diverse processes including pre-mRNA splicing, mRNA export, mRNA
37 translation, and nonsense-mediated mRNA decay (NMD) via various interacting peripheral
38 proteins. The EJC-binding protein RNPS1 might serve two functions: it suppresses mis-
39 splicing of cryptic splice sites and activates NMD in the cytoplasm. When analyzing the
40 transcriptome-wide effects of EJC and RNPS1 knockdowns in different human cell lines, we
41 find no evidence for RNPS1 being a globally essential NMD factor. However, various aberrant
42 splicing events strongly suggest that the main function of RNPS1 is splicing regulation. Rescue
43 analyses revealed that about half of these RNPS1-dependent splicing events was fully or
44 partially rescued by the expression of the isolated RRM domain of RNPS1, whereas other
45 splicing events are regulated by its C-terminal domain. We identified many splicing-regulatory
46 factors, including SR proteins and U1 snRNP components, that specifically interact with the C-
47 terminus or with the RRM of RNPS1. Thus, RNPS1 emerges as a multifunctional splicing
48 regulator that promotes correct and efficient splicing of different vulnerable splicing events via
49 the formation of diverse splicing-promoting complexes.

50

51

52

53 **Introduction**

54 The majority of human genes contain introns and their transcribed pre-mRNAs are subject to
55 (alternative) splicing (1). During splicing, intronic sequences are excised and exons are ligated
56 by the spliceosome, resulting in a mature mRNA that is subsequently exported to the
57 cytoplasm (2). The spliceosome has the critical as well as delicate task to identify the correct
58 splice sites, because frequently there is more than one possible splice site. These additional
59 sites can be designated alternative splice sites, but also so-called cryptic splice sites (3). While
60 the usage of the former can be employed to generate different transcript isoforms, the
61 erroneous utilization of cryptic splice sites leads to mis-splicing and the production of defective
62 transcripts (4). Both processes use similar mechanisms, but with opposing results. Alternative
63 splicing (AS) increases the number of protein isoforms produced from a single gene by the
64 varying usage of 5' and 3' splice sites, skipping of exons and inclusion of introns (5). In addition,
65 it is also an important step in gene expression regulation. In contrast, the use of cryptic splice
66 sites is often associated with the production of non-functional transcripts and the occurrence
67 of disease (4). Therefore, (alternative) splicing requires tight regulation and accurate operation
68 of the spliceosome to ensure the production of the correct mature mRNAs.

69 The recognition of the proper splice sites by the spliceosome is assisted by auxiliary proteins,
70 which bind to the pre-mRNA and guide the spliceosome to the correct positions (6). The group
71 of splicing regulatory proteins is quite diverse and includes several RNA-binding proteins that
72 interact with specific sequence motifs, such as the SR proteins (7). The class of SR proteins
73 is characterized by one or two N-terminal RNA binding domains (e.g. RRM), and a C-terminal
74 domain enriched in arginine and serine dipeptides (RS-domain). SR proteins bind to exonic
75 splicing enhancers (ESEs) and thereby define the exons to be maintained in the mature mRNA
76 (8). In contrast, hnRNPs bind mainly to intronic sequences (intronic splicing enhancers, ISEs)
77 and support their recognition and removal by the spliceosome (9,10). The loss or exchange of
78 ESE and ISE sequences can have dramatic effects on the splicing pattern and might even
79 result in the inactivation of genes (11). However, the splicing process is not only regulated by

80 SR proteins and hnRNPs, but also by many other RNA-binding proteins (12). Among these is
81 also the exon junction complex (EJC), an RNA-binding protein complex, which binds 24 nt
82 upstream of an emerging exon-exon junction, independent of the RNA sequence (13). The
83 EJC core consists of the three proteins EIF4A3, RBM8A and MAGOH that are deposited onto
84 the mRNA during splicing by interactions of the EJC proteins with spliceosome components
85 (14,15).

86 The EJC carries out diverse functions during post-transcriptional gene regulation and besides
87 regulating pre-mRNA splicing, it also facilitates the transport and translation of spliced mRNAs
88 (13). Furthermore, the EJC is critical for nonsense-mediated mRNA decay (NMD) and EJC
89 proteins were initially described to enable the detection of NMD substrates, particularly mRNAs
90 containing premature translation termination codons (PTCs) (16,17). NMD not only serves as
91 a quality control mechanism by ensuring the degradation of incorrect mRNAs, but is also
92 important for the regulation of gene expression (18). NMD substrates can be produced in
93 different ways: NMD-activating termination codons may result from AS or genomic mutations,
94 in other cases NMD is triggered by a long 3' UTR. Efficient NMD takes place when a ribosome
95 terminates at a PTC. EJCs bound to the mRNA downstream of that PTC will serve as a signal
96 for the NMD machinery to initiate the degradation of that mRNA. Therefore, the EJC is
97 essential for NMD to correctly identify transcripts that need to be degraded.

98 In order to carry out all its different tasks, the EJC functions as a binding platform for auxiliary
99 factors which itself have varying regulatory potentials. Two of the EJC-associated complexes,
100 the apoptosis and splicing associated protein complex (ASAP) and the PSAP complex are
101 known regulators of splicing (19). They share two of their components, RNA binding protein
102 with serine rich domain 1 (RNPS1) and Sin3A associated protein 18 (SAP18), but vary in the
103 third component, which is either Acinus (ACIN1) or Pinin (PNN), respectively (20,21). A
104 previous study has shown that knockdown (KD) of PNN and ACIN1 affects different splicing
105 events, suggesting that ASAP and PSAP complexes have non-overlapping functions in
106 splicing regulation (22). In *D. melanogaster*, retention of PIWI intron 4 relies on the EJC core,

107 ACIN1 and RNPS1 (23,24). Recent studies furthermore demonstrate the ability of the EJC
108 core and the PSAP to suppress the usage of cryptic 5' and 3' splice sites (25-28). While cryptic
109 3' splice sites are suppressed by direct masking by the EJC core, the suppression of cryptic 5'
110 splice sites involves an unknown mechanism requiring RNPS1 recruitment via the EJC core
111 and the PSAP complex (25,26). It is likely that RNPS1 represents the central functional
112 component in all these processes, whereas ACIN1 and probably PNN as well play a role in
113 RNPS1 recruitment. ACIN1, for example, directly binds to the EJC core (22), which would
114 explain how the interaction of RNPS1 and the EJC is established.

115 In addition to its function in splicing, RNPS1 also has the ability to activate NMD when tethered
116 to a reporter mRNA downstream of the termination codon (17,29). It has also been reported
117 that the presence of RNPS1 on NMD-targeted mRNAs leads to more pronounced degradation
118 (30). However, there are controversial results as to whether RNPS1 has an essential role in
119 NMD or not (31,32). Although the exact function of RNPS1 during NMD remains to be
120 determined, it is clear that RNPS1 interacts with the EJC and possibly also with components
121 of the NMD machinery, potentially forming a bridge between these two macromolecular
122 assemblies.

123 Although previous work had examined individual aspects of RNPS1, its function in the context
124 of the EJC is still not fully understood and therefore demands a more comprehensive
125 characterization of RNPS1. In this study, we uncover that RNPS1 only mildly affects a small
126 subset of NMD targets and its main function is the regulation of AS. To that end, the RNPS1
127 RRM, which is known to be required for ASAP/PSAP assembly, regulates splicing by binding
128 other splicing factors, including SR proteins and spliceosomal components (20). We identified
129 many components of the U1 snRNP that interact with the C-terminus and thus conclude that
130 RNPS1 is a part of and bridges different splicing competent complexes to the EJC to regulate
131 splicing of surrounding/adjacent introns. In our model RNPS1 acts as a multi-functional adapter
132 that recruits splicing factors independently of the mRNA sequence to the EJC binding site.

133

134

135 **Material and Methods**

136 **Cell Culture**

137 Flp-In-T-REx-293 (HEK 293; human, female, embryonic kidney, epithelial; Thermo Fisher
138 Scientific, RRID:CVCL_U427), HeLa Flp-In-T-REx (HeLa FT; human, female, cervix; Elena
139 Dobrikova and Matthias Gromeier, Duke University Medical Center) and HeLa Tet-Off (HTO;
140 human, female, cervix; Clontech, RRID: CVCL_V352) cells were cultured in high-glucose,
141 GlutaMAX DMEM (Gibco) supplemented with 9% fetal bovine serum (Gibco) and 1x Penicillin
142 Streptomycin (Gibco). The cells were cultivated at 37°C and 5% CO₂ in a humidified incubator.
143 The generation of stable cell lines is described below and all cell lines are summarized in
144 [Supplementary Table 1](#).

145

146 **Stable cell lines and plasmids**

147 RNPS1 point and deletion mutants were PCR amplified using Q5 polymerase (New England
148 Biolabs) and inserted into PB-CuO-MCS-BGH-EF1-CymR-Puro (modified from System
149 Biosciences), together with an N-terminal FLAG-emGFP-tag via NheI and NotI (both New
150 England Biolabs) restriction sites. As a control, FLAG-emGFP was equally cloned into the PB-
151 CuO-MCS-BGH-EF1-CymR-Puro vector.

152 HEK 293 and HTO cells were stably transfected using the PiggyBac Transposon system. 2.5-
153 3x10⁵ cells were seeded 24 h before transfection in 6-wells. 1 µg of PiggyBac construct was
154 transfected together with 0.8 µg of the Super PiggyBac Transposase expressing vector using
155 the calcium phosphate method. 48 h after transfection, the cells were transferred into 10 cm
156 dishes and selected with 2 µg ml⁻¹ puromycin (InvivoGen). After 7-10 days, the colonies were
157 pooled. Expression of the PiggyBac constructs was induced with 30 µg ml⁻¹ cumate.

158 RFX5 reporters were PCR amplified as described above and cloned into
159 pcDNA5/FRT/TO/FLAG. HeLa FT cells were stably transfected with the reporters using the
160 Flp-In-T-REx system. Transfection and selection was performed like for PiggyBac transfected

161 cells, with the following differences: 1.5 µg of pcDNA5 construct were co-transfected with
162 1.5 µg Flippase expression vector (pOG44) and cells were selected with 100 µg ml⁻¹
163 Hygromycin (InvivoGen). All cell lines generated and plasmids used in this study are listed in
164 [Supplementary Table 1](#).

165

166 **Co-Immunoprecipitation**

167 Expression of FLAG-emGFP tagged RNPS1 mutants and FLAG-emGFP control was induced
168 in stable cell lines (1.5 x 10⁶ cells per 10 cm dish) using cumate (as described above) 72 h
169 before cell lysis. The samples were lysed in 600 µl buffer E (20 mM HEPES-KOH (pH 7.9),
170 100 mM KCl, 10% glycerol, 1 mM DTT, Protease Inhibitor) in the presence of 1 µg ml⁻¹
171 RNase A and sonicated using the Bandelin Sonopuls mini20 with 10 pulses (2.5 mm tip, 1s
172 pulse, 50% amplitude). For immunoprecipitation, the protein concentration of the lysates was
173 measured using Pierce Detergent Compatible Bradford Assay Reagent (Thermo Fisher
174 Scientific) and adjusted in buffer E. Then, the lysates were loaded onto Anti-FLAG M2
175 Magnetic Beads (Sigma-Aldrich) and incubated for 2 h at 4°C with overhead shaking. After
176 that, the beads were washed four times for 3 min with mild wash buffer (20 mM HEPES-KOH
177 (pH 7.9), 137 mM NaCl, 2 mM MgCl₂, 0.2% Triton X-100, 0.1% NP-40). For elution, 2x 21.5 µl
178 (42.5 µl total) of a 200 mg ml⁻¹ dilution of FLAG peptides (Sigma) in 1x TBS was used.

179

180 **Label-free Mass Spec and computational analysis**

181 For Label-free Mass spec, samples were immunoprecipitated as described above and after
182 addition of 1 volume of 5% SDS in PBS reduced with DTT and alkylated with CAA (final
183 concentrations 5 mM and 55 mM, respectively). For tryptic protein digestion, a modified version
184 of the single pot solid phase-enhanced sample preparation (SP3) protocol was used as
185 described below (33). Samples were reduced with 5 mM Dithiothreitol followed by alkylation
186 using 40 mM Chloroacetamide. Afterwards, proteins were supplemented with paramagnetic

187 Sera-Mag speed beads (Cytiva) and mixed in a 1:1-ratio with 100% acetonitrile (ACN). After 8
188 min incubation, protein-beads-complexes were captured using an in-house build magnetic
189 rack, washed twice with 70% EtOH, and washed once with 100% ACN. After airdrying and
190 reconstitution in 5 μ l 50 mM triethylammonium bicarbonate, samples were supplemented with
191 0.5 μ g trypsin and 0.5 μ g LysC and incubated overnight at 37°C. The beads were resuspended
192 on the next day and mixed with 200 μ l ACN, followed by 8 min incubation. Subsequently, the
193 samples were placed on the magnetic rack to wash the tryptic peptides once with 100% ACN.
194 Samples were airdried, dissolved in 4% DMSO, transferred into new PCR tubes, and acidified
195 with 1 μ l of 10% formic acid. Proteomics analysis was performed by the proteomics core facility
196 at CECAD via data-dependent acquisition using an Easy nLC1200 ultra high-performance
197 liquid chromatography (UHPLC) system connected via nano electrospray ionization to a Q
198 Exactive Plus instrument (all Thermo Scientific) running in DDA Top10 mode. Based on their
199 hydrophobicity the tryptic peptides were separated using a chromatographic gradient of 60 min
200 with a binary system of buffer A (0.1% formic acid) and buffer B (80% ACN, 0.1% formic acid)
201 with a total flow of 250 nl/min. Separation was achieved on in-house made analytical columns
202 (length: 50 cm, inner diameter: 75 μ m) containing 2.7 μ m C18 Poroshell EC120 beads (Agilent)
203 heated to 50 °C in a column oven (Sonation). Over a time period of 41 min, Buffer B was
204 linearly increased from 3% to 30% followed by an increase to 50% in 8 min. Finally, buffer B
205 was increased to 95% within 1 min followed by 10 min washing step at 95% B. Full mass
206 spectrometry (MS) spectra (300-1,750 m/z) were recorded with a resolution of 70,000, a
207 maximum injection time of 20 ms and an AGC target of 3e6. In each full MS spectrum, the top
208 10 most abundant ions were selected for HCD fragmentation (NCE 27) with a quadrupole
209 isolation width of 1.8 m/z and 10 s dynamic exclusion. The MS/MS spectra were then
210 measured with a 35,000 resolution, an injection time of maximum 110 ms and an AGC target
211 of 5e5.

212 The MS RAW files were then analyzed with MaxQuant suite (version 1.5.3.8) on standard
213 settings. By matching against the human UniProt database the peptides were then identified
214 using the Andromeda scoring algorithm (34). Carbamidomethylation of cysteine was defined

215 as a fixed modification, while methionine oxidation and N-terminal acetylation were variable
216 modifications. The digestion protein was Trypsin/P. A false discovery rate (FDR) < 0.01 was
217 used to identify peptide-spectrum matches and to quantify the proteins. Data processing,
218 statistical analysis, as well as clustering and enrichment analysis were performed in the
219 Perseus software (version 1.6.15.0) (35).

220

221 **Immunoblot analysis**

222 Protein samples from co-immunoprecipitation were loaded onto SDS-polyacrylamide gels
223 using SDS-sample buffer, separated by gel-electrophoresis and analysed by immunoblotting.
224 All antibodies were diluted in 50 mM Tris [pH 7.2], 150 mM NaCl with 0.2% Tween-20 and 5%
225 skim milk powder. Antibodies and dilutions are listed in [Supplementary Table 1](#). For
226 visualization, we used Amersham ECL Prime or Select Western Blotting Detection Reagent
227 (GE Healthcare) in combination with the Fusion FX-6 Edge system (Vilber Lourmat).

228

229 **Protein structure modelling and visualization**

230 Chimera X Version 1.1 was used to visualize the structure of the ASAP complex (accession
231 number 4A8X on PDB, (20)).

232

233 **siRNA-mediated knockdowns**

234 $2-3 \times 10^5$ cells were seeded in 6-well plates well and reverse transfected using 2.5 μ l
235 Lipofectamine RNAiMAX and a total of 60 pmol of the respective siRNA(s) according to the
236 manufacturer's instructions. All siRNAs used in this study are listed in [Supplementary Table 1](#).

237

238

239 **RNA extraction, Reverse transcription, endpoint and quantitative RT-PCR**

240 RNA was extracted using different extraction methods. For endpoint or quantitative RT-PCR
241 (RT-qPCR), RNA was extracted using peqGOLD TriFast (VWR Peqlab) or RNA-Solv Reagent
242 (Omega Bio-Tek) following the manufacturer's instructions for TriFast but using 150 μ l 1-
243 bromo-3-chloropropane instead of 200 μ l chloroform and eluting the RNA in 20 μ l RNase-free
244 water. Reverse Transcription was performed using the GoScript Reverse Transcriptase
245 (Promega), 10 μ M VNN-(dT)₂₀ primer and 0.5-1 μ g of total RNA in a 20 μ l reaction volume.
246 RT-PCR and RT-qPCR were performed according to the manufacturer's protocols using
247 MyTaq™ Red Mix (Bioline/BIOCAT) for RT-PCR and GoTaq qPCR Master Mix (Promega) for
248 RT-qPCR. All primers used in this study are listed in [Supplementary Table 1](#).

249

250 **RNA-Sequencing and computational analysis**

251 HTO or HEK 293 cells and the indicated rescue cell lines were treated with siRNA as described
252 above. HTO sets were harvested using TriFast, the HEK 293 set was harvested using RNA
253 Solv Reagent. Extraction of total RNA was performed with DIRECTzol Miniprep Kit (Zymo
254 Research), according to the manufacturer's instructions.

255 For each sample, three biological replicates were analyzed. The Spike-In Control Mix (SIRV
256 Set1 SKU: 025.03, Lexogen), which enables performance assessment by providing a set of
257 external RNA controls, was added to the total RNA, as listed in [Supplementary Table 2](#). The
258 Spike-Ins were used for quality control purposes, but not used for the final analysis of
259 differential gene expression (DGE), differential transcript usage (DTU) or alternative splicing
260 (AS). The cDNA library was prepared using the TruSeq Stranded Total RNA kit (Illumina).
261 Library preparation involved the removal of ribosomal RNA using biotinylated target-specific
262 oligos combined with Ribo-Zero Gold rRNA removal beads from 1 μ g total RNA input. the Ribo-
263 Zero Human/Mouse/Rat kit depleted cytoplasmic and mitochondrial rRNA from the samples.
264 Following a purification step, the remaining RNA was fragmented and cleaved. The first strand

265 cDNA was synthesized using reverse transcriptase and random primers. Subsequently, the
266 second strand cDNA synthesis was performed using DNA Polymerase I and RNase H. To the
267 resulting double-stranded cDNA, a single 'A' base was added and the adapters were ligated.
268 After this, the cDNA was purified and amplified with PCR, followed by library validation and
269 quantification on the TapeStation (Agilent). Equimolar amounts of library were pooled and
270 quantified using the Peqlab KAPA Library Quantification Kit 587 and the Applied Biosystems
271 7900HT Sequence Detection System. Sequencing was performed on an Illumina
272 NovaSeq6000 sequencing instrument with an PE100 protocol.

273 The resulting reads were aligned against the human genome (version 38, GENCODE release
274 33 transcript annotations (36), supplemented with SIRVomeERCCome annotations from
275 Lexogen; (obtained from <https://www.lexogen.com/sirvs/download/>) using the STAR read
276 aligner (version 2.7.3a) (37). Salmon (version 1.3.0) (38) was used to compute estimates for
277 transcript abundance with a decoy-aware transcriptome. Transcript abundances were
278 imported, followed by differential gene expression analysis using the DESeq2 (39) R package
279 (version 1.28.1) with the significance thresholds $|\log_2\text{FoldChange}| > 1$ and adjusted p-value
280 (padj) < 0.05 . Differential splicing was detected with LeafCutter (version 0.2.9) (40) with the
281 significance thresholds $|\text{deltaPSI}| > 0.1$ and $\text{padj} < 0.001$. Alternatively, rMATS (version 4.1.1,
282 (41)) with novel splice site detection was used to identify alternative splicing (AS) classes,
283 followed by analysis using maser (version 1.8.0) and significance thresholds $|\text{deltaPSI}| > 0.2$
284 and $\text{padj} < 0.01$.

285 Differential transcript usage was computed with IsoformSwitchAnalyzeR (ISAR, version
286 1.10.0) and the DEXSeq method (42-47). Significance thresholds were delta isoform fraction
287 $|\text{dIF}| > 0.1$ and adjusted p-value ($\text{isoform_switch_q_value}$) < 0.05 . Intron retention was
288 computed with IRFinder (version 1.2.6, (48)) in FastQ mode and differential intron retention
289 was calculated using DESeq2 with the significance thresholds $|\log_2\text{FoldChange}| > 1$ and padj
290 < 0.001 . Sashimi plots were generated using ggsashimi (version 1.0.0, (49)).

291 **Results**

292 **RNPS1 plays a minor role in NMD**

293 RNPS1 was shown to regulate multiple types of AS in combination with other ASAP/PSAP
294 components in *D. melanogaster* and human cells (22-25). Furthermore, several studies
295 indicated that RNPS1 is able to activate NMD (29,30,50,51) and more recently RNPS1 was
296 reported to be involved in the recognition of many EJC-dependent NMD substrates (31) ([Figure](#)
297 [1A](#)). To investigate the role of RNPS1 in NMD in the context of the EJC, we performed RNA-
298 sequencing (RNA-Seq) analyses of cultured human cells depleted of either RNPS1 or the EJC
299 core factors EIF4A3, MAGOH or RBM8A ([Figure 1B](#)). Additionally, we sequenced RNA from
300 stable cell lines expressing siRNA-insensitive RNPS1 or EIF4A3 constructs to rescue the
301 respective siRNA mediated knockdown (KD) ([Figure 1B and Supplementary Figure 1A, B](#)). In
302 total, we generated three new RNA-Seq datasets from Flp-In-T-REx-293 (HEK 293) cells and
303 HeLa Tet-Off (HTO) cells and re-analyzed existing RNA-Seq datasets of RNPS1 KD-rescue in
304 HeLa Flp-In-T-REx (HeLa FT; E-MTAB-6564)(25).

305 Global differential gene expression (DGE) analysis using DESeq2 identified more than 1000
306 up- or downregulated genes in each RNPS1 KD condition ([Supplementary Figure 1C](#),
307 [Supplementary Table 3](#)). The expression of FLAG-emGFP-tagged full-length RNPS1 (RNPS1
308 FL) or the RRM domain rescued the levels of many genes in HEK 293 and HTO cells. FLAG-
309 tagged RNPS1 FL conferred an almost complete rescue of all events in HeLa FT cells. The
310 rescue with an RNPS1 mutant unable to interact with the ASAP/PSAP complex and the EJC
311 (RNPS1 176) (25), resulted in more mis-regulated genes than the KD alone ([Supplementary](#)
312 [Figure 1C](#)), which indicates that this mutant exerts a dominant-negative effect.

313 We hypothesized that if RNPS1 is indeed required for NMD, many NMD-targeted genes should
314 be upregulated upon RNPS1 KD. As a reference for NMD-targeted genes, we used the DGE
315 analysis of a recent RNA-Seq dataset from SMG7 knockout (KO) HEK 293 cells with additional
316 SMG6 KD ((52); E-MTAB-9330), which displayed nearly complete NMD inhibition. The
317 postulated role of RNPS1 as an NMD activator seems to be supported by a substantial overlap

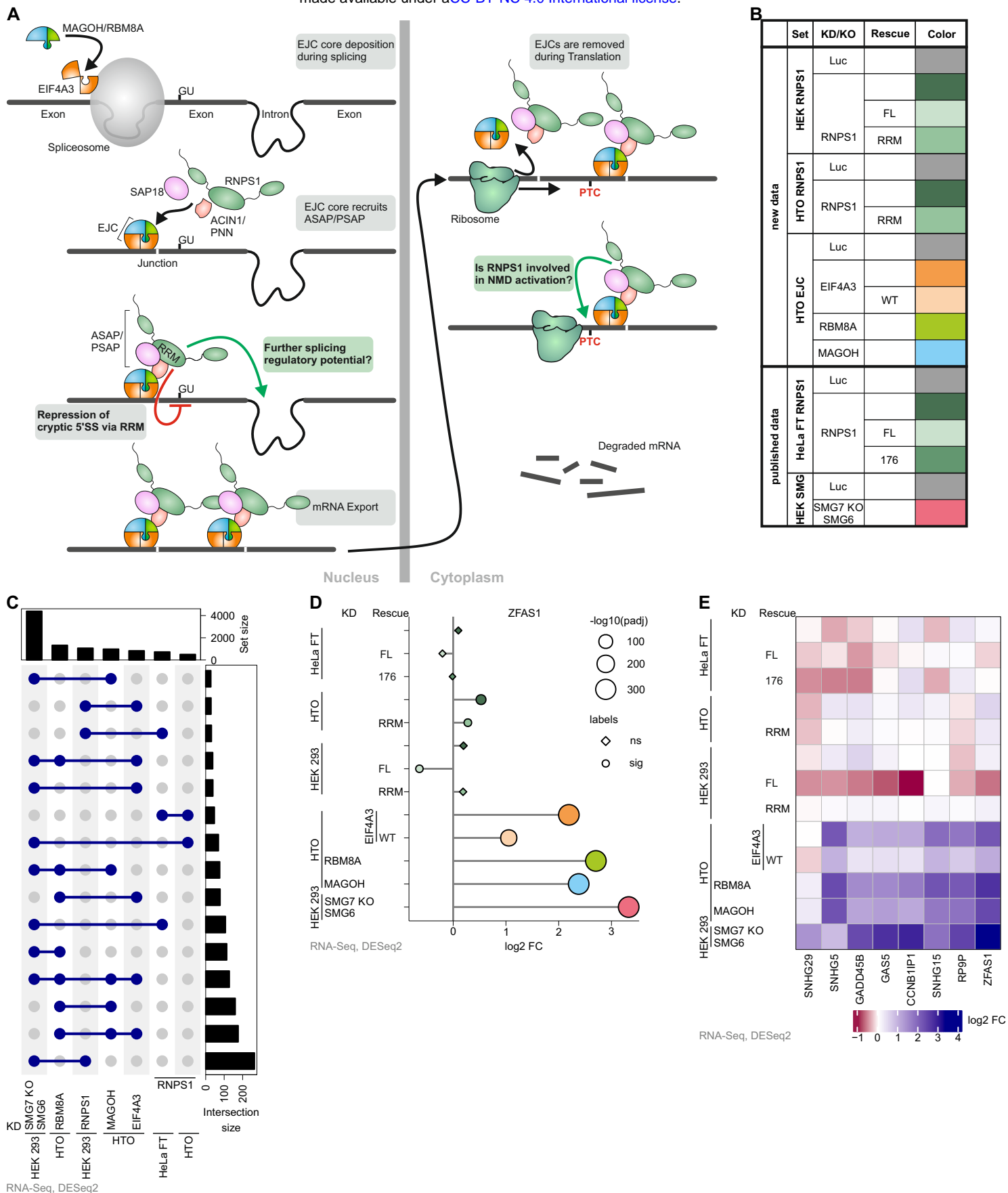


Figure 1: RNPS1 is not required for all EJC-dependent NMD events

(A) Schematic depiction of exon junction complex (EJC) deposition on mRNAs during splicing, recruitment of RNPS1-containing ASAP or PSAP complexes to EJCs, alternative splicing regulation (including cryptic 5' splice site suppression by RNPS1 RRM domain) and NMD activation by RNPS1. Grey boxes indicate established functions and green boxes indicate uncertain functions of RNPS1, which are investigated in this manuscript.

(B) Overview of published and newly generated RNA-Sequencing (RNA-Seq) data sets, indicating which human cell lines, siRNA-mediated knockdown (KD), CRISPR knockout (KO) and, if applicable, rescue construct was employed. Each condition is assigned to a specific color that is used throughout this manuscript.

(C) Differential gene expression (DGE) was analyzed using DESeq2 and upregulated genes identified, cutoffs were \log_2 fold change (\log_2 FC) > 1 and adjusted p-value (padj) < 0.05. Top 15 intersections between the selected RNA-Seq conditions are depicted in an UpSet plot. P-values were calculated by DESeq2 using a two-sided Wald test and corrected for multiple testing using the Benjamini-Hochberg method.

(D) DGE analysis of ZFAS1 \log_2 FCs in the indicated RNA-Seq conditions as compared to the corresponding control. Size depicts the $-\log_{10}(\text{padj})$, shape depicts whether the expression change is significant or non-significant (cutoff adjusted p-value (padj) < 0.001).

(E) Heatmap of \log_2 fold changes of selected, verified SMG6-SMG7 and EJC-dependent NMD target genes.

319 of upregulated genes (27-40%) between the three RNPS1 KD conditions and SMG7 KO with
320 SMG6 KD ([Figure 1C and Supplementary Figure 1D](#)). However, the overlap between RNPS1
321 KD with e.g. the KD of the EJC core factor RBM8A was substantially lower (14-28%). When
322 visualizing the extent of differential gene expression of known NMD targets, e.g. snoRNA host
323 genes ZFAS1 and GAS5 (53), we observed only very small effects of RNPS1 KD compared
324 to the strong upregulation upon EJC or SMG6-SMG7 depletion ([Figure 1D and Supplementary](#)
325 [Figure 1E](#)). Similar trends were observed for other NMD targets ([Figure 1E](#)). Of note, the
326 overexpression of RNPS1 FL robustly led to further downregulation of selected known NMD
327 targets, suggesting that elevated RNPS1 levels can enhance NMD (30).

328 To further characterize the role of RNPS1 in NMD with an orthogonal approach, we analyzed
329 differential transcript usage (DTU) using the IsoformSwitchAnalyzeR package (ISAR) (44).
330 This approach detects upregulated transcripts with annotated PTCs, which indicates NMD
331 inhibition. Depletion of RNPS1 in all three cell lines (HEK 293, HTO and HeLa FT) caused a
332 noticeable upregulation of transcripts bearing a PTC ([Figure 2A, Supplementary Table 4](#)),
333 which was quantitatively less pronounced than in SMG7 KO SMG6 KD. Although this in
334 principle supports a role of RNPS1 in the NMD process, we only found a minimal overlap of
335 PTC-containing isoforms between RNPS1 KD and either EJC or SMG6-SMG7 depletion
336 ([Figure 2B](#)). In contrast, the overlaps between EJC core factor KDs and SMG6-SMG7
337 depletion were robust. When plotted against each other, the event strength of differentially
338 used transcripts found in both RBM8A KD and SMG7 KO SMG6 KD showed good correlation
339 ([Supplementary Figure 2A](#)). In contrast, there were considerably fewer shared transcripts
340 between RNPS1 KD and SMG7 KO SMG6 KD conditions, which also showed weaker
341 correlation ([Supplementary Figure 2B, C](#)). To gain deeper insight into RNPS1's NMD function,
342 we visualized the RNA-Seq data for bona fide NMD targets, such as SRSF2, where the
343 inclusion of an alternative exon and splicing of an intron in the 3' untranslated region activates
344 NMD. (54). Both NMD-activating AS events were clearly visible in the EJC- and SMG6-SMG7
345 depleted conditions, but comparably weak in the RNPS1 KDs ([Supplementary Figure 2D](#)). A
346 likely explanation for the minimal overlap of RNPS1-dependent upregulated PTC-containing

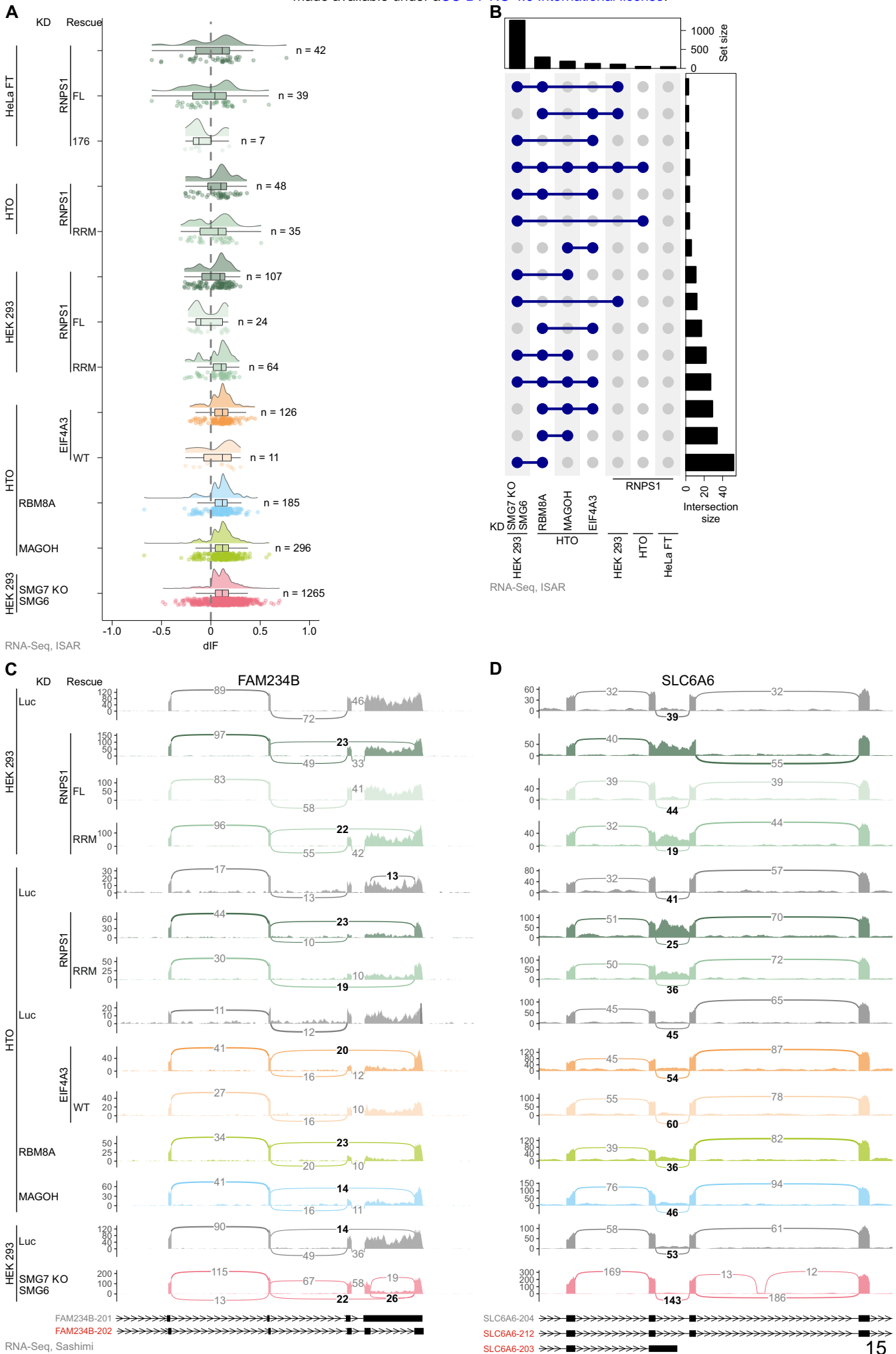


Figure 2: Differential transcript usage analysis reveals RNPS1 main role in alternative splicing rather than NMD activation.
Legend on next page.

Figure 2: Differential transcript usage analysis reveals RNPS1 main role in alternative splicing rather than NMD activation.

(A) Raincloud plot depicting the change in isoform fraction (dIF) for GENCODE (release 33) annotated premature translation termination codon (PTC)-containing isoforms (determined via IsoformSwitchAnalyzeR, ISAR) in the indicated RNA-Seq data. Number of individual events with $\text{padj} < 0.001$ cutoff is indicated on the right. P-values were calculated by IsoformSwitchAnalyzeR using a DEXSeq-based test and corrected for multiple testing using the Benjamini-Hochberg method.

(B) UpSet plot showing the top 15 overlap between the different RNA-Seq conditions with respect to differential transcript usage (DTU).

(C, D) Sashimi plots show the mean junction coverage of the indicated RNA-Seq data with the canonical and NMD-sensitive isoforms for **(C)** FAM234B and **(D)** SLC6A6 depicted below. NMD-relevant and alternatively spliced junctions are highlighted and NMD isoforms are labeled in red.

349 transcripts are many mRNA isoforms that seemed to be mis-classified as NMD targets by the
350 ISAR algorithm due to unannotated AS events. One example is FAM234B, showing the highest
351 delta isoform fraction (dIF) value in HEK293 RNPS1 KD, in which splicing of an intron in the 3'
352 untranslated region can produce an NMD isoform (Figure 2C). Accumulation of this PTC-
353 containing isoform is detected in SMG6-SMG7 conditions, whereas in RNPS1- and EJC-
354 depleted cells the skipping of two exons produces an isoform that does not undergo NMD
355 (Figure 2C). Another example is SLC6A6, for which a strictly RNPS1-dependent intron
356 retention event is erroneously counted as upregulation of a PTC-containing isoform (Figure
357 2D). We found many more cases where our closer inspection revealed that RNPS1 depletion
358 does not lead to the accumulation of NMD-targeted isoforms, but rather to the processing of
359 unannotated and NMD-irrelevant transcripts.

360 In conclusion, at first glance the results from both DGE and DTU analyses seemed to suggest
361 that RNPS1 indeed influences NMD, but in a specific rather than a global way. However, many
362 known NMD targets seemed to be unaffected by RNPS1 depletion, whereas some targets
363 such as ZFAS1 were degraded more efficiently when RNPS1 FL was overexpressed (Figure
364 1D, E). Especially the DTU in-depth analysis confirmed that the main role of RNPS1 is not in
365 promoting degradation of NMD targets, but in regulating AS (Figure 2C, D).

366

367 **Many, but not all RNPS1 dependent alternative splicing events are rescued by** 368 **RNPS1 RRM expression**

369 Encouraged by the findings of the DTU analysis, we next wanted to further characterize the
370 function of RNPS1 in regulating AS. Previous studies detected hundreds of RNPS1-dependent
371 AS events and experiments with a few individual transcripts demonstrated that the isolated
372 RRM can rescue AS events caused by RNPS1 KD (25). However, the RNPS1 RRM rescue
373 did not normalize splicing of FAM234B or SLC6A6 (Figure 2C, D), suggesting that the RRM
374 cannot perform the same splicing regulatory functions as RNPS1 FL.

375 To detect transcriptome-wide AS upon loss of RNPS1 or rescue with the RRM domain, we
376 analyzed the RNA-Seq datasets with the intron-centric LeafCutter software (40). Compared to
377 the almost complete restoration of normal splicing by RNPS1 FL, about two-thirds of AS events
378 could not be rescued by the RNPS1 RRM domain (Figure 3A, Supplementary Table 5).
379 However, it proved difficult to define whether an AS event is fully rescued or not since the
380 outcome relied in part on the chosen computational cutoffs. We also observed partially rescued
381 events when we visualized the AS strength as deltaPSI (dPSI) for all AS events found in both
382 the RNPS1 KD and the RRM or FL rescue data (Supplementary Figure 3A, B). These results
383 suggest that the RRM only incompletely rescues RNPS1-dependent alternatively spliced
384 junctions. We validated this partial rescue with selected transcripts (RER1 and FDPS) using
385 RT-PCR and RT-qPCR in both HEK 293 and HTO cells. Both transcripts are alternatively
386 spliced upon RNPS1 KD. RER1 splicing was rescued by RRM expression but in contrast,
387 splicing of FDPS was still impaired in the HEK 293 and HTO cells expressing the RRM (Figure
388 3B, Supplementary Table 6). This selectivity was also confirmed for two other targets (INTS3
389 and TAF15), of which INTS3 splicing was rescued, whereas TAF15 was not (Supplementary
390 Figure 3C). Since the RNPS1 RRM is required for ASAP/PSAP assembly, which is also
391 essential for EJC interaction, we speculated that the RRM rescues mainly EJC-dependent
392 splicing events. To this end, we examined exemplary RNPS1-dependent splice events in RNA-
393 Seq data from EJC-protein knockdowns and determined the effects of the RRM rescue.
394 MSTO1 and C5ORF22 are two transcripts with increased AS in the RNPS1 KD, which are also
395 found in EJC KD conditions (Figure 3C, D). While normal splicing of MSTO1 is almost
396 completely restored in RRM-overexpressing cells, the AS event in C5ORF22 remains
397 unchanged in the RRM rescue. Hence, RRM rescue and EJC-dependence do not correlate.
398 These results indicate that, irrespective of the cell line used for the rescue assay and the EJC-
399 dependence of the splice event, the RNPS1 RRM is able to rescue some, but not all AS events
400 in RNPS1 KD.

401

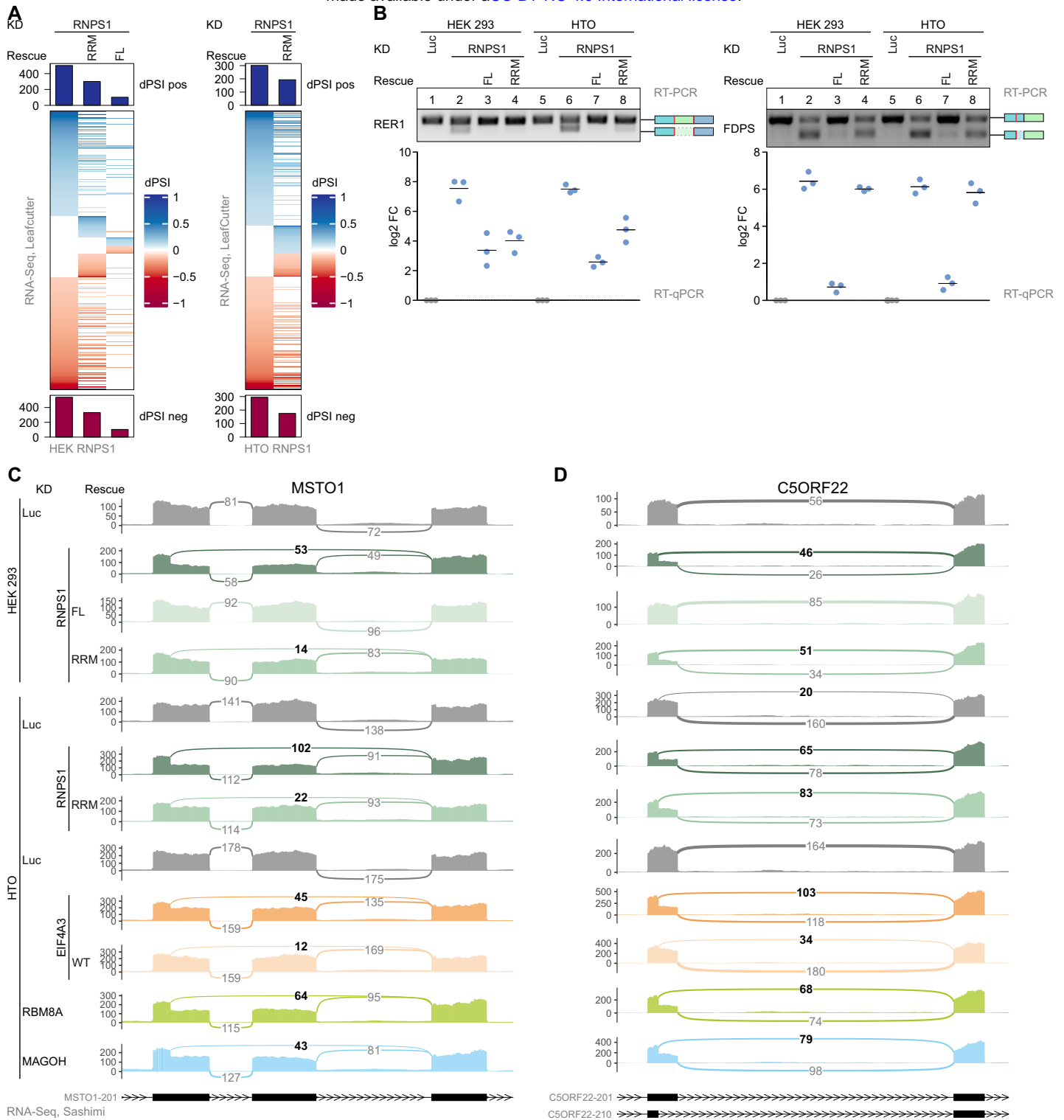


Figure 3: RNPS1 RRM domain partially rescues RNPS1-dependent alternative splicing events.

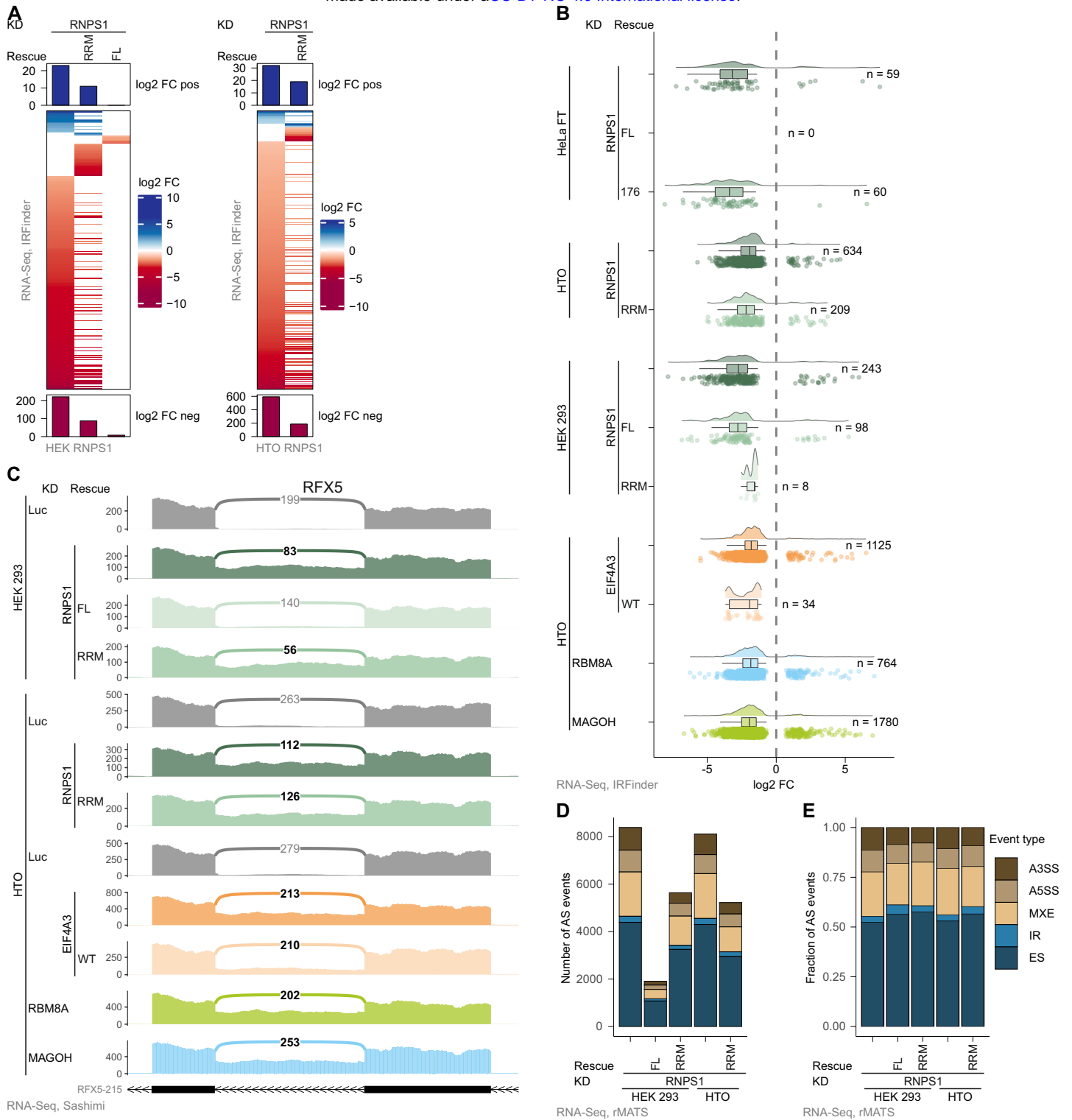
(A) DeltaPsi (Δ PSI) values of ASevents in the indicated RNA-Seq data were calculated with LeafCutter and depicted in a Heatmap (Cutoffs: $|\Delta$ PSI| > 0.1 and p_{adj} < 0.001). P-values were calculated by LeafCutter using an asymptotic Chi-squared distribution and corrected for multiple testing using the Benjamini-Hochberg method.

(B) Comparison of alternatively spliced transcript isoform abundance by RT-PCR and RT-qPCR of RER1 and FDPS in HEK293 or HTO KD and KD/rescue cells with the resulting PCR product indicated on the right. A representative replicate of the RT-PCRs ($n=3$) is shown. The \log_2 FC of RT-qPCRs is calculated as the ratio of alternatively spliced to normally spliced transcript and plotted as datapoints and means.

(C, D) Mean junction coverage in the indicated RNA-Seq condition is depicted as a sashimi plot with the annotated transcript isoforms indicated below and the relevant alternative splice junctions highlighted. Event types are (C) Exon skipping (ES) and alternative 5' splice site (A5SS) usage in MSTO1 and (D) A5SS usage in C5ORF22.

403 **RNPS1 regulates various types of alternative splicing**

404 Previously, RNPS1 was shown to regulate intron retention (IR) in *D. melanogaster* (23,24), but
405 LeafCutter is unable to detect IR events. Hence, we used the IRFinder software to identify
406 RNPS1-regulated retained introns (48). Overall, the RRM rescued more than half of the
407 RNPS1-dependent IR events in HTO and HEK 293 cells (Figure 4A, Supplementary Table 7).
408 However, the splicing of many introns was not rescued at all by the RRM (Figure 4B). The
409 EJC-dependent RFX5 intron 9 retention, for instance, is one of the strongest IR events found
410 in RNPS1 and its splicing was not affected by the RRM (Figure 4C). From a mechanistic point
411 of view, IR is especially interesting, since it represents a seemingly contradictory function of
412 RNPS1: On the one hand, RNPS1 suppresses recursive splicing of cryptic 5' splice sites, but
413 on the other hand it activates splicing of some introns and thereby represses IR. Therefore,
414 we aimed for a deeper analysis of RFX5 intron 9 splicing. We suspected that splicing of the
415 surrounding introns and subsequent EJC deposition and RNPS1 recruitment reinforces correct
416 RFX5 intron 9 splicing. To test this hypothesis, minigene-reporters in which either one or both
417 introns were deleted were designed and stably transfected into HeLa FT cells (Supplementary
418 Figure 4A Top). RT-PCR of the different reporters shows that RFX5 intron 9 splicing relies on
419 the splicing of both surrounding introns, but mostly on the subsequent intron 10
420 (Supplementary Figure 4A Bottom). This finding matches our hypothesis that EJCs deposited
421 at the surrounding junctions stimulate splicing of intron 9, similar to what has been described
422 for the PIWI pre-mRNA in *D. melanogaster* (23,24). In a tethering assay, a RFX5 reporter in
423 which intron 10 is replaced by MS2 stem loops was co-transfected with different RNPS1
424 MS2V5-tagged constructs (Supplementary Figure 4A Top, 4B Top). As seen in the RNA-Seq
425 data, the RNPS1 RRM was unable to rescue RFX5 splicing, even in the tethering assay (Figure
426 4C, Supplementary Figure 4B Bottom). Interestingly, the RNPS1 176 mutant, which cannot
427 assemble ASAP or PSAP and was unable to rescue most RNPS1-dependent AS events, was
428 able to rescue RFX5 intron 9 splicing when tethered to the mRNA. This indicates that IR
429 repression relies on RNPS1 as the effector molecule and does not require complete
430 ASAP/PSAP complexes or EJC recruitment, once RNPS1 is deposited on the mRNA.



432 Although RFX5 correct splicing was not rescued by RNPS1 RRM expression, several IR
433 events were substantially improved, like INTS2 ([Supplementary Figure 4C](#)). Therefore, we
434 were wondering whether we can detect discrepancies between RRM rescue of alternative
435 splice sites and of IR. To reveal possible differences, we classified all RNPS1-dependent splice
436 events into categories (exon skipping (ES), alternative 5' or 3' splice sites (A5SS/A3SS), exon
437 inclusion (EI) and IR) by using rMATS (41) and determined whether the RRM rescues certain
438 forms of AS. Absolute counts of the various types of AS events and also relative proportions
439 revealed that all types of RNPS1-dependent splicing events were equally well rescued by the
440 RRM ([Figure 4D, E, Supplementary Table 8](#)). Taken together, the results of three different
441 bioinformatic analyses show that the RNPS1 KD effects are only partially rescued by the
442 RNPS1 RRM.

443

444 **Many splicing associated factors are recruited by the RNPS1 RRM**

445 Despite the incomplete rescue of RNPS1-dependent AS events, our findings suggest that the
446 RRM can regulate different types of AS events, in addition to the previously shown rescue of
447 cryptic 5' splice sites (25). Therefore, the RNPS1 RRM presumably assembles a splicing-
448 regulatory complex that mediates at least part of the activity of RNPS1 FL. However, neither
449 the mechanism of cryptic 5' splice site suppression nor the factors interacting with the RRM
450 are known in detail. To characterize the components of this putative splicing-regulatory
451 complex, we set out to identify protein factors that interact with the RNPS1 RRM. We generated
452 HEK 293 cell lines expressing the N-terminally FLAG-emGFP-tagged RRM, confirmed its
453 expression by Western blot (WB) and identified co-purified proteins by mass spectrometry
454 (MS) after FLAG-immunoprecipitation (FLAG-IP; [Supplementary Figure 5A](#)). As expected, the
455 RNPS1 RRM efficiently pulled down the three nuclear EJC core components (EIF4A3,
456 RBM8A, MAGOH) and all proteins of the ASAP and PSAP complexes ([Figure 5A, B,](#)
457 [Supplementary Table 9](#)). Furthermore, RRM-containing complexes contained many factors
458 that are involved in splicing or splicing regulation ([Figure 5A](#)). This included spliceosomal or

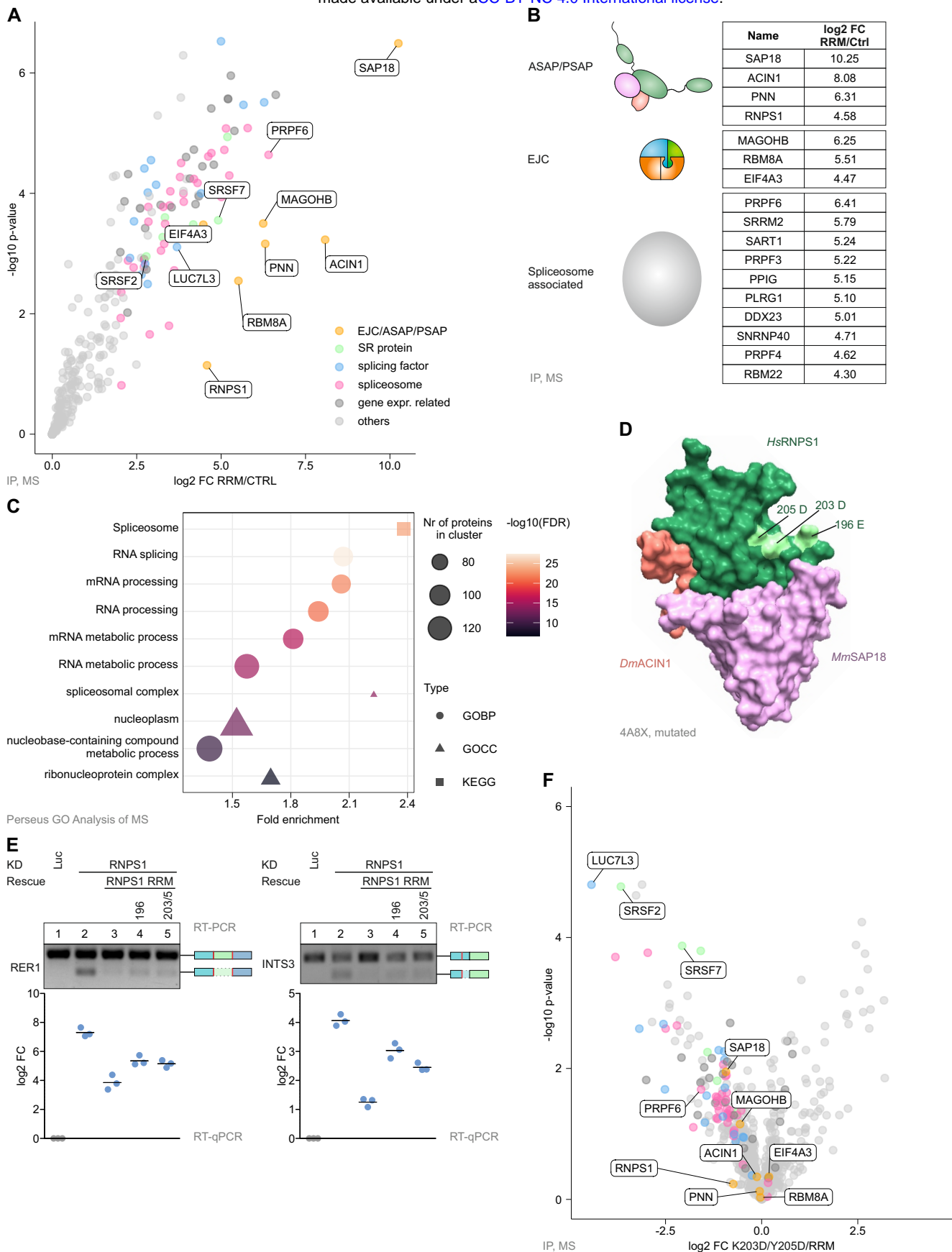


Figure 5: RNPS1 RRM interacts with a broad variety of splicing-regulatory proteins.

(A) FLAG-RNPS1 RRM construct was overexpressed in HEK 293 cells, followed by FLAG-immunoprecipitation (FLAG-IP) and label-free mass spectrometry (MS). The $-\log_{10}$ p-value of identified proteins is plotted against the \log_2 FC in a volcano plot (Cutoff: \log_2 FC ≥ 0). Proteins identified by MS were manually classified into EJC/ASAP/PSAP proteins, SR-Proteins, splicing factors, gene expression related and others (Supplementary Table 6).

(B) List of \log_2 FCs of ASAP/PSAP, EJC and top 10 spliceosome associated proteins found in the RNPS1 RRM MS data.

(C) Clustered RNPS1 RRM MS results were analyzed for enriched GO terms from GO biological processes (GOBP), GO cellular component (GOCC) and KEGG pathway using Perseus software. $-\log_{10}$ false discovery rate (FDR) of the top 10 terms (according to p-value) are plotted against fold enrichment.

(D) Structure of the ASAP complex with the indicated RNPS1 mutated residues highlighted in light green (PDB accession number 4A8X, (Murachelli, 2012 #16)).

(E) RT-PCR and RT-qPCR of RER1 and INTS3 from HEK 293 cells exposed to control or RNPS1 KD and expressing the indicated rescue construct. RT-PCR was performed in triplicates (n=3), one representative replicate is shown and the resulting PCR-product is depicted on the right. For RT-qPCR, the \log_2 FC of the alternatively spliced transcript to the normal transcript is calculated and plotted as datapoints and means (n=3).

(F) Label-free MS of FLAG-IP from cells overexpressing the RNPS1 RRM K203D/Y205D construct. The volcano plot shows \log_2 FC and the corresponding $-\log_{10}$ p-value.

460 spliceosome-associated proteins (e.g. PRPF6, SNRNP40, U2SURP), many SR proteins
461 (SRSF1, 2, 4, 6, 7 and 9) and other splicing factors, like TRA2A, TRA2B and LUC7L3 ([Figure](#)
462 [5A, B](#)). From the various proteins in the RRM interactome, we independently confirmed SRSF7
463 by Western blotting as a robust interaction partner ([Supplementary Figure 5A](#)). We used the
464 Perseus software to cluster proteins enriched in the RRM IP and to identify enriched Gene
465 Ontology (GO) terms for GO biological processes, cellular components, KEGG pathways or
466 PFAM ([Supplementary Table 9](#)) (35). Strikingly, the top 10 terms of this analysis are related to
467 mRNA metabolism or splicing ([Figure 5C](#)). This is in good agreement with the observation that
468 the RRM domain of RNPS1 is sufficient to regulate specific splicing events. However, it
469 remains to be investigated in more detail, whether all splicing factors interact directly with the
470 RRM or are recruited indirectly via the other proteins of the ASAP or PSAP complex. In
471 conclusion, the RRM of RNPS1 equips the EJC with splicing regulatory abilities by directly or
472 indirectly recruiting splicing-associated factors.

473 To address the question whether the RNPS1 RRM can directly recruit splicing-regulatory
474 proteins, we used available structural information (20) to mutate potential binding sites on the
475 RRM. To this end, we generated RRM mutants, in which single or multiple surface-exposed
476 amino acids were mutated ([Figure 5D](#)). We expected these mutations to allow the formation
477 of ASAP/PSAP complexes, but to disrupt interactions with other splicing effectors. To use
478 these mutants in rescue assays, we integrated them into the genome of HEK 293 cells using
479 the PiggyBac system and induced their expression simultaneously with the siRNA-mediated
480 depletion of endogenous RNPS1. Both mutations, R196E and K203D/Y205D ([Figure 5D](#))
481 reduced the ability of the RRM to rescue the cryptic splicing of 5' splice sites in RER1 and
482 INTS3, two RNPS1-dependent splice events ([Figure 5E, Supplementary Table 6](#)).
483 Interestingly, the two mutants are located close to each other on the surface of the RRM,
484 suggesting that they might interfere with the binding of the same protein(s). These mutants
485 enabled us to identify functionally important interaction partners of the RRM using
486 immunoprecipitation and MS. The analysis of the MS data showed that the mutants were still
487 able to interact equally well with the EJC and ASAP/PSAP complex as the wildtype RRM,

488 which was also validated by WB (Figure 5F, Supplementary Figure 5B, C). This indicates that
489 the ability of the RRM to rescue the AS of RER1 and INTS3 is not reduced because of impaired
490 binding to the ASAP/PSAP complex or the EJC. However, the K203D/Y205D pulled down
491 fewer of the splicing-related factors that are efficiently pulled down by the wildtype RRM (Figure
492 5F). Although the effect was not quite as pronounced in the R196E pull down, several splicing-
493 regulatory factors exhibited decreased binding to this mutant, too (Supplementary Figure 5C).
494 Two of the most altered factors were SRSF2 and SRSF7, which were far less efficiently pulled
495 down by both mutants.

496 We therefore conclude that the RRM of RNPS1 is able to regulate splicing by assembling a
497 splicing competent complex, containing inter alia the two SR proteins SRSF2 and SRSF7. The
498 formation of the splice complex is impaired by the mutation of surface patches on the RRM.

499

500 **RNPS1 C-terminus mediates interaction with U1 snRNP**

501 The incomplete splicing rescue by the RNPS1 RRM construct clearly demonstrated that other
502 regions of RNPS1 have to play a key role in the regulation of certain splicing events. Therefore,
503 we aimed for an in-depth analysis of the functional domains of RNPS1, for which we followed
504 the well-known domain architecture of RNPS1. As previously mentioned, the RRM domain
505 was shown to be required for ASAP/PSAP assembly (20). The S-Domain, which is located N-
506 terminally of the RRM, was shown to interact with SRP54, while the C-terminal arginine-
507 serine/proline-rich domain (RS/P) interacts with hTra2 β (55). No interaction partners are known
508 for the N-terminus. We generated various RNPS1 deletion mutants, lacking either one or two
509 of the RNPS1 domains, stably integrated them into the genome of HEK 293 cells and induced
510 their expression shortly after RNPS1 KD (Figure 6A). Subsequently, we analyzed if the mutants
511 were able to rescue the RNPS1-dependent AS events in FDPS and TAF15. These splicing
512 events were fully rescued by RNPS1 FL and the Del-N variant, but not by any other deletion
513 mutant (Figure 6B). The amount of mis-spliced transcript varied between the RNPS1 variants.
514 The differences in rescue activity indicated that different domains perform partially redundant

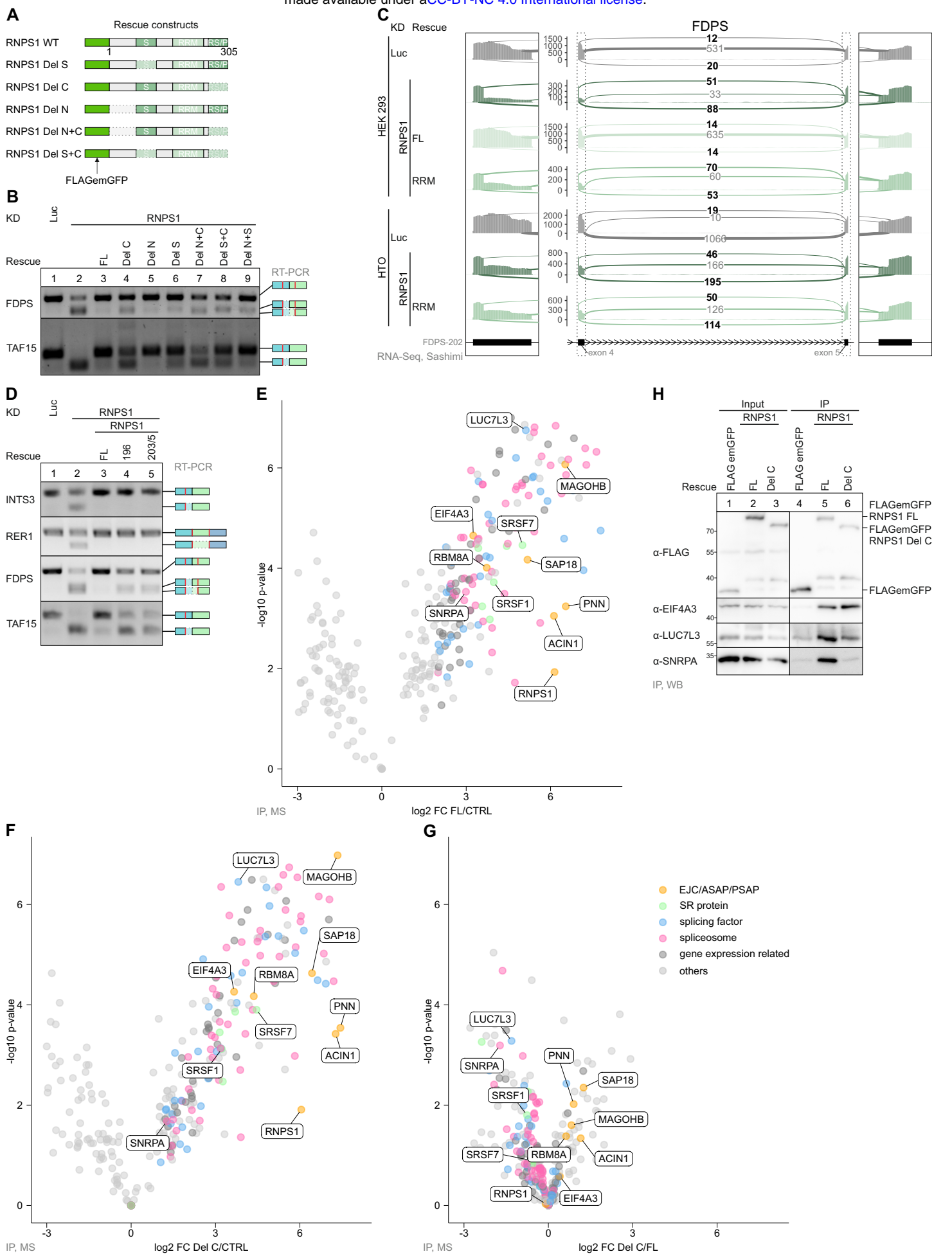


Figure 6: RNPS1 C-terminus is important for regulation of specific alternative splicing events.

(A) Schematic representation of RNPS1 rescue constructs and domain deletions.
(B) RT-PCR of FDPS and TAF15 AS in HEK 293 cells after control or RNPS1 KD with the rescue constructs depicted in (A). One representative replicate is shown (n=3).
(C) FDPS RNA-Seq mean junction coverage is shown as sashimi plot with alternatively spliced junction reads highlighted.
(D) RT-PCR of INTS3, RER1, FDPS and TAF15 was performed in triplicates (n=3) in the indicated KDs and KD/rescues. One representative replicate is shown with the resulting PCR product depicted on the right.
(E, F, G) The $-\log_{10} p$ -value of FLAG-IP MS plotted against $\log_2 FC$ in a volcano plot. For the FLAG-IP, HEK 293 cells overexpressing either a control (CTRL), full-length RNPS1 (FL) or C-terminally shortened RNPS1 (Del-C) were used. Clustering of proteins identified by MS was performed manually into the classes EJC/ASAP/PSAP proteins, SR-Proteins, splicing factors, gene expression related and others (Supplementary Table 7).
(H) FLAG-IPs of HEK 293 cells overexpressing either a CTRL, RNPS1 FL or RNPS1 Del-C analyzed by Western blot (WB). Antibodies used are shown on the left and a representative replicate is shown (n=3).

516 functions and are not equally important for the activity of RNPS1. However, there were
517 variations in detail and some splice events appeared to be domain-specific, which we could
518 observe, for example, for the AS of FDPS (Figure 6B, C). If rescued with the RNPS1 construct
519 lacking the S-domain, an A5SS in exon 4 was used. When the rescue construct lacked the C-
520 terminus, the same A5SS was combined with an A3SS in the fifth exon of FDPS, resulting in
521 a slightly faster migrating band in the gel. In the RNPS1 KD, approximately 30 % of the FDPS
522 transcripts resulted from only the A5SS while 50 % resulted from alternative 5' and 3' splicing
523 and the remaining transcripts were normally spliced. Interestingly, the sashimi plot also shows
524 that the expression of the RRM pushes AS more towards only A5SS usage, similar to the
525 rescue with the Del-S construct (Figure 6B, C).

526 These experiments have established that in some transcripts, certain domains of RNPS1 are
527 essential for normal splicing and their deletion cannot be compensated by other domains, for
528 example the RRM alone. Therefore, we wondered to what extent the splicing function mediated
529 by either the RRM or the other regions in RNPS1 substitute each other. Specifically, we asked
530 if the previously generated surface mutations in the RRM (Figure 5) would affect the function
531 of RNPS1 FL and expressed them in RNPS1 depleted HEK 293 cells (Supplementary Figure
532 6A). The same events in RER1 and INTS3 that were not efficiently rescued by the mutated
533 RRM (Figure 5E) were completely rescued by RNPS1 FL even if its RRM was mutated (Figure
534 6D, 5E). Unexpectedly, splicing of FDPS and TAF15 could not be rescued by RNPS1 FL
535 carrying the RRM mutations, although FDPS and TAF15 are not rescued by the RRM alone.
536 This leads to the paradoxical observation that in the full-length context, mutations in the RRM
537 seem to affect events that require other domains of RNPS1 for their correct splicing. This
538 suggests that for some splicing events the individual domains of RNPS1 not only perform
539 specific functions, but in some cases also act synergistically together. Even the mutation of
540 RRM surface patches is sufficient to severely disrupt some particularly sensitive events. This
541 suggests that one of the functions of RNPS1 is to locally accumulate a certain concentration
542 of splicing factors by recruiting them via its different domains.

543 Overall, the deletion of the C-terminus had the strongest effect on the AS events that we
544 analyzed in detail. Therefore, we set out to identify proteins that interact with RNPS1 via its C-
545 terminus. We expressed and immunoprecipitated FLAG-emGFP-tagged RNPS1 FL and the
546 C-terminally (Del-C) shortened version and analyzed the interactome by MS. As expected, we
547 found many splicing factors and components of the spliceosome. However, we did not detect
548 NMD factors (UPF or SMG proteins), supporting our view that RNPS1 has only a minor function
549 in NMD. Notably, several RNPS1-interacting proteins were pulled down less efficiently by
550 RNPS1 Del-C. For example, RNPS1 FL strongly interacted with several U1 snRNP
551 components, whose interaction was significantly reduced with RNPS1 Del-C ([Figure 6E, F, G,](#)
552 [Supplementary Table 10](#)). The interaction of other splicing factors was also affected by the
553 deletion of the C-terminus. We exemplarily confirmed the reduced interaction with the U1
554 component SNRPA and splicing factor LUC7L3 in a WB ([Figure 6H](#)). In contrast, no difference
555 in binding to FL and Del-C RNPS1 was observed for the EJC and the other ASAP-/PSAP
556 components ([Figure 6E, F, G, Supplementary Figure 6B](#)). In fact, the pull down of EJC and
557 ASAP/PSAP components even seemed to be slightly improved in the Del-C construct ([Figure](#)
558 [6E, G, Supplementary Figure 6B](#)). Overall, our results suggest that the C-terminus of RNPS1
559 mediates a direct or indirect interaction with the U1 snRNP. This interaction could play a role
560 in the selection of RNPS1-dependent 5' splice sites.

561

562 Discussion

563 Although RNPS1 has been the subject of several studies, it remained unclear which of its
564 various functions is most important in the context of the EJC. In this work, we analyze the roles
565 of RNPS1 in NMD and AS regulation and present new molecular details on how AS regulation
566 is mediated by RNPS1. Our analysis of several RNA-Seq datasets suggests that RNPS1 is
567 globally less important for NMD than for example the three core EJC factors EIF4A3, MAGOH
568 or RBM8A. Some NMD targets were upregulated upon RNPS1 depletion, but often only to a
569 low extent ([Figure 1D, 1E and Supplementary Figure 1E](#)). Rescue/overexpression of RNPS1,
570 on the other hand, appeared to slightly enhance NMD efficiency. This is in good agreement
571 with previous work in which NMD activation by tethering RNPS1 to reporter mRNAs was shown
572 (50,56). Similarly, the NMD activity of different HeLa cell strains was previously reported to
573 correlate with their RNPS1 expression levels (30). However, some recently reported RNPS1-
574 dependent NMD targets exhibited no consistent response to RNPS1 depletion and were either
575 slightly up- or even slightly downregulated in our RNA-Seq datasets (31). Furthermore, many
576 transcripts that were identified in the DTU analysis as upregulated NMD isoforms in RNPS1
577 KD turned out - after closer inspection - to be incorrectly classified by ISAR. RNPS1, as
578 previously reported, leads to many new AS events that are not annotated and therefore these
579 new transcript isoforms are apparently mistaken for real NMD isoforms by the ISAR analysis.
580 Altogether, we conclude that RNPS1 is able to increase the NMD efficiency of specific
581 transcripts, which is in good agreement with an NMD-activating function. However, RNPS1
582 does not interact with NMD factors, raising questions about the mechanism of this effect.
583 Overall, our data leave us with the puzzling observation that the overexpression of RNPS1
584 stimulates NMD, while the RNPS1 KD has virtually no effect on NMD.

585 Our finding that many, if not most, of the inspected transcripts identified as RNPS1-dependent
586 NMD targets likely result from AS further emphasizes the importance of splicing regulation by
587 RNPS1. We found previously that the RRM domain is sufficient to regulate some EJC-
588 dependent splicing events (25). Hence, we started our analysis with the initial hypothesis that

589 the RRM domain is sufficient for the regulation of many, if not all RNPS1-dependent splicing
590 events. Unexpectedly, RNA-Seq analyses of RNPS1-depleted cells rescued with the RRM
591 showed that it can only partially replace RNPS1, both quantitatively and qualitatively.
592 Expression of the RRM frequently resulted in incomplete rescue compared to full-length
593 RNPS1. Many other splice events were not rescued at all by the RRM. However, our attempt
594 to classify all RNPS1-dependent splicing events into RRM-rescued and RRM non-rescued did
595 not yield clear results.

596 During the detailed analyses of the RNA-Seq datasets, we found many of the previously
597 described EJC- and RNPS1-dependent splicing events, for example the use of cryptic 5' and
598 3' splice splices (25-28). In addition, we identified multiple examples of introns, for whose
599 efficient splicing RNPS1 was required. Previous studies in *D. melanogaster* showed that
600 splicing of intron 4 of the PIWI pre-mRNA depends on the deposition of EJCs on exon-exon
601 junctions upstream and downstream (23,24). Our findings demonstrate that this phenomenon
602 is also conserved in human cells. We propose a mechanism similar to that in the fruit fly,
603 namely that the splicing of some weak introns is delayed after the surrounding introns are
604 spliced. Such out-of-order splicing has already been shown for other EJC-dependent splicing
605 events (25). We hypothesize that this splice-activating function of RNPS1 is an important
606 contributor to genome maintenance and prevents inadvertent IR. Mechanistically, multiple
607 EJC-bound RNPS1 proteins appear to cooperate hereby, acting in 3' and 5' directions from
608 deposited EJCs.

609 Our global analysis confirmed that the RRM domain of RNPS1 can rescue different classes or
610 splice events. For this purpose, it interacts with a variety of splicing-related proteins including
611 spliceosomal proteins of different snRNPs. This was an unexpected result, since the main
612 splicing regulatory function of RNPS1 was originally attributed to other domains (55), whereas
613 the RRM seemed to be involved mainly in the formation of the ASAP or PSAP complex (20).
614 However, our results cannot exclude that some of the splicing-related proteins in the RRM
615 interactome are recruited via the other proteins of the ASAP or PSAP complex. PNN, ACIN1

616 and SAP18 have all been shown to interact with various splicing factors themselves (57-61).
617 Further analyses will therefore be necessary to identify direct interactions between the
618 individual proteins and to disentangle their precise function.

619 This raises the question if and how other regions of RNPS1 contribute to splicing regulation.
620 To answer this question, we followed previous classifications of the RNPS1 architecture and
621 examined the function of different deletion mutants (55). The deletion of the S-domain affected
622 some splice events, while deletion of the N-terminus had no effect on the examined events.
623 However, we observed the strongest effects with the deletion of the C-terminus and therefore
624 focused our further analyses on the C-terminal region of RNPS1. Our interactome data showed
625 that RNPS1 interacts with U1 snRNP components via its C-terminus ([Figure 6G, 7A,](#)
626 [Supplementary Figure 7A](#)). Since the U1 snRNP binds to and defines the 5' splice sites of
627 introns (62), this interaction could be mechanistically involved in the regulation of 5' splice sites
628 by RNPS1. Interestingly, the suppression of 5' splice sites was one of the most important
629 functions shown for RNPS1 and the PSAP complex in the context of the EJC (25,26). One
630 seemingly obvious explanation would be that other parts of RNPS1, like the RRM or maybe
631 also the S-domain, repress cryptic 5' splice sites, while the RNPS1 C-terminus enhances
632 splicing of nearby 5' splice sites by recruiting the U1 snRNP. Nevertheless, we can only
633 speculate about the exact mechanistic details, and it is also conceivable that the interaction of
634 U1 with the RNPS1 C-terminus prevents cryptic 5' splicing. Further experiments will be
635 required to unravel the molecular mechanism.

636 Apart from the interaction with the U1 snRNP, we were able to detect numerous other
637 interactions of RNPS1 and its C-terminus with splice proteins and the spliceosome. These
638 interactions are very interesting from a splicing regulatory point of view and indicate that full-
639 length RNPS1 seems to have a more diverse interactome than the RRM alone, which itself
640 maintains more interactions than its mutants ([Figure 7A, B, Supplementary Figure 7B](#)).
641 Considering previous data and the data from this work, we suggest that RNPS1 bridges
642 splicing factors and spliceosomal components to the EJC, thereby recruiting variable splicing

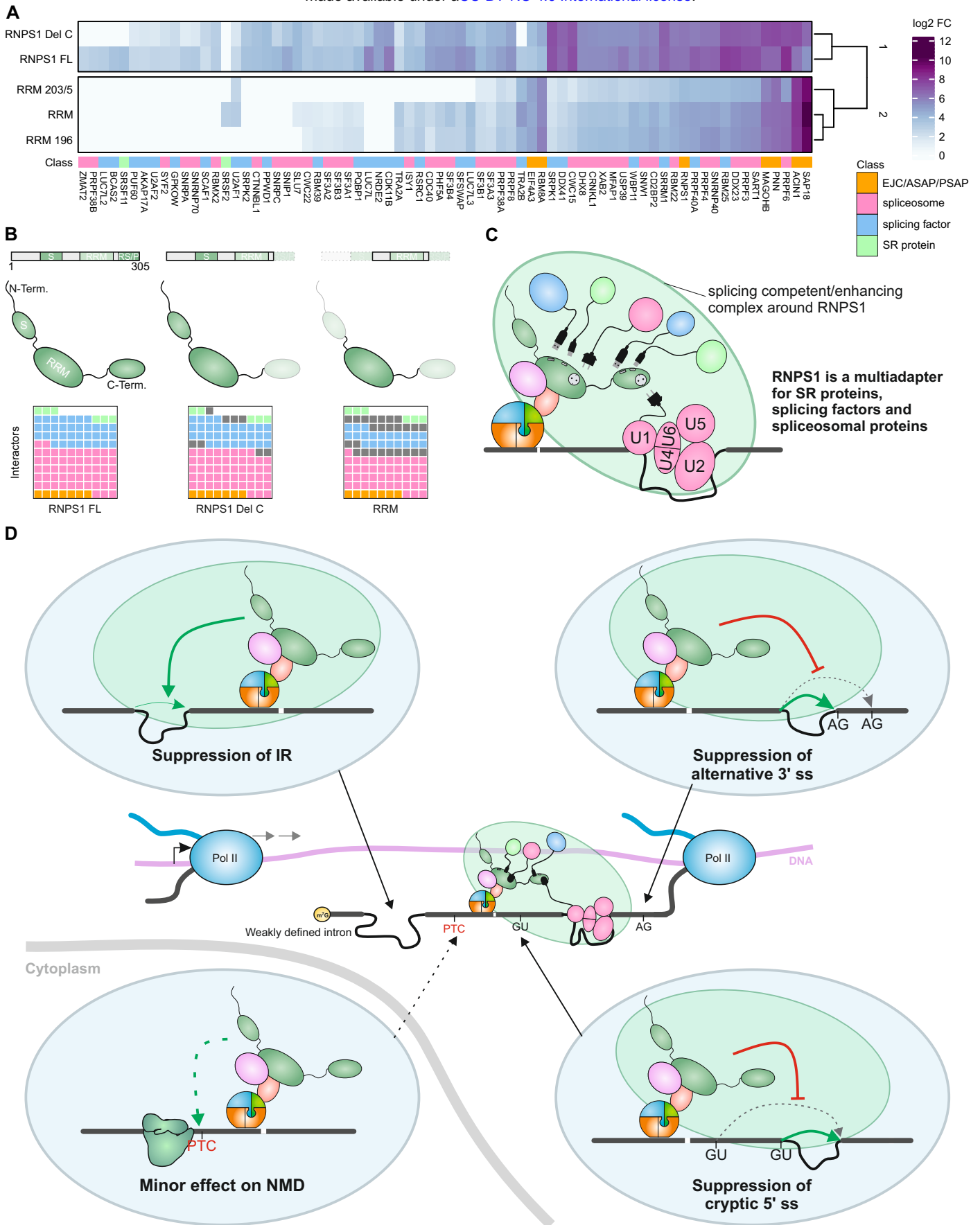


Figure 7: Model for alternative splicing regulation by RNPS1.

(A) Heatmap showing the log₂ FC of RNPS1 interactors compared to the control. Only interactors of the indicated classes are shown. The mean log₂ FC for each interacting protein was calculated across all conditions and the sum of the absolute differences of all conditions to this mean had to be > 4. EJC/ASAP/PSAP proteins are shown for comparison.

(B) Waffle-plots depicting the loss of interaction partners of the indicated classes in the MS of RNPS1 deletion and point mutants.

(C) RNPS1 interacts with spliceosomal proteins, SR-proteins and other splicing factors with its C-terminus and its RRM. A splicing competent or splicing enhancing complex is formed based on RNPS1.

(D) Correct splicing of not-well defined introns requires EJC deposition and RNPS1 recruitment via ASAP/PSAP. By assembling a splicing competent or splicing enhancing complex, RNPS1 prevents IR and represses alternative 3' and alternative 5' splice sites. NMD is mildly activated if RNPS1 is bound to an mRNA.

644 competent complexes to the RNA to guide splicing of nearby introns (Figure 7C). RNPS1
645 seems to act as a "multi-adaptor" that binds the U1 snRNP, SR proteins, LUC7L proteins or
646 other splicing factors as required. Due to the amount of different splicing factors that RNPS1
647 recruits to the pre-mRNA, it can also regulate several different AS events (Figure 7D). The
648 specific way in which RNPS1 acts on each splicing event is determined by the context and the
649 exact position of the EJC. For example, if there are poorly defined introns in its vicinity, RNPS1
650 stimulates their splicing. In the case of cryptic 5' splice sites located downstream of RNPS1, it
651 helps to define exonic regions of the mRNA and prevent their re-splicing. The formation of
652 splice-supporting, high-molecular-weight complexes can best explain the role of RNPS1 and
653 could also serve as a model for the mechanism of other multifunctional splicing factors.
654 Moreover, it also fits well with the higher-order mRNP complexes described in the context of
655 nuclear EJC-bound mRNAs (31). Although our model mainly considers the function of RNPS1
656 in the context of the EJC, it is possible that it can also bind directly to mRNA or is recruited by
657 some of the above proteins to mRNA. This would also explain why not all splicing events
658 regulated by RNPS1 are also EJC-dependent. Interestingly, we had previously observed that
659 the C-terminus can interact nonspecifically with RNA (25). This could reflect its multiple
660 interactions with other RNA-binding proteins, or indicate an intrinsic affinity for RNA.

661 RNPS1 is of great interest as a multifunctional splicing protein, because it can either suppress
662 or activate splice sites and enhance the splicing of weak introns, which might seem
663 contradictory at first (Figure 7D). It carries out these functions in conjunction with various other
664 proteins, especially the EJC. This network of interactions allows RNPS1 to regulate a variety
665 of AS events, as it does not rely on its own RNA-binding ability, unlike SR proteins for example.
666 Thus, RNPS1 could be the prototype of flexible, sequence-independent splice regulators,
667 which can be used in regions where no other splicing enhancers can be present due to
668 evolutionary constraints. It will be interesting to find out if other splicing factors can work in a
669 similar way. On the other hand, the interaction of RNPS1 with the EJC needs to be
670 characterized in more detail. So far, we only have some indications how the ASAP complex
671 might interact with the EJC, but more insights will be needed to better understand the 3D

672 structure of EJC-ASAP or EJC-PSAP assemblies. This would also allow us to understand their
673 effect on adjacent splice sites and introns.

674

675

676 **Data availability**

677 RNA-sequencing data generated for this manuscript have been deposited in the ArrayExpress
678 database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) (63) under accession number E-MTAB-
679 10768 [<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-10768>] for the RNPS1 HTO
680 dataset, accession number E-MTAB-10770
681 [<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-10770>] for the RNPS1 HEK 293
682 dataset and accession number E-MTAB-10770
683 [<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-10770>] for the EJC HTO dataset.

684 Published datasets analysed for this paper include: ArrayExpress accession number E-MTAB-
685 6564 [<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-6564>] for the RNPS1 HeLa
686 FT dataset (25) and ArrayExpress accession number E-MTAB-9330
687 [<https://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9330>] for the SMG7 KO and SMG6
688 KD HEK 293 dataset (52).

689 The mass spectrometry proteomics data have been deposited to the ProteomeXchange
690 Consortium via the PRIDE (64) partner repository with the dataset identifier PXD027251
691 [<https://www.ebi.ac.uk/pride/archive/projects/PXD027251>].

692 Published protein structure of the ASAP complex was used (PDB: 4A8X
693 [<http://doi.org/10.2210/pdb4A8X/pdb>]) (20).

694 All relevant data supporting the key findings of this study are available within the article and its
695 Supplementary Information files or from the corresponding author upon reasonable request.

696 **Funding**

697 This work was supported by grants from the Deutsche Forschungsgemeinschaft to C.D. (DI
698 1501/8-1, DI1501/8-2) and N.H.G (GE 2014/6-2 and GE 2014/10-1) and by the Center for
699 Molecular Medicine Cologne (CMMC, Project C 07; to N.H.G.). V.B. was funded under the
700 Institutional Strategy of the University of Cologne within the German Excellence Initiative.
701 N.H.G. acknowledges funding by a Heisenberg professorship (GE 2014/7-1 and GE 2014/13-

702 1) from the Deutsche Forschungsgemeinschaft. C.D was kindly supported by the Klaus Tschira
703 Stiftung gGmbH (00.219.2013). This work was supported by the DFG Research Infrastructure
704 as part of the Next Generation Sequencing Competence Network (project 423957469).

705 **Acknowledgements**

706 We thank members of the Gehring lab for discussions and reading of the manuscript. We also
707 thank Marek Franitza and Christian Becker (Cologne Center for Genomics, CCG) for preparing
708 the sequencing libraries and operating the sequencer. NGS analyses were carried out at the
709 production site WGGC Cologne.

710 **Author contributions**

711 Conceptualization: Lena P. Schlautmann, Volker Boehm, Niels H. Gehring;

712 Methodology: Lena P. Schlautmann, Volker Boehm, Niels H. Gehring;

713 Software: Lena P. Schlautmann, Volker Boehm;

714 Investigation: Lena P. Schlautmann, Volker Boehm, and Jan-Wilm Lackmann;

715 Resources and Data Curation: Volker Boehm, Janine Altmüller and Jan-Wilm Lackmann;

716 Writing – Original Draft, Review & Editing: Lena P. Schlautmann, Volker Boehm, Niels H.
717 Gehring;

718 Visualization: Lena P. Schlautmann and Volker Boehm;

719 Supervision: Niels H. Gehring;

720 Funding Acquisition: Christoph Dieterich and Niels H. Gehring;

721

722

723 References

- 724 1. Pan, Q., Shai, O., Lee, L.J., Frey, B.J. and Blencowe, B.J. (2008) Deep surveying of
725 alternative splicing complexity in the human transcriptome by high-throughput
726 sequencing. *Nat Genet*, **40**, 1413-1415.
- 727 2. Wahl, M.C., Will, C.L. and Luhrmann, R. (2009) The spliceosome: design principles of
728 a dynamic RNP machine. *Cell*, **136**, 701-718.
- 729 3. Roca, X., Sachidanandam, R. and Krainer, A.R. (2003) Intrinsic differences between
730 authentic and cryptic 5' splice sites. *Nucleic Acids Res*, **31**, 6321-6333.
- 731 4. Wang, G.S. and Cooper, T.A. (2007) Splicing in disease: disruption of the splicing code
732 and the decoding machinery. *Nat Rev Genet*, **8**, 749-761.
- 733 5. Nilsen, T.W. and Graveley, B.R. (2010) Expansion of the eukaryotic proteome by
734 alternative splicing. *Nature*, **463**, 457-463.
- 735 6. Fu, X.D. and Ares, M., Jr. (2014) Context-dependent control of alternative splicing by
736 RNA-binding proteins. *Nat Rev Genet*, **15**, 689-701.
- 737 7. Long, J.C. and Caceres, J.F. (2009) The SR protein family of splicing factors: master
738 regulators of gene expression. *Biochem J*, **417**, 15-27.
- 739 8. Liu, H.X., Zhang, M. and Krainer, A.R. (1998) Identification of functional exonic splicing
740 enhancer motifs recognized by individual SR proteins. *Genes Dev*, **12**, 1998-2012.
- 741 9. De Conti, L., Baralle, M. and Buratti, E. (2013) Exon and intron definition in pre-mRNA
742 splicing. *Wiley Interdiscip Rev RNA*, **4**, 49-60.
- 743 10. Busch, A. and Hertel, K.J. (2012) Evolution of SR protein and hnRNP splicing
744 regulatory factors. *Wiley Interdiscip Rev RNA*, **3**, 1-12.
- 745 11. Cartegni, L. and Krainer, A.R. (2002) Disruption of an SF2/ASF-dependent exonic
746 splicing enhancer in SMN2 causes spinal muscular atrophy in the absence of SMN1.
747 *Nat Genet*, **30**, 377-384.
- 748 12. Witten, J.T. and Ule, J. (2011) Understanding splicing regulation through RNA splicing
749 maps. *Trends Genet*, **27**, 89-97.
- 750 13. Schlautmann, L.P. and Gehring, N.H. (2020) A Day in the Life of the Exon Junction
751 Complex. *Biomolecules*, **10**.
- 752 14. Busetto, V., Barbosa, I., Basquin, J., Marquenet, É., Hocq, R., Hennion, M., Paternina,
753 J.A., Namane, A., Conti, E., Bensaude, O. *et al.* (2020) Structural and functional
754 insights into CWC27/CWC22 heterodimer linking the exon junction complex to
755 spliceosomes. *Nucleic Acids Research*, **48**, 5670-5683.
- 756 15. Gerbracht, J.V. and Gehring, N.H. (2018) The exon junction complex: structural
757 insights into a faithful companion of mammalian mRNPs. *Biochem Soc Trans*, **46**, 153-
758 161.
- 759 16. Le Hir, H., Gatfield, D., Izaurralde, E. and Moore, M.J. (2001) The exon-exon junction
760 complex provides a binding platform for factors involved in mRNA export and
761 nonsense-mediated mRNA decay. *EMBO J*, **20**, 4987-4997.
- 762 17. Lykke-Andersen, J., Shu, M.D. and Steitz, J.A. (2001) Communication of the position
763 of exon-exon junctions to the mRNA surveillance machinery by the protein RNPS1.
764 *Science*, **293**, 1836-1839.
- 765 18. Kishor, A., Fritz, S.E. and Hogg, J.R. (2019) Nonsense-mediated mRNA decay: The
766 challenge of telling right from wrong in a complex transcriptome. *Wiley Interdiscip Rev*
767 *RNA*, **10**, e1548.
- 768 19. Deka, B. and Singh, K.K. (2017) Multifaceted Regulation of Gene Expression by the
769 Apoptosis- and Splicing-Associated Protein Complex and Its Components. *Int J Biol*
770 *Sci*, **13**, 545-560.
- 771 20. Murachelli, A.G., Ebert, J., Basquin, C., Le Hir, H. and Conti, E. (2012) The structure
772 of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex.
773 *Nat Struct Mol Biol*, **19**, 378-386.

- 774 21. Schwerk, C., Prasad, J., Degenhardt, K., Erdjument-Bromage, H., White, E., Tempst,
775 P., Kidd, V.J., Manley, J.L., Lahti, J.M. and Reinberg, D. (2003) ASAP, a novel protein
776 complex involved in RNA processing and apoptosis. *Mol Cell Biol*, **23**, 2981-2990.
- 777 22. Wang, Z., Ballut, L., Barbosa, I. and Le Hir, H. (2018) Exon Junction Complexes can
778 have distinct functional flavours to regulate specific splicing events. *Scientific Reports*,
779 **8**, 9509.
- 780 23. Malone, C.D., Mestdagh, C., Akhtar, J., Kreim, N., Deinhard, P., Sachidanandam, R.,
781 Treisman, J. and Roignant, J.Y. (2014) The exon junction complex controls
782 transposable element activity by ensuring faithful splicing of the piwi transcript. *Genes*
783 *Dev*, **28**, 1786-1799.
- 784 24. Hayashi, R., Handler, D., Ish-Horowicz, D. and Brennecke, J. (2014) The exon junction
785 complex is required for definition and excision of neighboring introns in *Drosophila*.
786 *Genes Dev*, **28**, 1772-1785.
- 787 25. Boehm, V., Britto-Borges, T., Steckelberg, A.L., Singh, K.K., Gerbracht, J.V., Gueney,
788 E., Blazquez, L., Altmüller, J., Dieterich, C. and Gehring, N.H. (2018) Exon Junction
789 Complexes Suppress Spurious Splice Sites to Safeguard Transcriptome Integrity. *Mol*
790 *Cell*, **72**, 482-495.e487.
- 791 26. Blazquez, L., Emmett, W., Faraway, R., Pineda, J.M.B., Bajew, S., Gohr, A.,
792 Haberman, N., Sibley, C.R., Bradley, R.K., Irimia, M. *et al.* (2018) Exon Junction
793 Complex Shapes the Transcriptome by Repressing Recursive Splicing. *Mol Cell*, **72**,
794 496-509.e499.
- 795 27. Joseph, B. and Lai, E.C. (2021) The Exon Junction Complex and intron removal
796 prevent re-splicing of mRNA. *PLoS Genet*, **17**, e1009563.
- 797 28. Fukumura, K., Inoue, K. and Mayeda, A. (2018) Splicing activator RNPS1 suppresses
798 errors in pre-mRNA splicing: A key factor for mRNA quality control. *Biochem Biophys*
799 *Res Commun*, **496**, 921-926.
- 800 29. Gehring, N.H., Kunz, J.B., Neu-Yilik, G., Breit, S., Viegas, M.H., Hentze, M.W. and
801 Kulozik, A.E. (2005) Exon-junction complex components specify distinct routes of
802 nonsense-mediated mRNA decay with differential cofactor requirements. *Mol Cell*, **20**,
803 65-75.
- 804 30. Viegas, M.H., Gehring, N.H., Breit, S., Hentze, M.W. and Kulozik, A.E. (2007) The
805 abundance of RNPS1, a protein component of the exon junction complex, can
806 determine the variability in efficiency of the Nonsense Mediated Decay pathway.
807 *Nucleic Acids Res*, **35**, 4542-4551.
- 808 31. Mabin, J.W., Woodward, L.A., Patton, R.D., Yi, Z., Jia, M., Wysocki, V.H., Bundschuh,
809 R. and Singh, G. (2018) The Exon Junction Complex Undergoes a Compositional
810 Switch that Alters mRNP Structure and Nonsense-Mediated mRNA Decay Activity. *Cell*
811 *Reports*, **25**, 2431-2446.e2437.
- 812 32. Gerbracht, J.V., Boehm, V., Britto-Borges, T., Kallabis, S., Wiederstein, J.L., Ciriello,
813 S., Aschemeier, D.U., Krüger, M., Frese, C.K., Altmüller, J. *et al.* (2020) CASC3
814 promotes transcriptome-wide activation of nonsense-mediated decay by the exon
815 junction complex. *Nucleic Acids Res*, **48**, 8626-8644.
- 816 33. Hughes, C.S., Foehr, S., Garfield, D.A., Furlong, E.E., Steinmetz, L.M. and Krijgsveld,
817 J. (2014) Ultrasensitive proteome analysis using paramagnetic bead technology. *Mol*
818 *Syst Biol*, **10**, 757.
- 819 34. Cox, J., Hein, M.Y., Luber, C.A., Paron, I., Nagaraj, N. and Mann, M. (2014) Accurate
820 proteome-wide label-free quantification by delayed normalization and maximal peptide
821 ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*, **13**, 2513-2526.
- 822 35. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M. and
823 Cox, J. (2016) The Perseus computational platform for comprehensive analysis of
824 (prote)omics data. *Nature Methods*, **13**, 731-740.
- 825 36. Frankish, A., Diekhans, M., Ferreira, A.-M., Johnson, R., Jungreis, I., Loveland, J.,
826 Mudge, J.M., Sisu, C., Wright, J., Armstrong, J. *et al.* (2018) GENCODE reference
827 annotation for the human and mouse genomes. *Nucleic Acids Research*, **47**, D766-
828 D773.

- 829 37. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P.,
830 Chaisson, M. and Gingeras, T.R. (2013) STAR: ultrafast universal RNA-seq aligner.
831 *Bioinformatics*, **29**, 15-21.
- 832 38. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. and Kingsford, C. (2017) Salmon
833 provides fast and bias-aware quantification of transcript expression. *Nature Methods*,
834 **14**, 417-419.
- 835 39. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and
836 dispersion for RNA-seq data with DESeq2. *Genome Biology*, **15**, 550.
- 837 40. Li, Y.I., Knowles, D.A., Humphrey, J., Barbeira, A.N., Dickinson, S.P., Im, H.K. and
838 Pritchard, J.K. (2018) Annotation-free quantification of RNA splicing using LeafCutter.
839 *Nature Genetics*, **50**, 151-158.
- 840 41. Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q. and Xing, Y.
841 (2014) rMATS: robust and flexible detection of differential alternative splicing from
842 replicate RNA-Seq data. *Proc Natl Acad Sci U S A*, **111**, E5593-5601.
- 843 42. Anders, S., Reyes, A. and Huber, W. (2012) Detecting differential usage of exons from
844 RNA-seq data. *Genome Res*, **22**, 2008-2017.
- 845 43. Vitting-Seerup, K. and Sandelin, A. (2017) The Landscape of Isoform Switches in
846 Human Cancers. *Mol Cancer Res*, **15**, 1206-1220.
- 847 44. Vitting-Seerup, K. and Sandelin, A. (2019) IsoformSwitchAnalyzeR: analysis of
848 changes in genome-wide patterns of alternative splicing and its functional
849 consequences. *Bioinformatics*, **35**, 4469-4471.
- 850 45. Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W. and Smyth, G.K. (2015)
851 limma powers differential expression analyses for RNA-sequencing and microarray
852 studies. *Nucleic Acids Res*, **43**, e47.
- 853 46. Sonesson, C., Love, M.I. and Robinson, M.D. (2015) Differential analyses for RNA-seq:
854 transcript-level estimates improve gene-level inferences. *F1000Res*, **4**, 1521.
- 855 47. Robinson, M.D. and Oshlack, A. (2010) A scaling normalization method for differential
856 expression analysis of RNA-seq data. *Genome Biol*, **11**, R25.
- 857 48. Middleton, R., Gao, D., Thomas, A., Singh, B., Au, A., Wong, J.J., Bomane, A., Cosson,
858 B., Eyraas, E., Rasko, J.E. *et al.* (2017) IRFinder: assessing the impact of intron retention
859 on mammalian gene expression. *Genome Biol*, **18**, 51.
- 860 49. Garrido-Martín, D., Palumbo, E., Guigó, R. and Breschi, A. (2018) ggsashimi: Sashimi
861 plot revised for browser- and annotation-independent splicing visualization. *PLoS*
862 *Comput Biol*, **14**, e1006360.
- 863 50. Lykke-Andersen, J., Shu, M.-D. and Steitz, J.A. (2001) Communication of the Position
864 of Exon-Exon Junctions to the mRNA Surveillance Machinery by the Protein RNPS1.
865 *Science*, **293**, 1836-1839.
- 866 51. Singh, G., Jakob, S., Kleedehn, M.G. and Lykke-Andersen, J. (2007) Communication
867 with the exon-junction complex and activation of nonsense-mediated decay by human
868 Upf proteins occur in the cytoplasm. *Mol Cell*, **27**, 780-792.
- 869 52. Boehm, V., Kueckelmann, S., Gerbracht, J.V., Kallabis, S., Britto-Borges, T., Altmüller,
870 J., Krüger, M., Dieterich, C. and Gehring, N.H. (2021) SMG5-SMG7 authorize
871 nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. *Nat*
872 *Commun*, **12**, 3965.
- 873 53. Lykke-Andersen, S., Chen, Y., Ardal, B.R., Lilje, B., Waage, J., Sandelin, A. and
874 Jensen, T.H. (2014) Human nonsense-mediated RNA decay initiates widely by
875 endonucleolysis and targets snoRNA host genes. *Genes & development*, **28**, 2498-
876 2517.
- 877 54. Sureau, A., Gattoni, R., Dooghe, Y., Stévenin, J. and Soret, J. (2001) SC35
878 autoregulates its expression by promoting splicing events that destabilize its mRNAs.
879 *The EMBO journal*, **20**, 1785-1796.
- 880 55. Sakashita, E., Tatsumi, S., Werner, D., Endo, H. and Mayeda, A. (2004) Human
881 RNPS1 and its associated factors: a versatile alternative pre-mRNA splicing regulator
882 in vivo. *Mol Cell Biol*, **24**, 1174-1187.
- 883 56. Gehring, N.H., Neu-Yilik, G., Schell, T., Hentze, M.W. and Kulozik, A.E. (2003) Y14
884 and hUpf3b form an NMD-activating complex. *Mol Cell*, **11**, 939-949.

- 885 57. Singh, K.K., Erkelenz, S., Rattay, S., Dehof, A.K., Hildebrandt, A., Schulze-Osthoff, K.,
886 Schaal, H. and Schwerk, C. (2010) Human SAP18 mediates assembly of a splicing
887 regulatory multiprotein complex via its ubiquitin-like fold. *RNA*, **16**, 2442-2454.
- 888 58. Rodor, J., Pan, Q., Blencowe, B.J., Eyraes, E. and Cáceres, J.F. (2016) The RNA-
889 binding profile of Acinus, a peripheral component of the exon junction complex, reveals
890 its role in splicing regulation. *RNA*, **22**, 1411-1426.
- 891 59. Deka, B. and Singh, K. (2019) The arginine and serine-rich domains of Acinus modulate
892 splicing. *Cell Biol Int*, **43**, 954-959.
- 893 60. Wang, P., Lou, P.J., Leu, S. and Ouyang, P. (2002) Modulation of alternative pre-
894 mRNA splicing in vivo by pinin. *Biochem Biophys Res Commun*, **294**, 448-455.
- 895 61. Zhang, Y., Kwok, J.S., Choi, P.W., Liu, M., Yang, J., Singh, M., Ng, S.K., Welch, W.R.,
896 Muto, M.G., Tsui, S.K. *et al.* (2016) Pinin interacts with C-terminal binding proteins for
897 RNA alternative splicing and epithelial cell identity of human ovarian cancer cells.
898 *Oncotarget*, **7**, 11397-11411.
- 899 62. Roca, X., Krainer, A.R. and Eperon, I.C. (2013) Pick one, but be quick: 5' splice sites
900 and the problems of too many choices. *Genes Dev*, **27**, 129-144.
- 901 63. Athar, A., Füllgrabe, A., George, N., Iqbal, H., Huerta, L., Ali, A., Snow, C., Fonseca,
902 N.A., Petryszak, R., Papatheodorou, I. *et al.* (2019) ArrayExpress update - from bulk to
903 single-cell expression data. *Nucleic Acids Res*, **47**, D711-d715.
- 904 64. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu,
905 D.J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M. *et al.* (2019) The PRIDE
906 database and related tools and resources in 2019: improving support for quantification
907 data. *Nucleic Acids Res*, **47**, D442-d450.

908

909