Splicing-sensitive fusion transcripts associated with key tumor characteristics occur at high frequency in neuroblastoma

Yao Shi[1*], Vilma Rraklli[1], Eva Maxymovitz[1], Shuijie Li[2], Isabelle Westerlund[1], Oscar Bedoya Reina[2], Juan Yuan[1], Petra Bullova[2], C. Christofer Juhlin[3], Adam Stenman[3], Catharina Larsson[3], Per Kogner[4], Maureen J. O'Sullivan[5, 6], Susanne Schlisio[2] and Johan Holmberg[1,*]

[1]Department of Cell and Molecular Biology, Karolinska Institutet, Solnavägen 9, SE-171 65 Stockholm, Sweden

[2]Department of Microbiology, Tumor- and Cellbiology, Karolinska Institutet, Solnavägen 9, SE-171 65 Solna, Sweden

[3]Department of Oncology-Pathology, Karolinska Institutet, Cancer Center Karolinska (CCK), Karolinska University Hospital, Stockholm, Sweden

[4]Department of Women's and Children's Health, Karolinska Institutet, Karolinska University Hospital, SE-171 76 Stockholm, Sweden

[5]Department of Histopathology, Our Lady's Children's Hospital, Dublin, Ireland

[6]Trinity Translational Medicine Institute, Trinity College, Dublin, Ireland

[*]Corresponding authors:

Yao Shi, (yao.shi@ki.se)

Johan Holmberg, (johan.holmberg@ki.se)

Abstract

The paucity of recurrent mutations has hampered efforts to understand the pathogenesis of neuroblastoma. Through analysis of RNA-sequenced neuroblastoma, we identified >900 primarily intrachromosomal fusion transcripts generated by genes in close proximity. Fusions were enriched in chromosomal regions gained or lost in neuroblastoma and included well-known neuroblastoma oncogenes. The majority of fusions contained canonical splicing sites and a subset exhibited increased sensitivity to spliceosome inhibition. As a proof-of-principle that a gene product with altered properties can be produced by these fusions, we characterized the *ZNF451-BAG2* fusion which generates a truncated BAG2-protein capable of inhibiting retinoic acid-induced differentiation. Our findings elucidate a mechanism through which altered gene products, relevant for neuroblastoma pathogenesis and representing possible novel drug targets, can be generated.

Introduction

Despite intense sequencing efforts few recurrently mutated genes have been identified in neuroblastoma (1, 2), resulting in a deficiency of drug targets. Instead, high-risk neuroblastoma is characterized by large-scale chromosomal rearrangements such as chromothripsis and loss or gain of chromosomal regions (e.g. loss of 1p36 and 11q or gain of 17q and 2p) with or without *MYCN* amplification (1, 3). Certain other types of tumors harbor and are driven by fusion proteins generated by chromosomal translocations (e.g. BCR-ABL in chronic myelogenous leukemia) (4). An additional mode, through which fusion transcripts can be generated, is represented by *cis*-splicing of adjacent genes (5, 6). Besides a fusion resulting from small interstitial genomic deletions at 11q generating either a *MLL-FOXR1* or a *PAFAH1B2-FOXR2* fusion (7) no intra-chromosomal chimeric transcripts have been described in neuroblastoma. However, they have been shown to be present in different tumor types as well as in non-transformed tissues and be promoted by different types of cellular stress such as infections or mutations (8-12). In order to explore whether neuroblastoma tumors harbor previously undetected gene fusions, we analyzed a cohort of 172 sequenced neuroblastoma tumors. We

identified an abundance of fusion transcripts, of which a significant proportion exhibited a distinct genomic distribution according to tumor risk. Identified fusions were predominantly generated by genes in close proximity and flanked by canonical splicing donors and acceptors. This pattern was distinct to fusions unique for neuroblastoma, whereas fusions we identified in normal adrenal gland or in other tumors did not exhibit such a pattern. Furthermore, a subset of identified NB specific fusions was hypersensitive to pharmacologic spliceosome inhibition in comparison to their wild type cognates. High expression levels of spliceosome factors were strongly associated with high-risk disease and spliceosome inhibition also promoted apoptosis in neuroblastoma cells. As a proof of principle, that fusions can generate novel gene products with alternative properties, we cloned and characterized the *ZNF451-BAG2* fusion. The generated protein exhibited distinct protein-protein binding properties compared to wild-type BAG2 and impeded retinoic acid induced differentiation. This reveals how a fusion gene product can influence neuroblastoma response to a drug commonly used in the treatment of high-risk patients (13).

Results

Fusion transcripts are a common feature of neuroblastoma

To reveal novel gene fusions in neuroblastoma we analyzed a data set (National Cancer Institute TARGET, dbGap Study Accession: phs000218.v16.p6) comprising 172 paired-end RNA sequenced neuroblastoma tumors (referred to as "NB172"), out of which 139 were diagnosed as high-risk, 19 as intermediate-risk and 14 as low-risk according to the Children's Oncology Groups staging (COG), (Supplementary Table 1). We applied the fusion detection tool FusionCatcher (14) and identified chimeric transcripts in 163 out of 172 cases with an average of 31 distinct fusion transcripts per tumor (Supplementary Table 2). Short homologous sequences (SHS) have been suggested to serve as templates for reverse transcriptase dependent false positive chimeras/fusions (15). In order to avoid potential false positive fusions, we removed any fusion that contained genes with SHSs of five or more nucleotides, which reduced the number of identified fusions from 1073 to 924. The structural consequences of the fusions ranged from truncated proteins through bona fide fusion proteins to

deletion of genes. The majority of fusions (786/924) revealed by our analysis were intra-chromosomal fusion transcripts (Supplementary Table 2-3) many of which consisted of adjacent genes. Importantly, 114 fusions occurred at a frequency of 5% or more (Top 25 in Fig. 1a, all >10% in Fig. 1b and full list in Supplementary Table 3) and all of these were intrachromosomal. There was a significant enrichment of fusion junctions at chromosomes 17 and 22 (Supplementary Fig. 1a). Furthermore, there was a significant enrichment of fusions that occurred in >10% of tumors at the same chromosomes (Fig. 1b). Gain of 17q is the most frequently occurring genomic alteration in high-risk neuroblastoma and a marker for adverse clinical outcome (3), whereas 22q alterations have been reported to be involved in the transition to metastatic and more aggressive neuroblastoma (16). We analyzed the fusion transcripts occurring exclusively in low/intermediate-risk and exclusively in high-risk tumors as well as fusion transcripts common to low/intermediate-risk and high-risk levels (Supplementary Table 4-5). Chimeric transcripts unique to low/intermediate-risk tumors exhibited significant enrichment at several chromosomal arms including 11q, a region commonly lost in high-risk neuroblastoma (Fig. 1c). In contrast, both common and high-risk unique fusion transcripts were enriched at 17q and 22q but not at 11q (Fig. 1c), with a pronounced increase in frequency of 17q fusion transcripts in high-risk tumors (Fig. 1c). Thus, with increased risk the frequency of 17q located fusion junctions also increases and was more than seven times higher than the average fusion rate per chromosomal arm. Several fusion transcripts encompassed factors involved in neuroblastoma pathogenesis, including *ARID1B*, *CASZ1*, *HDAC8*, *LMO1, MYCN, BRCA1, TERT* and *PDE6G* (Supplementary Table 6). Notably, tumors harboring fusion transcripts of well-known neuroblastoma oncogenes (e.g. *MYCN* and *LMO1*) also exhibited high expression levels of their wild-type cognates (Fig. 1 d-e). The high-risk susceptibility locus in *LMO1* is significantly associated with MYCN-non amplified high-risk neuroblastoma but not with MYCN-amplified high-risk neuroblastoma (17), interestingly the *LMO1-RIC3* fusion (Supplementary Fig. 1b) was exclusively detected in MYCN-non amplified high-risk neuroblastoma (Supplementary Table 6). The *BRCA1-VAT1* fusion was also only detected in MYCN-non amplified high-

risk cases; previously it has been shown that copy number amplification of *BRCA1* in NB is restricted to cases lacking MYCN-amplification (18) (Supplementary Table 6).

Validation of fusion transcripts specific for neuroblastoma

To corroborate the fusion transcripts observed in the NB172 dataset in an independent cohort, we performed paired-end RNA-sequencing in an additional cohort containing 14 neuroblastoma patient samples, together with eight neuroblastoma cell lines, NB-validation (NB-v, Supplementary table 7, Materials and Methods). We identified 139 fusions, of which 82 (~59%) were present in the NB172 dataset (Fig. 2a and Supplementary Table 8). To investigate whether the identified fusions were neuroblastoma specific, we analyzed a cohort of 161 sequenced tissue samples from human normal adrenal glands (19). Out of 342 detected fusions in the adrenal gland cohort, only 23 (~6.7%) were present in the NB172 dataset and only 4 (~1.2%) of these were present in the NB-v cohort (Fig. 2a and Supplementary Table 9). This enrichment of common fusion transcripts in the neuroblastoma cohorts vs. the adrenal gland dataset was highly significant (chi-square test with Yate's correction, p-value<0.0001). Fusion transcript associated genes unique to and shared by the two neuroblastoma datasets were enriched at 17q and 2p, two chromosomal regions where gains are closely associated with high-risk neuroblastoma (Fig. 2b). Parametric analysis of gene set enrichment (PAGE) (20) comparing high-risk grade 4 tumors with low-risk grade 4s tumors (according to the International Neuroblastoma Staging System, INSS) in the R2 498-SEQC data base (21) showed that the NB172/NB-v common genes identified in Fig. 2a are enriched in the grade 4 high-risk tumors (Fig. 2c). To further investigate whether the identified fusions are distinct for neuroblastoma we analyzed a set of 65 sequenced rhabdoid tumors (National Cancer Institute TARGET, dbGap Study Accession: phs000470.v17.p7) wherein 2055 unique fusion transcripts were detected. However, the overlap with the fusions detected in NB172 dataset was limited to 44 transcripts (~2.1%) (Fig. 2d). In a cohort of 177 sequenced osteosarcoma tumors (National Cancer Institute TARGET, dbGap Study Accession: phs000468.v17.p7; Fig. 2d) we could detect 1650 unique fusion transcripts but there was no overlap

with the NB172 cohort (Fig. 2d). In contrast to fusions detected in neuroblastoma, the majority of detected fusions in rhabdoid tumors and osteosarcoma were inter-chromosomal (Supplementary Fig. 2a). Detected fusions that occur at higher frequencies than 5% are more abundant in neuroblastoma (in 12.3% of tumors) than in rhabdoid tumor (5.7%) and osteosarcoma (0.4%) (Supplementary Fig. 2b). In osteosarcoma there was a considerable number of tumors (~24.9%, 44/177) harboring fusions predicting substantial deletion of the P53 tumor suppressor, disruption of the gene or a truncation at the C-terminal (Fig. 2e, Supplementary Table 10). Notably, *TP53* is one of the most frequently altered genes in osteosarcoma (22). As a comparison we could not detect any fusions containing *TP53* in the neuroblastoma tumors (0/172) (Fig. 2e).

For further validation, we designed primers spanning the fusion junctions of selected chimeric transcripts *VPS45-PLEKHO1, METTL23-MFSD11, HDAC8-CITED1, ZNF451-BAG2, TRIM3-HPX* and *PRR11-SMG8* (Supplementary Table 11). We proceeded to perform RT-PCR in 10 neuroblastoma tumor samples and in a panel consisting of cDNA from 14 untransformed human tissues. Our expression analysis revealed that all selected candidates, except *PRR11-SMG8*, were expressed in a neuroblastoma specific manner (Fig. 2f). For validation, PCR-products including four additional fusions, *TAF15-AC015849.2, FADS1-TMEM258, CHCHD10-VPREB3* and *LMO1-RIC3*, were excised, inserted into the pCR-Blunt II-TOPO vector and subsequently sequenced. All the sequenced PCR-products exhibited an identical sequence of nucleotides to that of the fusions identified by FusionCatcher (Fig. 2f and Supplementary Table 11-12). A previous report identified the *VPS45-PLEKHO1* fusion in non-transformed tissue (12), but it was not detected in our panel of non-transformed tissues.

Transcriptional profile of *ZNF451-BAG2* positive neuroblastoma predicts poor clinical outcome and high tumor risk

One of the 25 most frequently occurring identified fusion transcripts encompassed the *BCL2 associated athanogene* (*BAG2*) which encodes a co-chaperone, BAG2, involved in targeting misfolded proteins for

157 degradation through an ubiquitin independent pathway (23). BAG2 levels have previously been shown
158 to increase upon neuronal differentiation in neuroblastoma cells (24). In addition, BAG2 clears
159 phosphorylated TAU from neuronal microtubule (23), potentially promoting stabilization of axons, an
160 important feature of neuronal differentiation. We thus selected *ZNF451-BAG2* to investigate if the
161 presence of a fusion transcript actually would generate a protein with altered functional properties.
162
163 To elucidate whether expression of the *ZNF451-BAG2* fusion correlates with altered expression levels
164 of transcripts predicting clinical outcome we analyzed nine RNA-sequenced neuroblastoma tumors of
165 the NB-v cohort that had been validated by RT-PCR (Fig. 2f). Tumors harboring the *ZNF451-BAG2* fusion
166 (3/9 tumors, A13, A14 and A15) had significantly elevated expression of 32 genes and 34 genes with
167 lower expression (Fig. 3a). To correlate these differentially expressed genes with clinical outcome we
168 utilized a cohort of 498 sequenced neuroblastoma tumors (498-SEQC) available in the R2 database
169 (21). The majority of genes with elevated expression was also enriched for in high-risk tumors whereas
170 the opposite was the case for genes with lower expression (Fig. 3b). Consequently, several genes with
171 elevated expression in *ZNF451-BAG2* neuroblastoma were strong indicators of shorter overall survival
172 e.g. *ENOSF1* (exemplified in Fig. 3c). In contrast, genes that are strong predictors of longer overall
173 survival showed decreased expression in *ZNF451-BAG2* expressing neuroblastoma cases (exemplified
174 in Fig. 3d). In addition, k-means analysis revealed that a subset of *ZNF451-BAG2* associated transcripts
175 clustered the 498-SEQC cohort into two groups (Fig. 3e). Group 1, with predominantly low expression
176 levels, consists mainly of low-risk tumors, with low stages according to INSS and lack of *MYCN*
177 amplification whereas the opposite is evident for group 2 wherein expression levels of *ZNF451-BAG2*
178 associated transcripts are high (Fig. 3e). Consequently, patients with group 2 tumors have a
179 significantly shorter overall survival (Fig. 3f). Gene set enrichment analysis (25) of transcriptional
180 differences between *ZNF451-BAG2* expressing NB and those lacking *ZNF451-BAG2* expression showed
181 a significant enrichment of cell cycle associated gene sets and a depletion of apoptosis related gene
182 set (Fig. 3g-h).

183

184 **The *ZNF451-BAG2* fusion generates a truncated BAG2 protein, present in a subset of neuroblastoma**
185 **tumors**

186 The *ZNF451-BAG2* fusion spans the 3' UTR or exon 14 of *ZNF451* and the second exon of *BAG2,*
187 potentially generating a truncated *BAG2* transcript lacking the first exon (Fig. 4a). Its first exon encodes
188 part of a coiled-coil domain that is absent in the *ZNF451-BAG2* fusion (Fig. 4a). Full length *BAG2*
189 encodes a 23.8 kDa protein, whereas the *ZNF451-BAG2* fusion transcript encodes a smaller 19.6 kDa
190 protein (ΔBAG2) (Fig. 4a). The *ZNF451-BAG2* chimera was present in 31 of the 172 sequenced tumors
191 (18%). Alignment of wild-type BAG2 protein (BAG2) across different species showed that in ΔBAG2 the
192 highly conserved N-terminal coiled-coil domain was truncated (Supplementary Fig. 3a), implying
193 functional relevance of the truncated region for BAG2. To understand if tumors wherein *ZNF451-BAG2*
194 was identified (Fig. 2f) also had detectable levels of ΔBAG2 protein, we performed immunoblotting
195 with an antibody targeting BAG2. All probed (n=13) tumors contained BAG2 protein at varying levels
196 however only five tumors also co-expressed detectable levels of ΔBAG2, while no detectable levels of
197 ΔBAG2 were observed in tissue from six human normal adrenal glands (Fig. 4b and Supplementary Fig.
198 3b).

199

200 ***ΔBAG2* impairs clearance of phosphorylated TAU and binding to HSC70**

201 BAG2 has been shown to be important for clearance of phosphorylated forms of the TAU protein
202 (pTAU) and thus been implicated as a stabilizer of microtubules (23). To test if this capacity was
203 attenuated by the presence of ΔBAG2 we probed the levels of pTAU in a panel of neuroblastoma
204 tumors and normal adrenal glands. This showed a clear association between the presence of pTAU and
205 endogenous ΔBAG2 (Fig. 4b and Supplementary Fig. 3b). To validate that this was caused by ΔBAG2
206 protein expression, we cloned and validated the *ΔBAG2* transcript where after we expressed it and
207 *BAG2* alone or in combination in SK-N-FI neuroblastoma cells. Upon *BAG2* overexpression, the levels
208 of pTAU were significantly reduced whereas total TAU was present in amounts similar to those in

control-transduced cells (Fig. 4c). *ΔBAG2* overexpressing cells retained pTAU levels and more importantly, upon co-expression BAG2 failed to clear pTAU (Fig. 4Cc), implying that ΔBAG2 can act as a negative regulator of BAG2 function. BAG2 has been shown to bind the heat shock cognate 70 (HSC70) (26), a chaperone protein important for pTAU ubiquitination (27) and axon outgrowth (28). Overexpression of *BAG2* and *ΔBAG2* in SK-N-FI neuroblastoma cells followed by BAG2 immunoprecipitation via FLAG revealed that BAG2 but not ΔBAG2 binds a 70 kDa protein (Supplementary Fig. 3c). Since BAG2 has been reported to bind HSC70 (26), we performed additional immunoprecipitation followed by immunoblotting with a HSC70 specific antibody. This revealed that BAG2 binds HSC70, whereas ΔBAG2 does not (Fig. 4d).

ΔBAG2 impedes differentiation of neuroblastoma in response to retinoic acid

To investigate whether ΔBAG2 impinges on the capacity of neuroblastoma cells to differentiate, we treated *CTRL* (empty vector)*, BAG2* or *ΔBAG2* expressing SK-N-FI cells (Fig. 4E) with retinoic acid (RA), a compound used in adjuvant therapy of high-risk neuroblastoma patients (13). After six days of RA treatment, cells transduced with either *CTRL* (Fig. 4f-g, l) or *BAG2* (Fig. 4h-i, l) acquired neuronal morphology with long neurites. In contrast, *ΔBAG2* transduced cells exhibited a weaker response to RA, with significantly less neurite formation (Fig. 4j-l). To validate this effect we transduced SK-N-BE(1) neuroblastoma cells with a doxycycline-inducible version of the *ΔBAG2* fusion genes (Fig. 4m). Upon doxycycline induction, *ΔBAG2* expressing cells exhibited a reduced capacity to respond to RA (Fig. 4n-r).

Alternative splicing affects the generation of neuroblastoma specific fusions

Previously it has been suggested that *cis*-splicing between adjacent genes can generate fusion transcripts in prostate cancer (5, 6). To elucidate whether there was a correlation between distance and frequency we plotted the distance between 5’ and 3’ of the fusion junction in the intrachromosomal fusion transcript versus the fusion frequency in the NB172 and NB-v data sets (Fig.

5a). Fusions occurring at high frequency in both data sets were enriched between 1 to 100kb, whereas transcripts separated by more than 100kb occurred at lower frequencies and almost exclusively in the larger NB172 data set (Fig. 5a). The spatial proximity of identified transcripts suggests that these fusions are the result of *cis*-splicing (5, 8). Inspection of nucleotide sequences located at 5' and 3' of the fusion junctions for canonical splicing donors (GT) and acceptors (AG) revealed that 81.9% carried GT and AG at the 5' and 3' fusion sites (GT*AG) (Fig. 5b). Notably, 87% of intra-chromosomal fusions had GT*AG at the fusion junction whereas only 45.9% of inter-chromosomal fusions contained canonical splice sites at the junctions. This pattern was unique to neuroblastoma as fusions detected in normal adrenal gland, osteosarcoma and rhabdoid tumor exhibited no enrichment of splice sites (Fig. 5b). Aberrant RNA splicing has been suggested as a driving event for several cancers and mutations in genes coding for components of the spliceosome have been identified in several tumors (29). Furthermore in breast cancer, it has been shown that an intact spliceosome is required to tolerate oncogenic MYC hyperactivation (30). We compared expression levels of genes of the KEGG, "Spliceosome" gene category between neuroblastoma and normal adrenal gland. Out of 134 genes 46 had significantly higher expression levels in the neuroblastoma, whereas only three had lower levels of expression (Fig. 5c). To further elucidate whether there were clinically relevant differences in expression of spliceosome genes between low- and high-risk neuroblastoma we performed k-means analysis of the 498-SEQC neuroblastoma cohort. Together with previous observations (31) our analysis elucidates how the differential expression of spliceosome factors clearly identifies tumors of different clinical outcome, with high expression levels of splicing factors predicting high-risk tumors with bleak clinical outcome and substantially shorter overall survival (Fig. 6d-e). Furthermore, previous genome sequencing studies of neuroblastoma patients revealed mutations in several spliceosome factors in primary tumors and *de novo* mutations in spliceosome factors occurred in relapsed tumors (1, 2, 32, 33) (summarized in Supplementary Table 13). To investigate whether inhibition of spliceosome activity would selectively impede the generation of fusion transcripts but not their wild type cognates, we treated neuroblastoma cells (LAN-1 and SK-N-BE(1)) with the spliceosome inhibitor pladienolide B (34).

Upon treatment, there was a loss of expression for a majority of selected high frequency fusion transcripts whereas expression of most wild type genes constituting the fusions were unaffected at these concentrations (Fig 6a-b and Supplementary Fig. S4). To elucidate if selective loss of fusion transcripts upon spliceosome inhibition was associated with increased apoptosis we treated the neuroblastoma cell lines LAN-1 and SK-N-BE(1) with increasing concentrations of pladienolide B. As controls we utilized non-transformed human diploid fibroblasts (HNDF). Already at 5nM the neuroblastoma cell lines exhibited increased levels of cleaved caspase-3 and cleaved PARP, whereas control cells showed no signs of cell death. At 20nM cell death in both neuroblastoma cell lines was further increased but control cells were still non-responsive (Fig. 6c).

Discussion

Our study shows that a high frequency of neuroblastoma specific fusion transcripts could constitute an overlooked process through which altered transcripts are generated. It has been shown that upon cellular stress (e.g. viral infection, replicative or osmotic stress and mutational events) transcriptional termination can be blocked, increasing the probability of generating “downstream of genes”-transcripts (11). Thus, a proportion of fusions could represent passengers that occur as a response to cellular stress combined with neuroblastoma associated events such as gain or loss of chromosomal regions (e.g. 17q and 2p). In contrast to such passengers, certain fusions could be early events actually preceding and promoting other transformative events. Pharmacologic inhibition of splicing selectively repressed expression of several top frequent fusion transcripts but not of their wild-type cognates and there was a high frequency of splicing donor/acceptor sites in neuroblastoma specific fusions but not in fusions detected in normal adrenal gland, osteosarcoma or rhabdoid tumors. This pattern implies that a substantial proportion of the detected fusions are of the same *cis*-splicing type as previously reported in prostate cancer (5). Our analysis reveals that high expression levels of splicing associated factors is a distinguishing feature of high-risk neuroblastoma, representing a strong predictor of tumor grade. Regardless of the mechanisms underlying the generation of these fusions, they are not

287 necessarily dependent on amino acid changing mutations but can still provide a source of modified
288 gene products with the potential to promote neuroblastoma but also reveal novel drug targets. Given
289 the low frequency of recurrent mutations in neuroblastoma, such a pool of altered gene products could
290 indeed be relevant for tumor pathophysiology. A background of expressed fusion transcripts
291 potentially augments the effect of oncogenic drivers. Interestingly, our analysis shows that when
292 established drivers of neuroblastoma (e.g. *MYCN* and *LMO1*) are part of the fusion transcripts the
293 expression levels of wild type cognates are elevated. In addition, a panel of neuroblastoma specific
294 fusions occurring at high frequency could serve as biomarkers for diagnosis and the presence of risk
295 specific fusions could sub-divide neuroblastoma patients for precision therapy. Hence, fusions that are
296 passenger events rather than oncogenic drivers can still be of clinical relevance. One concern with
297 previously reported fusions is the relatively few cases that have been independently validated. In 2015,
298 only 3% of fusions identified by deep sequencing could be reproducibly detected (4). Arguably, the
299 "non-genomic" characteristic of this type of intrachromosomal fusions potentially augments the
300 detection of false positives. The risk that a proportion of fusion transcripts constitute false positives as
301 the result of spurious transcription or of sequencing errors is reduced by our crosswise analysis with
302 other tumors and healthy tissues. There is a clear enrichment of common fusions unique for the
303 neuroblastoma data sets that do not appear in any of the other tumors nor in normal adrenal glands.
304 It should however be noted that the normal adrenal gland is not perfect as control tissue due to the
305 cellular heterogeneity of the organ. Nevertheless, neuroblastoma specific fusions are enriched for
306 genes located at chromosomal regions (2p or 17q) which are commonly gained in high-risk
307 neuroblastoma. The selective loss of several fusion transcripts upon spliceosome inhibition is an
308 additional strong indication that fusion transcripts indeed are present. Our validation of ten fusions
309 through PCR and subsequent sequencing further underscores that these fusions can generate
310 alternative gene products. Interestingly, tumor specific distribution of fusions is reflected in the
311 osteosarcoma tumors where the *TP53* tumor suppressor is a fusion partner in an disproportional
312 amount of the detected fusions, mirroring the importance of inactivating mutations in *TP53* for this

disease (22). It has previously been reported that the presence of short homology sequences (SHS) can generate false RNA-chimeras due to template switching during the reverse transcriptase reaction (35). Such a mechanism would presumably generate random fusions between transcripts containing matching SHSs with no preference for any particular chromosomal location nor any preference for intrachromosomal vs interchromosomal fusion transcripts. The non-random enrichment of fusions at chromosomal locations that mirror the disease as well as the enrichment of intrachromosomal fusions between closely located genes suggests that reverse transcriptase induced template switching is not the cause of these fusions. Furthermore, it is plausible that fusions detected due to random template switching should correlate with high expression levels, which we do not detect. Our study shows that an altered protein with novel properties can be generated as a consequence of an intrachromosomal, splicing dependent fusion and that this altered protein influences the response to a drug (RA) commonly used to treat high-risk neuroblastoma patients. However, to fully evaluate the importance of identified fusions further functional experiments are required. A previous study indeed shows how the *SLC45A3-ELK4* read-through fusion transcript is elevated in prostate cancer tissue, is androgen-regulated and can be detected in an non-invasive assay from biopsies of men at risk of having prostate cancer (9). Even though it is possible that a portion of individual fusions are passenger events rather than oncogenic drivers a continued effort is justified to understand the regulation of fusions and downstream consequences of fusions in neuroblastoma as well as in other tumors.

Materials and Methods

Human tissue Samples. Neuroblastoma primary tumors came either from the Swedish NB Registry (ethical permission (DN03-736) granted by Karolinska Institutets Forskningsetikommitté Nord, (clinical information described in Li et. al. (36), or from the Irish NB cohort (described in Supplementary table 7), with ethical approval of the Medical and Research Ethics Committee of Our Lady's Children's

Hospital, Crumlin, Dublin, Ireland. Informed consent from families of subjects was obtained for samples. Six histologically confirmed normal human adrenal glands were included as controls (covered by existing ethical approvals; Dnr 01-136 and Dnr 01-353 KI forskningsetikkommitté Nord). Human total RNA from different normal tissues from Clontech (Human Total RNA Master Panel II,Cat#. 636643).

Paired-end RNA-Seq. RNA was isolated using the PerfectPure RNA Cultured Cell Kit (cell lines) and PerfectPure RNA Tissue Kit (patient samples) from 5 PRIME. RNA-seq libraries were prepared using TruSeq RNA Library Preparation Kit v2 (Illumina); paired-end RNA sequencing (125 bp) were performed in SciLifeLab (Stockholm).

Data Analysis. FusionCatcher (14) was applied to detect the fusion transcripts in paired-end RNA-seq data. Reads were mapped to hg19 and differential expression analysis was performed as described in(37). For enrichment analysis in Figure 1c, only fusions occurring above ~3% in each risk group were included (1 case in Low/Intermediate-risk, 4 cases in High-risk and 5 cases for the common fusion transcripts shared by Low/Intermediate-risk and high-risk groups).

Cloning and expression of wildtype *BAG2* and *ZNF451-BAG2*. cDNA was synthesized from total RNA by the iScript cDNA Synthesis Kit (Bio-Rad)**.** Coding regions of wildtype *BAG2* and *ZNF451-BAG2* were amplified from SK-N-AS cells for subcloning into p3XFLAG-CMV14 (Sigma) and pLVX-EF1α-IRES-mCherry (Clontech) using primer pair 1 and 2 respectively (Supplementary Table 11), inserted to pCR-Blunt II-TOPO vector via zero Blunt TOPO PCR Cloning Kit (ThermoFisher Scientific) and sequenced. *BAG2* and *ZNF451-BAG2* were subcloned into p3XFLAG-CMV14 vector using EcoRI and BamHI and pLVX-EF1α-IRES-mCherry lentiviral vector using *Eco*RI and *Bam*HI. To generate pLVX-Tetone-puro-

IRES-mCherry-empty/BAG2/ZNF451-BAG2 vectors, pLVX-Tetone-puro-empty (Clontech) construct was linearized with *Age*I, blunted and digested with *Eco*RI; pLVX-EF1α-IRES-mCherry-empty/BAG2/ZNF451-BAG2 vectors were linearized with *Mlu*I, blunted and digested with *Eco*RI to release IRES-mCherry-empty/BAG2/ZNF451-BAG2; then linearized and blunted pLVX-Tetone-puro-empty construct was ligated with fragment of IRES-mCherry-empty/BAG2/ZNF451-BAG2 separately.

Lentiviruses expressing pLVX-EF1α-IRES-mCherry-empty/BAG2/ZNF451-BAG2 and pLVX-Tetone-puro-IRES-mCherry-empty/BAG2/ZNF451-BAG2 were produced in 293FT cells using lipofectamine 2000 based protocol.

Cell culture. All neuroblastoma cell lines (SH-SY5Y, SK-N-SH, SK-N-FI, SK-N-BE(1), SK-N-AS, LAN-1, CHP-212 and IMR-5) were maintained in RPMI1640 medium supplemented with 10% FBS, 1% penicillin/streptomycin and 1% L-glutamine and grown in 5%CO_2 at 37°C.

Differentiation assay and immunofluorescence staining. For short-term RA induced differentiation assay, SK-N-FI cells were seeded in 6 well plates; after 24 hours, cells were infected by lentiviruses expressing pLVX-EF1α-IRES-mCherry-empty/BAG2/ZNF451-BAG2 for 48 hours; then cells were trypsinized and 20,000 cells were reseeded into 6 well plates with coverslips; 24h later, cells were treated with 1 μM retinoic acid (RA) or DMSO as control for 6 days. For doxycycline-inducible system, 20,000 SK-N-BE(1) cells with stable overexpression of pLVX-Tetone-puro-IRES-mCherry-ZNF451-BAG2 were seeded in 6 well plates with coverslips, and after 24h, cells were pre-treated with 1 μM doxycycline or DMSO as a control for one day; then cells received one of the following 4 different treatment for 4 days: DMSO, 1 μM RA, 1 μM doxycycline or 1 μM RA+1 μM doxycycline.

Cells on coverslips were fixed in RPMI1640 medium containing 2% PFA for 5 minutes at room temperature (RT), washed once with PBS at RT, fixed with 4% PFA for another 15 minutes at RT and

386 washed twice with cold PBS; then cells were permeabilized with PBS containing 0.25% Triton X-100 for
387 15 minutes at RT and blocked with 3% BSA for 1h at RT. Immunofluorescent stainings were performed
388 with the following primary antibody Tuj1 (Covance, 1:1000) overnight at +4 °C.

389

390 **Quantification.** Cells were stained as in Fig. 4g-l and o-r. Images were taken by confocal microscopy
391 (10X). For quantification, images were coded and a researcher who had not participated in staining
392 and image acquisition manually counted the ratio of transduced cells extending $TUJ1^+$ neurites/total
393 number of transduced cells/microscopic field. For DOX^- SK-N-BE(1) the ratio of cells extending $TUJ1^+$
394 neurites/total number of cells/microscopic field were counted. Microscopic fields containing less than
395 three transduced cells were discarded. Grubb´s test was performed for outlier detection (alpha
396 $p<0.05$). For statistical analysis in Fig. 4l and r one-way ANOVA followed by a Tukey's multiple
397 comparison test was performed.

398

399 **Western blotting**. Immunoblotting was performed using standard protocols. Following primary
400 antibodies were used: BAG2(sc-390262, 1:200), phospho-ser396-Tau(sc-101815, 1:1000), HSC70(sc-
401 7298, 1:200) from Santa Cruz; phosphor-ser404-Tau(44-758G, 1:200) from ThermoFisher Scientific;
402 total Tau(A0024, 1:1000) from Dako; beta-actin(AC-15, 1:3000), FLAG(F1804, 1:1000) from Sigma;
403 PARP (9542s), Caspase-3 (14220s) and cleaved caspase-3 (9664s) from Cell Signaling.

404

405 **Co-immunoprecipitation.** SK-N-FI cells were seeded in 15 cm dishes 24 hours before transfection.
406 Transfection of p3XFLAG-CMV14-empty/BAG2/ZNF451-BAG2 constructs was performed using
407 lipofectamine 2000. 48 hours post-transfection, cells were harvested and proteins were extracted
408 using lysis buffer containing 10 mM TRIS-HCl (pH7.4), 150 mM NaCl, 0.5% NP40, complete protease
409 and phosphatase inhibitors (Roche). FLAG-tagged fusion proteins were immunoprecipitated using

Anti-FLAG M2 magnetic beads (Sigma) according to manufacturer's instructions. Bound proteins were examined by silver staining or Western blotting using standard protocols.

Acknowledgements: We thank members of the Holmberg and Schlisio groups for valuable comments.

Funding: The Johan Holmberg lab is supported by the Swedish Children Cancer Foundation, the Swedish Cancer foundation, Knut and Alice Wallenberg Foundation, Swedish Research Council (VR), The Strategic Research Programme in Cancer (StratCan, SFO) and The Swedish Brain Foundation.

Author contributions: Y.S. and J.H. designed the study. Y.S. performed the majority of the experiments with help from V.R., E.M., S.L., P.B., J.Y., and I.W. Y.S. generated the libraries for RNA-sequencing. Y.S. and J.H performed the analysis with help from O.B.R. C.C.J, A.S., C.L., P.K. and M.J.S supplied the clinical material. Y.S. and J.H. wrote the manuscript with input from all authors. **Data and materials availability:** Data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Raw and processed NGS data will be deposited in the GEO database.

Declaration of interests

The authors declare no competing interests.

References

1. Pugh TJ*, et al.* (2013) The genetic landscape of high-risk neuroblastoma. *Nat Genet* 45(3):279-284.
2. Molenaar JJ*, et al.* (2012) Sequencing of neuroblastoma identifies chromothripsis and defects in neuritogenesis genes. *Nature* 483(7391):589-593.
3. Brodeur GM (2003) Neuroblastoma: biological insights into a clinical enigma. *Nat Rev Cancer* 3(3):203-216.
4. Mertens F, Johansson B, Fioretos T, & Mitelman F (2015) The emerging complexity of gene fusions in cancer. *Nat Rev Cancer* 15(6):371-381.
5. Qin F*, et al.* (2015) Discovery of CTCF-sensitive Cis-spliced fusion RNAs between adjacent genes in human prostate cells. *PLoS Genet* 11(2):e1005001.
6. Zhang Y*, et al.* (2012) Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. *Cancer discovery* 2(7):598-607.

7. Santo EE, *et al.* (2012) Oncogenic activation of FOXR1 by 11q23 intrachromosomal deletion-fusions in neuroblastoma. *Oncogene* 31(12):1571-1581.
8. Jia Y, Xie Z, & Li H (2016) Intergenically Spliced Chimeric RNAs in Cancer. *Trends Cancer* 2(9):475-484.
9. Rickman DS, *et al.* (2009) SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. *Cancer Res* 69(7):2734-2738.
10. Maher CA, *et al.* (2009) Transcriptome sequencing to detect gene fusions in cancer. *Nature* 458(7234):97-101.
11. Proudfoot NJ (2016) Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* 352(6291):aad9926.
12. Babiceanu M, *et al.* (2016) Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res* 44(6):2859-2872.
13. Matthay KK, *et al.* (2009) Long-term results for children with high-risk neuroblastoma treated on a randomized trial of myeloablative therapy followed by 13-cis-retinoic acid: a children's oncology group study. *J Clin Oncol* 27(7):1007-1013.
14. D. Nicorici MS, H. Edgren, S. Kangaspeska, A. Murumagi, O. Kallioniemi, S. Virtanen, O. Kilkku, (2014) FusionCatcher – a tool for finding somatic fusion genes in paired-end RNA-sequencing data. *bioRxiv*.
15. Xie B, *et al.* (2016) Two RNAs or DNAs May Artificially Fuse Together at a Short Homologous Sequence (SHS) during Reverse Transcription or Polymerase Chain Reactions, and Thus Reporting an SHS-Containing Chimeric RNA Requires Extra Caution. *PLoS One* 11(5):e0154855.
16. Khan FH, *et al.* (2015) Acquired genetic alterations in tumor cells dictate the development of high-risk neuroblastoma and clinical outcomes. *BMC Cancer* 15:514.
17. Hungate EA, *et al.* (2017) Evaluation of Genetic Predisposition for MYCN-Amplified Neuroblastoma. *J Natl Cancer Inst* 109(10).
18. Lee SH, *et al.* (2017) ARID1B alterations identify aggressive tumors in neuroblastoma. *Oncotarget* 8(28):45943-45950.
19. Consortium GT (2013) The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 45(6):580-585.
20. Kim SY & Volsky DJ (2005) PAGE: parametric analysis of gene set enrichment. *BMC Bioinformatics* 6:144.
21. Koster J (2008) 'R2: Genomics Analysis and Visualization Platform (http://r2.amc.nl)'.
22. Martin JW, Squire JA, & Zielenska M (2012) The genetics of osteosarcoma. *Sarcoma* 2012:627254.
23. Carrettiero DC, Hernandez I, Neveu P, Papagiannakopoulos T, & Kosik KS (2009) The Cochaperone BAG2 Sweeps Paired Helical Filament-Insoluble Tau from the Microtubule. *Journal of Neuroscience* 29(7):2151-2161.
24. Santiago FE, Almeida MC, & Carrettiero DC (2015) BAG2 Is Repressed by NF-kappaB Signaling, and Its Overexpression Is Sufficient to Shift Abeta1-42 from Neurotrophic to Neurotoxic in Undifferentiated SH-SY5Y Neuroblastoma. *Journal of molecular neuroscience : MN* 57(1):83-89.
25. Subramanian A, *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102(43):15545-15550.
26. Qin L, Guo J, Zheng Q, & Zhang H (2016) BAG2 structure, function and involvement in disease. *Cellular & molecular biology letters* 21:18.
27. Shimura H, Schwartz D, Gygi SP, & Kosik KS (2004) CHIP-Hsc70 complex ubiquitinates phosphorylated tau and enhances cell survival. *J Biol Chem* 279(6):4869-4876.
28. DeGeer J, *et al.* (2015) Hsc70 chaperone activity underlies Trio GEF function in axon growth and guidance induced by netrin-1. *J Cell Biol* 210(5):817-832.
29. Wang BD & Lee NH (2018) Aberrant RNA Splicing in Cancer and Drug Resistance. *Cancers (Basel)* 10(11).

30. Hsu TY, *et al.* (2015) The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* 525(7569):384-388.
31. Laetsch TW, *et al.* (2014) Multiple components of the spliceosome regulate Mcl1 activity in neuroblastoma. *Cell Death Dis* 5:e1072.
32. Eleveld TF, *et al.* (2015) Relapsed neuroblastomas show frequent RAS-MAPK pathway mutations. *Nat Genet* 47(8):864-871.
33. Sausen M, *et al.* (2013) Integrated genomic analyses identify ARID1A and ARID1B alterations in the childhood cancer neuroblastoma. *Nat Genet* 45(1):12-17.
34. Mizui Y, *et al.* (2004) Pladienolides, new substances from culture of Streptomyces platensis Mer-11107. III. In vitro and in vivo antitumor activities. *J Antibiot (Tokyo)* 57(3):188-196.
35. Akiva P, *et al.* (2006) Transcription-mediated gene fusion in the human genome. *Genome Res* 16(1):30-36.
36. Li S, *et al.* (2016) The 1p36 Tumor Suppressor KIF 1Bbeta Is Required for Calcineurin Activation, Controlling Mitochondrial Fission and Apoptosis. *Dev Cell* 36(2):164-178.
37. Westerlund I, *et al.* (2017) Combined epigenetic and differentiation-based treatment inhibits neuroblastoma tumor growth and links HIF2alpha to tumor suppression. *Proc Natl Acad Sci U S A* 114(30):E6137-E6146.

Figure legends

Figure 1. Identification of fusion transcripts in neuroblastoma tumors.

(a) Identification of the 25 most frequent fusions by FusionCatcher in a cohort of 172 paired end RNA sequenced neuroblastoma patient samples derived from the NCI TARGET project (NB172).

(b) Circos plot of genomic distribution of top frequent fusions (> 10%) in NB172.

(c) Enrichment of fusion transcripts common or unique to low/intermediate-risk or high-risk tumors in chromosomal arms as calculated by a normalized enrichment score = (counts of fusion transcripts in each chromosomal arm / length of chromosomal arm (Mb)) / average enrichment.

(d-e) *MYCN/LMO1* expression levels are high in neuroblastoma tumors (red) bearing fusion transcripts of which one fusion partner is either *MYCN* or *LMO1,* as shown by expression value FPKM (Fragments Per Kilobase per Million mapped reads).

P-values in (b-c) were calculated from Z-values, assuming standard normal distribution. *p<0.05, **p<0.01, ***p<0.001.

Figure 2. Validation of fusion transcripts.

(a) Venn diagram of identified fusions in three datasets, NB172 (NCI TARGET), Validation-NB (NB-v, 14 neuroblastoma patients plus 8 neuroblastoma cell lines) and adrenal gland (161 samples from normal human adrenal gland).

(b) Enrichment of fusion transcript genes unique to the NB172 and the NB-v cohorts on chromosomal arms 2p and 17q.

(c) Parametric analysis of gene set enrichment (PAGE) of the 498-SEQC neuroblastoma dataset show an enrichment of NB172 and NB-v unique common fusions in high-risk grade 4 tumors compared to low-risk 4S tumors.

(d) Venn diagram of identified fusions in three datasets, NB172 (NCI TARGET), Rhabdoid (65 Rhabdoid tumor patient samples) and Osteosarcoma (177 Osteosarcoma patient samples).

(e) Comparison of the number of neuroblastoma (NB) and osteosarcoma (OS) tumors with at least one fusion transcript containing *TP53*.

(f) Validation of fusion transcripts by RT-PCR and sequencing in Validation-NB neuroblastoma patients, indicated normal tissues were used as controls; DRG, human dorsal root ganglion.

P-values in (a and b) were calculated by chi-square test with Yate's correction.

Figure 3. Differentially expressed genes in *ZNF458-BAG2* containing neuroblastoma predicts clinical outcome and tumor risk.

(a) Number of differentially expressed genes (DEGs) in RT-PCR validated neuroblastoma samples with or without *ZNF451-BAG2* fusion as shown in Fig. 2e.

(b) Distribution of DEGs from (a) in 498 sequenced neuroblastoma (498-SEQC) according to high-risk vs low-risk disease. Red dots designate genes up in (a) and green dots designates genes down in (a). The two most significant genes in high-risk (*ENOSF1*) vs low-risk (*RAMP3*) disease are indicated.

(c) Overall survival probability according to *ENOSF1* expression.

(d) Overall survival probability according to *RAMP3* expression.

(e) K-means analysis of the 498-SEQC dataset, utilizing the DEGs in tumors harboring the *ZNF451-BAG2* fusion (a), generates two clusters with significant differences in Risk, INSS and *MYCN* amplification.

(f) Overall survival probability of the two clusters identified in (e).

(g-h) Gene set enrichment analysis (GSEA) showing enrichment of genes in the cell cycle associated gene set "HALLMARK_G2M_CHECKPOINT" (g) and depletion of genes in the "HALLMARK_APOPTOSIS" gene set in ΔBAG2 containing neuroblastoma.

Figure 4. ΔBAG2 impairs the clearance of phosphorylated TAU and binding to HSC70 and inhibits retinoic acid (RA) induced differentiation in neuroblastoma cells.

(a) Schematic representation of *ZNF451-BAG2* fusion, the resulting truncated BAG2 is referred to as ΔBAG2.

(b) Endogenous tumor ΔBAG2 protein expression is associated with high levels of phosphorylated TAU (p-TAU) in neuroblastoma tumors as detected by immunoblotting.

(c) SKNFI cells were transfected with p3XFLAG-CMV14-empty/*BAG2*/*ΔBAG2*/*BAG2*+*ΔBAG2* for 48 hours. Cells were harvested and proteins were extracted; whole-cell lysates were used to detect BAG2, FLAG, p-TAU on Ser404, total TAU and ACTIN.

(d) FLAG-tagged proteins were immunoprecipitated from whole-cell lysates as prepared in (c) using Anti-FLAG M2 magnetic beads and eluted. Immunoprecipitated proteins were western blotting to detect HSC70, BAG2.

(e-l) Constitutive lentiviral overexpression of *ΔBAG2*, but not wildtype *BAG2* (backbone pLVX-EF1α-IRES-mCherry) inhibited RA-induced differentiation (6 days of treatment) in SK-N-FI cells.

(m-r) Doxycycline inducible lentiviral overexpression of *ΔBAG2* (backbone pLVX-Tet-one-puro-IRES-mCherry) inhibited RA-induced differentiation (4 days of treatment) in SK-N-BE(1) cells. Protein levels of BAG2 and ΔBAG2 were analyzed by Western blotting (e and m). Immunostaining was performed using antibody against neuronal marker β3-tubulin (TUJ1).

Data in l and r is represented as mean of transduced cells with $TUJ1^+$ neurites/total number of transduced cells +/- SEM, each data-point represents this ratio in a single 10x microscopic field (n=10-20). ***p<0.001, one-way ANOVA with Tukey's multiple comparison test.

Figure 5. Enrichment of canonical splicing pattern at fusion junctions is associated with aberrant spliceosome activity in high-risk neuroblastoma.

(a) Plot of the frequency versus the distance from 5' to 3' in the intrachromosomal fusion transcript identified in the NB discovery cohort (NB172); each dot represents a unique fusion transcript; fusion transcripts re-identified in the NB validation cohort (NB-v) were marked as red.

(b) Distribution of unique intrachromosomal vs interchromosomal fusion junctions flanked by GT_AG or other nucleotide motifs in neuroblastoma, rhabdoid tumor, osteosarcoma and normal adrenal gland.

(c) Differential expression of genes in the KEGG Spliceosome pathway between neuroblastoma dataset (NBL172) versus human normal adrenal gland dataset.

(d) K-means analysis of the 498-SEQC dataset, utilizing the genes in the KEGG spliceosome pathway, generates two clusters with significant differences in Risk, INSS and *MYCN*-amplification.

(e) Overall survival probability of the two clusters identified in (d).

Figure 6. Pharmaceutical inhibition of spliceosome activity reduces generation of fusion transcripts and induction of apoptosis in neuroblastoma cells.

595 (a-b) Splicing-dependent generation of several high frequent fusion transcripts in LAN-1 (a) and SK-N-
596 BE(1) (b) neuroblastoma cells. SK-N-BE(1) and LAN1 cells were treated with splicing inhibitor
597 Pladienolide B at 100 nM (SK-N-BE(1)) and 5-50 nM (LAN-1) for 6 hours; RNA was isolated from
598 harvested cells and reversely transcribed to cDNA; RT-PCR were performed with primers spanning the
599 fusion junctions (fusion transcript) or exon-exon boundary (wild-type cognate).

600 (c) Induction of apoptosis in LAN-1 and SK-N-BE(1) cells, but not in human normal diploid fibroblasts
601 (HDNF) upon pladienolide B treatment (0-20 nM) for 48 hours, as detected by western blotting of full-
602 length CASPASE-3, cleaved CASPASE-3, full-length PARP and cleaved PARP; β-ACTIN was used as a
603 loading control.

604

Figure 1

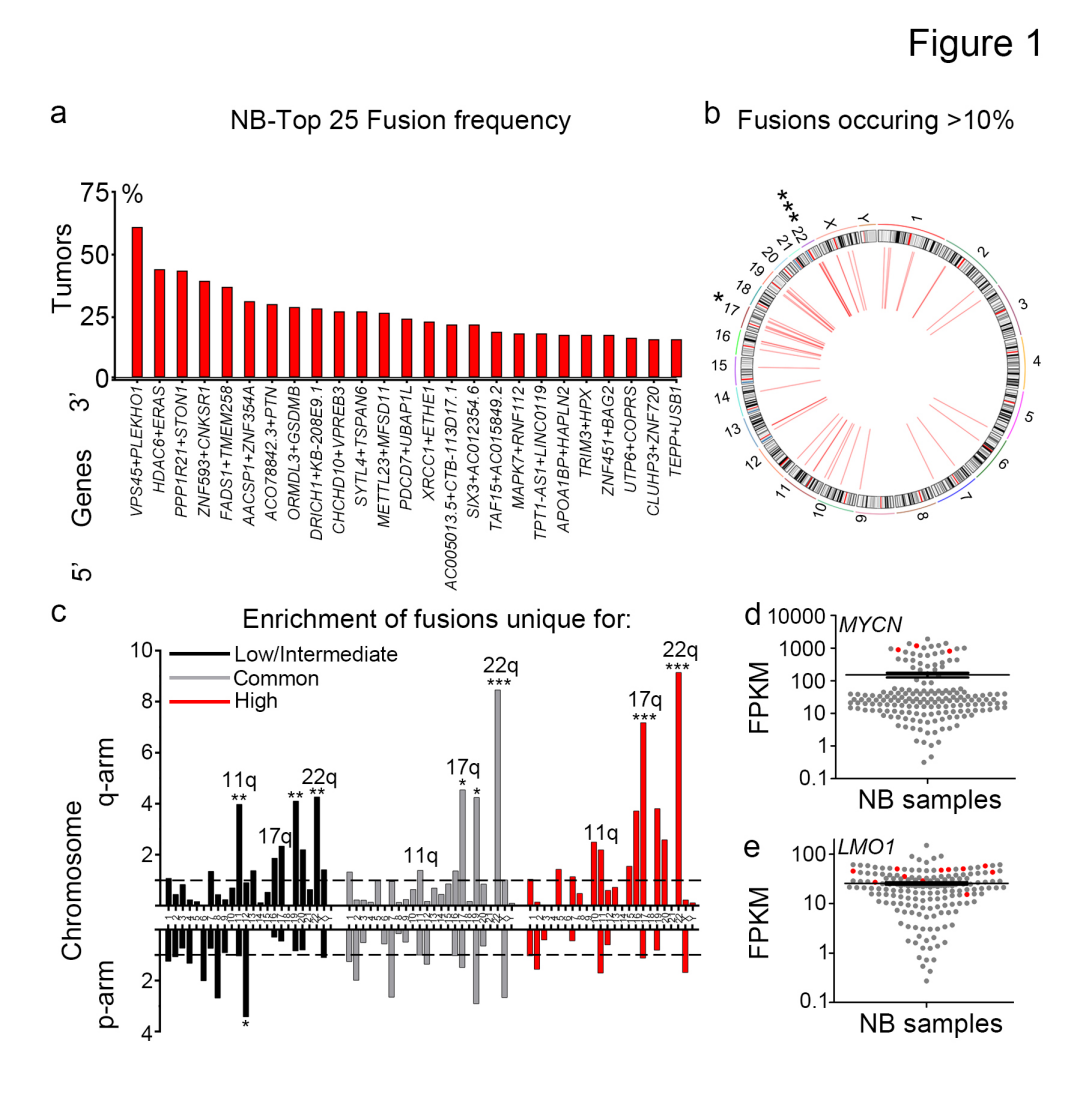

a
NB-Top 25 Fusion frequency
75 %
50
25
0
Tumors
3'
Genes
5'
VPS45+PLEKHO1
HDAC6+ERAS
PPP1R21+STON1
ZNF593+CNKSR1
FADS1+TMEM258
AACSP1+ZNF354A
ACO78842.3+PTN
ORMDL3+GSDMB
DRICH1+KB-208E9.1
CHCHD10+VPREB3
SYTL4+TSPAN6
METTL23+MFSD11
PDCD7+UBAP1L
XRCC1+ETHE1
AC005013.5+CTB-113D17.1
SIX3+AC012354.6
TAF15+AC015849.2
MAPK7+RNF112
TPT1-AS1+LINC0119
APOA1BP+HAPLN2
TRIM3+HPX
ZNF451+BAG2
UTP6+COPRS
CLUHP3+ZNF720
TEPP+USB1
b
Fusions occuring >10%
c
Enrichment of fusions unique for:
Low/Intermediate
Common
High
Chromosome
q-arm
p-arm
11q
17q
22q
d
MYCN
FPKM
NB samples
e
LMO1
FPKM
NB samples

Figure 2

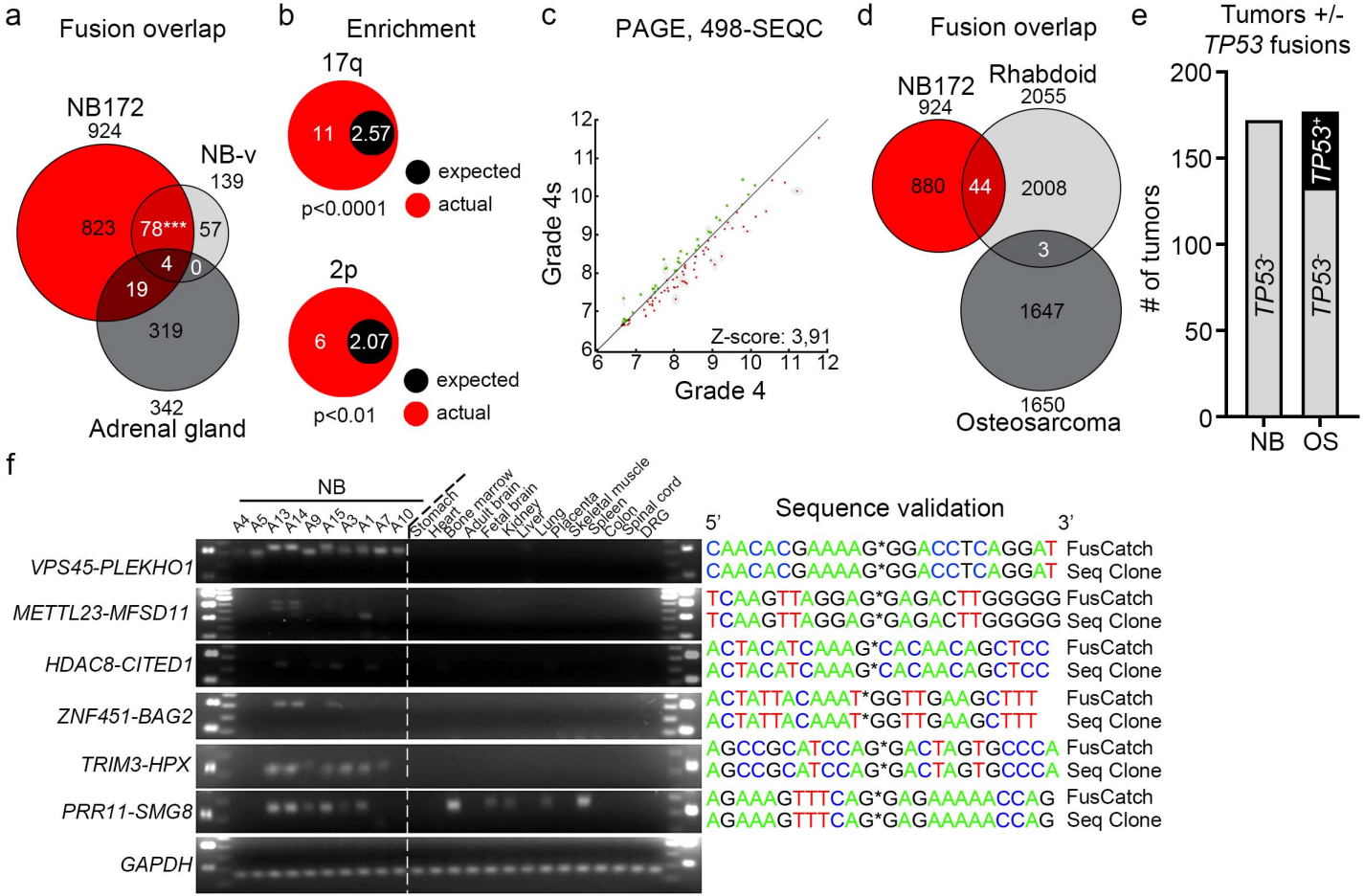

a
Fusion overlap
NB172
924
NB-v
139
823
78***
57
4
0
19
319
342
Adrenal gland
b
Enrichment
17q
11
2.57
expected
actual
p<0.0001
2p
6
2.07
expected
actual
p<0.01
c
PAGE, 498-SEQC
Grade 4s
Grade 4
Z-score: 3,91
d
Fusion overlap
NB172
924
Rhabdoid
2055
880
44
2008
3
1647
1650
Osteosarcoma
e
Tumors +/-
TP53 fusions
# of tumors
TP53-
TP53+
TP53-
NB
OS
f
NB
Stomach
Heart
Bone marrow
Adult brain
Fetal brain
Kidney
Liver
Lung
Placenta
Skeletal muscle
Spleen
Colon
Spinal cord
DRG
A4
A5
A13
A14
A9
A15
A3
A1
A7
A10
Sequence validation
5'
3'
VPS45-PLEKHO1
CAACACGAAAAG*GGACCTCAGGAT FusCatch
CAACACGAAAAG*GGACCTCAGGAT Seq Clone
METTL23-MFSD11
TCAAGTTAGGAG*GAGACTTGGGGG FusCatch
TCAAGTTAGGAG*GAGACTTGGGGG Seq Clone
HDAC8-CITED1
ACTACATCAAAG*CACAACAGCTCC FusCatch
ACTACATCAAAG*CACAACAGCTCC Seq Clone
ZNF451-BAG2
ACTATTACAAAT*GGTTGAAGCTTT FusCatch
ACTATTACAAAT*GGTTGAAGCTTT Seq Clone
TRIM3-HPX
AGCCGCATCCAG*GACTAGTGCCCA FusCatch
AGCCGCATCCAG*GACTAGTGCCCA Seq Clone
PRR11-SMG8
AGAAAGTTTCAG*GAGAAAAACCAG FusCatch
AGAAAGTTTCAG*GAGAAAAACCAG Seq Clone
GAPDH

Figure 3

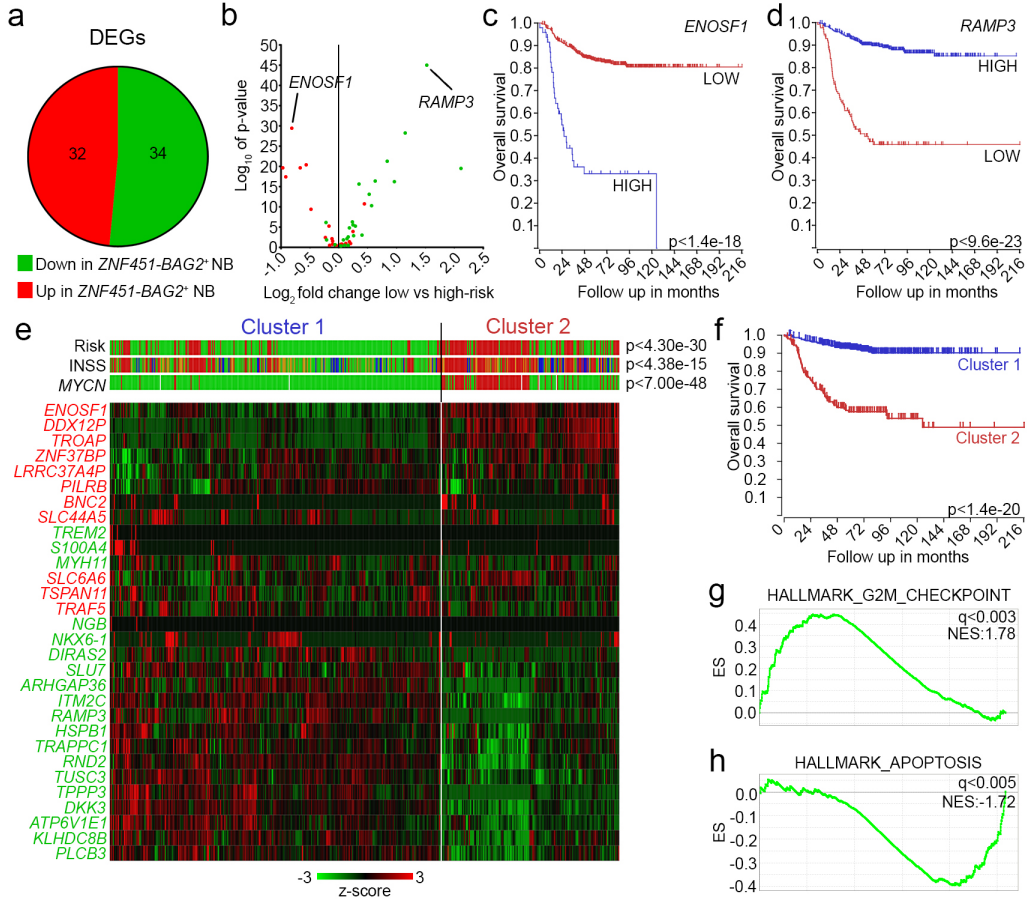

a
DEGs
32
34
Down in ZNF451-BAG2+ NB
Up in ZNF451-BAG2+ NB
b
Log10 of p-value
ENOSF1
RAMP3
Log2 fold change low vs high-risk
c
ENOSF1
LOW
HIGH
Overall survival
p<1.4e-18
Follow up in months
d
RAMP3
HIGH
LOW
Overall survival
p<9.6e-23
Follow up in months
e
Cluster 1
Cluster 2
Risk
INSS
MYCN
p<4.30e-30
p<4.38e-15
p<7.00e-48
ENOSF1
DDX12P
TROAP
ZNF37BP
LRRC37A4P
PILRB
BNC2
SLC44A5
TREM2
S100A4
MYH11
SLC6A6
TSPAN11
TRAF5
NGB
NKX6-1
DIRAS2
SLU7
ARHGAP36
ITM2C
RAMP3
HSPB1
TRAPPC1
RND2
TUSC3
TPPP3
DKK3
ATP6V1E1
KLHDC8B
PLCB3
-3
3
z-score
f
Cluster 1
Cluster 2
Overall survival
p<1.4e-20
Follow up in months
g
HALLMARK_G2M_CHECKPOINT
q<0.003
NES:1.78
ES
h
HALLMARK_APOPTOSIS
q<0.005
NES:-1.72
ES

Figure 4

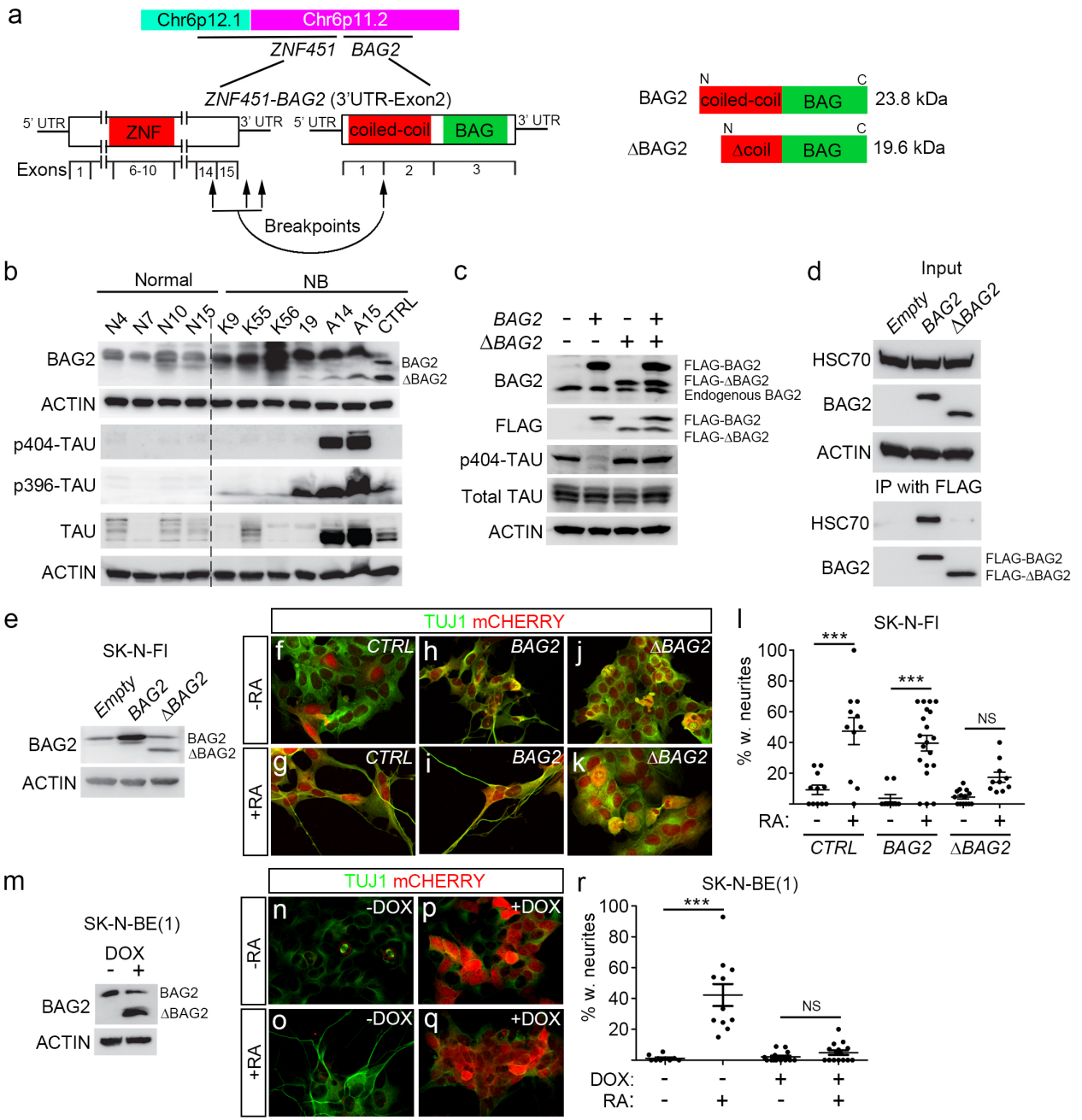

a
Chr6p12.1
Chr6p11.2
ZNF451
BAG2
ZNF451-BAG2 (3'UTR-Exon2)
5' UTR
ZNF
3' UTR
5' UTR
coiled-coil
BAG
3' UTR
Exons
1
6-10
14
15
1
2
3
Breakpoints
BAG2
N
C
coiled-coil
BAG
23.8 kDa
ΔBAG2
N
C
Δcoil
BAG
19.6 kDa
b
Normal
NB
N4
N7
N10
N15
K9
K55
K56
19
A14
A15
CTRL
BAG2
ACTIN
p404-TAU
p396-TAU
TAU
ACTIN
BAG2
ΔBAG2
c
BAG2
ΔBAG2
BAG2
FLAG
p404-TAU
Total TAU
ACTIN
FLAG-BAG2
FLAG-ΔBAG2
Endogenous BAG2
FLAG-BAG2
FLAG-ΔBAG2
d
Input
Empty
BAG2
ΔBAG2
HSC70
BAG2
ACTIN
IP with FLAG
HSC70
BAG2
FLAG-BAG2
FLAG-ΔBAG2
e
SK-N-FI
Empty
BAG2
ΔBAG2
BAG2
ACTIN
BAG2
ΔBAG2
TUJ1 mCHERRY
-RA
+RA
f
CTRL
h
BAG2
j
ΔBAG2
g
CTRL
i
BAG2
k
ΔBAG2
l
SK-N-FI
% w. neurites
100
80
60
40
20
***
***
NS
RA:
-
+
-
+
-
+
CTRL
BAG2
ΔBAG2
m
SK-N-BE(1)
DOX
-
+
BAG2
ACTIN
BAG2
ΔBAG2
TUJ1 mCHERRY
-RA
+RA
n
-DOX
p
+DOX
o
-DOX
q
+DOX
r
SK-N-BE(1)
% w. neurites
100
80
60
40
20
***
NS
DOX:
-
-
+
+
RA:
-
+
-
+

Figure 5

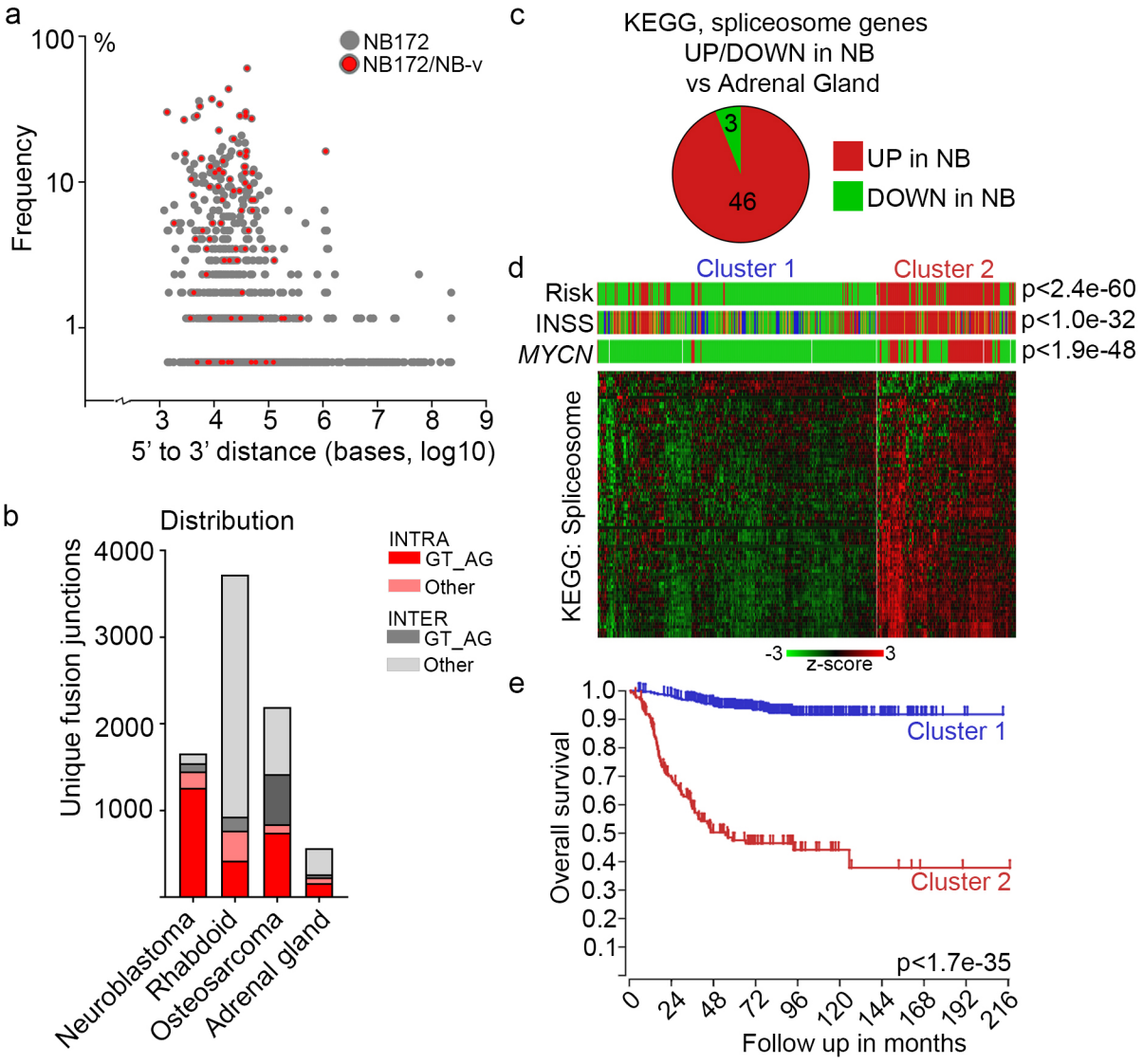
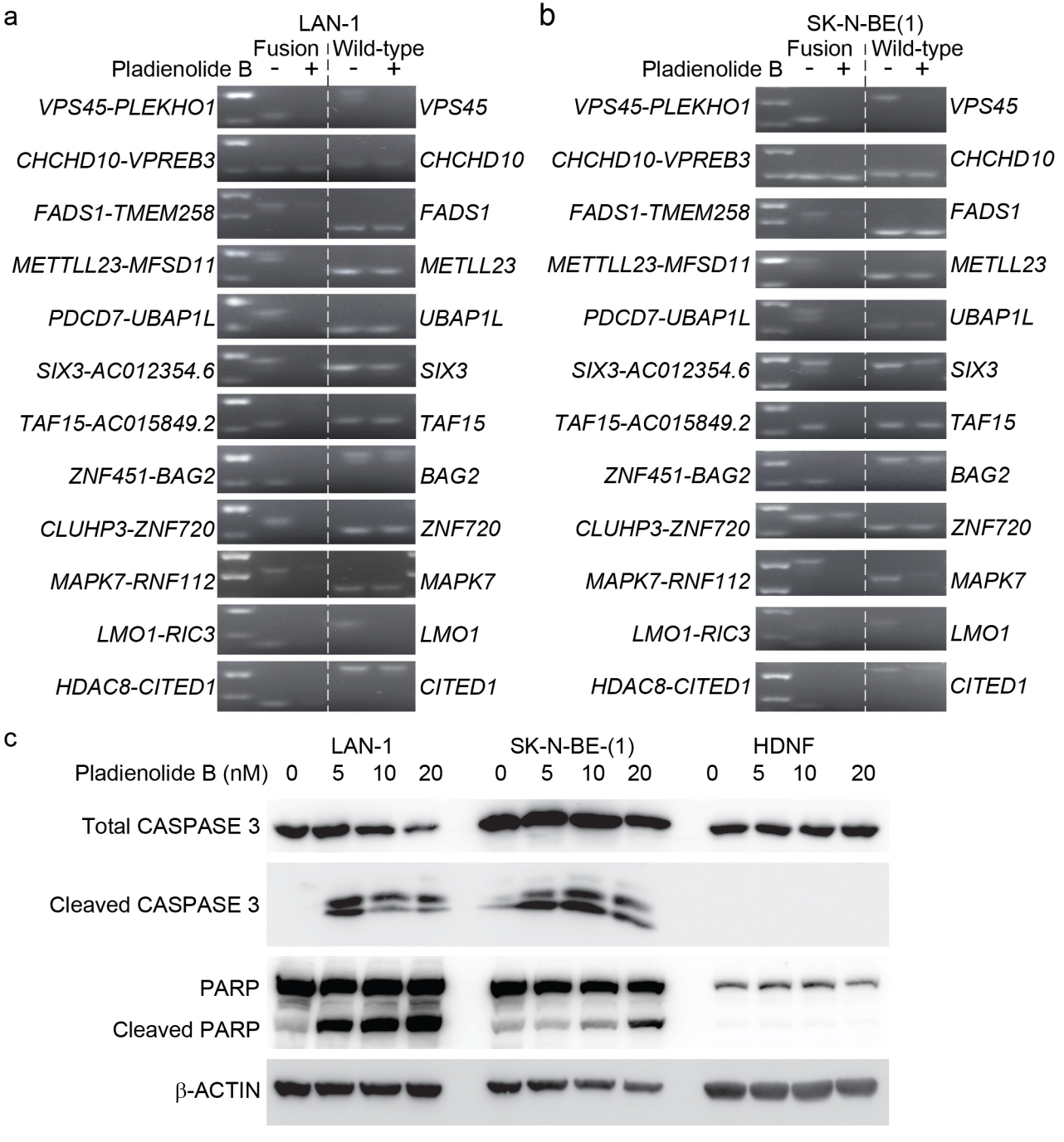

a
100
%
NB172
NB172/NB-v
Frequency
10
1
3
4
5
6
7
8
9
5' to 3' distance (bases, log10)
b
Distribution
Unique fusion junctions
4000
3000
2000
1000
INTRA
GT_AG
Other
INTER
GT_AG
Other
Neuroblastoma
Rhabdoid
Osteosarcoma
Adrenal gland
c
KEGG, spliceosome genes
UP/DOWN in NB
vs Adrenal Gland
3
46
UP in NB
DOWN in NB
d
Cluster 1
Cluster 2
Risk
INSS
MYCN
p<2.4e-60
p<1.0e-32
p<1.9e-48
KEGG: Spliceosome
-3
z-score
3
e
Overall survival
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Cluster 1
Cluster 2
p<1.7e-35
0
24
48
72
96
120
144
168
192
216
Follow up in months

Figure 6


a
LAN-1
Fusion
Wild-type
Pladienolide B - + - +
VPS45-PLEKHO1
VPS45
CHCHD10-VPREB3
CHCHD10
FADS1-TMEM258
FADS1
METTLL23-MFSD11
METTLL23
PDCD7-UBAP1L
UBAP1L
SIX3-AC012354.6
SIX3
TAF15-AC015849.2
TAF15
ZNF451-BAG2
BAG2
CLUHP3-ZNF720
ZNF720
MAPK7-RNF112
MAPK7
LMO1-RIC3
LMO1
HDAC8-CITED1
CITED1
b
SK-N-BE(1)
Fusion
Wild-type
Pladienolide B - + - +
VPS45-PLEKHO1
VPS45
CHCHD10-VPREB3
CHCHD10
FADS1-TMEM258
FADS1
METTLL23-MFSD11
METTLL23
PDCD7-UBAP1L
UBAP1L
SIX3-AC012354.6
SIX3
TAF15-AC015849.2
TAF15
ZNF451-BAG2
BAG2
CLUHP3-ZNF720
ZNF720
MAPK7-RNF112
MAPK7
LMO1-RIC3
LMO1
HDAC8-CITED1
CITED1
c
LAN-1
SK-N-BE-(1)
HDNF
Pladienolide B (nM) 0 5 10 20 0 5 10 20 0 5 10 20
Total CASPASE 3
Cleaved CASPASE 3
PARP
Cleaved PARP
β-ACTIN