




Review

Recurrent Spliceosome Mutations in Cancer: Mechanisms and Consequences of Aberrant Splice Site Selection

Carlos A. Niño ^{1,*} , Rossella Scotto di Perrotolo ¹  and Simona Polo ^{1,2,*} ¹ Fondazione Istituto FIRC di Oncologia Molecolare (IFOM), 20139 Milan, Italy; rossella.scotto@ifom.eu² Dipartimento di Oncologia ed Emato-Oncologia, Università degli Studi di Milano, 20122 Milan, Italy

* Correspondence: carlos.nino@ifom.eu (C.A.N.); simona.polo@ifom.eu (S.P.)

Simple Summary: The spliceosome ribonucleoprotein complex catalyzes the removal of introns and exons ligation, a fundamental post-transcriptional process that generates mature RNAs. Cancer-associated mutations in spliceosome components give rise to aberrant splice site selection and, therefore, the production of novel isoform variants that support tumorigenesis. In this review, we summarize the current research regarding cancer hotspot mutations identified in spliceosome components acting at the very first step of splicing, namely the U1 snRNA, SF3B1, and U2AF1.

Abstract: Splicing alterations have been widely documented in tumors where the proliferation and dissemination of cancer cells is supported by the expression of aberrant isoform variants. Splicing is catalyzed by the spliceosome, a ribonucleoprotein complex that orchestrates the complex process of intron removal and exon ligation. In recent years, recurrent hotspot mutations in the spliceosome components U1 snRNA, SF3B1, and U2AF1 have been identified across different tumor types. Such mutations in principle are highly detrimental for cells as all three spliceosome components are crucial for accurate splice site selection: the U1 snRNA is essential for 5' splice site recognition, and SF3B1 and U2AF1 are important for 3' splice site selection. Nonetheless, they appear to be selected to promote specific types of cancers. Here, we review the current molecular understanding of these mutations in cancer, focusing on how they influence splice site selection and impact on cancer development.

Keywords: splicing; cancer; spliceosome; U1 snRNA; SF3B1; U2AF1



Citation: Niño, C.A.; Scotto di Perrotolo, R.; Polo, S. Recurrent Spliceosome Mutations in Cancer: Mechanisms and Consequences of Aberrant Splice Site Selection. *Cancers* **2022**, *14*, 281. <https://doi.org/10.3390/cancers14020281>

Academic Editors: Maria Paola Paronetto and Claudia Ghigna

Received: 30 November 2021

Accepted: 4 January 2022

Published: 7 January 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Splicing is a fundamental process in gene expression regulation and is therefore essential for maintaining accurate cellular homeostasis, fitness, and fate. Splicing of precursor mRNAs (pre-mRNAs) into mature mRNAs is a two-step reaction carried out by the spliceosome, which catalyzes accurate intron removal and exon ligation [1]. The spliceosome is a ribonucleoprotein (RNP) complex composed of the small nuclear RNP subunits snRNP U1, snRNP U2, snRNP U4/U6, and snRNP U5, which associate with different splicing co-factors [1,2]. These snRNP subunits are generated through the association of the small nuclear RNAs (snRNAs) U1, U2, U4/U6, and U5 with Sm or Sm-like (LSm) proteins [2,3].

The spliceosome does not possess a preformed active site is assembled de novo directly onto each intron [1,2]. Intron removal relies on three pre-mRNA consensus sequences that are recognized by different spliceosome components: (i) the 5' and 3' splice sites, located at the 5' and 3' ends of the intron, (ii) the branch point sequence (BPS), located between 18 and 40 nucleotides upstream of the 3' splice site, and (iii) the polypyrimidine tract (PPT), located between the BPS and the 3' splice site (Figure 1A). Spliceosome proteins are essential for the correct recognition and base-pairing between spliceosome snRNAs and the splice sites and, as such, represent an important layer of fine regulation [1,3].

During the splicing cycle, the spliceosome undergoes a series of conformational changes that are mainly represented by base pairings occurring between pre-mRNA

and snRNAs [1,3], supported by DExD/H-type RNA-dependent ATPases/helicases [2] (Figure 1B).

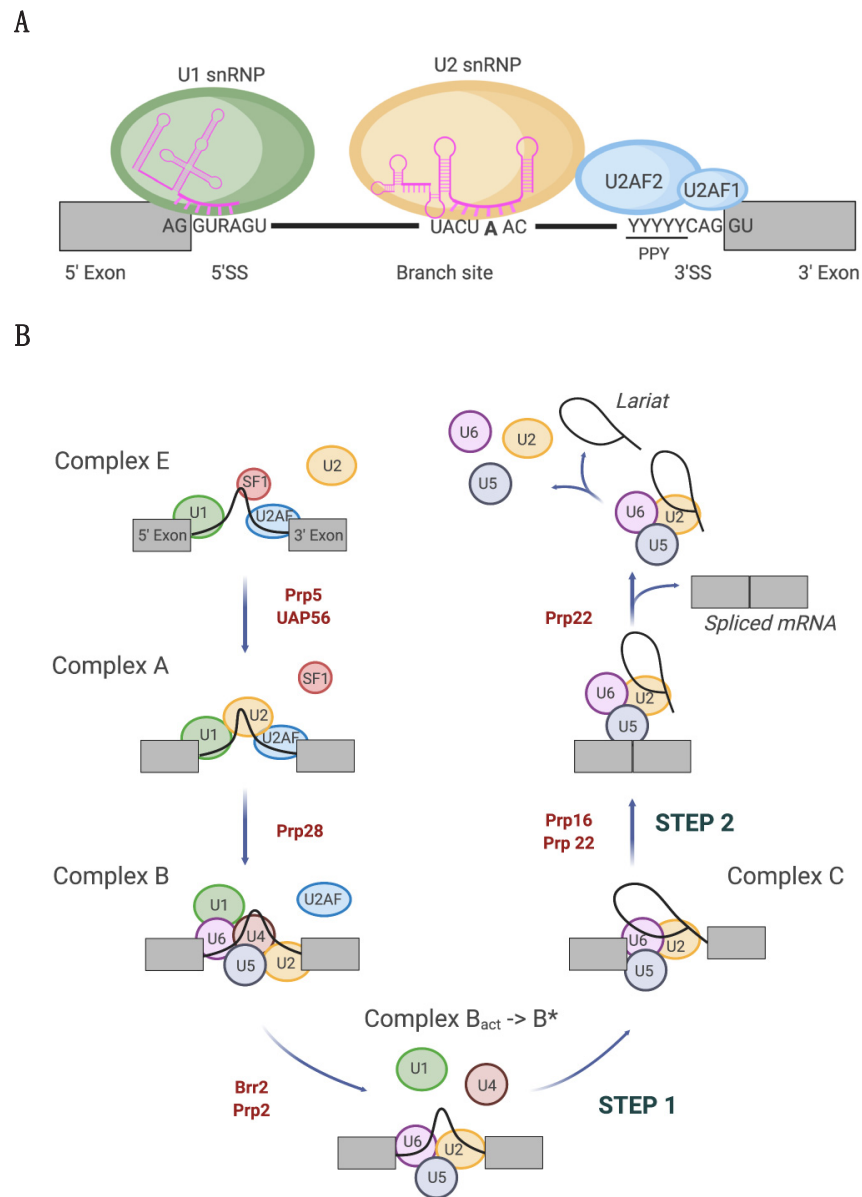


Figure 1. The splicing cycle. (A) Splicing complex assembly is initiated by the recognition of consensus sequence elements located at the exon (grey boxes)/intron (black line) boundaries. U1 snRNP and U2AF recognize the 5' (5' ss) and 3' (3' ss) splice sites, respectively, inducing U2 snRNP recruitment at the branch point. (B) Pre-mRNA splicing stepwise reaction performed by the spliceosome is depicted. U1, U2, U4, U5, and U6 subunits are composed of snRNAs and Sm/LSm proteins. Details of the process are given in the main text. Rearrangements and remodeling are assisted by ATPases/helicases (indicated in red) to allow the progression through complexes E, A, B, B*, and C.

The splicing reaction begins with the base-pair-mediated interaction between the U1 snRNA and the 5' splice site, forming the so-called E (early)-complex (Figure 1B). At the 3' splice site, direct RNA–protein interactions are established between splicing factor 1 (SF1) and the U2 auxiliary factor heterodimer (U2AF) with the BPS and PPT, respectively [4–6]. Following SF1 displacement from the BPS, the U2 snRNA pairs with the BPS, forming the pre-spliceosome or complex A (Figure 1B). U2 snRNP engagement is assisted by a direct interaction between U2 snRNP proteins and the BPS and BPS flanking regions. In particular, direct interactions are established between the U2 snRNP components p14

and SF3B1 with the adenine at the BPS and the U2AF heterodimer, respectively [7,8]. In the complex A, the interaction between U1 and U2 snRNPs brings the two splice sites close to each other, followed by the recruitment of the preassembled U4-U6-U5 tri-snRNP, forming the pre-catalytic spliceosome or complex B. This latter undergoes conformational and compositional rearrangements with the displacement of U1 and U4 snRNPs, thus determining the transition toward the complex B^{act} (activated) (Figure 1B). This complex is further remodeled into the catalytically active complex B*, which catalyzes the first transesterification reaction, generating complex C. After U2 snRNP conformational change and repositioning of the reaction intermediate, complex C catalyzes the second step of the splicing reaction. The post-spliceosome complex then dissociates, thus releasing the spliced mRNA product, the excised intron bound to U2-U5-U6, called lariat (Figure 1B), and the snRNPs that are recycled in another splicing reaction [1,2].

Aberrant splicing is commonly observed in cancer [9–11]. Indeed, cancer-associated alterations generate splice variants involved in all of the different hallmarks of cancer, including cancer cell survival, proliferation, metastasis, angiogenesis, and chemoresistance [12–14]. Aberrant splicing can arise from alterations in the components of the splicing cycle, such as mutations in splicing sites and/or spliceosome components, or from mutations or altered expression of non-spliceosome RNA binding proteins, which assist splice site selection and accurate splicing.

Three specific components of complex A have recently gained attention as they are recurrently mutated in various cancer types and affect the very first step in the splicing cycle: the U1 snRNA, which is essential for 5' splice site recognition, the U2 snRNP component SF3B1 and the U2 auxiliary factor subunit U2AF1, both of which are involved in 3' splice site selection [15–20]. Elegant studies that took advantage of whole-transcriptome analysis and in vivo and in vitro models have revealed the consequences of such spliceosome mutations and their impact on cancer development and progression [15–20]. Here, we review the current knowledge about the genetic characteristics of U1 snRNA, SF3B1, and U2AF1 somatic mutations and the mechanism by which they influence splicing decision and, as a consequence, tumor development.

2. Tumor-Associated U1 snRNA Mutations Give Rise to Aberrant 5' Splice Site Recognition

Cancer-related alterations in the noncoding component of the spliceosome, the snRNAs, have barely been recognized and studied. One exception is represented by U1 snRNA, whose somatic mutations were recently identified in different cancer types [18]. This snRNA forms a base-pair interaction with the 5' splice site of the pre-mRNA and is essential for the 5' splice site recognition operated by the spliceosome [21,22] (Figure 2).

The most recurrent mutation in cancer comprises an A > G or A > C substitution at the third base of the U1 snRNA [18,19] (Figure 2). Although corresponding to the same base, each substitution has been found to be restricted to specific cancer types. The 3A > C mutation has been identified in hepatocellular carcinoma (HCC, ~6%), chronic lymphocytic leukemia (CLL, ~4%) and B-cell non-Hodgkin lymphomas (~2%) [18]. In CLL, the 3A > C mutation is restricted to the unfavorable IGHV-unmutated CLL (U-CLL) subtype and is, thus, associated with an aggressive behavior [18]. Interestingly, the 3A > C mutation was detected already at the initial stages of the disease in both CLL and HCC, suggesting that this is an early event during tumorigenesis. Instead, the 3A > G mutation has been identified in medulloblastoma (MBL, ~19%) and with a very low frequency in pancreatic adenocarcinoma (PAAD, 0.4%) [18,19]. In MBL, the U1 mutation is restricted to the sonic hedgehog (SHH) subgroup tumors and with a mutation frequency up to 50% [19]. The 3A > G mutation in SHH was found to be highly prevalent in adults (97%), less frequent in adolescents (25%) and absent in infants suffering from MBL [19]. In the case of adolescent patients presenting a mutation in SHH, the U1 snRNA mutation was associated with high risk of relapse.

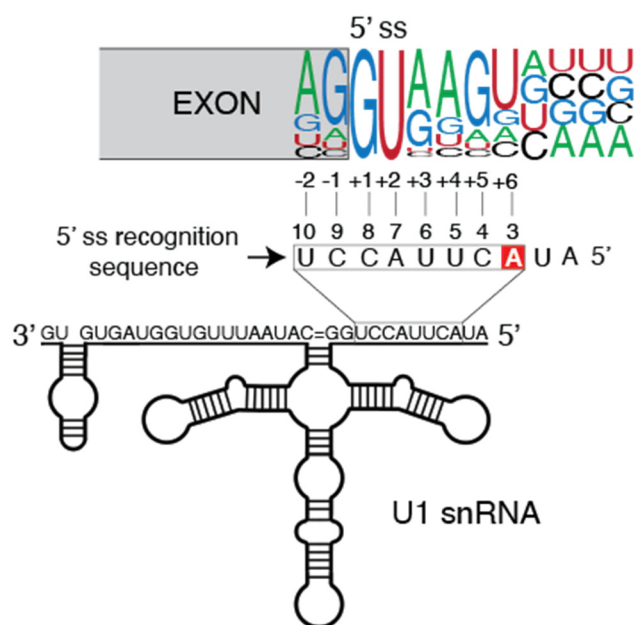


Figure 2. Canonical U1 snRNA and 5' splice site base-pairing. Schematic representation of the base-pair interaction between the U1 5' splice site recognition sequence (box) and the 5' splice site consensus motif. The numbers above the box indicate the base position with respect to the 5' end. The adenine at the third base of the U1 snRNA, which is recurrently mutated in cancer, is highlighted in a red box. The 5' splice site consensus motif is shown and the numbers below the sequence indicate the base position at the exon (–) and the intron (+) with respect to the exon/intron junction.

2.1. Molecular Basis of Altered Splicing Mediated by the U1 snRNA 3A > C/G Mutant

Tumors harboring the U1 snRNA A > C/G mutation present altered splicing patterns as the mutation generates many novel 5' splice sites that are usually absent in non-mutated tissues [18,19]. Splicing alterations observed in cancer patients mainly correspond to: (i) 5' splice site usage variations at exons and (ii) 5' cryptic splice sites generating new exons [18,19].

The 3A > C/G mutation is located in the highly conserved 5' splice site recognition sequence that matches the pre-mRNA [21,22] (Figure 2). The adenine in the third position directly base-pairs with the +6 base at the 5' splice site, which is highly represented by an uracil in the 5' splice site consensus motif (Figure 2). The 3A > C and 3A > G mutations are, therefore, predicted to impact the A:U base pairing specificity toward C:G and G:C base-pairing, respectively. Indeed, 5' splice site which usage was increased in U1 snRNA 3A > C mutant CLL and HCC tumors shown a clear enrichment of G at the +6 position that is not observed in canonical or inefficiently spliced 5' splice sites in mutant tumors [18]. Similarly, the cryptic 5' splice sites observed in MBL harboring the 3A > G mutation are enriched in C at the +6 position [19]. As a consequence of the mismatch with the canonical sequence, intron retention was observed in MBL tumors harboring the U1 snRNA A > G mutation [19]. Indeed, 5' splice sites from the retained introns are enriched in uracil at +6 position, the preferred 5' splice site consensus motif that is well recognized by wild-type U1 snRNA. This observation suggests that, at least in MBL, mutant U1 snRNA could inhibit the canonical recognition by wild-type U1 snRNA on some specific 5' splice sites [19]. However, how mutant U1 snRNA inhibits wild-type U1 snRNA recognition and why only a subset of 5' splice sites are subjected to this inhibition remain elusive.

2.2. Oncogenic Roles of Mutant U1 snRNA 3A > C/G

One of the important questions that remains unanswered is whether the mutations described above are required/sufficient to drive tumor onset or sustain tumor progression. Important clues about the oncogenic potential of U1 snRNA mutations came from

transcriptome analysis of U1 snRNA mutant tumors that revealed oncogenic programs potentially regulated by U1 snRNA mutations [18]. CLL tumors harboring the 3A > C mutation showed the up-regulation of genes belonging to RNA processing (transcription and splicing), protein ubiquitination, and telomere maintenance, whereas genes belonging to apoptosis, B-cell receptor signaling, and cytoplasmic ribosome processes were down-regulated. Moreover, splicing pattern analysis of the tumors identified aberrant splicing events in genes known to be involved in cancer, such as CD44, POLD1, MSI2, and ABCD3. In the case of SHH medulloblastomas, the U1 snRNA A > G mutation was associated with splicing alterations in components of the sonic hedgehog signaling pathway, such as PTCH1, GLI2, and CCND2 [19]. This observation is relevant as sonic hedgehog signaling activation is a known driver of SHH medulloblastoma [23], thus directly associating U1 snRNA mutations with SHH tumorigenesis. Detailed functional *in vivo* analysis of the contribution of mutant U1 snRNA and the role of the generated cancer isoforms is required to uncover the oncogenic role of the mutant U1 snRNA and its effectors, which will be instrumental to exploit novel therapeutic strategies.

3. Spliceosomes Containing Mutant SF3B1 Promote Tumorigenesis through the Use of Cryptic 3' Splice Sites

One of the most exciting discoveries of the recent years has been the recurring somatic mutations in SF3B1, in particular in hematopoietic and lymphoid malignancies [24]. SF3B1 is an essential component of the U2 snRNP, which is critical to BPS recognition and essential for 3' splice site selection [1,2]. SF3B1 mutations are highly prevalent in myelodysplasia syndromes (MDS, ~30%), reaching mutational rates up to 83% in the refractory anemia with ring sideroblasts MDS subtype (RARS) [25]. SF3B1 is also the most frequently mutated gene in CLL (~14%) [26,27], and one of the few genes found to be frequently mutated in uveal melanoma (UVM) (~18%) [28]. Interestingly, SF3B1 mutations correlate with good prognosis in MDS and UVM patients [28,29], while a reduced overall survival was observed in CLL and luminal B and progesterone receptor-negative breast cancer patients [30–32].

Alterations of the SF3B1 gene, such as non-silent mutations, could alter both splice site recognition and splicing decision. Indeed, SF3B1 harbors unique characteristics that support its function. On the one hand, SF3B1 acts on splice site recognition, crosslinking with RNA at both flanking sites of the BPS [7]. On the other hand, SF3B1 can directly interact, via its N-terminal region, with the BPS binding protein and U2 snRNP subunit p14 [33] and with the PPT binding protein U2AF2 [7].

Remarkably, mutations in SF3B1 are heterozygous and the vast majority are single nucleotide missense mutations located at the C-terminal of the protein inside the HEAT-repeat domain (HD) (Figure 3A). The most frequent SF3B1 mutations are K700E, K666N, and R625H. The K700E mutation is the prevalent mutation in MDS and CLL [25], whereas mutations in R625 are highly frequent in UVM [28]. Hereafter, these hotspot mutations will be referred to as SF3B1^{MUT}.

3.1. Molecular Basis of Cryptic 3' Splice Site Usage by SF3B1^{MUT} Spliceosomes

Transcriptome analysis of SF3B1^{WT} and SF3B1^{MUT} tumors from CLL, UVM, and breast cancer patients allowed the identification of splicing alterations associated with SF3B1 hotspot mutations and revealed that SF3B1^{MUT} stimulates the usage of cryptic 3' splice sites by recognizing a cryptic BPS [34–36]. Indeed, the most frequently altered splicing events in SF3B1^{MUT} tumors are cryptic 3' splice site usage (up to 60–76%) [34–36], with 3' cryptic splicing sites located both close to (proximal, ~55%) or far away from (distal, ~45%) the canonical 3' splice site [34].

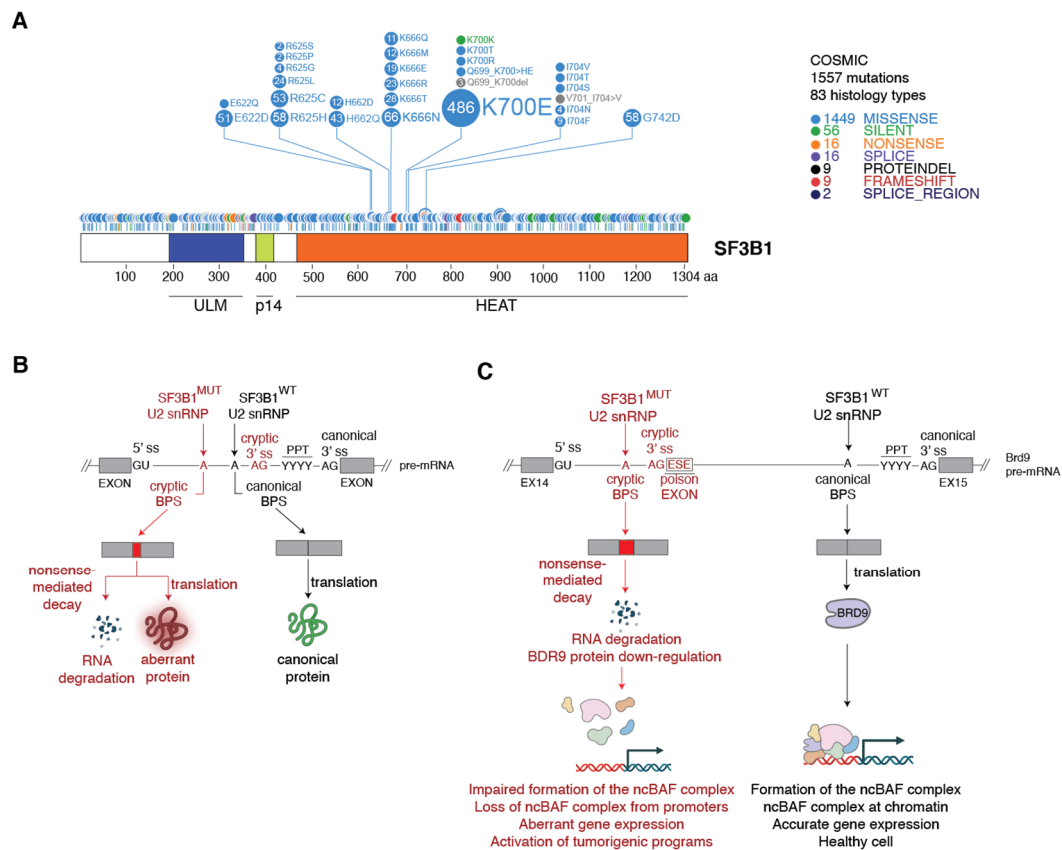


Figure 3. Mutant SF3B1 binds cryptic BPS and promotes the use of cryptic 3' splice sites. (A) Schematic representation of the SF3B1 protein domains and its mutation profile. SF3B1 domains are indicated: U2AF-homology ligand motif (ULM), p14 binding domain (p14), and HEAT-repeat domain (HEAT). The mutation profile was obtained from the COSMIC database (<https://cancer.sanger.ac.uk/cosmic>, accessed on 25 October 2021) and visualized using ProteinPaint (<https://pecan.stjude.cloud/proteinpaint>, accessed on 25 October 2021). The number of samples harboring each mutation is indicated inside each circle and their relative abundance is represented by the disc size. (B) Sequence requirements for cryptic BPS usage by mutant SF3B1. A schematic exon-intron-exon region is shown together with the characteristic sequence features present in the splicing events altered by SF3B1^{MUT}: the 5' splice site (5' ss), the canonical and cryptic 3' splice sites (3' ss), the canonical and cryptic BPS, and the PPT. Red indicates the SF3B1 mutant, the adenine at the cryptic BPS, and the AG dinucleotide at the 3' cryptic splice site. (C) Mutant SF3B1 promotes the inclusion of a poison exon in the BRD9 transcript. Schematic representation of the BRD9 pre-mRNA containing the exon 14-intron-exon 15 region. The required splicing sequences are indicated as in (B). ESE, exon splicing enhancer. Poison exon inclusion targets the BRD9 transcript to NMD, promoting BRD9 protein down-regulation and impaired formation of the ncBAF complex.

Analysis of specific splicing events in model cell lines showed that neither SF3B1 knock-down (KD) nor SF3B1 over-expression recapitulate the splicing alterations observed in SF3B1^{MUT} cells, indicating that hotspot mutations in SF3B1 are likely change-of-function mutations [36]. The main consequence of the cryptic 3' splice site usage by SF3B1^{MUT} is the loss of the open reading frame, with ~50% of the affected transcripts being targeted for degradation by non-sense mediated decay (NMD), leading to reduced expression of the codified protein [35] (Figure 3B).

Proximal cryptic 3' splice sites recognized by SF3B1^{MUT}-U2 snRNPs are characterized by a sequence signature that includes a cryptic AG dinucleotide (cryptic 3' splice site) localized upstream the canonical 3' splice site and downstream the canonical BPS [35,36] (Figure 3B). In addition, a hidden BPS, rich in adenines, is present upstream the cryptic AG

3' splice site. SF3B1^{MUT}-U2 snRNP recognizes an adenine present in this cryptic BPS, thus stimulating the usage of the cryptic 3' splice site [35,36].

Interestingly, SF3B1^{MUT} likely stimulates the usage of distal and proximal 3' splice sites by the very same mechanism [34]. Indeed, inspection of the distal cryptic 3' sites exploited by SF3B1^{MUT} led to the identification of an upstream adenosine-rich sequence, very similar to the cryptic BPS sequence used by SF3B1^{MUT} at proximal 3' splice sites. As not all introns contain such cryptic BPS required for SF3B1^{MUT} recognition, only a subset of transcripts is actually aberrantly spliced in SF3B1^{MUT} tumors.

Why the SF3B1^{MUT}-U2 snRNP binds preferentially to a cryptic BPS instead of the canonical one remains unclear. It has been proposed that SF3B1^{MUT} may possess an enhanced affinity towards the specific nucleotides at flanking regions of the cryptic BPS [35]. Another hypothesis suggests that SF3B1^{MUT} could induce a conformational change in the U2 snRNP, which in turn promotes binding to the cryptic BPS [36]. Unfortunately, at present, no data validated these models. Using purified SF3b complexes and crosslinking experiments, SF3B1^{WT} and SF3B1^{MUT} were shown to bind equally efficiently to a synthetic BPS-PPT RNA [37]. Moreover, SF3B1^{MUT} did not seem to affect the structure and molecular interaction of the SF3b complex [37], including the interaction with U2AF2 in vivo [36] and in vitro [37]. Thus, SF3B1^{MUT} shows no major effects on the topological architecture of the splicing complex which, however, does not exclude minor but critical conformational changes in the context of a full U2 snRNP complex where the stability of the U2 snRNP:BPS interaction could be affected by SF3B1 mutations.

A third possibility is that SF3B1 may affect binding with interactors outside the U2 snRNP, such as the intron itself or additional splicing factors. Indeed, the hotspot residue K700 has been found to be exposed to the solvent in the SF3b complex structure [37]. By purifying and analyzing the composition of U2 snRNP complex wild-type and mutant from K562 myelogenous leukemia cells, Zhang and colleagues found that the SUGP1 protein is lost in SF3B1^{MUT}-containing complexes [38]. Strikingly, SUGP1 KD in SF3B1^{WT} cell lines recapitulates the usage of cryptic BPS and cryptic 3' splice sites observed in SF3B1^{MUT} cell lines [38]. Furthermore, the over-expression of SUGP1 in SF3B1^{MUT} cells rescues the interaction between SUGP1 and SF3B1^{MUT}-containing complexes and partially rescue 3' splice site usage alterations [38]. Both observations suggest that the loss of SUGP1 from SF3B1^{MUT} spliceosomes could dictate the usage of the cryptic 3' splicing site observed in SF3B1^{MUT} tumors.

How SUGP1 achieves such a critical role remains an open question. Interestingly, point mutations in the G-patch domain at the C-terminal of SUGP1 is sufficient to induce the usage of cryptic 3' splice sites in cell lines [38]. Based on the role of G-patch proteins as activators of DEAH-box RNA helicases [39], it has been proposed that SUGP1 could act by modulating RNA helicase activity during splicing [38]. By doing so, SUGP1 could mediate the remodeling and exchange of splicing factors at the conventional BPS, such as the displacement of SF1 by the U2 snRNP. Thus, in the context of the SF3B1^{MUT}, SUGP1 would not be recruited at the spliceosome, forcing U2 snRNP deposition at the cryptic BPS [38]. Again, this intriguing model lacks experimental validation, as the hypothetical RNA helicase modulated by SUGP1 remains unsettled.

3.2. Mechanism of SF3B1^{MUT}-Driven Tumorigenesis

SF3B1 mutations have multiple consequences in tumorigenesis. It has been suggested that SF3B1 mutations arise later in leukemia development and contribute to disease progression [27]. Accordingly, mutations in SF3B1 correlate with a rapid disease progression in CLL [40]. Mice expressing heterozygous and hematopoietic-restricted SF3B1 K700E mutation do not develop leukemia but exhibit macrocytic anemia due to a block in terminal erythroid maturation, erythroid dysplasia, and long-term hematopoietic stem cell expansion, all of which are characteristic features of MDS [15]. Transcriptome analysis of CLL samples from SF3B1^{MUT} tumors, identified expression alterations in genes belonging to biological pathways associated with poor prognosis [41]. An example is represented by

the Notch pathway, whose aberrant signaling has been linked to CLL with reduced overall survival [42]. Myeloid and leukemic cell lines expressing SF3B1^{MUT} exhibit a higher Notch pathway activity [26], possibly due to the expression of a DVL2 exon 11-aberrant isoform unable to repress Notch signaling [41]. Correlative evidence was also found for a few altered splice variants induced by SF3B1^{MUT} (CHD1L, GAK, RAD9A, JMY) involved in the DNA damage response [41,43].

Recently, Inoue and colleagues identified one key oncogenic effector of the SF3B1^{MUT} spliceosome, the non-canonical chromatin remodeling BAF complex (ncBAF) subunit BRD9 [17]. By using a distal cryptic 3' splice site, a cryptic BPS, and an exon splicing enhancer (ESE), SF3B1^{MUT} induces the inclusion of a new exon upstream the BRD9 exon 15 (Figure 3C). This poison exon (14a) interrupts the open reading frame and targets the aberrant BRD9 transcript to NMD, thereby down-regulating BRD9 expression. Remarkably, the inclusion of exon14a was observed in all cohorts of CLL, UVM, and MDS patients harboring SF3B1 hotspot mutations [17]. At the functional level, SF3B1^{MUT}, by inducing BRD9 loss, disrupts the ncBAF complex and, as a consequence, its loss from chromatin affects the expression of genes involved in apoptosis, cell growth, adhesion, and migration [17]. Validating experiments showed that BRD9 KD in non-tumorigenic Melan-a melanocytes results in tumor growth and the induction of melanomagenesis in vivo while BRD9 depletion in melanoma cells increases the number of pulmonary metastases [17]. Finally, the inhibition of BRD9 poison exon inclusion by using specific antisense oligonucleotides (ASOs) rescues BRD9 expression in SF3B1^{MUT} melanoma cells in vitro and inhibits tumor growth by inducing tumor necrosis in vivo [17]. Thus, the loss of BRD9 represents the first SF3B1^{MUT}-induced pro-oncogenic event driving tumorigenesis. While more in-depth functional studies are needed to uncover further critical oncogenic effectors of SF3B1^{MUT}, these studies have validated ASOs as a valuable therapeutic option to treat SF3B1^{MUT} cancers.

4. Cancer-Associated U2AF1 Mutations Influence U2AF1 3' Splice Site Recognition and Splicing Outcome

A third cancer-associated mutated splicing component is U2AF1, which is recurrently mutated in myelodysplasia syndromes (MDS, ~9%) and hematological malignancies such as chronic myelomonocytic leukemia (CMML, ~8%) and acute myeloid leukemia (AML, ~4%) [25,44,45]. U2AF1 mutations are more frequent in high-risk MDS patients and are associated with reduced survival in CMML patients [45]. Similar bad prognosis and higher frequency of disease persistence after chemotherapy were observed in intermediate-risk AML patients with U2AF1 mutations [46]. Finally, U2AF1 mutations are present in lung adenocarcinoma (LUAD, ~5%), where they are associated with reduced patient survival [35,47].

From the molecular point-of-view, U2AF1 heterodimerizes with U2AF2, forming the U2 snRNP auxiliary factor U2AF, which is essential for 3' splice site recognition [5,6]. U2AF1 specifically recognizes the AG dinucleotide at the 3' splice site [48–50], whereas U2AF2 binds the PPT and interacts with the U2 snRNP [7,50] (Figure 1A).

In this case also, U2AF1 mutations are heterozygous and the retained wild-type allele is expressed, suggesting that mutant cancer cells may require the residual normal protein to be viable. Remarkably, the mutations are mostly located at two hotspot positions, Ser34 (S34F/Y) and Glu157 (Q157P/R) (Figure 4A). S34F/Y is the most recurrent U2AF1 mutation in MDS and AML, and is the unique mutation found in LUAD [25,44,51,52]. The two cancer-associated residues map in the two zinc finger domains (ZnF) present in U2AF1 (Figure 4A). Recent structural studies performed with fission yeast U2AF1 in complex with a 3' splice site RNA sequence suggested that both ZnFs bind and recognize the AG dinucleotide [53]. This biochemical study has illuminated the molecular mechanism of the recognition of 3' splice site by U2AF1 suggesting a dysfunctional role of U2AF1^{MUT} in RNA recognition and, as a consequence, in the splicing outcome.

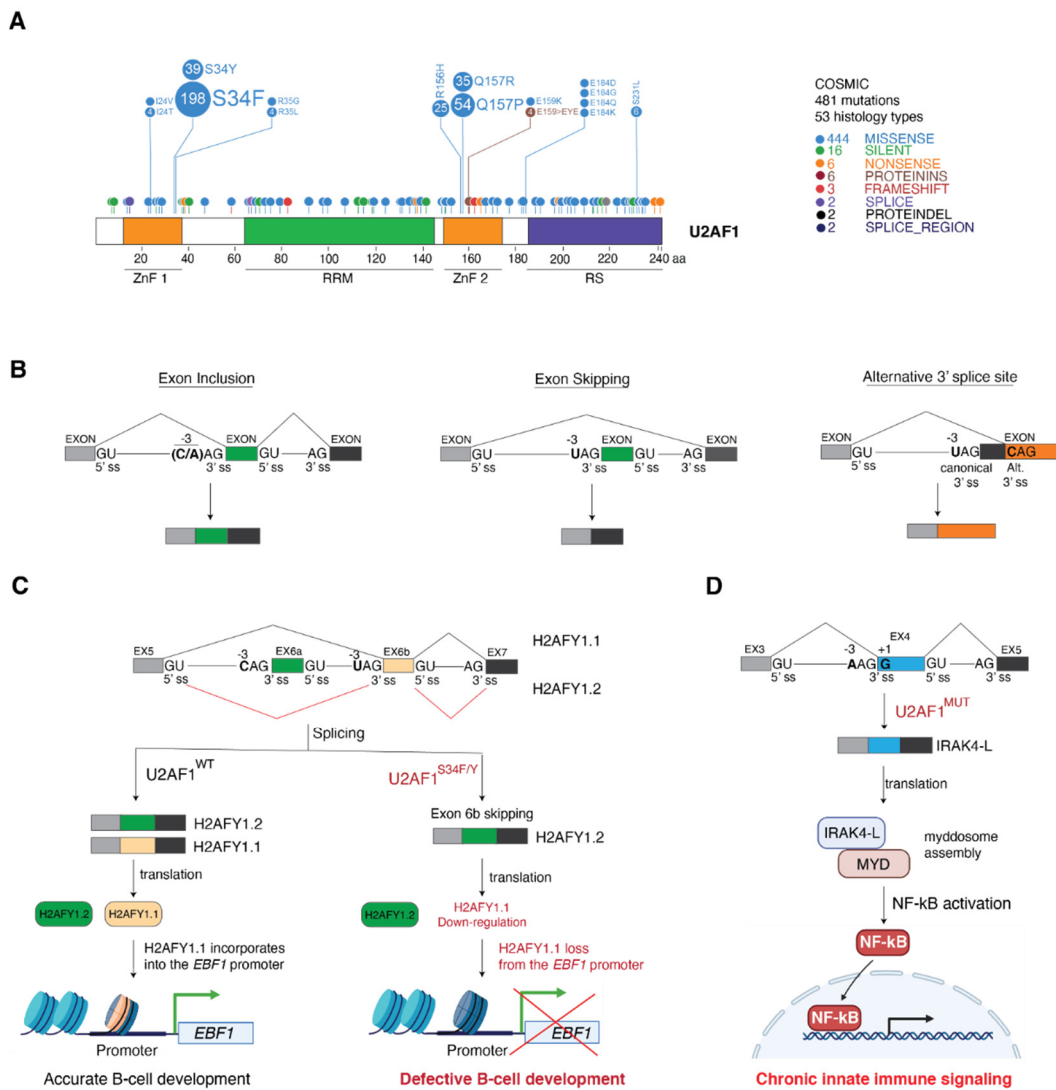


Figure 4. U2AF1 mutations influence U2AF1 3' splice site recognition and splicing outcome. (A) Schematic representation of the U2AF1 protein with its mutation profile. U2AF1 domains are indicated: zinc finger domains 1 (ZnF 1) and 2 (ZnF 2), the RNA recognition motif (RRM) and the arginine/serine rich domain (RS). The mutation profile was obtained from the COSMIC database (<https://cancer.sanger.ac.uk/cosmic>, accessed on 25 October 2021) and visualized using ProteinPaint (<https://pecan.stjude.cloud/proteinpaint>, accessed on 25 October 2021). The number of samples harboring each mutation is indicated inside each circle and their relative abundance is represented by the disc size. (B) Features of splicing events affected by the S34F/Y U2AF1 mutant. Schematic exon-intron-exon regions are shown together with the characteristic sequence features present in the splicing events altered by U2AF1^{S34F/Y}: exon inclusion (left), exon skipping (center), and alternative 3' splice site usage. The positions of the 5' splice site (5' ss), the canonical 3' splice sites (3' ss), and the alternative 3' splice site (Alt 3' ss) are indicated. The -3 nucleotide, which impacts on 3' splice site selection by mutant U2AF1^{S34F/Y}, is highlighted in bold. (C) U2AF1^{S34F/Y} down-regulates the H2AFY1.1 transcript, leading to B-cell developmental defects. Schematic representation of the H2AFY pre-mRNA from exon 5 to exon 7, containing the mutually exclusive exons 6a and 6b. Splicing sequences are indicated as in (B). Both H2AFY isoforms are expressed in U2AF1^{WT} cells. In U2AF1^{S34F/Y} mutant cells, exon 6b is excluded; thus, H2AFY1.1 is not expressed. H2AFY1.1 loss down-regulates *Ebfl* expression and leads to B-cell developmental defects. (D) Mutant U2AF1 up-regulates IRAK4 long isoform expression and NF-κB activation. Schematic representation of the IRAK4 pre-mRNA from exon 3 to exon 5, containing the alternative splice exon 4. Splicing sequences are indicated as in (B). The +1 G base is indicated in bold. IRAK4-L promotes the myddosome complex (MYD) assembly and activates NF-κB signaling.

4.1. Molecular Basis of Altered 3' Splice Site Recognition by Mutant U2AF1

Deep analysis of the transcriptome from patient-derived primary samples (MDS, AML, LUAD) and human and mouse cell lines expressing U2AF1^{WT} or U2AF1^{MUT} have allowed the identification of rare U2AF1^{MUT}-dependent splicing alterations [16,51,52,54–56]. In agreement with its moderate effect on splicing, iCLIP data from HCC78 LUAD cells ectopically expressing U2AF1^{WT} or U2AF1^{S34F} showed only ~20% difference in occupancy between U2AF1^{WT} and U2AF1^{S34F} [56]. The most common splicing events altered in U2AF1^{MUT} cancers are exon cassettes (~40–60%) followed by alternative 3' splice sites usage (~20%).

In contrast to SF3B1 mutations, U2AF1^{MUT} does not feed the NMD pathway with new substrates [54], but alters both exon usage and polyadenylation site selection and translation efficiency by de-regulating U2AF1 binding to polyadenylation sites and 5' UTR regions, respectively [57,58]. Interestingly, the S34F/Y and Q157P/R mutations present different patterns of exon inclusion [54], indicating a differential impact on cancer cell physiology. U2AF1^{MUT}-regulated exons are characterized by specific sequence features. Exons included by the U2AF1^{S34F/Y} mutant are enriched in C/A nucleotide at position -3, whereas frequently skipped exons often have a U in the same position [16,51,52,54–56] (Figure 4B). Importantly, adjacent exons do not show such enrichment and have an equal probability of C and U enrichment at the -3 position [51]. In line with the sequence analysis, iCLIP data from HCC78 LUAD cells, ectopically expressing U2AF1^{WT} or U2AF1^{S34F}, demonstrated that U2AF1^{S34F} preferentially binds (C/A)AG over UAG [56]. This sequence signature was also observed inside exon sequences subjected to 3' alternative splice site usage by mutant U2AF1^{S34F/Y} [52,55] (Figure 4B).

In the case of the U2AF1^{Q157P/R} mutant, preferentially included exons were enriched in G over A at the +1 nucleotide [54] (Figure 4B). The same splice site sequence selectivity was also observed in K562 erythroleukemic cell lines ectopically expressing U2AF1^{WT}, U2AF1^{S34F/Y} or U2AF1^{Q157P/R} [54]. Importantly, splicing events altered by U2AF1 KD in this cellular model did not show any particular sequence enrichment, and did not mimic U2AF1^{MUT}. Thus, splicing alterations observed in U2AF1^{MUT} tumors are the consequence of change-of-function mutations [54], with the U2AF1^{S37F/Y} mutant preferentially splicing (C/A)AG over UAG splice sites (Figure 4B). It is worth mentioning that the S37F/Y and Q157P/R mutations have not been investigated at the structural and biochemical level and future molecular studies on U2AF1^{MUT}:RNA interaction are predicted to shed light on the mechanism underlying U2AF1 activity.

4.2. Consequences of U2AF1 Mutations on Tumorigenesis

Transgenic mice expressing the U2AF1^{S37F} mutation display various abnormalities associated with MDS, including peripheral blood leukopenia, reduction in B cells and monocytes in the bone marrow, and bone marrow progenitor cell expansion [16,20]. In addition, transcriptome analysis of common myeloid progenitors (CMPs) from transgenic mice identified U2AF1^{S34F}-dependent gene expression alteration in genes belonging to immune response and leukocytic activation processes [16]. Nevertheless, U2AF1^{S34F} transgenic mice do not develop MDS or AML [16]. Although U2AF1^{MUT} does not generate widespread splicing alterations, altered splicing events driven by U2AF1^{MUT} belong to mutated genes in MDS and AML (ASXL1, GNAS, PICALM) and involved in cancer hallmarks, including stem cell biology (MED24), transcription regulation (H2AFY), replication stress response pathway (ATR), and innate immune pathway (IRAK4) [16,20,54,59].

The defects in B cell development observed in U2AF1^{S34F} transgenic mice [20] have been associated with U2AF1^{S34F}-mediated splicing alterations of the transcript of histone variant H2AFY (macroH2A1), which is associated with both transcription repression and transcription activation [60]. Two different isoforms are generated by the mutually exclusive inclusion of exon 6a and exon 6b, leading to H2AFY1.2 and H2AFY1.1, respectively (Figure 4C). Wild-type human MDS bone marrow expresses both isoforms, while U2AF1^{S34F/Y} MDS samples express only the H2AFY1.2 isoform, indicating that

U2AF1^{S34F/Y} promotes exon 6b skipping and H2AFY1.1 down-regulation [16] (Figure 4C). Sequence feature analyses support this idea: exon 6a presents a CAG trinucleotide at the 3' splice site, whereas exon 6b presents a UAG trinucleotide, characteristic sequences for U2AF1^{S34F/Y} exon inclusion and exon skipping, respectively (Figure 4C).

Interestingly, H2AFY knock-out (KO) mice present the same B cell development defects, but not the additional MDS features observed in U2AF1^{S34F} transgenic mice [20]. Moreover, the expression of H2AFY1.1, but not H2AFY1.2, rescues the B cell development defects observed in both H2AFY KO and U2AF1^{S34F} transgenic mice [20]. These results mechanistically link the U2AFY splicing alteration to the defects in B cell development observed in U2AF^{S34F} transgenic mice. Molecularly, the down-regulation of the U2AFY1.1 isoform leads to its loss at the promotor of the transcription factor EBF1, which is critical for B cell development, and is, therefore, the key responsible for the phenotype observed [20].

Another oncogenic effector of U2AF1^{MUT} is the interleukin-1 receptor-associated kinase 4 (IRAK4) (Figure 4D). IRAK4 is a Serine/Threonine kinase that acts downstream the Toll-like receptor superfamily, activating NF-κB and MAPK signaling in inflammation [61]. In AML, U2AF1^{MUT} mediates the inclusion of IRAK4 exon 4 and, therefore, the expression of the IRAK4 long isoform (IRAK4-L), which is frequently expressed in AML patients with poor outcome [59]. The recognition of the 3' splice site at exon 4 by U2AF1^{MUT} is favored by the presence of an adenine at the -3 position and a guanine at the +1 position, which are the preferred binding splice site sequences for U2AF1^{S34F/Y} and U2AF1^{Q157P/R}, respectively (Figure 4D). Molecularly, exon 4 codifies for the N-terminal death-domain (DD) that is essential for protein-protein interactions and the formation of the myddosome signaling complex. In AML, the U2AF1^{MUT}-mediated expression of IRAK4-L stabilizes IRAK4 interaction with myddosome components, resulting in maximal activation of NF-κB signaling, which is essential for leukemic cell fitness [59]. Importantly, *in vivo* xenografts from MDS patients harboring U2AF1 mutations were strongly affected when treated with the IKAR4 inhibitor CA-4948, showing a 50% decrease in MDS cell engraftment. These observations uncover a previously unknown vulnerability to IRAK4 inhibitors of AML tumors harboring U2AF1 mutations, which might be relevant for therapeutical approaches such as targeted ASOs.

5. Conclusions and Outlook

Genome-wide transcriptome analysis identified 119 splicing factor genes with non-silent mutations across different tumors, including hotspot and loss-of-function mutations [24]. Tumors appear to modify the splicing outcome by selecting alterations in key factors involved in early stages of the splicing process. Indeed, hotspot mutations in the U2 spliceosome components SF3B1, U2AF1 [24], and U1 snRNA [18,19] were found to be exceptionally well represented. On the other hand, to restrict the otherwise broad and deleterious effects generated by defects on splice site selection, cancer cells appear to select specific mutations to achieve a "balanced alteration". This is ensued in two ways: first, U1 snRNA, SF3B1, and U2AF1 mutations occur in specific hotspot bases/residues that restrict their binding and activity to a specific group of transcripts; second, the heterozygous nature of the mutations guaranties the presence of the wild-type spliceosomes, which can carry out splicing on canonical splice sites. Cells cannot tolerate a dramatic and widespread alteration in splicing, as demonstrated by the fact that spliceosome mutations are mutually exclusive [18,25]. In addition, cancer cells with SF3B1 or U2AF1 mutations require the presence of the wild-type spliceosome to survive [16,62,63]. Thus, cancer cells exploit spliceosome mutations to drive cancer-specific and tolerable splicing changes in order to support transformation and tumor development.

Importantly, SF3B1^{MUT} or U2AF1^{MUT} cancer cells are highly susceptible to pharmacology-driven splicing perturbations [15,16], a phenomenon that has been therapeutically exploited. Unfortunately, the failure of the phase I clinical trial of the SF3B1 inhibitor E7107 in solid tumors [64] has underlined the need for specific therapies that are non-toxic for normal cells. A more selective strategy is represented by ASOs, which, by targeting

specific splicing events, allow the restoration of non-oncogenic isoforms [65]. The use of nusinersen in the treatment of patients with spinal muscular atrophy is paradigmatic of this therapeutic approach [66]. Preclinical work on the SF3B1 target BRD9 suggests that ASOs might also have a therapeutic value in oncology [17]. A similar strategy could be exploited to restore H2AFY1.1 expression and inhibit IRAK4-L expression in CLL U2AF1^{S34F} cells. It is predicted that uncovering additional oncogenic isoform effectors of the mutant spliceosome will not only improve our understanding of how the mutant spliceosome drives tumor development but will also allow the design of more selective therapeutic strategies.

Author Contributions: C.A.N. and R.S.d.P. wrote the manuscript and prepared the figures; S.P. conceptualized and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Associazione Italiana per la Ricerca sul Cancro (AIRC IG. IG19875) and the Italian Ministry of Education, Universities and Research (PRIN 20108MXN2J). C.A.N.'s work was supported by the Marie Skłodowska-Curie Actions (MSCA-IF-2016 #752553).

Acknowledgments: Figures were made using BioRender and ProteinPaint.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ABCD3	ATP Binding Cassette Subfamily D Member 3
AML	Acute myeloid leukemia
ASO	Antisense oligonucleotide
ASXL1	ASXL Transcriptional Regulator 1
ATP	Adenosine triphosphate
ATR	ATR Serine/Threonine Kinase
BPS	Branch point sequence
BRD9	Bromodomain Containing 9
CCDN2	Cyclin D2
CD44	Cluster of differentiation 44
CHD1L	Chromodomain Helicase DNA Binding Protein 1 Like
CLIP	UV cross-linking and immunoprecipitation
CLL	Chronic lymphocytic leukemia
CMML	Chronic myelomonocytic leukemia
CMPs	Common myeloid progenitors
DD	Death domain
DVL2	Dishevelled-2
ESE	Exon splicing enhancer
GAK	Cyclin G Associated Kinase
GLI2	GLI Family Zinc Finger 2
GNAS	GNAS Complex Locus
H2AFY	MacroH2A.1 Histone
H2AFY1.1	MacroH2A.1 Histone isoform 1
H2AFY1.2	MacroH2A.1 Histone isoform 2
HCC	Hepatocellular carcinoma
HD	HEAT repeat domain
iCLIP	Individual-nucleotide resolution CLIP
IGHV	Immunoglobulin heavy-chain variable region
IRAK4	Interleukin 1 Receptor Associated Kinase 4
IRAK4-L	Interleukin 1 Receptor Associated Kinase 4 isoform long
JMY	Junction Mediating and Regulatory Protein, P53 Cofactor
KD	Knockdown
KH	K-homology
LSm	Sm-like
LUAD	Lung adenocarcinoma

Abbreviations

MAPK	Mitogen-activated protein kinase
MBL	Medulloblastoma
MDS	Myelodysplastic syndromes
MED24	Mediator Complex Subunit 24
MSI2	Musashi RNA Binding Protein 2
ncBAF	non canonical chromatin remodelling BAF complex
NF-κB	Nuclear Factor Kappa B Subunit 1
NMD	Non-sense mediated decay
NOTCH	Notch homolog 1, translocation-associated (Drosophila)
PAAD	Pancreatic adenocarcinoma
PICALM	Phosphatidylinositol Binding Clathrin Assembly Protein
POLD1	DNA Polymerase Delta 1, Catalytic Subunit
PPT	Polypyrimidine tract
pre-mRNA	precursor messenger RNA
PTCHI	Patched homolog 1
RAD9A	RAD9 Checkpoint Clamp Component A
RARS	Refractory anemia with ring sideroblasts MDS subtype
RNP	Ribonucleoprotein
RRM	RNA recognition motif
RS domain	Serine-arginine rich domain
SF1	Splicing factor 1
SF3B	Splicing factor 3B
SHH	Sonic hedgehog
snRNA	small nuclear RNA
snRNP	small nuclear RNP
SUGP1	SURP and G-Patch Domain-Containing Protein 1
U-CLL	IGHV-unmutated CLL
U2AF1/2	U2 auxiliary factor subunit 1/2
ULM	U2AF ligand motif
UTR	Untranslated region
UVM	Uveal melanoma
ZNF	Zinc finger domain

References

- Will, C.L.; Luhrmann, R. Spliceosome structure and function. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a003707. [[CrossRef](#)]
- Matera, A.G.; Wang, Z. A day in the life of the spliceosome. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 108–121. [[CrossRef](#)] [[PubMed](#)]
- Kastner, B.; Will, C.L.; Stark, H.; Luhrmann, R. Structural Insights into Nuclear pre-mRNA Splicing in Higher Eukaryotes. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a032417. [[CrossRef](#)] [[PubMed](#)]
- Berglund, J.A.; Chua, K.; Abovich, N.; Reed, R.; Rosbash, M. The splicing factor BBP interacts specifically with the pre-mRNA branchpoint sequence UACUAAC. *Cell* **1997**, *89*, 781–787. [[CrossRef](#)]
- Ruskin, B.; Zamore, P.D.; Green, M.R. A factor, U2AF, is required for U2 snRNP binding and splicing complex assembly. *Cell* **1988**, *52*, 207–219. [[CrossRef](#)]
- Zamore, P.D.; Green, M.R. Identification, purification, and biochemical characterization of U2 small nuclear ribonucleoprotein auxiliary factor. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 9243–9247. [[CrossRef](#)]
- Gozani, O.; Potashkin, J.; Reed, R. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell Biol.* **1998**, *18*, 4752–4760. [[CrossRef](#)]
- Schellenberg, M.J.; Edwards, R.A.; Ritchie, D.B.; Kent, O.A.; Golas, M.M.; Stark, H.; Luhrmann, R.; Glover, J.N.; MacMillan, A.M. Crystal structure of a core spliceosomal protein interface. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1266–1271. [[CrossRef](#)]
- Kahles, A.; Lehmann, K.V.; Toussaint, N.C.; Huser, M.; Stark, S.G.; Sachsenberg, T.; Stegle, O.; Kohlbacher, O.; Sander, C.; The Cancer Genome Atlas Research Network. Comprehensive Analysis of Alternative Splicing across Tumors from 8705 Patients. *Cancer Cell* **2018**, *34*, 211–224.e216. [[CrossRef](#)]
- Jayasinghe, R.G.; Cao, S.; Gao, Q.; Wendl, M.C.; Vo, N.S.; Reynolds, S.M.; Zhao, Y.; Climente-Gonzalez, H.; Chai, S.; Wang, F.; et al. Systematic Analysis of Splice-Site-Creating Mutations in Cancer. *Cell Rep.* **2018**, *23*, 270–281.e273. [[CrossRef](#)]
- Group, P.T.C.; Calabrese, C.; Davidson, N.R.; Demircioglu, D.; Fonseca, N.A.; He, Y.; Kahles, A.; Lehmann, K.V.; Liu, F.; Shiraishi, Y.; et al. Genomic basis for RNA alterations in cancer. *Nature* **2020**, *578*, 129–136. [[CrossRef](#)]
- Paronetto, M.P.; Passacantilli, I.; Sette, C. Alternative splicing and cell survival: From tissue homeostasis to disease. *Cell Death Differ.* **2016**, *23*, 1919–1929. [[CrossRef](#)]

13. Di Matteo, A.; Belloni, E.; Pradella, D.; Cappelletto, A.; Volf, N.; Zacchigna, S.; Ghigna, C. Alternative splicing in endothelial cells: Novel therapeutic opportunities in cancer angiogenesis. *J. Exp. Clin. Cancer Res.* **2020**, *39*, 275. [[CrossRef](#)]
14. Mitra, M.; Lee, H.N.; Collier, H.A. Splicing Busts a Move: Isoform Switching Regulates Migration. *Trends Cell Biol.* **2020**, *30*, 74–85. [[CrossRef](#)]
15. Obeng, E.A.; Chappell, R.J.; Seiler, M.; Chen, M.C.; Campagna, D.R.; Schmidt, P.J.; Schneider, R.K.; Lord, A.M.; Wang, L.; Gambe, R.G.; et al. Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell* **2016**, *30*, 404–417. [[CrossRef](#)]
16. Shirai, C.L.; Ley, J.N.; White, B.S.; Kim, S.; Tibbitts, J.; Shao, J.; Ndonwi, M.; Wadugu, B.; Duncavage, E.J.; Okeyo-Owuor, T.; et al. Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing in vivo. *Cancer Cell* **2015**, *27*, 631–643. [[CrossRef](#)]
17. Inoue, D.; Chew, G.L.; Liu, B.; Michel, B.C.; Pangallo, J.; D’Avino, A.R.; Hitchman, T.; North, K.; Lee, S.C.; Bitner, L.; et al. Spliceosomal disruption of the non-canonical BAF complex in cancer. *Nature* **2019**, *574*, 432–436. [[CrossRef](#)] [[PubMed](#)]
18. Shuai, S.; Suzuki, H.; Diaz-Navarro, A.; Nadeu, F.; Kumar, S.A.; Gutierrez-Fernandez, A.; Delgado, J.; Pinyol, M.; Lopez-Otin, C.; Puente, X.S.; et al. The U1 spliceosomal RNA is recurrently mutated in multiple cancers. *Nature* **2019**, *574*, 712–716. [[CrossRef](#)] [[PubMed](#)]
19. Suzuki, H.; Kumar, S.A.; Shuai, S.; Diaz-Navarro, A.; Gutierrez-Fernandez, A.; De Antonellis, P.; Cavalli, F.M.G.; Juraschka, K.; Farooq, H.; Shibahara, I.; et al. Recurrent noncoding U1 snRNA mutations drive cryptic splicing in SHH medulloblastoma. *Nature* **2019**, *574*, 707–711. [[CrossRef](#)]
20. Kim, S.P.; Srivatsan, S.N.; Chavez, M.; Shirai, C.L.; White, B.S.; Ahmed, T.; Alberti, M.O.; Shao, J.; Nunley, R.; White, L.S.; et al. Mutant U2AF1-induced alternative splicing of H2afy (macroH2A1) regulates B-lymphopoiesis in mice. *Cell Rep.* **2021**, *36*, 109626. [[CrossRef](#)] [[PubMed](#)]
21. Zhuang, Y.; Weiner, A.M. A compensatory base change in U1 snRNA suppresses a 5′ splice site mutation. *Cell* **1986**, *46*, 827–835. [[CrossRef](#)]
22. Kondo, Y.; Oubridge, C.; van Roon, A.M.; Nagai, K. Crystal structure of human U1 snRNP, a small nuclear ribonucleoprotein particle, reveals the mechanism of 5′ splice site recognition. *eLife* **2015**, *4*, e04986. [[CrossRef](#)]
23. Northcott, P.A.; Jones, D.T.; Kool, M.; Robinson, G.W.; Gilbertson, R.J.; Cho, Y.J.; Pomeroy, S.L.; Korshunov, A.; Lichter, P.; Taylor, M.D.; et al. Medulloblastomics: The end of the beginning. *Nat. Rev. Cancer* **2012**, *12*, 818–834. [[CrossRef](#)] [[PubMed](#)]
24. Seiler, M.; Peng, S.; Agrawal, A.A.; Palacino, J.; Teng, T.; Zhu, P.; Smith, P.G.; The Cancer Genome Atlas Research Network. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep.* **2018**, *23*, 282–296.e284. [[CrossRef](#)] [[PubMed](#)]
25. Yoshida, K.; Sanada, M.; Shiraiishi, Y.; Nowak, D.; Nagata, Y.; Yamamoto, R.; Sato, Y.; Sato-Otsubo, A.; Kon, A.; Nagasaki, M.; et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **2011**, *478*, 64–69. [[CrossRef](#)]
26. Wang, L.; Lawrence, M.S.; Wan, Y.; Stojanov, P.; Sougnez, C.; Stevenson, K.; Werner, L.; Sivachenko, A.; DeLuca, D.S.; Zhang, L.; et al. SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. *N. Engl. J. Med.* **2011**, *365*, 2497–2506. [[CrossRef](#)] [[PubMed](#)]
27. Landau, D.A.; Carter, S.L.; Stojanov, P.; McKenna, A.; Stevenson, K.; Lawrence, M.S.; Sougnez, C.; Stewart, C.; Sivachenko, A.; Wang, L.; et al. Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* **2013**, *152*, 714–726. [[CrossRef](#)]
28. Harbour, J.W.; Roberson, E.D.; Anbunathan, H.; Onken, M.D.; Worley, L.A.; Bowcock, A.M. Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. *Nat. Genet.* **2013**, *45*, 133–135. [[CrossRef](#)]
29. Malcovati, L.; Karimi, M.; Papaemmanuil, E.; Ambaglio, I.; Jadersten, M.; Jansson, M.; Elena, C.; Galli, A.; Walldin, G.; Della Porta, M.G.; et al. SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood* **2015**, *126*, 233–241. [[CrossRef](#)]
30. Oscier, D.G.; Rose-Zerilli, M.J.; Winkelmann, N.; Gonzalez de Castro, D.; Gomez, B.; Forster, J.; Parker, H.; Parker, A.; Gardiner, A.; Collins, A.; et al. The clinical significance of NOTCH1 and SF3B1 mutations in the UK LRF CLL4 trial. *Blood* **2013**, *121*, 468–475. [[CrossRef](#)]
31. Landau, D.A.; Tausch, E.; Taylor-Weiner, A.N.; Stewart, C.; Reiter, J.G.; Bahlo, J.; Kluth, S.; Bozic, I.; Lawrence, M.; Bottcher, S.; et al. Mutations driving CLL and their evolution in progression and relapse. *Nature* **2015**, *526*, 525–530. [[CrossRef](#)] [[PubMed](#)]
32. Fu, X.; Tian, M.; Gu, J.; Cheng, T.; Ma, D.; Feng, L.; Xin, X. SF3B1 mutation is a poor prognostic indicator in luminal B and progesterone receptor-negative breast cancer patients. *Oncotarget* **2017**, *8*, 115018–115027. [[CrossRef](#)] [[PubMed](#)]
33. Spadaccini, R.; Reidt, U.; Dybkov, O.; Will, C.; Frank, R.; Stier, G.; Corsini, L.; Wahl, M.C.; Luhrmann, R.; Sattler, M. Biochemical and NMR analyses of an SF3b155-p14-U2AF-RNA interaction network involved in branch point definition during pre-mRNA splicing. *RNA* **2006**, *12*, 410–425. [[CrossRef](#)] [[PubMed](#)]
34. DeBoever, C.; Ghia, E.M.; Shepard, P.J.; Rassenti, L.; Barrett, C.L.; Jepsen, K.; Jamieson, C.H.; Carson, D.; Kipps, T.J.; Frazer, K.A. Transcriptome sequencing reveals potential mechanism of cryptic 3′ splice site selection in SF3B1-mutated cancers. *PLoS Comput. Biol.* **2015**, *11*, e1004105. [[CrossRef](#)] [[PubMed](#)]
35. Darman, R.B.; Seiler, M.; Agrawal, A.A.; Lim, K.H.; Peng, S.; Aird, D.; Bailey, S.L.; Bhavsar, E.B.; Chan, B.; Colla, S.; et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3′ Splice Site Selection through Use of a Different Branch Point. *Cell Rep.* **2015**, *13*, 1033–1045. [[CrossRef](#)] [[PubMed](#)]

36. Alsafadi, S.; Houy, A.; Battistella, A.; Popova, T.; Wassef, M.; Henry, E.; Tirode, F.; Constantinou, A.; Piperno-Neumann, S.; Roman-Roman, S.; et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **2016**, *7*, 10615. [[CrossRef](#)] [[PubMed](#)]
37. Cretu, C.; Schmitzova, J.; Ponce-Salvatierra, A.; Dybkov, O.; De Laurentiis, E.I.; Sharma, K.; Will, C.L.; Urlaub, H.; Luhrmann, R.; Pena, V. Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Mol. Cell* **2016**, *64*, 307–319. [[CrossRef](#)]
38. Zhang, J.; Ali, A.M.; Lieu, Y.K.; Liu, Z.; Gao, J.; Rabadan, R.; Raza, A.; Mukherjee, S.; Manley, J.L. Disease-Causing Mutations in SF3B1 Alter Splicing by Disrupting Interaction with SUGP1. *Mol. Cell* **2019**, *76*, 82–95.e87. [[CrossRef](#)]
39. Bohnsack, K.E.; Ficner, R.; Bohnsack, M.T.; Jonas, S. Regulation of DEAH-box RNA helicases by G-patch proteins. *Biol. Chem.* **2021**, *402*, 561–579. [[CrossRef](#)]
40. Stilgenbauer, S.; Schnaiter, A.; Paschka, P.; Zenz, T.; Rossi, M.; Dohner, K.; Buhler, A.; Bottcher, S.; Ritgen, M.; Kneba, M.; et al. Gene mutations and treatment outcome in chronic lymphocytic leukemia: Results from the CLL8 trial. *Blood* **2014**, *123*, 3247–3254. [[CrossRef](#)]
41. Wang, L.; Brooks, A.N.; Fan, J.; Wan, Y.; Gambe, R.; Li, S.; Hergert, S.; Yin, S.; Freeman, S.S.; Levin, J.Z.; et al. Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. *Cancer Cell* **2016**, *30*, 750–763. [[CrossRef](#)]
42. Puente, X.S.; Pinyol, M.; Quesada, V.; Conde, L.; Ordonez, G.R.; Villamor, N.; Escaramis, G.; Jares, P.; Bea, S.; Gonzalez-Diaz, M.; et al. Whole-genome sequencing identifies recurrent mutations in chronic lymphocytic leukaemia. *Nature* **2011**, *475*, 101–105. [[CrossRef](#)]
43. Te Raa, G.D.; Derks, I.A.; Navrkalova, V.; Skowronska, A.; Moerland, P.D.; van Laar, J.; Oldreive, C.; Monsuur, H.; Trbusek, M.; Malcikova, J.; et al. The impact of SF3B1 mutations in CLL on the DNA-damage response. *Leukemia* **2015**, *29*, 1133–1142. [[CrossRef](#)]
44. Graubert, T.A.; Shen, D.; Ding, L.; Okeyo-Owuor, T.; Lunn, C.L.; Shao, J.; Krysiak, K.; Harris, C.C.; Koboldt, D.C.; Larson, D.E.; et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat. Genet.* **2011**, *44*, 53–57. [[CrossRef](#)]
45. Makishima, H.; Visconte, V.; Sakaguchi, H.; Jankowska, A.M.; Abu Kar, S.; Jerez, A.; Przychodzen, B.; Bupathi, M.; Guinta, K.; Afile, M.G.; et al. Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. *Blood* **2012**, *119*, 3203–3210. [[CrossRef](#)] [[PubMed](#)]
46. Saygin, C.; Hirsch, C.; Przychodzen, B.; Sekeres, M.A.; Hamilton, B.K.; Kalaycio, M.; Carraway, H.E.; Gerds, A.T.; Mukherjee, S.; Nazha, A.; et al. Mutations in DNMT3A, U2AF1, and EZH2 identify intermediate-risk acute myeloid leukemia patients with poor outcome after CR1. *Blood Cancer J.* **2018**, *8*, 4. [[CrossRef](#)]
47. Imielinski, M.; Berger, A.H.; Hammerman, P.S.; Hernandez, B.; Pugh, T.J.; Hodis, E.; Cho, J.; Suh, J.; Capelletti, M.; Sivachenko, A.; et al. Mapping the hallmarks of lung adenocarcinoma with massively parallel sequencing. *Cell* **2012**, *150*, 1107–1120. [[CrossRef](#)] [[PubMed](#)]
48. Wu, S.; Romfo, C.M.; Nilsen, T.W.; Green, M.R. Functional recognition of the 3' splice site AG by the splicing factor U2AF35. *Nature* **1999**, *402*, 832–835. [[CrossRef](#)]
49. Merendino, L.; Guth, S.; Bilbao, D.; Martinez, C.; Valcarcel, J. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **1999**, *402*, 838–841. [[CrossRef](#)]
50. Zorio, D.A.; Blumenthal, T. Both subunits of U2AF recognize the 3' splice site in *Caenorhabditis elegans*. *Nature* **1999**, *402*, 835–838. [[CrossRef](#)] [[PubMed](#)]
51. Przychodzen, B.; Jerez, A.; Guinta, K.; Sekeres, M.A.; Padgett, R.; Maciejewski, J.P.; Makishima, H. Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. *Blood* **2013**, *122*, 999–1006. [[CrossRef](#)]
52. Brooks, A.N.; Choi, P.S.; de Waal, L.; Sharifnia, T.; Imielinski, M.; Saksena, G.; Pedamallu, C.S.; Sivachenko, A.; Rosenberg, M.; Chmielecki, J.; et al. A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. *PLoS ONE* **2014**, *9*, e87361. [[CrossRef](#)]
53. Yoshida, H.; Park, S.Y.; Sakashita, G.; Nariai, Y.; Kuwasako, K.; Muto, Y.; Urano, T.; Obayashi, E. Elucidation of the aberrant 3' splice site selection by cancer-associated mutations on the U2AF1. *Nat. Commun.* **2020**, *11*, 4744. [[CrossRef](#)]
54. Ilagan, J.O.; Ramakrishnan, A.; Hayes, B.; Murphy, M.E.; Zebari, A.S.; Bradley, P.; Bradley, R.K. U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Res.* **2015**, *25*, 14–26. [[CrossRef](#)]
55. Okeyo-Owuor, T.; White, B.S.; Chatrikhi, R.; Mohan, D.R.; Kim, S.; Griffith, M.; Ding, L.; Ketkar-Kulkarni, S.; Hundal, J.; Laird, K.M.; et al. U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. *Leukemia* **2015**, *29*, 909–917. [[CrossRef](#)]
56. Esfahani, M.S.; Lee, L.J.; Jeon, Y.J.; Flynn, R.A.; Stehr, H.; Hui, A.B.; Ishisoko, N.; Kildebeck, E.; Newman, A.M.; Bratman, S.V.; et al. Functional significance of U2AF1 S34F mutations in lung adenocarcinomas. *Nat Commun* **2019**, *10*, 5712. [[CrossRef](#)] [[PubMed](#)]
57. Park, S.M.; Ou, J.; Chamberlain, L.; Simone, T.M.; Yang, H.; Virbasius, C.M.; Ali, A.M.; Zhu, L.J.; Mukherjee, S.; Raza, A.; et al. U2AF35(S34F) Promotes Transformation by Directing Aberrant ATG7 Pre-mRNA 3' End Formation. *Mol. Cell* **2016**, *62*, 479–490. [[CrossRef](#)] [[PubMed](#)]
58. Palangat, M.; Anastasakis, D.G.; Fei, D.L.; Lindblad, K.E.; Bradley, R.; Hourigan, C.S.; Hafner, M.; Larson, D.R. The splicing factor U2AF1 contributes to cancer progression through a noncanonical role in translation regulation. *Genes Dev.* **2019**, *33*, 482–497. [[CrossRef](#)] [[PubMed](#)]

59. Smith, M.A.; Choudhary, G.S.; Pellagatti, A.; Choi, K.; Bolanos, L.C.; Bhagat, T.D.; Gordon-Mitchell, S.; Von Ahrens, D.; Pradhan, K.; Steeples, V.; et al. U2AF1 mutations induce oncogenic IRAK4 isoforms and activate innate immune pathways in myeloid malignancies. *Nat. Cell Biol.* **2019**, *21*, 640–650. [[CrossRef](#)] [[PubMed](#)]
60. Gamble, M.J.; Frizzell, K.M.; Yang, C.; Krishnakumar, R.; Kraus, W.L. The histone variant macroH2A1 marks repressed autosomal chromatin, but protects a subset of its target genes from silencing. *Genes Dev.* **2010**, *24*, 21–32. [[CrossRef](#)] [[PubMed](#)]
61. Zarrin, A.A.; Bao, K.; Lupardus, P.; Vucic, D. Kinase inhibition in autoimmunity and inflammation. *Nat. Rev. Drug Discov.* **2021**, *20*, 39–63. [[CrossRef](#)] [[PubMed](#)]
62. Zhou, Q.; Derti, A.; Ruddy, D.; Rakiec, D.; Kao, I.; Lira, M.; Gibaja, V.; Chan, H.; Yang, Y.; Min, J.; et al. A chemical genetics approach for the functional assessment of novel cancer genes. *Cancer Res.* **2015**, *75*, 1949–1958. [[CrossRef](#)] [[PubMed](#)]
63. Wadugu, B.A.; Nonavinkere Srivatsan, S.; Heard, A.; Alberti, M.O.; Ndonwi, M.; Liu, J.; Grieb, S.; Bradley, J.; Shao, J.; Ahmed, T.; et al. *U2af1* is a haplo-essential gene required for hematopoietic cancer cell survival in mice. *J. Clin. Investig.* **2021**, *131*, e141401. [[CrossRef](#)] [[PubMed](#)]
64. Eskens, F.A.; Ramos, F.J.; Burger, H.; O'Brien, J.P.; Piera, A.; de Jonge, M.J.; Mizui, Y.; Wiemer, E.A.; Carreras, M.J.; Baselga, J.; et al. Phase I pharmacokinetic and pharmacodynamic study of the first-in-class spliceosome inhibitor E7107 in patients with advanced solid tumors. *Clin. Cancer Res.* **2013**, *19*, 6296–6304. [[CrossRef](#)]
65. Bennett, C.F. Therapeutic Antisense Oligonucleotides Are Coming of Age. *Annu. Rev. Med.* **2019**, *70*, 307–321. [[CrossRef](#)] [[PubMed](#)]
66. Meylemans, A.; De Blecker, J. Current evidence for treatment with nusinersen for spinal muscular atrophy: A systematic review. *Acta Neurol. Belg.* **2019**, *119*, 523–533. [[CrossRef](#)]