Splicing factor mutations and cancer.

Yoshida, Kenichi; Ogawa, Seishi

CITATION:

ISSUE DATE:
2014-02-12

URL:
http://hdl.handle.net/2433/198762

This is the peer reviewed version of the following article: Yoshida, K. and Ogawa, S. (2014), Splicing factor mutations and cancer. WIREs RNA, 5: 445–459, which has been published in final form at http://dx.doi.org/10.1002/wrna.1222. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.  この論文は出版社版ではありません。引用の際には出版社版をご確認ご利用ください。; This is not the published version. Please cite only the published version.
Review

Splicing factor mutations and cancer

Kenichi Yoshida and Seishi Ogawa

Department of Pathology and Tumor Biology, Graduate School of Medicine, Kyoto University

Correspondence should be addressed to S.O. (sogawa-tky@umin.ac.jp).
**ABSTRACT**

Recent advance in high-throughput sequencing technologies unexpectedly revealed that somatic mutations of splicing factor genes frequently occurred in several types of hematological malignancies including myelodysplastic syndromes (MDS) and other myeloid neoplasms and chronic lymphocytic leukemia (CLL). Splicing factor mutations have been also reported in solid cancers such as breast cancer, pancreatic cancer, uveal melanoma and lung adenocarcinoma. These mutations were heterozygous and mainly affected *U2AF1* (*U2AF35*), *SRSF2* (*SC35*), *SF3B1* (*SF3B155* or *SAP155*) and *ZRSR2* (*URP*), which are engaged in the initial steps of RNA splicing including 3’ splice site recognition, and occurred in a largely mutually exclusive pattern, suggesting a common impact of these mutations on RNA splicing. Here we describe splicing factor mutations reported in various types of cancers, their functional/biological effect and their potential as therapeutic targets.

**Keywords:**
Splicing factor, mutation, cancer
INTRODUCTION

RNA splicing is the process of generating a mature mRNA (messenger RNA) from pre-mRNA during which exon-intron borders are recognized and the intervening intronic sequences are removed.[1] RNA splicing enables production of multiple different mRNA species by alternative splicing of exons or exon segments, which are found in nearly 95% of mammalian genes to generate tissue- and species-specific differentiation patterns.[2] RNA splicing is catalyzed by the spliceosome, which is composed of five small nuclear ribonucleoproteins (snRNPs) and a large number of auxiliary proteins.[1] On the other hand, there are two types of introns, termed U2 and U12 type; U2-type introns account for more than 99.5% of introns, whereas U12-type introns constitute less than 0.5% of introns and reside in 700-800 genes. These introns are spliced by the U2-type and U12-type dependent spliceosomes, which are referred to as ‘major’ and ‘minor’ spliceosomes, respectively.[3] The U1, U2, U4/U6, U5 snRNPs are the main components of the major spliceosomes and each snRNP consists of a small nuclear RNA (snRNA) and a varying number of associated proteins. Catalytic process of pre-mRNA splicing is a two-step process, during which spliceosomes are assembled in a stepwise fashion as splicing proceeds.[1, 2] In the first step, the 5’ splice site is cleaved and, in a coordinated reaction, the 5’ G residue of the intron is linked to an A residue near the 3’ splice junction. In the second step, the 3’ splice site is cleaved, producing ligated exons and excised "lariat" introns (Figure 1).

Recently, several lines of evidence suggest that RNA splicing is also associated with transcription and splicing factors also play an important role in the transcriptional regulation[2], although transcription from DNA to pre-mRNA and pre-mRNA splicing were previously considered as separated processes. For example, splicing factor SC35 (SRSF2) has been shown to affect transcriptional elongation in a gene specific manner.[4] In addition, close interactions between the alternative splicing mechanism and chromatin modifications also has been suggested.[2]

One of the recent, rather unexpected, findings in cancer genetics using high-throughput sequencing is frequent mutations of splicing factor genes found in myeloid dysplasias and to a lesser extent, in other hematopoietic neoplasms and solid cancers.[5-7] Among major targets of these mutations are U2AF1 (U2AF35), SRSF2 (SC35), SF3B1 (SF3B155 or SAP155) and ZRSR2 (URP), which are known to be commonly involved in the initial steps of RNA splicing. This review summarizes the splicing factor mutations in various types of cancers and discusses the mechanisms by which these mutations lead to carcinogenesis and their potential as therapeutic targets.

RNA splicing and disease
Alterations in RNA splicing have been implicated in various types of human diseases through disruption of either cis-acting elements within affected genes or trans-acting factors that are required for normal splicing or splicing regulation.[8] In general, disruption of cis-acting elements affects only a single gene, whereas defects in trans-acting factors could impact on multiple genes. Disruption of cis-acting elements is a well-known mechanism causing abnormal translation of protein; as many as 50% of disease-causing mutations are estimated to be cis-acting mutations that affect RNA splicing, leading to altered protein translation.[9] Other alterations of trans-acting elements, in which germline mutations affect major spliceosome components, including four components of U4/U5/U6 tri-snRNP complex, PRPF31, PRPF8, PRPF3 and RP9, have been recognized as the causes of autosomal dominant forms of retinitis pigmentosa.[10] Germline mutations of SF3B4, a component of U2 snRNP, have been shown to be the cause of Nager syndrome.[11] More recently, mutation of the minor spliceosome component, RNU4ATAC, which codes U4atac snRNA, is also reported as a cause of MOPD I (Taybi-Linder syndrome).[12, 13] Therefore mutations in ubiquitous spliceosome components actually cause only tissue-specific phenotypes. This cell type specificity is likely to result from sensitivity of some cell-specific pre-mRNAs to dysfunction of each spliceosome component.[8, 10] For example, pre-mRNA of RDS and FSCN2, whose expression are specific to photoreceptor, were shown to be sensitive to PRPF31 mutations.[14]

**RNA splicing and cancer**

Deregulated or abnormal RNA splicing has been considered as a common finding in cancer cells and causally related to carcinogenesis[15], although a more recent analysis revealed that the prevalence of alternatively spliced genes was actually slightly lower in tumors compared to normal tissues.[16] For example, a variety of cancer cell types show increased expression of the longer, anti-apoptotic splicing form of the Bcl-x gene, while another shorter isoform (Bcl-x(s)) promotes apoptosis.[15][17] Caspase-2 also generates both proapoptotic (caspase-2L) and anti-apoptotic isoforms (caspase-2S) by alternative splicing.[18] Similarly, short isoform of Fas (sFas) also have been shown to be antiapoptotic and elevated production of that has been observed in various types of cancers.[19] Although the molecular mechanism of these abnormal splicing is largely unknown, a recent report suggests a possible role of deregulated trans-acting factor. A member of the SR protein family, SRSF1 (SF2/ASF), is involved in RNA splicing and highly expressed in various types of human cancers, where an oncogenic role of the protein was indicated by transformation of SRSF1-overexpressed fibroblasts, although no oncogenic SRSF1 mutation has been thus far reported.[20] SRSF1 is also shown to control alternative splicing of tumor
suppressor BIN1, MNK1 and S6K1 kinases, which could contribute to the oncogenic action of SRSF1. However, more direct evidence suggesting oncogenic role of trans-acting splicing factors has been obtained through comprehensive analysis of somatic mutations in a variety of cancer types using massively parallel sequencing, in which multiple components of the RNA splicing machinery were shown to be recurrently mutated.

**Splicing factor mutations in hematological malignancies**

MDS and related disorders are a heterogeneous group of myeloid neoplasms characterized by varying degrees of cytopenia and transition to acute myeloid leukemia (AML).[21] Recently, several groups identified frequent somatic mutations of the RNA splicing machinery in MDS using whole exome/genome sequencing.[22-25] Mutations are heterozygous and highly frequent (45%-85%) in and specific to MDS and related myeloid neoplasms with myelodysplastic features, but less common (10%) in patients with de novo AML[22, 26] and myeloproliferative neoplasms (MPN) [22, 27] (Table 1) and rarely found in pediatric myeloid malignancies[28, 29] and lymphoid neoplasms[30] except for chronic lymphocytic leukemia (CLL).[31-33] Although more than 30 splicing factors are known to be mutated in at least one patient[22, 26, 32, 34, 35] (Table 1), 4 genes represent the major mutational targets, including *U2AF1, SRSF2, SF3B1* and *ZRSR2*. Most of the mutations affected the components that are involved in the initial steps of pre-mRNA splicing including 3’-splice site recognition and occurred in a mutually exclusive manner, suggesting a common impact of these mutations on RNA splicing in the pathogenesis of myelodysplasia (Figure 2). Another conspicuous feature of spliceosome mutations is the presence of mutational hot spots in *U2AF1, SRSF2* and *SF3B1* and the lack of nonsense/frameshift changes, indicating that functional consequence of these mutations should not be a simple loss-of-function, but accompanies some gain-of-function. As mentioned below, however, the frequency of mutations differs significantly across disease types; *SF3B1* mutations are extremely frequent in those MDS phenotypes showing ring sideroblasts compared to other MDS subtypes, while *SRSF2* mutations are more common in chronic myelomonocytic leukemia (CMML) (28-47%). Thus, despite common association of different splicing factor mutations to MDS phenotypes, the functional/biological impact of each mutation may not be completely identical, but probably reflecting discrete function of each spliceosome component, critically different, which might contribute to associated disease phenotypes. *SF3B1* mutations were also reported in 9.7-15% of CLL cases[31-33], in which these mutations tend to represent subclonal mutations and were significantly associated with poor clinical outcome.[36] The functional basis
for the common SF3B1 mutations in chronic myeloid neoplasms and lymphoid malignancies is still unknown but this may indicate a possibility of the common origin of these malignancies.

**Mutated spliceosome components**

**U2AF1**

U2AF1 (U2AF35: 35 kDa subunit of U2AF) is tightly bound to U2AF2 (U2AF65: 65 kDa subunit of U2AF), consisting of the U2 auxiliary factor (U2AF) heterodimer.[1] U2AF1 interacts with the AG dinucleotide of the 3’ splice site through its RNA recognition motif (RRM) and also play a role in juxtaposing the 5’ and 3’ splice sites by binding to SR proteins such as SRSF2 through its arginine-serine-rich (RS) domain. In addition, the pre-mRNA branch site and polypyrimidine tract (Py-tract) are bound by SF1/BBP and U2AF2, respectively and these interactions yield the spliceosomal E complex and play crucial roles in the recognition of the 3’ splice site of an intron. U2AF1 is mutated in 5-16% of MDS and 5-17% of CMML. More recently, recurrent U2AF1 mutations have been reported in hairy-cell leukemia-variant (HCVv).[37] Most of reported U2AF1 mutations were observed at two amino acid positions, S34 and Q157, which are located within evolutionally conserved zinc finger domains (Figure 3).[22]

**SRSF2**

SRSF2 (SC35) is a member of the SR protein family of proteins and thought to interact with exonic splicing enhancers (ESEs) thorough its RRM domain and recruits U2AF to the upstream 3’ splice site and U1 snRNP to the downstream 5’ splice site, giving positive effects on splice-site usage.[38] SRSF2 mutations were more frequent in CMML (28-47%). Most of reported mutations or insertion-deletions of SRSF2 exclusively involved P95 within the intervening sequence between the RRM and RS domains, indicating the gain-of-function nature of these mutations (Figure 3).[22] SRSF2 mutations are also reported in other myeloid malignancies including MPN transformed to secondary AML[39] and systemic mastocytosis.[40]

**SF3B1 (SF3b155)**

SF3B1 is a component of the SF3B complex, which consists of U2 snRNP together with the SF3A complex. U2 snRNP bind to a branch point sequence (BPS) via SF3B14 (p14) and also to U2AF2 via SF3B1 (SF3b155) to stabilize the interaction between the U2 snRNA and pre-mRNA’s BPS, leading to the establishment of the spliceosomal A complex.[1] Although SF3B1 mutations are found in a variety of myeloid malignancies, extremely high mutation rates have been reported in those MDS subtypes showing increased ring sideroblasts (>15% of all erythroblasts in bone marrow), including refractory anemia with ring sideroblasts (RARS) (64-83%) and refractory cytopenia with multilineage dysplasia and ring sideroblasts (RCMD-RS) (37-76%).[22, 25, 41, 42]
SF3B1 mutations were also common in RARS associated with marked thrombocytosis (RARS-T) (67-85%)[25, 41, 43], which has been classified as a unique form of myelodysplastic/myeloproliferative neoplasms (MDS/MPN) showing increased ring sideroblasts and an elevated platelet count. Moreover, even in other MDS subtypes, SF3B1 mutations were associated with significant increased, although less than 15%, ring sideroblasts, in which the presence of SF3B1 mutations showed very high positive and negative predictive values (~97.7% and 98.7%, respectively) for the presence of ring sideroblasts.[41] SF3B1 mutations clustered in several hot spots, such as K700, E622, R625, H662 and K666, located in HEAT domains that extend through exon12 to 15 (Figure 3).[22, 23]

**ZRSR (URP)**
ZRSR2 functions at 3' splice site of both U2- and U12-type introns[44] and participates in the formation of the splicing A complex on 3’ splice sites recognition. In addition, ZRSR2 is also engaged in the second step of splicing at U2-type intron.[44] ZRSR2 mutations were widely distributed along the entire coding region and most of them were nonsense mutations or frameshift changes, or involved splicing donor/acceptor sites that caused either a premature truncation or a large structural change of the protein, leading to loss-of-function (Figure 3).[22] ZRSR2 mutations are less frequent in MDS compared to mutations of U2AF1, SRSF2, and SF3B1, and were associated with refractory anemia with excess blast (RAEB) and the international prognostic scoring system (IPSS) int-1 or int-2 risk-groups.[45]

**Other splicing factors**
Other splicing factors have been also reported to be recurrently mutated in myeloid malignancies, although low frequencies of these mutations make their role in myeloid leukemogenesis being less established. These include an additional two components of splicing A complex (SF3A1 and U2AF2), PRPF40B, whose functions in RNA splicing is poorly defined[22], LUC7L2, a homologue of the putative RNA-binding protein Luc7, PRPF8 and other factors.[26, 32] LUC7L2 is thought to be involved in recognition of 5’ splice site and its inactivation mutations have been reported in MDS and secondary AML.[35, 46] PRPF8 (pre-mRNA processing factor 8) is a highly conserved component of both major and minor spliceosomes and its mutations also have been reported in de novo and secondary AML. [26, 35] Of note is that a recent study showed that truncating mutation of prpf8 impaired myeloid differentiation in Zebrafish, suggesting its essential role in hematopoietic development.[47]

**Clinical impact of spliceosome mutations**
Despite their common role in myeloid leukemogenesis anticipated from their non-overlapped distribution and functional implications, they seemed to have different impacts on clinical outcome in MDS and other myeloid malignancies. It seems that SF3B1 mutations are consistently associated with better overall survival (OS) and leukemia free survival (LFS) in MDS and RARS-T,[23, 41] although initial reports showed conflicting observations.[42, 48]

SRSF2 mutations are more common in advanced MDS subtypes (RAEB) and associated with a significantly shorter OS and LFS,[35, 49] while these mutations seemed to have no significant impact on prognosis among CMML patients.[50, 51] SRSF2-mutated AML transformed from MPN also have a shorter OS.[39] In addition, Lasho et al. and Vannucchi et al. reported SRSF2 mutations were associated with poor prognosis in primary myelofibrosis (PMF) patients.[52, 53] U2AF1 mutated MDS cases are associated with faster progression to secondary AML with no significant effects on OS or event free survivals[24][54], whereas another group reported a significantly shorter OS for U2AF1 mutated myeloid malignancies.[35] ZRSR2 mutations did not seem to affect prognosis.[55] The IPSS and the revised IPSS (IPSS-R) have been widely used for prognostication of MDS patients, however, this does not take into account for somatic mutations. Therefore, there are several attempts to integrate the mutational information of these genes into the prognostic model of MDS.[55, 56]

On the other hand, SF3B1 mutations on CLL were associated with poor prognosis[31, 32, 57] and more advanced fludarabine refractory disease.[33] Rossi et al. has proposed an integrated prognostic model of cytogenetic and molecular markers including SF3B1 mutations.[58]

**Splicing factor mutations in solid cancers**

Splicing factor mutations also have been reported in a wide variety of solid cancers, although generally at much lower frequencies. SF3B1 has been reported to be mutated in less than 4% of breast cancer, endometrial cancer, bladder cancer, pancreatic ductal adenocarcinoma and adenoid cystic carcinoma, where mutations were distributed at the same hot spots as reported in MDS and CLL, such as K700E and K666Q.[23, 59-62] SF3B1 mutations were also common in uveal melanoma, in which most of the mutations almost exclusively involved R625.[63, 64] In uveal melanoma, the R625 mutations were associated with a lower metastasis rate and good prognosis, whereas those at other nucleotide positions were found in uveal melanomas with disomy 3 that developed metastases.[64] The U2AF1 S34F mutation was found in 3% of the cases with lung adenocarcinoma.[65]
Splicing factor mutations and disease progression

Deep sequencing of mutant alleles in cancer specimens provides a clue to speculate the origin of splicing factor mutations during clonal evolution. [24] [41] Thus, in the majority of cases with myeloid malignancies analyzed, SF3B1 and U2AF1 mutations generally showed a high allelic burden in both CD34-positive cell and granulocyte derived DNA, suggesting that these mutations were acquired during early development of multipotent hematopoietic stem cells and transmitted to their myeloid progenitors. [24] [41] Recent high-throughput targeted sequencing of 94-111 genes in a large cohort of patients with MDS and related neoplasms also confirmed that splicing factor mutations were typically acquired early. [55, 56] Similarly, SRSF2 mutations, together with KIT D816V and TET2 mutations, also showed a high allelic burden in advanced systemic mastocytosis and could be early events. [40] Analysis of multiple mutations using single-cell-derived colonies provides more direct information about the order of mutations, which clarified that SRSF2 mutation is thought to be one of the initiating mutations of CMML pathogenesis together with TET2 mutation according to their higher prevalence. [66]

However, the founder role of the splicing factor mutation could be variable depending on cases. In the initial report in RARS-T, which is characterized by frequent coexisting mutations of SF3B1, and JAK2 and/or MPL, SF3B1 mutations showed higher allelic burden compared to coexisting JAK2/MPL mutations, suggesting that the JAK2 or MPL mutations occurred in a pre-existing SF3B1-mutated cell. [41] In a recent report, in contrast, higher allelic burden of JAK2 than SF3B1 mutant alleles indicated that SF3B1 mutations in a pre-existing MPN clone having a JAK2 mutation was implicated in the development of RARS-T. [67] Also, in the case of CLL, estimation of multiple driver mutations/chromosomal lesions using whole-exome sequencing and copy number analyses revealed that SF3B1 mutations represented subclonal lesions in an approximately half of the SF3B1-mutated CLL cases. [36] Similarly, U2AF1 mutations in HCLv were also thought to be subclonal mutations because of their lower allele frequencies. [37]

Functional impacts of splicing factor mutations

Despite strong genetic evidence of their involvement in clonal selection, the molecular mechanism of splicing factor mutations remains largely unknown. When overexpressed in HeLa cells, mutant U2AF1, but not the wild-type allele, induces severe suppression, rather than promotion, of cell proliferation with severe cell cycle arrest at G2/M transition and increased apoptosis. [22] Mutant U2AF1-transduced hematopoietic stem/progenitor cells show compromised repopulating capacity. [22] However, the interpretation of these observations may be complicated by the fact that transduction of oncogenic alleles potentially induces cellular stress that could result in cell...
senescence or induction of apoptosis.[68] While the hot-spot mutations always exists in heterozygous state, overexpression of these mutant alleles may rather mimic homozygous mutations in their expression level. Thus more sophisticated evaluation would be required to evaluate the oncogenic function of the splicing factor mutations, possibly using appropriate mouse model. On the other hand, some evidence suggested that other driver mutations or cytogenetic changes could contribute to the clonal dominance of cells with splicing factor mutations cooperatively. For example, in MDS, SRSF2 mutations frequently co-occurred with IDH2[56] and TET2 mutations.[55] Similarly, in CLL, SF3B1 mutations were frequently co-occurred with del(11q).[31]

Given the critical role of these splicing factors in pre-mRNA processing, one would speculate that their oncogenic actions are through de-regulated pre-mRNA splicing, leading to abnormal protein expression of relevant genes. In fact, when transduced in HeLa cells, the U2AF1 S34F mutant allele induces a variety of abnormal RNA splicing in a global manner, including exon skipping, unspliced intron, altered splice site recognition and altered exon usage ([22] and unpublished observation) (Figure 4). Abnormal exon skipping in mutant U2AF1-transduced cells was also demonstrated using a minigene reporter assay.[24] Of interest, an increased splicing activity was demonstrated for U2AF1 S34F mutant-transduced cells compared to wild-type U2AF1-transduced cells in a double-reported splicing assay[24], supporting the putative gain of function of this allele. Visconte et al. performed knock-down experiment of SF3B1 gene on K562 cell line and showed dysfunction of splicing exclusively affected U2-type introns as expected.[69]

Of note is mutations in ubiquitous spliceosome components actually cause specific phenotypes such as myelodysplasia and CLL. In this context, mutated splicosome components are ubiquitous but not necessarily essential for all introns or splice sites. For examples, AG-independent 3’ splice site which has a long stretch of pyrimidines can bind to U2AF2 without U2AF1.[70, 71] Therefore, one or more sensitive pre-mRNA to these mutations may cause such disease phenotypes. However, currently, little is known about the functionally relevant gene targets for abnormal splicing that would be caused by these mutations, if ever it play a major role in the development of MDS and other cancers. Przychodzen et al. analyzed RNA sequencing data of AML patients with and without U2AF1 mutation and identified U2AF1 mutant-specific splicing patterns, including exon skipping and exon retention in 35 genes, which are involved in functionally important pathways such as cell cycle and RNA processing.[72] Quesada et al. performed exon arrays and RNA sequencing of CLL samples with SF3B1 mutations and showed impaired RNA splicing activities of genes such as FOXP1, whose splicing variants are known to be associated with diffuse large B cell lymphoma (DLBCL).[32] However, the relevance of abnormal splicing of these
genes to leukemogenesis has not been confirmed and needs further investigation in the future. Papaemmanuil et al. performed expression analyses of CD34+ bone marrow cells derived from MDS patients and identified 94 genes which were down-regulated in SF3B1 mutated cases.[23] Among these, 7 of 50 most significantly down-regulated genes were related to mitochondrial functions, deregulation of which could be implicated in abnormal iron-metabolism that is a characteristic of RARS. Furney et al. performed RNA sequencing of uveal melanoma samples and revealed that SF3B1 mutations were associated with differential alternative splicing of ABCC5, UQCC and the long non-coding RNA (lncRNA) CRNDE.[73]

Finally, abnormal pre-mRNA splicing may not be the only mechanism that explains the oncogenic capacity of these splicing factor mutations. Several lines of evidence indicate close coupling between transcription and pre-mRNA splicing[4], in which possible interactions between epigenetic regulators and the splicing machinery.[74, 75] In addition, close associations between mutation types and disease subtypes could be better explained by the discrete function of each mutation other than the common function in pre-mRNA splicing. For example, sf3b1 is localized on the 5’ region of the Hox gene cluster and physically interacts with different components in polycomb repressive complex 1 (PRC1), such as Zfp144 and Rnf2, which seem to be functionally relevant interactions. In fact, sf3b1+/− mice showed posterior transformation of vertebrae with abnormal Hox gene expression but not with abnormal splicing of the relevant Hox gene transcripts, which is repressed by the background of trithorax mutation (mll mutation).[76] It has been demonstrated that SRSF2 was also involved in the regulation of DNA stability and that depletion of SRSF2 can lead to genomic instability.[77] Of interest in this context, regardless of disease subtypes, samples with SRSF2 mutations were reported to have significantly more mutations of other genes compared with U2AF1 mutations.[22]

**Therapeutic implication of splicing factor mutations**

Given the essential role of pre-mRNA splicing in cell homeostasis, inhibition of splicing factors using low molecular weight compounds that targets mutated spliceosome components could be an interest of possible therapeutics, because cancer cells carrying these splicing factor mutations are likely to have compromised pre-mRNA splicing and could be more sensitive to these compounds than unmutated cells (synthetic lethality) (Figure 5).[78] Homozygous deletion of sf3b1 allele results in lethality in the very early step during development.[76] Several compounds have been known to inhibit pre-mRNA splicing. Spliceostatin A is the methylketal derivative of natural product FR901464, who has anti-cancer activity, and inhibit RNA splicing through the connection with SF3B complex.[79] E7107 is the derivative of natural product pladienolide B with
anticancer effect and also inhibit RNA splicing through the binding with SAP130 (SF3B3). In addition, Sudemycin, which is the analog of FR901464, and Meayamycin, that is totally synthesized FR901464 analog, are also shown to inhibit RNA splicing and show anticancer activity. Although currently, there is no proof of concept for this therapeutic possibility shown experimentally, it would be an intriguing possibility to be tested.

CONCLUDING REMARKS
Splicing factors represent a novel class of mutational targets in cancer, especially in MDS and other hematopoietic malignancies, which were unexpectedly revealed by massively parallel sequencing of cancer genome. Most of these mutations involved the spliceosome components involved in the 3’ splice site recognition during early steps of RNA splicing, including SF3B1, U2AF1, SRSF2 and ZRSR2, indicating their oncogenic actions through de-regulated RNA splicing, although the exact mechanism of their oncogenic actions is largely unknown. A potential role of spliceosome inhibition in novel therapeutics for these intractable neoplasms is also an interest of the investigation in the future.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Disease type</th>
<th>Frequency</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF3B1</td>
<td>MDS</td>
<td>16.0%</td>
<td>Haematologica, 2013[85]</td>
</tr>
<tr>
<td></td>
<td>MDS, MDS/MPN</td>
<td>32.9%</td>
<td>Haferlach T, et al. Leukemia 2013[55]</td>
</tr>
<tr>
<td></td>
<td>CMML</td>
<td>6.0%</td>
<td>Itzykson R, et al. JCO, 2013[86]</td>
</tr>
<tr>
<td></td>
<td>RARS-T</td>
<td>84.8%</td>
<td>Broséus J, et al. Leukemia, 2013[43]</td>
</tr>
<tr>
<td></td>
<td>MPN</td>
<td>2.0%</td>
<td>Nangalia J, et al. NEJM, 2013[27]</td>
</tr>
<tr>
<td></td>
<td>PMF</td>
<td>6.5%</td>
<td>Lasho TL, et al. Leukemia, 2011[87]</td>
</tr>
<tr>
<td></td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td></td>
<td>CLL</td>
<td>15.0%</td>
<td>Wang L, et al. NEJM, 2011[31]</td>
</tr>
<tr>
<td></td>
<td>Fludarabine-refractory CLL</td>
<td>17.0%</td>
<td>Rossi D, et al. Blood, 2011[33]</td>
</tr>
<tr>
<td>Disease</td>
<td>Incidence</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Pancreatic ductal adenocarcinoma</td>
<td>4.0%</td>
<td>ICGC. Nature, 2012[61]</td>
<td></td>
</tr>
<tr>
<td>MDS</td>
<td>10.0%</td>
<td>Mian SA, et al. Haematologica, 2013[85]</td>
<td></td>
</tr>
<tr>
<td>MDS, MDS/MPN</td>
<td>7.7%</td>
<td>Haferlach T, et al. Leukemia 2013[55]</td>
<td></td>
</tr>
<tr>
<td>CMML</td>
<td>5.0%</td>
<td>Itzykson R, et al. JCO, 2013[86]</td>
<td></td>
</tr>
<tr>
<td>MPN</td>
<td>2.6%</td>
<td>Nangalia J, et al. NEJM, 2013[27]</td>
<td></td>
</tr>
<tr>
<td>Therapy-related MDS and AML</td>
<td>2.6%</td>
<td>Shih AH, et al. Haematologica, 2013[88]</td>
<td></td>
</tr>
<tr>
<td>de novo AML</td>
<td>4.0%</td>
<td>TCGA. NEJM, 2013[26]</td>
<td></td>
</tr>
<tr>
<td>Advanced systemic mastocytosis</td>
<td>5.6%</td>
<td>Schwaab et al. blood, 2013[40]</td>
<td></td>
</tr>
<tr>
<td>Childhood MDS, JMML</td>
<td>0.3%</td>
<td>Hirabayashi S, et al. Blood, 2012[28]</td>
<td></td>
</tr>
<tr>
<td>JMML</td>
<td>3.7%</td>
<td>Takita J, et al. Leukemia, 2012[29]</td>
<td></td>
</tr>
<tr>
<td>SRSF2</td>
<td>MDS, CMML, MPN</td>
<td>28.4%(CMML), 11.6%(MDS without RS), 5.5%(RARS, RCMD-RS), 6.5%(AML/MDS), 0.7%(AML), 1.9%(MPN)</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>----------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS, CMML</td>
<td>28%(CMML)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lower-risk MDS</td>
<td>13.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS</td>
<td>11.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS</td>
<td>12.4%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS</td>
<td>14.6%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS</td>
<td>13.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS, MDS/MPN</td>
<td>14%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MDS, MDS/MPN</td>
<td>17.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SRSF2</td>
<td>CMML 47.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CMML 40.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MPN 1.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMF 17.0%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>PMF 8.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Secondary AML derived from MPNs 18.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Therapy-related MDS and AML 7.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>de novo AML 0.5%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Advanced systemic mastocytosis 35.9%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Childhood MDS, JMML 0.3%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>JMML 3.7%</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZRSR2</td>
<td>MDS, CMML 8%(CMML), 7.7%(MDS without RS), 1.4%(RARS, RCMD-RS), 1.6%(AML/MDS), 1.9%(MPN)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZRSR2</td>
<td>MDS, CMML 0.8%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDS 11.1%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>MDS 3.1%</td>
<td></td>
</tr>
</tbody>
</table>


Nangalia J, et al. NEJM, 2013[27]


Vannucchi AM, et al. Leukemia 2013[53]


TCGA. NEJM, 2013[26]

Schwaab J, et al. blood, 2013[40]


Takita J, et al. Leukemia, 2012[29]
<table>
<thead>
<tr>
<th>Gene</th>
<th>Association</th>
<th>Frequency</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDS, secondary AML</td>
<td>2.5%</td>
<td>Walter MJ, et al. Leukemia, 2013[84]</td>
<td></td>
</tr>
<tr>
<td>MDS</td>
<td>1.0%</td>
<td>Mian SA, et al. Haematologica, 2013[85]</td>
<td></td>
</tr>
<tr>
<td>MDS, MDS/MPN</td>
<td>4.5%</td>
<td>Papaemmanuil E, et al. Blood, 2013[56]</td>
<td></td>
</tr>
<tr>
<td>MDS, MDS/MPN</td>
<td>7.6%</td>
<td>Haferlach T, et al. Leukemia 2013[55]</td>
<td></td>
</tr>
<tr>
<td>CMML</td>
<td>8.0%</td>
<td>Itzykson R, et al. JCO, 2013[86]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>de novo AML</td>
<td>0.5%</td>
<td>Quesada V, et al. Nat Gent, 2012[32]</td>
</tr>
<tr>
<td></td>
<td>CLL</td>
<td>1.9%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>SF1</td>
<td>MDS, CMML, AML</td>
<td>0.9%</td>
<td>Yoshida K, et al. Nature, 2011[22]</td>
</tr>
<tr>
<td>LUC7L2</td>
<td>MDS, CMML</td>
<td>0.8%</td>
<td>Makishima H, et al. Blood, 2012[35]</td>
</tr>
<tr>
<td></td>
<td>MDS, secondary AML</td>
<td>0.6%</td>
<td>Walter MJ, et al. Leukemia, 2013[84]</td>
</tr>
<tr>
<td>PRPF8</td>
<td>MDS, CMML</td>
<td>0.8%</td>
<td>Makishima H, et al. Blood, 2012[35]</td>
</tr>
<tr>
<td></td>
<td>de novo AML</td>
<td>2.0%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>SNRNP200</td>
<td>MDS, secondary AML</td>
<td>0.6%</td>
<td>Walter MJ, et al. Leukemia, 2013[84]</td>
</tr>
<tr>
<td></td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>CSTF2T</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>DDX1</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>DDX23</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>DHX32</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>HNRNPK</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>METTL3</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>PLRG1</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>POLR2A</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>PRPF3</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>RBMX</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>SRRM2</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>SRSF2</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>SRSF6</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>SUPT5H</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>TRA2B</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>Gene</td>
<td>Cancer Type</td>
<td>Frequency</td>
<td>Reference</td>
</tr>
<tr>
<td>--------</td>
<td>-------------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>U2AF1L4</td>
<td>de novo AML</td>
<td>0.5%</td>
<td>TCGA. NEJM, 2013[26]</td>
</tr>
<tr>
<td>CELF4</td>
<td>CLL</td>
<td>1.9%</td>
<td>Quesada V, et al. Nat Gent, 2012[32]</td>
</tr>
<tr>
<td>SFRS1</td>
<td>CLL</td>
<td>1.9%</td>
<td>Quesada V, et al. Nat Gent, 2012[32]</td>
</tr>
<tr>
<td>SFRS7</td>
<td>CLL</td>
<td>1.0%</td>
<td>Quesada V, et al. Nat Gent, 2012[32]</td>
</tr>
</tbody>
</table>

Abbreviations: JMML, juvenile myelomonocytic leukemia.
Figure Legends

**Figure 1**
**Pre-mRNA and major spliceosomes**
The U1, U2, U4/U6, U5 snRNPs are the main components of the major spliceosomes and each snRNP consists of a small nuclear RNA (snRNA) and a varying number of associated proteins. Assembly of spliceosomes begins with binding of U1 snRNP to the 5’ splice site of intron. Subsequently, U2 snRNP interacts with the BPS (A complex) and then U4/U6.U5 tri-snRNP are recruited, forming the B complex. After the release of U1 and U4 snRNP, activated B complex undergo the catalytic process of splicing (C complex). Catalytic process of pre-mRNA splicing is a two-step process, during which spliceosomes are assembled in a stepwise fashion as splicing proceeds.[1, 2] In the first step, the 5’ splice site is cleaved and the 5’ residue of the intron is linked to near the 3’ splice junction. In the second step, the 3’ splice site is cleaved, producing ligated exons and excised "lariat" introns. mRNA is released in the form of mRNA and associated protein (mRNP).

**Figure 2**
**Frequent splicing factor mutations in myelodysplasia**
Most of the mutations affected the components that are involved in the initial steps of pre-mRNA splicing including 5’-splice site recognition (LUC7L2) and 3’-splice site recognition (SF3B1, SRSF2, U2AF1, ZRSR2, U2AF2, SF3A1 and SF1) (upper panel). These mutations are almost mutually exclusive, suggesting a common impact of these mutations on RNA splicing in the pathogenesis of myelodysplasia. Distribution of mutations in 8 spliceosome genes, where diagnosis of each sample is shown by indicated colors (lower panel).

**Figure 3**
**Reported mutations of splicing factor genes including U2AF1, SRSF2, ZRSR2 and SF3B1 in myelodysplasia**
Mutations of U2AF1, SRSF2 and SF3B1 clustered in several hot spots (S34 and Q157 of U2AF1; P95 of SRSF2; E622, R625, H662, K666 and K700 of SF3B1), suggesting the gain-of-function nature of these mutations. On the other hand, ZRSR2 mutations were widely distributed along the entire coding region and most of them were nonsense mutations or frameshift changes, or involved splicing donor/acceptor sites that caused either a premature truncation or a large structural change of the protein, leading to loss-of-function.

**Figure 4**
**A variety of splicing alterations observed in mutant U2AF1-transduced HeLa cells**
U2AF1 S34F mutant allele induces a variety of abnormal RNA splicing in a global manner, including exon skipping, unspliced intron, altered splice site recognition and altered exon usage.

**Figure 5**
**Potential of splicing factor mutations as therapeutic targets (synthetic lethality)**
Cancer cells carrying splicing factor mutations (lower panel) are likely to have compromised pre-mRNA splicing and could be more sensitive to spliceosome inhibitors than unmutated cells (upper panel).
Fig 2

A Self-archived copy in
Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp
Fig 3

A Self-archived copy in Kyoto University Research Information Repository
https://repository.kulib.kyoto-u.ac.jp

U2AF1
(21q22.3)

Zn
UHM
Zn
RS

A26V
R156Q/H
R158H

S34F/Y
Q157R/P/G/ks

240 aa

SRSF2
(17q25.1)

RRM
RS

Y44H
R86_G93dup

221 aa

ZRSP2
(Xp22.1)

Zn
UHM
Zn
RS

E28X
S40X
I85T
R68sp
A96fs
E118fs
K230X
E148X
H191Y
L337fs
G251F
G256fs
N360fs
R295X
D345fs
N360fs
N32K
C332fs
C326R
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs
C326fs

483 aa

SF3B1
(2q33.1)

K182E
G347V
E491G

N626D
Y623C
R590K

HD
HD
HD
HD
HD
HD
HD

K700E/Q699_K700del
H662Q/D/Y
D781G

A1188V

E592K
E622D
R625L/C/G

* A744P
* G740R
** V701F
* K666N/T/E/R/Q/M

1304 aa
Fig 4

- Normal splicing
- Unspliced intron
- Altered splice site recognition
- Exon skipping
- Altered exon usage

https://repository.kulib.kyoto-u.ac.jp
Fig 5

Splicing factor
Non-Mutated Cells

Splicing intact

Spliceosome inhibition

Splicing

splicing

Splicing factor
Mutated Cells

splicing

敏度？

敏感？
References


