Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis

Hideki Makishima,1 Valeria Visconte,1 Hirotoshi Sakaguchi,2 Anna M. Jankowska,1 Sarah Abu Kar,1 Andres Jerez,1 Bartlomiej Przychodzen,1 Manoj Bupathi,1 Kathryn Guinta,1 Manuel G. Afable,1 Mikkael A. Sekeres,1 Richard A. Padgett,3 Ramon V. Tiu,1 and Jaroslaw P. Maciejewski1

1Department of Translational Hematology and Oncology Research, Taussig Cancer Institute, Cleveland Clinic, Cleveland, OH; 2Department of Pediatrics, Nagoya University Graduate School of Medicine, Nagoya, Japan; and 3Department of Molecular Genetics, Lerner Research Institute, Cleveland Clinic, Cleveland, OH

Myelodysplastic syndromes (MDSs) are chronic and often progressive myeloid neoplasms associated with remarkable heterogeneity in the histomorphology and clinical course. Various somatic mutations are involved in the pathogenesis of MDS. Recently, mutations in a gene encoding a spliceosomal protein, SF3B1, were discovered in a distinct form of MDS with ring sideroblasts. Whole exome sequencing of 15 patients with myeloid neoplasms was performed, and somatic mutations in spliceosomal genes were identified. Sanger sequencing of 310 patients was performed to assess phenotype/genotype associations. To determine the functional effect of spliceosomal mutations, we evaluated pre-mRNA splicing profiles by RNA deep sequencing. We identified additional somatic mutations in spliceosomal genes, including SF3B1, U2AF1, and SRSF2. These mutations alter pre-mRNA splicing patterns. SF3B1 mutations are prevalent in low-risk MDS with ring sideroblasts, whereas U2AF1 and SRSF2 mutations are frequent in chronic myelomonocytic leukemia and advanced forms of MDS. SF3B1 mutations are associated with a favorable prognosis, whereas U2AF1 and SRSF2 mutations are predictive for shorter survival. Mutations affecting spliceosomal genes that result in defective splicing are a new leukemogenic pathway. Spliceosomal genes are probably tumor suppressors, and their mutations may constitute diagnostic biomarkers that could potentially serve as therapeutic targets. (Blood. 2012;119(14):3203-3210)

Introduction

The myelodysplastic syndromes (MDSs) are characterized by clonal hematopoiesis, a variety of chromosomal abnormalities, bone marrow (BM) failure, and a propensity for evolution to acute myeloid leukemia (AML). Because of their often protracted course, MDSs recapitulate the stages of acquisition of a malignant phenotype, thereby offering insights into leukemogenesis. Although, traditionally, histomorphology-based schemes have been applied to subclassify patients with MDS, this approach is unlikely to be reflective of the underlying pathogenesis. Instead, a better molecular characterization of MDS on the genomic, epigenetic, and genetic levels probably more objectively diagnoses conditions, determines patients’ prognosis and, based on the underlying molecular defects, directs the application of targeted therapies. The emerging realization of the molecular diversity of MDS parallels the clinical and phenotypic heterogeneity of this disease. Moreover, molecular defects have the potential to serve as biomarkers and probably are more suitable for the identification of therapy targets and responsiveness/refractoriness to treatment.

The application of high-throughput molecular technologies, including high-density single nucleotide polymorphism arrays (SNP-As)3 and new sequencing technologies5,5 has led to the improved characterization of genomic lesions such as chromosomal aberrations and of somatic mutations affecting specific classes of genes,5 including signal transducers (eg, CBL),7-10 apoptotic genes (eg, TP53 and RAS),11-13 genes involved in epigenetic regulation of DNA (eg, DNMT3A, IDH1/2, and TET2),14-18 and histone modifiers (eg, EZH2, UTX, and ASXL1).19-24 Although some mutations in these factors are activating, most are loss-of-function or hypomorphic mutations and affect bona fide tumor suppressor genes (TSGs). Of greatest diagnostic effect are recurrent mutations found in specific genes. Most TSG mutations are not canonical, though, making systematic clinical diagnostics more difficult.

We and other groups recently identified a variety of new mutations present at distinct frequencies in subgroups of patients with MDS.25,26 Analogous to previously applied strategies that identified TET2, CBL, EZH2, and other mutations,7,18,21 we have identified mutations affecting, in a recurrent fashion, genes of the spliceosome machinery, using a combination of targeted search approaches and unbiased mass sequencing. Such mutations probably constitute a new class of TSGs ubiquitously involved in leukemogenesis.

Methods

Patient population

BM aspirates or blood samples were collected from 315 patients with MDS (n = 88). MDS/myeloproliferative neoplasms (MDS/MPNs; n = 66), MPN (n = 52), secondary AML (sAML; n = 54) that evolved from these conditions and primary AML (pAML; n = 55) seen at Cleveland Clinic or Nagoya University between 2003 and 2010 (Table I). Informed consent for sample collection was obtained according to protocols approved by the

Submitted December 22, 2011; accepted February 6, 2012. Prepublished online as Blood First Edition paper, February 9, 2012; DOI 10.1182/blood-2011-12-399774. The online version of this article contains a data supplement.

© 2012 by The American Society of Hematology

BLOOD, 5 APRIL 2012 • VOLUME 119, NUMBER 14 3203
**Table 1. Clinical characteristics of patients participating in this study**

<table>
<thead>
<tr>
<th>Category</th>
<th>No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDS</td>
<td>88</td>
</tr>
<tr>
<td>Low risk</td>
<td>58</td>
</tr>
<tr>
<td>MDS/MPN</td>
<td>77</td>
</tr>
<tr>
<td>RARS</td>
<td>38</td>
</tr>
<tr>
<td>High risk</td>
<td>20</td>
</tr>
<tr>
<td>RAEB 1/2*</td>
<td>30</td>
</tr>
<tr>
<td>MDS/MPN</td>
<td>66</td>
</tr>
<tr>
<td>CML</td>
<td>48</td>
</tr>
<tr>
<td>MPN</td>
<td>18 (11)</td>
</tr>
<tr>
<td>AML</td>
<td>52</td>
</tr>
<tr>
<td>Primary AML</td>
<td>55</td>
</tr>
<tr>
<td>Secondary AML</td>
<td>54</td>
</tr>
<tr>
<td>PV/PMF/ET</td>
<td>16</td>
</tr>
<tr>
<td>CMML</td>
<td>36</td>
</tr>
<tr>
<td>JMML</td>
<td>109</td>
</tr>
<tr>
<td>CML</td>
<td>55</td>
</tr>
<tr>
<td>MDS/MPN</td>
<td>54</td>
</tr>
</tbody>
</table>

MDS indicates myelodysplastic syndromes; RCUD, refractory cytopenia with unilineage dysplasia; RCMD, refractory cytopenia with multilineage dysplasia; MDS-U, MDS unclassifiable; RARS, refractory anemia with ring sideroblasts; RAEB, refractory anemia with excess blasts; MDS/MPN, MDS/myeloproliferative neoplasms; CMML, chronic myelomonocytic leukemia; acML, atypical chronic myeloid leukemia; JMML, juvenile myelomonocytic leukemia; RARS-T, RARS associated with marked thrombocytosis; PV, polycythemia vera; PMF, primary myelofibrosis; ET, essential thrombocythemia; and AML, acute myeloid leukemia.

*Eight cases included with therapy-related myeloid malignancies.

institutional review boards and in accordance with the Declaration of Helsinki. Diagnosis was confirmed and assigned according to World Health Organization classification criteria. Low-risk MDS was defined as patients having <5% myeloblasts. Patients with >5% myeloblasts constituted those with advanced disease. Serial samples were obtained for 38 patients. To study the germline genotype, immunoselected CD3-sorted cells were used. Cyto genetic analysis was performed according to standard banding techniques on the basis of 20 metaphases. Clinical parameters studied included age, sex, overall survival (OS), blood counts, and metaphase cytogenetics. The median follow-up of the cohort was 18 months (1-168 months).

**Cytogenetics and SNP-A analyses**

Technical details about sample processing for SNP-A were previously described. Affymetrix 250K and 6.0 Kit (Affymetrix) were used. A stringent algorithm was applied for the identification of SNP-A lesions. All other lesions were confirmed as somatic or germline by analysis of CD3-sorted cells.

**Whole exome sequencing**

Genomic DNA was extracted from BM or peripheral blood with the use of standard methods and subjected to agarose gel and optical density ratio tests to confirm the purity and concentration before Covaris fragmentation. Fragmented genomic DNA (0.5-2.5 µg) was tested for size distribution and concentration with the use of an Agilent 2100 Bioanalyzer (Agilent Technologies). Illumina libraries were made from qualified fragmented gDNA with the use of NEBNext reagents (New England Biolabs), and the resulting libraries were subjected to exome enrichment with the use of NimbleGen SeqCap EZ Human Exome Library Version 2.0 (Roche NimbleGen) according to the manufacturer’s instructions. Enriched libraries were tested for enrichment by quantitative PCR and for size distribution and concentration by an Agilent 2100 Bioanalyzer. The samples were then sequenced on an Illumina HiSeq2000 that generated paired end reads of 100 nucleotides. Paired BM mononuclear cells and CD3+ peripheral blood lymphocytes were used as germline controls. DNAnexus software (DNAnexus Inc; https://dnanexus.com) was used to visualize single nucleotide changes, insertions, and/or deletions at the gene, exon, and base pair levels. A rational bioanalytic algorithm was applied to identify candidate nonsynonymous alterations. Multiple steps were performed to reduce the false-positive rate within reported results. First, whole exome assembly was nonredundantly mapped with the reference genome hg19. Next, the analytic algorithm within DNAnexus called all the positions that vary from a reference genome. Each potential mutation was compared against databases of known SNPs, including Entrez Gene and the Ensembl Genome Browser. These candidate alterations were subtracted by the results of CD3+ peripheral blood DNA and subsequently validated with Sanger sequencing (see “Sanger sequencing analysis”). Moreover, spliceosome-associated gene mutations were screened with whole exome sequencing results available through The Cancer Genome Atlas (National Cancer Institute).

**Whole RNA deep sequencing**

Total RNA was extracted from BM mononuclear cells with the use of the Nucleospin RNA II Kit (Macherey-Nagel) with forward treatment. The integrity and purity of total RNA were assessed with Agilent Bioanalyzer. cDNA (1-2 µg) was generated with Clontech SmartPCR CDNA kit (Clontech Laboratories) from 100 ng of total RNA. cDNA was fragmented with Covaris, profiled with Agilent Bioanalyzer, and subjected to Illumina library preparation with NEBNext reagents (New England Biolabs). The quality and quantity and the size distribution of the Illumina libraries were determined with an Agilent Bioanalyzer. The libraries were then submitted for Illumina HiSeq2000 sequencing according to standard procedures. Paired-end 90-bp reads were generated and subjected to data analysis with the use of the platform provided by DNAnexus. DNAnexus software allowed visualization of reads derived from spliced mRNA and those that completely match the genome, including both sense and antisense.

**Statistical analysis of clinical data**

The Kaplan-Meier method was used to analyze survival outcomes (OS) of subgroups characterized by the presence of mutant versus wild-type (WT) variants of specific spliceosome-associated gene mutations with the log-rank test (JMP9; SAS Institute). Significance was determined at a 1-sided α level of 0.05.

**Results**

**Detection of somatic spliceosomal mutations in myeloid malignancies**

Initially, we performed whole exome sequencing of 15 index cases with various forms of myeloid malignancies and identified distinct somatic mutations in genes encoding components of the spliceosomal machinery (supplemental Table 1, available online). A heterozygous missense mutation in U2AF1 (Q157R) was found in a patient with refractory cytopenia with unilineage dysplasia (RCUD) and uniparental disomy of 2q (UPD2q; Figure 1A). Analysis of DNA from CD3+ cells showed a much lower frequency of the base change (2 of 15 reads),
Figure 1. Somatic spliceosomal gene (U2AF1, SF3B1, SRSF2, LUC7L2, PRPF8, and ZRSR2) mutations as detected by next-generation sequencing (NGS) and Sanger sequencing technologies. (A) With the use of an NGS-based whole exome sequencing analysis of whole BM DNA from a patient with refractory cytopenia with unilineage dysplasia (left), a mutation of U2AF1 (21q22.3) at position 44,514,777 (T>G) was detected in 13 of 18 reads. Analysis of DNA from CD3+-cells showed a much lower frequency of the base change (2 of 15 reads; right), highlighting the somatic nature of this alteration. The finding was confirmed by Sanger sequencing. Arrows and bars indicate the specific nucleotide and predicted codon, respectively. It should be noted that U2AF1 is expressed from the minus strand; therefore, the NGS presentation (top panels) is complementally reversed in comparison to the Sanger sequencing results (middle panels). This heterozygous somatic mutation results in the predicted nucleotide change 470 A>G in exon 6 of the coding region, which lead to the amino acid change Q157R in the second zinc finger domain. In the entire cohort, 27 mutations were observed in 26 patients, including a whole gene deletion. All 26 missense mutations were located in 1 of the 2 zinc finger domains (ZNFs); 2 residues, S34 or Q157, were frequently affected (bottom panels). RRM indicates RNA recognition motif. (B) With the use of an NGS analysis of a patient with CMML (middle left), a mutation of SF3B1 (2q33.1) at position 198,267,491 (C>T) was detected in 9 of 24 reads. The somatic nature of this alteration was confirmed by an analogous analysis of the CD3+-fraction, with the change being less frequent (2 of 23; middle right). The mutation was confirmed by Sanger sequencing (bottom). This heterozygous somatic mutation results in the nucleotide change 1866 G>T in exon 14 of SF3B1, resulting in the amino acid change E622D in the HSH155 domain. Analysis of the entire cohort identified mutations in 33 patients, including a case with a whole gene deletion. (C) Further screening by NGS led to the detection of a nonsense mutation (R27X) in LUC7L2 (top) which participates in the recognition of splice donor sites in association with the U1 snRNP spliceosomal subunit, and a missense mutation (M1307I) in PRPF8 (17p13.3; bottom) which is a large US snRNP-specific protein essential for pre-mRNA splicing. RS indicates serine/arginine-rich domain; US 2-snRNA bdg, US-snRNA binding site 2; and MPN, Mpr1p, Pad1p N-terminal domain. (D) Mutations of SRSF2, an arginine/serine-rich splicing factor, were detected in 29 cases among the entire cohort, including 2 whole gene deletions and a microdeletion within the gene (top). All mutations were heterozygous and affected P95. The somatic nature of the P95R mutation was confirmed with whole BM and T-cell rich fraction DNAs (bottom). (E) A nonsense mutation (W153X) was found in ZRSR2, another arginine/serine-rich splicing regulatory factor, in a case of CMML. ZRSR2 is located at Xp22.2, and the nonsense mutation was hemizygous in this male case (BM).
highlighting the somatic nature of this alteration. The 2 reads originated from contamination by a small amount of the mutated clone. The contamination was confirmed by flow cytometric analysis after magnetic bead sorting. Similarly, heterozygous, somatic SF3B1 mutations (E622D) were detected in a patient with refractory anemia with ring sideroblasts (RARS) associated with marked thrombocytosis (RARS-T; Figure 1B). In a patient with sAML, we identified a somatic mutation (M1307I) in PRPF8 and a heterozygous mutation (R27X) in LUC7L2, respectively (Figure 1C). On the basis of these findings, we screened other rationally selected genes within the spliceosomal machinery in patients with MDS and related disorders (n/H11005 120). With the use of this targeted approach, we identified mutations in ZRSR2 (W153X) and SRSF2 (P95R; Figure 1D-E). All of these mutations were somatic as confirmed by sequencing of the corresponding germline-derived DNA. Moreover, screening of whole exome sequencing results for AML available through the Cancer Genome Atlas showed the presence of mutations in genes encoding other components of the spliceosome, including mutations in HCF1 (P72L), SAP30 (T247I), SRSF6 (W123S), SON (M1024V), and U2AF26 (C116X; supplemental Figure 1). We found that among the initially screened cohort of patients with myeloid malignancies, mutations were highly recurrent in U2AF1, SF3B1, and SRSF2, whereas mutations of other spliceosomal genes were each detected in 1 of 120 cases, indicating a frequency of <1%. Although we identified 2 mutant copies of U2AF1 (21q22.3) in a sAML case with trisomy 21 (supplemental Figure 2), all other 19 cases of trisomy 21 screened for this mutation were negative. Moreover, no homozygous mutations were found in patients with UPD21q (n/H11005 8). Similarly, all mutations of SF3B1 and SRSF2 were heterozygous, and patients with somatic UPD2q33.1 or UPD17q25 (regions containing SF3B1 and SRSF2, respectively) did not harbor homozygous mutations of the associated genes.

Clinical associations and frequencies of spliceosomal mutations in myeloid malignancies

We subsequently screened a large cohort of patients (n/H11005 310) with MDS and related disorders (including the initially screened 120 cases) in a stepwise fashion to determine the frequency of spliceosomal mutations discovered in the index cases. On the basis of the initial screen, we noted that mutations in U2AF1, SF3B1, and SRSF2 were the most frequent. Consequently, we sequenced 120 cases for ZRSR2, LUC7L2, and PRPF8 and a total 310 for SF3B1, U2AF1, and SRSF2. All SF3B1 mutations were located in exon 14 or 15, with the K700 mutation being the most recurrent (Figure 1B). Similarly, all mutations in SRSF2 affected position P95 (Figure 1D). In contrast, mutations in U2AF1 affected exons 2 and 6, corresponding to the
2 zinc finger domains of this protein (Figure 1A; supplemental Figure 3). The extended cohort of patients was used to identify phenotype/genotype associations (supplemental Table 2). To avoid any misunderstanding, we included information as to which cases were analyzed by whole exome sequencing in supplemental Table 2. In addition, in supplemental Table 1, we clarified the mutational status of spliceosomal genes. In low-risk MDS, mutations of any 1 of these 3 genes were found in 39% of patients, and further analysis found that mutations in SF3B1 were highly associated with RARS. Among patients with MDS/MPN, SF3B1 mutations were not common in chronic myelomonocytic leukemia (CMML), but they were frequent in patients with RARS-T; thus, the presence of RS was found to correlate highly with SF3B1 mutations, irrespective of other clinical or morphologic features (Figure 2A). In contrast, U2AF1 mutations were most
frequent in the high-risk MDS/AML cohort (11%), whereas SRSF2 was most frequently mutated in MDS/MPN (24%) particularly in CMML (28%; Figure 2A).

Effect of spliceosomal mutations on clinical outcomes

Subsequently, we studied the effect of the most common spliceosomal mutations on clinical outcomes. We first analyzed the entire cohort of patients (Table 1) and determined the survival of patients in whom the 3 most common spliceosomal mutations were present. We included the information as to SF3B1 mutation status of 39 cases reported in our previous study.35 When 310 patients genotyped for these mutations were analyzed, the presence of SF3B1 mutations was associated with longer survival and U2AF1 mutations with shorter survival, whereas SRSF2 mutations had no effect on survival. We then analyzed the effect of these mutations in more clinically uniform subgroups to more precisely determine their clinical consequences. As expected, in sAML and pAML, because of overall poor prognosis, the presence of spliceosomal mutations did not further affect survival (supplemental Figure 4). However, in low-risk MDS, patients with SF3B1 mutations showed a tendency toward better prognosis (P = 0.09), whereas those with SRSF2 mutations had worse survival. In MDS/MPN, SRSF2 mutations were more common than U2AF1; however, U2AF1 mutations were associated with shorter survival (Figure 2B). When serial samples were analyzed, we found that an U2AF1 mutation detected at the sAML stage was present from initial MDS presentations, suggesting an ancestral origin of this mutation (supplemental Figure 5). Overall, SF3B1 mutations were less prevalent in patients with advanced forms of MDS, indicating that mutation of this factor does not contribute to progression (Figure 2A).

Effects of spliceosomal mutations on spliceosomal function

Conceptually, mutations of spliceosomal proteins could result in defective splicing, including altered splice site recognition, or altered alternative splicing. To determine the functional consequences of spliceosomal mutations on splicing, we performed whole mRNA deep sequencing. In the presence of functional spliceosomal machinery, sequencing reads are expected not to cross the intron/exon boundaries and therefore should not contain any intronic sequences. We analyzed RNA sequencing results in patients with mutations in U2AF1 (n = 3), SRSF2 (n = 2), SF3B1 (n = 2), and U2AF26 (n = 1), as well as in a healthy control and 1 patients with MDS with a WT configuration of these genes. No genome-wide increase in intron retention was observed in the patients with mutations. However, we found several specific genes in which the splicing pattern was altered. For instance, U2AF1 mutations were associated with defective splicing of intron 5 of TET2 at both splice sites (Figure 3A; supplemental Figure 6), whereas splicing of other TET2 introns were less affected. Another gene in which splicing was affected was RUNX1. At both the 3' and 5' splice sites of RUNX1 intron 6, unsliced reads were more frequent than spliced reads (Figure 3B). Such a splicing abnormality was more prominent in cases with SRSF2 mutations but not detected in the cases with SF3B1 mutations or WT spliceosomal genes. Similarly, U2AF26 mutations resulted in an alteration of RUNX1 splicing (supplemental Figure 7). Moreover, alternative splicing analysis showed that exon 9 of FECH was skipped in U2AF1 mutant cases but not in U2AF1 WT cases, including those with SF3B1 mutations (supplemental Figure 8).

Discussion

In recent years, a number of somatic mutations have been associated with various myeloid malignancies, including MDS.15-17,19,22 Systematic screening studies found that such mutations do not only occur as sole abnormalities but are often found in combination, probably contributing to the phenotypic heterogeneity found in MDS.25-26 During our search for new molecular lesions associated with MDS, we identified somatic mutations affecting various components of the spliceosomal machinery, similar to previous reports.35-39 Until these recent results were obtained, spliceosomal dysfunction was only rarely associated with malignant transformation.

In our studies in myeloid malignancies, the most commonly affected spliceosomal genes included SF3B1, U2AF1, and SRSF2. Although these mutations were common and found in a wide spectrum of myeloid diseases, some are strongly associated with specific phenotypic features, such as SF3B1 mutations with MDS or MDS/MPN with RS and SRSF2 mutations with CMML and advanced forms of MDS such as sAML and RA with excess blasts. Serial studies performed on patients from the initial diagnosis of low-risk MDS through subsequent transformation indicate that U2AF1 mutations may represent early ancestral events. A higher cross-sectional prevalence among advanced cases suggests that both U2AF1 and SRSF2 constitute high-risk defects. Indeed, U2AF1 and SRSF2 mutations were associated with worse survival in CMML and low-risk MDS, respectively. In contrast, SF3B1 mutations are associated with generally good prognosis, compatible with their association with RARS, which has a protracted clinical course.37,39 That these mutations were not overlapping and also not found in cases with other spliceosomal defects suggests that the effect of a spliceosomal mutation on myeloid malignancies might be inhibited by mutation in another factor or that they might not have a synergistic effect on leukemogenesis.

The prognostic effect of SF3B1 mutations has been investigated in several studies. Although initial reports (354 patients) suggested SF3B1 mutations convey significantly longer OS, leukemia-free survival (LFS), and event-free survival in their cohort,7 subsequent analysis of 323 MDS cases, patients carrying an SF3B1 mutation showed a significantly better OS and no effect on LFS in RARS and refractory cytopenia with multilineage dysplasia with ringed sideroblasts (RCMD-RS). The “protective effect” of SF3B1 mutation on OS was lost in multivariate analysis, including clinical variables such as International Prognostic Scoring System score.39 Our results in the whole cohort of 310 patients with MDS were generally consistent with those studies. However, we could not detect a significant difference in OS when subgroups with RS were analyzed separately, perhaps because of low number of WT cases. More recently, a large series of patients with MDS (N = 317) showed that the SF3B1 mutation status was not associated with time to AML progression or OS, regardless if all patients or only the subgroup of patients with RS were included in the analyses.40 Similar results were reported for myelo-fibrosis.41 When the survival benefit of SF3B1 mutations was examined in patients with MDS with increased RS, the presence of SF3B1 mutations was associated with better OS and LFS in univariate analysis; however, significance was completely accounted for by the World Health Organization categorization for morphologic risk, and no additional prognostic value from the presence or absence of SF3B1 mutations was observed when RARS and RCMD-RS were analyzed separately.42 In sum, most of
the available results suggest that SF3B1 mutations convey better survival in the cohort of whole myeloid malignancies, including RARS and RCMD-RS. However, such significant effects seem less in each background-matched specific disease phenotype by multivariate analysis, including, for example, International Prognostic Scoring System score.

Most of the mutations we detected were heterozygous, indicating that homozygous mutations may lead to cell death or that some of the functional consequences are related to dominant negative effects. Other mutations, including those in LUC7L2, ZRSR2, and PRPF8, appear to be less prevalent, and, because of the small number of positive cases within the already large cohort studied, we were unable to establish whether these mutations correlated with characteristic phenotypic features and survival. It is possible that the mutations lead to distinct phenotypes by affecting different stages of splicing and/or by causing defective splicing of specific gene transcripts. It is also possible that the phenotypic features are related to the transcriptional spectrum of cells in which the mutations occur. For instance, SF3B1 mutations may affect splicing of transcripts coding for proteins associated with iron handling in erythroid precursors, leading to RS. Such phenotypes would then be absent in myeloid cells lacking the corresponding transcripts.

The precise pathogenetic mechanisms associated with facilitation of clonal evolution remain unclear. Mechanistically, defective splicing of specific genes may have similar consequences to loss of function mutations through the retention of introns (supplemental Figure 9). We have performed deep sequencing of mRNA in hematopoietic cells derived from mutant cases and compared the results obtained in cells with a corresponding WT form of the gene. For U2AF1 and U2AF26 we found that accumulation of unspliced transcripts rather than abnormal alternative splicing is the main consequence of the mutations. We did not observe widespread defective splicing, but rather distinct changes in specific introns, as seen in RUNX1, that would functionally have the same effect as RUNX1 mutations. This splicing abnormality of RUNX1 is also observed in cases with SRSF2 mutations but not in SF3B1 mutants. U2AF1 and SRSF2 are frequently mutated in high-risk MDS and CMML whereby RUNX1 mutations are relevant, which explains our theory that spliceosomal mutations results in similar phenotype to the corresponding loss-of-function mutations. Such an indirect alteration of key proteins known to be involved in malignant transformation could explain the leukemogenic effects of spliceosomal mutations, in particular how spliceosomal mutations may result in a phenocopy of features associated with known TSG mutations.

Among the 3 most frequently mutated spliceosomal genes, U2AF1 mutations might result in loss of function. Further studies will clarify the whole alteration of mRNA splicing profiles induced by spliceosomal gene mutations.

Interestingly, spliceosome mutations, specific in SF3B1, were reported in chronic lymphocytic leukemia.44-46 The discovery of recurrent somatic mutations in various genes encoding spliceosomal proteins indicates that spliceosomal defects constitute an important and ubiquitous pathway in malignant transformation. Further mutational screening in other malignancies might uncover the pathogenesis of common or unique tumorgenesis in various tissues.

The most relevant question, both biologically and clinically, is which other gene mutations are associated with the mutation status of individual spliceosomal genes. According to our preliminary results, U2AF1 mutations seem to be most commonly associated with ASXL1 and TET2 mutations, whereas SF3B1 can occur in the context of RUNX1 mutations (data not shown). Some presentations by other groups at the 2011 meeting of the American Society of Hematology suggested a correlation between SF3B1 and DNMT3A mutations47,48 and a favorable survival effect of RUNX1 mutations in CMML with SRSF2.49 Because of the large numbers of diverse mutations, large studies will be needed to fully evaluate the correlation between the mutation status of spliceosomal genes and other genes in a proper fashion.

In sum, our studies found the widespread presence of mutations in genes involved in splicing in myeloid neoplasms. Such mutations may lead to distinct phenotypes and, because there is effect on survival, their detection may have future diagnostic utility.

Acknowledgments

The authors thank The Cancer Genome Atlas for access to the whole genome sequencing results described in the text.

This work was supported by the National Institutes of Health (grants RO1HL-082983, J.P.M.; U54 RR019391, J.P.M.; and K24 HL-077522, J.P.M.); by a grant from the AA & MDS International Foundation, and by the Robert Duggan Charitable Fund (J.P.M.).

Authorship

Contribution: H.M. designed and performed research, collected data, performed statistical analysis, and wrote the manuscript; V.V. performed research, collected data, and wrote the manuscript; H.S., A.M.J., S.A.K., A.J., and B.P. performed research; M.B., K.G., and T.V.V. performed research, collected data, and wrote the manuscript; J.P.M. designed research, contributed analytical tools, collected data, analyzed and interpreted data, and wrote the manuscript; M.G.A. collected data; M.A.S. and R.V.T. collected data, analyzed and interpreted data, and wrote the manuscript. Conflict-of-interest disclosure: The authors declare no competing financial interests.

Correspondence: Jaroslaw P. Maciejewski, Taussig Cancer Institute/R40, 9500 Euclid Ave, Cleveland, OH 44195; e-mail: maciejj@ccf.org.

References


Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis

Hideki Makishima, Valeria Visconte, Hirotoshi Sakaguchi, Anna M. Jankowska, Sarah Abu Kar, Andres Jerez, Bartlomiej Przychodzen, Manoj Bupathi, Kathryn Guinta, Manuel G. Afable, Mikkael A. Sekeres, Richard A. Padgett, Ramon V. Tiu and Jaroslaw P. Maciejewski