**Cancer Cell**

*Article*

**SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition**

**Graphical Abstract**

**Highlights**

- *Srsf2P95H/wild-type* mice develop myelodysplasia but *Srsf2*-deficient mice do not
- Proline 95 mutations change the RNA binding specificity of SRSF2
- Mutant SRSF2 promotes an isoform of *EZH2* that undergoes nonsense-mediated decay
- Restoring EZH2 expression partially rescues hematopoiesis in *Srsf2* mutant cells

**Authors**

Eunhee Kim, Janine O. Ilagan, ..., Robert K. Bradley, Omar Abdel-Wahab

**Correspondence**

rbradley@fhcrc.org (R.K.B.), abdelwao@mskcc.org (O.A.-W.)

**In Brief**

Kim et al. report that myelodysplastic syndrome-associating *SRSF2* mutations alter SRSF2’s sequence-specific RNA binding activity, leading to recurrent mis-splicing of key hematopoietic regulators such as *EZH2* and impaired hematopoietic differentiation.

**Accession Numbers**

GSE65349

Kim et al., 2015, Cancer Cell 27, 617–630

May 11, 2015 ©2015 Elsevier Inc.

http://dx.doi.org/10.1016/j.ccell.2015.04.006


SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition

Eunhee Kim, Janine O. Ilagan, Yang Liang, Gerrit M. Daubner, Stanley C.-W. Lee, Aravind Ramakrishnan, Yue Li, Young Rock Chung, Jean-Baptiste Miclo, Michele E. Murphy, Hana Cho, Min-Kyung Kim, Ahmad S. Zebari, Shlomzion Aumann, Christopher Y. Park, Silvia Buonamici, Peter G. Smith, H. Joachim Deeg, Camille Lobry, Iannis Aifantis, Yongo Modir, Frederic H.-T. Allain, Stephanie Halene, Robert K. Bradley, and Omar Abdel-Wahab

1Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
2Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
3Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
4Hematology, Yale Comprehensive Cancer Center and Department of Internal Medicine, Yale University School of Medicine, New Haven, CT 06520, USA
5Institute for Molecular Biology and Biophysics, ETH, 8093 Zürich, Switzerland
6Clinical Research Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA
7Division of Medical Oncology, School of Medicine, University of Washington, Seattle, WA 98109, USA
8Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA
9Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
10H3 Biomedicine, Cambridge, MA 03129, USA
11Institut National de la Sante et de la Recherche Medicale (INSERM) U1009, Institut Gustave Roussy, 94805 Villejuif, France
12Université Paris-Sud, 91400 Orsay, France
13Howard Hughes Medical Institute and Department of Pathology, New York University School of Medicine, New York, NY 10016, USA
14Department of Medicine, University of Cambridge, MRC Laboratory of Molecular Biology, Cambridge CB2 0QH, UK
15Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
16Co-first author
17Co-senior author
*Correspondence: rbradley@fhcrc.org (R.K.B.), abdellwao@mskcc.org (O.A.-W.)
http://dx.doi.org/10.1016/j.ccel.2015.04.006

SUMMARY

Mutations affecting spliceosomal proteins are the most common mutations in patients with myelodysplastic syndromes (MDS), but their role in MDS pathogenesis has not been delineated. Here we report that mutations affecting the splicing factor SRSF2 directly impair hematopoietic differentiation in vivo, which is not due to SRSF2 loss of function. By contrast, SRSF2 mutations alter SRSF2’s normal sequence-specific RNA binding activity, thereby altering the recognition of specific exonic splicing enhancer motifs to drive recurrent mis-splicing of key hematopoietic regulators. This includes SRSF2 mutation-dependent splicing of EZH2, which triggers nonsense-mediated decay, which, in turn, results in impaired hematopoietic differentiation. These data provide a mechanistic link between a mutant spliceosomal protein, alterations in the splicing of key regulators, and impaired hematopoiesis.

INTRODUCTION

Somatic mutations in genes encoding components of the spliceosome have been identified in a spectrum of human malignancies, including ~60% of patients with myelodysplastic syndromes (MDS) (Bejar et al., 2012; Papaemmanuil et al., 2013; Yoshida et al., 2011). These mutations occur most commonly in SF3B1 (Splicing Factor 3b Subunit 1), SRSF2 (Serine/arginine-Rich Splicing Factor 2), and U2AF1 (U2 Small Nuclear RNA Auxiliary Factor 1) and almost always as

Significance

Frequent somatic mutations affecting components of the spliceosome have been identified in hematologic malignancies; however, the functional role of these mutations is not known. Here we identify that commonly occurring mutations in the spliceosomal gene SRSF2 impair hematopoietic differentiation and promote myelodysplasia by altering SRSF2’s preference for specific exonic splicing enhancer motifs. This results in consistent mis-splicing in a manner that promotes the expression of abnormal isoforms of a number of key hematopoietic regulators, some of which have been linked previously to leukemogenesis (including BCOR and EZH2). These data provide a mechanistic basis for the enrichment of spliceosomal mutations in myelodysplasia and identify altered RNA recognition as an important driver of leukemogenesis.
heterozygous missense mutations that are mutually exclusive (Papaemmanuil et al., 2011; Wang et al., 2011; Yoshida et al., 2011). Although the genetic data in MDS suggest that these alterations are critical to disease pathogenesis, it remains unknown how these mutations contribute to MDS and whether they are sufficient to induce MDS.

Recent studies have suggested that mutations in the splicosomal gene U2AF1 after RNA splicing (Brooks et al., 2014; Graubert et al., 2012; Ilagan et al., 2015; Przychodzen et al., 2013; Quesada et al., 2012), and studies of gene expression in primary patient samples with and without U2AF1 mutations have been performed in an effort to identify downstream mis-spliced genes that might contribute to abnormal hematopoiesis (Brooks et al., 2014; Graubert et al., 2012; Ilagan et al., 2015). However, it remains unknown how these mutations contribute to hematopoietic transformation. To date, no studies have investigated the in vivo effects of splicosomal mutations expressed from the endogenous locus in the correct cellular context, which might allow delineation of how these alleles contribute to MDS pathogenesis.

To test whether splicosomal gene mutations are sufficient to drive MDS and determine how altered RNA splicing contributes to transformation in vivo, we studied the biological and transcriptional consequences of mutations in SRSF2. SRSF2 mutations occur in 20%–30% of MDS and ~50% of chronic myelomonocytic leukemia (CML) patients (Papaemmanuil et al., 2013; Yoshida et al., 2011). SRSF2 is a member of the serine/arginine-rich (SR) protein family that contributes to both constitutive and alternative splicing by binding to exonic splicing enhancer (ESE) sequences within pre-mRNA through its RNA recognition motif domain (RRM) (Graveley and Maniatis, 1998; Liu et al., 2000; Schaal and Maniatis, 1999; Zahler et al., 2004). SRSF2 mutations are consistently associated with adverse outcomes among MDS and acute myeloid leukemia (AML) patients (Papaemmanuil et al., 2013; Vannucci et al., 2013; Zhang et al., 2012). Despite the clinical importance of SRSF2 mutations, to date there have been no studies of the functional impact of SRSF2 mutations on hematopoiesis or splicing. Here we studied the biological and transcriptional effects of somatic expression of the common SRSF2P95H mutation in the hematopoietic compartment.

RESULTS

Srsf2P95H Mutant Mice Develop MDS, a Phenotype Distinct from Mice with Heterozygous or Homozygous Loss of Srsf2

Given the genetic heterogeneity of primary patient samples as well as the fact that stable overexpression of splicosomal proteins, even in wild-type (WT) form, is poorly tolerated (Lareau et al., 2007), we first generated a murine model for conditional expression of the commonly occurring SRSF2P95H mutation from the endogenous murine locus of Srsf2 (Figure 1A; Figures S1A and S1B). Mice heterozygous for the Srsf2P95H allele (Srsf2P95H/WT) were crossed to Mx1-cre transgenic mice (Kühn et al., 1995) on a C57BL/6 background to allow for inducible expression of Cre recombinase following intraperitoneal injection of polyinosine-polyribosine (pIpC) (12 µg/g every other day for 3 days by injection, as described previously [Moran-Cru-
Given that mutations in SRSF2 occur as early genetic events in MDS pathogenesis (Papaemmanuil et al., 2013) and that MDS is characterized by expansion of HSPCs, we next examined HSPC numbers and function in Srsf2P95H mice. Analysis of CD45.2+ HSPC subsets from Mx1-cre Srsf2P95H/WT mice and littermate controls 14 weeks after plpC injection revealed expansion of lineage-negative Sca1+ c-Kit+ (LSK) and restricted hematopoietic progenitor cells (LSK CD48+ CD150+; hematopoietic

Figure 1. Conditional Expression of Srsf2P95H Results in Myeloid Dysplasia, a Phenotype Distinct from Heterozygous or Homozygous Loss of Srsf2

(A) Depiction of the Srsf2P95H allele. (B) RNA-seq of LSK cells in Mx1-cre Srsf2WT and Mx1-cre Srsf2P95H/WT mice. (C–E) White blood cell (WBC) count (C), hemoglobin (Hb) (D), and MCV (E) of red blood cells of CD45.1 recipient mice 18 weeks following noncompetitive transplantation of bone marrow from CD45.2+ Mx1-cre Srsf2WT, Mx1-cre Srsf2fl/WT, and Mx1-cre Srsf2P95H/WT mice (n = 10 mice/genotype for all genotypes except Mx1-cre Srsf2fl/WT, where n = 5; plpC was administered to recipient mice 4 weeks following transplantation). (F and G) H&E staining of femurs (scale bars, 50 μm) (F) and peripheral blood smears (G) from Mx1-cre Srsf2WT, Mx1-cre Srsf2fl/fl, or Mx1-cre Srsf2P95H/WT mice (scale bars, 10 μm). A representative neutrophil (left) and erythroid precursor (right) is shown for Srsf2 WT and KO mice. Mx1-cre Srsf2P95H cells were marked by hypolobated and hypogranulated neutrophils (left two photos) and nuclear irregularities as well as cytoplasmic vacuolization and blebbing of erythroid precursors (right two photos). Error bars represent mean ± SD. ***p < 0.001; ****p < 0.0001. See also Figure S1.
progenitor cell fraction 2 [HPC-2; Oguro et al., 2013] in mutant mice relative to controls (Figures 2A and 2B). A similar LSK expansion was seen in spleens of Srsf2P95H mutant mice (although splenomegaly was not observed up to 20 weeks post-plpC) (Figures S2A and S2B). Because the detection of increased HSPCs in Srsf2P95H mutant mice appeared paradoxical given the decreased peripheral blood counts in these same mice, we next examined the cell cycle kinetics and apoptosis of Srsf2 mutant HSPCs. Indeed, Srsf2P95H LSK cells were characterized by an increase in the proportion of cells in S-phase as well as in early apoptosis (Figures 2C–2E). Despite HSPC expansion in Srsf2P95H mutant mice, purified LSK cells from mice with a homozygous Srsf2 deletion or heterozygous Srsf2P95H mutation had similarly impaired colony formation and serial re-plating capacity in vitro (Figure S2C).

To assess the functional effects of Srsf2 alterations on HSC self-renewal in vivo, we next compared Srsf2 heterozygous KO, homozygous KO, and heterozygous P95H mutant mice in competitive transplantation assays (Figure 3A). Equal numbers of BM MNCs from CD45.1 WT mice and CD45.2 Mx1-cre Srsf2 WT, Mx1-cre Srsf2fl/fl WT, Mx1-cre Srsf2fl/fl, or Mx1-cre Srsf2P95H/WT mice were transplanted into lethally irradiated CD45.1 mice, followed by plpC injection 4 weeks later. An assessment of peripheral blood chimerism monthly thereafter revealed a complete loss of CD45.2 chimerism in mice transplanted with Mx1-cre Srsf2fl/fl cells and a significant decrease in chimerism in mice transplanted with Mx1-cre Srsf2P95H/WT cells (Figure 3B; Figures S3A and S3B). However, an analysis of BM LSK chimerism 18 weeks post-transplant revealed an increase in CD45.2+ HSPCs derived from Srsf2P95H mice relative to other
groups and a near complete absence of CD45.2+ HSPCs from Srsf2fl/fl mice (Figures 3C and 3D; Figure S3C). Serial competitive transplantation of whole bone marrow from Srsf2P95H, Srsf2 heterozygous KO, and Srsf2 WT primary recipient transplanted mice continued to reveal an impaired reconstitution capacity of Srsf2P95H mutant mice relative to Srsf2 heterozygous KO or control mice (Figure S3D). Of note, colony assays and competitive transplantation experiments were performed using multiple genotypes of control mice (Cre-negative Srsf2P95H mice as well as Mx1-cre Srsf2 WT mice; Figures S2C and S3E) to control for any possible confounding effect of Cre expression or the presence of the unexcised P95H knockin allele.

The fact that Mx1-cre Srsf2P95H/WT mice had an increase in HSPCs despite impaired formation of mature peripheral blood cells suggested that mutant Srsf2 was associated with impaired HSPC differentiation. Flow cytometric analysis of mature and intermediate precursor cell subsets in Srsf2P95H mice was therefore performed to identify the stage of impaired hematopoiesis. This revealed that peripheral leukopenia was predominantly due to decreased peripheral blood B cells, evident at all stages of B lymphopoiesis following the transition of pre-proB to proB cells, in Srsf2P95H mice relative to controls (Figures S3F and S3G). Moreover, immunophenotypic analysis of intermediate hematopoietic progenitors (Pronk et al., 2007) revealed deficits in early erythroid progenitors in Srsf2P95H mice relative to controls, initiating at the pre-MegE and pre-colony-forming units, erythroid, stages (Figures S3H and S3I). Given prior data showing that homozygous deletion of Srsf2 resulted in defective T cell maturation and CD45 splicing (Wang et al., 2001), we also examined thymic T cell differentiation and CD45 isoform expression in Srsf2P95H mice relative to controls (Figures S3J and S3K). This revealed no effect of Srsf2P95H mutation on thymic
T cell maturation or protein expression of the specific CD45 isoforms identified previously to be downregulated with homozygous deletion of Srsf2 (Wang et al., 2001).

Collectively, the biological analysis of Srsf2P95H mutant mice identified phenotypes distinct from mice with a partial or complete loss of Srsf2, suggesting that SRSF2 mutations alter SRSF2’s normal function rather than resulting in haploinsufficiency or a dominant-negative function. Of note, despite the impaired hematopoietic differentiation, increase in HSPC subsets, and morphologic dysplasia in Srsf2P95H/WT mice, no Srsf2P95H mutant mice developed acute myeloid leukemia in up to 70 weeks of observation.

**SRSF2 Mutations Are Associated with Global Alterations of Gene Expression and Splicing**

We next sought to identify the transcriptional and post-transcriptional alterations caused by SRSF2 mutations through RNA-seq of purified LSK and myeloid progenitor (MP, lineage-negative Sca1- c-Kit+) populations. This was performed 4 weeks after plpC administration. In an unsupervised cluster analysis based on coding gene expression, samples clustered first by cell type and then by genotype (Figure S4A). The expression of several hematopoietic regulators was altered in Srsf2P95H mutant cells, including upregulation of Gfi1, Cebpe, and Hoxb2 in LSK cells; downregulation of Gata1 and Gata2 in MP cells; and downregulation of Cdkn1a in both populations. In addition, we observed preferential down- versus upregulation of the expression of coding genes in Srsf2 mutant cells relative to the WT (Figures S4B and S4C). Gene ontology (GO) analysis revealed an enrichment for the downregulation of genes in both LSK and MP cells involved in the regulation of cell cycle, proliferation, differentiation, and apoptosis (upregulated genes were not enriched for these processes; Figure S4-D).

To identify changes in splicing driven by SRSF2 mutations that might contribute to disease, we augmented our mouse data with RNA-seq data from primary CMML (n = 13, 3 with SRSF2 mutation) and AML (n = 9, 5 with SRSF2 mutation) patient samples (Table S1) as well as K562 cells ectopically expressing an empty vector or a single allele of SRSF2 (WT, P95H, P95L, and P95R). In all sequenced patients with SRSF2 mutations, the WT and mutant alleles were expressed at similar levels (Table S1), as was the case for the Srsf2P95H mouse cells (Figure 1). Similarly, SRSF2's normal function rather than resulting in haploinsufficiency or a dominant-negative function. Of note, despite the impaired hematopoietic differentiation, increase in HSPC subsets, and morphologic dysplasia in Srsf2P95H/WT mice, no Srsf2P95H mutant mice developed acute myeloid leukemia in up to 70 weeks of observation.

**Figure 4. SRSF2 Mutations Alter Exonic Splicing Enhancer Preference**

(A) Scatterplot of cassette exon inclusion in K562 cells expressing empty vector or SRSF2P95R. Percentages indicate the percent of alternatively spliced cassette exons with increased or decreased inclusion. Red and blue dots represent individual cassette exons that are promoted or repressed in SRSF2P95R versus empty vector cells, respectively. Promoted and repressed cassette exons are defined as those whose inclusion levels are increased or decreased by ≥5% with a Bayes factor of ≥5, as estimated by Wagenmakers’ framework (Wagenmakers et al., 2010).

(B) Enriched (right) and depleted (left) k-mers in cassette exons promoted versus repressed in SRSF2P95R versus WT cells.

(C) Scatterplot of cassette exon inclusion in TF-1 cells following transfection with a siRNA against SRSF2 or a control non-targeting siRNA (KD, knockdown). Percentages indicate the percent of alternatively spliced cassette exons with increased or decreased inclusion.

(D) Enriched (right) and depleted (left) k-mers in cassette exons promoted versus repressed in SRSF2 KD versus control cells.

(E) Mean enrichment of all variants of the SSNG motif in cassette exons promoted versus repressed in TF-1 cells following SRSF2 knockdown and K562, LSK, and MP cells expressing WT or mutant SRSF2. Error bars indicate 95% confidence intervals estimated by bootstrapping.

(F) Relative frequency of CCNG and GGNG motifs in cassette exons promoted versus repressed by SRSF2 mutations in LSK and MP cells (top), K562 cells (left), and primary AML and CMML samples with or without SRSF2 mutations (right) (the sample numbers correspond to the patient identifiers in Table S1). Shading indicates 95% confidence interval by bootstrapping. The schematic illustrates a portion of a metagene containing the differentially spliced cassette exon. From left to right, the features are the upstream exon (gray box) and intron (black line), the cassette exon (black box, vertical dashed lines), and the downstream intron (black line) and exon (gray box). Horizontal axis, genomic coordinates defined with respect to the 5' and 3' splice sites where 0 is the splice site itself. Vertical axis, relative frequency of the indicated motifs over genomic loci containing cassette exons promoted versus repressed by SRSF2 mutations (log scale).

See also Figure S4.
C or G) and efficiently recognizes both CCNG and GGNG (Daubner et al., 2012). Therefore, our ab initio analysis suggested that mutations affecting the P95 residue may alter SRSF2’s ability to recognize variants of its normal SSNG motif.

To further explore this hypothesis, we compared the relative enrichment of all four SSNG variants in cassette exons that were differentially spliced upon depletion of SRSF2, overexpression of WT SRSF2, or expression of mutant SRSF2. SRSF2 depletion—achieved by knockdown of endogenous SRSF2 in the absence of mutant protein expression (Figure S4K)—caused preferential skipping of cassette exons, consistent with SRSF2’s canonical role in promoting exon recognition (Figure 4C). Ab initio motif analyses identified both C- and G-rich variants of the SSNG motif as the most enriched motifs in cassette exons that were repressed following SRSF2 depletion (Figure 4D). Quantitation of the enrichment of each SSNG variant revealed that all were associated with exon repression following knockdown. In contrast, overexpression of WT SRSF2 was associated with enrichment of each SSNG variant (Figure 4E). These data suggest that different SSNG variants function as equally efficacious SRSF2-dependent ESEs, consistent with SRSF2’s in vitro binding specificity (Daubner et al., 2012). In contrast, K562 cells as well as LSK and MP cells expressing mutant Srsf2 exhibited enrichment for CCNG and depletion for GGNG in exons that were promoted versus repressed (Figure 4E).

To test whether this motif enrichment and depletion was due to ESE activity, we computed the spatial distribution of CCNG and GGNG motifs across genomic loci containing cassette exons that were promoted or repressed in association with SRSF2 mutations. CCNG and GGNG were, respectively, enriched and depleted specifically over cassette exons and not over the flanking introns or exons. We observed similar motif preferences and distributions in patient transcriptomes (Figure 4F). Because CCNG/GGNG motifs were not consistently enriched/depleted in introns flanking differentially spliced cassette exons, and because we were unable to identify enriched motifs with ab initio searches in introns, we conclude that differential cassette exon splicing is likely due primarily to altered recognition of exonic motifs. Together, these data reveal spatially restricted enrichment of specific ESEs in association with SRSF2 mutations and suggest that SRSF2 mutations cause alteration rather than loss of normal ESE recognition activity.

**SRSF2 Proline 95 Mutations Alter RNA Binding Specificity by Changing the Conformation of Both RRM Termin**

We next tested whether this association between SRSF2 mutations and enrichment/depletion of specific ESEs was due to altered SRSF2:RNA interactions. We purified SRSF2’s RNA RRM as described previously and performed isothermal titration calorimetry (ITC) with the RNA ligand 5’-uCCAGu-3’ , an optimal SRSF2 target according to the SSNG consensus sequence (Daubner et al., 2012). All three P95 mutations resulted in an increase in binding affinity of 3.9- to 4.5-fold relative to WT SRSF2 (Figures 5A and 5B; Figure S5A), consistent with the enrichment for CCNG motifs that we observed in exons promoted by SRSF2 mutations (Figure 4B). We next tested whether P95 mutations resulted in altered RNA binding specificity. In contrast to 5’-uCCAGu-3’ RNA, ITC measurements revealed that all three P95 mutants exhibited a 1.2- to 2.1-fold decrease in binding affinity to the 5’-uGGAGu-3’ RNA relative to WT SRSF2 (Figures 5A and 5B; Figure S5B). ITC measurements using the RNA sequences 5’-uCCAGu-3’ and 5’-uGGAGu-3’ revealed that G > C substitutions at the second motif position resulted in larger increases in binding affinity than at the first motif position (2.6- to 3.4-fold versus 1.1- to 1.8-fold; Figure 5B; Figures S5C and SSD). The RNA binding preferences measured by ITC were remarkably consistent with the ESE enrichment identified by RNA-seq. For each mutant, the level of motif enrichment (Figure 4E) was roughly proportional to the affinity increase (Figure 5C), and the enrichment and affinity measurement supported the same relative preference for each specific motif (CC > GC > CG > GG). This strongly supports the notion that the splicing changes caused by P95 mutations are the result of an altered sequence specificity of the SRSF2 RRM.

P95 is located at the C-terminal end of the SRSF2 RRM, and the published solution structure of SRSF2 in complex with 5’-uCCAGu-3’ revealed extensive contacts of P95 with the second cytosome (Figure S5E), emphasized by several intermolecular nuclear Overhauser effects (NOEs) (Daubner et al., 2012). To test whether SRSF2’s RNA binding surface was altered by P95 mutations, we conducted nuclear magnetic resonance (NMR) titration with the SRSF2 P95H RRM and the 5’-uCCAGu-3’ RNA and assigned the backbone of this complex using standard heteronuclear NMR experiments. Mapping of the chemical shift perturbations revealed that the RNA-binding surface of the RRM is not disturbed by the P95H mutation. However, both termini experienced large changes in their environment (Figure 5D), an observation that held true for all three P95 mutations (Figure S5F). Consistent with our ESE and ITC analyses, this relocation of termini primarily affected the second cytosome, which exhibited the largest chemical shift perturbations of its proton resonances (Figures S5G and SSH). Smaller changes of chemical shifts were observed when P95 mutants were bound to 5’-uGGAGu-3’ (Figure S5H). Together, our experiments indicate that SRSF2 mutations change SRSF2’s normal RNA-binding affinity and specificity in vitro, likely explaining the widespread alterations in ESE preference we observed in vivo.

**Mutant SRSF2 Promotes Mis-splicing and Degradation of EZH2**

We next used our transcriptome data to identify common changes in splicing driven by SRSF2 mutations that might contribute to disease. Intersection of differentially spliced genes in LSK, MP, CMML, and AML samples identified 75 genes differentially spliced in association with SRSF2 mutations in both LSK and MP cells and at least one primary patient cohort as well as an additional 97 (LSK) and 87 (MP) genes differentially spliced in one mouse cell population, but not the other, as well as a patient cohort (Figure 6A; Tables S2–S5). Many of these genes have a known importance in myeloid malignancies. For example, SRSF2 mutations promoted the inclusion of a highly conserved “poison” cassette exon of EZH2 (Enhancer of zeste homolog 2) and repressed a frame-preserving cassette exon of BCOR (BCL6 corepressor) (Figure 6A; Figures S6A and S6B). Of note, we did not identify altered splicing of CD45 in SRSF2 mutant cells (Tables S2–S5), which has been noted previously as being...
altered in murine Srsf2 KO hematopoietic cells (Wang et al., 2001).

To identify potential functional consequences of recurrent mis-splicing, we focused on the splicing event in EZH2. SRSF2 mutant cells exhibited preferential inclusion of a poison cassette exon that introduces a premature termination codon predicted to result in nonsense-mediated decay (NMD) of EZH2 (Figures 6B and 6C). Both the poison exon itself and its flanking intronic sequences exhibited high sequence conservation across vertebrates, exceeding the sequence conservation exhibited by the upstream and downstream constitutive coding exons themselves, which is a common feature of physiologically important splicing events (Lareau et al., 2007; Ni et al., 2007; Figure 6B).

We validated this EZH2 splicing change using both qualitative and quantitative isoform-specific RT-PCR in leukemia cell lines that were WT or mutant for SRSF2 (Figures S6C and S6D) as well as in an independent panel of primary AML patient samples with or without SRSF2 mutations (n = 8, 4 with SRSF2 mutations; Figure 6D; Figure S6E).

Next, to confirm whether the cassette exon promoted by SRSF2 mutations triggers degradation by NMD, we measured the half-life of the inclusion isoform of EZH2 in SRSF2-P95H cells transfected with a control or anti-UPF1 (a required NMD factor) short hairpin RNA (shRNA) following transcriptional shutoff with actinomycin D (1 Hoen et al., 2011; Figure 6F; Figures S6F and S6G). The fact that the mRNA half-life of the inclusion isoform of EZH2 was lengthened by UPF1 knockdown in these
experiments suggests that this particular isoform of EZH2, which is promoted by mutant SRSF2, undergoes NMD. The half-life of a well-characterized NMD substrate of SRSF3 (Lareau et al., 2007; Ni et al., 2007) increased similarly following UPF1 knockdown, confirming that UPF1 knockdown effectively inhibited NMD (Figure S6H).

Next, to identify whether the protein product of EZH2 is altered in SRSF2 mutant cells, we performed WB analysis of a panel of human AML cell lines WT (TF-1, K562) or mutant for SRSF2 (K052) (all WT for EZH2). This revealed lower EZH2 protein levels as well as lower global levels of histone H3 lysine 27 trimethylation (H3K27me3, a methylation mark placed by EZH2) in SRSF2 mutant K052 cells (Figure 6F). To further validate this finding in an isogenic context, we performed WB analysis in K562 cells ectopically expressing WT SRSF2 or SRSF2P95H/L/R mutant cDNA. This analysis revealed a consistent downregulation of EZH2 protein expression as well as global H3K27me3 in all three SRSF2 mutant samples compared with SRSF2 WT K562 cells (Figure 6G).

Consistent with SRSF2 mutations promoting a disabling splicing change in EZH2, EZH2 loss-of-function mutations are common in MDS. In an analysis of >1,800 MDS patients where EZH2 and SRSF2 were both sequenced, EZH2 loss-of-function mutations were mutually exclusive with SRSF2 mutations (p < 0.0001) (Bejar et al., 2012; Ernst et al., 2010; Haferlach et al., 2014; Muto et al., 2013; Papaemmanuil et al., 2013; Figure 6H).

The above data strongly link SRSF2 mutations to disabling splicing of EZH2. We next sought to examine whether the change in RNA ESE preference induced by SRSF2 mutations caused EZH2 mis-splicing. We therefore cloned the genomic locus containing the EZH2 poison exon and flanking introns and constitutive exons to create a minigene that recapitulates this splicing event. We identified three potential SRSF2-dependent SSGN motifs in the poison exon (CCTG, CCTG, and GCAG), one or more of which we expected to be better recognized by mutant SRSF2 than WT SRSF2. We then mutated each motif to the corresponding GG equivalent, both separately and in combination (Figure 6I). Measuring cassette exon recognition in K562 cells expressing WT or mutant SRSF2, we found that the first motif was required for robust splicing change in SRSF2 mutant cells, such that the mutation CCTG > GGTG prevented an increase in poison exon recognition (Figure 6J). We conclude that SRSF2 mutations induce a disabling splicing change in EZH2 in an ESE-dependent manner, consistent with altered RNA recognition activity.

We next sought to test whether restoring normally spliced EZH2 mRNA could rescue hematopoiesis in SRSF2 mutant cells. EZH2 full-length cDNA or an empty vector (both in a retroviral ZsGreen1 vector) were overexpressed in c-Kit+ Srsf2P95H or WT cells, followed by assessment of methylcellulose colony formation of c-Kit+/ZsGreen1+ cells. EZH2 cDNA was equally overexpressed in Srsf2 mutant and WT cells (Figure S6I), and Srsf2P95H mutant cells overexpressing full-length EZH2 experienced an ∼50% increase in colony formation relative to Srsf2P95H mutant cells expressing an empty vector (Figure 6K; Figure S6J). In contrast, EZH2 overexpression had no substantial effect on initial colony formation in Srsf2 WT cells (Figure 6K; Figure S6I). These data identify that restoration of normally spliced EZH2 mRNA in SRSF2 mutant cells at least partially rescues the hematopoietic defects induced by mutant SRSF2.

**DISCUSSION**

The consistent occurrence of heterozygous point mutations affecting highly restricted residues of spliceosomal proteins strongly suggests a gain-of-function or dominant-negative activity for these mutations in malignant transformation. Here we identify an effect of the SRSF2P95H mutation distinct from loss of SRSF2 and reveal that mutations in SRSF2 confer an alteration in function that results in key aspects of MDS. This includes an increase in HSPCs in Srsf2P95H mutant mice with impaired differentiation, altered cell cycle kinetics, and increased apoptosis resulting in peripheral cytopenias and morphologic dysplasia. By contrast, WT Srsf2 appears to be constitutively required for hematopoiesis.

---

**(A) Intersection of genes exhibiting differential splicing in SRSF2 mutant versus WT mouse LSK and MP cells and primary AML and CML samples (restricted to orthologous genes). (B) Integrative Genomics Viewer (IGV)/Sashimi plot illustrating the EZH2 cassette exon promoted by SRSF2 mutations in multiple datasets analyzed here (top) (the patient numbers listed in the Sashimi plot correspond to the numbers in Table S1 detailing patient characteristics). The DNA sequence conservation of the locus, as estimated by phastCons (Siepel et al., 2005), across 30 vertebrate species is shown in the track below the Sashimi plot. (C) Bar plot describing the percentage of EZH2 transcripts harboring a specific cassette exon in the SRSF2 mutant relative to WT primary AML samples from RNA-seq data. Error bars indicate 95% confidence intervals. (D) RT-PCR of an EZH2 exon inclusion event in an independent set of SRSF2 WT and mutant AML samples. (E) Quantitative RT-PCR of EZH2 inclusion isoform in SRSF2P95H mutant cell line K052 cells with or without UPF1 knockdown and actinomycin D treatment. (F) WB analysis for EZH2 and H3K27me3 in SRSF2EZH2 WT (TF-1, K562) and SRSF2P95H mutant/EZH2 WT (K052) AML cell lines. (G) WB analysis for EZH2, H3K27me3, and FLAG epitope in K562 cells with lentiviral overexpression of N-terminal FLAG-tagged SRSF2 WT, SRSF2P95H, SRSF2P95L, or SRSF2P95R (left). Relative quantification of EZH2 protein expression by WB to total histone H3 expression in K562 cells expressing SRSF2 mutants relative to WT is shown on the right. (H) EZH2 and SRSF2 mutations are mutually exclusive in the sequencing of DNA from >1,000 MDS patients (Bejar et al., 2012; Ernst et al., 2010; Haferlach et al., 2014; Muto et al., 2013; Papaemmanuil et al., 2013). (I) Schematic of the EZH2 cassette exon with SSGN motifs highlighted and mutations to GG equivalents shown. (J) EZH2 cassette exon inclusion for minigenes containing the endogenous cassette exon or a cassette exon with mutation of motifs 1, 2, and/or 3 to the GG equivalent. (K) Photographs (left) and enumeration (right) of c-Kit+/ZsGreen1+ cells from Srsf2 WT or Srsf2P95H mice 14 days after overexpression of empty vector or EZH2 cDNA and platting in methylcellulose medium. **p < 0.01, ****p < 0.0001. Error bars represent mean ± SD unless stated otherwise. See also Figure S6 and Tables S1–S5.
Transcriptional analysis of SRSF2 mutant cells revealed that SRSF2 mutations result in genome-wide alterations in ESE preference in both human and murine cells. Biochemical analysis of the interaction of SRSF2 with RNA in cell-free in vitro assays identified an analogous change in specificity of interactions between SRSF2 and pre-mRNA induced by SRSF2 mutations. This altered interaction of mutant SRSF2 with RNA appears to be due to an effect of SRSF2P95H/L/R mutations on the conformations of the termini of SRSF2’s RRM domain, as revealed by NMR spectroscopy. Our genomic and biochemical assays indicate that SRSF2 mutations cause alteration rather than loss-of-function, driving preferential recognition of cassette exons containing C- versus G-rich ESEs.

The altered pre-mRNA recognition activity of mutant SRSF2 likely underlies the mis-splicing of key transcriptional regulators—several of which have been implicated previously in MDS pathogenesis. This includes promotion of a poison exon of EZH2 that undergoes NMD and results in reduced EZH2 protein expression in SRSF2 mutant cells. Loss-of-function mutations in EZH2 occur in the same exact spectrum of myeloid malignancies as SRSF2 mutations (Ernst et al., 2010; Nikoloski et al., 2010) and loss of Ezh2 has been functionally linked to MDS development in vivo (Muto et al., 2013). Moreover, SRSF2 and EZH2 mutations are mutually exclusive in MDS patients (Haferlach et al., 2014; Papaemmanuil et al., 2013), but the basis for this observation was previously unknown. The data here provide a mechanistic basis for this mutual exclusivity as SRSF2 mutations functionally reduce EZH2 protein expression.

In addition to the effects of mutant SRSF2 on EZH2 splicing and protein expression, a number of other genes of known importance in hematopoiesis and malignancy were also consistently differentially spliced in isogenic human cells, primary patient samples, and murine cells bearing mutant SRSF2. These include additional genes mutated in MDS (such as BCO1), genes with an importance in hematopoietic stem cell self-renewal (such as IKAROS), and genes critical for cell survival (such as CASPASE 8). Future efforts to understand the functional effects of each of these specific splicing events will be important in further delineating the effects of mutant SRSF2 on MDS pathogenesis as well as possibly providing novel means for therapeutic targeting of SRSF2 mutant cells.

Our studies, which reveal both mechanistic splicing alterations and specific mis-spliced isoforms in SRSF2 mutant cells, may provide insights into therapeutic opportunities for targeting SRSF2 mutant cells. For example, the observations that mutant SRSF2 promotes the inclusion of a poison exon in an ESE-dependent manner and that restoration of normally spliced EZH2 mRNA partially rescues defective hematopoiesis in SRSF2 mutant cells suggest that normal cellular function may be at least partially restored by manipulating specific pathologic splicing events.

**EXPERIMENTAL PROCEDURES**

Generation of the Srsf2P95H conditional knockin mice is described in the Supplemental Experimental Procedures. All animal procedures were conducted in accordance with the Guidelines for the Care and Use of Laboratory Animals and approved by the Institutional Animal Care and Use Committees at Memorial Sloan Kettering Cancer Center.

**Patient Samples**

Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center and Fred Hutchinson Cancer Research Center and conducted in accordance with the Declaration of Helsinki protocol. Informed consent was obtained from all human subjects.

**mRNA Sequencing**

For sorted mouse cell populations, K562 cells, and primary AML and CMML samples, RNA was extracted using Qiagen RNeasy columns, polyA-selected, unstranded Illumina libraries were prepared with a modified TruSeq protocol. 0.5 x AMPure XP beads were added to the sample library to select for fragments of ~200 bp, followed by 1 x beads to select for fragments of >100 bp. These fragments were then amplified with PCR (15 cycles) and separated by gel electrophoresis (2% agarose). 300-bp DNA fragments were isolated and sequenced on the Illumina HiSeq 2000 (100 million 2 x 49 bp reads/sample).

**RNA-Seq Read Mapping**

Reads were mapped to the University of California, Santa Cruz (UCSC) hg19 (NCBI GRCh37) human genome or UCSC mm10 (NCBI GRCh38) genome assemblies. First, a modified version of RNA-seq by expectation maximization (RSEM) that called Bowtie v1.0.0, with the --v 2 argument was created. This modified RSEM was then called with the arguments --bowtie-m 100--bowtie-chunksize 500 --calc-ci --output-genome-bam on the gene annotation file. Read alignments with MAPping Quality (MAQ) scores of 0 and/or a splice junction overhang of less than 6 bp were then filtered out. The remaining unaligned reads were then aligned by TopHat v2.0.8b with the arguments -b bowtie1 --read-mismatches 2 --read-edit-dist 2 --no-mixed --no-discordant --min-anchor-length 6 --splice-mismatches 0 --min-intron-length 10 --max-intron-length 1000000 --min-isozyme-fraction 0.0 --no-novel-juncs --no-novel-indels --raw-juncs on the splice junction file (--mate-inner-dist and --mate-std-dev were calculated by mapping to constitutive coding exons with the Mixture of Isoforms (MISO) exon,utils.py utility). The resulting TopHat alignments were then filtered as for the RSEM-generated alignments. Finally, the RSEM- and TopHat-created binary sequence alignment/map (BAM) files were merged to create final BAM files.

**Isoform Expression Measurements**

Reads were mapped to the University of California, Santa Cruz (UCSC) hg19 (NCBI GRCh37) human genome or UCSC mm10 (NCBI GRCh38) genome assemblies. First, a modified version of RNA-seq by expectation maximization (RSEM) that called Bowtie v1.0.0, with the --v 2 argument was created. This modified RSEM was then called with the arguments --bowtie-m 100--bowtie-chunksize 500 --calc-ci --output-genome-bam on the gene annotation file. Read alignments with MAPping Quality (MAQ) scores of 0 and/or a splice junction overhang of less than 6 bp were then filtered out. The remaining unaligned reads were then aligned by TopHat v2.0.8b with the arguments -b bowtie1 --read-mismatches 2 --read-edit-dist 2 --no-mixed --no-discordant --min-anchor-length 6 --splice-mismatches 0 --min-intron-length 10 --max-intron-length 1000000 --min-isozyme-fraction 0.0 --no-novel-juncs --no-novel-indels --raw-juncs on the splice junction file (--mate-inner-dist and --mate-std-dev were calculated by mapping to constitutive coding exons with the Mixture of Isoforms (MISO) exon,utils.py utility). The resulting TopHat alignments were then filtered as for the RSEM-generated alignments. Finally, the RSEM- and TopHat-created binary sequence alignment/map (BAM) files were merged to create final BAM files.

**ACCESSION NUMBERS**

The accession number for the RNA sequencing data reported in this paper is Gene Expression Omnibus (GEO); GSE65349.

**SUPPLEMENTAL INFORMATION**

Supplemental Information includes Supplemental Experimental Procedures, six figures, and five tables and can be found with this article online at http://dx.doi.org/10.1016/j.ccell.2015.04.006.

**AUTHOR CONTRIBUTIONS**

ACKNOWLEDGMENTS

E.K. is supported by the Worldwide Cancer Research Fund. A.R. was supported by the NIH/NHLBI (U01 HL09993), NIH/NIDDK (K08 DK082783), the J.P. McCarthy Foundation, and the Storb Foundation. S.H. and O.A.-W. are supported by grants from the Edward P. Evans Foundation. S.H. was supported by Yale Comprehensive Cancer Center institutional funds. R.K.B. was supported by the Hartwell Innovation Fund, Damon Runyon Cancer Research Foundation (DFS 04-12), Eliison Medical Foundation (AG-NS-1030-13), NIH/NIDDK (R56 DK103854), NIH/NCI recruitment support (P30 CA015704), and Fred Hutchinson Cancer Research Center institutional funds. J.O.I. was supported by an NIH/NCI training grant (T32 CA009657) and NIH/NIDDK pilot study (P30 DK056465). C.L. is supported by a career development award grant from the Leukemia and Lymphoma Society and an ATIP-Avenir grant from the French government. O.A.-W. is supported by an NIH K08 clinical investigator award (K08CA160647-01), a Department of Defense Postdoctoral Fellow Award in Bone Marrow Failure Research (W81XWH-12-1-0041), the Josie Robertson Investigator Program, and a Damon Runyon Clinical Investigator Award with support from the Evans Foundation. F.H.-T.A. acknowledges support from the NCCR RNA and Disease funded by the Swiss National Science Foundation and the SNF Sinergia CRSII3_127454. Y.L. and Y.M. were supported by NIH/NIGMS grant R01 GM102869 and Senior Research Fellowship Grant 101908/Z/13/Z (to Y.M.) from the Wellcome Trust. J.D. acknowledges assistance from Dr. Nezih Cereb, HistoGenetics (Ossining, NY).

Received: December 2, 2014
Revised: February 19, 2015
Accepted: April 10, 2015
Published: May 11, 2015

REFERENCES


