Award Number: **W81XWH-13-1-0472**

TITLE: The Role of U2AF1 Mutations in the Pathogenesis of Myelodysplastic Syndromes

PRINCIPAL INVESTIGATOR: **Matthew Walter**

CONTRACTING ORGANIZATION: **Washington University**
St. Louis MO 63110

REPORT DATE: **December 2016**

TYPE OF REPORT: **Final**

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for public release; distribution unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
December 2016	Final	30 Sep 2013 - 29 Sep 2016

4. TITLE AND SUBTITLE
The Role of U2AF1 Mutations in the Pathogenesis of Myelodysplastic Syndromes

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-13-1-0472

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Matthew J. Walter

Email:
miwalter@dom.wustl.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Washington University
St. Louis, MO 63130-4862

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
U2AF1 mutations occur in up to 11% of myelodysplastic syndrome (MDS) patients. To study the effects of the most common *U2AF1* mutation, U2AF1(S34F), on hematopoiesis and pre-mRNA splicing *in vivo*, we created doxycycline-inducible U2AF1(WT) and U2AF1(S34F) transgenic mice. Following transgene induction, U2AF1(S34F) mice have reduced WBCs, increased hematopoietic stem/progenitor cells, and increased HSC cell cycling compared to U2AF1(WT) mice. U2AF1(S34F) stem cells are at a competitive disadvantage compared to control cells, suggesting that the increase in HSC cell cycling following U2AF1(S34F) expression may lead to stem cell exhaustion. Next, we compared RNA splicing in progenitor cells from U2AF1(S34F) and U2AF1(WT) mice using whole transcriptome RNA-seq. We identified 460 splicing junctions that were differentially expressed in U2AF1(S34F) samples compared to U2AF1(WT). We validated several homologous dysregulated junctions (i.e., across species) in MDS patient bone marrow samples that have mutant U2AF1(S34F) versus U2AF1(WT). Together, these results suggest that mutant U2AF1 expression contributes to the altered hematopoiesis and pre-mRNA splicing observed in patients with *U2AF1* mutations. This study also identifies changes in gene isoform expression unique to *U2AF1* mutations that may have functional significance for MDS pathogenesis, which is being investigated in ongoing studies.

15. SUBJECT TERMS
Myelodysplastic Syndromes, Splicing, Spliceosome, Mouse Model, Hematopoiesis, RNA-seq, U2AF1

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	32	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS:

1. INTRODUCTION:

The goal of this project is to understand the mechanism of disease pathogenesis induced by U2AF1 mutations in myelodysplastic syndromes (MDS). U2AF1 is a key spliceosome protein that binds the AG dinucleotide of the 3' splice acceptor site in pre-mRNA introns during splicing and is mutated in up to 11% of MDS patients, making it one of the most commonly mutated genes in MDS. Overall, mutations in spliceosome genes occur in up to ~50% of patients with MDS, further implicating altered pre-mRNA splicing in disease pathogenesis. We hypothesize that U2AF1 mutations result in altered mRNA splicing in hematopoietic cells, and thereby lead to altered progenitor/stem cell function and ineffective hematopoiesis. In this project, we will test our hypothesis in the following Specific Aims. Specific Aim 1. We will determine whether the U2AF1(S34F) mutation alters hematopoiesis in vivo. We will inducibly express wild-type and S34F mutant (resulting from the most common U2AF1 mutation) human U2AF1 cDNAs in mice and determine the contribution of mutant U2AF1 to MDS pathogenesis by comprehensively evaluating basal hematopoiesis and stem cell function. Specific Aim 2. We will use RNA-Seq to identify alternatively spliced genes in primary hematopoietic progenitor cells harvested from U2AF1(S34F) mutant mice. We will identify alternatively spliced genes induced by U2AF1 mutations by performing transcriptome sequencing (RNA-Seq) using RNA isolated from wild-type and mutant bone marrow progenitors. Candidate genes with alternative splicing will be interrogated in MDS patient samples with and without U2AF1 mutations and tested for their contribution to disease.

2. KEYWORDS:

Myelodysplastic Syndromes
Splicing
Spliceosome
Mouse model
Hematopoiesis
RNA-seq
U2AF1

3. ACCOMPLISHMENTS:

Task 1. Seek IACUC and DoD ACURO approval for the use of animals.
Objectives: Obtain approval.
Progress and Accomplishments: IACUC and DoD ACURO approved.
Opportunities for training and professional development: Nothing to Report.
Dissemination of results: Nothing to Report.
Plans for next reporting period: Nothing to Report.

Task 2. Specific Aim 1. We will determine whether the U2AF1(S34F) mutation alters hematopoiesis in vivo.
Objectives: We will determine whether expression mutant U2AF1(S34F) induces ineffective hematopoiesis in mice.
Progress and Accomplishments: Figures from Task 2 results were reported in our prior year 1 update and published in Cancer Cell in 2015. Therefore, I will summarize these findings here without figures. U2AF1(S34F)-recipient mice have reduced total WBCs in the peripheral blood compared to U2AF1(WT)- and rtTA only-recipient controls (4.3 vs 7.11 and 7.13 K/μl, respectively, $p \leq 0.01$), but no significant changes in bone marrow cellularity or spleen size (n=9-11). U2AF1(S34F)-recipient mice have a perturbed mature cell lineage distribution, including reduced monocytes and B cells in both peripheral blood ($p \leq 0.05$) and bone marrow ($p \leq 0.01$) when compared to control mice (n=9-11). Reduction of bone marrow monocytes occurs as early as 5 days and is associated with increased Annexin V+ ($p \leq 0.05$) and phospho-H2AX ($p \leq 0.05$) compared to controls, suggesting loss of these cells may be due to apoptosis. In addition, U2AF1(S34F)-recipient mice have increased numbers of progenitors in both bone marrow and spleen by CFU-C methylcellulose assay and flow cytometry for c-Kit+/Lineage- cells, as well as common myeloid progenitors (CMPs), when compared to U2AF1(WT) and rtTA only controls ($p \leq 0.05$, n=5-10). U2AF1(S34F)-recipient mice also have an increase in the frequency of bone marrow hematopoietic stem cells (HSCs) measured by flow cytometry for bone marrow KLS (c-Kit+/Lineage-/Sca-1+) cells ($p \leq 0.05$). The increase in bone marrow KLS cells in U2AF1(S34F)-recipient mice seen as early as 5 days is associated with higher levels of intracellular Ki67 (a marker of cell proliferation) in KLS cells compared to controls ($p < 0.05$, n=8-13). We hypothesized that further perturbation of the spliceosome using compounds that modulate splicing would produce a synthetic interaction in mutant cells but not wild-type cells. Indeed, we observed that the progenitor cell expansion

induced by U2AF1(S34F) could be attenuated by a pre-mRNA splicing modulator (sudemycin) (**Figure 1**). Collectively, these data indicate U2AF1(S34F) expression alters hematopoiesis *in vivo*.
Opportunities for training and professional development: Trainees learned experimental design and analysis including how to characterize hematopoiesis in pre-clinical models using flow cytometry, progenitor assays, and bone marrow transplantation. In addition, trainees created plots, performed statistical analysis, wrote manuscripts, and prepared and presented the data orally to other scientists.
Dissemination of results: These data have been published in *Cancer Cell*. 2015 May 11;27(5):631-43. doi: 10.1016/j.ccell.2015.04.008. PubMed PMID: 25965570; PubMed Central PMCID: PMC4430854. These data are in press at *Nature Communications.* 8, 14060 doi: 10.1038/ncomms14060 (in press).
Plans for next reporting period: Nothing to Report.

Task 3. Specific Aim 2. We will use RNA-Seq to identify alternatively spliced genes in primary hematopoietic progenitor cells harvested from U2AF1(S34F) mutant mice.
Current Objectives: We will identify alternatively spliced genes in primary murine hematopoietic progenitor cells expressing mutant U2AF1 using RNA-Seq. We will validate alternatively spliced genes in primary MDS bone marrow cells expressing mutant U2AF1 and test their contribution to disease.
Progress and Accomplishments: We presented the results from unbiased RNA sequencing on sorted bone marrow CMPs following 5 days of transgene induction in U2AF1(S34F)- and U2AF1(WT)-transplanted mice (n=3 each) in our year 2 update and results have been published in Cancer Cell in 2015. In summary, we identified alternative splicing of 742 splicing junctions that were differentially expressed in 633 genes in U2AF1(S34F)/rtTA samples compared to U2AF1(WT)/rtTA controls (FDR<0.1). These genes were enriched in genes involved in RNA binding (FDR<0.1). We intersected junctions that were significant (FDR<0.1) across 3 datasets: mouse CMP samples (n=219 junctions), AML patient samples with and without U2AF1 mutations (n=162 junctions), and our previously-described primary human CD34+ cells over-expressing U2AF1(S34F) or U2AF1(WT) (n=1652 junctions). The intersection included 17 homologous dysregulated junctions (p<0.0001; simulation) in 13 genes, including *H2AFY*, a gene known to be involved in stem cell biology. Alternative splicing of *H2AFY*, and 6 other genes, was confirmed in primary MDS patient bone marrow samples. The H2AFY 1.1 isoform is reduced by 3-fold in U2AF1(S34F) samples compared to U2AF1(WT).

We investigated the contribution of *H2afy* to hematopoietic alterations induced by U2AF1(S34F) by generating U2AF1(S34F) mice that lack *H2afy*. We hypothesize that a further reduction of H2afy 1.1 expression to zero will exacerbate the hematopoietic phenotypes induced by U2AF1(S34F). *H2afy* null mice lack expression of both 1.1 and 1.2 isoforms. We observed that compound U2AF1(S34F);H2afy$^{-/-}$ mice have enhanced hematopoietic phenotypes (**Figure 2**). Collectively, the data suggests that reduction of H2afy 1.1 expression may contribute to U2AF1(S34F)-induced phenotypes. Adding back H2afy isoform 1.1 in null mice will verify its direct contribution to this phenotype. Collectively, these data indicate U2AF1(S34F) expression alters pre-mRNA splicing *in vivo*. These splicing changes may contribute to disease pathogenesis in MDS.
Opportunities for training and professional development: Trainees learned experimental design and analysis including RNA-seq analysis. In addition, trainees created plots, performed statistical analysis, wrote manuscripts, and prepared and presented the data orally to other scientists.
Dissemination of results: These data have been published in Cancer Cell. 2015 May 11;27(5):631-43. doi: 10.1016/j.ccell.2015.04.008. PubMed PMID: 25965570; PubMed Central PMCID: PMC4430854.
Plans for next reporting period: Nothing to Report.

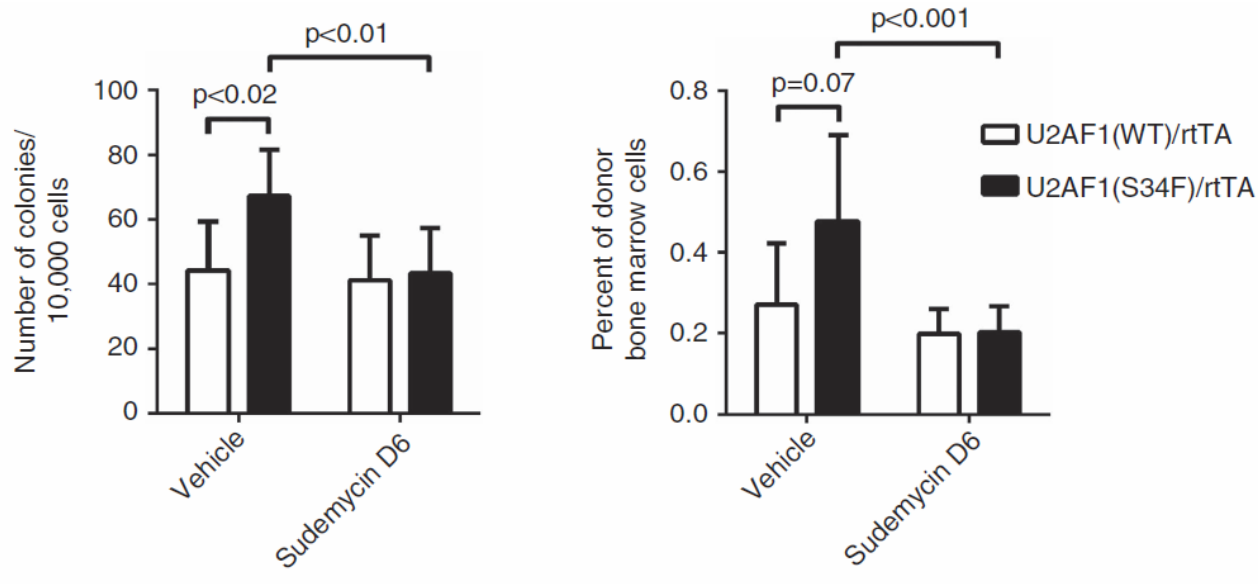


Figure 1. Sudemycin D6 treatment attenuates mutant U2AF1-associated progenitor cell expansion in U2AF1(S34F) transgenic mice. (left) Hematopoietic progenitor CFU-C colony forming assay and (right) flow cytometry for HSPC cell surface markers (c-Kit+, lineage-, Sca1+, KLS, right panel) on U2AF1(WT)- or (S34F) mutant-recipient mouse bone marrow following treatment is shown [n=6 for both U2AF1(WT) conditions, n=7 for vehicle-treated U2AF1(S34F), n=11 for sudemycin D6-treated U2AF1(S34F)]. Data pooled from 2 independent experiments; statistics calculated using two-tailed t-tests for each comparison shown; mean values with s.d. shown.

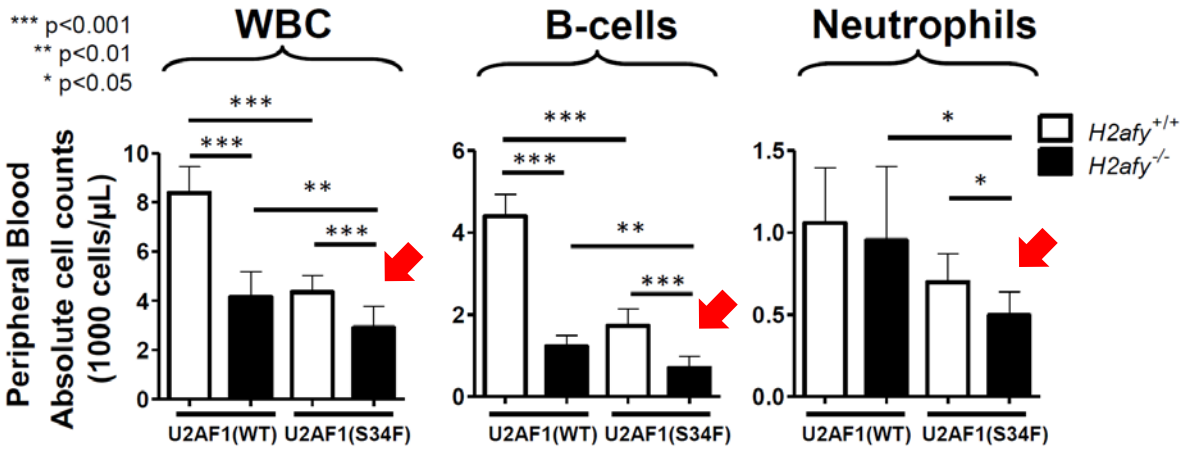


Figure 2. Deletion of H2afy enhances mutant U2AF1-induced hematopoietic alterations. (left) Peripheral blood white blood cell (WBC) counts are lowest in U2AF1(S34F) mice lacking H2afy (red arrow). (Middle) Peripheral blood B-cell counts are lowest in U2AF1(S34F) mice lacking H2afy (red arrow). (Right) Peripheral blood neutrophil counts are lowest in U2AF1(S34F) mice lacking H2afy (red arrow). N=8-10, each group, At 4 month of Doxycycline treatment; statistics calculated using two-tailed t-tests for each comparison shown; mean values with s.d. shown.

Task 4. Data analysis and report generation
Objectives: Analyze data.
Progress and Accomplishments: Analysis is complete, as reported above.
Opportunities for training and professional development: Trainees learned data analysis including hematopoiesis characterization, flow cytometry, progenitor assays, and RNA-seq. In addition, trainees created plots, performed statistical analysis, wrote manuscripts, and prepared and presented the data orally to other scientists.
Dissemination of results: These data have been published in *Cancer Cell*. 2015 May 11;27(5):631-43. doi: 10.1016/j.ccell.2015.04.008. PubMed PMID: 25965570; PubMed Central PMCID: PMC4430854. These data are in press at *Nature Communications.* 8, 14060 doi: 10.1038/ncomms14060 (in press).
Plans for next reporting period: Nothing to Report.

4. IMPACT:
What was the impact on the development of the principal discipline(s) of the project and on other disciplines? The results provide evidence that spliceosome gene mutations, specifically *U2AF1* mutations, affect hematopoiesis and contribute to bone marrow failure. Given that spliceosome gene mutations are the most common family of genes mutated in MDS, a better understanding of the underlying mechanisms of disease pathogenesis could impact novel treatment approaches in the future. The mice generated here are available to the scientific community to study U2AF1 mutations in a variety of tissues. Our long-term goal is to identify therapeutic approaches to target spliceosome mutant cells that contribute to bone marrow failure.
What was the impact on other disciplines? Spliceosome gene mutations are found in other solid tumors (eg, lung cancer, breast cancer). Therefore, knowledge gained here could have broad implications for cancer.
What was the impact on technology transfer? Nothing to Report.
What was the impact on society beyond science and technology? Nothing to Report.

5. CHANGES/PROBLEMS: Nothing to Report.

6. PRODUCTS:
Published Manuscripts:
1. Shirai CL, Ley JN, White BS, Kim S, Tibbitts J, Shao J, Ndonwi M, Wadugu B, Duncavage EJ, Okeyo-Owuor T, Liu T, Griffith M, McGrath S, Magrini V, Fulton RS, Fronick C, O'Laughlin M, Graubert TA, Walter MJ. Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo. Cancer Cell. 2015 May 11;27(5):631-43. doi: 10.1016/j.ccell.2015.04.008. PubMed PMID: 25965570; PubMed Central PMCID: PMC4430854. Federal support acknowledged.
2. Cara Lunn Shirai, Brian S. White, Manorama Tripathi, Roberto Tapia, James N. Ley, Matthew Ndonwi, Sanghyun Kim, Jin Shao, Alexa Carver, Borja Saez, Robert S. Fulton, Catrina Fronick, Michelle O'Laughlin, Chandraiah Lagisetti, Thomas R. Webb, Timothy A. Graubert, and Matthew J. Walter. Mutant U2AF1-Expressing Cells Are Sensitive to Pharmacological Modulation of the Spliceosome. Nature Communications. 8, 14060 doi: 10.1038/ncomms14060 (in press). Federal support acknowledged.

Abstracts:
1. Walter MJ, American Society of Hematology, 58th Annual meeting, San Diego, CA, USA. December, 2016. "Clinical Implications of Spliceosome Mutations: Epidemiology, Clonal Hematopoiesis, and Potential Therapeutic Strategies".

Presentations:
1. Shirai CL, Workshop on Splicing Factor Mutations in Cancer at the Broad Institute in Cambridge, MA. October 2016. "Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome."*
2. Shirai CL, First International Caparica Conference on Splicing in Caparica, Lisbon, Portugal. September, 2016. "Biologic and Therapeutic Implications for U2AF1 Mutations in Blood Cancers."*
3. Walter MJ, American Society of Hematology, 58th Annual meeting, San Diego, CA, USA. December, 2016. "Clinical Implications of Spliceosome Mutations: Epidemiology, Clonal Hematopoiesis, and Potential Therapeutic Strategies".

Website(s) or other Internet site(s). Nothing to Report.
Technologies or techniques. Nothing to Report.
Inventions, patent applications, and/or licenses. Nothing to Report.

Other Products. We generated U2AF1(S34F) and U2AF1(WT) transgenic mice. RNA-seq data is available at NCBI Gene Expression Omnibus accession number GSE66793.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

Individuals who worked on the project

Name: Matthew Walter
Project Role: PI
Researcher Identifier (e.g. ORCID ID): 0000-0002-7753-1091
Nearest person month worked: 5
Contribution to Project: Dr. Walter has overall responsibility for all aspects of the project.
Funding Support: The Leukemia and Lymphoma Society Scholar Award provides salary support for Dr. Walter and does not include supply costs. Therefore, there is no overlap in personnel costs, supplies, or other expenses between the LLS Scholar Award and the DoD award.

Name: Matthew Ndonwi
Project Role: Research laboratory supervisor
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 2
Contribution to Project: Dr. Ndonwi performed animal care, handling, and cell manipulation.
Funding Support: Effort was funded by this award.

Name: Brian Wadugu
Project Role: Graduate Student
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 1
Contribution to Project:Mr. Wadugu performed animal care, handling, and characterization of transgenic mice and RNA-Seq experiments.
Funding Support: Effort was funded by this award.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period? Nothing to report.
What other organizations were involved as partners? Nothing to report.

8. SPECIAL REPORTING REQUIREMENTS. Nothing to report.

COLLABORATIVE AWARDS: Nothing to report.
QUAD CHARTS: Nothing to report.

9. APPENDICES: Nothing to report.

nature COMMUNICATIONS

ARTICLE

Received 25 Mar 2016 | Accepted 24 Nov 2016 | Published 9 Jan 2017

DOI: 10.1038/ncomms14060 OPEN

Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome

Cara Lunn Shirai[1,*], Brian S. White[1,*], Manorama Tripathi[1,*], Roberto Tapia[1], James N. Ley[1], Matthew Ndonwi[1], Sanghyun Kim[1], Jin Shao[1], Alexa Carver[1], Borja Saez[2], Robert S. Fulton[3], Catrina Fronick[3], Michelle O'Laughlin[3], Chandraiah Lagisetti[4], Thomas R. Webb[4], Timothy A. Graubert[2] & Matthew J. Walter[1]

Somatic mutations in spliceosome genes are detectable in ~50% of patients with myelodysplastic syndromes (MDS). We hypothesize that cells harbouring spliceosome gene mutations have increased sensitivity to pharmacological perturbation of the spliceosome. We focus on mutant U2AF1 and utilize sudemycin compounds that modulate pre-mRNA splicing. We find that haematopoietic cells expressing mutant U2AF1(S34F), including primary patient cells, have an increased sensitivity to *in vitro* sudemycin treatment relative to controls. *In vivo* sudemycin treatment of U2AF1(S34F) transgenic mice alters splicing and reverts haematopoietic progenitor cell expansion induced by mutant U2AF1 expression. The splicing effects of sudemycin and U2AF1(S34F) can be cumulative in cells exposed to both perturbations—drug and mutation—compared with cells exposed to either alone. These cumulative effects may result in downstream phenotypic consequences in sudemycin-treated mutant cells. Taken together, these data suggest a potential for treating haematological cancers harbouring *U2AF1* mutations with pre-mRNA splicing modulators like sudemycins.

[1] Division of Oncology, Washington University School of Medicine, St Louis, Missouri 63110, USA. [2] Massachusetts General Hospital Cancer Center, Boston, Massachusetts 02114, USA. [3] McDonnell Genome Institute, Washington University, St Louis, Missouri 63108, USA. [4] SRI International, Bioscience Division, Menlo Park, California 94025, USA. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to M.J.W. (email: mjwalter@wustl.edu).

Myelodysplastic syndromes (MDS) are the most common adult myeloid malignancy with up to 40,000 new cases diagnosed each year in the United States[1,2]. MDS are a heterogeneous group of clonal haematopoietic stem cell disorders characterized by peripheral blood cytopaenias and progenitor expansion; approximately one-third of patients will transform to a secondary acute myeloid leukaemia (AML) that has a poor prognosis[3]. The only curative therapy is bone marrow transplantation, which is often not an option because of patient comorbidities[3]. New treatment approaches are greatly needed. At least half of all MDS patient bone marrow samples harbour a mutation in one of several spliceosome genes[4–10], highlighting a potential genetic vulnerability. In addition, spliceosome gene mutations often occur in the founding clones of MDS tumours, providing an attractive target for elimination of all tumour cells[10,11]. Spliceosome gene mutations are mutually exclusive of each other in patients[4,10–12], implying either a redundancy in pathogenic function or that a cell cannot tolerate two spliceosome perturbations at once. With this in mind, we hypothesized that cells harbouring a spliceosome gene mutation would have increased sensitivity to further perturbation of the spliceosome by splicing modulator drugs. To examine this, we utilized sudemycin compounds that bind the SF3B1 spliceosome protein and modulate pre-mRNA splicing[13–15]. We used sudemycin D1 and D6, which are synthetic compounds that have been optimized by several rounds of medicinal chemistry for their potent *in vivo* antitumour activity[13]. We examined the sensitivity of spliceosome mutant cells to sudemycin treatment, focusing on mutations in the spliceosome gene *U2AF1*, which have been identified in 11% of MDS patients, utilizing the S34F missense mutation most commonly found in our studies[4,5]. Mutant U2AF1(S34F) expression has been shown by our group and others to cause altered pre-mRNA splicing in a variety of cell types, as well as altered haematopoiesis and pre-mRNA splicing in mice[4,5,16–19].

In this manuscript, we provide evidence that U2AF1(S34F)-expressing cells are sensitive to the splicing modulator drug sudemycin. Haematopoietic cells expressing mutant U2AF1 show reduced survival and altered cell cycle in response to sudemycin D6 *in vitro*. *In vivo* treatment of U2AF1(S34F) transgenic mice with sudemycin results in an attenuation of mutant U2AF1-induced haematopoietic progenitor cell expansion that is associated with increased cell death. In addition, unsupervised analysis of whole-transcriptome sequencing (RNA-seq) finds that sudemycin D6 perturbs RNA splicing in both mutant U2AF1(S34F)- and U2AF1(WT)-expressing bone marrow cells; however, sudemycin D6 treatment further modulates mutant U2AF1(S34F)-induced splicing changes to create cumulative effects on cells *in vivo*. The cumulative RNA-splicing effects of sudemycin and mutant U2AF1 may contribute to the downstream phenotypic consequences we observe *in vivo*.

Results

Sudemycin alters RNA splicing in primary human CD34+ cells. We first examined the pre-mRNA splicing alterations induced by sudemycin D6 in primary human haematopoietic cells. We treated CD34+ haematopoietic progenitor cells isolated from human umbilical cord blood with 1,000 nM of sudemycin D6 or dimethylsulphoxide (DMSO) vehicle control for 6 h *in vitro* and performed whole-transcriptome (RNA-seq) analysis ($n=6$ each, Supplementary Fig. 1). We identified robustly altered gene expression and pre-mRNA splicing patterns induced by sudemycin, as shown by unsupervised clustering of samples using expressed genes (Fig. 1a) and pre-mRNA splice junctions (Fig. 1b), respectively. Our analysis identified 1,030 differentially expressed genes (FDR < 5%, $|\log_2 FC| > 1$) and 18,833 dysregulated splicing events (FDR < 5%, |delta per cent spliced in or PSI ($\Delta\Psi$)| > 10%, Supplementary Data 1 and 2, respectively) that discriminated between sudemycin D6-treated samples and controls. Sudemycin D6 treatment induced altered pre-mRNA splicing with a bias towards increased exon skipping and intron retention (Fig. 1b). However, there was no apparent bias in the sequence motif surrounding splice acceptor sites of cassette exons that were alternatively spliced (Fig. 1c), in contrast to previously observed biases in sequences surrounding alternatively spliced junctions induced by expression of mutant spliceosome proteins U2AF1, SF3B1 and SRSF2 (refs 16–25).

To determine whether particular pathways are enriched for splicing perturbations, we applied GOseq to 6,278 genes with junctions significantly altered by sudemycin D6 treatment (FDR < 5%, $|\log_2 FC| > 2$). While pathway enrichment was minimal (enrichment scores < 2), GOseq analysis indicated that pathways involved in pre-mRNA splicing, RNA processing and transport, cell cycle, as well as ATPase and helicase activity were enriched in splice junctions altered by sudemycin D6 treatment (FDR < 10%; Supplementary Data 3). Genes with sudemycin-altered expression were enriched in pathways involved in receptor and signal transduction activities (FDR < 10%; Supplementary Data 4).

Mutant U2AF1 cells have increased sensitivity to sudemycin. To examine the effects of sudemycin D6 on haematopoietic cells expressing mutant U2AF1, we generated K562 human erythroleukaemia and OCI-AML3 AML cell lines that have stably integrated doxycycline-inducible, FLAG-tagged U2AF1(S34F) or FLAG-tagged U2AF1(WT) to control for U2AF1 overexpression (Supplementary Fig. 2a,b for K562; Fig. 2c,d for OCI-AML3). Mutant U2AF1(S34F)-expressing K562 cells showed reduced survival and lower IC50 ($P<0.0001$, extra sum-of-squares F-test) relative to uninduced mutant U2AF1(S34F) and U2AF1(WT)-expressing control cells (Fig. 2a). These effects were also observed in human OCI-AML3 cell lines expressing mutant U2AF1(S34F) compared with U2AF1(WT)-expressing cells and other control cells ($P<0.003$, extra sum-of-squares F-test; Fig. 2b). Reduced survival of K562 cells in the presence of sudemycin D6 is associated with an altered cell cycle profile: U2AF1(S34F)-expressing K562 cells had a decrease of cells in the S-phase and an increase of cells in the sub-G0/G1 and G2/M phases (Fig. 2c). Furthermore, MDS or AML cells with U2AF1(S34F) mutations treated *in vitro* with sudemycin D1, a sudemycin compound very similar to D6, showed an increased sensitivity to sudemycin (reduced S-phase) relative to control MDS/AML cells without spliceosome gene mutations (Fig. 2d). In contrast, treatment of MDS/AML patient cells with the chemotherapeutic drug daunorubicin (not predicted to disrupt splicing) showed no specificity for mutant U2AF1(S34F) samples compared with controls (Supplementary Fig. 2e). In addition, human CD34+ cells expressing U2AF1(S34F) showed increased sensitivity to another splicing modulator drug (E7107) similar to sudemycin (Supplementary Fig. 2f).

Sudemycin reduces mutant U2AF1 progenitor expansion *in vivo*. We next examined the effect of sudemycin treatment *in vivo* on mutant U2AF1(S34F)-induced phenotypes using our previously described U2AF1(S34F) transgenic mouse model[19]. We induced U2AF1(S34F) or U2AF1(WT) transgenes for 7 days in the bone marrow cells of transplanted mice (to study haematopoietic cell-intrinsic effects) and treated mice concurrently with sudemycin D6 (50 mg kg^{-1} per day) or vehicle for 5 of

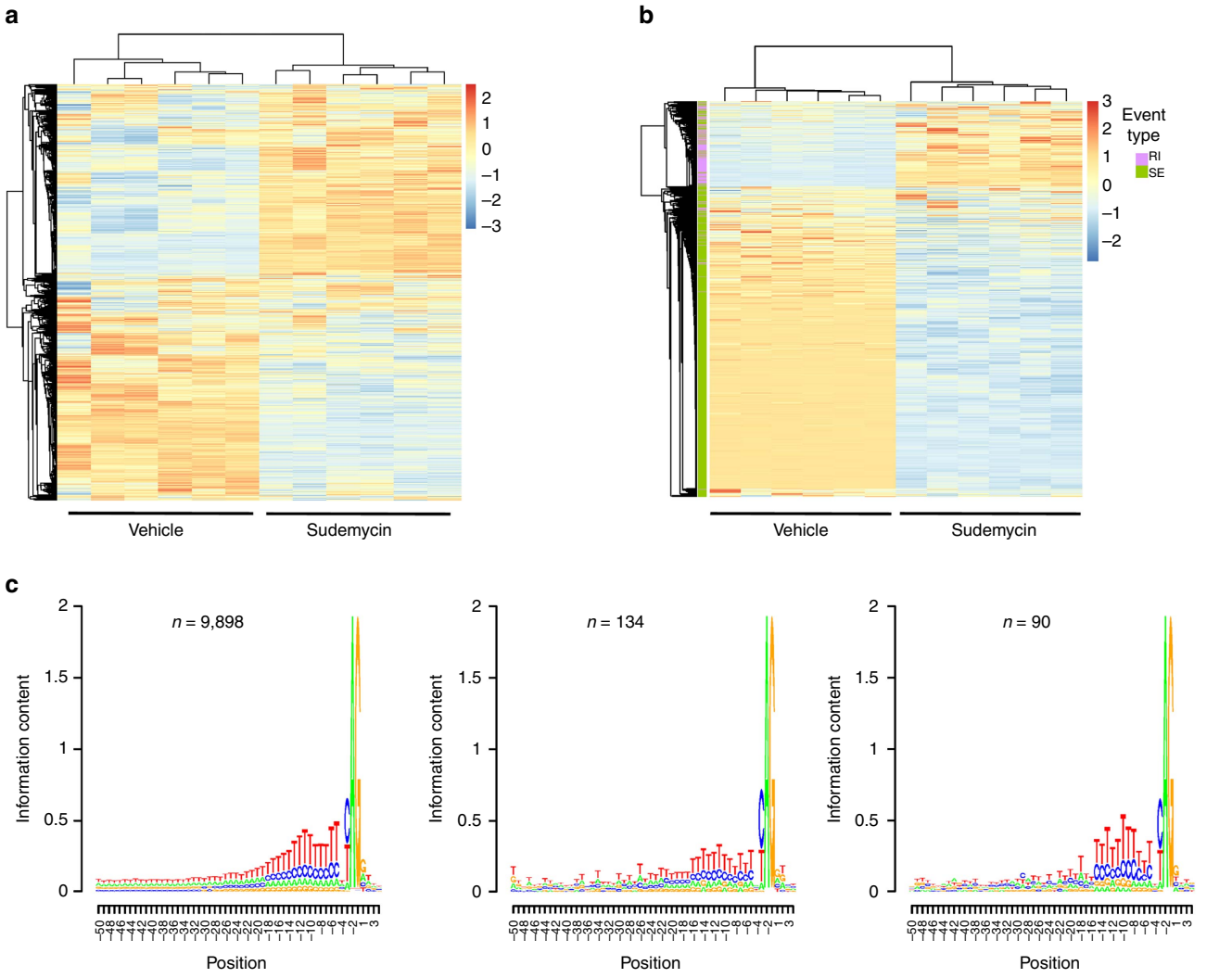


Figure 1 | Sudemycin D6 alters gene expression and pre-mRNA splicing in primary human CD34+ haematopoietic cells. Whole-transcriptome (that is, RNA-seq) analysis was performed on CD34+ cells isolated from human umbilical cord blood following treatment of samples with 1,000 nM Sudemycin D6 or DMSO vehicle for 6 h ($n=6$). Unsupervised hierarchical clustering of (**a**) expressed genes and (**b**) splice junctions. Skipped exons (SE, green) and retained introns (RI, purple) event types are visualized. Values are *z*-scores computed from regularized logarithm values for genes and from per cent spliced in (PSI or Ψ) values for splicing events. (**c**) Intronic sequence contexts of cassette exon 3′ splice sites skipped more often in sudemycin- relative to vehicle-treated cells (FDR < 5%, $|\Delta\Psi| > 10\%$, left panel) or skipped more often in vehicle-relative to sudemycin-treated cells (FDR < 5%, $|\Delta\Psi| > 10\%$, middle panel), along with a context of unperturbed control exons (FDR > 50%, $|\Delta\Psi| < 0.1\%$, right panel). Position is relative to the first base in the exon.

those days; see schema (Fig. 3a). Sudemycin D6 treatment of transplanted mice showed an attenuation of the previously described[19] mutant U2AF1(S34F)-induced haematopoietic progenitor cell expansion by colony-forming unit (CFU-C) assay (Fig. 3b) and by flow cytometry for lineage-, c-Kit+, Sca1+ (KLS) cells (Fig. 3c) when compared with control U2AF1 mutant mice treated with vehicle and mice transplanted with U2AF1(WT)-expressing bone marrow. The attenuation of mutant U2AF1-induced progenitor expansion by sudemycin-treated mice is associated with increased Annexin V+ staining of KLS cells (Fig. 3d).

Sudemycin and U2AF1 (S34F) splicing effects can be cumulative. To investigate the potential genotype-specific effects of sudemycin treatment on splice isoform expression, we performed whole-transcriptome sequencing (RNA-seq) on U2AF1(S34F)- and U2AF1(WT)-recipient mouse bulk bone marrow cells following *in vivo U2AF1* transgene induction and treatment with sudemycin D6 (50 mg kg^{-1} per day for 5 days) or vehicle (Supplementary Fig. 3a,b). RNA was harvested 18 h after the last drug treatment (similar to described above; schema shown in Fig. 3a). Sudemycin D6 treatment at this dose and schedule does not markedly skew the mature lineage distribution within bulk bone marrow of mutant or wild-type (WT) U2AF1 transgenic mice (Supplementary Fig. 3c). Using an unsupervised approach, we observed that sudemycin D6 perturbs splicing in both mutant U2AF1(S34F) and U2AF1(WT)-expressing bone marrow cells (Supplementary Data 5–9); this is visualized by the segregation of samples according to genotype and treatment within a principal component analysis (PCA) of cassette exon (Fig. 4a) and retained intron (Supplementary Fig. 4a) splicing events. Furthermore, the splicing bias observed in human cells treated with sudemycin (described above) is recapitulated in U2AF1(WT) mouse cells: sudemycin D6 induces exon skipping more often than exon inclusion relative to vehicle (388 of 657 significant (FDR < 10%, $|\Delta\Psi| > 1\%$) events; $P < 2 \times 10^{-6}$, one-sided binomial test), as well as intron retention more often than removal (98 of 145 significant (FDR < 10%,

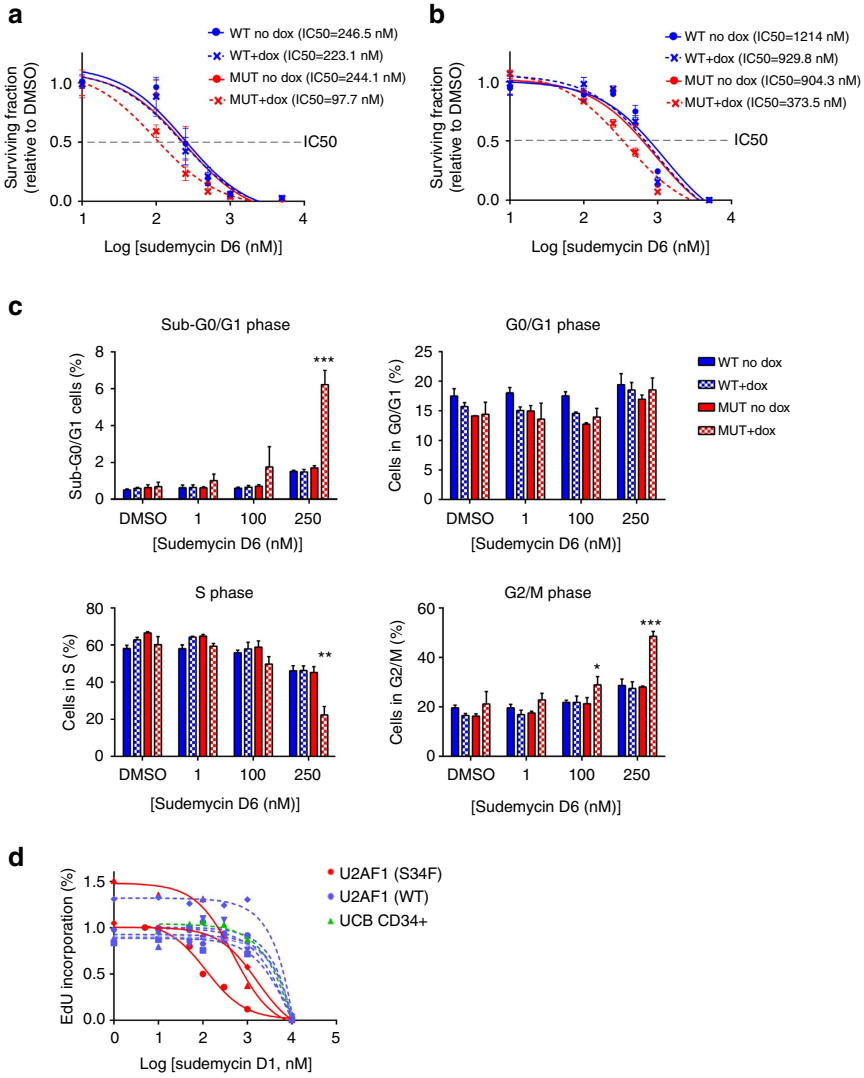


Figure 2 | Mutant U2AF1(S34F)-expressing cells display increased sensitivity to sudemycin D *in vitro*. (**a**) K562 cells ($n=6$ for control groups, $n=9$ for U2AF1(S34F) treated with doxycycline) or (**b**) OCI-AML3 cells ($n=3$ for all groups) with stably integrated, doxycycline-inducible U2AF1(WT) or mutant U2AF1(S34F) were cultured with increasing concentrations of sudemycin D6 concurrently with doxycycline (250 ng ml^{-1}, where indicated) for 5 days following 2 days of initial induction of mutant or WT U2AF1; total cell numbers were measured. The surviving fraction of cells is shown. IC50, inhibitory concentration at 50% of maximum cell survival. (**c**) K562 cell cycle phases were determined using BrdU/7AAD ($n=3$, representative of two experiments; $^{*}P<0.05$, $^{**}P<0.01$, $^{***}P<0.001$ statistics calculated with two-tailed *t*-tests of MUT + dox samples compared with each control group at a given concentration of sudemycin D6, and the least significant value is given for each group; mean values with s.d. shown). (**d**) Primary human MDS or AML cells (both mutant U2AF1(S34F) samples ($n=3$) and those wild type for *U2AF1* ($n=6$)) or normal umbilical cord blood CD34+ cells ($n=1$) were cultured on irradiated HS27 stroma, and proliferation (EdU incorporation) was measured after 3 days of exposure to increasing concentrations of sudemycin D1.

$|\Delta\Psi|>1\%$) events; $P<1.4\times10^{-5}$, one-sided binomial test). As in human CD34+ cells, the sudemycin-induced changes were not associated with an apparent sequence motif (Supplementary Fig. 4b); however, we did observe the previously reported increase in a T in the -3 position of the intronic 3′ splice acceptor site of exons more commonly skipped in mutant U2AF1(S34F) cells[16–19] (Supplementary Fig. 4c). In addition, we defined 'high-confidence' sets of U2AF1(S34F) and sudemycin targets, and subsets of those had a high validation rate in orthogonal experimental (NanoString[26]) and statistical (edgeR[27]) platforms (Supplementary Information and Supplementary Data 10 and 11).

Next, we examined potential interactions between the drug and mutation within cells, focusing on exon skipping events. Along these lines, we observed that the exon skipping effects induced by sudemycin D6 (relative to vehicle) within U2AF1(WT)-expressing cells are highly correlated with the drug effects in U2AF1(S34F)-expressing cells ($R^2=0.8$, $P<2.2\times10^{-16}$, F-test, events significant in both comparisons (FDR<10%), Fig. 4b). The vast majority of these events are concordant (in the same direction of induced change with similar magnitude) across genotypes (slope of the regression line = 0.75), suggesting that sudemycin treatment results in similar splicing alterations in these targets in both mutant U2AF1 and WT cells (Fig. 4b). We further assessed drug/genotype interaction using a statistical linear model: of 32,529 dysregulated splicing events (across all event types), only 136 showed statistically significant evidence of interaction (that is, synergy or antagonism; DEXSeq, FDR<10%). However, when sudemycin D6 and mutant U2AF1(S34F) dysregulate a

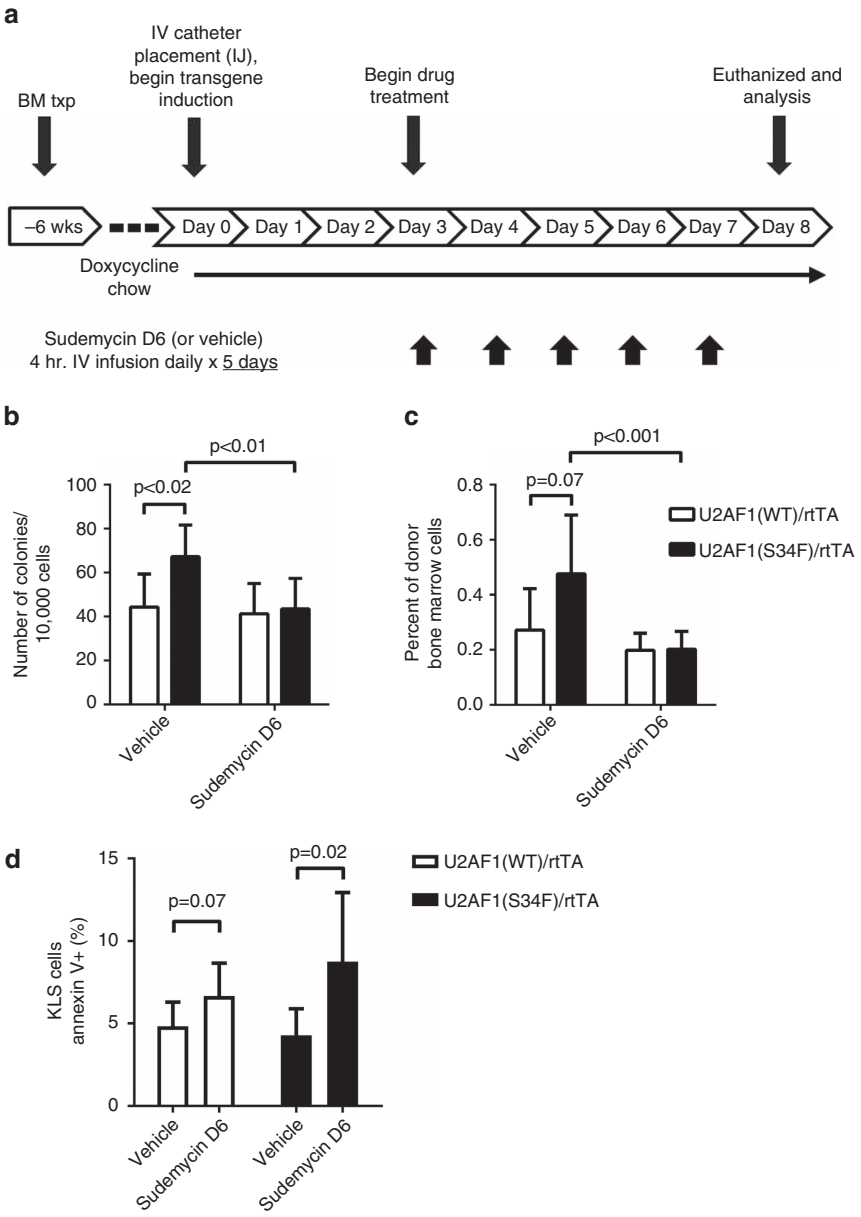


Figure 3 | Sudemycin D6 treatment attenuates mutant U2AF1-induced progenitor cell expansion in U2AF1(S34F) transgenic mice. (**a**) Schema of sudemycin treatment of transgenic mice *in vivo*. Doxycycline-inducible U2AF1(S34F) or U2AF1(WT) transgenic mouse bone marrow was transplanted into recipient mice. Intrajugular (IJ) catheters were placed for IV drug infusions, which were performed over 4 h for 5 days. (**b**) Haematopoietic progenitor CFU-C colony-forming assay and (**c**) flow cytometry for haematopoietic stem and progenitor (HSPC) cell surface markers (c-Kit +, lineage −, Sca1+, KLS, right panel) on U2AF1(WT)- or (S34F) mutant-recipient mouse bone marrow following treatment is shown ($n=6$ for both U2AF1(WT) conditions, $n=7$ for vehicle-treated U2AF1(S34F), $n=11$ for sudemycin D6-treated U2AF1(S34F)). (**d**) Annexin V + KLS cells were quantified in the bone marrow of mice following *in vivo* sudemycin treatment as described above ($n=7$–9). Data pooled from two independent experiments (**b**–**d**); statistics calculated using two-tailed *t*-tests for each comparison shown; mean values with s.d. shown.

junction in the same direction (for example, both increasing exon skipping), this results in a cumulative effect in a sudemycin-treated mutant cell that is greater than the effect induced by sudemycin treatment of WT cells (indicated by red and blue colour in Fig. 4b–d). As an example, U2AF1(S34F) expression induces increased exon skipping in a *4932438A13Rik* splice junction compared with U2AF1(WT) in vehicle-treated cells ($\Delta\Psi=0.205$, middle two columns in Fig. 4c). Sudemycin D6 also increases exon skipping of the same junction in U2AF1(WT) ($\Delta\Psi=0.188$) and U2AF1(S34F) ($\Delta\Psi=0.204$) cells (that is, sudemycin-induced exon skipping is independent of genotype; Fig. 4c). Therefore, as expected, the cumulative effect induced by both mutation and drug treatment ($\Delta\Psi=0.409$) exceeds the individual effects of mutant U2AF1 expression or sudemycin treatment alone, ultimately resulting in different levels of exon skipping in sudemycin-treated cells expressing mutant U2AF1 versus WT cells (Fig. 4c). The cumulative effect of mutation and drug can also be observed in splicing events that result in increased exon inclusion (Fig. 4d). Together, these data signify that the effects of sudemycin and U2AF1(S34F) on splicing in the same cell can be cumulative—that is, greater than the separate effects of drug or mutant.

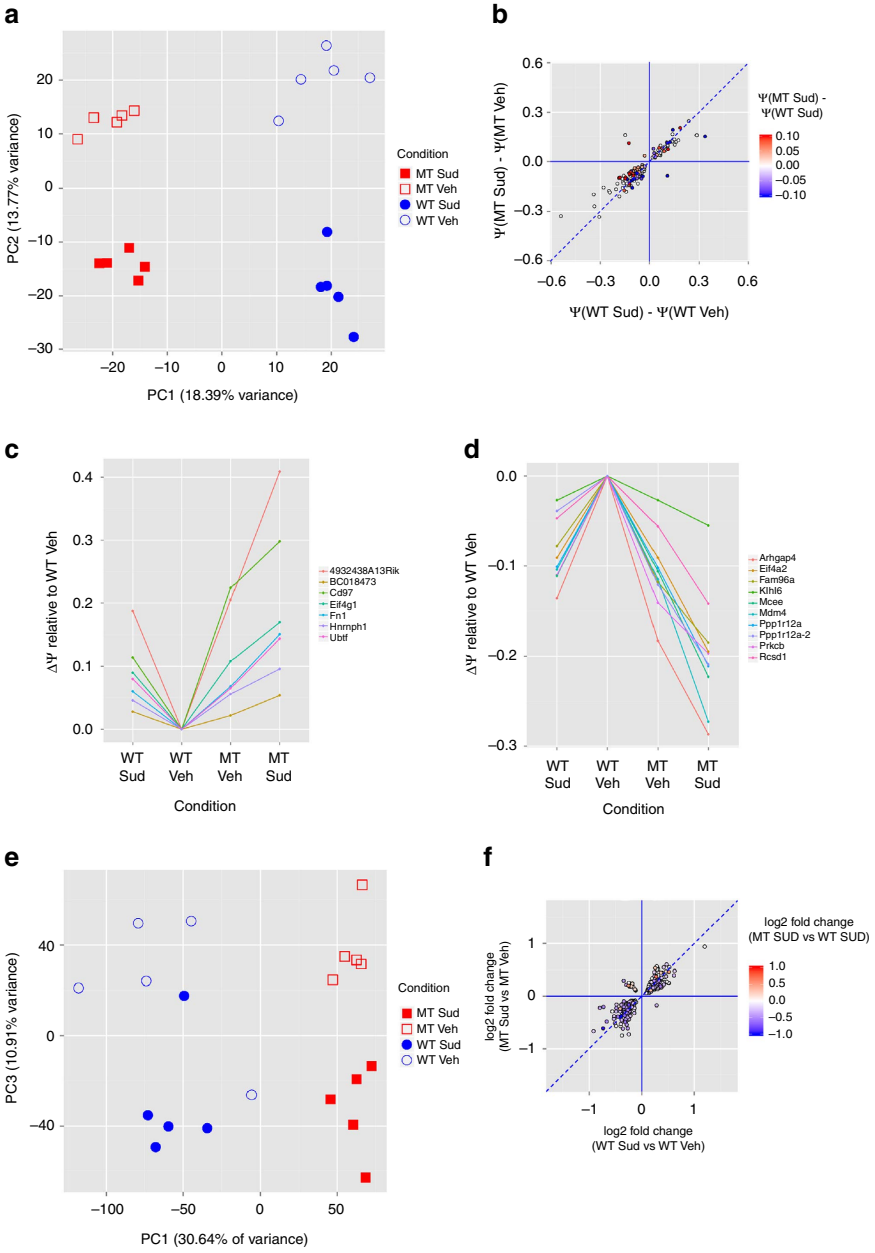


Figure 4 | Sudemycin D6 treatment alters splicing and gene expression in mutant U2AF1-haematopoietic cells. RNA sequencing and transcriptome analysis was performed on RNA harvested from mouse bone marrow cells expressing mutant U2AF1(S34F) or U2AF1(WT) following treatment of mice with sudemycin D6 (50 mg kg^{-1}) or vehicle control for 5 days ($n=5$ per genotype and treatment). (**a**) PCA of normalized expression of skipped cassette exon events. (**b**) Correlation between sudemycin-induced changes relative to vehicle treatment in cassette exons of U2AF1(WT) cells $\left(\Delta\Psi^{WT}_{Sud}=\Psi_{WT,Sud}-\Psi_{WT,Veh}\right)$ (FDR<10%) and of mutant U2AF1(S34F) cells $\left(\Delta\Psi^{MT}_{Sud}=\Psi_{MT,Sud}-\Psi_{MT,Veh}\right)$ (FDR<10%; $R^2=0.8$; $P<2.2\times10^{-16}$, F-test). Dashed line indicates similar effects in U2AF1(WT) and U2AF1(S34F) cells $\left(\Delta\Psi^{WT}_{Sud}=\Delta\Psi^{MT}_{Sud}\right)$. Colour scale indicates cumulative splicing changes of U2AF1(S34F) expression with sudemycin treatment (that is, $\Delta\Psi^{Sud}_{MT}=\Psi_{MT,Sud}-\Psi_{WT,Sud}$), with red being a positive change, blue being a negative change, and white being no difference between sudemycin-treated U2AF1(S34F) and U2AF1(WT) cells. (**c**,**d**) Delta PSI ($\Delta\Psi$) for each condition on the horizontal axis relative to vehicle-treated U2AF1(WT) cells ($(\Psi_{WT,Veh})=0$) for events that are significantly dysregulated across all five comparisons (see Supplementary Methods), have concordant sudemycin-induced dysregulation across genotype (that is, $\Delta\Psi^{WT}_{Sud}$ and $\Delta\Psi^{MT}_{Sud}$ have the same direction), and in which the sudemycin effect is exacerbated by mutation. The cumulative effect on cassette exons induced by both mutant U2AF1 and sudemycin D6 treatment relative to vehicle-treated WT cells, $\Delta\Psi_{cum}$, may result in increased exon skipping (positive values, **c**) or increased exon inclusion (negative values, **d**). (**e**) PCA of normalized expression of expressed genes. (**f**) Correlation between sudemycin-induced changes relative to vehicle treatment (expressed as $\log_2$ fold changes, FDR<10%) in U2AF1(WT) and U2AF1(S34F) cells ($R^2=0.8$; $P<2.2\times10^{-16}$, F-test). Dashed line indicates $\log_2$ fold changes induced by sudemycin are the same in U2AF1(WT) and U2AF1(S34F) cells. Colour scale indicates cumulative contribution of mutant U2AF1 expression to sudemycin-induced gene expression changes, that is, $\log_2$ fold change of gene expression altered by sudemycin in mutant U2AF1(S34F) cells relative to U2AF1(WT) cells, with red being a positive change, blue being a negative change and white being no difference. U2AF1 mutant (MT), sudemycin D6 (Sud), U2AF1 WT, vehicle (Veh), principal component 1 (PC1), principal component 2 (PC2), principal component 3 (PC3), per cent spliced in (PSI, Ψ).

Sudemycin induces gene expression alterations. Sudemycin treatment *in vivo* also results in altered gene expression in mutant U2AF1 mouse bone marrow cells compared with U2AF1(WT)-expressing control cells (Supplementary Data 12). As with splicing alterations, mouse bone marrow samples segregate by genotype and by treatment in an unsupervised analysis of gene expression (Fig. 4e). As seen at the junction level, there is a high-degree of correlation between sudemycin-induced, gene-level effects (direction and magnitude relative to vehicle) across mutant and WT genotypes ($R^2 = 0.8$, $P < 2.2 \times 10^{-6}$, F-test, events significant (FDR < 10%) in both comparisons, Fig. 4f). We also observed cumulative effects of mutation and drug on gene expression (colour in Fig. 4f), which may result in downstream cellular pathway changes. Using GOseq, we identified pathways that were most enriched for differentially expressed genes in mutant cells treated with sudemycin D6 relative to the other genotype and treatment groups (FDR < 10%). We found mutant cells treated with sudemycin D6 were enriched in biologic pathways related to immune and inflammatory responses, antigen processing and presentation, cytokine production, leukocyte differentiation, cell death and apoptotic processes, when compared with the other genotype and treatment groups (Supplementary Data 13).

Discussion

We provide evidence suggesting that U2AF1(S34F)-expressing cells are sensitive to the splicing modulator drug sudemycin. Haematopoietic cells expressing mutant U2AF1 have reduced survival and cell cycle changes following sudemycin D6 treatment *in vitro*. *In vivo*, sudemycin D6 is capable of attenuating mutant U2AF1-associated expansion of haematopoietic progenitors in transgenic mice that is associated with increased cell death. In addition, while the effects of sudemycin D6 treatment on splicing and gene expression can be independent of the effects of mutant U2AF1 expression, sudemycin D6 treatment can also modulate splicing changes induced by mutant U2AF1(S34F) to create cumulative effects on cells *in vivo*. Together, these data indicate that mutant U2AF1(S34F)-expressing cells may have a therapeutic vulnerability to splicing modulator drugs such as sudemycin.

Previous studies using splicing modulator drugs, as well as non-pharmacological methods to target splicing factors, have indicated that the spliceosome is a promising therapeutic target for many cancer cell types[13,14,28–31]. Recent studies have suggested that cancers with spliceosome gene mutations may have an increased sensitivity to splicing modulator drugs like sudemycin;[32–34] this study provides further evidence to support this hypothesis. Specifically, we show that mutant U2AF1 cells are sensitive to *in vivo* treatment with sudemycin. Using unbiased RNA-seq, we show that sudemycin D6 has effects on mutant U2AF1-associated splicing changes, resulting in different levels of transcript isoforms in cells expressing mutant U2AF1 compared with WT following sudemycin treatment. Ultimately, cumulative changes in isoform expression could be the cause of the mutant U2AF1-specific responses to sudemycin treatment that we observe *in vitro* and *in vivo*. Identification of critical downstream targets of sudemycin *in vivo* will likely require examination of various cellular populations within the bone marrow (including haematopoietic stem and progenitor cells) and harvesting cells at more immediate time points following drug treatment—both limitations of the current study. Future studies will focus on generating mutant U2AF1-expressing leukaemias in mice to test the efficacy of sudemycin on fully transformed haematopoietic tumours, as has recently been reported using E7107 to treat Srsf2(P95H)-expressing leukaemias[33].

Our data and others highlight several possible mechanisms for the sensitivity of spliceosome mutant samples to splicing modulator therapy. Pathway analysis of differentially expressed genes induced by sudemycin revealed enrichment in inflammatory signalling pathways. This is consistent with the enrichment in biological pathways related to cytokine and immune signalling observed following *in vivo* treatment of Srsf2(P95H) mutant mice with the splicing modulator E7107 (ref. 33), raising the possibility that spliceosome mutant-specific phenotypes observed in mice following splicing modulator drug treatment may be driven by an altered inflammatory response in mutant cells. Whether the altered inflammatory response in mutant cells treated with splicing modulator drugs is a direct result of mutant-altered splicing or gene expression could be explored in future studies. Alternatively, it is also possible that the cumulative effect of sudemycin and mutant U2AF1 expression on pre-mRNA splicing may simply create a state of 'spliceosome sickness' in cells by exceeding a tolerable threshold of splicing perturbations and their downstream consequences. Along these lines, other splicing modulator drugs have been shown to cause increased intron retention, R loop formation, DNA damage and cell death[28,29,31,35]. Spliceosome gene mutations and splicing modulator drugs may both induce these consequences in a cell, and the cumulative effect of the drug and mutation may create a toxic intracellular milieu.

Exploring the clinical utility of splicing modulator therapies in MDS patients with spliceosome mutations who have failed current therapies is warranted and currently being pursued (NCT02841540), as these patients have few treatment options. Whether excessive toxicity will occur in WT cells treated with newer splicing modulators is a major question. Initial phase I clinical trial studies using a splicing modulator, E7107, showed toxicities in some patients with solid tumours, including ocular toxicity (NCT00459823, NCT00499499). The mechanism for the toxicity is not known. Moving forward, mutant U2AF1 and other preclinical models of spliceosome mutations will be valuable in further testing the *in vivo* efficacy and toxicity of drugs that modulate splicing through various mechanisms. Collectively, our data suggest that a mutant splicing factor together with a splicing modulator drug like sudemycin may create a unique cellular toxicity that could be exploited for therapeutic purposes.

Methods

Isolation and drug treatment of CD34+ haematopoietic cells. Mononuclear cells from human umbilical cord blood were separated by Ficoll gradient centrifugation (each n value is a pooled set of four individual umbilical cord blood samples). CD34+ cells were enriched using autoMACS-positive selection (CD34 MicroBead Kit, Miltenyi Biotec) according to the manufacturer's instructions to achieve >90% purity of CD34+ cells. Following isolation, CD34+ cells were cultured in *X-Vivo* media (Lonza) with cytokines (SCF, FLT3L, IL3 and TPO) overnight. For whole-transcriptome sequencing with sudemycin, either sudemycin D6 (1,000 nM) or the vehicle DMSO was added to cells for 6 h, and cells were then harvested for RNA. Tissue acquisition was performed per protocol approved by the Washington University School of Medicine Institutional Review Board.

RNA-seq of CD34+ cells treated with sudemycin. RNA was isolated from human CD34+ cells using the miRNEasy Kit (Qiagen), and removal of genomic DNA was performed via a Turbo DNA-Free Kit (Ambion). Ribosomal RNA was removed using Ribozero (Epicenter), followed by cDNA preparation and generation of stranded libraries using the TruSeq Stranded Total RNA Sample Prep Kit (Illumina). Sequencing was performed on the HiSeq2500 platform (Illumina) to generate 2 × 125 bp paired-end reads. RNA-seq data were deposited in NCBI dbGAP (phs000159.v9).

RNA-seq analysis of human CD34+ cells treated with sudemycin. Analysis of (stranded) RNA-seq data generated from primary human CD34+ haematopoietic cells treated with sudemycin D6 was performed using the genome modelling system[36]. Reads were aligned to the human genome (hg19/NCBI build 37) using TopHat[37] version 2.0.8, with annotations provided by Ensembl[38] version 67. All downstream bioinformatic and statistical analyses, including calculation of *P* values using Fisher's exact test and simulation, were performed in R[39] and python. The heatmap of gene expression was created using normalized expression values output by DESeq2 package (version 1.6.3 (ref. 40)) that was subsequently *z*-scored. DESeq2 was used to detect differentially expressed genes, and gene-set enrichment analysis was then performed by applying GOseq[41] (version 1.18.0) to differentially expressed genes (FDR < 5%, $|\log_2 FC| > 1$, DESeq2). Additional details of RNA-seq analysis of gene expression are further described in the Supplementary Methods.

Alternative splicing events were computed via rMATS[42] (version 3.0.8). We generated the heatmap of splice junctions using 'per cent spliced in' (PSI or Ψ) values for exon skipping and intron retention events. For downstream analyses, namely reported dysregulated events and visualization of sequence contexts below, *P* values for events passing the filter were re-adjusted for multiple hypothesis testing using the method of Benjamini and Hochberg[43]. A positive ΔPSI value indicates an event that was spliced in more often in the vehicle-treated relative to the sudemycin-treated samples. Sequence contexts of splice sites in exon skipping and intron retention events were visualized using seqLogo version 1.32.1. Dysregulated events visualized were those with $|\Delta PSI| > 10\%$ and having a post-filtering re-adjusted FDR < 5% from rMATS. Dysregulated splicing junctions were determined by DEXSeq[4] version 1.12.2. Only expressed junctions were analysed, and gene-set enrichment analysis was performed by applying GOseq to dysregulated junctions (FDR < 5%, $|\log_2 FC| > 1$, DEXSeq). All of these are described in more detail in the Supplementary Methods.

Creation of inducible U2AF1(S34F) and U2AF1(WT) cell lines. Doxycycline-inducible FLAG-tagged WT U2AF1 or FLAG-tagged mutant U2AF1(S34F) lentiviral expression plasmids were previously described[4], and lentivirus was generated in 293T cells with the packaging plasmids pMD-G, pMD-Lg and REV. Concentrated virus (multiplicity of infection (MOI) of 3 for K562, MOI of 5 for OCI-AML3) was used to transduce K562 cells (ATCC, CCL243) or OCI-AML3 (DSMZ, ACC 582). Transduced cells (marked by green fluorescent protein) were isolated by flow cytometry cell sorting. Expression of mutant or WT U2AF1 was induced using the indicated concentrations of doxycycline hyclate (Sigma, St Louis, MO) in water.

Cell line counting and BrdU incorporation assays. K562 (or OCI-AML3) cells in culture were seeded at ~200,000 cells per ml with different concentrations of Sudemycin D6 (or with DMSO control) following 48 h of initial doxycycline treatment to induce mutant U2AF1(S34F) or U2AF1(WT) expression. Cells were counted on Days 0, 3 and 5 of drug treatment using flow cytometry counting particles (Spherotech) as per the manufacturer's recommendations, along with propidium iodide (Millipore) to exclude dead cells from counts. Day 5 data were graphed in Graphpad Prism (Graphpad Software Inc.) using a nonlinear regression best fit line of the log (drug) versus response for each genotype and treatment. Statistical differences between curves were examined by the extra sum-of-squares F-test function in Graphpad Prism. For K562 cells, 5-bromodeoxyuridine (BrdU) incorporation was performed using the BrdU Flow Kit (APC, #559619, BD Biosciences) as per the manufacturer's instructions using a 45 min pulse of BrdU of cells in culture on Day 5 of drug treatment.

Culture of MDS/AML cells and EdU incorporation assay. Primary MDS and AML cells were cultured on irradiated (4,000 cGy) HS27 stroma as described previously[45]. Briefly, MDS or AML cells were cultured on stroma for 2–3 days and then treated with increasing doses of sudemycin D1 (or daunorubicin in Supplementary Information). MDS and AML cells were cultured with drug for 3 days and analysed for cell proliferation by EdU incorporation assay (Invitrogen) via flow cytometry. All experiments were performed on 96-well plates. Studies were performed per protocol approved by the Washington University School of Medicine Institutional Review Board and with patients' consent for sample use.

Murine bone marrow transplant and drug infusion. To generate mice for each experiment, 1×10^6 transgenic mouse donor bone marrow cells from two to three mice pooled (CD45.2) were transplanted into at least five lethally irradiated (1,100 rads) congenic WT recipient mice (C57BL/6 × 129S4Sv/Jae)F1 (CD45.1/CD45.2) per genotype, as previously described[19]. Donor mice were between 8 and 12 weeks of age, and recipient mice ranged from 6 to 12 weeks of age; donor and recipient mice were sex-matched (both sexes were used in experiments). Donor chimerism was confirmed ≥6 weeks post transplantation to ensure engraftment of transgenic bone marrow. Post-engraftment, intrajugular catheters were surgically placed for intravenous drug infusions of mice over 4 h daily for 5 days with either sudemycin D6 (50 mg kg^{-1}) or vehicle control (HP-β-CD (2-hydroxypropyl)-β-cyclodextrin in phosphate buffer pH 7.4). This dose was determined by prior studies[13]. Mice received doxycycline chow to induce U2AF1(S34F) or U2AF1(WT) transgene on the day of surgery, and drug infusion began 2 days later. Mice of each group were randomly selected for treatment with sudemycin or vehicle control. Investigators were not blinded to the group allocation during the experiment or analysis. Animals were excluded from analysis if their catheters did not remain patent during the entire treatment period. Mice were euthanized for analysis the day following the last drug infusion. Sample sizes for experiments were chosen to allow for statistical comparison between groups. All mouse procedures were performed according to the protocols approved by the Washington University Animal Studies Committee.

Mouse haematopoietic progenitor cell assay. Methylcellulose progenitor CFU-C assays were performed using Methocult GF M3434 (Stem Cell Technologies). Bulk bone marrow cells were obtained from experimental mice, and red blood cells were lysed before plating of 10,000 bone marrow cells per 1.3 ml media; each sample was evaluated in duplicate. Progenitor colonies (defined as ≥40 cells per colony) were counted following 7 days culture at 37 °C with 5% CO_2.

Mouse haematopoietic cell flow cytometry. For flow cytometry, all antibodies are from eBioscience (unless indicated), and catalogue number provided (if available). For haematopoietic progenitor/stem cells, we used the following antibodies (volume of antibody used in 200 µl of fluorescence-activated cell sorting (FACS) buffer for staining is also indicated): CD45.1-APC (#17-0453, 3 µl), CD45.2-PE (#12-0454, 3 µl), Biotin-conjugated lineage (Gr-1 (#13-5931, 0.25 µl), Cd3e (#13-0032, 0.5 µl), B220 (#13-0452, 0.5 µl), Ter119 (#13-5921, 0.5 µl) and CD41(#13-0411, 1 µl)), streptavidin secondary-eFluor605NC, c-Kit-APCeFluor780 (#47-1172, 1.5 µl), Sca1-PerCP-Cy5.5 (#45-5981, 0.5 µl). Flow cytometry for Annexin V+ staining of KLS cells was performed following initial staining of KLS cells as described above using Annexin V-APC (2.5 µl, BD Biosciences, #550474) incubated in 1× Annexin V binding buffer (BD Biosciences). All other incubations occurred in FACS buffer. All flow cytometry analyses were performed using FACScan or Gallios cytometers (BD Biosciences) and analysed using the FlowJo software (FlowJo, LLC, Ashland, OR, USA).

RNA-seq of mouse bone marrow cells *in vivo*. RNA was isolated and prepared as described above for human cells, using the miRNEasy Kit (Qiagen) followed by removal of genomic DNA via a Turbo DNA-Free Kit (Ambion) and ribosomal RNA using Ribozero (Epicenter). The cDNA preparation and generation of stranded libraries was performed using the TruSeq Stranded Total RNA Sample Prep Kit (Illumina). Sequencing was performed on the HiSeq2500 platform (Illumina) to generate 2 × 126 bp paired-end reads. The RNA-seq data have been deposited in NCBI's Gene Expression Omnibus[46] and are accessible through the GEO Series accession number GSE89834.

RNA-seq analysis of mouse bone marrow cells *in vivo*. Analysis of (stranded) RNA-seq data generated from U2AF1(S34F) and U2AF1(WT) murine cells treated with sudemycin D6 or vehicle was performed similarly to the human CD34+ cell analysis described above, with some differences as follows. Alignment again utilized genome model system, and reads were aligned to the mouse genome (mm9/NCBI build 37). 'Per cent spliced in' (PSI or Ψ) values were calculated for all four conditions ({U2AF1(S34F), U2AF1(WT)} × {sudemycin D6, vehicle}) using rMATS. PCA was performed independently on events annotated by rMATS as skipping cassette exons or retaining introns using the *z*-scored Ψ values of these events.

To quantitate the simultaneous effect of treatment and genotype for each splicing event, we calculated the change in per cent spliced in values for the four pairwise comparisons in which the genotype (alternately, treatment) was the same in the pair, but in which the treatment (alternately, genotype) differed. We refer to the unchanged condition as the 'context' and to the two conditions that differ as 'A' and 'B' and denote the corresponding change in the per cent spliced in values as $\Delta\Psi^{\text{context}}_{A-B} = \Psi_{\text{context},A} - \Psi_{\text{context},B}$. In addition, we defined the cumulative effect in a mutant, drug-treated cell relative to a WT, vehicle-treated cell as $\Delta\Psi_{\text{cum}} = \Psi_{\text{Sud,MT}} - \Psi_{\text{Veh,WT}}$. Each of the five comparisons described were evaluated using rMATS, and *P* values for these events were then adjusted for multiple hypothesis testing as described above. Scatterplots were then plotted of cassette exon-skipping events dysregulated by sudemycin in both U2AF1(WT) and U2AF1(S34F) contexts ($|\Delta\Psi^{\text{WT}}_{\text{SudVeh}}| > 1\%$, $|\Delta\Psi^{\text{MT}}_{\text{SudVeh}}| > 1\%$, FDR < 10% in both comparisons). To highlight cumulative cassette exon-skipping effects, 'trajectories' comparing Ψ values to the 'baseline' $\Psi_{\text{Veh,WT}}$ were plotted. Events that were significant in all five comparisons (that is, $|\Delta\Psi^{\text{WT}}_{\text{SudVeh}}| > 1\%$, $|\Delta\Psi^{\text{MT}}_{\text{Sud-Veh}}| > 1\%$, $|\Delta\Psi^{\text{Veh}}_{\text{MT-WT}}| > 1\%$, $|\Delta\Psi^{\text{Sud}}_{\text{MT-WT}}| > 1\%$, $|\Delta\Psi_{\text{Cum}}| > 1\%$, with FDR < 10% in all comparisons) were plotted; the latter condition ensures that visually discernible differences are statically significant and not attributable to statistical noise. Additional details of this can be found in the Supplementary Methods.

Non-additive (that is, super-additive synergistic or sub-additive antagonistic) interactions between drug and mutation were assessed in DEXSeq by comparing a generalized linear model that included main effects for drug and mutation with a second model that additionally included an interaction term representing drug/mutation synergy or antagonism.

As above, splice site sequence contexts of exon-skipping events dysregulated by sudemycin D6 ($|\Delta\Psi^{WT}_{Sud-Veh}| > 1\%$, FDR < 10%) and corresponding unperturbed control events ($|\Delta\Psi^{WT}_{Sud-Veh}| > 0.1\%$, FDR > 50%) were plotted using seqLogo. Similarly, sequence contexts were displayed for events dysregulated ($|\Delta\Psi^{Veh}_{MT-WT}| > 1\%$, FDR < 10%) or unaffected ($|\Delta\Psi^{Veh}_{MT-WT}| > 1\%$, FDR > 50%) by U2AF1(S34F).

At the gene-level, PCA was performed as described above for human expressed genes. Genes differentially expressed within the above-described five conditions were determined via DESeq2. Scatterplots were then made comparing $\log_2$ fold changes of genes dysregulated by sudemycin in both U2AF1(WT) and U2AF1(S34F) contexts (FDR < 10% in both comparisons). Pathways enriched (FDR < 10%) for genes differentially expressed between U2AF1(S34F), sudemycin-treated cells and U2AF1(WT) and/or vehicle-treated cells were independently determined in a pairwise manner using GOseq. Again, expanded details of this approach can be found in the Supplementary Methods.

Data availability. All relevant data generated in this study are available at data-deposition sites. For human CD34+ cells treated with sudemycin D6 *in vitro*, data are available at NCBI dbGAP (phs000159.v9). For transgenic mice expressing mutant U2AF1(S34F) or U2AF1(WT) and treated with sudemycin D6 *in vivo*, data are available at Gene Expression Omnibus (GSE89834).

References

1. Ma, X. Epidemiology of myelodysplastic syndromes. *Am. J. Med.* **125,** S2–S5 (2012).
2. Cogle, C. R., Craig, B. M., Rollison, D. E. & List, A. F. Incidence of the myelodysplastic syndromes using a novel claims-based algorithm: high number of uncaptured cases by cancer registries. *Blood* **117,** 7121–7125 (2011).
3. Troy, J. D., Atallah, E., Geyer, J. T. & Saber, W. Myelodysplastic syndromes in the United States: an update for clinicians. *Ann. Med.* **46,** 283–289 (2014).
4. Yoshida, K. *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478,** 64–69 (2011).
5. Graubert, T. A. *et al.* Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat. Genet.* **44,** 53–57 (2011).
6. Papaemmanuil, E. *et al.* Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl. J. Med.* **365,** 1384–1395 (2011).
7. Visconte, V. *et al.* SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. *Leukemia* **26,** 542–545 (2011).
8. Damm, F. *et al.* Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. *Blood* **119,** 3211–3218 (2012).
9. Thol, F. *et al.* Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood* **119,** 3578–3584 (2012).
10. Walter, M. J. *et al.* Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. *Leukemia* **27,** 1275–1282 (2013).
11. Papaemmanuil, E. *et al.* Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122,** 3616–3627 quiz 3699 (2013).
12. Haferlach, T. *et al.* Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* **28,** 241–247 (2014).
13. Lagisetti, C. *et al.* Optimization of antitumor modulators of pre-mRNA splicing. *J. Med. Chem.* **56,** 10033–10044 (2013).
14. Fan, L., Lagisetti, C., Edwards, C. C., Webb, T. R. & Potter, P. M. Sudemycins, novel small molecule analogues of FR901464, induce alternative gene splicing. *ACS Chem. Biol.* **6,** 582–589 (2011).
15. Convertini, P. *et al.* Sudemycin E influences alternative splicing and changes chromatin modifications. *Nucleic Acids Res.* **42,** 4947–4961 (2014).
16. Okeyo-Owuor, T. *et al. U2AF1* mutations alter sequence specificity of pre-mRNA binding and splicing. *Leukemia* **29,** 909–917 (2014).
17. Brooks, A. N. *et al.* A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. *PLoS ONE* **9,** e87361 (2014).
18. Ilagan, J. O. *et al.* U2AF1 mutations alter splice site recognition in hematological malignancies. *Genome Res.* **25,** 14–26 (2014).
19. Shirai, C. L. *et al.* Mutant U2AF1 expression alters hematopoiesis and pre-mRNA splicing *in vivo*. *Cancer Cell* **27,** 631–643 (2015).
20. Przychodzen, B. *et al.* Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. *Blood* **122,** 999–1006.
21. DeBoever, C. *et al.* Transcriptome sequencing reveals potential mechanism of cryptic 3′ splice site selection in SF3B1-mutated cancers. *PLoS Comput. Biol.* **11,** e1004105 (2015).
22. Darman, R. B. *et al.* Cancer-associated SF3B1 hotspot mutations induce cryptic 3′ splice site selection through use of a different branch point. *Cell Rep.* **13,** 1033–1045 (2015).
23. Alsafadi, S. *et al.* Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **7,** 10615 (2016).
24. Kim, E. *et al.* SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* **27,** 617–630 (2015).
25. Zhang, J. *et al.* Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. *Proc. Natl Acad. Sci. USA* **112,** E4726–E4734 (2015).
26. Geiss, G. K. *et al.* Direct multiplexed measurement of gene expression with color-coded probe pairs. *Nat. Biotechnol.* **26,** 317–325 (2008).
27. Robinson, M. D. & Smyth, G. K. Moderated statistical tests for assessing differences in tag abundance. *Bioinformatics* **23,** 2881–2887 (2007).
28. Kaida, D. *et al.* Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol.* **3,** 576–583 (2007).
29. Kotake, Y. *et al.* Splicing factor SF3b as a target of the antitumor natural product pladienolide. *Nat. Chem. Biol.* **3,** 570–575 (2007).
30. Bonnal, S., Vigevani, L. & Valcarcel, J. The spliceosome as a target of novel antitumour drugs. *Nat. Rev. Drug Discov.* **11,** 847–859 (2012).
31. Hsu, T. Y. *et al.* The spliceosome is a therapeutic vulnerability in MYC-driven cancer. *Nature* **525,** 384–388 (2015).
32. Xargay-Torrent, S. *et al.* The splicing modulator sudemycin induces a specific antitumor response and cooperates with ibrutinib in chronic lymphocytic leukemia. *Oncotarget* **6,** 22734–22749 (2015).
33. Lee, S. C. *et al.* Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. *Nat. Med.* **22,** 672–678 (2016).
34. Obeng, E. A. *et al.* Physiologic expression of Sf3b1(K700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. *Cancer Cell* **30,** 404–417 (2016).
35. Wan, Y. *et al.* Splicing function of mitotic regulators links R-loop-mediated DNA damage to tumor cell killing. *J. Cell Biol.* **209,** 235–246 (2015).
36. Griffith, M. *et al.* Genome modeling system: a knowledge management platform for genomics. *PLoS Comput. Biol.* **11,** e1004274 (2015).
37. Kim, D. *et al.* TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14,** R36 (2013).
38. Flicek, P. *et al.* Ensembl 2013. *Nucleic Acids Res.* **41,** D48–D55 (2013).
39. Gentleman, R. C. *et al.* Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* **5,** R80 (2004).
40. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11,** R106 (2010).
41. Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11,** R14 (2010).
42. Shen, S. *et al.* rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc. Natl Acad. Sci. USA* **111,** E5593–E5601 (2014).
43. Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate-a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B Methodol.* **57,** 289–300 (1995).
44. Anders, S., Reyes, A. & Huber, W. Detecting differential usage of exons from RNA-seq data. *Genome Res.* **22,** 2008–2017 (2012).
45. Klco, J. M. *et al.* Functional heterogeneity of genetically defined subclones in acute myeloid leukemia. *Cancer Cell* **25,** 379–392 (2014).
46. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30,** 207–210 (2002).

Acknowledgements

Support was provided by NIH/NHLBI (T32HL007088 to C.L.S.), Barnes-Jewish Hospital Foundation (B.S.W., T.A.G. and M.J.W.), an NIH/NCI SPORE in Leukemia (P50CA171963 to C.L.S., B.S.W., T.A.G. and M.J.W.), an NIH/NCI grant (K12CA167540 to C.L.S. and B.S.W.), a Clinical and Translational Award from the NIH National Center for Advancing Translational Sciences (UL1 TR000448 to B.S.W.), the Edward P Evans Foundation (T.A.G. and M.J.W.), the Lottie Caroline Hardy Trust (T.A.G. and M.J.W.), a Leukemia and Lymphoma Society Scholar Award (M.J.W.) and Translational Research Award (T.A.G.), by NIH grant CA140474 (T.R.W.) and Department of Defense (BM120018; M.J.W.). Support for procurement of human samples was provided by an NIH/NCI grant (P01 CA101937). Technical assistance was provided by the Alvin J Siteman Cancer Center High Speed Cell Sorting Core, the Tissue Procurement Core supported by an NCI Cancer Center Support Grant (P30CA91842), Carla Weinheimer and Mouse Cardiovascular Phenotyping Core in the Center for Cardiovascular Research at Washington University School of Medicine, and the McDonnell Genome Institute (Director, Richard Wilson and Co-Director, Elaine Mardis) and Chris Markovic for sequencing and NanoString experiments, respectively.

E7107 and technical assistance were kindly provided by Silvia Buonamici and Peter Smith (H3 Biomedicine). We are grateful to Drs Tim Ley, Dan Link, and John DiPersio for helpful scientific discussions.

Author contributions

The study was designed by: C.L.S., B.S.W., M.T., T.A.G. and M.J.W. Primary cell/cell line experiments performed by: C.L.S., M.T., R.T., S.K., J.S., A.C. and B.S. Mouse model experiments by: C.L.S., R.T., J.N.L., M.N. and S.K. RNA sequencing by: R.S.F., C.F. and M.O. Bioinformatics analysis by: B.S.W. Sudemycin drug development and synthesis by: C.L. and T.R.W. The manuscript was written and edited by: C.L.S., B.S.W., T.A.G. and M.J.W. All co-authors reviewed and approved the submission.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Shirai, C. L. *et al.* Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. *Nat. Commun.* **8,** 14060 doi: 10.1038/ncomms14060 (2017).

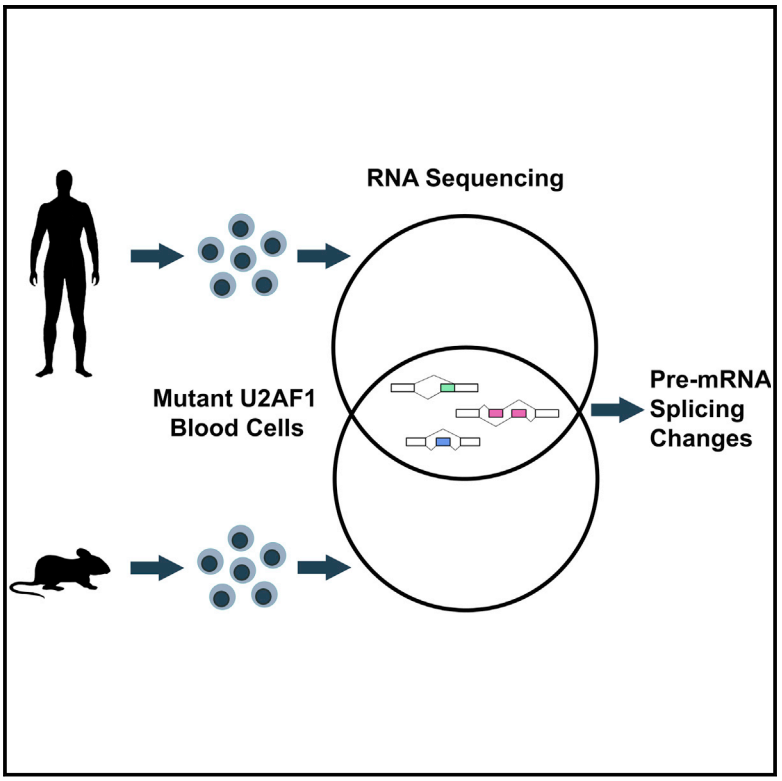
Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Article

Cancer Cell

Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo

Graphical Abstract



Authors

Cara Lunn Shirai, James N. Ley, ..., Timothy A. Graubert, Matthew J. Walter

Correspondence

mjwalter@dom.wustl.edu

In Brief

Shirai et al. show that the *U2AF1* mutation most commonly found in myelodysplastic syndromes (MDS) alters pre-mRNA splicing in RNA processing genes, ribosomal genes, and recurrently mutated MDS and acute myeloid leukemia-associated genes in hematopoietic progenitor cells and affects hematopoiesis.

Highlights

- U2AF1 mutant mice develop leukopenia and hematopoietic progenitor cell expansion
- U2AF1 mutant stem cells are outcompeted by wild-type stem cells in vivo
- Mutant U2AF1 alters splicing similarly in mouse and human hematopoietic cells
- Mutant U2AF1 alters splicing of RNA processing and recurrently mutated MDS/AML genes

Accession Numbers

GSE66793

Shirai et al., 2015, Cancer Cell *27*, 631–643
May 11, 2015
http://dx.doi.org/10.1016/j.ccell.2015.04.008

CellPress

Cancer Cell
Article

Mutant U2AF1 Expression Alters Hematopoiesis and Pre-mRNA Splicing In Vivo

Cara Lunn Shirai,[1] James N. Ley,[1] Brian S. White,[1,2] Sanghyun Kim,[1] Justin Tibbitts,[1] Jin Shao,[1] Matthew Ndonwi,[1] Brian Wadugu,[1] Eric J. Duncavage,[3] Theresa Okeyo-Owuor,[1] Tuoen Liu,[1] Malachi Griffith,[2] Sean McGrath,[2] Vincent Magrini,[2] Robert S. Fulton,[2] Catrina Fronick,[2] Michelle O'Laughlin,[2] Timothy A. Graubert,[4] and Matthew J. Walter[1,*]

[1]Division of Oncology, Department of Medicine, Washington University, St. Louis, MO 63110, USA
[2]The Genome Institute, Washington University, St. Louis, MO 63110, USA
[3]Department of Pathology and Immunology, Washington University, St. Louis, MO 63110, USA
[4]Massachusetts General Hospital/Harvard Medical School, Boston, MA 02114, USA
*Correspondence: mjwalter@dom.wustl.edu
http://dx.doi.org/10.1016/j.ccell.2015.04.008

SUMMARY

Heterozygous somatic mutations in the spliceosome gene *U2AF1* occur in ~11% of patients with myelodysplastic syndromes (MDS), the most common adult myeloid malignancy. It is unclear how these mutations contribute to disease. We examined in vivo hematopoietic consequences of the most common *U2AF1* mutation using a doxycycline-inducible transgenic mouse model. Mice expressing mutant U2AF1(S34F) display altered hematopoiesis and changes in pre-mRNA splicing in hematopoietic progenitor cells by whole transcriptome analysis (RNA-seq). Integration with human RNA-seq datasets determined that common mutant U2AF1-induced splicing alterations are enriched in RNA processing genes, ribosomal genes, and recurrently mutated MDS and acute myeloid leukemia-associated genes. These findings support the hypothesis that mutant U2AF1 alters downstream gene isoform expression, thereby contributing to abnormal hematopoiesis in patients with MDS.

INTRODUCTION

Myelodysplastic syndromes (MDS) are the most common myeloid malignancy of the elderly, with approximately 10,000 new cases occurring in the United States annually (Ma, 2012). MDS are a heterogeneous group of clonal hematopoietic stem cell disorders characterized by peripheral blood cytopenias, with up to 30% of patients experiencing progression to secondary acute myeloid leukemia (AML) (Troy et al., 2014). Mutations in spliceosome genes have been identified in over half of MDS patient bone marrow samples, making it the most common class of genes mutated in MDS (Damm et al., 2012; Graubert et al., 2012; Papaemmanuil et al., 2011; Thol et al., 2012; Visconte et al., 2012; Walter et al., 2013; Yoshida et al., 2011). The recurrently-mutated spliceosome genes encode factors that are involved in the recognition of the 3′-intronic splice site and are mutually exclusive of one another in patient samples (Haferlach et al., 2014; Papaemmanuil et al., 2013; Walter et al., 2013; Yoshida et al., 2011), implying that they may contribute similarly to MDS pathogenesis or, alternatively, may not be tolerated by a cell when they co-occur.

Our group and others identified mutations in *U2AF1* (U2 small nuclear RNA auxiliary factor 1) in 11% of patients with MDS, making it one of the most commonly mutated genes in this disease (Graubert et al., 2012; Yoshida et al., 2011). In addition, *U2AF1* mutations typically occur in the founding clone, suggesting they may play an important role in disease initiation (Haferlach et al., 2014; Papaemmanuil et al., 2013; Walter et al.,

Significance

Mutations in spliceosome genes occur in up to ~50% of patients with myelodysplastic syndromes (MDS), suggesting that perturbations in pre-mRNA splicing contribute to disease pathogenesis. We generated a murine model of the most common mutation in the spliceosome gene *U2AF1* and observed hematopoietic phenotypes and pre-mRNA splicing alterations that also occur in patients with MDS or acute myeloid leukemia. Concordant changes in isoform expression of RNA processing genes, ribosomal genes, and recurrently mutated genes in myeloid cancers in the mouse and human highlight cellular processes and pathways that may functionally contribute to mutant U2AF1-associated diseases. Determining whether splicing changes in the same genes are induced by other MDS-associated spliceosome gene mutations may further prioritize key target genes in MDS.

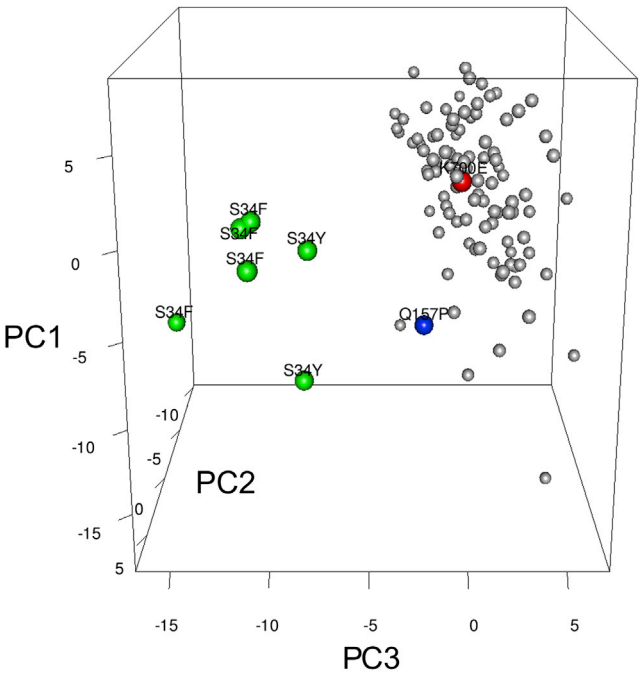


Figure 1. Global Alterations in Pre-mRNA Splicing Are Distinct in AML Patients with *U2AF1* Mutations

Unsupervised principal component analysis of standardized splicing ratios of expressed 3′ splice sites of cassette and mutually exclusive exons (greater than five reads in half or more of samples) in patients with AML (n = 6 mutant U2AF1[S34F/Y], green dots; n = 102 normal for spliceosome genes, gray dots; n = 1 U2AF1[Q157P], blue dot; n = 1 SF3B1[K700E], red dot). Putative spliceosome genes were used to denote "normal"; see also Table S1.

2013). *U2AF1* mutations have also been identified across major cancer types, including AML, lung adenocarcinoma, head and neck squamous cell carcinoma, uterine corpus endometrial carcinoma, bladder urothelial carcinoma, breast adenocarcinoma, and colorectal carcinoma, further implicating *U2AF1* broadly in cancer pathogenesis (Kandoth et al., 2013). During normal pre-mRNA splicing activity, the U2AF1 protein binds to the AG dinucleotide at the 3′ end of the intron, thereby assisting its cofactor U2AF2 in recruitment of the U2 snRNP for spliceosome activation as part of the spliceosome E/A complex (Wahl et al., 2009; Wu et al., 1999).

It is unknown how *U2AF1* mutations contribute to MDS pathogenesis. Several groups have reported splicing changes in cells expressing mutant U2AF1, including cell lines (HeLa, K562), primary human $CD34^+$ cells, and patient AML and MDS cells (Brooks et al., 2014; Graubert et al., 2012; Ilagan et al., 2014; Okeyo-Owuor et al., 2014; Przychodzen et al., 2013; Yoshida et al., 2011). However, the downstream targets of mutant U2AF1 identified in these studies are variable, perhaps due to differences in cell types, co-occurring mutations, and experimental methods used. Initial in vivo studies using a retroviral overexpression model showed that mouse bone marrow cells expressing mutant U2AF1 have reduced repopulation ability (Yoshida et al., 2011). Together, these data lead us to hypothesize that mutant U2AF1-induced splicing alterations and subsequent changes in gene isoform expression result in abnormal hematopoiesis.

RESULTS

Splicing Is Altered in Primary Human AML Cells Expressing Mutant U2AF1

To examine the effects of *U2AF1* mutations on global splicing in primary patient samples, we used RNA-seq data from the The Cancer Genome Atlas (TCGA) AML cohort (Cancer Genome Atlas Research Network, 2013). We identified eight samples with a spliceosome gene mutation, including six samples with a *U2AF1* mutation affecting the S34 amino acid (four S34F, two S34Y), one sample with U2AF1(Q157P), one sample with SF3B1(K700E), and 102 samples without a mutation or copy number alteration in spliceosome genes (Table S1). Unsupervised clustering using the splicing ratio of cassette and mutually exclusive exon splice junctions segregated the six mutant U2AF1(S34) patient samples from the 102 control, the U2AF1(Q157P), and SF3B1(K700E) samples (Figure 1), indicating that splicing is distinctly altered in primary patient cells with U2AF1 mutations affecting the S34 amino acid.

Generation of U2AF1(S34F) Transgenic Mice

To study the in vivo consequences of *U2AF1* mutations on splicing and hematopoiesis in an isolated genetic system and to prioritize splicing alterations identified in primary patient samples, we created a mouse model to study the most abundant *U2AF1* mutation found in MDS patients [U2AF1(S34F)]. Using a previously-validated, single-copy, site-specific integration approach (Beard et al., 2006), we generated doxycycline-inducible U2AF1(S34F) and control U2AF1(WT) transgenic mice. We integrated human cDNA (human and mouse U2AF1 proteins differ in only one amino acid in the polyglycine tract of the C-terminal RS domain) coding for U2AF1(S34F) or U2AF1(WT) into the *Col1a1* locus of KH2 mouse embryonic stem (ES) cells (Figure S1A). KH2 ES cells contain the M2rtTA tetracycline-responsive transactivator protein (rtTA) ubiquitously expressed from the *Rosa26* locus, which allows for induction of the integrated U2AF1(S34F) or U2AF1(WT) transgene following treatment of cells with the tetracycline derivative doxycycline (Figure 2A). Integration was validated in both mouse lines by Southern blot analysis (Figure 2B). These transgenic mouse lines display dose-dependent, inducible, in vivo expression of the integrated *U2AF1* transgene (S34F or WT) in bone marrow cells (Figure 2C). For subsequent experiments, we used a doxycycline dose (625 ppm doxycycline chow) that induced levels of exogenous transgene expression similar to endogenous mouse *U2af1* levels in bone marrow cells, consistent with the heterozygous expression of *U2AF1* mutations in MDS patient bone marrow samples (Graubert et al., 2012). This dose of doxycycline induced similar levels of transgene expression in U2AF1(S34F)/rtTA and U2AF1(WT)/rtTA double transgenic mice, as determined by pyrosequencing (data not shown). U2AF1(WT) and U2AF1(S34F) expression was in excess of normal mouse *U2af1* expression levels, resulting in approximately double the normal protein level in cells (Figure S1B); however, overexpression of human U2AF1 transgene(s) at this dose of doxycycline has no impact on endogenous mouse *U2af1* expression (Figure S1C). *U2AF1* transgene expression was induced in myeloid and lymphoid cell lineages, as well as in stem and progenitor cells (Figure S1D).

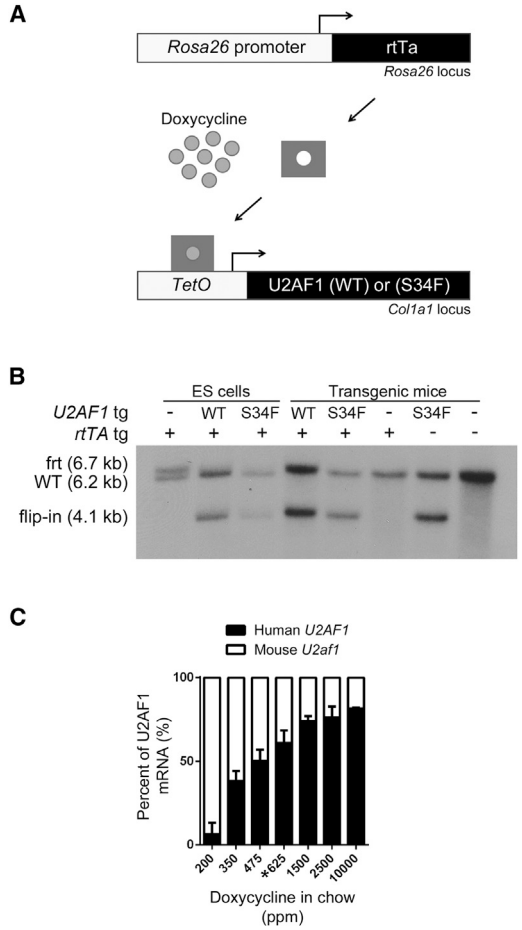


Figure 2. Generation of Doxycycline-Inducible U2AF1 Transgenic Mice

(A) Schematic of doxycycline-inducible *U2AF1* transgene system. Human U2AF1(S34F) or U2AF1(WT) cDNA was integrated into a modified *Col1a1* locus of mouse ES cells containing the reverse tetracycline transactivator (M2rtTA) expressed from the *Rosa26* locus. Addition of doxycycline induces *U2AF1* transgene expression.

(B) Southern blot of ES cell or transgenic mouse tail genomic DNA using the 5′ *Col1a1* probe to detect the integration of U2AF1(S34F) or U2AF1(WT) transgenes. Flip-in of the pBS31' vector is detected by the appearance of a 4.1 kb band. Frt is the modified frt locus (6.7 kb), and WT is the wild-type locus (6.2 kb).

(C) Doxycycline-dose response curve of *U2AF1* transgene expression relative to endogenous mouse *U2af1* in bone marrow cells measured by RT-PCR followed by pyrosequencing assay (n = 2–7). Data are represented as mean ±SD. The 625 ppm doxycycline chow (*) was used in all subsequent experiments; ppm, parts per million.

See also Figure S1.

Hematopoietic Cell Lineage Is Altered in U2AF1(S34F)-Recipient Mice

To examine the cell-autonomous effects of mutant U2AF1(S34F) expression on hematopoiesis, we transplanted transgenic mouse bone marrow into lethally-irradiated wild-type congenic recipient mice, allowing for ≥6 weeks of engraftment prior to induction of the transgene. Initial experiments were conducted with all six possible genotypes produced from transgenic colonies. Control genotypes of wild-type (nontransgenic) mice, single transgenic rtTA-only, U2AF1(WT)-only, and U2AF1(S34F)-only, and double transgenic U2AF1(WT)/rtTA mice were similar in parameters measured (Figure S2A). Given the similarity of all controls, subsequent experiments were performed with the U2AF1(S34F)/rtTA double transgenic mice and two control genotypes: U2AF1(WT)/rtTA double transgenic mice and single transgenic rtTA-only mice (littermates from the U2AF1[S34F] transgenic mouse colony).

Examination of hematopoietic compartments in recipient mice following one month of doxycycline to induce transgene expression revealed a reduction in peripheral blood total white blood cell (WBC) counts in U2AF1(S34F)/rtTA-recipient mice compared to U2AF1(WT)/rtTA- and rtTA only-recipient controls (4.3 versus 7.11 and 7.13 K/μl, respectively, p = 0.014) (Figure 3A), but no difference in red blood cell parameters or platelet counts (Figure S2B). This peripheral blood leukopenia was stable and persisted for up to 12 months (Figure 3A); it was also dependent on mutant U2AF1 expression, since the WBC counts in mutant U2AF1(S34F)/rtTA-recipient mice recovered to levels similar to U2AF1(WT)/rtTA-recipient mice after halting doxycycline treatment, even after 6 months of transgene expression (Figure S2C). Flow cytometry analysis of mature lineage cells in the peripheral blood revealed a reduction in B cells and monocytes in U2AF1(S34F)/rtTA-recipient mice compared to controls (Figure 3B).

There were no differences in bone marrow cellularity or spleen weights in U2AF1(S34F)/rtTA-recipient mice compared to controls (Figure S2D). However, mature cell lineage distribution was altered in the bone marrow of U2AF1(S34F)/rtTA-recipient mice after one month of doxycycline. Both monocytes and B cells were reduced in U2AF1(S34F)/rtTA-recipient mouse bone marrow, and neutrophils were increased (Figure 3C). The reduction of monocytes in the bone marrow of U2AF1(S34F)/rtTA-recipient mice occurred as soon as 5 days after transgene induction (Figure S2E) and was associated with an increase in Annexin V^+ and phospho-H2AX staining of monocytes (Figure 3D). In vitro culture of transgenic mouse bone marrow revealed a similar trend in Annexin V^+ and phospho-H2AX flow cytometry in U2AF1(S34F)/rtTA-expressing cells compared to controls (Figure S2F). There was no evidence of bone marrow dysplasia in U2AF1(S34F)/rtTA-recipient mice (Figure S2G). The overall survival of U2AF1(S34F)/rtTA- and U2AF1(WT)/rtTA-recipient mice was similar (Figure S2H), and mutant mice did not develop MDS or AML after at least 1 year of continuous doxycycline.

U2AF1(S34F)-Recipient Mice Have Increased Progenitor Cells in Bone Marrow and Spleen, but Stem Cells Have a Competitive Disadvantage in Repopulation Assays

Following 1 month of doxycycline to induce transgene expression, U2AF1(S34F)/rtTA-recipient mice displayed an increased frequency of myeloid progenitors in both bone marrow (Figure 4A) and spleen (Figure 4B) by progenitor colony forming assay (CFU-C) compared to U2AF1(WT)/rtTA- and rtTA-only-recipient control mice. This increase was also seen by flow cytometry analysis for c-Kit^+/$lineage^-$ (KL) cells in the bone marrow (Figure S3A), as well as for common myeloid progenitors (CMP) in bone marrow (Figure 4C) and spleen (Figure 4D) of U2AF1(S34F)/rtTA-recipient mice compared to controls. Granulocyte/macrophage progenitors (GMP) were variably increased (Figure S3B)

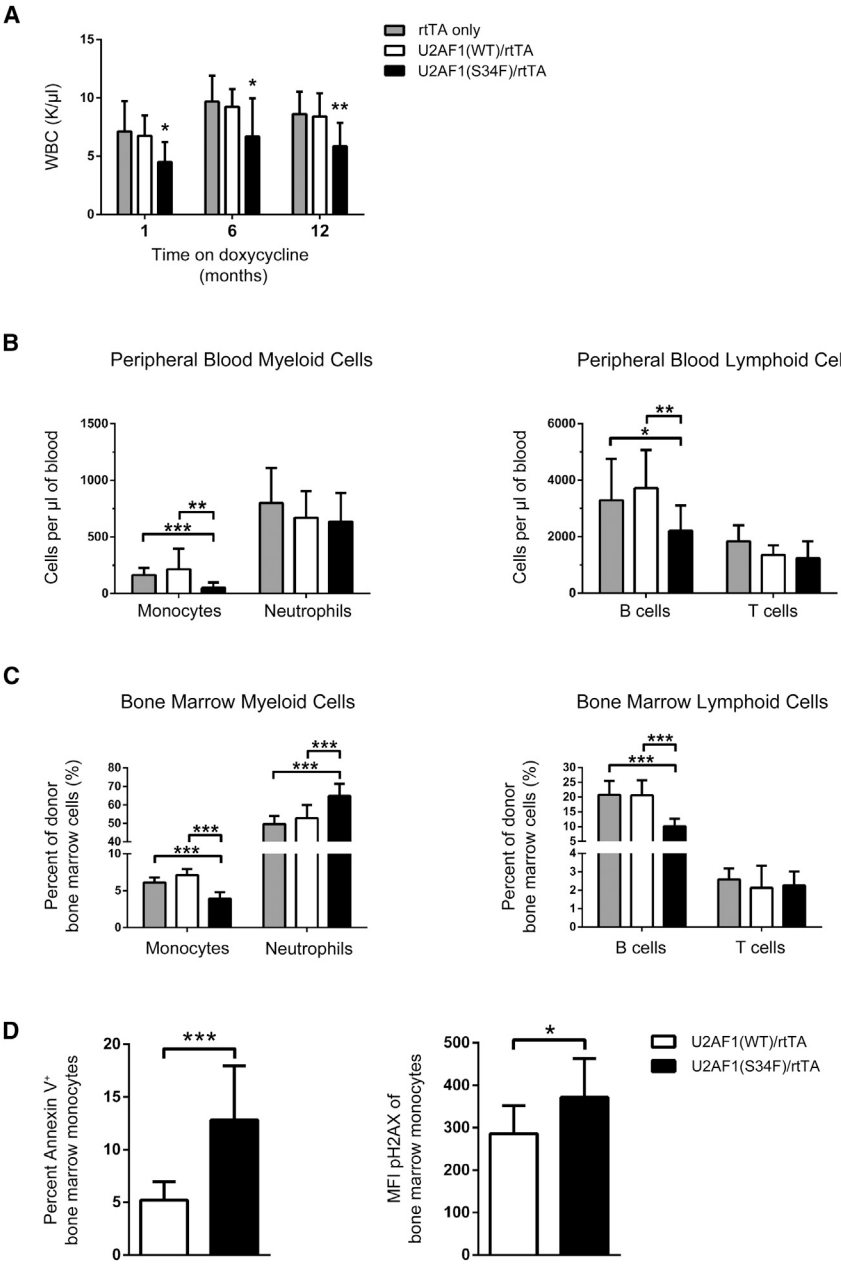


Figure 3. Mature Hematopoietic Cell Lineages Are Altered by In Vivo Expression of U2AF1(S34F)

(A) Peripheral blood white blood cell counts of mice transplanted with transgenic donor marrow following 1, 6, and 12 months of doxycycline induction of transgene expression (n = 5–12).

(B) Absolute count of peripheral blood monocytes (Gr-1$^+$, CD115$^+$), neutrophils (Gr-1$^+$, CD115$^-$), B cells (B220$^+$), and T cells (CD3e$^+$) based on WBC and flow cytometry following 1 month doxycycline (n = 9–11).

(C) Mature lineage distribution of bone marrow cells by flow cytometry for monocytes, neutrophils, B cells, and T cells following 1 month doxycycline (n = 9–11).

(D) Detection of apoptosis in bone marrow monocytes (Gr-1$^+$, CD115$^+$) following 5 days of doxycycline by flow cytometry for Annexin V$^+$ (left, n = 9–14) and intracellular phospho-H2AX (right, n = 8–13). All data are represented as mean ±SD. *p $\leq$ 0.05, **p $\leq$ 0.01, ***p $\leq$ 0.001; MFI, mean fluorescence intensity.

See also Figure S2.

and megakaryocyte/erythroid progenitors (MEP) were unchanged in mutant mice compared to controls (Figure S3C).

Flow cytometry for the stem cell-enriched compartment defined by c-Kit$^+$/lineage$^-$/Sca-1$^+$ (KLS) cells also showed an increase in the bone marrow of U2AF1(S34F)/rtTA-recipient mice compared to controls (Figure 4E) following 1 month of doxycycline. Further characterization of stem cells by flow cytometry using SLAM and CD34/Flk2 revealed a trend toward increased multipotent progenitors, but no difference in short-term and long-term stem cell populations (Figure S3D). The increase in bone marrow stem-enriched and progenitor cells (KLS, KL, CMP) measured by flow cytometry and by CFU-C in U2AF1(S34F)/rtTA-recipient mice was detectable following only 5 days of doxycycline exposure (Figures S3E–S3H). The increased frequency of KLS cells after 5 days of transgene induction was associated with increased Ki67$^+$ staining in KLS cells from U2AF1(S34F)/rtTA-recipient mice compared to controls (Figure 4F); a similar trend was seen for BrdU incorporation (Figure S3I). There was no increase in hematopoietic stem-enriched and progenitor cells measured by flow cytometry and methylcellulose in the absence of doxycycline in U2AF1(S34F)/rtTA-recipient mice (Figures S3E, S3G, and S3H), confirming this early phenotype at 5 days was not due to doxycycline-independent transgene expression. Next, we evaluated stem cell function using a competitive repopulation assay. Although we observed an increase in the frequency of stem-enriched and progenitor cells in mutant U2AF1-recipient mice, competitive repopulation assays revealed a disadvantage for U2AF1(S34F)/rtTA-expressing stem cells compared to U2AF1(WT)/rtTA-expressing cells in primary recipient mice (Figure 4G). The competitive disadvantage was sustained when the bone marrow from primary recipients was transplanted into secondary and tertiary recipient mice (Figure 4H).

Comprehensive RNA Sequencing of U2AF1(S34F)-Expressing Mouse Bone Marrow Progenitors Reveals Splicing Alterations

MDS patient bone marrow samples have an increased fraction of common myeloid progenitors (CMPs) compared to normal bone marrow donors (Pang et al., 2013; Woll et al., 2014), and we observed an increase in CMPs in mutant U2AF1 mouse bone marrow as early as 5 days following doxycycline exposure with

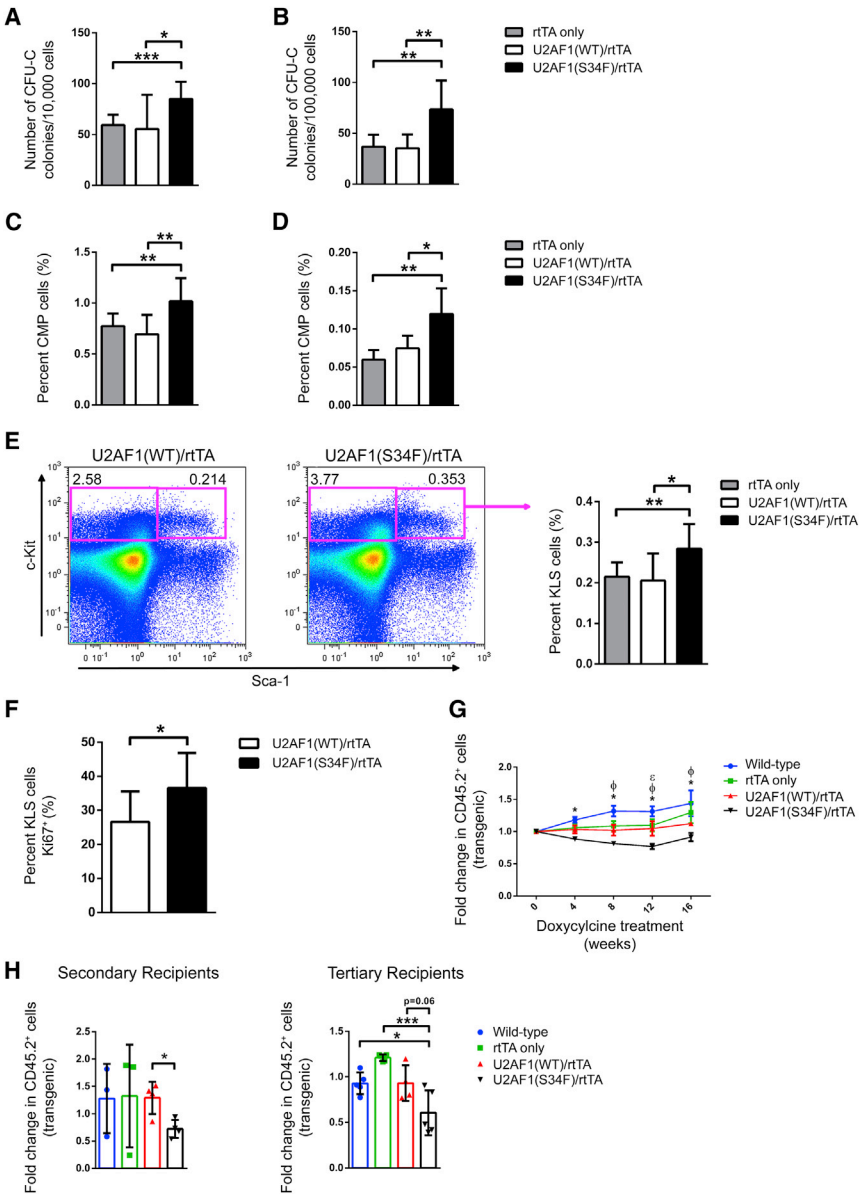


Figure 4. Hematopoietic Stem-Enriched and Progenitor Cells Are Increased by In Vivo Expression of U2AF1(S34F) but Have a Competitive Disadvantage in Repopulation Assays

(A and B) Myeloid progenitor colony forming CFU-C assays of bone marrow (A, n = 9–11) and spleen cells (B, n = 5–10) from mice transplanted with transgenic donor marrow following 1 month of doxycycline.

(C and D) Flow cytometry for donor-derived common myeloid progenitors (CMP: CD45.1$^-$, CD45.2$^+$, lin$^-$, c-Kit$^+$, Sca-1$^-$, CD34$^+$, Fcγ^-) in bone marrow (C, n = 9–11) and spleen (D, n = 5–6) of mice transplanted with transgenic donor marrow following 1 month of doxycycline.

(E) Flow cytometry for donor-derived, stem-cell enriched fractions of bone marrow (KLS: CD45.1$^-$, CD45.2$^+$, lin$^-$, c-Kit$^+$, Sca-1$^+$ cells [n = 9–11]) following 1 month of doxycycline. Representative flow cytometry plots of U2AF1(WT)/rtTA and U2AF1(S34F)/rtTA-recipient mice showing c-Kit and Sca-1 antigen expression on CD45.2$^+$, lin$^-$ gated cells (left); quantification of multiple mice (right).

(F) Intracellular flow cytometry of Ki67$^+$ levels in bone marrow KLS cells following 5 days of doxycycline (n = 9).

(G) Peripheral blood chimerism of primary recipient mice transplanted with equal portions of transgenic (CD45.2$^+$) and wild-type (CD45.1$^+$/CD45.2$^+$) bone marrow cells in competitive repopulation assays. Data are shown as fold change of transgenic (CD45.2$^+$ chimerism) from the start of doxycycline (n = 5–12). Differences (two-way ANOVA) are indicated at each time point as different from the U2AF1(S34F) genotype by *wild-type nontransgenic mice, $^{\Phi}$rtTA only, and $^{\varepsilon}$U2AF1(WT)/rtTA.

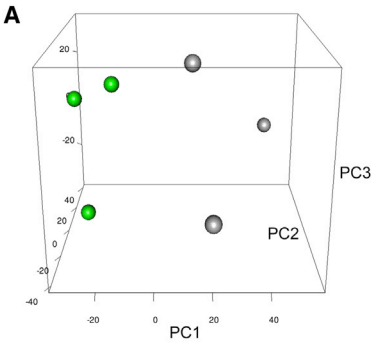
(H) Bone marrow cells from primary recipient mice of the same genotype were pooled equally and transplanted into secondary recipient mice (left, n = 3–5); likewise, bone marrow from secondary recipients was transplanted into tertiary recipient mice (right, n = 4–5). Fold change of transgenic cells (CD45.2$^+$ chimerism) measured in peripheral blood $\geq$4 months post-secondary and tertiary transplants.

All data are represented as mean ± SD, except those in (C), which are represented as mean ± SEM. *p $\leq$ 0.05, **p $\leq$ 0.01, ***p $\leq$ 0.001. See also Figure S3.

a concomitant increase in CFU-C formation. Therefore, we sorted donor-derived CMPs from U2AF1(S34F)/rtTA- and U2AF1(WT)/rtTA-recipient mice following 5 days of transgene induction and performed transcriptome sequencing (RNA-seq) to detect U2AF1(S34F)-induced splicing alterations. The levels of transgene induction were similar among samples (Figure S4A). On average, CMP samples yielded 3.2 × 10^8 sequencing reads (Figure S4B), with an average of 2.0 × 10^8 uniquely mapped reads across samples (Figure S4C). The distribution of reads from various species of RNA was similar, with mutant samples having a slightly higher percentage of intergenic bases ($p < 0.05$) and lower percentage of coding bases ($p = 0.05$) (Figure S4D).

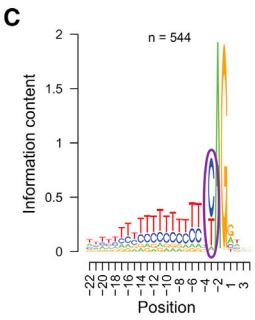
Unsupervised clustering based on the splicing ratio of cassette and mutually exclusive exon junctions segregated U2AF1(S34F)-expressing CMP samples from U2AF1(WT) controls (Figure 5A), as it did for patients with AML (Figure 1). DESeq (Anders and Huber, 2010), an algorithm that accounts for RNA-seq replicate data and biologic variation, identified changes in overall gene expression in 128 genes (FDR < 0.1) in U2AF1(S34F)/rtTA samples compared to U2AF1(WT)/rtTA controls (Table S2). Using these genes, we performed gene set enrichment and pathway analyses with GOseq (Young et al., 2010) and identified an enrichment of genes involved in immune response and leukocyte activation processes (FDR < 0.1) (Table S3).

DEXSeq (Anders et al., 2012), an algorithm that identifies altered junctions while accounting for changes in overall gene expression, identified 742 splicing junctions that were



B

Molecular Function (GO:MF)			
Category	Term	FDR	Enrichment
GO:0003723	RNA binding	0.003	1.77
GO:0044822	poly(A) RNA binding	0.005	1.90



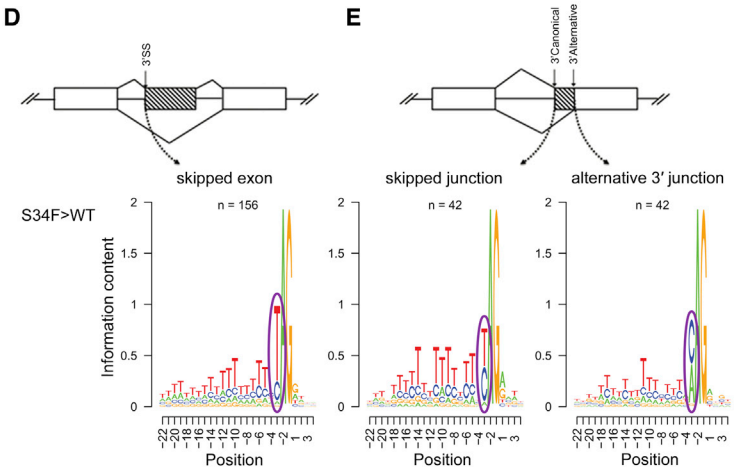


Figure 5. U2AF1(S34F) Expression Alters Splicing in Mouse Progenitors Cells In Vivo

(A) Unsupervised principal component analysis of standardized splicing ratios of expressed 3′ splice sites of cassette and mutually exclusive exons (greater than five reads in half or more of samples) in donor-derived bone marrow common myeloid progenitor cells (CMP) sorted from mice transplanted with U2AF1(S34F)/rtTA or U2AF1(WT)/rtTA bone marrow (n = 3 pools of five to seven mice each) following 5 days doxycycline.

(B) Pathways enriched (GOseq FDR < 0.1) in differentially spliced genes in U2AF1(S34F)/rtTA CMPs.

(C) The consensus sequence surrounding the AG dinucleotide of the 3′ splice site of 544 randomly selected junctions not altered by U2AF1(S34F).

(D and E) The 3′ splice site consensus sequence of U2AF1(S34F)-skipped exons (D, n = 156) and of skipped canonical splice junctions and their associated alternative 3′ junctions (E, n = 42) identified by DEXSeq (FDR < 0.1).

Results are depicted in Logos plots. ss, splice site. See also Figure S4 and Tables S2, S3, and S4.

differentially expressed in 633 genes in U2AF1(S34F)/rtTA samples compared to U2AF1(WT)/rtTA controls (FDR < 0.1) (Table S4). Using these 633 genes, we again performed gene set enrichment and pathway analyses and identified an enrichment of genes involved in RNA binding (FDR < 0.1) (Figure 5B). The majority of differentially spliced junctions in U2AF1(S34F)/rtTA samples compared to U2AF1(WT)/rtTA controls were classified as cassette exon, simple alternative 3′ splice acceptor site, and coordinate cassette exon events (Figure S4E).

U2AF1(S34F) Has Altered 3′ Splice Acceptor Sequence Preference

To further examine the effects of U2AF1(S34F) expression on splicing in vivo, we determined the consensus sequence flanking the AG dinucleotide at the 3′ splice acceptor site known to be recognized by U2AF1 during splicing (Wu et al., 1999). In contrast to the sequence motif seen in non-dysregulated control junctions (Figure 5C), exons skipped more frequently by mutant U2AF1(S34F) relative to U2AF1(WT) were enriched for uracil (indicated by a T) in the −3 position relative to the AG dinucleotide (Figure 5D). We observed a similar pattern in alternative 3′ splice site usage; there is an enrichment of T in the −3 position of canonical 3′ splice sites whose alternative 3′ splice site was used more frequently in mutant U2AF1(S34F)/rtTA samples relative to U2AF1(WT)/rtTA controls (Figure 5E). This pattern is the same as seen in AML patient samples with *U2AF1* mutations and in primary human CD34$^+$ cells expressing U2AF1(S34F) or U2AF1(WT), and similar to previously reported data in human samples (Brooks et al., 2014; Ilagan et al., 2014; Okeyo-Owuor et al., 2014; Przychodzen et al., 2013). These data indicate that the sequence-specific pattern of altered splicing induced by mutant U2AF1 is similar in mouse and human cells.

U2AF1(S34F)-Induced Splicing Changes Are Enriched in Genes Involved in RNA Processing and Splicing, Protein Translation, and Recurrently Mutated Genes in MDS/AML

To prioritize altered splicing events for further analysis, we intersected significant junctions (DEXSeq; FDR < 0.1) across three datasets: mouse CMP samples (n = 219 junctions), AML patient samples with and without *U2AF1* mutations (n = 162 junctions), and our previously described primary human CD34$^+$ cells

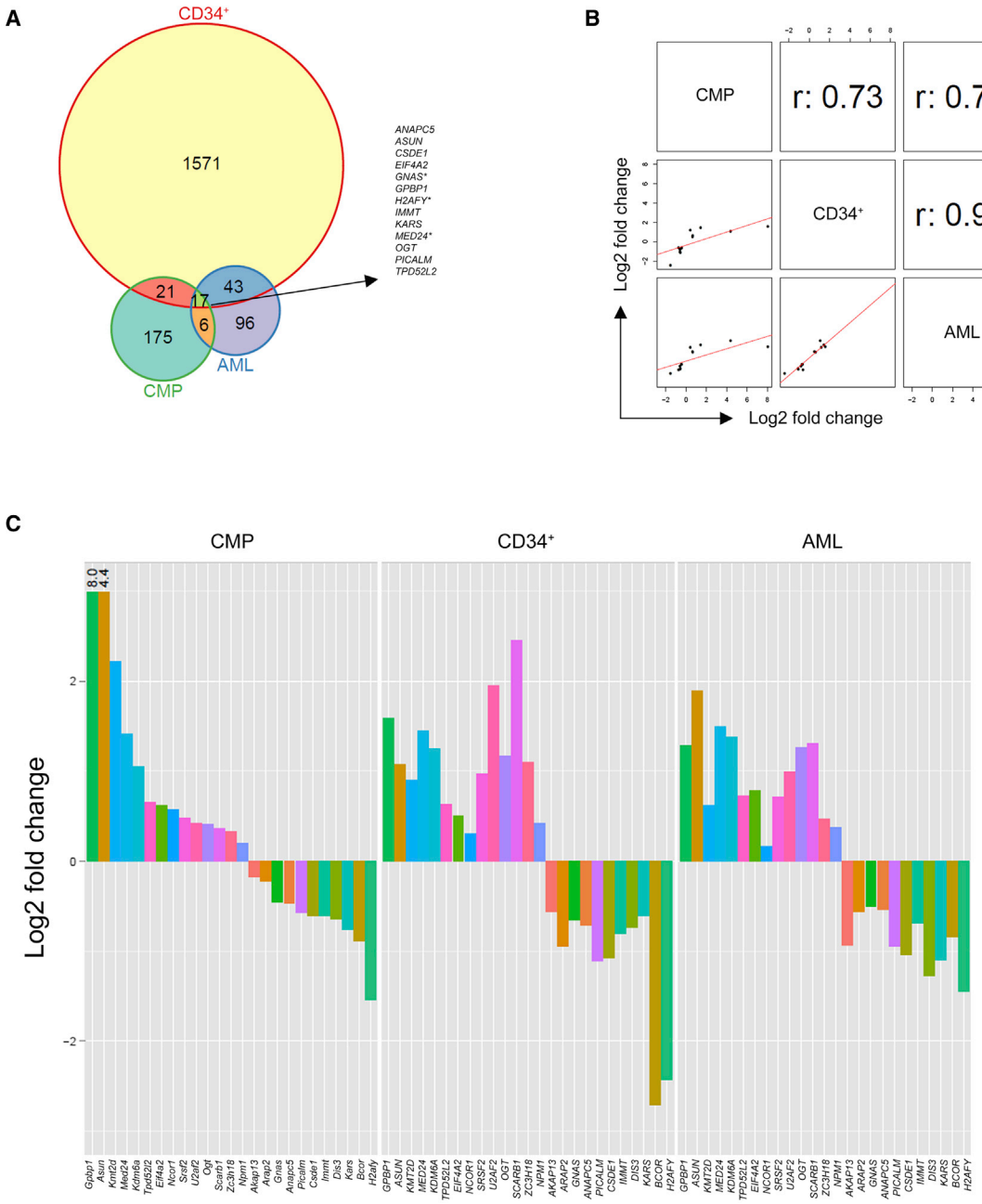


Figure 6. Mutant U2AF1 Alters Splicing in Common Targets across Species

(A) Venn diagram of the overlap of independently discovered mutant U2AF1-induced splice junction changes by DEXSeq analysis (FDR < 0.1) in three RNA-seq datasets: transgenic mouse CMPs, primary human hematopoietic CD34$^+$ cells, and human AML patient samples. *Indicates multiple significant junctions detected within the gene.

(B) Correlation plots and r values for junctions identified by overlap of DEXSeq datasets shown in the Venn diagram. One junction per gene is shown.

(C) Log2 fold change in junction expression [U2AF1(WT) versus U2AF1(S34F)] for overlapping genes (A) and recurrently mutated genes in MDS and AML identified by Fisher's meta-analysis (Table 1), plotted in descending order based on CMP results. *U2AF1* was excluded due to an inability to differentiate endogenous versus exogenous U2AF1(WT) transcript in CD34$^+$ cells.

See also Figure S5 and Tables S5 and S6.

overexpressing U2AF1(S34F) or U2AF1(WT) (n = 1,652 junctions) (Okeyo-Owuor et al., 2014). The intersection was (necessarily) composed of homologous junctions and was additionally restricted to include only concordantly dysregulated (same direction of log fold change) junctions; 17 homologous dysregulated junctions ($p < 0.0001$; simulation) were present in 13 genes (Figures 6A and 6B). Several of the dysregulated junctions corresponded to the same splicing event in a transcript and were

Table 1. Spliceosome and Recurrently Mutated MDS/AML Genes with Dysregulated Splicing

Spliceosome Genes	Recurrently Mutated Genes in MDS/AML
CLK1	*AKAP13*
CLK3	*ARAP2*
CPSF4	*BCOR*
HNRNPA1	*DIS3*
HNRNPA2B1	*GNAS*
HNRNPC	*KDM6A*
HNRNPD	*KMT2D*
HNRNPH1	*NCOR1*
HNRNPM	*NPM1*
PCBP2	*SCARB1*
PRPF40A	*SRSF2*
PTBP1	*U2AF1*
PTBP2	*U2AF2*
RBM39	*ZC3H18*
RNMT	
SF3B3	
SNRNP70	
SNRPG	
SRSF10	
SRSF11	
SRSF2	
SRSF7	
TCERG1	
THOC1	
THOC2	
U2AF1	
U2AF1L4	
U2AF2	
YBX1	
ZMAT2	

Fisher FDR < 0.1. See also Table S7.

identified as reciprocal splice junctions. All of these junctions occur in known isoforms, though the function of most of these isoforms remains largely unknown. Several of these junctions occur in genes mutated in MDS and AML (*GNAS, PICALM*) (Bejar et al., 2011; Borel et al., 2012) or known to be involved in stem cell biology (*H2AFY, MED24*) (Buschbeck et al., 2009; Chia et al., 2010; Gaspar-Maia et al., 2013).

To broaden the list of potentially concordant changes across species, we performed a meta-analysis of the three independent datasets using Fisher's combined probability test (Brown, 1975) and identified 555 significantly dysregulated homologous junctions in 415 genes across all three datasets (FDR < 0.1) (Table S5). The larger number of dysregulated junctions relative to the conservative intersection above enables both consensus sequence analysis and pathway/gene-set enrichment analyses. Skipped exons from the Fisher dataset also show an enrichment of T at the −3 position of 3′ splice sites, as in the CMP data (Figure S5). We hypothesized that splicing in biologic pathway(s) could be widely disrupted by mutation of an upstream splicing factor. To test this, we performed gene set enrichment and pathway analyses on the 415 splicing-dysregulated genes using GOseq. This resulted in 28 pathways/categories identified as significant (FDR < 0.1, Enrichment > 2) by one of the three independent analyses of KEGG pathways or GO molecular function (MF) or biological processes (BP) categories; 14 of these contain genes involved in RNA processing, RNA splicing, RNA localization/transport, and RNA binding, whereas 11 were involved in protein translation processes and ribosomal pathways (Table S6). Additionally, spliceosome genes ($p = 4.68 \times 10^{-8}$, Table S1) and genes recurrently mutated in MDS and AML (Bejar et al., 2011; Cancer Genome Atlas Research Network, 2013; Haferlach et al., 2014; Papaemmanuil et al., 2013; Walter et al., 2013) ($p = 0.03$, Table S7) were enriched in the 415 splicing-dysregulated genes (Table 1 and Figure 6C). Oncogenes and tumor suppressors (Supek et al., 2014) were also highly represented ($p = 0.09$, data not shown).

U2AF1(S34F)-Induced Splicing Changes Are Present in Human MDS Bone Marrow Samples

Next, we examined MDS patient bone marrow samples for mutant U2AF1-induced splicing changes discovered by RNA-seq analysis in a subset of affected genes. Dysregulated splicing events were prioritized for validation by overlap in the three datasets, splicing ratio change, or known biological function of the gene, with particular focus on genes mutated in MDS. Mutant U2AF1-induced splice isoform changes identified by RNA-seq analysis were concordant with RT-PCR of MDS patient bone marrow samples for seven of eight splicing events examined, including *H2AFY*, *BCOR*, *PICALM*, *GNAS*, *KDM6A*, *KMT2D* (*MLL2*), and *MED24*; only *EIF4A2* showed no difference (Figure 7 and data not shown).

DISCUSSION

In this study, we provide evidence that mutant U2AF1 expression alters hematopoiesis and pre-mRNA splicing in the primary hematopoietic progenitor cells of mice. U2AF1(S34F) expression in mice results in leukopenia and changes in the distribution of mature hematopoietic lineages in the peripheral blood and bone marrow. In addition, U2AF1(S34F) expression increases the frequency of progenitor cells in the bone marrow and spleen of mice as well as the frequency of progenitor cells that are cycling in the bone marrow. We identify U2AF1(S34F)-specific changes in splice isoforms in 633 genes in mouse common myeloid progenitors. Finally, through an integrative meta-analysis of three unique RNA-seq datasets, we identify splice junctions that are consistently altered across species by mutant U2AF1 expression and that are enriched in RNA processing genes, translational processes/ribosomal genes, and recurrently mutated genes in MDS/AML. Collectively, these results suggest that U2AF1(S34F)-induced alternative splicing may contribute to the altered hematopoiesis that is characteristic of patients with MDS.

Mutant U2AF1(S34F) mice display several phenotypes associated with MDS, including leukopenia, increased apoptosis of maturing cells in the bone marrow, and progenitor cell expansion (Catenacci and Schiller, 2005; Hamblin, 1992; Pang et al., 2013;

Raza et al., 1995a, 1995b; Shetty et al., 2000; Troy et al., 2014; Woll et al., 2014; Yoshida, 1993). However, these mice do not show dysplasia, nor do they develop MDS or AML. Similar to another genetic mouse model of MDS mutations (*ASXL1*) (Abdel-Wahab et al., 2013), mutant U2AF1(S34F)-expressing stem cells appear to have a competitive disadvantage and increased apoptosis compared to wild-type cells, highlighting a conundrum for MDS biology: how is clonal dominance achieved by a cell containing mutation(s) that result in reduced competitiveness and increased apoptosis? Although *U2AF1* mutations are typically early genetic events, they may occur in cells already harboring a mutation that is permissive for clonal expansion. Cooperativity studies with significantly co-occurring mutations, such as del(20) or *ASXL1* mutations (Bacher et al., 2014; Damm et al., 2012; Papaemmanuil et al., 2013; Walter et al., 2013) (or with other common mutations in MDS, like those in *TET2* or *DNMT3A*), may be necessary to determine whether *U2AF1* mutations require a specific genetic context to cause MDS or AML. Ultimately, *U2AF1* mutations may contribute to ineffective hematopoiesis that is typical in MDS, whereas other mutations confer clonal dominance. This mouse model could be used to address these questions.

In this study, we show U2AF1(S34F)-induced splicing alterations in primary hematopoietic cells in an isolated genetic system in vivo. By using primary mouse CMPs and human $CD34^+$ cells, we capitalize on identifying early splicing changes that are induced by mutant U2AF1 expression in the absence of transformation. Importantly, expression of mutant U2AF1 in mouse bone marrow cells results in an altered 3′ splice acceptor site sequence preference that is identical to the pattern observed in human cells, suggesting that the same mechanism of altered splicing is conserved across species and credentialing the mouse model for mechanistic studies.

We identified a list of consistently altered candidate genes that may contribute to myeloid disease pathogenesis. Comparison of these datasets with dysregulated junctions found in lung adenocarcinoma patient samples with *U2AF1* mutations suggests that at least some genes with altered splicing are shared across cancer and cell types harboring *U2AF1* mutations; indeed, 13 of 29 genes with significant alterations associated with *U2AF1* mutations in both lung adenocarcinoma and AML are also present in our list of junctions identified by meta-analysis (Brooks et al., 2014).

Further studies will be needed to identify the consequences of mutant U2AF1-induced splicing alterations for hematopoiesis and myeloid disease. It is still unknown if the contribution of *U2AF1* mutations to MDS pathology involves pre-mRNA splicing alterations of one, a select few, or many genes of interest. Future rescue experiments using target mRNAs/shRNAs to alter isoform expression in mice will be necessary to address this question. However, the observation that dysregulated splicing junctions induced by mutant U2AF1(S34F) are enriched in recurrently mutated genes in MDS/AML raises the possibility that subtle changes in isoform expression in one or more genes may contribute to mutant U2AF1-associated disease. Changes in isoform expression of cancer-associated genes may have important consequences for a cell. For example, different isoforms of *NRAS* have been shown to activate different signaling molecules, and expression of a rare isoform results in increased cell proliferation and transformation when compared to other *NRAS* isoforms (Eisfeld et al., 2014). We identified consistent splicing alterations in several genes commonly affected by loss of function mutations in MDS and cancer, including *BCOR* and *KDM6A*; however, splicing changes in both genes resulted in frame-preserving isoforms. Although the functional consequences of these changes are unknown, the dysregulated isoform of *BCOR* has been shown to lack a domain necessary to interact with the transcription factor AF9 (a common fusion partner of MLL) and repress AF9 transcriptional activity (Srinivasan et al., 2003).

We report here that perturbations in spliceosome and RNA processing genes occur in three independent mutant U2AF1 RNA-seq datasets. It raises the possibility that mutation of a spliceosome gene may result in autoregulatory changes in splicing machinery by altering isoform expression (Jangi and Sharp, 2014). This may account for the mutual exclusivity of spliceosome gene mutations in patients with MDS (Haferlach et al., 2014; Papaemmanuil et al., 2013; Walter et al., 2013; Yoshida et al., 2011). Perturbations in the splicing process may have consequences for genome stability, as missplicing has been implicated in co-transcriptional R-loop formation (RNA:DNA hybrid) (Chan et al., 2014; Skourti-Stathaki and Proudfoot, 2014). This may contribute to the increased phospho-H2AX staining we observe in monocytes by inducing DNA damage through R-loop formation. Further studies are needed to address this. In addition, the identification of genes involved in ribosome function and translational processes are an interesting consequence of mutant U2AF1 expression, because perturbations in the ribosome have been linked to MDS (Barlow et al., 2010; Ebert et al., 2008; Pellagatti et al., 2008). If other spliceosome gene mutations (e.g., *SRSF2*, *SF3B1*) perturb either of these two pathways, it would suggest a unifying mechanism of action for all these mutations. Ultimately, examination of dysregulated splicing caused by other spliceosome gene mutations and comparison with these findings may help clarify which gene pathways are important.

By performing all of our experiments in a cell-autonomous manner using a transplantation model, transgenic donor bone marrow cells necessarily undergo the stress of repopulating the hematopoietic system in the irradiated microenvironment of recipient mice. However, phenotypes observed in the transplanted mice were also seen in non-transplanted transgenic mice, including bone marrow progenitor expansion and monocyte apoptosis (data not shown). While the transgenic model we describe here could be used to test the dependency of tumors on mutant U2AF1 expression (by removing doxycycline once tumors develop), the overexpression of U2AF1(S34F) or U2AF1(WT) transgene, albeit low, is expected to result in non-stoichiometric levels of U2AF1 relative to other splicing factors. However, no detectable phenotypes were observed in U2AF1(WT)-expressing transgenic mice compared to rtTA only controls. Additional models of *U2AF1* mutations (e.g., knockin of S34F at the endogenous *U2af1* locus and mice with other *U2AF1* mutations) will be important to further characterize the effect of *U2AF1* mutations in vivo.

Although this study improves our understanding of mutant U2AF1 biology, our ultimate goal is to improve treatment of patients with *U2AF1* mutations. The pervasive nature of

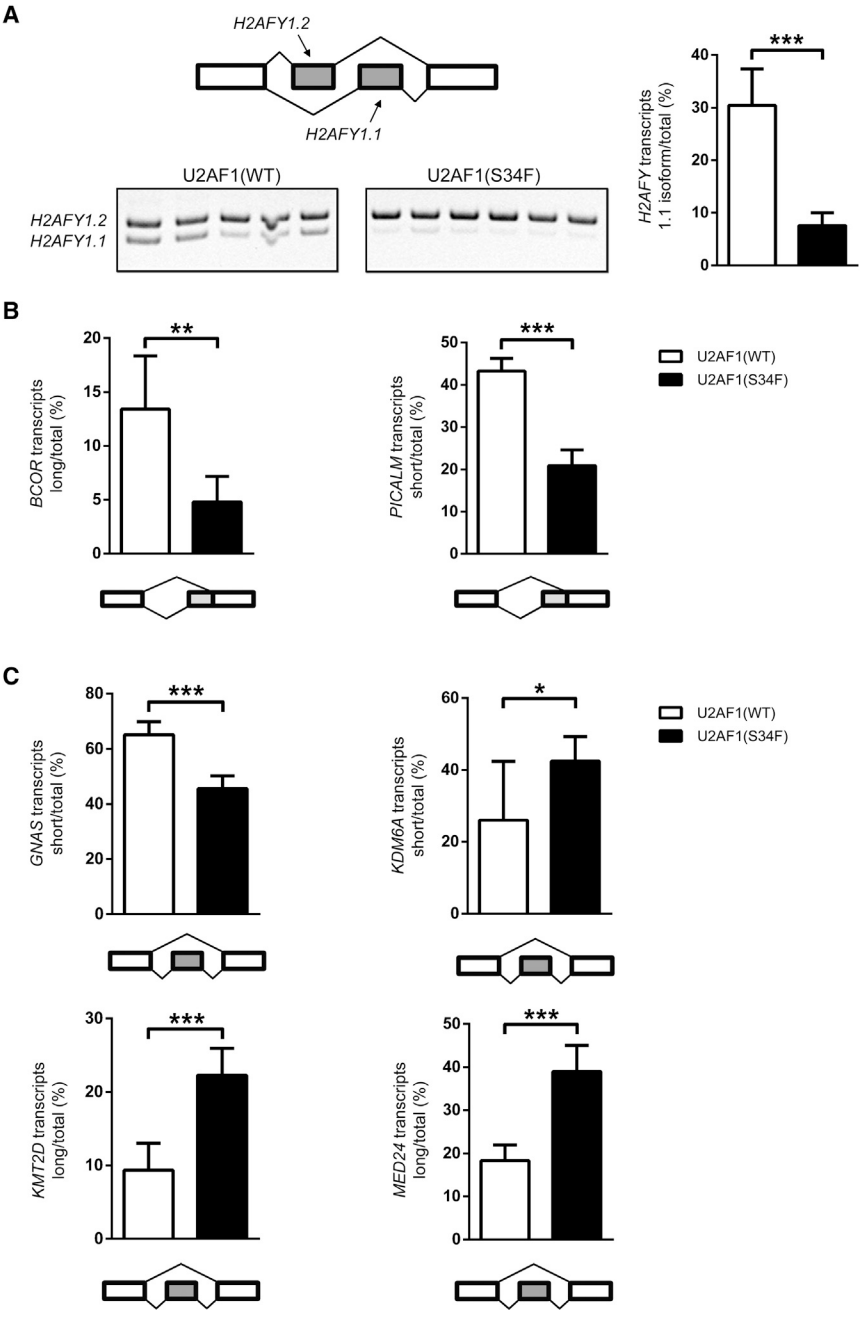


Figure 7. Mutant U2AF1 Alters Splicing in MDS Bone Marrow Cells

(A) Mutually exclusive exons in *H2AFY* detected by RT-PCR and gel electrophoresis (quantified in right); the diagram of the event measured and the gel image (left).

(B and C) Quantification of alternative splice site utilization (B) and altered cassette exon events (C). All data are represented as mean ±SD. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001; n = 5–6.

spliceosome gene mutations in MDS highlights the need to understand and develop effective therapeutic approaches, perhaps by using splicing modulator drugs. This *U2AF1* mouse model may be a useful pre-clinical tool to test potential therapies in a variety of cancers with *U2AF1* mutations.

EXPERIMENTAL PROCEDURES

Generation of Doxycycline-Inducible U2AF1(S34F) and U2AF1(WT) Transgenic Mice

cDNAs that code for human U2AF1(S34F) and U2AF1(WT) were cloned from an MDS patient bone marrow sample (UPN 571656) by RT-PCR using Superscript III kit (Invitrogen) and introduced individually into the pBS31′ vector (Beard et al., 2006) using EcoRI restriction sites. U2AF1(S34F) and (WT) cDNA sequences differed only in the one nucleotide that produced the S34F amino acid conversion. U2AF1(S34F)-pBS31′ and U2AF1(WT)-pBS31′ plasmids were introduced into KH2 ES cells by the FLP/FRT system (Beard et al., 2006) to produce U2AF1(S34F) and U2AF1(WT)-expressing ES cell lines. ES cell clones were selected by hygromycin, and integration of U2AF1 (S34F) or (WT)-pBS31′ was confirmed using PCR and Southern blot approaches. Integration at the 5′ end was detected using standard Southern blotting techniques as described (Beard et al., 2006). PCR screening of ES cells and genotyping of mice were performed with the primers described (see Supplemental Experimental Procedures). KH2 ES cells are (C57BL/6 × 129S4Sv/Jae)F1-derived V6.5 embryonic stem (ES) cells (Beard et al., 2006); therefore, transgenic mouse colonies were maintained by breeding heterozygous, doubly transgenic mice (U2AF1(S34F)/rtTA and U2AF1(WT)/rtTA) to (C57BL/6 × 129S4Sv/Jae)F1 mice. Doxycycline was administered via doxycycline-containing rodent chow (Pico 5053 base, TestDiet) or via doxycycline in water (in UV light-protected red water bottles, changed every 2 days).

All mouse procedures were performed according to protocols approved by the Washington University Animal Studies Committee.

RT-PCR and Pyrosequencing

Following doxycycline induction of transgene expression in vivo, bone marrow cells were harvested from the leg bones. RNA was isolated from bone marrow cell pellets using the RNeasy kit (QIAGEN) followed by DNA removal with Turbo DNA-free kit (Ambion). Reverse transcription PCR was then performed using the Superscript III RT-PCR kit (Invitrogen). PCR for *U2AF1* and *U2af1* to detect exogenous human transgene and endogenous mouse transcripts expressed in transgenic cells, respectively, was performed, and pyrosequencing for single nucleotide polymorphisms that differ between mouse and human cDNAs was performed on PCR products (primer sequences are listed in the Supplemental Experimental Procedures).

Murine Bone Marrow Transplant and Competitive Repopulation Assays

To generate mice for each experiment, 1×10^6 transgenic mouse donor bone marrow cells from two to three mice pooled (CD45.2) were transplanted into at least five lethally irradiated (1,100 rads) wild-type recipient mice (C57BL/6 × 129S4Sv/Jae)F1 (CD45.1/CD45.2) per genotype. Donor mice were between 8 and 12 weeks of age, and recipient mice ranged from 6 to 12 weeks of age; donor and recipient mice were sex-matched. Competitive repopulation assays were performed by mixing transgenic donor "test" mouse bone marrow cells (CD45.2) from two to three mice in equal proportions with

(C57BL/6 × 129S4Sv/Jae)F1 (CD45.1/CD45.2) donor "competitor" mouse bone marrow cells from two to three mice, and then transplanted into (C57BL/6 × 129S4Sv/Jae)F1 (CD45.1/CD45.2) recipient mice. Transgene induction was started ≥6 weeks after transplantation.

Mouse Hematopoietic Progenitor Assay

Methylcellulose progenitor colony forming assays (CFU-C) were performed using Methocult GF M3434 (Stem Cell Technologies). Bone marrow and spleen cells were obtained from transplanted mice following doxycycline induction of transgene expression. Red blood cells were lysed prior to plating of 10,000 bone marrow cells or 100,000 spleen cells per 1.3 ml media; each sample was evaluated in duplicate. Progenitor colonies (defined as ≥40 cells/colony) were counted following 7 days culture at 37°C with 5% CO_2.

Flow Cytometry

Flow cytometry of donor-derived mature cell lineage distribution was performed on bone marrow, peripheral blood, and spleen using the cell surface receptors: CD45.1, CD45.2, CD115, Gr-1, B220, and CD3e (antibodies and procedures described in the Supplemental Experimental Procedures). Apoptosis was evaluated by flow cytometry for Annexin V (APC, BD Biosciences) and intracellular Phospho-H2AX (FITC, JBW301, EMD Millipore) using the BD Cytofix/Cytoperm kit (BD Biosciences) following the manufacturer's recommendations. Donor-derived hematopoietic progenitor and stem cell flow cytometry was performed on bone marrow and spleen using the following cell surface markers (antibodies described in the Supplemental Experimental Procedures): CD45.1, CD45.2, mature lineage (Gr-1, Cd3e, B220, Ter119, and CD41), c-Kit, Sca-1, CD34, Fcγ, CD150, CD48, and Flk2. All flow cytometry was performed using FACScan or Gallios cytometers (BD Biosciences) and analyzed using FlowJo software (FlowJo). Progenitor cells were isolated for RNA-seq using an iCyt Synergy flow sorter (Sony).

RNA Sequencing of Common Myeloid Progenitors

Donor-derived common myeloid progenitors (CMPs; lin^-, c-Kit^+, Sca-1^-, $CD34^+$, $Fc\gamma^-$) cells were sorted from pooled bone marrow of five to seven U2AF1(S34F)/rtTA- or U2AF1(WT)/rtTA-transplanted mice. Cells were sorted into DMEM media, and RNA was prepared from cell pellets using a miRNeasy kit (QIAGEN). Genomic DNA was removed by Turbo DNA-free kit (Ambion). Ribosomal RNA was depleted (Ribozero, Epicenter), followed by cDNA preparation and TruSeq stranded Illumina library production; 2 × 100 bp paired-end sequencing was performed on the HiSeq2000 platform (Illumina).

RNA-Seq Analysis

Reads were aligned to the mouse mm9 reference genome using TopHat (version 2.0.8) (Kim et al., 2013). All subsequent analyses were performed in R (Gentleman et al., 2004). Differential gene expression was determined using DESeq2 (version 1.6.1) (Anders and Huber, 2010), which applies a negative binomial model to counts of reads mapped to genes. Gene models were provided by Ensembl version 67 (Flicek et al., 2013). Differential junction expression was determined using DEXSeq (version 1.12.1) (Anders et al., 2012), which applies a negative binomial model to (novel or Ensembl-annotated) junction-spanning read counts while controlling for potential changes in overall gene expression between conditions. Gene enrichment analysis was performed using GOseq (version 1.18.0) (Young et al., 2010), which accounts for gene-length induced biases. Regions spanning junctions were downloaded from the UCSC Genome Browser (Kent et al., 2002) using BSgenome (version 1.34.0), and their sequence logos were created using seqLogo (1.32.1). Principal component analysis was performed using prcomp, with normalization performed by ggbiplot (version 0.55). Additional bioinformatic analysis of RNA-seq data, including splicing ratio calculation, is described in the Supplemental Experimental Procedures.

RT-PCR and Gel Electrophoresis for Validation of Discovered Altered Junctions

RNA extraction from MDS patient samples, cDNA preparation, and RT-PCR followed by gel electrophoresis were carried out as described elsewhere (Okeyo-Owuor et al., 2014). Briefly, total RNA was treated with DNase using the Turbo DNA-free kit (Ambion), and cDNA was generated using the Superscript III reverse transcriptase kit (Invitrogen). PCR products were electrophoresed in TBE-polyacrylamide gels and stained with ethidium bromide for imaging. Densitometry was performed using ImageJ. PCR primers are described in the Supplemental Experimental Procedures.

All patients provided written consent on a protocol approved by the Washington University Human Studies Committee.

Statistical Analyses

Statistical analysis and generation of graphical presentations of data were performed using analysis functions in GraphPad Prism 6 and as described in the Supplemental Experimental Procedures for RNA-seq analysis.

ACCESSION NUMBERS

The NCBI Gene Expression Omnibus accession number for the RNA-seq data reported in this paper is GSE66793.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, five figures, and seven tables and can be found with this article online at http://dx.doi.org/10.1016/j.ccell.2015.04.008.

AUTHOR CONTRIBUTIONS

The study was designed by C.L.S., T.A.G., and M.J.W.; mouse characterization by C.L.S., J.N.L., S.K., J.T., M.N., E.J.D., and T.L.; RNA sequencing by S.M., V.M., R.S.F., C.F., and M.O.; bioinformatics analysis was performed by B.S.W. and M.G.; clinical sample validation was performed by S.K., J.S., B.W., and T.O.O.; the manuscript was written and edited by C.L.S., J.N.L., B.S.W., S.K., T.A.G., M.J.W. All authors reviewed and approved the submission.

ACKNOWLEDGMENTS

Support was provided by NIH/NHLBI (T32HL007088 to C.L.S.), Barnes-Jewish Hospital Foundation (to B.S.W., T.A.G., and M.J.W.); an NIH/NCI SPORE in Leukemia (P50CA171963 to B.S.W., T.A.G., and M.J.W.); an NIH/NCI grant (K12CA167540 to B.S.W); the Edward P. Evans Foundation (to T.A.G. and M.J.W.); a Clinical and Translational Award from the NIH National Center for Advancing Translational Sciences (UL1 TR000448 to B.S.W.); and a Howard Hughes Medical Institute Physician-Scientist Early Career Award, Leukemia and Lymphoma Society Scholar Award, Department of Defense (BM120018), and Aplastic Anemia and MDS International Foundation (to M.J.W.). Technical assistance was provided by the Alvin J. Siteman Cancer Center High Speed Cell Sorting Core and the Tissue Procurement Core supported by an NCI Cancer Center Support Grant (P30CA91842) and the Genome Institute. The authors thank Tianjiao Wang for technical assistance and Drs. Tim Ley and Dan Link for helpful scientific discussions.

Received: December 24, 2014
Revised: March 16, 2015
Accepted: April 13, 2015
Published: May 11, 2015

REFERENCES

Abdel-Wahab, O., Gao, J., Adli, M., Dey, A., Trimarchi, T., Chung, Y.R., Kuscu, C., Hricik, T., Ndiaye-Lobry, D., Lafave, L.M., et al. (2013). Deletion of Asxl1 results in myelodysplasia and severe developmental defects in vivo. J. Exp. Med. *210*, 2641–2659.

Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biol. *11*, R106.

Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. Genome Res. *22*, 2008–2017.

Bacher, U., Haferlach, T., Schnittger, S., Zenger, M., Meggendorfer, M., Jeromin, S., Roller, A., Grossmann, V., Krauth, M.T., Alpermann, T., et al. (2014). Investigation of 305 patients with myelodysplastic syndromes and

20q deletion for associated cytogenetic and molecular genetic lesions and their prognostic impact. Br. J. Haematol. *164*, 822–833.

Barlow, J.L., Drynan, L.F., Hewett, D.R., Holmes, L.R., Lorenzo-Abalde, S., Lane, A.L., Jolin, H.E., Pannell, R., Middleton, A.J., Wong, S.H., et al. (2010). A p53-dependent mechanism underlies macrocytic anemia in a mouse model of human 5q- syndrome. Nat. Med. *16*, 59–66.

Beard, C., Hochedlinger, K., Plath, K., Wutz, A., and Jaenisch, R. (2006). Efficient method to generate single-copy transgenic mice by site-specific integration in embryonic stem cells. Genesis *44*, 23–28.

Bejar, R., Stevenson, K., Abdel-Wahab, O., Galili, N., Nilsson, B., Garcia-Manero, G., Kantarjian, H., Raza, A., Levine, R.L., Neuberg, D., and Ebert, B.L. (2011). Clinical effect of point mutations in myelodysplastic syndromes. N. Engl. J. Med. *364*, 2496–2506.

Borel, C., Dastugue, N., Cances-Lauwers, V., Mozziconacci, M.J., Prebet, T., Vey, N., Pigneux, A., Lippert, E., Visanica, S., Legrand, F., et al. (2012). PICALM-MLLT10 acute myeloid leukemia: a French cohort of 18 patients. Leuk. Res. *36*, 1365–1369.

Brooks, A.N., Choi, P.S., de Waal, L., Sharifnia, T., Imielinski, M., Saksena, G., Pedamallu, C.S., Sivachenko, A., Rosenberg, M., Chmielecki, J., et al. (2014). A pan-cancer analysis of transcriptome changes associated with somatic mutations in U2AF1 reveals commonly altered splicing events. PLoS ONE *9*, e87361.

Brown, M.B. (1975). Method for Combining Non-Independent, One-Sided Tests of Significance. Biometrics *31*, 987–992.

Buschbeck, M., Uribesalgo, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H., and Di Croce, L. (2009). The histone variant macroH2A is an epigenetic regulator of key developmental genes. Nat. Struct. Mol. Biol. *16*, 1074–1079.

Cancer Genome Atlas Research Network (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Catenacci, D.V., and Schiller, G.J. (2005). Myelodysplasic syndromes: a comprehensive review. Blood Rev. *19*, 301–319.

Chan, Y.A., Hieter, P., and Stirling, P.C. (2014). Mechanisms of genome instability induced by RNA-processing defects. Trends Genet. *30*, 245–253.

Chia, N.Y., Chan, Y.S., Feng, B., Lu, X., Orlov, Y.L., Moreau, D., Kumar, P., Yang, L., Jiang, J., Lau, M.S., et al. (2010). A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature *468*, 316–320.

Damm, F., Kosmider, O., Gelsi-Boyer, V., Renneville, A., Carbuccia, N., Hidalgo-Curtis, C., Della Valle, V., Couronné, L., Scourzic, L., Chesnais, V., et al.; Groupe Francophone des Myélodysplasies (2012). Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. Blood *119*, 3211–3218.

Ebert, B.L., Pretz, J., Bosco, J., Chang, C.Y., Tamayo, P., Galili, N., Raza, A., Root, D.E., Attar, E., Ellis, S.R., and Golub, T.R. (2008). Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. Nature *451*, 335–339.

Eisfeld, A.K., Schwind, S., Hoag, K.W., Walker, C.J., Liyanarachchi, S., Patel, R., Huang, X., Markowitz, J., Duan, W., Otterson, G.A., et al. (2014). NRAS isoforms differentially affect downstream pathways, cell growth, and cell transformation. Proc. Natl. Acad. Sci. USA *111*, 4179–4184.

Flicek, P., Ahmed, I., Amode, M.R., Barrell, D., Beal, K., Brent, S., Carvalho-Silva, D., Clapham, P., Coates, G., Fairley, S., et al. (2013). Ensembl 2013. Nucleic Acids Res. *41*, D48–D55.

Gaspar-Maia, A., Qadeer, Z.A., Hasson, D., Ratnakumar, K., Leu, N.A., Leroy, G., Liu, S., Costanzi, C., Valle-Garcia, D., Schaniel, C., et al. (2013). MacroH2A histone variants act as a barrier upon reprogramming towards pluripotency. Nat. Commun. *4*, 1565.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. *5*, R80.

Graubert, T.A., Shen, D., Ding, L., Okeyo-Owuor, T., Lunn, C.L., Shao, J., Krysiak, K., Harris, C.C., Koboldt, D.C., Larson, D.E., et al. (2012). Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. Nat. Genet. *44*, 53–57.

Haferlach, T., Nagata, Y., Grossmann, V., Okuno, Y., Bacher, U., Nagae, G., Schnittger, S., Sanada, M., Kon, A., Alpermann, T., et al. (2014). Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. Leukemia *28*, 241–247.

Hamblin, T. (1992). Clinical features of MDS. Leuk. Res. *16*, 89–93.

Ilagan, J.O., Ramakrishnan, A., Hayes, B., Murphy, M.E., Zebari, A.S., Bradley, P., and Bradley, R.K. (2014). U2AF1 mutations alter splice site recognition in hematological malignancies. Genome Res. *25*, 14–26.

Jangi, M., and Sharp, P.A. (2014). Building robust transcriptomes with master splicing factors. Cell *159*, 487–498.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M.A., et al. (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333–339.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. *12*, 996–1006.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. *14*, R36.

Ma, X. (2012). Epidemiology of myelodysplastic syndromes. Am. J. Med. *125* (7, Suppl), S2–S5.

Okeyo-Owuor, T., White, B.S., Chatrikhi, R., Mohan, D.R., Kim, S., Griffith, M., Ding, L., Ketkar-Kulkarni, S., Hundal, J., Laird, K.M., et al. (2014). U2AF1 mutations alter sequence specificity of pre-mRNA binding and splicing. Leukemia *29*, 909–917.

Pang, W.W., Pluvinage, J.V., Price, E.A., Sridhar, K., Arber, D.A., Greenberg, P.L., Schrier, S.L., Park, C.Y., and Weissman, I.L. (2013). Hematopoietic stem cell and progenitor cell mechanisms in myelodysplastic syndromes. Proc. Natl. Acad. Sci. USA *110*, 3011–3016.

Papaemmanuil, E., Cazzola, M., Boultwood, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J.S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., et al.; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium (2011). Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N. Engl. J. Med. *365*, 1384–1395.

Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., Yoon, C.J., Ellis, P., Wedge, D.C., Pellagatti, A., et al.; Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood *122*, 3616–3627, quiz 3699.

Pellagatti, A., Hellström-Lindberg, E., Giagounidis, A., Perry, J., Malcovati, L., Della Porta, M.G., Jädersten, M., Killick, S., Fidler, C., Cazzola, M., et al. (2008). Haploinsufficiency of RPS14 in 5q- syndrome is associated with deregulation of ribosomal- and translation-related genes. Br. J. Haematol. *142*, 57–64.

Przychodzen, B., Jerez, A., Guinta, K., Sekeres, M.A., Padgett, R., Maciejewski, J.P., and Makishima, H. (2013). Patterns of missplicing due to somatic U2AF1 mutations in myeloid neoplasms. Blood *122*, 999–1006.

Raza, A., Gezer, S., Mundle, S., Gao, X.Z., Alvi, S., Borok, R., Rifkin, S., Iftikhar, A., Shetty, V., Parcharidou, A., et al. (1995a). Apoptosis in bone marrow biopsy samples involving stromal and hematopoietic cells in 50 patients with myelodysplastic syndromes. Blood *86*, 268–276.

Raza, A., Mundle, S., Iftikhar, A., Gregory, S., Marcus, B., Khan, Z., Alvi, S., Shetty, V., Dameron, S., Wright, V., et al. (1995b). Simultaneous assessment of cell kinetics and programmed cell death in bone marrow biopsies of myelodysplastics reveals extensive apoptosis as the probable basis for ineffective hematopoiesis. Am. J. Hematol. *48*, 143–154.

Shetty, V., Hussaini, S., Broady-Robinson, L., Allampallam, K., Mundle, S., Borok, R., Broderick, E., Mazzoran, L., Zorat, F., and Raza, A. (2000). Intramedullary apoptosis of hematopoietic cells in myelodysplastic syndrome patients can be massive: apoptotic cells recovered from high-density fraction of bone marrow aspirates. Blood *96*, 1388–1392.

Skourti-Stathaki, K., and Proudfoot, N.J. (2014). A double-edged sword: R loops as threats to genome integrity and powerful regulators of gene expression. Genes Dev. *28*, 1384–1396.

Srinivasan, R.S., de Erkenez, A.C., and Hemenway, C.S. (2003). The mixed lineage leukemia fusion partner AF9 binds specific isoforms of the BCL-6 corepressor. Oncogene *22*, 3395–3406.

Supek, F., Miñana, B., Valcárcel, J., Gabaldón, T., and Lehner, B. (2014). Synonymous mutations frequently act as driver mutations in human cancers. Cell *156*, 1324–1335.

Thol, F., Kade, S., Schlarmann, C., Löffeld, P., Morgan, M., Krauter, J., Wlodarski, M.W., Kölking, B., Wichmann, M., Görlich, K., et al. (2012). Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. Blood *119*, 3578–3584.

Troy, J.D., Atallah, E., Geyer, J.T., and Saber, W. (2014). Myelodysplastic syndromes in the United States: an update for clinicians. Ann. Med. *46*, 283–289.

Visconte, V., Makishima, H., Jankowska, A., Szpurka, H., Traina, F., Jerez, A., O'Keefe, C., Rogers, H.J., Sekeres, M.A., Maciejewski, J.P., and Tiu, R.V. (2012). SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. Leukemia *26*, 542–545.

Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The spliceosome: design principles of a dynamic RNP machine. Cell *136*, 701–718.

Walter, M.J., Shen, D., Shao, J., Ding, L., White, B.S., Kandoth, C., Miller, C.A., Niu, B., McLellan, M.D., Dees, N.D., et al. (2013). Clonal diversity of recurrently mutated genes in myelodysplastic syndromes. Leukemia *27*, 1275–1282.

Woll, P.S., Kjällquist, U., Chowdhury, O., Doolittle, H., Wedge, D.C., Thongjuea, S., Erlandsson, R., Ngara, M., Anderson, K., Deng, Q., et al. (2014). Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. Cancer Cell *25*, 794–808.

Wu, S., Romfo, C.M., Nilsen, T.W., and Green, M.R. (1999). Functional recognition of the 3′ splice site AG by the splicing factor U2AF35. Nature *402*, 832–835.

Yoshida, Y. (1993). Hypothesis: apoptosis may be the mechanism responsible for the premature intramedullary cell death in the myelodysplastic syndrome. Leukemia *7*, 144–146.

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. Nature *478*, 64–69.

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. Genome Biol. *11*, R14.