AD_____

AWARD NUMBER: W81XWH-16-1-0059

TITLE: Therapeutic targeting of spliceosomal-mutant acquired bone marrow failure disorders

PRINCIPAL INVESTIGATOR: Omar Abdel-Wahab MD

RECIPIENT: Sloan Kettering Institute for Cancer Research

REPORT DATE:
August 2019

TYPE OF REPORT: Final Technical

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

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
August 2019	Final	1 May 2016 – 30 April 2019

4. TITLE AND SUBTITLE

Therapeutic targeting of spliceosomal-mutant acquired bone marrow failure disorders

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-16-1-0059

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Omar Abdel-Wahab M.D.

E-Mail: abdelwao@mskcc.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Sloan Kettering Institute for Cancer Research
1275 York Avenue
New York, NY 10065-6007

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

Genes encoding core components of the RNA splicing machinery are the most common mutational targets in acquired bone marrow failure (BMF) due to myelodysplastic syndromes (MDS). The goals of this proposal were to identify (1) the subset of biologically and therapeutically relevant targets that link spliceosomal mutations to MDS and (2) therapeutic strategies that interfere with the altered function of mutant spliceosomal proteins. Since award of this grant we have found that different spliceosomal mutations, despite imparting distinct effects on splicing and gene expression, are negatively selected for when co-expressed in the same cell or in a homozygous state. At the same time, aberrant splicing of distinct targets/events by mutant SF3B1 and SRSF2 results in hyperactivated NF-kB signaling, thereby identifying convergent biological consequences of splicing factor mutations and the basis for their mutual exclusivity and heterozygous nature. In addition, we have completed both a negative selection shRNA screen and a genome-wide CRISPR dropout screen to identify genes selectively required in spliceosomal mutant cells. This effort has also highlighted a requirement for innate immune signaling in SF3B1-mutant MDS and has implicated a few specific proteins as potential novel therapeutic targets for spliceosomal mutant MDS.

15. SUBJECT TERMS
MDS, SF3B1, Splicing, SRSF2, U2AF1, ZRSR2.

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	18	**19b. TELEPHONE NUMBER** *(include area code)*

TABLE OF CONTENTS

	Page No.
1. Introduction	4
2. Keywords	4
3. Accomplishments	4
4. Impact	8
5. Changes/Problems	9
6. Products	10
7. Participants & Other Collaborating Organizations	13
8. Special Reporting Requirements	16
9. Appendices	18

1. **INTRODUCTION:** Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.

Genes encoding core components of the RNA splicing machinery are the most common mutational targets in acquired bone marrow failure (BMF) due to myelodysplastic syndromes (MDS). MDS associated "spliceosomal mutations" most frequently affect SF3B1, SRSF2, and U2AF1. Mutations in each of these genes conspicuously occur as heterozygous point mutations at highly restricted residues and are mutually exclusive, suggesting that they are oncogenic gain-of-function alterations. While these data suggest that each RNA splicing factor mutation may share downstream targets, currently few mis-spliced genes have been conclusively identified as downstream targets of any mutant splicing protein, and no shared targets downstream of all three spliceosomal gene mutations have been identified. Thus, the goals of this proposal were to identify (1) the subset of biologically and therapeutically relevant targets that link spliceosomal mutations to MDS and (2) therapeutic strategies that interfere with the altered function of mutant spliceosomal proteins.

2. **KEYWORDS:** Provide a brief list of keywords (limit to 20 words).

MDS, SF3B1, Splicing, SRSF2, U2AF1, ZRSR2

3. **ACCOMPLISHMENTS:** The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency Grants Officer whenever there are significant changes in the project or its direction.

What were the major goals of the project?
List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

The major goals of this project were as follows:
Specific Aim 1. Identify unifying molecular abnormalities across spliceosomal mutations using combined genomic and proteomic approaches.
Specific Aim 2. Identify additional genes that are required for the survival of cells carrying different spliceosomal mutations with synthetic lethality screens.
Specific Aim 3. Identify novel therapeutic strategies specifically targeting mutant spliceosomal protein function.

What was accomplished under these goals?
For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

Aim 1 accomplishments: We identified that different spliceosomal mutations, despite imparting distinct effects on splicing and gene expression, are negatively selected for when co-expressed in the same cell or in a homozygous state. At the same time, aberrant splicing of distinct targets/events by mutant SF3B1 and SRSF2 results in hyperactivated NF-κB signaling. Specifically, mutations in SF3B1 result in aberrant splicing and consequent downregulation of MAP3K7 while mutations in SRSF2 generate a stable truncated form of Caspase 8. Each of these alterations result in activation of NFk-B signaling. These data identify convergent biological consequences of splicing factor mutations and the basis for their mutual exclusivity and heterozygous nature. Our shared manuscript on these data were published since the last reporting period in Cancer Cell as follows:

Lee SC, North K, Kim E, Jang E, Obeng E, Lu SX, Liu B, Inoue D, Yoshimi A, Ki M, Yeo M, Zhang XJ, Kim MK, Cho H, Chung YR, Taylor J, Durham BH, Kim YJ, Pastore A, Monette S, Palacino J, Seiler M, Buonamici S, Smith PG, Ebert BL, Bradley RK*, Abdel-Wahab O*. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018 Aug 13;34(2):225-241.e8. doi: 10.1016/j.ccell.2018.07.003. PubMed PMID: 30107174; PubMed Central PMCID: PMC6373472. ***Co-corresponding authors.**

Aim 2 accomplishments: *SRSF2*-mutant cells are preferentially sensitive to RBM39-degrading compounds. We previously identified that spliceosomal mutant cells are preferentially sensitive to genetic or pharmacologic perturbations to the splicing process. Our initial work on this focused on pharmacologic inhibition of the SF3b complex of the core spliceosome. In an effort to identify additional RNA binding proteins (RBPs) required in myeloid malignancy cells, we performed an RNA binding domain-focused CRISPR screen to systematically target RBPs across a variety of cancer cell lines. This identified a unique genetic requirement for *RBM39* in myeloid leukemia cells. Serendipitously, several drugs (known as anticancer "sulfonamides," as they all contain a sulfonamide moiety) were recently discovered to specifically degrade RBM39 protein by physically linking RBM39 to the Ddb1/CUL4 E3 ubiquitin ligase complex via an adaptor protein known as DCAF15. We therefore tested those compounds to discover that splicing factor mutant cells were especially sensitive to RBM39 degraders. Importantly, the safety of RBM39-degrading compounds has been established in previous phase I clinical trials in solid tumor and leukemia patients (NCT00078637, NCT01773421). We are now initiating a phase II clinical trial of the RBM39-degrading agent E7820 in patients with myeloid neoplasms *and an RNA splicing factor mutation* who have relapsed or are refractory to standard therapies. Our manuscript on these data were published this year:

Wang E, Lu SX, Pastore A, Chen X, Imig J, Chun-Wei Lee S, Hockemeyer K, Ghebrechristos YE, Yoshimi A, Inoue D, Ki M, Cho H, Bitner L, Kloetgen A, Lin KT, Uehara T, Owa T, Tibes R, Krainer AR, Abdel-Wahab O*, Aifantis I*. Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell. 2019 Mar 18;35(3):369-384.e7. doi: 10.1016/j.ccell.2019.01.010. Epub 2019 Feb 21. PubMed PMID: 30799057; PubMed Central PMCID: PMC6424627. ***Co-corresponding authors.**

Aim 3 accomplishments: Preferential effect of PRMT inhibition on spliceosomal mutant leukemias. We recently identified that inhibiting symmetric (SDMA) or asymmetric dimethyl arginine methylation (ADMA), mediated by PRMT5 and type I protein arginine methyltransferases (PRMTs) respectively, reduces splicing fidelity and results in strong preferential killing of spliceosomal mutant leukemias over WT counterparts. Consistent with this, proteomic analyses identified RNA binding proteins and splicing factors as the most enriched PRMT5 and/or PRMT Type I substrates in leukemia. Accordingly, combined PRMT5/type I PRMT inhibition resulted in synergistic cell killing and pronounced effects on splicing compared with inhibiting either enzymatic activity alone. These data identify genetic subsets of cancer most likely to respond to PRMT inhibition and a mechanistic basis for therapeutic efficacy of PRMT inhibition in cancer. These data were published in the following publication:

Fong JY, Pignata L, Goy PA, Kawabata KC, Lee SC, Koh CM, Musiani D, Massignani E, Kotini AG, Penson A, Wun CM, Shen Y, Schwarz M, Low DH, Rialdi A, Ki M, Wollmann H, Mzoughi S, Gay F, Thompson C, Hart T, Barbash O, Luciani GM, Szewczyk MM, Wouters BJ, Delwel R, Papapetrou EP, Barsyte-Lovejoy D, Arrowsmith CH, Minden MD, Jin J, Melnick A, Bonaldi T, Abdel-Wahab O*, Guccione E*. Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. Cancer Cell. 2019 Aug 12;36(2):194-209.e9. doi: 10.1016/j.ccell.2019.07.003. PubMed PMID: 31408619. ***Co-corresponding authors.**

Other Accomplishments:

Discovery that RNA splicing and epigenomic alterations act combinatorially to drive myeloid neoplasm development: We studied the transcriptomes of 982 patients with myeloid neoplasms to discover frequently co-occurring mutations in *SRSF2* and *IDH2* that together promote leukemogenesis by coordinately dysregulating RNA splicing and the epigenome. While mutations in either *SRSF2* or *IDH2* alone imparted distinct splicing changes, co-expression of mutant IDH2 altered the splicing effects of mutant SRSF2 and resulted in more profound splicing changes than did either mutation alone. Co-expression of mutant IDH2 and SRSF2 resulted in lethal MDS with proliferative features *in vivo* and enhanced self-renewal in a manner not observed with either mutation alone. These data identify a pathogenic crosstalk between altered splicing and epigenetic state in a subset of leukemias, provide functional evidence that mutations in splicing factors drive myeloid malignancy development, and identify splicing changes as mediators of *IDH2*-mutant leukemogenesis. These data were published in the following recent paper:

Yoshimi A, Lin KT, Wiseman DH, Rahman MA, Pastore A, Wang B, Lee SC, Micol JB, Zhang XJ, de Botton S, Penard-Lacronique V, Stein EM, Cho H, Miles RE, Inoue D, Albrecht TR, Somervaille TCP, Batta K, Amaral F, Simeoni F, Wilks DP, Cargo C, Intlekofer AM, Levine RL, Dvinge H, Bradley RK, Wagner EJ, Krainer AR, Abdel-Wahab O. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. Nature. 2019 Oct 2. doi: 10.1038/s41586-019-1618-0. [Epub ahead of print] PubMed PMID: 31578525.

***SF3B1* mutations disrupt the non-canonical BAF (ncBAF) complex:** Unlike other spliceosomal mutations, which primarily occur in MDS and related hematologic diseases, *SF3B1* mutations are also common in non-hematologic malignancies. In order to understand this distinction, we used the above models of spliceosomal mutations to discover that *SF3B1* mutations found in both hematologic and non-hematologic disease converge on repression of BRD9, a core component of the recently described ncBAF complex, to promote cell transformation. These data explained the pervasive nature of *SF3B1* mutations and also elucidated a potential mechanism-based therapeutic. These data were published in the following recent paper:

Inoue D, Chew GL, Liu B, Michel BC, Pangallo J, D'Avino AR, Hitchman T, North K, Lee SC, Bitner L, Block A, Moore AR, Yoshimi A, Escobar-Hoyos L, Cho H, Penson A, Lu SX, Taylor J, Chen Y, Kadoch C, Abdel-Wahab O*, Bradley RK*. Spliceosomal disruption of the non-canonical BAF complex in cancer. Nature. 2019 Oct 9. doi: 10.1038/s41586-019-1646-9. [Epub ahead of print] PubMed PMID: 31597964. ***Co-corresponding authors**

What opportunities for training and professional development has the project provided?
If the project was not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency. Training activities may include, for example, courses or one-on-one work with a mentor. "Professional development" activities result in increased knowledge or skill in one's area of expertise and may include workshops, conferences, seminars, study groups, and individual study. Include participation in conferences, workshops, and seminars not listed under major activities.

On the basis of the work completed from this award:
-Dr. Stanley Chun-Wei Lee (former postdoctoral fellow in the Abdel-Wahab lab) received an NIH K99/R00, a prestigious Scholar Award from the American Society of Hematology (ASH), and has now started his own independent laboratory as a tenure-track faculty member at the Fred Hutchinson Cancer Research Center.
-Dr. Akihide Yoshimi (former postdoctoral fellow in the Abdel-Wahab lab) received a prestigious Special Fellow Award from the Leukemia & Lymphoma Society and started his own independent laboratory as a tenure-track faculty member at the Japanese National Cancer Institute (Tokyo, Japan).
-Dr. Heidi Dvinge (former Bradley lab member) became appointed as an independent faculty member as an Assistant Professor in the Dept. of Biomolecular Chemistry at the University of Wisconsin, Madison.
-Dr. Daichi Inoue (former Abdel-Wahab lab member) received a Special Fellow Award from the Leukemia & Lymphoma Society and started his own laboratory as a tenure-track faculty member at Kobe University (Kobe, Japan).

How were the results disseminated to communities of interest?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the results were disseminated to communities of interest. Include any outreach activities that were undertaken to reach members of communities who are not usually aware of these project activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

Results from the last year of our research were highlighted in the following lay press research articles online by MSKCC and the Fred Hutchinson Cancer Research Center:
https://www.mskcc.org/blog/splicing-may-be-effective-target-fight-against
https://www.fredhutch.org/en/news/center-news/2019/10/sf3b1-cancer-mutation.html

What do you plan to do during the next reporting period to accomplish the goals?
If this is the final report, state "Nothing to Report."

Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

Nothing to Report

4. **IMPACT:** Describe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:

What was the impact on the development of the principal discipline(s) of the project?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project. Summarize using language that an intelligent lay audience can understand (Scientific American style).

One of the main findings from our work thus far is that cells which contains mutations in RNA splicing factors have enhanced sensitivity to drugs that further alter the RNA splicing process. This has led to a clinical trial of a novel RNA splicing inhibitor in patients with MDS and other forms of myeloid leukemia (clinicaltrials.gov identifier NCT02841540). This is a phase I clinical trial sponsored by the company H3 Biomedicine Inc. and is designed to test the safety of this approach and ability to block splicing in patients. In addition, GSK has launced another phase I clinical trial of their PRMT5 inhibitor for MDS and spliceosomal mutant MDS patients based on this work (NCT03614728). Finally, we are working with Eisai Pharmaceuticals to initiate a phase II trial of the RBM39 degrading compound E7820 in spliceosomal mutant MDS and related myeloid neoplasms now.

What was the impact on other disciplines?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

Nothing to report.

What was the impact on technology transfer?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

- *transfer of results to entities in government or industry;*
- *instances where the research has led to the initiation of a start-up company; or*
- *adoption of new practices.*

Nothing to report.

What was the impact on society beyond science and technology?

If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

- *improving public knowledge, attitudes, skills, and abilities;*
- *changing behavior, practices, decision making, policies (including regulatory policies), or social actions; or*
- *improving social, economic, civic, or environmental conditions.*

Nothing to report.

5. **CHANGES/PROBLEMS:** The Project Director/Principal Investigator (PD/PI) is reminded that the recipient organization is required to obtain prior written approval from the awarding agency Grants Officer whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, "Nothing to Report," if applicable:

Changes in approach and reasons for change
Describe any changes in approach during the reporting period and reasons for these changes. Remember that significant changes in objectives and scope require prior approval of the agency.

Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them
Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

Nothing to report.

Changes that had a significant impact on expenditures
Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period. If required, were these changes approved by the applicable institution committee (or equivalent) and reported to the agency? Also specify the applicable Institutional Review Board/Institutional Animal Care and Use Committee approval dates.

Significant changes in use or care of human subjects

Nothing to report.

Significant changes in use or care of vertebrate animals.

Nothing to report.

Significant changes in use of biohazards and/or select agents

Nothing to report.

6. PRODUCTS: List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state "Nothing to Report."

- **Publications, conference papers, and presentations**
  Report only the major publication(s) resulting from the work under this award.

  Journal publications. *List peer-reviewed articles or papers appearing in scientific, technical, or professional journals. Identify for each publication: Author(s); title; journal; volume: year; page numbers; status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no)*

  - Inoue D, Chew GL, Liu B, Michel BC, Pangallo J, D'Avino AR, Hitchman T, North K, Lee SC, Bitner L, Block A, Moore AR, Yoshimi A, Escobar-Hoyos L, Cho H, Penson A, Lu SX, Taylor J, Chen Y, Kadoch C, Abdel-Wahab O*, Bradley RK*. Spliceosomal disruption of the non-canonical BAF complex in cancer. Nature. 2019 Oct 9. doi: 10.1038/s41586-019-1646-9. [Epub ahead of print] PubMed PMID: 31597964. ***Co-corresponding authors.**

  - Yoshimi A, Lin KT, Wiseman DH, Rahman MA, Pastore A, Wang B, Lee SC, Micol JB, Zhang XJ, de Botton S, Penard-Lacronique V, Stein EM, Cho H, Miles RE, Inoue D, Albrecht TR, Somervaille TCP, Batta K, Amaral F, Simeoni F, Wilks DP, Cargo C, Intlekofer AM, Levine RL, Dvinge H, Bradley RK, Wagner EJ, Krainer AR, Abdel-Wahab O. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. Nature. 2019 Oct 2. doi: 10.1038/s41586-019-1618-0. [Epub ahead of print] PubMed PMID: 31578525.

- Fong JY, Pignata L, Goy PA, Kawabata KC, Lee SC, Koh CM, Musiani D, Massignani E, Kotini AG, Penson A, Wun CM, Shen Y, Schwarz M, Low DH, Rialdi A, Ki M, Wollmann H, Mzoughi S, Gay F, Thompson C, Hart T, Barbash O, Luciani GM, Szewczyk MM, Wouters BJ, Delwel R, Papapetrou EP, Barsyte-Lovejoy D, Arrowsmith CH, Minden MD, Jin J, Melnick A, Bonaldi T, Abdel-Wahab O*, Guccione E*. Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. Cancer Cell. 2019 Aug 12;36(2):194-209.e9. doi: 10.1016/j.ccell.2019.07.003. PubMed PMID: 31408619. ***Co-corresponding authors.**

- Wang E, Lu SX, Pastore A, Chen X, Imig J, Chun-Wei Lee S, Hockemeyer K, Ghebrechristos YE, Yoshimi A, Inoue D, Ki M, Cho H, Bitner L, Kloetgen A, Lin KT, Uehara T, Owa T, Tibes R, Krainer AR, Abdel-Wahab O*, Aifantis I*. Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell. 2019 Mar 18;35(3):369-384.e7. doi: 10.1016/j.ccell.2019.01.010. Epub 2019 Feb 21. PubMed PMID: 30799057; PubMed Central PMCID: PMC6424627. ***Co-corresponding authors.**

- Lee SC, North K, Kim E, Jang E, Obeng E, Lu SX, Liu B, Inoue D, Yoshimi A, Ki M, Yeo M, Zhang XJ, Kim MK, Cho H, Chung YR, Taylor J, Durham BH, Kim YJ, Pastore A, Monette S, Palacino J, Seiler M, Buonamici S, Smith PG, Ebert BL, Bradley RK*, Abdel-Wahab O*. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018 Aug 13;34(2):225-241.e8. doi: 10.1016/j.ccell.2018.07.003. PubMed PMID: 30107174; PubMed Central PMCID: PMC6373472. ***Co-corresponding authors.**

- Wang E, Lu SX, Pastore A, Chen X, Imig J, Chun-Wei Lee S, Hockemeyer K, Ghebrechristos YE, Yoshimi A, Inoue D, Ki M, Cho H, Bitner L, Kloetgen A, Lin KT, Uehara T, Owa T, Tibes R, Krainer AR, Abdel-Wahab O*, Aifantis I*. Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell. 2019 Mar 18;35(3):369-384.e7. doi: 10.1016/j.ccell.2019.01.010. Epub 2019 Feb 21. PubMed PMID: 30799057; PubMed Central PMCID: PMC6424627. ***Co-corresponding authors.**

- Lee SC, North K, Kim E, Jang E, Obeng E, Lu SX, Liu B, Inoue D, Yoshimi A, Ki M, Yeo M, Zhang XJ, Kim MK, Cho H, Chung YR, Taylor J, Durham BH, Kim YJ, Pastore A, Monette S, Palacino J, Seiler M, Buonamici S, Smith PG, Ebert BL, Bradley RK*, Abdel-Wahab O*. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018 Aug 13;34(2):225-241.e8. doi: 10.1016/j.ccell.2018.07.003. PubMed PMID: 30107174; PubMed Central PMCID: PMC6373472. ***Co-corresponding authors.**

Books or other non-periodical, one-time publications. *Report any book, monograph, dissertation, abstract, or the like published as or in a separate publication, rather than a periodical or series. Include any significant publication in the proceedings of a one-time conference or in the report of a one-time study, commission, or the like. Identify for each one-time publication: Author(s); title; editor; title of collection, if applicable; bibliographic information; year; type of publication (e.g., book, thesis or dissertation); status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

Nothing to report.

Other publications, conference papers, and presentations. *Identify any other publications, conference papers and/or presentations not reported above. Specify the status of the publication as noted above. List presentations made during the last year (international, national, local societies, military meetings, etc.). Use an asterisk (*) if presentation produced a manuscript.*

2018 AACR Annual Meeting
2018 American Society for Biochemistry and Molecular Biology Annual Meeting
2018 23rd Congress of the European Hematology Association, Stockholm, Sweden
2018 9th Japanese Society of Hematology International Symposium, Kyoto, Tokyo
2018 Danny Thomas Lecture Series, St. Jude's Children's Research Hospital
2018 Edward P. Evans MDS Foundation Summit, Broad Institute, Cambridge, MA
2018 Cold Spring Harbor Laboratories Seminar Series, Cold Spring Harbor, NY

- **Website(s) or other Internet site(s)**
List the URL for any Internet site(s) that disseminates the results of the research activities. A short description of each site should be provided. It is not necessary to include the publications already specified above in this section.

Results from the last year of our research were highlighted in the following lay press research articles online by MSKCC and the Fred Hutchinson Cancer Research Center:

Article from MSKCC:
https://www.mskcc.org/blog/splicing-may-be-effective-target-fight-against

Article from Fred Hutchinson Cancer Research Center:
https://www.fredhutch.org/en/news/center-news/2019/10/sf3b1-cancer-mutation.html

- **Technologies or techniques**
Identify technologies or techniques that resulted from the research activities. In addition to a description of the technologies or techniques, describe how they will be shared.

Nothing to report.

- **Inventions, patent applications, and/or licenses**
  Identify inventions, patent applications with date, and/or licenses that have resulted from the research. State whether an application is provisional or non-provisional and indicate the application number. Submission of this information as part of an interim research performance progress report is not a substitute for any other invention reporting required under the terms and conditions of an award.

Nothing to report.

- **Other Products**
  Identify any other reportable outcomes that were developed under this project. Reportable outcomes are defined as a research result that is or relates to a product, scientific advance, or research tool that makes a meaningful contribution toward the understanding, prevention, diagnosis, prognosis, treatment, and/or rehabilitation of a disease, injury or condition, or to improve the quality of life. Examples include:
  - *data or databases;*
  - *biospecimen collections;*
  - *audio or video products;*
  - *software;*
  - *models;*
  - *educational aids or curricula;*
  - *instruments or equipment;*
  - *research material (e.g., Germplasm; cell lines, DNA probes, animal models);*
  - *clinical interventions;*
  - *new business creation; and*
  - *other.*

We have deposited RNA-seq, ChIP–seq,and eRRBS data generated as part of this study were deposited in the Gene Expression Omnibus under accession numbers GSE124720 and the NCBI Sequence Read Archive under accession number SRP133673.

7. PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?
Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate "no change."

Example:

Name:	*Mary Smith*
Project Role:	*Graduate Student*
Researcher Identifier (e.g. ORCID ID):	*1234567*
Nearest person month worked:	*5*
Contribution to Project:	*Ms. Smith has performed work in the area of combined error-control and constrained coding.*
Funding Support:	*The Ford Foundation (Complete only if the funding support is provided from other than this award).*

Name: Omar Abdel-Wahab
Project Role: PI
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 1
Contribution to Project: Overseeing all experiments in conjunction with Dr. Bradley.
Funding Support: US National Institutes of Health (NIH)-NHLBI grant R01 HL128239

Name: Robert Bradley
Project Role: co-PI
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 1
Contribution to Project: Overseeing all experiments in conjunction with Dr. Abdel-Wahab.
Funding Support: US National Institutes of Health (NIH)-NHLBI grant R01 HL128239

Name: Stanley Chun-Wei Lee
Project Role: Postdoctoral fellow
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 2
Contribution to Project: Stanley has performed all in vivo experiments related to Specific Aim 1 and the shRNA screen in Aim 2.
Funding Support: Leukemia and Lymphoma Society Special Fellow Award, NIH K99/R00, ASH Scholar

Name: Bo Liu
Project Role: Postdoctoral fellow
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 2
Contribution to Project: Bo generated the splicing reporters described in Aim 3.
Funding Support: Josie Robertson Award (MSKCC)

Name: Young Rock Chung
Project Role: Sr. Research Assistant
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 1
Contribution to Project: Assisted postdoctoral fellows in the Abdel-Wahab lab with experiments.
Funding Support: NIH (R01-CA201247), Pershing Square, W81XWH-16-1-0059

Name: Khrystyna Dilai
Project Role: Graduate Student
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 3
Contribution to Project: Khrystyna performed the informatics analyses to identify convergent effects of spliceosomal gene mutations in Aim 1.
Funding Support: US National Institutes of Health (NIH)-NHLBI grant R01 HL128239

Name: Akihide Yoshimi
Project Role: Postdoctoral fellow
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 2
Contribution to Project: Dr. Yoshimi completed the experiments on H3B-8800.
Funding Support: Leukemia and Lymphoma Society Special Fellow Award

Name: Daichi Inoue
Project Role: Postdoctoral fellow
Researcher Identifier (e.g. ORCID ID): None
Nearest person month worked: 2
Contribution to Project: Dr. Inoue is work on validating results from the CRISPR Screen in Aim 2 and the RBM39 work.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

If the active support has changed for the PD/PI(s) or senior/key personnel, then describe what the change has been. Changes may occur, for example, if a previously active grant has closed and/or if a previously pending grant is now active. Annotate this information so it is clear what has changed from the previous submission. Submission of other support information is not necessary for pending changes or for changes in the level of effort for active support reported previously. The awarding agency may require prior written approval if a change in active other support significantly impacts the effort on the project that is the subject of the project report.

SEE APPENDIX FOR CHANGES IN ACTIVE SUPPORT

What other organizations were involved as partners?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe partner organizations – academic institutions, other nonprofits, industrial or commercial firms, state or local governments, schools or school systems, or other organizations (foreign or domestic) – that were involved with the project. Partner organizations may have provided financial or in-kind support, supplied facilities or equipment, collaborated in the research, exchanged personnel, or otherwise contributed.
Provide the following information for each partnership:

Organization Name:
Location of Organization: (if foreign location list country)
Partner's contribution to the project (identify one or more)

- *Financial support;*
- *In-kind support (e.g., partner makes software, computers, equipment, etc., available to project staff);*
- *Facilities (e.g., project staff use the partner's facilities for project activities);*
- *Collaboration (e.g., partner's staff work with project staff on the project);*
- *Personnel exchanges (e.g., project staff and/or partner's staff use each other's facilities, work at each other's site); and*
- *Other.*

Organization Name: Fred Hutchinson Cancer Research Center
Location of Organization: (if foreign location list country): Seattle, WA
Partner's contribution to the project (identify one or more)
☒ Collaboration (e.g., partner's staff work with project staff on the project).

8. SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS: For collaborative awards, independent reports are required from BOTH the Initiating PI and the Collaborating/Partnering PI. A duplicative report is acceptable; however, tasks shall be clearly marked with the responsible PI and research site. A report shall be submitted to **https://ers.amedd.army.mil** for each unique award.

QUAD CHARTS: If applicable, the Quad Chart (available on https://www.usamraa.army.mil) should be updated and submitted with attachments.

9. APPENDICES:

Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.

- Inoue D, Chew GL, Liu B, Michel BC, Pangallo J, D'Avino AR, Hitchman T, North K, Lee SC, Bitner L, Block A, Moore AR, Yoshimi A, Escobar-Hoyos L, Cho H, Penson A, Lu SX, Taylor J, Chen Y, Kadoch C, Abdel-Wahab O*, Bradley RK*. Spliceosomal disruption of the non-canonical BAF complex in cancer. Nature. 2019 Oct 9. doi: 10.1038/s41586-019-1646-9. [Epub ahead of print] PubMed PMID: 31597964. ***Co-corresponding authors.**

- Yoshimi A, Lin KT, Wiseman DH, Rahman MA, Pastore A, Wang B, Lee SC, Micol JB, Zhang XJ, de Botton S, Penard-Lacronique V, Stein EM, Cho H, Miles RE, Inoue D, Albrecht TR, Somervaille TCP, Batta K, Amaral F, Simeoni F, Wilks DP, Cargo C, Intlekofer AM, Levine RL, Dvinge H, Bradley RK, Wagner EJ, Krainer AR, Abdel-Wahab O. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. Nature. 2019 Oct 2. doi: 10.1038/s41586-019-1618-0. [Epub ahead of print] PubMed PMID: 31578525.

- Fong JY, Pignata L, Goy PA, Kawabata KC, Lee SC, Koh CM, Musiani D, Massignani E, Kotini AG, Penson A, Wun CM, Shen Y, Schwarz M, Low DH, Rialdi A, Ki M, Wollmann H, Mzoughi S, Gay F, Thompson C, Hart T, Barbash O, Luciani GM, Szewczyk MM, Wouters BJ, Delwel R, Papapetrou EP, Barsyte-Lovejoy D, Arrowsmith CH, Minden MD, Jin J, Melnick A, Bonaldi T, Abdel-Wahab O*, Guccione E*. Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. Cancer Cell. 2019 Aug 12;36(2):194-209.e9. doi: 10.1016/j.ccell.2019.07.003. PubMed PMID: 31408619. ***Co-corresponding authors.**

- Wang E, Lu SX, Pastore A, Chen X, Imig J, Chun-Wei Lee S, Hockemeyer K, Ghebrechristos YE, Yoshimi A, Inoue D, Ki M, Cho H, Bitner L, Kloetgen A, Lin KT, Uehara T, Owa T, Tibes R, Krainer AR, Abdel-Wahab O*, Aifantis I*. Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell. 2019 Mar 18;35(3):369-384.e7. doi: 10.1016/j.ccell.2019.01.010. Epub 2019 Feb 21. PubMed PMID: 30799057; PubMed Central PMCID: PMC6424627. ***Co-corresponding authors.**

- Lee SC, North K, Kim E, Jang E, Obeng E, Lu SX, Liu B, Inoue D, Yoshimi A, Ki M, Yeo M, Zhang XJ, Kim MK, Cho H, Chung YR, Taylor J, Durham BH, Kim YJ, Pastore A, Monette S, Palacino J, Seiler M, Buonamici S, Smith PG, Ebert BL, Bradley RK*, Abdel-Wahab O*. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018 Aug 13;34(2):225-241.e8. doi: 10.1016/j.ccell.2018.07.003. PubMed PMID: 30107174; PubMed Central PMCID: PMC6373472. ***Co-corresponding authors.**

OMAR I. ABDEL-WAHAB, M.D.

ACTIVE SUPPORT

W81XWH-18-1-0383 (PI: Abdel-Wahab) 8/15/2018 - 8/14/2021 0.91 calendar
Congressionally Directed Medical Research Programs $ 100,000
Mechanistic and Therapeutic Implications of Spliceosomal Gene Mutations in ER+ Breast Cancer
Goals: Our grant aims to (1) identify what drives breast cancer growth; determine how to stop it and (2) revolutionize treatment regimens by replacing them with ones that are more effective, less toxic, and impact survival.
Specific Aims: (1) Determine the mechanism by which the SF3B1 K700E mutation promotes hormoneindependent growth of breast cancer, (2) Examine the effect of SF3B1 K700E mutation on mammary tumorigenesis in vivo, (3) Determine the therapeutic benefit of modulating RNA splicing in genetically defined breast cancers with and without spliceosomal gene mutations.
Agency Contract: Grants Specialist: Jodi Cardoza Phone: 301-619-2693 Email: jodi.l.cardoza.civ@mail.mil
Overlap: None

1 UG1 CA233332-01 (PI: Abdel-Wahab) 3/7/2019 - 2/28/2025 0.46 calendar
NCI $ 737,015
ECOG-ACRIN Integrated Leukemia Translational Science Center (LTSC)
Advances in the treatment of hematologic malignancies have been disappointing. The ECOG-ACRIN
Goals: Leukemia Translational Science Center will create, support, and synergize partnerships between clinical and laboratory investigators and foster the continuous, near-term translation of state-of-the art laboratory studies into the clinical context to improve outcomes for leukemia patients.
Specific Aims: 1. To identify the molecular signatures which define clinical, prognostic, and therapeutically distinct leukemia subtypes, and can be used to develop risk-stratified therapeutic protocols and/or to therapy with mechanism-based therapies for biomarker-defined disease subsets. 2. To evaluate the mechanism of action, therapeutic efficacy, and predictors of response/resistance of mechanism based therapies for leukemia patients. 3. To perform dynamic, biomarker based molecular assessment of response to leukemia therapies which can be used to monitor disease response, inform mechanisms of sensitivity and resistance to leukemia therapies, and to design clinical trials which incorporate dynamic biomarker profiling into clinical trial design and clinical care.
Agency Contract: Grants Management Specialist: Barbara Hodgkins Email: barb.hodgkins@nih.gov Phone: (240) 276-6294
Overlap: None

I12-0063 (PI: Abdel-Wahab) 1/1/2019 - 12/31/2020 0.24 calendar
Starr Cancer Consortium $ 75,000
Identification of transcriptional determinants of asparaginase sensitivity in leukemias
Goals: Asparaginase, which selectively depletes asparagine from serum, kills asparagine-dependent leukemia cells and is a first line chemotherapeutic agent. However, the precise mechanisms underlying the asparagine response of leukemic cells and whether there are other cancer types dependent on particular nutrients remain poorly understood.
Specific Aims: Aim 1: Determine the precise mechanism by which ZBTB1 enables ALLs to survive under asparagine depletion. Aim 2: Examine the clinical significance of ZBTB1 in ALL and other lymphoid malignancies. Aim 3: Map the amino acid dependencies of blood cancer cell lines using DNA-barcoding technology.
Agency Contract: Sylvie Le Blancq, PhD, Executive Director, Sloan Kettering Institute, 1275 York Avenue, New York, NY 10065, lesblanc2@mskcc.org, 646-888-3773
Overlap: None

1 UG1 CA233290-01 (PI: Aghajanian / Lee / Morris / Tallman / Zivanovic) 3/6/2019 - 2/28/2025 0.60 calendar
NCI $ 9,615
Network Lead Academic Participating Site: Memorial Sloan Kettering Cancer Center
Goals: This goal will be achieved through the continued successful development and execution of definitive, randomized, clinical treatment and advanced imaging trials across a broad range of diseases and diverse patient populations.
Specific Aims: 1. Scientific Leadership in Development of National Clinical Trials Network (NCTN) Trials 2. Scientific

Leadership in the Activities of the NCTN and NCI Scientific Steering Committees 3. Patient Accrual on NCTN Trials 4. Mentoring Junior Investigators in Clinical Trial Research
Agency Contract: Grants Management Specialist: Barbara Hodgkins Email: barb.hodgkins@nih.gov Phone: (240) 276-6294
Overlap: None

7018-19 (PI: Cleveland) 10/1/2018 - 9/30/2023 0.60 calendar
Leukemia and Lymphoma Society $ 178,333
Regulation and Targeting of Inflammatory Circuits in Myelodysplastic Syndromes
Goals & Specific Aims: Aim 1: Determine the mechanistic basis for hyper-activated innate immune signaling in spliceosome-mutant MDS bearing mutations in RNA splicing factors. Aim 2: Define the therapeutic potential for MAP3K7, p38/MAPK, and TGF-□ pathway inhibition in spliceosome mutant MDS. Aim 3: Define the mechanistic basis for preferential dependency on IFN-JAK-STAT signaling in SF3B1 mutant cells.
Agency Contract: Director of Research Administration, 3 International Drive, Suite 200, Rye Brook, New York 10573, researchprograms@lls.org
Overlap: None

DRG2018 (PI: Figueroa) 9/1/2018 - 8/31/2020 0.24 calendar
Edward P. Evans Foundation $ 90,909
Determining the Contribution of MDS-associated RNA Splicing Factor Mutations to Altered DNA Methylation
Goals & Specific Aims: As part of this collaboration, our lab will perform analyses to integrate DNA methylation and hydroxy-cytosine methylation data with RNA splicing data. We will also attempt to understand how alterations in DNA methylation alter splicing through focused studies of splicing and expression of potential regulators of DNA methylation. In addition, we will perform locus-specific biochemical experiments to understand how alterations in the epigenome impact splicing.
Agency Contract: Michael Lewis, Ph.D. President,mdl@epefoundation.org 978 494-6009
Overlap: None

1R01HL145283-01 (PI: Landau) 2/1/2019 - 11/30/2022 0.20 calendar
NHLBI $ 87,279
The role of DNA methylation modifiers in shaping the hematopoietic differentiation topology
Goals & Specific Aims: The Abdel-Wahab will generate and maintain the murine models used in Aims 1-2 as well as perform purification of mouse and human cell populations in all aims of the grant. The Abdel-Wahab lab will also perform the cell transduction and differentiation experiments described in Aim 2.
Agency Contract: Grants Management Specialist: Laurel Katherine Kennedy Email: laurel.kennedy@nih.gov Phone: 301-82704777
Overlap: None

W81-XWH-18-1-0265 (PI: Liu) 7/1/2018 - 6/30/2020 0.56 calendar
Congressionally Directed Medical Research Programs $ 12,136
Understanding and Targeting Mutant p53 in Myelodysplastic Syndromes
Goals and Specific Aims: Task 1. We will assist in the generation of patient-derived xenografts from patients with myelodysplastic syndromes (MDS). Task 2. We will generate p53R248W/+SRSF2P95H/+ -Mx1-Cre+ mice.
Agency Contract: Grant Specialist: Karen L. Petrore Phone: 301-619-3590 Email: Karen.l.petrore.civ@mail.mil
Overlap: None

GC238654 (PI: Shlush) 11/1/2018 - 10/31/2023 1.60 calendar
Leukemia and Lymphoma Society/ Rising Tide Foundation $ 363,636
Early diagnosis and treatment of pre-leukemia
Goals: Successful completion of our aims could have major ramifications on how subjects with ARCH are evaluated and treated, which could, in turn, have a major impact on development of MDS and AML. We now propose to utilize this knowledge to treat high-risk individuals with ARCH, at a time point before they have developed disease, by targeting the driving alterations most associated with AML

development.
Specific Aims: Aim 1. Develop novel efficient strategies to identify individuals with preleukemia. Aim 2. Determine the molecular and functional consequences of SMMs in human preleukemia. Aim 3. Identify targeted therapies effective at eliminating preL-HSPCs carrying SMMs.
Agency Contract: Director of Research Administration, 3 International Drive, Suite 200, Rye Brook, New York 10573, researchprograms@lls.org
Overlap: None

5 R01 HL138090-03 (PI: Geissmann) 8/16/2017 - 6/30/2021 0.60 calendar
NHLBI $ 415,910
Lineages and pathophysiology of clonal histiocytic disorders
Goals: This project aims to characterize at the molecular, cellular and organismal levels the pathophysiology of clonal histiocytic disorders, exemplified by Langerhans Cell Histiocytosis (LCH) and Erdheim-Chester Disease (ECD).
Specific Aims: We will (Aim 1) investigate the pathological, molecular and cellular mechanisms that underlie the development of the brain neurodegenerative disease we observe in our genetic model Csf1rMeriCreMer; BRAFV600E; Rosa26LSL-YFP in which we introduce a BRAFV600E allele in a small number of EMPs. Next (Aim 2) we will determine the pathological consequences of targeting BRAFV600E outside the brain and in particular test whether environmental factors can influence the pathogenesis of histiocytosis in the liver. We will also investigate the extent to which lineage(s) of origin of histiocytoses determines pathological outcomes using a murine model to conditionally target BRAFV600E expression to a small faction of HSCs. Finally (Aim 3) we will investigate the efficiency of BRAF inhibitor administration to prevent or treat the neurodegenerative disease in our models.
Agency Contract: Renee Livshin, livshinr@nhlbi.nih.gov, 301-435-0174
Overlap: None

LLS-02 (PI: Nimer) 10/1/2017 - 9/30/2022 0.20 calendar
Leukemia and Lymphoma Society $ 77,500
Interventional Epigenetics in Myeloid Malignancies
Goals: Our ultimate goal is to define the mechanistic epigenetic basis for the development and progression of myeloid malignancies and to develop novel and effective epigenetic-focused therapies for patients with these disorders.
Specific Aims: Major Task 1: Determine genome-wide localization of ASXL1 versus ASXL2. Major Task 2: Characterize the effects of ASXL2 loss in the setting of AML1-ETO. Major Task 3: Determine the effect of p300 inhibition in AML1-ETO cells with or without Asxl2 loss. Major Task 4: Test LSD1 inhibition in ASXL1/2-deficient AML cells. Major Task 5: Identify genes which are synthetic lethal with ASXL1-mutant cells by performing functional genomic screens.
Agency Contract: Director of Research Administration, 3 International Drive, Suite 200, Rye Brook, New York 10573, researchprograms@lls.org
Overlap: None

6533-18 (PI: Park) 10/1/2017 - 9/30/2020 0.30 calendar
Leukemia and Lymphoma Society $ 180,018
Chemotherapy-Free Targeted Therapeutic Approaches for New and Relapsed Hairy Cell Leukemia
Goals: This proposed study has a potential to challenge and change the current standard treatment practice of frontline HCL and offer chemotherapy-free, rationally targeted therapeutic options in HCL.
Specific Aims: Aim 1: Determine the efficacy of combined RAF inhibition and anti-CD20 immunotherapy as frontline therapy for classical hairy cell leukemia (cHCL). Aim 2: Determine the safety and efficacy of ERK inhibition in relapsed/refractory (R/R) cHCL and hairy cell leukemia variant (HCLv).
Agency Contract: Director of Research Administration, 3 International Drive, Suite 200, Rye Brook, New York 10573, researchprograms@lls.org
Overlap: None

I10-0074 (PI: Abdel-Wahab) 1/1/2017 - 12/31/2019 NCE 0.10 calendar
Starr Cancer Consortium $ 197,500
Dissecting micro-environmental contributions to the pathogenesis of myelodysplastic syndromes
Goals: Our proposal aims at highlighting the molecular alterations in the HMEV of MDS patients that could serve as novel therapeutic targets.
Specific Aims: Specific Aim 1: Identify the functional impact of molecular alterations in the microenvironment of MDS patients on hematopoiesis. We will investigate the mechanisms through which molecular alterations identified in our gene

expression analysis of mesenchymal subsets in MDS patients contribute to MDS pathogenesis.
Agency Contract: Sylvie Le Blancq, PhD, Executive Director, Sloan Kettering Institute, 1275 York Avenue, New York, NY 10065, lesblanc2@mskcc.org, 646-888-3773
Overlap: None

GC238545 (PI: Abdel-Wahab) 1/15/2017 - 1/14/2020 0.60 calendar
Henry & Marilyn Taub Foundation $ 181,818
Understanding and manipulating pathologic splicing preferences and nonsense-mediated mRNA decay (NMD) in spliceosomal-mutant MDS
Goals: Here we aim to develop two novel mechanism-based approaches targeting spliceosomal-mutant MDS based on: (1) selectively targeting the pathologic RNA-binding preferences of mutant spliceosomal proteins; and (2) restoring the expression of key transcripts targeted for nonsense-mediated decay (NMD) due to aberrant splicing. Understanding the mechanistic link between NMD and spliceosomal gene mutations will be critical for this latter aim.
Specific Aims: Aim I. Determine the precise mechanism by which mutant SRSF2 promotes NMD. Aim II. Utilize antisense techniques to restore expression of mRNAs pathologically degraded by NMD in spliceosomal-mutant MDS cells. Aim III. Determine the therapeutic potential of oligonucleotide decoys that exploit the RNAbinding preferences of WT versus mutant SRSF2 in SRSF2-mutant malignant cells.
Agency Contract: Erin Johnstone, Program Officer, The Medical Foundation, ejohnstone@hria.org
Overlap: None

1R01HL128239-04 (PI: Bradley) 8/1/15 - 5/31/2020 2.26 calendar
NIH $ 181,189
Genetic and molecular basis for SRSF2 mutations in myelodysplasia
Goals: Measure the therapeutic efficacy of drugs that inhibit RNA splicing, and test whether drug-induced splicing inhibition specifically affects SRSF2 mutation-responsive exons versus causing widespread splicing failure.
Specific Aims: Perform functional studies to determine the effect of specific mis-spliced isoforms on myeloid differentiation. Test the hypothesis that a wildtype copy of SRSF2 s necessary for hematopoietic cell survival and maintenance of functional splicing. Determine whether SRSF2 mutations do not co-occur with other spliceosomal mutations because they are functionally redundant, or whether multiple spliceosomal mutations are deleterious to the cell. Test the hypothesis that commonly co-occurring SRSF2 and epigenetic factor mutations are synergistic, leading to more aggressive disease and/or more extreme splicing changes. Measure the therapeutic efficacy of drugs that inhibit RNA splicing, and test whether drug-induced splicing inhibition specifically affects SRSF2 mutation-responsive exons versus causing widespread splicing failure. Generate cells to be utilized in CRISPR/Cas9 screen and perform validation of sgRNAs from the screen.
Agency Contract: Kelly Stewart, Fred Hutchinson Cancer Res. Center, kstewart@fredhutch.org, 206-667-6925
Overlap: None

5 R01 CA201247-04 (PI: Abdel-Wahab / Park) 9/1/16 - 8/31/2021 1.81 calendar months
NCI $ 298,536
Origins of BRAF-mutant hematologic malignancies and their therapeutic resistance
Goals: We expect our studies to elucidate mechanisms that underlying the development of these disparate blood cancers driven by a shared mutation, provide the first description of the genetic events that likely cooperate with the BRAFV600E mutation to drive these disorders and therapy resistance.
Specific Aims: Aim 1. Determine the mechanistic basis for the divergent phenotypes of BRAFV600E mutant hematopoietic disorders. Aim 2. Determine whether recurrent mutations co-occurring with the BRAFV600E mutation HCL and SH alter the disease phenotype. Aim 3: Define mechanisms of vemurafenib resistance in HCL using patient samples and murine models of BRAFV600E-mutant hematopoietic malignancies.
Agency Contract: Funmi Elesinmogun, Grants Management Specialist, elesinmf@mail.nih.gov, 240-276-6313
Overlap: None

2314-17 (PI: Abdel Wahab) 7/1/2016 6/30/2021 0.0 calendar months*
Leukemia and Lymphoma Society $ 104,762
Investigating and Targeting Diverse Kinase Alterations Driving Systemic Histiocytic Neoplasms
Goals: The discovery of BRAFV600E mutations in 50% of patients with systemic histiocytoses provided the first molecular target in these disorders. More recently, MAP2K1, ARAF, and N/KRAS mutations as well as kinase fusions

have been identified in BRAFV600 wildtype patients. Here we propose to understand the clinical and biological importance of diverse kinase alterations in BRAFV600E wildtype histiocytic neoplasms.
Specific Aims: Aim 1: Determine the clinical activity of cobimetinib in adult systemic histiocytic disorder patients as measured by radiologic (RECIST) response criteria, metabolic (FDG-PET) response criteria, and longitudinal assessment of somatic mutation burden in cell-free DNA (cfDNA). Aim 2: Identify the biological importance of mutations outside of BRAFV600E on histiocytosis pathogenesis and the molecular determinants of response to MEK inhibition.
Agency Contract: Director of Research Administration, 3 International Drive, Suite 200, Rye Brook, New York 10573, researchprograms@lls.org
Overlap: None
**This grant provides salary support only. Effort is not required per the sponsor.*

6581-20 (PI: Liu) 7/1/2019 - 6/30/2022 0.24 calendar Leukemia and Lymphoma Society $ 60,000
Development of therapeutic strategy for the treatment of MDS
Goals & Specific Aims: AIM 1: Elucidate the mechanisms by which mutant p53 alters pre-mRNA splicing in HSCs and MDS cells. Dr. Abdel-Wahab will provide assistance with analysis of splicing in hematopoietic cells with or without mutations in p53. AIM 2. Determine the impact of inhibition of EZH2 and the spliceosome on MDS cells with TP53 mutations. Dr. Abdel-Wahab will test several therapies targeting RNA splicing in vitro and in vivo in primary human MDS and AML cells with or without mutations in TP5
Agency Contract: Director of Research Administration, 3 International Drive, Suite 200, Rye Brook, New York 10573, researchprograms@lls.org
Overlap: None

PENDING SUPPORT

R01HL128239 (PI: Bradley) 7/1/20 - 6/30/2024 1.8 calendar
NIH $ 212,885
Genetic and molecular basis for SRSF2 mutations in myelodysplasia
Goals: Measure the therapeutic efficacy of drugs that inhibit RNA splicing, and test whether drug-induced splicing inhibition specifically affects SRSF2 mutation-responsive exons versus causing widespread splicing failure.
Specific Aims: Perform functional studies to determine the effect of specific mis-spliced isoforms on myeloid differentiation. Test the hypothesis that a wildtype copy of SRSF2 s necessary for hematopoietic cell survival and maintenance of functional splicing. Determine whether SRSF2 mutations do not co-occur with other spliceosomal mutations because they are functionally redundant, or whether multiple spliceosomal mutations are deleterious to the cell. Test the hypothesis that commonly co-occurring SRSF2 and epigenetic factor mutations are synergistic, leading to more aggressive disease and/or more extreme splicing changes. Measure the therapeutic efficacy of drugs that inhibit RNA splicing, and test whether drug-induced splicing inhibition specifically affects SRSF2 mutation-responsive exons versus causing widespread splicing failure. Generate cells to be utilized in CRISPR/Cas9 screen and perform validation of sgRNAs from the screen.
Agency Contract: Kelly Stewart, Fred Hutchinson Cancer Res. Center, kstewart@fredhutch.org, 206-667-6925
Overlap: None

1 R01 CA251138-01 (PI: Abdel-Wahab) 7/1/2020 - 6/30/2025 1.80 calendar
NCI $ 621,000
Interrogating the minor spliceosome to understand and treat leukemia
Goals:Leukemias are cancers in which the bone marrow produces too many abnormal blood cells at the expense of normal blood cells. Here we will determine how a commonly occurring change in a gene called ZRSR2, which encodes a protein that participates in the process of RNA splicing, gives rise to leukemia. We will determine how ZRSR2 mutati ons drive leukemia and find new ways to treat leukemias with ZRSR2 mutations.
Agency Contract: N/A
Overlap: None

CA190171P2 (PI: Abdel-Wahab) 9/30/2020 - 9/29/2024 0.60 calendar
Congressionally Directed Medical Research Programs $ 100,000
Identifying Novel Therapeutic Targets and Combination Strategies for Patients with BPDCN
Goals:Test the hypothesis that in vivo resistance to tagraxofusp in BPDCN cells is mediated by epigenetic silencing o f diphthamide synthesis genes.

Agency Contract: N/A
Overlap: None

1 R01 CA242020-01A1 (PI: Abdel-Wahab/Aifantis) 4/1/2020 - 3/31/2025 1.8 calendar
NCI $ 515,704
Targeting an RNA-binding Protein Network in Acute Myeloid Leukemia
Goals:Recent studies have identified that proteins that bind RNA to regulate RNA translation, stability or splicing, pl aykey roles in human cancer. We recently identified a group of such RNA binding proteins (RBPs) that are required for t he survival of acute myeloid leukemia (AML), a deadly form of cancer in great need of improved therapy. In this prop osal we aim to understand the role of one such protein called RBM39, which is a putative, novel, and exciting new AML d rug target.
Specific Aims: Aim 1: Determine the biological role of RBM39 in normal and malignant hematopoiesis. Aim 2: Identify the mechanistic basis for cell-type and disease-specific roles for RBM39 in normal and malignant hematopoietic cells.
Agency Contract: N/A
Overlap: None

GC239063 (PI: Akerman) 4/5/2019 - 4/5/2024 0.24 calendar
NIH $ 36,192
SpliceIO: A machine learning approach to discover novel cancer specific antigens through splicing analysis using RNA seq data
Goals: Here, we propose to leverage and extend this resource to develop SpliceIO, a novel discovery platform that integrates the SpliceCore knowledge base with machine learning (ML) to identify novel TAEs resulting from aberrant splicing. This strategic development will lead Envisagenics to disrupt IO with an unprecedent modality for RNAseq analysis, creating partnership opportunities for co-develop of IO products with pharma.
Specific Aims: Aim 1. Develop SpliceIO to identify and predict viable alternatively spliced ectodomains. Aim 2. Identify and predict antigenicity of AS-derived TAE from public pediatric AML data to assess utility of SpliceIO in-silico. Aim 3. Leverage SpliceIO to discover viable AS-derived ectodomains associated with AML relapse and validate targets in-vitro.
Agency Contract: PI Name: Martin Akerman, PhD Email: makerman@envisagenics.com Tel: (631)697-9471
Overlap: None

GC240554 (PI: Guccione) 4/1/2020 - 3/31/2025 0.60 calendar
NIH $ 200,000
Therapeutic targeting of RNA splicing catalysis through inhibition of Protein Arginine Methylation
Goals & Specific Aims: Aim 1: Assess the synergistic therapeutic potential of drugs targeting the extended splicing network. Aim 2: Identify the molecular basis for the synergistic cooperation of PRMT5 and Type I PRMT inhibition.
Agency Contract: N/A
Overlap: None

GC241577 (PI: Liu) 7/1/2020 - 6/30/2025 1.20 calendar
NIH
Molecular mechanisms underlying age-related clonal hematopoiesis and diseases
Aims: SPECIFIC AIM 1: Determine the impact of mutant p53-
mediated chronic inflammation on ARCH progression and MDS pathogenesis during aging. SPECIFIC AIM 2: Determine the impact of mutant p53-driven dysregulation of pre-mRNA splicing on MDS pathogenesis with age.
Agency Contract: N/A
Overlap: None

COMPLETED SUPPORT

GC235782 (PI: Abdel-Wahab) 9/1/2016 - 8/31/2019 0.20 calendar
Edward P. Evans Foundation $ 181,818
Elucidating Critical Targets, Transcripts, and Collaborating Events in Spliceosomal-Mutant MDS
Goals & Specific Aims: The project goal is to understand which abnormally spliced products are functionally important for the development of myelodysplastic syndrome phenotypes.
Agency Contract: Michael Lewis, Ph.D. President,mdl@epefoundation.org 978 494-6009
Overlap: None

190454 (PI: Varmus) 9/1/2017 - 8/31/2019 0.24 calendar
Edward P. Evans Foundation $ 159,207
The Roles of Mutant U2AF1 in the Initiation and Maintenance of MDS
Goals: Our work will improve our basic understanding of key genetic alterations in MDS & provide crucial insights into genetic targets for therapeutic intervention in MDS. Importantly, the majority of our proposed experiments will be conducted on primary human hematopoietic tissues from MDS patients or healthy donors, combining the pre--clinical and clinical expertise of the Varmus and Abdel--Wahab labs. Therefore, the knowledge gained will be directly relevant to MDS patients carrying U2AF1 mutations.
Specific Aims: Specific Aim I: Determine if mutant U2AF1 is required for MDS disease maintenance. Specific Aim II: Determine the molecular targets of U2AF1(S34F) responsible for MDS development.
Agency Contract: Michael Lewis, Ph.D. President,mdl@epefoundation.org 978 494-6009
Overlap: None

GC231058 (PI: Abdel-Wahab) 1/1/2018 - 12/31/2018
Histiocytosis Association $ 50,000
Functional Evaluation of Immunologic Checkpoint Blockade in the Therapy of Histiocytosis
Goals: Thus, our proposal will provide greater insight into recurrent genomic alterations within the pediatric histiocytosis population as well as inform novel therapeutic approaches for children and adults with histiocytosis.
Specific Aims: The recent success of immune checkpoint blockade therapy in other cancer types combined with the expression of immune checkpoint antigens and abundant immune infiltrates in histiocytosis lesions suggests that such therapies may be effective for histiocytosis patients. We therefore seek here to (1) test the efficacy of checkpoint inhibitors (anti-CTLA4, anti-PD-1/PD-L1) alone or in combination with BRAF inhibition in in vivo models of LCH and (2) determine molecular correlates of PD-L1 expression.
Agency Contract: Grants Office, grants@histio.org, 856-589-6614
Overlap: None

1R56AG052501-01A1 (PI: Liu) 9/15/2017 - 8/31/2018
NIA $ 7,007
Mutant p53 rejuvenates aged stem cells through modulating epigenetic regulators
Goals & Specific Aims: Milestone Task 1: Determine if MDS cells expressing mutant TP53 will accelerate the onset of MDS in NSG-SGM3 mice. MDS cells are difficult to engraft immunodeficient mice. Milestone Task 2: Determine whether mutant SRSF2 collaborates with gain-of-function mutations in TP53 in vivo.
Agency Contract: Grant Management Specialist: Eva Lawson-lipchin Email: lawsonlipchine@nia.nih.gov Phone: 301-496-9350
Overlap: None

GA-2017-B0149 (PI: Park) 1/1/2018 - 12/31/2018
Hairy Cell Leukemia Foundation $ 25,000
Investigating the Basis of Hairy Cell Leukemia and Its Response to Therapy Through Single-Cell Approaches
Goals & Specific Aims: The Abdel-Wahab lab will perform flow cytometric sorting of primary patient samples from HCL patients for single cell RNA sequencing studies. In addition, they will perform bulk cell
RNA-sequencing in addition to mutational analysis by whole exome sequencing and/or MSKCC
IMPACT targeted mutational analysis.
Agency Contract: Grants Management Specialist: Kwesi Wright Email: wrightnk@mail.nih.gov Phone: (301) 451-4789
Overlap: None

W81XWH-16-1-0059 (PI: Abdel Wahab) 5/1/2016 - 4/30/2019 NCE
Congressionally Directed Medical Research Programs $ 180,000
Goals: Therapeutic targeting of spliceosomal mutant acquired bone marrow failure disorders
The purpose of the current DOD application is to test inhibitors of TGF-beta signaling pathway in preclinical mouse models as possible therapy for FA.
Specific Aims:
Agency Contract:
Overlap: None

GC231183(PI: Abdel-Wahab) 7/1/16 - 6/30/19

Pershing Square Sohn Cancer Research Alliance $ 181,818
Identification of novel transcripts, pathways, and therapeutic strategies to target spliceosomal-mutant malignancies
Goals:
Specific Aims: Aim I. Determine the functionally relevant splice isoforms linking spliceosomal gene mutations to cancer development. Aim II: Explore novel therapeutic approaches for spliceosomal-mutant cancers through targeting altered RNA binding preferences and new spliceosomal inhibitors. Aim III. Investigate the intersection of spliceosomal mutations and epigenetic alterations in myeloid malignancies.
Overlap: None
Agency Contract:
*Effort commitment is not required per the sponsor.

1 R13 TR001695-01 (PI: Diamond) 8/18/16 - 8/17/17
NCATS $ 10,000
Fourth International Medical Symposium on Erdheim-Chester Disease
Goals: Because of the rarity of the Erdheim-Chester disease (ECD), clinical experience is limited worldwide and international collaboration has proven vital to advancement in research and clinical care for these patients.
The proposed conference will be an international symposium to gather ECD specialists from the US and Europe to discuss clinical and research efforts, and specifically to derive a consensus approach to response criteria for implementation in ECD clinical trials.
Specific Aims: Aim 1. To bring together clinicians and investigators from an international setting who treat and study patients with Erdheim-Chester disease (ECD). Aim 2. To assess the current and updated collective experience regarding genomic alterations and targeted therapies for ECD. Aim 3. To define, by consensus approach of ECD experts, response criteria for therapeutic clinical trials for ECD.
Agency Contact: Leslie Le, Grants Management Specialist, (e) leleslie@mail.nih.gov, (P) 301-435-0856 Overlap: None

GC228236 (PI: J. Park) 7/1/14-6/30/17
Conquer Cancer Foundation $62,466
A Phase II Study of the BRAF Inhibitor, Vemurafenib, in Patients with Relapsed or Refractory Hairy Cell Leukemia (HCL): Evaluating Clinical Efficacy and Molecular Mechanisms of BRAF Inhibition in HCL
Specific Aims: Aim 1: To determine the clinical efficacy of the BRAF inhibitor vemurafenib in patients with BRAFV600E+ relapsed or refractory HCL, as assessed by overall response rates, duration of response, kinetics of minimal residual disease (MRD), disease-free and overall survival. Aim 2: To assess the effect of the BRAF inhibitor vemura.
Agency Contact: grants@conquercancerfoundation.org, 571-483-1700
Overlap: None

GC226666 (PI: Abdel-Wahab) 1/1/16 - 12/31/16
Histiocytosis Association $ 50,000
MEK Inhibition in the Therapy of Histiocytic Neoplasms
Goals: We have recently completed a clinical trial of vemurafenib for the adult BRAFV600E-mutant patients and we now seek to identify novel therapies that will benefit those patients that are BRAFV600-wildtype. This grant would provide critical funding to (1) allow us to understand clinical responses in this clinical trial in greater depth and (2) provide critical support which will allow us to get additional data which will serve as the basis of a multi-PI R01 application with other already identified investigators on the biology and therapy of MAP2K1-mutant malignancies.
Specific Aims: Aim 1: Determine the clinical activity of cobimetinib in adult systemic histiocytic disorder patients as measured by radiologic (RECIST) response criteria, metabolic (FDG-PET) response criteria, and longitudinal assessment of somatic mutation burden in cell-free DNA (cfDNA). Aim 2: Functionally characterize the novel kinase mutations and fusions identified in histiocytic disorders without the BRAFV600E mutation and evaluate the in vitro efficacy of MEK inhibition in the activating mutants. Results of these studies will be compared to clinical outcome data generated by the cobimetinib trial.
Agency contact: Grants Office, grants@histio.org, 856-589-6614
Overlap: None

GC223032 (PI: Abdel-Wahab) 10/1/14-9/30/16
The V Foundation $200,000
Understanding and Targeting Spliceosomal Mutations in Leukemia

Goals & Specific Aims: Aim 1: Determine how SRSF2 mutations alter SRSF2's normal roles in pre-mRNA splicing Aim 2: Utilize Srsf2 P95H mutant murine models to further identify transcriptional events and splicing changes critical to hematopoietic alterations driven by SRSF2 mutations Aim 3: Identify and test therapeutic strategies for targeting leukemic clones with spliceosomal gene mutations.
Agency contact: Carole C. Wegner, Director of Grants, 106 Towerview Court, Cary, NC 27513, 919-380-95005 Overlap: None

GC221801 (PI: Abdel-Wahab) 10/1/13-9/30/14
Erdheim-Chester Disease Global Alliance $49,261
Somatic Genetic Alterations in the Pathogenesis and Therapy of Histiocytic Disorders
Goals: Histiocytic disorders, including Langerhans cell histiocytosos (LCH) and Erdheim Chester disease (ECD), represent a wide range of rare diseases with heterogeneous clinical course and prognoses. The relative rarity and protean clinical nature of the histiocytic disorders has made it difficult to clearly delineate the pathophysiology of these conditions.
Agency Contact: Kathy Brewer, President, support@erdheim-chester.org.
Overlap: None

5 U10 CA021115-39 (PI: Comis) 2/1/14 - 1/31/15
NIH $ 48,548
Investigating the importance of MAP kinase signaling and novel mutations activating NOTCH signaling in CLL.
Goals: We propose to assess the status of somatic genetic abnormalities in CLL patient samples from the Integroup trial E2997 as performed in our pilot experiments.
Specific Aims: 1. Targeted resequencing and exome sequencing to identify novel somatic mutations that impact outcome and therapeutic response in CLL. 2. RNA-seq analysis to identify prognostically relevant expression changes in CLL. 3. Epigenetic studies to determine the impact of epigenomic alterations on CLL prognosis and therapeutic response.
Agency contact: Debra Strandberg, Strandberg.debra@jimmy.harvard.edu
Overlap: None

1 U10 CA180827-01 (PI: Paietta) 4/14/14 – 2/28/15
NIH $ 45,885
Genetic and Epigenetic Determinants of response to therapy and outcome in patients with chronic lymphocytic leukemia
Goals: Perform targeted sequencing on a panel of ~300 recurrently mutated genes in genomic DNA from diagnostic CLL patient samples from E2997 clinical trial. Identification of samples with mutations activating MAP kinas signaling and NOTCH signaling will allow us to perform further functional analysis of the therapeutic and biological relevance of these 2 pathways in CLL.
Agency contact: Mary LaRocca, 111 East 210th Street, Bronx, NY 10467, mlarocca@montefiore.org
Overlap: None

5 K08CA160647-05 (Abdel-Wahab) 9/20/11 – 8/31/16
NIH/NCI $157,300
Clinical Scientist Research Career Development Award "Role of ASXL1 mutations in myeloid malignancies." Goals and Specific Aims: Aim #1: Identify target genes of ASXL1 loss and alterations in chromatin state due to ASXL1 loss in normal and leukemic hematopoietic cells using in vitro assays. Aim #2: Determine the in vivo effects of ASXL loss1 on hematopoiesis by targeted disruption of ASXL1. Aim #3: Determine the role of ASXL1 mutations in the pathogenesis of MPN and AML through studies in patient samples and in in vivo models in cooperation with known MPN and AML disease alleles.
Agency Contact: Esther Young, youngel@mail.nih.gov, 240-276-6325
Overlap: None

00003930 (PI: Janku) 1/1/16 - 12/31/16
Erdheim-Chester Disease Global Alliance $ 24,631
Understanding and targeting novel molecular alterations in Erdheim-Chester disease without the BRAFV600E mutation
Goals: We plan to extend our effort in furthering the treatment and understanding of histiocytic disorders through preclinical and clinical correlative studies to identify and target novel kinase alterations identified in BRAFV600E-wildtype ECD patients. These studies will be critical for (i) establishing diagnostic approaches to identifying the diverse, clinically actionable genomic alterations in ECD and (ii) promoting novel therapeutic studies for ECD patients.

Specific Aims: 1) Perform genomic analysis of Erdheim Chester Disease (ECD) tissue biopsy samples. 2) Functionally characterize the novel kinase alterations recently identified in systemic histiocytic disorder. 3) Determine the sensitivity of kinase mutations present in BRAFV600E-wildtype ECD to targeted therapeutics in preclinical models.
Agency contact: Jaime Farias, Assistant Director, OSP, 1515 Holcombe Boulevard, Unit 1676, Houston, TX 77030 (P) 713-792-3220
Overlap: None

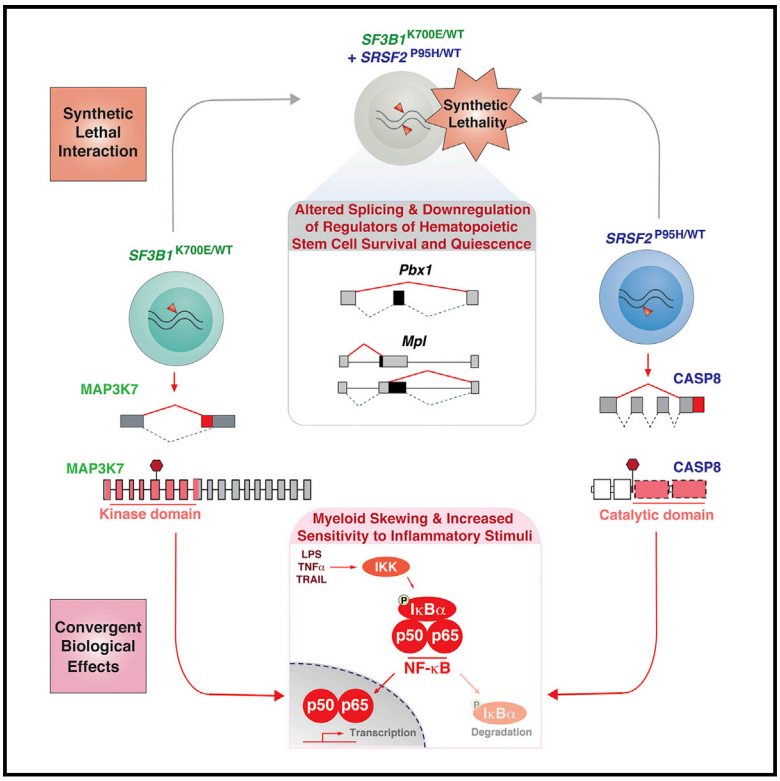
GC227311 (PI: Abdel Wahab) 1/1/16-12/31/16
Hairy Cell Leukemia Foundation $ 58,000
HCL Grant Proposal 2015
Goals: To identify the spectrum of recurrent genomic alterations co-existing with BRAFV600E mutations in HCL and delineating their clnical relevance and functional contribution to HCL development
Specific Aims: Aim I: Identify genetic alterations co-existing with the BRAFV600E mutation in HCL patients and their clinical importance. Aim II: Identify the functional importance of recurrent mutations coexisting with the BRAFV600E mutations in HCL pathogenesis.
Agency contact: Michael Margolin, MD. President, 575 Madison Avenue, 10th Floor, New York, NY 10022 Overlap: None

Article

Cancer Cell

Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations

Graphical Abstract



Highlights

- Mutations in *SF3B1* and *SRSF2* have a synthetic lethal interaction
- Mutations in RNA splicing factors are not tolerated in a homozygous state
- Mutations in *SF3B1* and *SRSF2* have distinct effects on pre-mRNA splicing
- Both *SF3B1* and *SRSF2* mutations result in hyperactive NF-κB signaling

Authors

Stanley Chun-Wei Lee, Khrystyna North, Eunhee Kim, ..., Benjamin L. Ebert, Robert K. Bradley, Omar Abdel-Wahab

Correspondence

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

In Brief

Lee et al. report that *SF3B1* and *SRSF2* mutations elicit distinct effects on splicing and are synthetically lethal due to the cumulative impact on hematopoietic stem cell survival and quiescence. These mutations share convergent effects on promoting NF-κB signaling to drive myelodysplastic syndrome.

Lee et al., 2018, Cancer Cell *34*, 225–241
August 13, 2018 © 2018 Elsevier Inc.
https://doi.org/10.1016/j.ccell.2018.07.003

CellPress

Cancer Cell
Article

Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations

Stanley Chun-Wei Lee,[1,11] Khrystyna North,[2,3,4,11] Eunhee Kim,[5,11] Eunjung Jang,[5] Esther Obeng,[6,10] Sydney X. Lu,[1] Bo Liu,[1] Daichi Inoue,[1] Akihide Yoshimi,[1] Michelle Ki,[1] Mirae Yeo,[5] Xiao Jing Zhang,[1] Min Kyung Kim,[1] Hana Cho,[1] Young Rock Chung,[1] Justin Taylor,[1] Benjamin H. Durham,[1] Young Joon Kim,[1] Alessandro Pastore,[1] Sebastien Monette,[7] James Palacino,[8] Michael Seiler,[8] Silvia Buonamici,[8] Peter G. Smith,[8] Benjamin L. Ebert,[6,10] Robert K. Bradley,[2,3,4,11,*] and Omar Abdel-Wahab[1,9,11,12,*]

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, Zuckerman 701, 408 East 69th Street, New York, NY 10065, USA
[2]Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Mailstop: M1-B514, Seattle, WA 98109-1024, USA
[3]Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[4]Department of Genome Sciences, University of Washington, Seattle, WA, USA
[5]School of Life Sciences, Ulsan National Institute of Science and Technology, Ulsan, Republic of Korea
[6]Division of Hematology, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA
[7]Laboratory of Comparative Pathology, Memorial Sloan Kettering Cancer Center, Weill Cornell Medicine, The Rockefeller University, New York, NY, USA
[8]H3 Biomedicine, Cambridge, MA, USA
[9]Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA
[10]Broad Institute of MIT and Harvard, Cambridge, MA, USA
[11]These authors contributed equally
[12]Lead Contact
*Correspondence: rbradley@fredhutch.org (R.K.B.), abdelwao@mskcc.org (O.A.-W.)
https://doi.org/10.1016/j.ccell.2018.07.003

SUMMARY

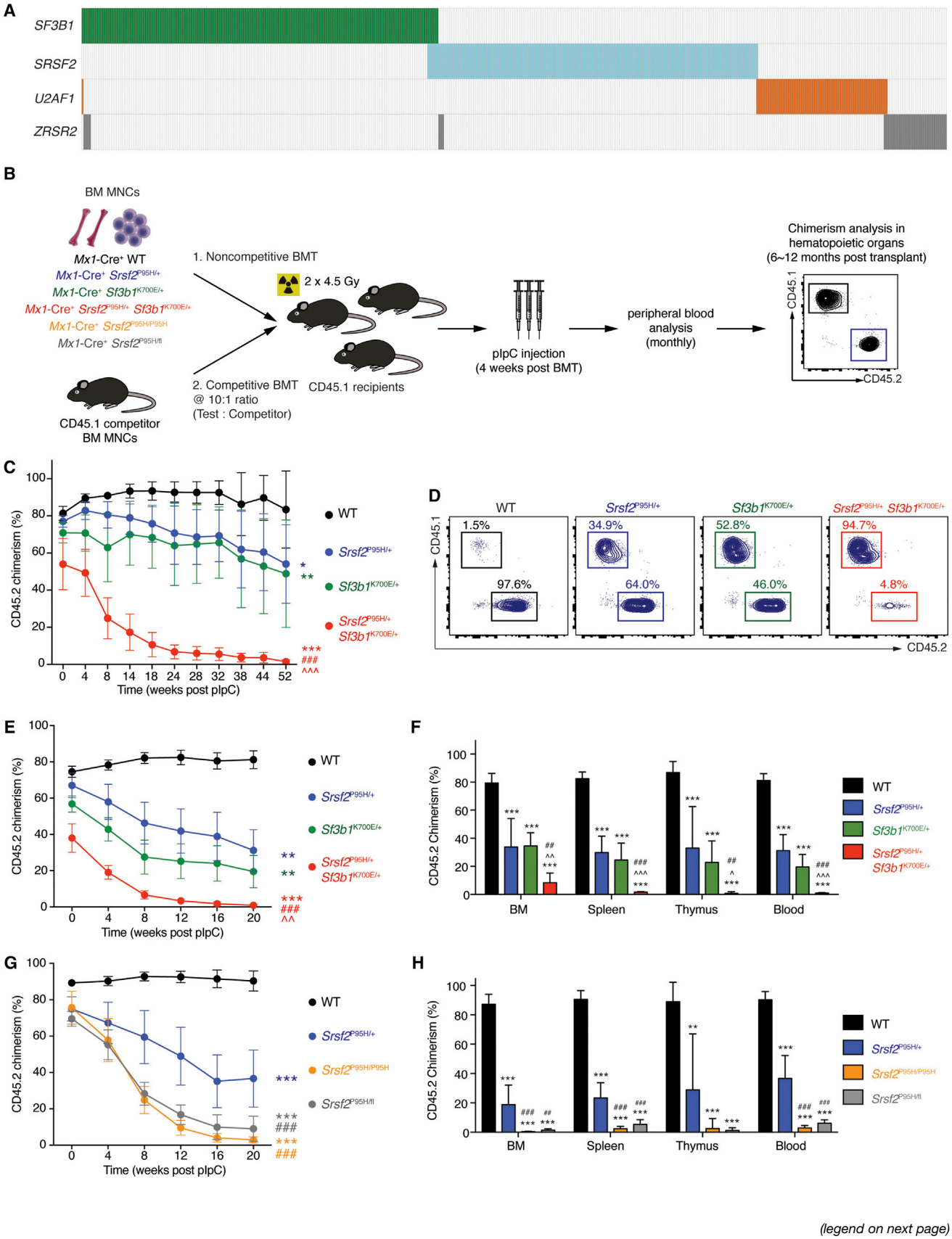
Mutations affecting RNA splicing factors are the most common genetic alterations in myelodysplastic syndrome (MDS) patients and occur in a mutually exclusive manner. The basis for the mutual exclusivity of these mutations and how they contribute to MDS is not well understood. Here we report that although different spliceosome gene mutations impart distinct effects on splicing, they are negatively selected for when co-expressed due to aberrant splicing and downregulation of regulators of hematopoietic stem cell survival and quiescence. In addition to this synthetic lethal interaction, mutations in the splicing factors SF3B1 and SRSF2 share convergent effects on aberrant splicing of mRNAs that promote nuclear factor κB signaling. These data identify shared consequences of splicing-factor mutations and the basis for their mutual exclusivity.

INTRODUCTION

Recurrent mutations in genes encoding spliceosome components have been identified across multiple cancer types. Spliceosomal gene mutations most commonly occur in myelodysplastic syndromes (MDS), chronic myelomonocytic leukemia (CMML), acute myeloid leukemia (AML) (Papaemmanuil et al., 2011; Yoshida et al., 2011), chronic lymphocytic leukemia (CLL) (Wang et al., 2011), and uveal melanoma (Harbour et al., 2013; Martin et al., 2013). *SF3B1*, *U2AF1*, and *SRSF2* are the

Significance

Mutual exclusivity of different mutations that affect a single pathway in cancer is commonly thought to indicate convergent effects of these mutations. RNA splicing-factor mutations constitute the most common class of alterations in patients with myelodysplastic syndromes and are also frequent in several additional cancer types, and occur as heterozygous mutations at restricted residues in a mutually exclusive manner. The mutual exclusivity of spliceosomal mutations suggests synthetic lethal and/or convergent biological effects of these mutations; however, there is currently no functional evidence supporting either of these possibilities. Here we report that individual spliceosomal mutations have non-overlapping effects on splicing and are mutually exclusive due to both synthetic lethal interactions and convergent effects on hyperactivation of innate immune signaling.



(legend on next page)

most commonly mutated genes. Mutations in each of these occur as heterozygous point mutations at specific residues, suggesting gain-of-function alterations. In addition, mutations in spliceosome genes are mutually exclusive with one another, presumably due to redundant effects and/or a limit on cellular tolerance of disrupted spliceosome function.

Despite insights into the effects of each mutation on pre-mRNA splicing, the basis for their mutual exclusivity, and functionally convergent effects they may have are unknown. For example, recent work identified that mutations affecting the core splicing factor SF3B1 are associated with cryptic 3′ splice site selection and altered branchpoint recognition (Darman et al., 2015; DeBoever et al., 2015). In contrast, mutations affecting SRSF2, an auxiliary splicing factor that binds exonic splicing enhancers to promote splicing, alter its RNA binding preference in a sequence-specific manner and thereby alter the efficiency of exon inclusion (Kim et al., 2015; Zhang et al., 2015). Finally, mutations affecting U2AF1 either promote or repress 3′ splice site based on sequences flanking the AG dinucleotide (Ilagan et al., 2014). Given that the effects of *SF3B1*, *U2AF1*, and *SRSF2* mutations on splicing mechanisms are distinct, it is unclear why these mutations are mutually exclusive. Moreover, evaluation of the effects of these mutations on splicing and gene expression in an isogenic manner has not been performed. Here we directly assessed the functional basis for the mutual exclusivity of spliceosome gene mutations using models expressing single- or double-mutant splicing factors simultaneously. These data provide a functional explanation for the genetic configuration of spliceosome gene mutations and identify a convergent effect of these mutations on a pathway of established relevance to MDS pathogenesis.

RESULTS

Simultaneous Expression of *Srsf2* and *Sf3b1* Mutations Is Incompatible with Hematopoiesis

Evaluation of sequencing data from >4,000 patients with myeloid neoplasms revealed that while ~48% (1,935/4,032) have a mutation in an RNA splicing factor, only ~2% of patients (86/4,032) have >1 splicing-factor mutation (Figure 1A and Table S1). To understand the basis of this mutual exclusivity, we generated mice for inducible heterozygous expression of two of the most common splicing-factor mutations in MDS (*SF3B1*^{K700E} and *SRSF2*^{P95H}) simultaneously (Figure 1B). We performed noncompetitive bone marrow transplantation (BMT), whereby each mutation was induced following stable engraftment in recipient mice (Figure 1B). Bone marrow (BM) cells co-expressing *Srsf2*^{P95H} and *Sf3b1*^{K700E} mutations were severely defective in multi-lineage reconstitution compared with other groups (Figures 1C, 1D, and S1A–S1D). Chimerism analysis and evaluation for recombination prior to polyinosinic-polycytidylic acid (pIpC) administration confirmed minimal spontaneous excision prior to BMT (Figures S1E and S1F). In competitive BMT (Figure 1E), BM cells co-expressing *Sf3b1*^{K700E} and *Srsf2*^{P95H} mutations were more readily outcompeted by competitor BM cells relative to single-mutant or wild-type (WT) controls. Analyses of hematopoietic organs 6 months post BMT revealed a near-complete absence of *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells, which was distinct from single-mutant groups (Figures 1E, 1F, and S1C). These data provide functional evidence that co-expression of mutant RNA splicing factors is not tolerated.

Expression of Splicing-Factor Mutations in a Homozygous or Hemizygous State Is Incompatible with Hematopoiesis

Prior studies identified that cells bearing mutant splicing factors require the WT allele for survival (Lee et al., 2016; Zhou et al., 2015). While these observations potentially explain the heterozygous nature of splicing-factor mutations in patients, the effect of expressing splicing-factor mutation in a homozygous manner was not assessed. To test this, we generated mice with conditional homozygous expression of the *SRSF2*^{P95H} mutation (*Mx1*-Cre^{+} *Srsf2*$^{P95H/P95H}$). In competitive BMT assays, hematopoietic stem and progenitor cells (HSPCs) from these mice showed severe defects in multi-lineage reconstitution relative to *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ or *Mx1*-Cre^{+} WT HSPCs, similar to defects seen with hemizygous *Srsf2*^{P95H} expression (*Mx1*-Cre^{+} *Srsf2*$^{P95H/fl}$) (Figures 1G, 1H, and S1G–S1I). These data firmly establish that neither hemizygous nor homozygous expression of splicing-factor mutations is tolerated, highlighting the unique dependency of spliceosome-mutant cells on residual WT spliceosome function.

Figure 1. Simultaneous Expression of Mutations in *Srsf2* and *Sf3b1* or Expression in the Homozygous State Is Incompatible with Hematopoiesis

(A) Heatmap of the four most commonly mutated genes encoding RNA splicing factors across 11 studies in myeloid malignancies (Bejar et al., 2012; Damm et al., 2012; Haferlach et al., 2014; Lasho et al., 2012; Makishima et al., 2012; Meggendorfer et al., 2012; Papaemmanuil et al., 2013; Patnaik et al., 2013; Thol et al., 2012; Yoshida et al., 2011; Zhang et al., 2012). Each column represents a patient, and each colored bar represents the presence of the specified mutation.

(B) Schema of competitive and noncompetitive bone marrow transplantation (BMT) using bone marrow mononuclear cells (BM MNCs) from 8-week-old *Mx1*-Cre^{+} wild-type (WT), *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/P95H}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/fl}$ mice. Polyinosinic-polycytidylic acid (pIpC) was administered to recipients 4 weeks post BMT to induce expression of mutant alleles.

(C) Percentage of CD45.2 chimerism in peripheral blood of recipients (n = 8–10 mice per genotype) in noncompetitive BMT.

(D) Representative fluorescence-activated cell sorting (FACS) plots of CD45.2^{+} cells in peripheral blood of recipients in noncompetitive BMT 52 weeks post pIpC.

(E) Percentage of CD45.2 chimerism in peripheral blood of recipients (n = 10 mice per genotype) in competitive BMT.

(F) Analysis of CD45.2 chimerism in the BM, spleen, thymus, and blood of recipients (n = 5–10 mice per genotype) in competitive BMT 20 weeks post pIpC.

(G) Percentage of CD45.2 chimerism in peripheral blood of recipients (n = 5–10 mice per genotype) in competitive BMT from *Mx1*-Cre^{+} WT, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/P95H}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/fl}$ mice.

(H) Analysis of CD45.2 chimerism in BM, spleen, thymus, and blood of recipients (n = 5–10 mice per genotype) in competitive BMT 20 weeks post pIpC.

Error bars represent mean ± SD. ANOVA and Tukey's post hoc test were used to compare groups. *p < 0.05, **p < 0.01, ***p < 0.001 versus *Mx1*-Cre^{+} WT mice; ^p < 0.05, ^^p < 0.01, ^^^p < 0.001 versus *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$ mice; ##p < 0.01, ###p < 0.001 versus *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ mice. See also Figure S1 and Table S1.

Simultaneous Expression of *Srsf2* and *Sf3b1* Mutations Resulted in Increased Apoptosis and Reduced Quiescence of HSPCs

To identify why co-expression of *Srsf2*^{P95H} and *Sf3b1*^{K700E} resulted in severe HSPC dysfunction, we performed cell-cycle and apoptosis analysis. Two-weeks following recombination, there was a significant increase in lineage$^-$ Sca-1$^+$ c-Kit$^+$ (LSK) cells undergoing apoptosis and in the S phase of cell cycle in *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ mice relative to controls (Figures 2A and 2B).

Selection against co-expression of two mutant splicing factors was strongly supported by the fact that mice with hematopoietic-specific expression of *Srsf2*^{P95H} and *Sf3b1*^{K700E} (*Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$) mutations were associated with 100% lethality at weaning, while single-mutant mice were present at the expected frequencies (Figure 2C). Analysis of 139 embryos at embryonic day 14.5 (E14.5) revealed that *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ fetuses were detectable at the expected frequency (Figures S2A and S2B). E14.5 *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ fetal livers had increased number of LSK cells as well as phenotypic long-term hematopoietic stem cell (LT-HSC, LSK CD150$^+$ CD48$^-$), multi-potent progenitor (MPP, LSK CD150$^-$ CD48$^-$), HPC-1 (LSK CD150$^-$ CD48$^+$), and HPC-2 (LSK CD150$^+$ CD48$^+$) populations compared with all other genotypes (Figures 2D, 2E, and S2C–S2G). Despite a significant increase in HSPCs, *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ fetal liver cells had significantly impaired colony-forming ability *in vitro* (Figure 2F) and markedly increased percentage of cycling and apoptotic cells (Figures 2G–2I). These derangements in HSPC phenotypes were manifest by E18.5 when *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ fetuses had near absence of hematopoietic cells in the BM as well as increased apoptotic cells in fetal liver (Figures S2H and S2I). Taken together, these data demonstrate that co-expression of splicing-factor mutations compromises hematopoiesis, driven by aberrant cell-cycle progression and increased apoptosis of HSPCs.

Srsf2 and *Sf3b1* Mutations Have Largely Distinct Effects on Gene Expression

To understand the mechanistic basis for mutual exclusivity of *SF3B1* and *SRSF2* mutations, we performed RNA sequencing (RNA-seq) on lineage$^-$ c-Kit$^+$ (LK) cells from *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ mice as well as single-mutant and WT controls (Figure 3A). The mean allelic ratio of *Sf3b1*^{K700E} and *Srsf2*^{P95H} expressed in double-mutant cells was 20.7% and 33.5%, markedly lower than the ~50% expression in single-mutant controls, illustrating the intolerability of combining *SF3B1* and *SRSF2* mutations (Figure 3B). Expression of either mutation resulted in dysregulation of hundreds of coding genes (Figure 3C and Table S2). *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells exhibited significantly more gene dysregulation than single-mutant cells, consistent with their more dramatic biological phenotype (Figure 3D). Double-mutant cells shared a greater proportion of differentially expressed genes with each single-mutant group than single mutants shared with one another (Figure 3E). Gene ontology (GO) analysis of coding genes differentially expressed within each genotype relative to WT control revealed strong signatures of impaired hematopoiesis in double-mutant cells, consistent with the multi-lineage defects in these cells (Figure 3F and Table S3).

We next tested whether mutations in SF3B1 and SRSF2 have independent effects on gene expression: that is, whether double-mutant cells recapitulated the gene dysregulation observed for each single-mutant genotype relative to WT. Approximately 80% and 40% of genes dysregulated in *Mx1*-Cre$^+$ *Sf3b1*$^{K700E/+}$ and *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ single-mutant cells were also dysregulated in *Mx1*-Cre *Sf3b1*$^{K700E/+}$ *Srsf2*$^{P95H/+}$ double-mutant cells (Figure 3E and Table S2). That degree of recapitulation of gene dysregulation was highly similar to and statistically indistinguishable ($p = 0.90$ and 0.44 for *Sf3b1*^{K700E} and *Srsf2*^{P95H} mutations, respectively) from expected recapitulation under the assumption of independence (Figure S3A). We therefore conclude that SF3B1 and SRSF2 mutations have independent effects on gene expression even when present in the same cell. Interestingly, in addition to recapitulating the dysregulation of specific genes expected based on single-mutant cells, double-mutant cells exhibited additional gene dysregulation (Figures S3B and S3C), consistent with the severe hematopoietic phenotype of these cells. These results suggest that SF3B1 and SRSF2 mutations have independent but compound effects on gene expression when present in the same cell.

Srsf2 and *Sf3b1* Mutations Have Distinct and Independent Effects on Splicing

We next assessed the consequences of single versus double mutations in *Sf3b1* and *Srsf2* on splicing (Table S4). Mutations in SF3B1 have been proposed to alter 3′ splice site recognition (Darman et al., 2015; DeBoever et al., 2015). In contrast, mutations in SRSF2 alter exon recognition via preferential recognition of C-rich exonic splicing enhancer (ESE) motifs relative to G-rich ESEs, while WT SRSF2 recognizes both classes of ESEs (Kim et al., 2015; Zhang et al., 2015). Our current understanding of SF3B1 and SRSF2 mutations therefore indicates that they induce distinct changes in splicing. However, the consequences of these mutations on splicing have not been compared in an isogenic context.

We used our double-mutant system to directly compare the effects of SF3B1 and SRSF2 mutations on RNA splicing. Consistent with prior studies, *Sf3b1* and *Srsf2* mutations affected 3′ splice site and exon recognition, respectively (Figures 4A, 4B, S4A, and S4B). We observed a modest enrichment of adenosines upstream of intron-proximal 3′ splice site promoted by the *Sf3b1*^{K700E} mutation, independent of the *Srsf2*^{P95H} mutation (Figure S4C). This enrichment was absent from cells expressing *Srsf2*^{P95H} alone. Cassette exons promoted versus repressed in cells expressing *Srsf2*^{P95H} were respectively enriched for CCNG and GGNG ESEs, independent of the presence or absence of *Sf3b1*^{K700E}, but were absent from cells expressing *Sf3b1*^{K700E} alone (Figures 4C and S4D). These analyses are consistent with previous reports that *SF3B1* and *SRSF2* mutations affect branchpoint recognition and ESE preference, respectively, and further demonstrate that these changes in splicing are specific to *SF3B1* and *SRSF2* mutations.

As *Sf3b1* and *Srsf2* mutations cause distinct changes in splicing, even when simultaneously present, we hypothesized that the effects of these mutations on splicing were independent. Consistent with this hypothesis, double-mutant cells

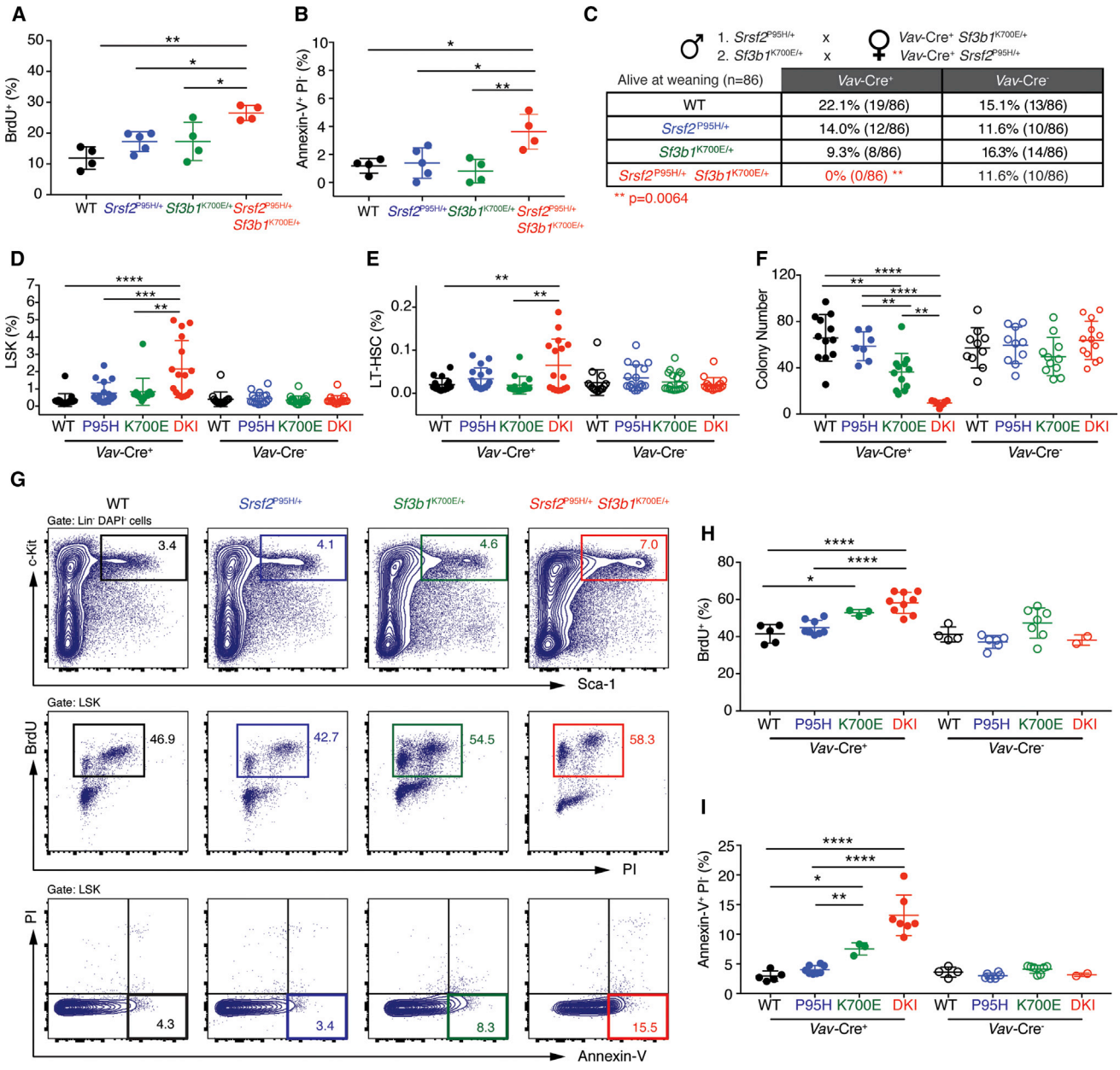


Figure 2. Combined Expression of Mutations in *Srsf2* and *Sf3b1* Results in Hematopoietic Stem and Progenitor Cell Apoptosis and Loss of Quiescence

(A and B) Percentage of bromodeoxyuridine$^+$ (BrdU$^+$) (A) or annexin-V$^+$ propidium iodide$^-$ (PI$^-$) (B) LSK cells from *Mx1*-Cre$^+$ WT (n = 4), *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ (n = 5), *Mx1*-Cre$^+$ *Sf3b1*$^{K700E/+}$ (n = 4), and *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ (n = 4) mice 2 weeks post plpC administration.

(C) Number of live mice at weaning from crossing *Sf3b1*$^{K700E/+}$ mice to *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ mice or by crossing *Srsf2*$^{P95H/+}$ mice to *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ mice. **p = 0.0064; two-sided Chi-square test.

(D and E) Percentage of LSK (D) and long-term hematopoietic stem cells (LT-HSC; LSK CD150$^+$ CD48$^-$) (E) in E14.5 fetal livers from *Vav*-Cre$^+$ WT (WT; n = 17), *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ (P95H; n = 17), *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ (K700E; n = 15), and *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ double-knockin (DKI; n = 16) fetuses and *Vav*-Cre$^-$ WT (n = 14), *Vav*-Cre$^-$ *Srsf2*$^{P95H/+}$ (P95H; n = 17), *Vav*-Cre$^-$ *Sf3b1*$^{K700E/+}$ (K700E; n = 19), and *Vav*-Cre$^-$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ DKI (n = 17) fetuses.

(F) Colony numbers from E14.5 fetal liver cells from *Vav*-Cre$^+$ WT (n = 12), *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ (P95H; n = 7), *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ (K700E; n = 12), and *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ DKI (n = 7) fetuses and from *Vav*-Cre$^-$ WT (n = 10), *Vav*-Cre$^-$ *Srsf2*$^{P95H/+}$ (P95H; n = 10), *Vav*-Cre$^-$ *Sf3b1*$^{K700E/+}$ (K700E; n = 11), and *Vav*-Cre$^-$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ DKI (n = 13) fetuses.

(G–I) Representative FACS plots (G) and quantitation of BrdU$^+$ (H) and annexin-V$^+$ PI$^-$ (I) LSK cells from *Vav*-Cre$^+$ WT (n = 5), *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ (P95H; n = 8), *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ (K700E; n = 3), and *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ DKI (n = 9) fetuses and from *Vav*-Cre$^-$ WT (n = 4), *Vav*-Cre$^-$ *Srsf2*$^{P95H/+}$ (P95H; n = 6), *Vav*-Cre$^-$ *Sf3b1*$^{K700E/+}$ (K700E; n = 7), and *Vav*-Cre$^-$ *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ DKI (n = 2) fetuses.

Error bars represent mean ± SD. ANOVA and Tukey's post hoc test was used to compare groups. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001. See also Figure S2.

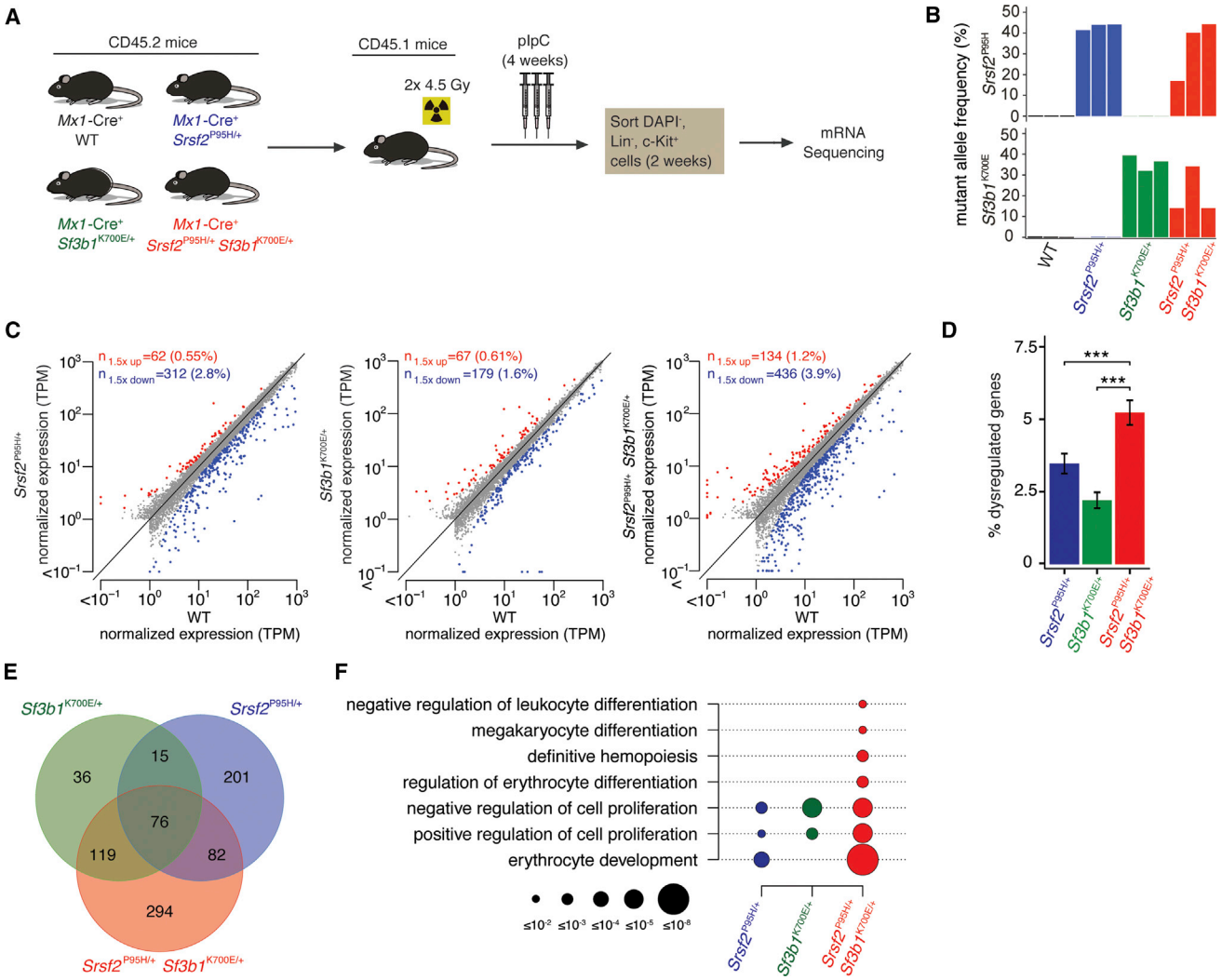


Figure 3. *Srsf2* and *Sf3b1* Mutations Have Distinct and Independent Effects on Gene Dysregulation

(A) Schema for RNA-seq.

(B) Expression of *Srsf2*^{P95H} and *Sf3b1*^{K700E} alleles as percentage of mRNAs expressed from *Srsf2* and *Sf3b1*. Color indicates genotype; the three biological replicates are A to C from left to right.

(C) Scatterplots comparing coding gene expression in *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells for replicate B. Red and blue indicate coding genes significantly up- or downregulated, respectively, in mutant relative to *Mx1*-Cre^{+} WT cells. TPM, transcripts per million (TMM-normalized).

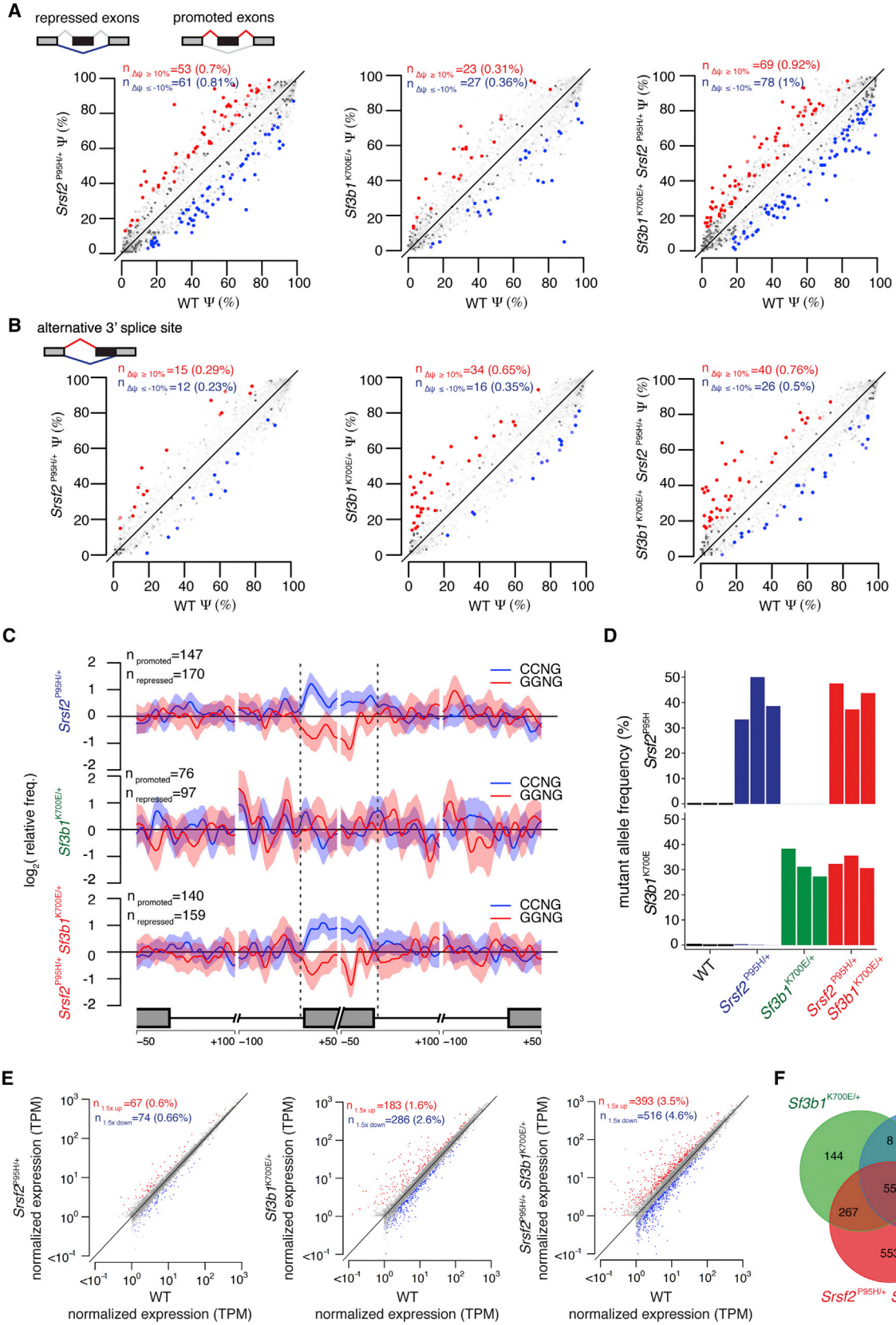
(D) Bar plots comparing percentage of significantly dysregulated coding genes (percentage of total coding genes expressed) in *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells for replicate B. Error bars represent mean ± SD. A two-sided binomial proportion test was used to compare groups; ***$p < 0.0001$.

(E) Venn diagram showing overlap between coding genes significantly dysregulated in *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells for replicate B.

(F) GO enrichment analysis of *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells for replicate B. Circle size indicates the magnitude of the p value for each term and comparison.

See also Figure S3; Tables S2 and S3.

recapitulated the preferential effects of *Sf3b1* and *Srsf2* mutations on 3′ splice site recognition and cassette exon recognition, respectively, that we observed in single-mutant cells (Figure S4A). We did not observe widespread decreases in splicing efficiency, such as retention of constitutive introns, in double-mutant cells (Figure S4E). Instead, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells exhibited modest increases in mis-spliced genes relative to cells bearing single mutations in *Sf3b1* or *Srsf2* (Figures 4A and 4B). *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$ *Srsf2*$^{P95H/+}$ cells recapitulated 40% and 32% of splicing dysregulation driven by expression of the single mutations in *Sf3b1*^{K700E} and *Srsf2*^{P95H}, respectively (Figure S4F). We identified only seven genes mis-spliced in both *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ and *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$ single-mutant as well as double-mutant cells, and



(legend on next page)

only six genes mis-spliced in both *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ and *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$ single-mutant but not double-mutant cells. We therefore conclude that SF3B1 and SRSF2 mutations typically have distinct and independent consequences on splicing.

To confirm these findings, we performed RNA-seq on purified LK cells from E14.5 *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ fetal liver and controls. The allelic frequencies of both mutations in double-mutant cells were near 50% and comparable with those of *Vav*-Cre^{+} single-mutant controls (Figure 4D). Similar to the *Mx1*-Cre system, the greatest alterations in gene expression were seen in *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ HSPCs compared with single-mutant controls, and double-mutant cells recapitulated much of the gene expression dysregulation in single-mutant cells (Figures 4E and 4F). GO analysis revealed signatures of impaired hematopoiesis in double-mutant cells, consistent with those seen in *Mx1*-Cre mice (Figure S4G and Table S5). Moreover, effects on splicing similar to those seen in the *Mx1*-Cre system were seen in HSPCs from *Vav*-Cre mice, including preferential effects of *Sf3b1* and *Srsf2* mutations on 3′ splice site and cassette exon recognition, respectively (Figures S4H and S4I).

Combined Expression of *Srsf2* and *Sf3b1* Mutations Impairs Expression of Regulators of HSPC Survival and Increases Sensitivity to Inflammatory Stimulation

Given the largely non-overlapping effects of *Srsf2*^{P95H} and *Sf3b1*^{K700E} mutations on gene expression and splicing, we evaluated transcripts that exhibited concomitant dysregulation in gene expression and splicing in the double-mutant state (Figure 5A and Table S6). Multiple regulators of HSPC survival and quiescence were significantly dysregulated and mis-spliced in *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ HSPCs, including the thrombopoietin receptor c-Mpl, integrin αIIb (CD41), and the transcription factor Pbx1 (Figures 5A and S5A). After confirming reduced expression of these transcripts in LK cells in an independent cohort of *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ mice (Figure 5B), we evaluated the consequences of short hairpin RNA (shRNA)-mediated depletion of these mRNAs in WT HSPCs (Figures S5B and S5C). Consistent with prior published data (Alexander et al., 1996; Ficara et al., 2008; Gekas and Graf, 2013; Qian et al., 2007; Yoshihara et al., 2007), silencing any of these individual genes was associated with severe impairment of HSPC clonogenicity (Figure S5D).

In addition to gene expression changes consistent with impaired hematopoiesis in *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells (Figure 3F), GO analysis also revealed a strong signature associated with immune signaling that was augmented in double- relative to single-mutant cells (Figures 5C and S4G). To functionally evaluate these gene expression changes, we studied the response of *Srsf2* and *Sf3b1* mutant cells to lipopolysaccharide (LPS). HSPCs from *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ and *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$ mice exhibited increased baseline nuclear factor κB (NF-κB) activation (marked by increased nuclear accumulation of phosphorylated p65 [p-p65]) relative to *Vav*-Cre^{+} WT HSPCs (Figures 5D and 5E), and this was further enhanced following *ex vivo* LPS stimulation in double-mutant cells relative to single-mutant or WT cells (Figures S5E and S5F).

Given prior data identifying that chronic LPS exposure impairs repopulating potential of HSPCs (Esplin et al., 2011; Zhao et al., 2013), we next evaluated the effects of chronic inflammation on the function of *Srsf2* or *Sf3b1* mutant BM HSPCs (Figure 5F). *Vav*-Cre^{+} WT, *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$, and *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$ mice were treated with LPS (1 mg/kg) every second day for 30 days followed by serial BMT into lethally irradiated recipient mice. Chronic LPS treatment had a mild effect on the repopulating potential of WT HSPCs, evidenced by similar peripheral blood and BM HSPC chimerism in primary recipients (Figures 5G and 5H). In contrast, LPS-treated *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ and *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$ BM HSPCs showed significant reduction in repopulating ability relative to vehicle-treated counterparts, and this defect was further exacerbated following serial transplantation (Figures 5G and 5H).

The above observations suggest that *Srsf2* and *Sf3b1* mutant HSPCs are intrinsically hypersensitive to inflammatory stimuli that contribute to defective hematopoietic function, rendering features observed in MDS patients. In addition, LPS-treated *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ mice had enhanced myeloid skewing and reduced B lymphopoiesis relative to vehicle-treated mice or *Vav*-Cre^{+} *Srsf2*$^{+/+}$ mice treated with LPS (Figure S5G). Moreover, acute *ex vivo* LPS stimulation of BM LSK cells from *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ mice also resulted in significant increase in nuclear p-p65 relative to *Vav*-Cre^{+} *Srsf2*$^{+/+}$ LSK cells (Figures S5H and S5I). The enhanced response to inflammatory stimuli in *Srsf2*

Figure 4. *Srsf2* and *Sf3b1* Mutations Have Distinct and Independent Effects on RNA Splicing

(A) Scatterplots of cassette exon inclusion in *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells. Axes indicate the fraction of mRNAs containing each cassette exon in the indicated sample. Red and blue indicate cassette exons whose inclusion is promoted or repressed, respectively, in mutant relative to WT cells.

(B) As in (A), but for alternative 3′ splice site events. Axes indicate the fraction of mRNAs that use the intron-proximal 3′ splice site in the indicated sample. Red and blue indicate intron-proximal 3′ splice sites whose usage is promoted or repressed, respectively, in mutant relative to WT cells.

(C) Plots illustrating the spatial distribution of the CCNG and GGNG (N = any nucleotide) exonic splicing enhancers adjacent to differentially spliced cassette exons promoted versus repressed in mutant relative to WT cells. Vertical axis indicates the relative frequency of each motif, averaged over all promoted versus repressed cassette exons for the indicated genotype comparisons.

(D) Expression of *Srsf2*^{P95H} and *Sf3b1*^{K700E} alleles as percentage of mRNAs expressed from *Srsf2* and *Sf3b1* in lineage^{-} c-Kit^{+} cells from fetal livers of *Vav*-Cre mice at E14.5. Color indicates genotype; the three biological replicates are A to C from left to right.

(E) Scatterplots comparing mean coding gene expression for all replicates in *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$, *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Vav*-Cre^{+} WT cells. Red and blue indicate coding genes significantly up- or downregulated, respectively, in mutant relative to *Vav*-Cre^{+} WT cells. TPM, transcripts per million (TMM-normalized).

(F) Venn diagram showing the overlap between coding genes significantly dysregulated in *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$, *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Vav*-Cre^{+} WT cells in all replicates.

See also Figure S4; Tables S4 and S5.

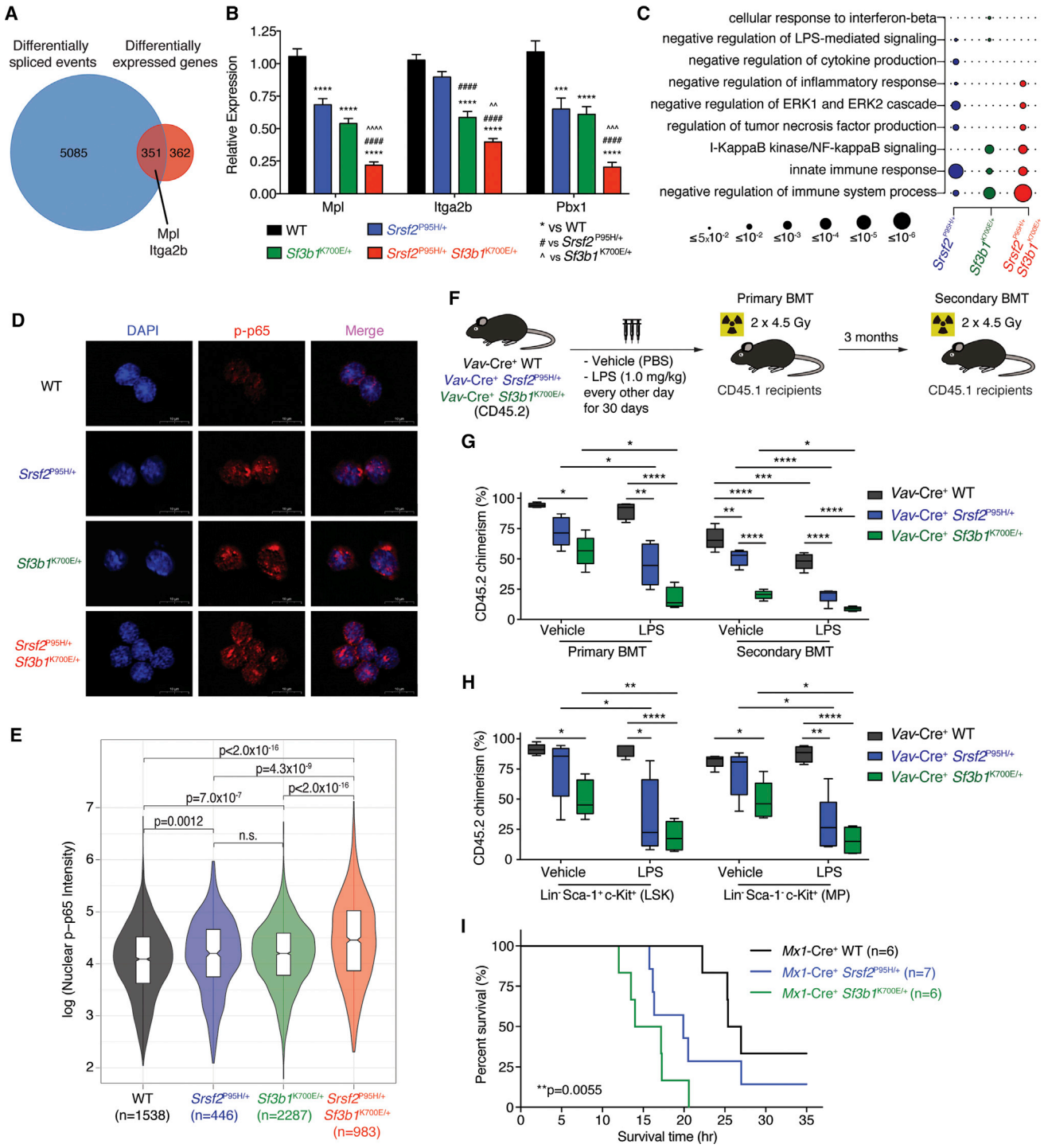


Figure 5. Co-expression of *Srsf2* and *Sf3b1* Mutations Results in Aberrant Splicing and Expression of Key Regulators of Hematopoietic Stem Cell Survival and Quiescence

(A) Venn diagram of genes differentially expressed and spliced in *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells in any replicate.

(B) Expression of Mpl, Itga2b, and Pbx1 in lineage^{-} c-Kit^{+} (LK) cells from *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ mice relative to control groups (n = 8–10 mice per genotype). Error bars represent mean ± SD. ***p < 0.001, ****p < 0.0001 versus *Mx1*-Cre^{+} WT; $^{\#\#\#\#}$p < 0.0001 versus *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$; $^{\wedge\wedge}$p < 0.01, $^{\wedge\wedge\wedge}$p < 0.001, $^{\wedge\wedge\wedge\wedge}$p < 0.0001 versus *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$.

(C) GO enrichment analysis of *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ cells relative to *Mx1*-Cre^{+} WT cells. Circle size indicates the magnitude of the p value for each term and comparison.

(D) Immunofluorescence of nuclear phosphorylated-p65 (p-p65) level in LK cells from *Vav*-Cre^{+} WT, *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$, *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$, and *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ mice following LPS stimulation *ex vivo*. Scale bars, 10 μm.

(legend continued on next page)

and *Sf3b1* mutant mice *in vivo* was evident upon LPS-induced sepsis. Exposure of primary *Mx1*-Cre$^+$ WT, *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$, and *Mx1*-Cre$^+$ *Sf3b1*$^{K700E/+}$ mice with LPS (15 mg/kg) resulted in accelerated death in both *Srsf2*- and *Sf3b1*-mutant mice relative to WT controls (Figure 5I).

SF3B1 Mutations Promote Mis-splicing of MAP3K7, Resulting in Hyperactivation of NF-κB Signaling

The above results identify that spliceosomal mutations are intolerable when co-expressed but they are independently more sensitive to inflammatory stimuli that converge on NF-κB signaling. To understand how spliceosome gene mutations activate immune signaling pathways, we first focused on events that are mis-spliced within SF3B1-mutant hematopoietic cells in both human and mouse (Table S7). Although recent studies reported few such shared mis-spliced events (Mupo et al., 2016; Obeng et al., 2016), we identified a larger set of 205 transcripts mis-spliced in both *Sf3b1*-mutant murine hematopoietic cells and *SF3B1*-mutant MDS patients (Figure 6A). One of the most robust mis-spliced events was MAP3K7, for which an intron-proximal 3′ splice site was promoted by mutant SF3B1 in both human and mouse hematopoietic cells. Reanalysis of RNA-seq data from CLL patient cohorts (DeBoever et al., 2015) and isogenic cell lines with endogenous mutations in *SF3B1*, as well as RT-PCR and Sanger sequencing of cDNA from *SF3B1*-mutant cells, validated this isoform (Figures 6B, 6C, and S6A). This aberrant 3′ splice site recognition occurred in exon 5 of MAP3K7, which encodes part of the kinase domain, and is predicted to result in an out-of-frame transcript that undergoes nonsense-mediated decay (Figures S6B and S6C). Consistent with this, we observed reduced MAP3K7 protein in isogenic cell lines (Figures 6D and 6E) as well as patient BM mononuclear cells (MNCs) and peripheral blood MNCs from MDS and CLL patients, respectively (Figures 6F and 6G), with *SF3B1* mutations versus those lacking splicing-factor mutations.

MAP3K7 encodes a kinase that mediates tumor necrosis factor α (TNFα), interleukin-1β (IL-1β), and Toll-like receptor signaling through the NF-κB, JNK, and MAPK pathways. The effects of MAP3K7 loss have been extensively studied and can result in loss or promotion of inflammation depending on cellular context (Ajibade et al., 2012; Lamothe et al., 2012; Sato et al., 2005; Tang et al., 2008; Vink et al., 2013; Xin et al., 2017). Here, we observed that *SF3B1*^{K700E} human myeloid or lymphoid leukemia cells stimulated with LPS had enhanced NF-κB activation compared with *SF3B1* WT controls (Figures 6D, 6E, S6D, and S6E). Increased p-p65 level was also evident at baseline in isogenic cells as well as primary patient BM MNCs from *SF3B1*-mutant MDS and peripheral blood MNCs from *SF3B1*-mutant CLL relative to *SF3B1* WT counterparts (Figures 6F and 6G). Increased nuclear p-p65 levels (Figure 7A) and enhanced NF-κB transcriptional activity using a reporter assay (Figures 7B and S7A) were also evident in *SF3B1*^{K700E} cells compared with *SF3B1* WT cells following LPS or TNFα stimulation. This was further confirmed by marked induction in NF-κB targets IL-1β and TNFα following LPS treatment in *SF3B1*^{K700E} NALM-6 cells relative to parental and *SF3B1*^{K700K} control cells (Figure 7C).

Given that a number of aberrant splicing and gene expression events occur in *SF3B1*-mutant cells, we next sought to understand the contribution of MAP3K7 loss to hyperactivated NF-κB signaling. First, hyperactive NF-κB signaling was confirmed using shRNA-mediated downregulation of MAP3K7 in NALM-6 and K562 parental cells at the level of p-p65 signaling and NF-κB transcriptional activity at both baseline and after LPS stimulation (Figures 7B and S7B). In addition, re-expression of MAP3K7 in K562 *SF3B1*^{K700E} cells resulted in a significant decrease in p-p65 in both resting state as well as following LPS exposure based on immunoblotting, NF-κB luciferase reporter, and nuclear p-p65 level (Figures 7D, 7F, and S7C). At a biological level, restoration of Map3k7 expression in *Sf3b1*$^{K700E/+}$ HSPCs resulted in mild rescue of HSPC clonogenicity (Figure S7D). Overall, these data suggest that the effects of SF3B1^{K700E} mutation on induction of NF-κB signaling are indeed, in part, mediated through aberrant splicing of MAP3K7.

SRSF2 Mutations Promote Aberrant Splicing of Caspase-8, Resulting in a Truncated Protein that Hyperactivates NF-κB Signaling

Consistent with the distinct effects of *SF3B1* and *SRSF2* mutations on splicing, MAP3K7 aberrant splicing was restricted to *SF3B1*-mutant cells (Figure S8A). We therefore searched for aberrant splicing events in *SRSF2*-mutant cells that might affect NF-κB signaling. One such event was aberrant splicing of caspase-8, encoded by *CASP8*, that was recurrently mis-spliced in AML and CMML patients with *SRSF2* mutations, but not those bearing *SF3B1* mutations (Figures 8A and S8B). Caspase-8 is a cysteine protease that initiates death-receptor-mediated apoptosis (Shu et al., 1997; Thome et al., 1997) in addition to regulating necroptosis and serving as a key activator of NF-κB (Chaudhary et al., 2000; Hu et al., 2000; Shikama et al., 2003). *SRSF2* mutations repressed a cassette exon of caspase-8, as was evident by RNA-seq, RT-PCR, qRT-PCR, and cDNA sequencing from cell lines and patient samples (Figures 8A, 8B, and S8C–S8E). CASP8 normally encodes a 54/55-kDa protein containing two death-effector domains (DED) at the N terminus and a C-terminal catalytic domain. Exclusion of this cassette exon results in an mRNA encoding a truncated caspase-8 protein lacking the C-terminal catalytic domains, which was readily

(E) Violin plots quantifying nuclear p-p65 intensity of LK cells from (D). ANOVA and Kruskal-Wallis ranked test was performed and adjusted for false discovery rate.
(F) Schema of competitive BMT using BM MNCs from *Vav*-Cre$^+$ WT, *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$, and *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ mice after chronic LPS exposure.
(G) Percentage of CD45.2 chimerism in peripheral blood of recipients (n = 4–5 mice per group) from primary and secondary BMT (both at 10 weeks post BMT).
(H) Chimerism of LSK or myeloid progenitor (MP; lineage$^-$ Sca-1$^-$ c-Kit$^+$) fractions from primary recipients 14 weeks post BMT.
(I) Kaplan-Meier analysis of *Mx1*-Cre$^+$ WT, *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$, and *Mx1*-Cre$^+$ *Sf3b1*$^{K700E/+}$ mice after a single dose of LPS (15 mg/kg) *in vivo*. Log-rank Mantel-Cox test was performed. **p = 0.0055.
ANOVA and Tukey's post hoc test were used to compare groups. *p < 0.05, **p < 0.01, ***p < 0.001, ****p < 0.0001 versus *Vav*-Cre$^+$ WT mice ("#" denotes versus *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ mice and "^" denotes versus *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ mice). In (G) and (H), the top and bottom lines of the box represent the upper and lower quartiles, respectively; the line inside the box represents the median; the lines above and below the box represent the maximum and minimum values, respectively. See also Figure S5 and Table S6.

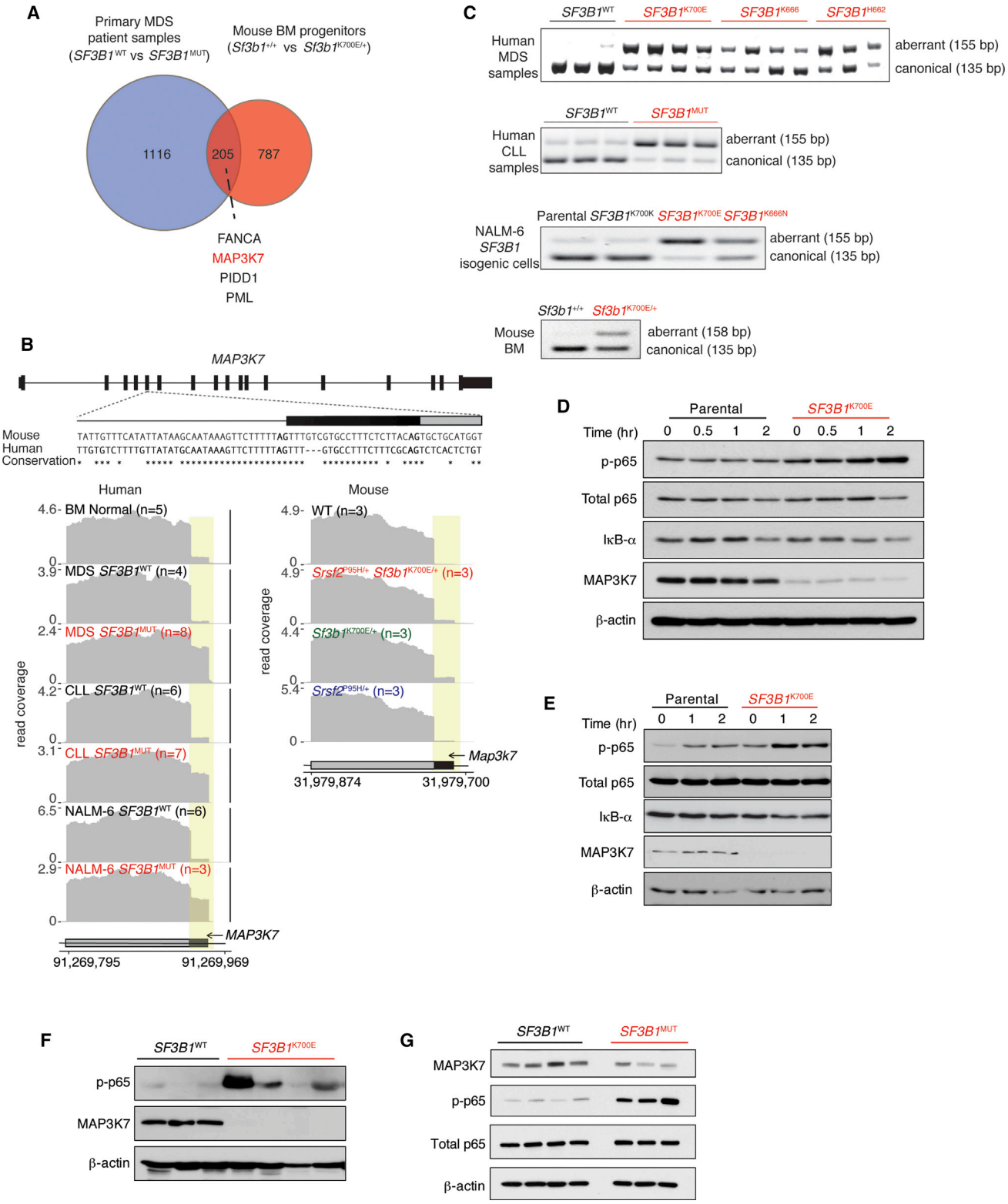


Figure 6. *SF3B1* Mutations Promote Mis-splicing of MAP3K7, Resulting in Hyperactivation of NF-κB Signaling

(A) Venn diagram illustrating overlap of differentially spliced genes in MDS patient samples mutant versus WT for *SF3B1* and murine hematopoietic progenitors in *Sf3b1*$^{K700E/+}$ versus *Sf3b1*$^{+/+}$ mice.

(B) From top to bottom, conservation of mouse and human MAP3K7 sequences adjacent to the competing 3′ splice site affected by SF3B1 mutations, and RNA-seq coverage plots in human and mouse samples.

(legend continued on next page)

detectable in *SRSF2*-mutant cells using an N-terminal anti-caspase-8 antibody (Figures 8C and 8D). Although several caspase-8 isoforms have been described (Xu et al., 2009; Yuan et al., 2012), the specific isoform detected in *SRSF2*-mutant cells is distinct from those previously described.

Previously described DED-only forms of caspase-8 have been suggested to serve either as competitive inhibitors or promoters of apoptosis (Xu et al., 2009; Yuan et al., 2012). Given this, we tested the effect of this *SRSF2*-mutant-specific truncated isoform (hereafter referred to as CASP8^{TR}) in an *SRSF2* WT cell line followed by TRAIL (TNF-related apoptosis-inducing ligand) stimulation to engage the death-receptor pathway. Overexpression of both caspase-8 full-length (CASP8^{FL}) and CASP8^{TR} isoforms resulted in robust expression of proteins at the expected size (Figures 8E and S8F) with no effect on cell growth (Figure S8G). Moreover, both isoforms promoted TRAIL-mediated cell death, suggesting that the truncated isoform did not affect cell death relative to CASP8^{FL} (Figure S8H). Given this, we next evaluated the effect of overexpressing CASP8^{TR} on NF-κB transcriptional activity, signaling, and nuclear localization. In both K562 and HAP1 cells with endogenous caspase-8 expression, overexpression of the CASP8^{TR}, but not the CASP8^{FL} isoform, resulted in robust induction of NF-κB activity upon increasing concentration of TRAIL (Figures 8E, 8F, S8I, and S8J). These data were also confirmed in K562 cells with or without endogenous *SRSF2*^{P95H} mutation (Figures 8G, S8K, and S8L). To rule out an effect of endogenous caspase-8, we assessed the effect of CASP8^{TR} on cell death and NF-κB activity in *CASP8*^{KO} HAP1 cells. Consistent with the lack of the catalytic domain, *CASP8*^{KO} HAP1 cells expressing CASP8^{TR} isoform were unable to undergo TRAIL-mediated cell death (Figures S8M and S8N); however, the CASP8^{TR} isoform was able to induce robust NF-κB signaling in the absence of WT caspase-8 (Figures 8F, S8O, and S8P).

DISCUSSION

Mutations affecting RNA splicing factors are the most common genetic alterations in MDS but the basis for their significant enrichment in this disease remains largely unexplained. Here we identify that two of the most commonly mutated splicing factors in MDS converge on activation of innate immune signaling through aberrant splicing of mRNAs encoding distinct enzymes.

Although the initial description of spliceosomal mutations predicted that the mutually exclusive pattern of these mutations might be due to a common impact on MDS pathogenesis (Yoshida et al., 2011), the data here identify that these mutations are mutually exclusive, in part due to synthetic lethal interaction. This provides one of the few examples of mutually exclusive oncogenic alterations due to synthetic lethal interactions in cancer. Interestingly, co-expression of SF3B1^{K700E} and SRSF2^{P95H} mutations resulted in impaired HSPC self-renewal, differentiation, and survival, but this is not due to widespread inhibition of splicing efficiency. Instead we identified aberrant splicing and dysregulation of key regulators of HSPC function in double-mutant cells, including the thrombopoietin receptor c-Mpl, the homeodomain transcription factor Pbx1, and integrin αIIb. Hematopoietic-specific deletion of these factors individually has previously been shown to result in failure of hematopoiesis due to reduced HSC self-renewal, increased apoptosis, and loss of quiescence (Ficara et al., 2008; Gekas and Graf, 2013; Qian et al., 2007; Yoshihara et al., 2007), all features characteristic of the double-mutant state.

Multiple lines of evidence implicate innate immune signaling in MDS pathogenesis (Basiorka et al., 2016; Fang et al., 2014; Wei et al., 2013), including experiments demonstrating increased innate immune signaling contributes to MDS development *in vivo* (Fang et al., 2017; Varney et al., 2015). Increased activation of Toll-like receptor and IL-1 receptor signaling with downstream activation of MAPK and NF-κB pathways is widely reported in MDS, but the mechanistic basis for this activation have mostly been restricted to MDS with deletion of chromosome 5q (Fang et al., 2014, 2017; Starczynowski, 2014; Varney et al., 2015). Our work identifies a mechanism for hyperactivated NF-κB signaling in a wider spectrum of MDS.

Although recent studies of *SF3B1*^{K700E} mutation identified relatively few mRNAs mis-spliced in both mouse and human cells (Mupo et al., 2016; Obeng et al., 2016), we identified a much greater overlap of aberrantly spliced transcripts shared across mouse and human *SF3B1*^{K700E} mutant hematopoietic cells than previously reported. This includes aberrant splicing of MAP3K7 through the use of an alternative 3′ splice site promoted by mutant SF3B1. As downregulation of Map3k7 in myeloid cells promotes myeloid neoplasms, this finding is likely to be relevant to *SF3B1*-mutant MDS. For example, *in vivo* knockdown of Map3k7 resulted in splenomegaly, myeloproliferation, and extramedullary hematopoiesis, as well as increased immune activation that was exacerbated by LPS (Vink et al., 2013). Similarly, myeloid-specific deletion of *Map3k7* heightened the response to inflammatory stimuli and causes a clonal myeloid leukemia (Ajibade et al., 2012; Eftychi et al., 2012; Lamothe et al., 2012; Xin et al., 2017). Interestingly, the pro-inflammatory and leukemogenic effects of *Map3k7* loss appear to be restricted to myeloid lineages (Ajibade et al., 2012), in stark contrast to pan-hematopoietic deletion of *Map3k7*, which results in complete failure of hematopoiesis (Tang et al., 2008). It will therefore be important to understand whether there is a differential requirement for hematopoiesis in *SF3B1* mutant cells given the partial loss of MAP3K7 induced by mutant *SF3B1*.

While aberrant splicing and downregulation of MAP3K7 was specific to SF3B1 mutant cells, cells expressing mutant SRSF2 shared similar elevated innate immune signaling and hypersensitivity to LPS. This led us to identify a gain-of-function effect of *SRSF2* mutations through generation of a C-terminal truncated caspase-8 isoform that promotes NF-κB signaling. This is

(C) RT-PCR of the MAP3K7 competing 3′ splice site in MDS and CLL patient samples with or without *SF3B1* mutations as well as isogenic human and mouse cells.
(D and E) Immunoblot of phosphorylated p65 (p-p65), IκB-α, MAP3K7, and loading controls in isogenic K562 (D) and NALM-6 (E) cells. "Time (hr)" refers to hours following LPS (5 μg/mL) exposure.
(F and G) Immunoblot analysis of p-p65, MAP3K7, and loading controls in BM MNCs from MDS (F) and peripheral blood MNCs from CLL patients (G) with or without *SF3B1* mutations.
See also Figure S6 and Table S7.

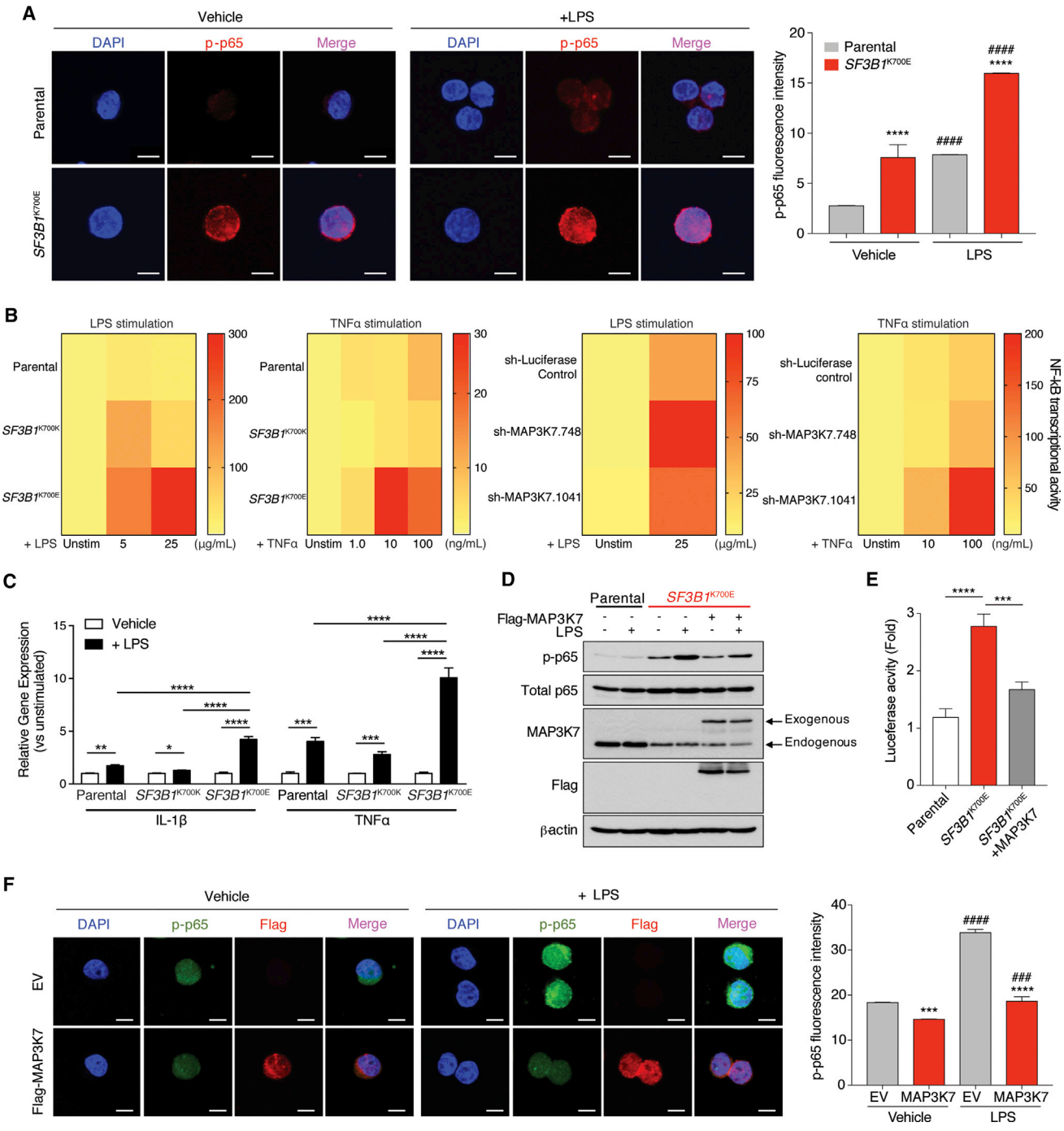


Figure 7. MAP3K7 Loss Results in Hyperactive NF-κB Signaling in *SF3B1*-Mutant Cells

(A) Immunofluorescence of phosphorylated p65 (p-p65) in K562 cells with or without *SF3B1*^{K700E} mutation 2 hr following LPS stimulation. Quantitation of p-p65 intensity is shown on the right (n = 3 independent experiments). Scale bars, 10 μm.

(B) Heatmap of NF-κB reporter signal in NALM-6 *SF3B1*-isogenic cells (left two panels) or parental NALM-6 cells with MAP3K7 shRNAs (right two panels) following LPS or TNFα stimulation for 24 hr.

(C) qRT-PCR analysis of IL-1β and TNFα 8 hr post LPS stimulation in NALM-6 *SF3B1*-isogenic cells (n = 2 independent experiments).

(D) Immunoblot of p-p65 in K562 *SF3B1*-isogenic cells ± FLAG-MAP3K7 cDNA and/or LPS (5 μg/mL) exposure for 2 hr.

(E) Quantitation of NF-κB reporter signal in cells from (D) (n = 3 independent experiments).

(F) Immunofluorescence of nuclear p-p65 and FLAG (MAP3K7) in cells from (D) (quantitation of p-p65 intensity [n = 3 independent experiments] on the right). Scale bars, 10 μm.

Error bars represent mean ± SD. ANOVA followed by Tukey's post hoc test were used to compare groups. *p < 0.05, **p < 0.005, ***p < 0.0002, ****p < 0.0001, ####p < 0.0001 versus Vehicle; ###p < 0.001 versus Vehicle. See also Figure S7.

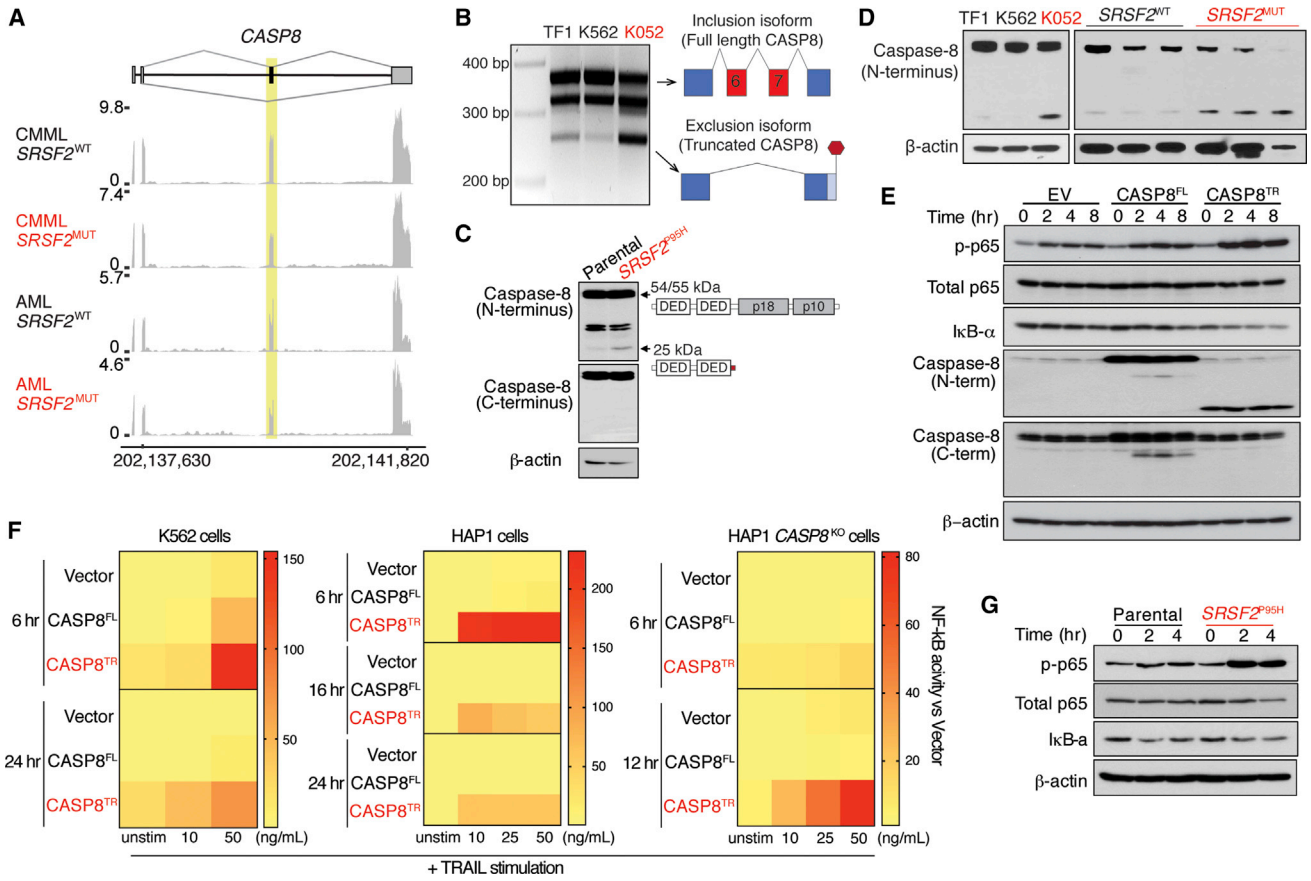


Figure 8. *SRSF2* Mutations Promote Aberrant Splicing of Caspase-8, Resulting in Expression of a Truncated Protein that Hyperactivates NF-κB Signaling

(A) RNA-seq coverage plots of caspase-8 splicing in CMML and AML patients WT or mutant (MUT) for *SRSF2*.

(B) RT-PCR analysis of caspase-8 splicing in *SRSF2* WT (TF1 and K562) or *SRSF2*-mutant (K052) cells.

(C and D) Immunoblot of caspase-8 using an N-terminal anti-caspase-8 antibody in K562 *SRSF2*-isogenic cells (C) or human leukemia cells (left panel) and primary AML patient samples (right panel) WT or mutant for *SRSF2* (D).

(E) Immunoblot of phosphorylated p65 (p-p65), IκB-α, caspase-8 in K562 cells expressing empty vector (EV), full-length caspase-8 ($CASP8^{FL}$), or the truncated caspase-8 isoform ($CASP8^{TR}$) after exposure to TRAIL (50 ng/mL).

(F) Heatmap of NF-κB reporter signal following TRAIL stimulation in K562, HAP1, or $CASP8^{KO}$ HAP1 cells expressing EV, $CASP8^{FL}$, or $CASP8^{TR}$.

(G) Immunoblot of p-p65 and IκB-α in K562 cells with or without $SRSF2^{P95H}$ following exposure to TRAIL (50 ng/mL).

See also Figure S8.

consistent with prior work showing that N-terminal prodomain-only containing isoforms of procaspase-8 activate NF-κB signaling through interactions with upstream regulators of NF-κB, a function not mediated by canonical full-length caspase-8 (Chaudhary et al., 2000; Hu et al., 2000; Shikama et al., 2003).

Together, our data demonstrate that different splicing-factor mutations alter distinct targets at the level of pre-mRNA splicing that nonetheless converge on the same downstream signaling node to hyperactivate innate immune signaling.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Animals
  - Primary Human MDS and CLL Samples
  - Cell Lines
- METHOD DETAILS
  - Peripheral Blood Analysis
  - Bone Marrow Transplantation Assays
  - *In Vivo* LPS Stimulation Experiment
  - *In Vitro* Colony-Forming Assays
  - Flow Cytometry Analyses
  - Cell Cycle and Apoptosis Analyses
  - Histological Analysis
  - Immunoblot
  - Immunofluorescence
  - Caspase-8 and MAP3K7 Constructs
  - mRNA Stability Assay
  - Luciferase Reporter Assay

- RT-PCR and Quantitative RT-PCR (qRT-PCR)
- shRNA Experiments
- mRNA Isolation, Sequencing, and Analysis
- Genome Annotations
- RNA-seq Read Mapping
- Differential Gene Expression Analysis
- Differential Splicing Analysis
- Gene Ontology (GO) Enrichment Analysis
- Motif Enrichment Analysis and Sequence Logos
- Analysis of Expected and Observed Gene Expression Convergence

● QUANTIFICATION AND STATISTICAL ANALYSIS
- Statistical Analyses

● DATA AND SOFTWARE AVAILABILITY
- Publicly Available RNA-seq Data
- Accession Codes

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and eight tables and can be found with this article online at https://doi.org/10.1016/j.ccell.2018.07.003.

ACKNOWLEDGMENTS

This work was supported by the Leukemia and Lymphoma Society (S.C.-W.L., D.I., R.K.B., O.A.-W.), the NCI K99 CA218896 (S.C.-W.L.), Aplastic Anemia and MDS International Foundation (A.Y.), Lauri Strauss Leukemia Foundation (A.Y.), US Dept. of Defense Bone Marrow Failure Research Program grant W81XWH-12-1-0041 (R.K.B., O.A.-W.), the National Research Foundation of Korea (NRF) grant funded by the Korean Government (Young Researcher Program, NRF-2017R1C1B2001991) (E.K.), the Settlement Research Fund (1.160104.01) of UNIST (Ulsan National Institute of Science and Technology) (E.K.), the NRF grant (the Individual Research in Basic Science and Engineering program, NRF-2017R1D1A1B03034094) (E.J.), the American Society of Hematology (S.C.-W.L., J.T., B.H.D., O.A.-W.), the Edward P. Evans Foundation (R.K.B., O.A.-W.), the Taub Foundation (O.A.-W.), grant R01 HL128239 (R.K.B., O.A.-W.), the Ellison Medical Foundation grant AG-NS-1030-13 (R.K.B.), grant R01 DK103854 (R.K.B.), the Starr Foundation grants I8-A8-075 and I9-A9-059 (O.A.-W.), and The Pershing Square Sohn Foundation. We would like to acknowledge the Molecular Cytology, Flow Cytometry, and Laboratory of Comparative Pathology core facilities for their technical assistance, and the MSK Cancer Center Support grant P30 CA008748.

AUTHOR CONTRIBUTIONS

S.C.-W.L., K.N., E.K., R.K.B., and O.A.-W. designed the study. S.C.-W.L., K.N., E.K., E.J., S.X.L., B.L., D.I., A.Y., M.K., M.Y., X.J.Z., M.K.K., H.C., Y.R.C., J.T., B.H.D., S.M., and Y.J.K. performed experiments. S.C.-W.L., K.N., E.K., E.J., S.X.L., B.L., D.I., A.P., J.P., M.S., S.B., P.G.S., and R.K.B. analyzed data. E.O. and B.L.E. provided the *Sf3b1*^{K700E} mouse. S.C.-W.L., K.N., R.K.B., and O.A.-W. prepared the manuscript with help from all co-authors.

DECLARATION OF INTERESTS

J.P., M.S., S.B., and P.G.S. are employees of H3 Biomedicine.

Received: May 11, 2017
Revised: April 25, 2018
Accepted: July 12, 2018
Published: August 13, 2018

REFERENCES

Ajibade, A.A., Wang, Q., Cui, J., Zou, J., Xia, X., Wang, M., Tong, Y., Hui, W., Liu, D., Su, B., et al. (2012). TAK1 negatively regulates NF-kappaB and p38 MAP kinase activation in Gr-1^{+}CD11b^{+} neutrophils. Immunity *36*, 43–54.

Alexander, W.S., Roberts, A.W., Nicola, N.A., Li, R., and Metcalf, D. (1996). Deficiencies in progenitor cells of multiple hematopoietic lineages and defective megakaryocytopoiesis in mice lacking the thrombopoietic receptor c-Mpl. Blood *87*, 2162–2170.

Basiorka, A.A., McGraw, K.L., Eksioglu, E.A., Chen, X., Johnson, J., Zhang, L., Zhang, Q., Irvine, B.A., Cluzeau, T., Sallman, D.A., et al. (2016). The NLRP3 inflammasome functions as a driver of the myelodysplastic syndrome phenotype. Blood *128*, 2960–2975.

Bejar, R., Stevenson, K.E., Caughey, B.A., Abdel-Wahab, O., Steensma, D.P., Galili, N., Raza, A., Kantarjian, H., Levine, R.L., Neuberg, D., et al. (2012). Validation of a prognostic model and the impact of mutations in patients with lower-risk myelodysplastic syndromes. J. Clin. Oncol. *30*, 3376–3382.

Chaudhary, P.M., Eby, M.T., Jasmin, A., Kumar, A., Liu, L., and Hood, L. (2000). Activation of the NF-kappaB pathway by caspase 8 and its homologs. Oncogene *19*, 4451–4460.

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. (2012). Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. Blood *119*, 3211–3218.

Darman, R.B., Seiler, M., Agrawal, A.A., Lim, K.H., Peng, S., Aird, D., Bailey, S.L., Bhavsar, E.B., Chan, B., Colla, S., et al. (2015). Cancer-associated SF3B1 hotspot mutations induce cryptic 3′ splice site selection through use of a different branch point. Cell Rep. *13*, 1033–1045.

DeBoever, C., Ghia, E.M., Shepard, P.J., Rassenti, L., Barrett, C.L., Jepsen, K., Jamieson, C.H., Carson, D., Kipps, T.J., and Frazer, K.A. (2015). Transcriptome sequencing reveals potential mechanism of cryptic 3′ splice site selection in SF3B1-mutated cancers. PLoS Comput. Biol. *11*, e1004105.

Dolatshad, H., Pellagatti, A., Fernandez-Mercado, M., Yip, B.H., Malcovati, L., Attwood, M., Przychodzen, B., Sahgal, N., Kanapin, A.A., Lockstone, H., et al. (2015). Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. Leukemia *29*, 1092–1103.

Dvinge, H., Ries, R.E., Ilagan, J.O., Stirewalt, D.L., Meshinchi, S., and Bradley, R.K. (2014). Sample processing obscures cancer-specific alterations in leukemic transcriptomes. Proc. Natl. Acad. Sci. USA *111*, 16802–16807.

Eftychi, C., Karagianni, N., Alexiou, M., Apostolaki, M., and Kollias, G. (2012). Myeloid TAKI [corrected] acts as a negative regulator of the LPS response and mediates resistance to endotoxemia. PLoS One *7*, e31550.

Esplin, B.L., Shimazu, T., Welner, R.S., Garrett, K.P., Nie, L., Zhang, Q., Humphrey, M.B., Yang, Q., Borghesi, L.A., and Kincade, P.W. (2011). Chronic exposure to a TLR ligand injures hematopoietic stem cells. J. Immunol. *186*, 5367–5375.

Fang, J., Barker, B., Bolanos, L., Liu, X., Jerez, A., Makishima, H., Christie, S., Chen, X., Rao, D.S., Grimes, H.L., et al. (2014). Myeloid malignancies with chromosome 5q deletions acquire a dependency on an intrachromosomal NF-kappaB gene network. Cell Rep. *8*, 1328–1338.

Fang, J., Bolanos, L.C., Choi, K., Liu, X., Christie, S., Akunuru, S., Kumar, R., Wang, D., Chen, X., Greis, K.D., et al. (2017). Ubiquitination of hnRNPA1 by TRAF6 links chronic innate immune signaling with myelodysplasia. Nat. Immunol. *18*, 236–245.

Fellmann, C., Hoffmann, T., Sridhar, V., Hopfgartner, B., Muhar, M., Roth, M., Lai, D.Y., Barbosa, I.A., Kwon, J.S., Guan, Y., et al. (2013). An optimized microRNA backbone for effective single-copy RNAi. Cell Rep. *5*, 1704–1713.

Ficara, F., Murphy, M.J., Lin, M., and Cleary, M.L. (2008). Pbx1 regulates self-renewal of long-term hematopoietic stem cells by maintaining their quiescence. Cell Stem Cell *2*, 484–496.

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.

Gekas, C., and Graf, T. (2013). CD41 expression marks myeloid-biased adult hematopoietic stem cells and increases with age. Blood *121*, 4463–4472.

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.

Harbour, J.W., Roberson, E.D., Anbunathan, H., Onken, M.D., Worley, L.A., and Bowcock, A.M. (2013). Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. Nat. Genet. *45*, 133–135.

Hu, W.H., Johnson, H., and Shu, H.B. (2000). Activation of NF-kappaB by FADD, casper, and caspase-8. J. Biol. Chem. *275*, 10838–10844.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with bioconductor. Nat. Methods *12*, 115–121.

Hubert, C.G., Bradley, R.K., Ding, Y., Toledo, C.M., Herman, J., Skutt-Kakaria, K., Girard, E.J., Davison, J., Berndt, J., Corrin, P., et al. (2013). Genome-wide RNAi screens in human brain tumor isolates reveal a novel viability requirement for PHF5A. Genes Dev. *27*, 1032–1045.

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.

Katz, Y., Wang, E.T., Airoldi, E.M., and Burge, C.B. (2010). Analysis and design of RNA sequencing experiments for identifying isoform regulation. Nat. Methods *7*, 1009–1015.

Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C., Ramakrishnan, A., Li, Y., Chung, Y.R., Micol, J.B., Murphy, M.E., et al. (2015). SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. Cancer Cell *27*, 617–630.

Lamothe, B., Lai, Y., Hur, L., Orozco, N.M., Wang, J., Campos, A.D., Xie, M., Schneider, M.D., Lockworth, C.R., Jakacky, J., et al. (2012). Deletion of TAK1 in the myeloid lineage results in the spontaneous development of myelomonocytic leukemia in mice. PLoS One *7*, e51228.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. *10*, R25.

Lasho, T.L., Jimma, T., Finke, C.M., Patnaik, M., Hanson, C.A., Ketterling, R.P., Pardanani, A., and Tefferi, A. (2012). SRSF2 mutations in primary myelofibrosis: significant clustering with IDH mutations and independent association with inferior overall and leukemia-free survival. Blood *120*, 4168–4171.

Lee, S.C., Dvinge, H., Kim, E., Cho, H., Micol, J.B., Chung, Y.R., Durham, B.H., Yoshimi, A., Kim, Y.J., Thomas, M., et al. (2016). Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. Nat. Med. *22*, 672–678.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics *12*, 323.

Makishima, H., Visconte, V., Sakaguchi, H., Jankowska, A.M., Abu Kar, S., Jerez, A., Przychodzen, B., Bupathi, M., Guinta, K., Afable, M.G., et al. (2012). Mutations in the spliceosome machinery, a novel and ubiquitous pathway in leukemogenesis. Blood *119*, 3203–3210.

Martin, M., Masshofer, L., Temming, P., Rahmann, S., Metz, C., Bornfeld, N., van de Nes, J., Klein-Hitpass, L., Hinnebusch, A.G., Horsthemke, B., et al. (2013). Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. Nat. Genet. *45*, 933–936.

Meggendorfer, M., Roller, A., Haferlach, T., Eder, C., Dicker, F., Grossmann, V., Kohlmann, A., Alpermann, T., Yoshida, K., Ogawa, S., et al. (2012). SRSF2 mutations in 275 cases with chronic myelomonocytic leukemia (CMML). Blood *120*, 3080–3088.

Meyer, L.R., Zweig, A.S., Hinrichs, A.S., Karolchik, D., Kuhn, R.M., Wong, M., Sloan, C.A., Rosenbloom, K.R., Roe, G., Rhead, B., et al. (2013). The UCSC Genome Browser database: extensions and updates 2013. Nucleic Acids Res. *41*, D64–D69.

Mupo, A., Seiler, M., Sathiaseelan, V., Pance, A., Yang, Y., Agrawal, A.A., Iorio, F., Bautista, R., Pacharne, S., Tzelepis, K., et al. (2016). Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. Leukemia *31*, 720–727.

Obeng, E.A., Chappell, R.J., Seiler, M., Chen, M.C., Campagna, D.R., Schmidt, P.J., Schneider, R.K., Lord, A.M., Wang, L., Gambe, R.G., et al. (2016). Physiologic expression of Sf3b1(K700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell *30*, 404–417.

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. (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. (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. Blood *122*, 3616–3627, quiz 3699.

Patnaik, M.M., Lasho, T.L., Finke, C.M., Hanson, C.A., Hodnefield, J.M., Knudson, R.A., Ketterling, R.P., Pardanani, A., and Tefferi, A. (2013). Spliceosome mutations involving SRSF2, SF3B1, and U2AF35 in chronic myelomonocytic leukemia: prevalence, clinical correlates, and prognostic relevance. Am. J. Hematol. *88*, 201–206.

Pelossof, R., Fairchild, L., Huang, C.H., Widmer, C., Sreedharan, V.T., Sinha, N., Lai, D.Y., Guan, Y., Premsrirut, P.K., Tschaharganeh, D.F., et al. (2017). Prediction of potent shRNAs with a sequential classification algorithm. Nat. Biotechnol. *35*, 350–353.

Qian, H., Buza-Vidas, N., Hyland, C.D., Jensen, C.T., Antonchuk, J., Mansson, R., Thoren, L.A., Ekblom, M., Alexander, W.S., and Jacobsen, S.E. (2007). Critical role of thrombopoietin in maintaining adult quiescent hematopoietic stem cells. Cell Stem Cell *1*, 671–684.

Robinson, M.D., and Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biol. *11*, R25.

Sato, S., Sanjo, H., Takeda, K., Ninomiya-Tsuji, J., Yamamoto, M., Kawai, T., Matsumoto, K., Takeuchi, O., and Akira, S. (2005). Essential function for the kinase TAK1 in innate and adaptive immune responses. Nat. Immunol. *6*, 1087–1095.

Shikama, Y., Yamada, M., and Miyashita, T. (2003). Caspase-8 and caspase-10 activate NF-kappaB through RIP, NIK and IKKalpha kinases. Eur. J. Immunol. *33*, 1998–2006.

Shu, H.B., Halpin, D.R., and Goeddel, D.V. (1997). Casper is a FADD- and caspase-related inducer of apoptosis. Immunity *6*, 751–763.

Starczynowski, D.T. (2014). Errant innate immune signaling in del(5q) MDS. Blood *124*, 669–671.

t Hoen, P.A., Hirsch, M., de Meijer, E.J., de Menezes, R.X., van Ommen, G.J., and den Dunnen, J.T. (2011). mRNA degradation controls differentiation state-dependent differences in transcript and splice variant abundance. Nucleic Acids Res. *39*, 556–566.

Tang, M., Wei, X., Guo, Y., Breslin, P., Zhang, S., Zhang, S., Wei, W., Xia, Z., Diaz, M., Akira, S., and Zhang, J. (2008). TAK1 is required for the survival of hematopoietic cells and hepatocytes in mice. J. Exp. Med. *205*, 1611–1619.

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

Thome, M., Schneider, P., Hofmann, K., Fickenscher, H., Meinl, E., Neipel, F., Mattmann, C., Burns, K., Bodmer, J.L., Schroter, M., et al. (1997). Viral FLICE-inhibitory proteins (FLIPs) prevent apoptosis induced by death receptors. Nature *386*, 517–521.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105–1111.

Varney, M.E., Niederkorn, M., Konno, H., Matsumura, T., Gohda, J., Yoshida, N., Akiyama, T., Christie, S., Fang, J., Miller, D., et al. (2015). Loss of Tifab, a del(5q) MDS gene, alters hematopoiesis through derepression of Toll-like receptor-TRAF6 signaling. J. Exp. Med. *212*, 1967–1985.

Vink, P.M., Smout, W.M., Driessen-Engels, L.J., de Bruin, A.M., Delsing, D., Krajnc-Franken, M.A., Jansen, A.J., Rovers, E.F., van Puijenbroek, A.A., Kaptein, A., et al. (2013). In vivo knockdown of TAK1 accelerates bone marrow proliferation/differentiation and induces systemic inflammation. PLoS One *8*, e57348.

Wagenmakers, E.J., Lodewyckx, T., Kuriyal, H., and Grasman, R. (2010). Bayesian hypothesis testing for psychologists: a tutorial on the Savage-Dickey method. Cogn. Psychol. *60*, 158–189.

Wang, L., Lawrence, M., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D., Zhang, L., et al. (2011). SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N. Engl. J. Med. *365*, 2497–2506.

Wei, Y., Dimicoli, S., Bueso-Ramos, C., Chen, R., Yang, H., Neuberg, D., Pierce, S., Jia, Y., Zheng, H., Wang, H., et al. (2013). Toll-like receptor alterations in myelodysplastic syndrome. Leukemia *27*, 1832–1840.

Xin, J., Breslin, P., Wei, W., Li, J., Gutierrez, R., Cannova, J., Ni, A., Ng, G., Schmidt, R., Chen, H., et al. (2017). Necroptosis in spontaneously-mutated hematopoietic cells induces autoimmune bone marrow failure in mice. Haematologica *102*, 295–307.

Xu, Z., Tang, K., Wang, M., Rao, Q., Liu, B., and Wang, J. (2009). A new caspase-8 isoform caspase-8s increased sensitivity to apoptosis in Jurkat cells. J. Biomed. Biotechnol. *2009*, 930462.

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.

Yoshihara, H., Arai, F., Hosokawa, K., Hagiwara, T., Takubo, K., Nakamura, Y., Gomei, Y., Iwasaki, H., Matsuoka, S., Miyamoto, K., et al. (2007). Thrombopoietin/MPL signaling regulates hematopoietic stem cell quiescence and interaction with the osteoblastic niche. Cell Stem Cell *1*, 685–697.

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.

Yuan, R.T., Young, S., Liang, J., Schmid, M.C., Mielgo, A., and Stupack, D.G. (2012). Caspase-8 isoform 6 promotes death effector filament formation independent of microtubules. Apoptosis *17*, 229–235.

Zhang, J., Lieu, Y.K., Ali, A.M., Penson, A., Reggio, K.S., Rabadan, R., Raza, A., Mukherjee, S., and Manley, J.L. (2015). Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. Proc. Natl. Acad. Sci. USA *112*, E4726–E4734.

Zhang, S.J., Rampal, R., Manshouri, T., Patel, J., Mensah, N., Kayserian, A., Hricik, T., Heguy, A., Hedvat, C., Gonen, M., et al. (2012). Genetic analysis of patients with leukemic transformation of myeloproliferative neoplasms shows recurrent SRSF2 mutations that are associated with adverse outcome. Blood *119*, 4480–4485.

Zhao, Y., Ling, F., Wang, H.C., and Sun, X.H. (2013). Chronic TLR signaling impairs the long-term repopulating potential of hematopoietic stem cells of wild type but not Id1 deficient mice. PLoS One *8*, e55552.

Zhou, Q., Derti, A., Ruddy, D., Rakiec, D., Kao, I., Lira, M., Gibaja, V., Chan, H., Yang, Y., Min, J., et al. (2015). A chemical genetics approach for the functional assessment of novel cancer genes. Cancer Res. *75*, 1949–1958.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-B220 Alexa Fluor 700 (RA3-6B2)	eBioscience	Cat# 56-0452-82; RRID: AB_891458
Anti-CD19 PECy7 (1D3)	eBioscience	Cat# 25-0193-82; RRID: AB_657663
Anti-CD3 BV605 (17A2)	BioLegend	Cat# 100237; RRID: AB_2562039
Anti-CD4 BV711 (GK1.5)	BioLegend	Cat# 100447; RRID: AB_2564586
Anti-CD8a PerCP/Cy5.5 (53-6.7)	BioLegend	Cat# 100734; RRID: AB_2075238
Anti-Gr-1 PECy7 (RB6-8C5)	eBioscience	Cat# 25-5931-82; RRID: AB_469663
Anti-CD11b/Mac-1 PE (M1/70)	eBioscience	Cat# 12-0112-82; RRID: AB_465546
Anti-NK1.1 APCCy7 (PK136)	BioLegend	Cat# 108724; RRID: AB_830871
Anti-Ter119	BioLegend	Cat# 116223; RRID: AB_2137788
Anti-c-Kit APC (2B8)	eBioscience	Cat# 17-1171-82; RRID: AB_469430
Anti-Sca-1 PECy7 (D7)	eBioscience	Cat# 25-5981-82; RRID: AB_469669
Anti-FcγRII/III Alexa Fluor 700 (93)	eBioscience	Cat# 56-0161-82; RRID: AB_493994
Anti-CD34 FITC (RAM34)	eBioscience	Cat# 11-0341-82; RRID: AB_465021
Anti-CD45.1 PerCP/Cy5.5 (A20)	BioLegend	Cat# 110728; RRID: AB_893346
Anti-CD45.2 BV605 (104)	BioLegend	Cat# 109841; RRID: AB_2563485
Anti-CD48 PerCP/Cy5.5 (HM48-1)	BioLegend	Cat# 103422; RRID: AB_2075051
Anti-CD150 PE (9D1)	eBioscience	Cat# 12-1501-82; RRID: AB_465873
Anti-CD44 FITC (IM7)	eBioscience	Cat# 11-0441-82; RRID: AB_465045
Anti-IgM PE (II/41)	eBioscience	Cat# 12-5790-82; RRID: AB_465940
Anti-CD25 PE (PC61.5)	eBioscience	Cat # 12-0251-82; RRID: AB_465607
Anti-CD43 FITC (S11)	BioLegend	Cat# 143204; RRID: AB_10960745
Anti-BrdU FITC Flow Kit	BD Pharmingen	Cat# 559619; RRID: AB_2617060
Annexin V FITC Apoptosis Detection Kit	BD Pharmingen	Cat# 556547
Anti-NF-κB/p65 (D14E12)	Cell Signaling Technology	Cat# 8242; RRID: AB_10859369
Anti-phospho NF-κB/p65-Ser536 (93H1)	Cell Signaling Technology	Cat# 3033; RRID: AB_331284
Anti-IκBα, rabbit polyclonal	Cell Signaling Technologies	Cat# 9242; RRID: AB_823540
Anti-TAK1/MAP3K7 (D94D7), rabbit monoclonal	Cell Signaling Technologies	Cat# 5206; RRID: AB_10694079
Anti-Caspase-8, N-terminal (E6), rabbit monoclonal	Abcam	Cat# ab32125; RRID: AB_2068469
Anti-Caspase-8, C-terminal (12F5), mouse monoclonal	Enzo Life Science	Cat# ALX-804-242-C100; RRID: AB_2050949
Anti-β-actin (AC-15), mouse monoclonal	Sigma-Aldrich	Cat# A5441; RRID: AB_476744
Anti-Flag (M2), mouse monoclonal	Sigma-Aldrich	Cat# F1804; RRID: AB_262044
Anti-mouse CD45, rat monoclonal	BD Pharmingen	Cat# 550539; RRID: AB_2174426
Anti-Myeloperoxidase, rabbit polyclonal	Dako	Cat# A0398; RRID: AB_2335676
Anti-Cleaved Caspase-3, rabbit polyclonal	Cell Signaling Technologies	Cat# 9661; RRID: AB_2341188
Anti-mouse Alexa Fluor 488, goat polyclonal	ThermoFisher Scientific	Cat# A-11001; RRID: AB_2534069
Anti-rabbit Alexa Fluor 594, goat polyclonal	ThermoFisher Scientific	Cat# A-11012; RRID: AB_2534079
Prolong Gold Antifade mounting agent	ThermoFisher Scientific	Cat# P36930
Biological Samples		
MDS and CLL primary patient samples	MSKCC	
Chemicals, Peptides, and Recombinant Proteins		
Escherichia coli 0111:B5 LPS	Sigma Aldrich	Cat# L2880
Polybrene	Millipore	Cat# TR-1003-G
Recombinant TRAIL (soluble, human)	Enzo Life Science	Cat# ALX-201-073-3020

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant human TNFα	PeproTech	Cat# 300-01A
Recombinant mouse IL-3	R&D	Cat# 403-ML
Recombinant mouse IL-6	R&D	Cat# 406-ML
Recombinant mouse SCF	R&D	Cat# 455-MC
Recombinant mouse TPO	R&D	Cat# 488-TO
Normal goat serum	ThermoFisher Scientific	Cat# PCN500
Actinomycin D	Life Technologies	Cat# 11805-017
Critical Commercial Assays		
MethoCult™ GF M3434	StemCell Technologies	Cat# M3434
NF-κB Cignal™ Reporter Assay	Qiagen	Cat# CCS-013L
Deposited Data		
RNAseq data	This paper	GSE97452
Experimental Models: Cell Lines		
Human: NALM-6 cells, Parental, *SF3B1*^{K700K}, *SF3B1*^{K700E}, *SF3B1*^{K666N}	Horizon Discovery	Cat# N/A
Human: K562 cells, Parental, *SF3B1*^{K700E}, *SRSF2*^{P95H}	Horizon Discovery	Cat# N/A
Human: HAP1 Parental cells	Horizon Discovery	Cat# C631
Human: HAP1 *CASP8*^{KO} cells	Horizon Discovery	Cat# HZGHC001511c007
Human: 293T cells	ClonTech	Cat# 632180
Human: 293 GPII cells	Clontech	Cat# 631530
Experimental Models: Organisms/Strains		
Mice: *Mx1*-Cre transgenic mice (C.Cg-Tg(Mx1-cre)1Cgn/J)	The Jackson Laboratory	Cat# JAX: 005673
Mice: *Vav*-Cre transgenic mice (B6.Cg-Tg(Vav1-icre)A2Kio/J)	The Jackson Laboratory	Cat# JAX: 008610
Mice: *Srsf2*$^{P95H/+}$ (B6J.B6NTac(SJL)-*Srsf2tm1.1Oaw*/J)	Kim et al. 2015	
Mice: *Srsf2*$^{fl/fl}$ (B6;129S4-*Srsf2*^{tm1Xdfu}/J)	The Jackson Laboratory	Cat# JAX: 018019
Mice: *Sf3b1*$^{K700E//+}$	Obeng et al. 2016	
Oligonucleotides		
Please Refer to Table S8		Table S8
Recombinant DNA		
pVSV.G	Addgene	Cat# 12259
psPAX2	Addgene	Cat# 12260
MSCV-IRES-GFP empty vector	Addgene	Cat# 52107
MSCV-Flag-CASP8^{FL}-IRES-GFP	This paper	
MSCV-Flag-CASP8^{TR}-IRES-GFP	This paper	
MSCV-Flag-mMap3k7-IRES-GFP	This paper	
MSCV-Flag-hMAK3K7-IRES-GFP	This paper	
T3G-dsRED-mirE-PGK-Neo-IRES-rtTA3 (LT3RENIR)	Fellmann et al., 2013	
pMSCV-LTR-mirE-PGK-SV40-IRES-GFP (MLS-E)	Fellmann et al., 2013	
pLKO.Empty Vector-Puro	Kim et al. 2015	
pLKO.sh-UPF1-Puro	Kim et al. 2015	
pNL3.2.NF-κB-RE Reporter Plasmid	Promega	Cat# N1111
Software and Algorithms		
Bowtie v1.0.0	Langmead et al. 2009	
RSEM v1.2.4	Li and Dewey, 2011	
TopHat v2.0.8b	Trapnell et al. 2009	
TMM method	Robinson and Oshlack 2010	
Wagenmakers's Bayesian framework	Wagenmakers et al. 2010	
MISO v2.0	Katz et al. 2010	

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
GOseq	Young et al. 2010	
Bioconductor	Huber et al. 2015	
Other		
FASTQ files from published RNA-seq studies of patients with MDS	Dolatshad et al. 2015	GSE63569
FASTQ files from published RNA-seq studies of patients with CLL	Darman et al. 2015	GSE72790

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Omar Abdel-Wahab (abdelwao@mskcc.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Animals

All animals were housed at Memorial Sloan Kettering Cancer Center (MSKCC). All animal procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC. Generation and genotyping of the *Srsf2*$^{P95H/+}$, *Sf3b1*$^{K700E/+}$ and *Srsf2*$^{fl/+}$ as well as the *Mx1*-Cre and *Vav*-Cre transgenic mice have been previously described (Kim et al., 2015; Obeng et al., 2016). 8-week-old female CD45.1 C57BL/6J mice (The Jackson Laboratory) were used as recipients for bone marrow transplantation assays.

Primary Human MDS and CLL Samples

Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center and conducted in accordance to the Declaration of Helsinki protocol. Patients provided samples after their informed consent and primary human de-identified MDS, AML, and CLL samples derived from whole peripheral blood or BM mononuclear cells were utilized.

Cell Lines

The NALM-6 isogenic cell lines (NALM-6 cells engineered to express the single mutations *SF3B1*^{K700E} or *SF3B1*^{K700K} from the endogenous locus) were cultured in RPMI/10% FCS and K562 isogenic cell lines (engineered to express *SF3B1*^{K700E} or *SRSF2*^{P95H} mutations from each respective endogenous locus) were cultured in IMDM/10% FCS. HAP1 and *CASP8*^{KO} HAP1 cells (obtained from Horizon Discovery) were cultured in IMDM/10% FCS.

METHOD DETAILS

Peripheral Blood Analysis

Blood was collected by submandibular bleeding using heparinized microhematocrit capillary tubes (Thermo Fisher Scientific). Automated peripheral blood counts were obtained using a ProCyte Dx Hematology Analyzer (IDEXX).

Bone Marrow Transplantation Assays

Primary mouse bone marrow (BM) cells were isolated from *Mx1*-Cre^{+} wild-type (WT), *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/P95H}$ or *Mx1*-Cre^{+} *Srsf2*$^{P95H/fl}$ mice (aged 8 weeks) into cold phosphate-buffered saline (PBS), without Ca^{2+} and Mg^{2+}, and supplemented with 2% bovine serum albumin (BSA) to generate single cell suspensions. Red blood cells (RBCs) were removed using ammonium chloride-potassium bicarbonate (ACK) lysis buffer, resuspended in PBS/2% BSA, and filtered through a 40μm cell strainer. Total nucleated cells were quantified by the Vi-Cell XR cell counter (Beckman Coulter). For competitive transplantation experiments, a total of 1.8 x 10^6 BM cells from donor mice (*Mx1*-Cre^{+} WT, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/P95H}$, *Mx1*-Cre^{+} *Srsf2*$^{P95H/fl}$, *Vav*-Cre^{+} WT, *Vav*-Cre^{+} *Srsf2*$^{P95H/+}$, and *Vav*-Cre^{+} *Sf3b1*$^{K700E/+}$ CD45.2^{+} mice were mixed with 0.2 x 10^6 wild-type CD45.1^{+} BM and transplanted via tail vein injection into 8-week old lethally irradiated (2 x 450 cGy) CD45.1^{+} recipient mice. For noncompetitive transplantation experiments, 2 x 10^6 total BM cells from *Mx1*-Cre^{+} WT, *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$, *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$, or *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ mice were injected into lethally irradiated (2 x 450 cGy) CD45.1^{+} recipient mice. To induce the conditional alleles on *Mx1*-Cre background, mice were treated with 3 doses of polyinosinic:polycytidylic acid (pIpC; 12mg/kg/day; GE Healthcare) every other day via intra-peritoneal injection. Peripheral blood chimerism of mature blood cell lineages was assessed routinely by flow cytometry.

In Vivo LPS Stimulation Experiment

For *in vivo* LPS stimulation, *Escherichia coli* 055:B5 LPS (Sigma Aldrich) was used. For chronic LPS exposure, *Vav*-Cre$^+$ WT, *Vav*-Cre$^+$ *Srsf2*$^{P95H/+}$ and *Vav*-Cre$^+$ *Sf3b1*$^{K700E/+}$ received intra-peritoneal injection of LPS (1 mg/kg) every other day for 30 days. For acute LPS exposure, *Mx1*-Cre$^+$ WT, *Mx1*-Cre$^+$ *Srsf2*$^{P95H/+}$ and *Mx1*-Cre$^+$ *Sf3b1*$^{K700E/+}$ mice that have received pIpC for 8 weeks prior to activate the mutant alleles were given a single dose of LPS (15 mg/kg) via intra-peritoneal injection.

In Vitro Colony-Forming Assays

Single-cell suspension was prepared from E14.5 fetal livers, and 25,000 cells from each embryo were plated in duplicates in cytokine-supplemented methylcellulose medium (MethoCult^{TM} GF M3434; StemCell Technologies), and colonies were enumerated 10-14 days later. To assess the effect of shRNA-mediated knockdown of target genes, 8-12 week-old C57BL/6 male mice were treated with 5-fluorouracil (150 mg/kg) via intra-peritoneal injection. Five days after injection, BM cells were harvested from the legs (femora and tibiae) and hip bones, and lineage-depletion was performed with biotin-conjugated antibodies against B220 (RA3-6B2), CD19 (1D3), CD3 (17A2), CD4 (GK1.5), CD8a (53-6.7), CD11b (M1/70), Gr-1 (RB6-8C5), NK1.1 (PK136) and Ter119, labeled with anti-biotin MicroBeads (130-090-485; Miltenyi Biotec), and lineage-negative (Lin$^-$) cells were magnetically separated using MACS columns according to the manufacturer's instructions. Lin$^-$ BM cells were cultured overnight in IMDM/10% FCS supplemented with mIL-3 (10 ng/mL), mIL-6 (10 ng/mL) and mSCF (50 ng/mL). The next day, cells were subjected to spinfection (2,700 rpm for 1 hr) with retroviral supernatants containing shRNAs or cDNAs of interests in the presence of polybrene (5 μg/mL; Millipore). 24 hr after spinfection, cells that were successfully infected with retrovirus were marked with GFP, and were purified by flow cytometry. FACS-sorted cells were cultured in cytokine-supplemented methylcellulose medium (MethoCult^{TM} GF M3434; StemCell Technologies), and colonies were enumerated 10-14 days later.

Flow Cytometry Analyses

Cells were incubated with antibodies in PBS/2% BSA (without Ca^{2+} and Mg^{2+}) for 45 min on ice. For hematopoietic stem and progenitor cell analysis from adult mouse bone marrow, cells were stained with a lineage cocktail of monoclonal antibodies including B220 (RA3-6B2), CD19 (1D3), CD3 (17A2), CD4 (GK1.5), CD8a (53-6.7), CD11b (M1/70), Gr-1 (RB6-8C5), NK1.1 (PK136) and Ter119, allowing for mature lineage exclusion from the analysis. For fetal liver analysis, CD11b (M1/70) was excluded from the lineage depletion cocktail. Cells were also stained with antibodies against c-Kit (2B8), Sca-1 (D7), FcγRII/III (93), CD34 (RAM34), CD45.1 (A20), CD45.2 (104), CD48 (HM48-1) and CD150 (9D1). DAPI was used to exclude dead cells. The composition of mature hematopoietic cell lineages in the bone marrow, spleen, thymus and peripheral blood was assessed using a combination of antibodies against B220, CD19, CD3, CD4, CD8a, CD11b, CD25 (PC61.5), CD44 (IM7), Gr-1, IgM (II/41), CD43 (S11). All FACS sorting was performed on FACS Aria, and analysis was performed on an LSR Fortessa (BD Biosciences). Data analysis was performed using the FlowJo software.

Cell Cycle and Apoptosis Analyses

For apoptosis assays, freshly harvested bone marrow or fetal liver cells were first stained with antibodies against cell surface markers of interests, and then stained with FITC-conjugated Annexin-V in Annexin-V binding buffer (BD Pharmingen) according to manufacturer instructions. For assessment of cell cycle status in adult bone marrow HSPCs, BrdU (1 mg/kg) was administered via intra-peritoneal injection to adult mice 48 hr prior to sacrifice. For cell cycle analysis of E14.5 fetal HSPCs, BrdU (1 mg/kg) was administered to pregnant mice via intra-peritoneal injection 3 hr prior to harvesting fetal livers. Assessment of BrdU incorporation was performed following manufacturer instructions (BD Pharmingen) and data was acquired on a LSR Fortessa (BD Biosciences).

Histological Analysis

Tissues, embryos and pups were fixed in 4% paraformaldehyde, processed routinely in alcohol and xylene, embedded in paraffin, sectioned at 5-micron thickness, and stained with hematoxylin-eosin (H&E). Multiple sections were obtained through the head in the coronal plane, trunk in the transverse plane, and fore and hind limbs in the longitudinal plane. Immunohistochemistry (IHC) was performed on a Leica Bond RX automated stainer (Leica Biosystems, Buffalo Grove, IL). Following HIER at pH 6.0, the primary antibody against mouse CD45 (BD Pharmingen; 550539), myeloperoxidase (Dako; A0398), and cleaved caspase-3 (Cell Signaling; 9661) were applied at a concentration of 1:250, 1:1000, and 1:250 respectively, followed by application of a polymer detection system (DS9800, Novocastra Bond Polymer Refine Detection, Leica Biosystems) in which the chromogen was 3,3 diaminobenzidine tetrachloride (DAB) and counterstain was hematoxylin. For quantification of cleaved caspase-3 by image analysis, whole slide digital images were generated on a slide scanner (Pannoramic 250 Flash III, 3DHistech, 20x/0.8NA objective, Budapest, Hungary) at a resolution of 0.2431 μm per pixel. Staining quantification was performed with QuPath 0.1.2 software (Centre for Cancer Research & Cell Biology, Queen's University Belfast, UK). The region of interest (ROI) was defined as the liver parenchyma. The number of DAB positive cells per mm^2 was measured with the positive cell detection algorithm. ROI selection and algorithm optimization and validation, and qualitative examination of all H&E and IHC slides were performed by a board-certified veterinary pathologist (S.M.).

Immunoblot

For immunoblotting, the following antibodies were used: NF-κB/p65 (CST; 8242), phosphorylated NF-κB/p65-Ser536 (CST; 3033), IκB-α (CST; 9242), TAK1/MAP3K7 (CST; 5206), Caspase-8, N-terminal (Abcam; clone E6), Caspase-8, C-terminal (Enzo Life Science; clone 12F5), Flag (Sigma-Aldrich; F-1804), β-actin (Sigma-Aldrich; A-5441).

Immunofluorescence

Following stimulation with LPS (Sigma-Aldrich), TNFα (PeproTech), or TRAIL (Enzo Life Science), cells were fixed with 4% paraformaldehyde/PBS for 10 min at room temperature (RT), permeabilized with PBS-T (PBS/1% BSA/0.2% Triton-X) for 15 min, blocked with PBS-T/5% goat serum (ThermoFisher PCN500) for 1 hr at RT, and incubated with primary antibodies (1:50 dilution for p-p65 and 1:100 for Flag) in PBS-T/5% goat serum overnight at 4°C. Cells were washed three times with PBS-T for 10 minutes at RT with gentle agitation, and were incubated with goat anti-rabbit Alexa Fluor 594 (ThermoFisher Scientific; A-11012) or goat anti-mouse Alexa Fluor 488 (ThermoFisher Scientific; A-11001) secondary antibodies (1:500 dilution in PBS-T/5/% normal goat serum) for 2 hr at RT in the dark. Cells were then washed twice with PBS-T for 10 min at RT, and counter-stained with DAPI (0.5 μg/mL; Sigma-Aldrich; D-9542) for 20 min at RT. Cells were coverslipped with ProLong Gold anti-fade (ThermoFisher Scientific; P36930). Images were captured using either a confocal microscope (Leica TCS SP5, Upright; Leica MicroSystems), or were digitally scanned with a Pannoramic Confocal Scanner (3DHistech, Budapest Hungary) using a 20x/0.8NA objective. The projected images were exported into tif format using CaseViewer software (3DHistech) and analyzed with ImageJ/FIJI. A macro was written that segments each nucleus using the DAPI channel and measures the p-p65 intensity within the nucleus after appropriate threshold was set.

Caspase-8 and MAP3K7 Constructs

MSCV-Flag-CASP8^{FL}-IRES-GFP, MSCV-Flag-CASP8^{TR}-IRES-GFP, MSCV-Flag-MAK3K7-IRES-GFP (human), MSCV-Flag-Map3k7-IRES-GFP (mouse) and MSCV-IRES-GFP empty vector constructs were used for overexpression studies. Retroviral supernatants were produced by transfecting 293 GPII cells with cDNA constructs and the packaging plasmid VSV.G using XtremeGene9 (Roche), and were used to transduce HAP1, *CASP8*^{KO} HAP1 and K562 parental and *SF3B1*^{K700E} isogenic cells in the presence of polybrene (5 μg/mL; Millipore). Successfully transduced cells expressing GFP were purified by flow cytometry. Cells were stimulated with LPS (Sigma-Aldrich), TNFα (PeproTech) or TRAIL (Enzo Life Science).

mRNA Stability Assay

For mRNA half-life measurement using qRT-PCR, UPF1 shRNA and control lentivirus infected K562 *SF3B1*^{K700E} cells were treated with 2.5 μg/ml Actinomycin D (Life Technologies) and harvested at 0, 2, 4, 6, and 8 hr (following protocols established previously (t Hoen et al., 2011)). MAP3K7 inclusion, MAP3K7 exclusion and 18s rRNA mRNA levels were measured by qRT-PCR.

Luciferase Reporter Assay

We generated K562, NALM-6 *SF3B1* isogenic cells, HAP1 and *CASP8*^{KO} HAP1 cells expressing the luciferase reporter for NF-κB response elements by following the manufacturer instructions (Cignal^{TM} Reporter Assay, Qiagen). Cells were stimulated with LPS, TNFα or TRAIL as described above, and NF-κB activity was assessed by luciferase intensity using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer instructions. To verify that the luciferase reporter assay was not aberrantly activated by basal leakiness, we used another NF-κB reporter plasmids (Promega; N1111) with known NF-κB response elements (RE), and performed mutagenesis in the NF-κB-RE using the QuiKChange II Site-Directed Mutagenesis Kit (Agilent Technologies; #200522). There are five putative NF-κB-RE binding sites in this reporter plasmid (5'-GGGRNTTTCC-3', where R is a purine, Y is a pyrimidine and N is any nucleotide). To mutagenize the binding sequence, the "TTTC" sequence was mutated to "AAAA". The primers used to create the two mutant plasmids are:

Mutant-Fwd

GGTACCTGAGCTCGCTAGCGGGAAAAAACGGGGACAAAACGGGAAAAAACGGGGACAAAACGGGAAAAAACAGATCTGGCCTC GGCGGCCAAGCTTA.

Mutant-Rev

TAAGCTTGGCCGCCGAGGCCAGATCTGTTTTTTCCCGTTTTGTCCCCGTTTTTTCCCGTTTTGTCCCCGTTTTTTCCCGCTAGCGAG CTCAGGTACC.

RT-PCR and Quantitative RT-PCR (qRT-PCR)

Total RNA was isolated using RNeasy Mini kit (Qiagen). For cDNA synthesis, total RNA was reverse transcribed to cDNA with SuperScript VILO cDNA synthesis kit (Life Technologies). The resulting cDNA was diluted 10-20 fold prior to use. Quantitative RT-PCR (qRT-PCR) was performed in 10 μL reactions with either SYBR Green PCR Master Mix or Taqman Gene Expression Master Mix with AmpErase (ThermoFisher Scientific). All qRT-PCR analysis was performed on an Applied Biosystems QuantStudio 6 Flex Cycler (ThermoFisher Scientific). Relative gene expression levels were calculated using the comparative CT method.

Primers used in RT-PCR reactions were:

MAP3K7 (human) – Fwd: GATGGAATATGCTGAAGGGG, Rev: CACTCCTTGGGAACACTGTA

Map3k7 (mouse) – Fwd: GATGGAATATGCAGAGGGGG, Rev: CACTCCTTGGGAACACTGTA

CASP8 (human) – Fwd: GAACTTCAGACACCAGGC, Rev: CTTTGTCCAAAGTCTTTGCTG

Primers used in qRT-PCR reactions were:
CASP8 exclusion isoform (aberrant):
Fwd: GATGAATTTTCAAATGACTTTGGAC
Rev: TGATCAGACAGTATCCCCGAG
CASP8 inclusion isoform (canonical):
Fwd: TGATGAATTTTCAAATGGGGAGGA
Rev: ATCCTGTTCTCTTGGAGAGTCC
MAP3K7 mRNA half-life experiment (human):
Fwd (common): GCGTTTATTGTAGAGCTTCGG
Rev (canonical): GCACCATGCAGCACATTATATAAAG
Rev (aberrant): CATGCAGCACTGCGAAAGAAAG
Taqman probes were used for gene expression analysis of TNF (Hs00174128_m1), IL-1β (Hs01555410_m1), GAPDH (Hs02786624_g1), Mpl (Mm00440310_m1), Pbx1 (Mm04207617_m1), Itga2b (Mm00439741_m1), and Hprt (Mm03024075_m1).

shRNA Experiments

NALM-6 parental cells carrying the NF-κB luciferase reporter were transduced with a doxycycline-inducible lentiviral vector, T3G-dsRED-mirE-PGK-Neo-IRES-rtTA3 (Fellmann et al., 2013), expressing shRNAs for MAP3K7 or a non-targeting renilla or firefly luciferase control. Transduced cells were selected with G418 (0.5 mg/mL; ThermoFisher Scientific), and the short hairpins were induced with the addition of doxycycline (2.0 μg/mL; Sigma Aldrich). All mouse shRNAs used in clonogenic assays were cloned into the retroviral pMSCV-LTR-mirE-PGK-SV40-IRES-GFP (MLS-E) backbone (Fellmann et al., 2013). All shRNAs were designed using the SplashRNA algorithm (Pelossof et al., 2017). The short hairpin sequences are:
sh-MAP3K7.748: TTAGGTAAATTTTTTATCAGTG
sh-MAP3K7.1041: TTTTCAACAATTTTGATTCTAA
sh-Luciferase control: TTAATCAGAGACTTCAGGCGGT
sh-Ren.713 control: CAGGAATTATAATGCTTATCTA
sh-Mpl.2121: TTATATAATAAACAGTGTCTAA
sh-Mpl.2368: TCAAATAAATAGATGACAGCAA
sh-Pbx1.824: TTCATCCAAACTCTGGTCTGTG
sh-Pbx1.1393: TCATTCAGAATTTCTGTGGCTT
sh-Itga2b.2279: TTCTCTTTCTTCTGAGTGCAGA
sh-Itga2b.3380: TTAGGAAAAGGGATGCACCCGG
sh-UPF1 (TRCN0000022254): CCGGGCATCTTATTCTGGGTAATAACTCGAGTTATTACCCAGAATAAGATGCTTTTT

mRNA Isolation, Sequencing, and Analysis

RNA was extracted from sorted mouse cell populations using Qiagen RNeasy columns. Poly(A)-selected, unstranded Illumina libraries were prepared with a modified TruSeq protocol. 0.5X AMPure XP beads were added to the sample library to select for fragments <400 bp, followed by 1X beads to select for fragments >100 bp. These fragments were then amplified with PCR (15 cycles) and separated by gel electrophoresis (2% agarose). 300 bp DNA fragments were isolated and sequenced on the Illumina HiSeq 2000 (∼100M 101 bp reads per sample).

Genome Annotations

Genome annotations for the human (NCBI GRCh37/UCSC hg19) and mouse (NCBI GRCm38/UCSC mm10) genomes were created as previously described (Dvinge et al., 2014). Genome annotations from the Ensembl (Flicek et al., 2013) and UCSC (Meyer et al., 2013) databases were merged with splicing event annotations from MISO v2.0 (Katz et al., 2010). An additional annotation of all possible combinations of annotated 5' and 3' splice sites found in the merged annotation was created for read mapping. Constitutive introns were defined as those whose associated 5' and 3' splice sites were alternatively spliced in the UCSC annotation.

RNA-seq Read Mapping

RNA-seq reads were sequentially mapped to the transcriptome and genome as previously described (Dvinge et al., 2014). Reads were first mapped to the transcriptome using Bowtie v1.0.0 (Langmead et al., 2009) and RSEM v1.2.4 (Li and Dewey, 2011). The resulting read alignments were then filtered to require that reads spanning splice junctions overlapped the flanking exons by at least six nt. The remaining unaligned reads were then mapped to the genome and splice junctions using TopHat v2.0.8b(Trapnell et al., 2009), where reads were only allowed to align to the splice junctions present in the file of all possible combinations of annotated 5' and 3' splice sites described above. The resulting read alignments were then merged with the output of RSEM's alignment to create a final file of aligned reads.

Differential Gene Expression Analysis

Gene expression analysis was performed using the gene expression estimates computed by RSEM in units of transcripts per million (TPM). Those estimates were then further normalized using the TMM method (Robinson and Oshlack, 2010), with a reference set of all protein-coding genes. Differentially expressed genes were defined as those with an associated Bayes factor $\geq$ 100 (computed using Wagenmakers's Bayesian framework (Wagenmakers et al., 2010)) and an associated fold-change $\geq$ 1.5.

Differential Splicing Analysis

Isoform ratios for annotated splicing events (cassette exons, competing 5' and 3' splice sites, and annotated retained introns) were calculated using MISO v2.0 (Katz et al., 2010). Splicing of constitutively spliced introns and junctions was quantified using only junction-spanning reads, as previously described (Hubert et al., 2013). Differential splicing in the murine data was identified by comparing samples from different genotypes for a single replicate in a pairwise fashion. The analysis was restricted to splicing events with at least 20 informative reads, where an informative read is defined as a read that distinguishes between isoforms. Differentially spliced events were defined as those with an associated Bayes factor was $\geq$ 5 (computed using Wagenmakers's Bayesian framework (Wagenmakers et al., 2010)) and absolute change in isoform ratio of $\geq$ 10%. Differential splicing in the human patient cohorts was identified using a group statistical test to identify differences between patient samples with or without defined splicing factor mutations. Differentially spliced events were defined as those with an associated p value $\leq$ 0.05 (computed using the Wilcoxon rank-sum test) and an absolute change in median per-group isoform ratio of $\geq$10%.

Gene Ontology (GO) Enrichment Analysis

GO enrichment analysis was performed using the GOseq method (Young et al., 2010) to correct for sequencing depth biases. The background set of genes was defined as all protein-coding genes. False discovery rates were calculated using the Wallenius method and corrected using the Benjamini-Hochberg method. We restricted reporting of enriched terms to those with at least two ancestors and fewer than 500 associated genes.

Motif Enrichment Analysis and Sequence Logos

The relative enrichment of different ESEs was computed by comparing motif occurrence within and adjacent to cassette exons that were promoted versus repressed in cells or samples with versus without defined splicing factor mutations. These analyses were performed using all cassette exons that were differentially spliced in at least one mouse replicate for a given genotype comparison. 95% confidence intervals were calculated by bootstrapping with 500 resampling steps. Sequence logos centered on intron-proximal or intron-distal 3' splice sites were created using all competing 3' splice sites that were differentially spliced in at least one mouse replicate for a given genotype comparison. The analysis was restricted to events with canonical GT and AG dinucleotides at the 5' and 3' splice sites. These analyses relied upon the GenomicRanges package in Bioconductor (Huber et al., 2015).

Analysis of Expected and Observed Gene Expression Convergence

Expected gene expression programs for *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ *Sf3b1*$^{K700E/+}$ samples were calculated by computing the mean of individual gene expression levels between *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ and *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$ samples for each replicate. This method of computing the expected gene expression program was motivated by the assumption that mutations in *Srsf2* and *Sf3b1* have independent consequences for individual gene expression. Gene dysregulation for the expected gene expression program was then calculated as described previously for the actual, observed gene expression values.

The numbers of dysregulated genes that were convergent between the expected gene expression program and the two *Mx1*-Cre^{+} *Srsf2*$^{P95H/+}$ and *Mx1*-Cre^{+} *Sf3b1*$^{K700E/+}$ single mutants was then determined for each replicate, again using a method identical to that for the actual, observed gene expression values (Figure S4A). A one-sided binomial proportion test was then used to test whether there were more dysregulated genes that were observed than expected (Figure S4B).

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical Analyses

All data are presented as mean±standard deviation, unless otherwise stated. The replicate for each experiment was stated in the figure legend or indicated in the figure. Statistical significance was determined by analysis of variance (ANOVA) after testing for normal distribution and equal variance, followed by Tukey's post-hoc test for multiple group comparisons. A p value of <0.05 was considered statistically significant. For non-normally distributed data, a non-parametric test (Kruskal-Wallis) was used, followed by multiple group comparisons using false-discovery rate (FDR). For Kaplan Meier survival analysis, Mantel-Cox log-ranked test was used to determine statistical significance. For offspring frequency analysis, a Chi-Square test was performed to test the difference between observed and expected frequencies from different genotypes. No blinding or randomization was used. Unless otherwise noted, all immunoblot quantitation and immunofluorescence image quantitation were representative of at least three biological replicates from independent experiments. Data were plotted using GraphPad Prism 7 software.

DATA AND SOFTWARE AVAILABILITY

Publicly Available RNA-seq Data

FASTQ files from published RNA-seq studies of patients with MDS (Dolatshad et al., 2015) and CLL (Darman et al., 2015) were downloaded from GEO series GSE63569 and GSE72790.

Accession Codes

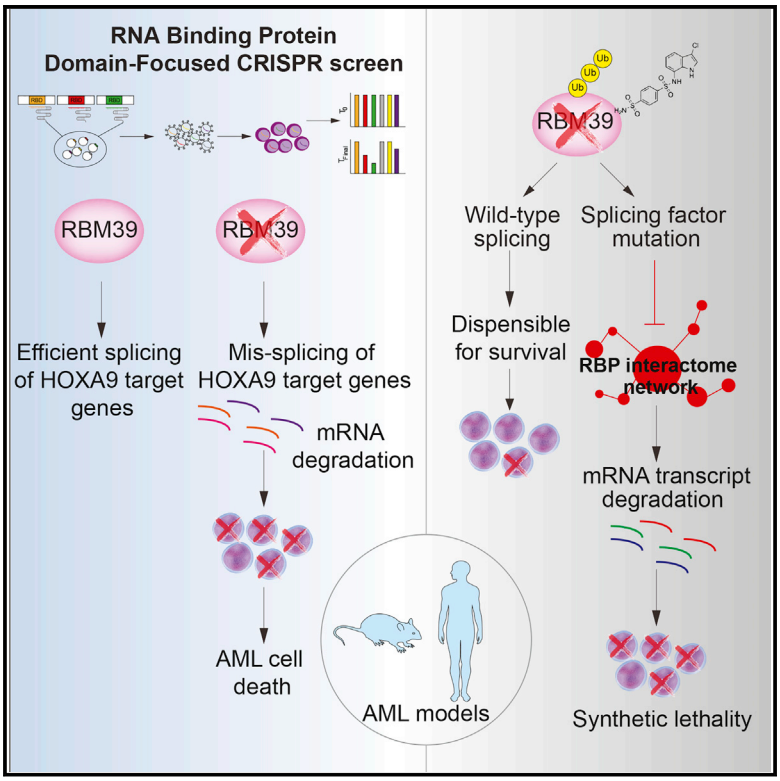
Gene Expression Omnibus: The accession number for all newly generated RNA-seq data reported in this paper are deposited into the GEO database (accession number GSE97452).

Article

Cancer Cell

Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia

Graphical Abstract



Highlights

- CRISPR/Cas9 domain screen reveals RBP dependencies in AML
- RBM39 is required for AML maintenance through mis-splicing of HOXA9 target genes
- Proteomic studies identify an essential RBP splicing network in AML
- Pharmacologic RBM39 degradation leads to broad anti-leukemic effects

Authors

Eric Wang, Sydney X. Lu, Alessandro Pastore, ..., Adrian R. Krainer, Omar Abdel-Wahab, Iannis Aifantis

Correspondence

abdelwao@mskcc.org (O.A.-W.), iannis.aifantis@nyumc.org (I.A.)

In Brief

Using a CRISPR/Cas9 screen targeting RNA-binding domains of classical RNA-binding proteins (RBPs), Wang et al. uncover a network of interacting RBPs upregulated in acute myeloid leukemia (AML) and crucial for RNA splicing and AML survival, highlighting RBM39 as a central, targetable component of the RBP network.

Wang et al., 2019, Cancer Cell *35*, 369–384
March 18, 2019 © 2019 Elsevier Inc.
https://doi.org/10.1016/j.ccell.2019.01.010

CellPress

Cancer Cell
Article

Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia

Eric Wang,[1,5] Sydney X. Lu,[2,5] Alessandro Pastore,[2,5] Xufeng Chen,[1] Jochen Imig,[1] Stanley Chun-Wei Lee,[2] Kathryn Hockemeyer,[1] Yohana E. Ghebrechristos,[1] Akihide Yoshimi,[2] Daichi Inoue,[2] Michelle Ki,[2] Hana Cho,[2] Lillian Bitner,[2] Andreas Kloetgen,[1] Kuan-Ting Lin,[3] Taisuke Uehara,[4] Takashi Owa,[4] Raoul Tibes,[1] Adrian R. Krainer,[3] Omar Abdel-Wahab,[2,*] and Iannis Aifantis[1,6,*]

[1]Department of Pathology and Laura & Isaac Perlmutter Cancer Center, NYU School of Medicine, New York, NY 10016, USA
[2]Human Oncology and Pathogenesis Program and Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[3]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA
[4]Tsukuba Research Laboratories, Eisai Company, Ltd, Tsukuba, Ibaraki 300-4352, Japan
[5]These authors contributed equally
[6]Lead Contact
*Correspondence: abdelwao@mskcc.org (O.A.-W.), iannis.aifantis@nyumc.org (I.A.)
https://doi.org/10.1016/j.ccell.2019.01.010

SUMMARY

RNA-binding proteins (RBPs) are essential modulators of transcription and translation frequently dysregulated in cancer. We systematically interrogated RBP dependencies in human cancers using a comprehensive CRISPR/Cas9 domain-focused screen targeting RNA-binding domains of 490 classical RBPs. This uncovered a network of physically interacting RBPs upregulated in acute myeloid leukemia (AML) and crucial for maintaining RNA splicing and AML survival. Genetic or pharmacologic targeting of one key member of this network, RBM39, repressed cassette exon inclusion and promoted intron retention within mRNAs encoding HOXA9 targets as well as in other RBPs preferentially required in AML. The effects of RBM39 loss on splicing further resulted in preferential lethality of spliceosomal mutant AML, providing a strategy for treatment of AML bearing RBP splicing mutations.

INTRODUCTION

Eukaryotic cells employ a wide range of mechanisms to regulate the fine-tuning of mRNA expression (Glisovic et al., 2008). These co- and post-transcriptional processes are orchestrated by interactions between RNA molecules and RNA-binding proteins (RBPs). RBPs are a diverse class of proteins containing unique RNA-binding domains (RBDs) (Cook et al., 2011). Classical RBPs encompass one or multiple conventional RBDs that include, but are not limited to, RNA-recognition motifs (RRMs), K-homology (KH) domains, and DEAD-box domains, which have been extensively determined structurally and biochemically (Lunde et al., 2007). Considering that RBPs are key regulators of gene expression, alterations of these proteins are also implicated in several human genetic diseases, including multiple cancers (Cooper et al., 2009; Kapeli et al., 2017; Lukong et al., 2008).

Acute myeloid leukemia (AML) is an aggressive hematologic malignancy with a dismal survival rate (<30% 5-year overall survival rate), highlighting the need for improved therapeutic interventions (Dohner et al., 2015). Advances in next-generation sequencing have exposed the genetic and epigenetic heterogeneity in AML pathogenesis (Papaemmanuil et al., 2016). Analyses of the genomic landscape in myeloid leukemia patients have discovered several somatic genetic lesions, including frequent mutations in RBPs that serve as splicing factors (Yoshida et al., 2011). Although the mechanisms by which

Significance

RNA-binding proteins (RBPs) regulate many aspects of transcription and translation and, as such, are thought to elicit cell- and tissue-type specific functions. Here, through systematic evaluation of RBPs across several cancer types, we identify RBPs specifically required in individual forms of cancer. In so doing we identify a network of functionally and physically interacting RBPs upregulated in AML over normal hematopoietic precursors and required for AML maintenance. Pharmacologic degradation of one such RBP, RBM39, led to aberrant splicing of multiple members of this RBP network as well as of transcriptional regulators required for AML survival. These data therefore identify RBPs with cancer-specific roles and illuminate a therapeutic approach targeting RBPs required for AML maintenance.

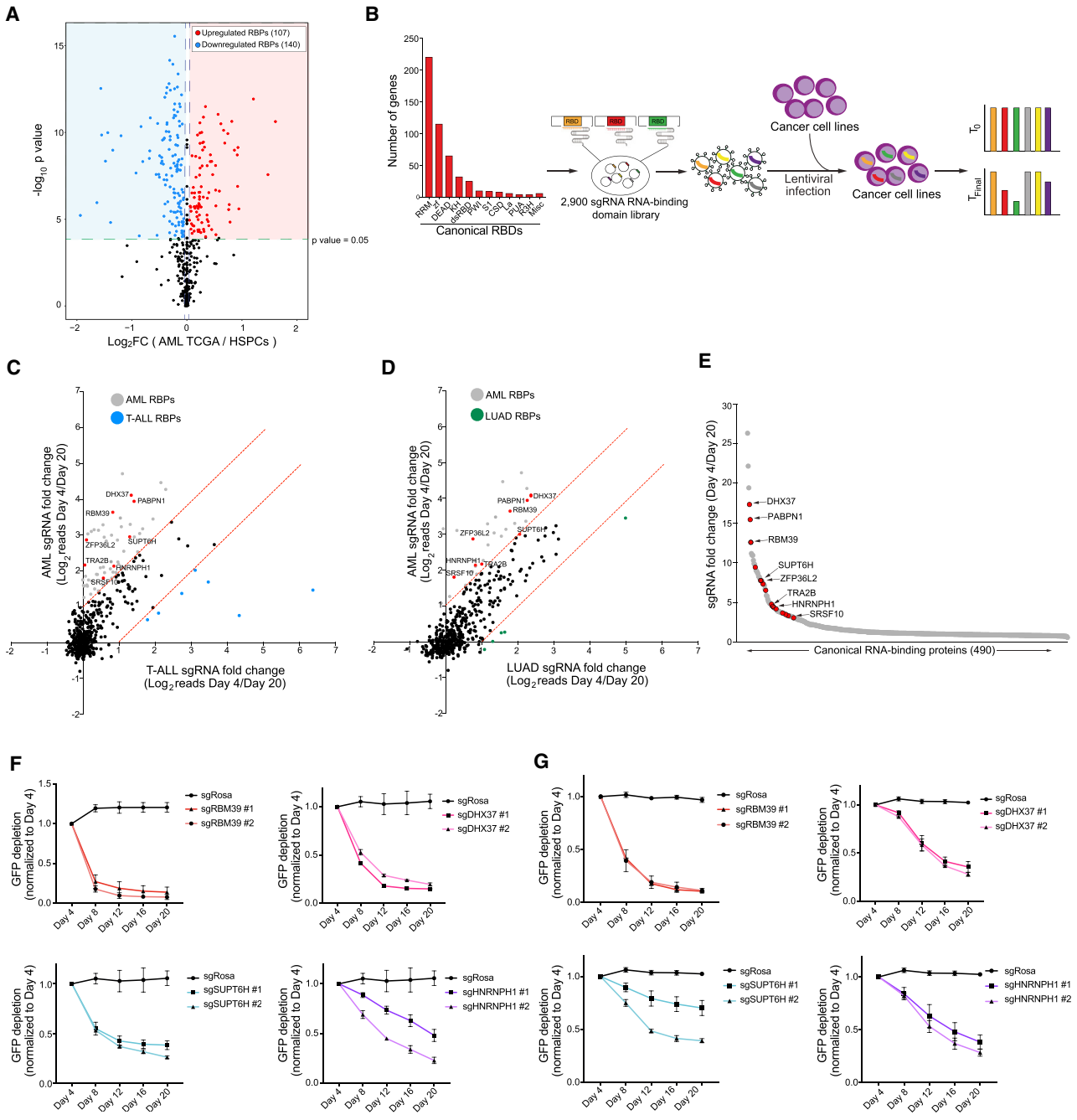


Figure 1. A CRISPR Domain-Targeted Screen Identifies RBP Dependencies in AML

(A) Volcano plot of differentially expressed classical RBPs (483 genes) in AML patient samples (n = 195) compared with normal human CD34$^+$ hematopoietic stem and progenitor cells (n = 28). Blue vertical lines indicate $\log_2$ fold change (FC) = 0.5/−0.5 cutoff ($p > 0.05$).

(B) Schematic depicting pooled RBP-focused CRISPR screen. RBD, RNA-binding domains; RRM, RNA-recognition motifs; zf, zinc finger; DEAD, DEAD-box ATPase; KH, KH homology; dsRBD, double-stranded RBDs; PWI, PWI motif; CSD, cold shock; La, La motif; PUA, PseudoUridine synthase and archaeosine transglycosylase; R3H, R-x3-H domain.

(C) Scatterplot comparison of CRISPR RBP-domain screen in AML and T-ALL. Plotted is the $\log_2$ fold change of sgRNA abundance (day 4/day 20) for each cell line. Each dot represents the average of all sgRNAs targeting an RBP. Red dots indicate RBPs that are significantly overexpressed in AML and preferentially required in AML.

(D) Scatterplot representation of CRISPR RBP-domain screen in AML and lung adenocarcinoma (LUAD). Red dots indicate RBPs that are significantly over-expressed in AML and preferentially required in AML.

(E) Fold change (day 4/day 20) in sgRNA abundance in pooled RBP-focused negative selection screen in MOLM-13 AML cells. Red dots indicate RBPs that are overexpressed in AML patients and exhibit greater than 3-fold depletion in CRISPR negative selection screen. Each dot represents the average of all sgRNAs targeting an RBP.

(legend continued on next page)

spliceosome mutations promote transformation are still being investigated, several studies have demonstrated that such mutations alter RNA-binding preferences to promote altered splicing as key in the molecular pathogenesis of these mutations (Dvinge et al., 2016). AML cells bearing splicing factor mutations exhibit increased sensitivity to pharmacologic modulation of splicing, thereby providing a therapeutic strategy for patients harboring spliceosome mutations (Lee and Abdel-Wahab, 2016; Lee et al., 2016; Obeng et al., 2016; Seiler et al., 2018). This has led to an ongoing phase I clinical trial to determine the safety and efficacy of a modulator of core splicing function H3B-8800 in patients with refractory leukemias (Seiler et al., 2018).

In the past few years, there has also been a deeper understanding of "non-oncogene addiction" of RBPs in cancers, which is the scenario in which RBPs are essential for cancer cell survival but are not targeted by genetic mutations. A few such RBP dependencies have already been identified by studying RBPs that are upregulated in diverse tumors (Anczukow et al., 2012; Fox et al., 2016; Wurth et al., 2016). Despite recent progress in uncovering essential RBPs in leukemia maintenance (Kharas et al., 2010; Palanichamy et al., 2016), the promise of targeting RBPs therapeutically is still limited by a lack of systematic evaluation for required RBPs in cancer. Here, our goal was to use functional screens to dissect the roles of classical RBPs in AML.

RESULTS

Aberrant RBP Expression in AML Patient Samples

Aberrant RBP expression has been commonly linked to promoting cancer progression through co- and post-transcriptional mechanisms. Analysis of datasets from The Cancer Genome Atlas (TCGA) uncovered numerous differentially expressed RBPs across various solid cancers, suggesting that RBPs are often dysregulated in tumors (Sebestyen et al., 2016). To define the extent to which RBP genes are altered in AML, we analyzed the mRNA expression of 484 classical RBPs in AML patients using publicly available data from the TCGA and Leucegene (Cancer Genome Atlas Research Network et al., 2013; Lavallee et al., 2015a). Comparison of RBP mRNA expression in AML patients with normal human $CD34^+$ human hematopoietic stem/progenitor cells identified approximately 51% (247/484) of the genes to be differentially expressed (Figure 1A and Table S1). Among the dysregulated RBPs in AML, we identified 107 RBP genes that were significantly upregulated ($p < 0.05$), whereas 140 genes were found to be downregulated. These data support previous findings that aberrant expression of RBPs occurs ubiquitously in cancers (Sebestyen et al., 2016).

A CRISPR/Cas9 Domain-Targeted Screen Reveals RBP Dependencies in AML

The majority of cancer dependencies can be predicted based on gene expression and therefore serve as clinically relevant biomarkers or therapeutic targets (Tsherniak et al., 2017). Based on our TCGA analysis, we hypothesized that several of these dysregulated RBPs may act as genetic vulnerabilities in AML and represent potential therapeutic targets. To investigate potential RBP dependencies in AML, we used the type II CRISPR (clustered, regularly interspaced, short palindromic repeat) system, which has been widely adopted in high-throughput methods to identify essential genes for cancer survival (Shalem et al., 2014; Wang et al., 2017). Given the established biochemical function of RBDs in RNA biology, we used a previously described single-guide RNA (sgRNA) domain-focused approach that enhances CRISPR/Cas9-negative selection by targeting functional protein domains (Shi et al., 2015). As a proof of principle, we evaluated domain essentiality by designing sgRNAs against the coding exons of three classical RBPs required to sustain cell growth based on a previous genome-wide CRISPR/Cas9 knockout screen (Hart et al., 2015). Individual sgRNAs were subcloned into a lentivirus-based GFP-tagged (LRG) sgRNA vector and cell proliferation was measured using competition assays. Notably, CRISPR scanning of the selected RBPs led to robust negative selection of sgRNAs that specifically targeted the RBDs in comparison with non-domain exon sgRNAs, which demonstrated modest depletion (Figure S1A). Based on these findings, we built a custom 2,900 sgRNA (~6–8 per gene) library against 490 RBPs composed of well-defined RBDs based on previous studies (Gerstberger et al., 2014; Lunde et al., 2007) (Figure 1B). Pooled sgRNAs were subcloned into the LRG vector and next-generation sequencing was carried out to confirm optimal sgRNA representation (Figure S1B). Subsequently, we performed a loss-of-function pooled screen in a Cas9-expressing AML cell line, MOLM-13 (MLL-AF9, $FLT3^{ITD}$), to identify RBPs required for AML proliferation. Genomic DNA was harvested from cells on day 4 and day 20 post transduction of sgRNA library, and individual sgRNA read counts were evaluated by deep sequencing. Changes in sgRNA abundance were assessed by measuring the average fold change (day 4/day 20) of all sgRNAs targeting a given gene. This negative selection screen identified 71 RBPs that strongly dropped out with >3-fold depletion. These also included positive controls of domain-targeting sgRNAs against known chromatin regulators (*BRD4*, *DOT1L*, and *KMT2D*) required for AML survival and the essential gene, *RPA3*, thereby confirming the quality of our negative selection screen (Figure S1C). In parallel, we performed counter-screens in three other cancer cell lines, a T cell acute lymphoblastic leukemia (T-ALL) line (CUTLL-1), a lung adenocarcinoma line (A549), and a melanoma line (501MEL), to delineate potential cancer lineage-specific RBPs. Pairwise comparison of domain CRISPR/Cas9 screens revealed several RBP dependencies unique to each cell line (Figures 1C, 1D, and S1D). We narrowed our focus on RBPs that were >2-fold depleted in our AML screen relative to T-ALL, melanoma, and lung adenocarcinoma (LUAD) to identify AML RBP candidates. Using this criterion, we uncovered 23 RBPs preferentially required in AML.

(F and G) Screen validation of RBP candidates using a competition-based proliferation assay in MOLM-13 (F) and THP-1 (G) AML cell lines. Plotted are GFP percentages measured during 20 days in culture and normalized to day 4. Negative control, sgRosa, and two independent sgRNAs targeting each RBP are shown in the graphs (mean ± SD, n = 3).
See also Figure S1 and Table S1.

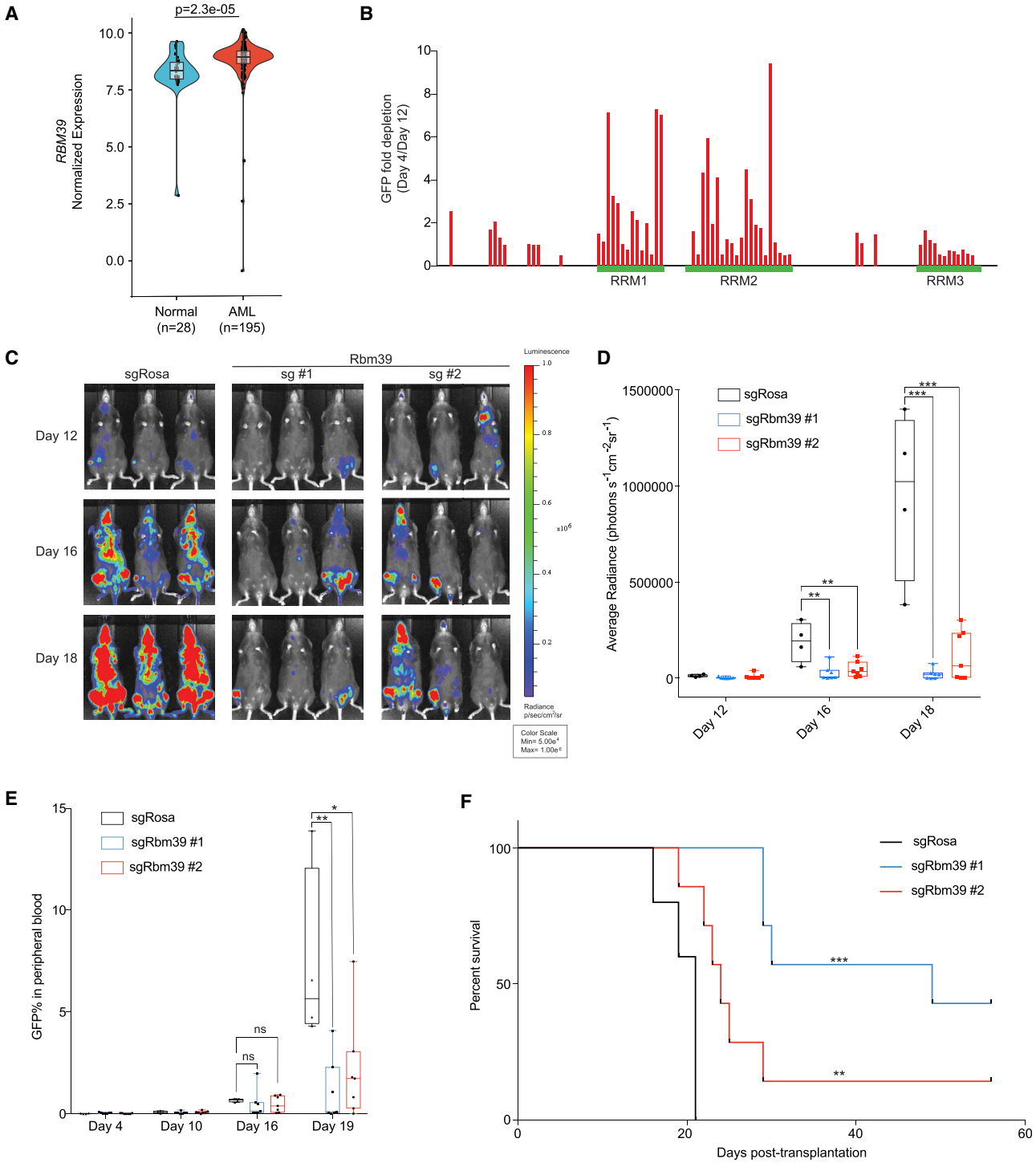


Figure 2. RBM39 is Required to Sustain AML Growth *In Vitro* and *In Vivo*

(A) Violin plot of *RBM39* normalized expression in AML patients (red) and normal human CD34^{+} hematopoietic stem and progenitor cells (blue). Horizontal line inside the box represent the mean with 25th–75th percentiles, showing all data points. Statistical analysis was performed using Wilcoxon rank-sum test.

(B) CRISPR mutagenesis of *RBM39* exons in MOLM-13 AML cells using a competition-based assay. Green boxes represent annotated RNA-binding domains of RBM39.

(C) Bioluminescent images of mice transplanted with MLL-AF9 Nras^{G12D} cells transduced with sgRosa (n = 4) or two independent sgRbm39 (n = 7/group). Representative images of three mice per sgRNA construct is shown. The same mice are depicted at each time point.

(legend continued on next page)

We next integrated the domain CRISPR screen with our transcriptome analysis in AML patients to identify RBPs that are both required for AML survival and dysregulated in expression in AML. From this analysis, we identified genes encoding 21 RBPs that were among the most highly depleted (>3-fold depletion) in our AML screen and significantly overexpressed in patient samples ($p < 0.05$) (Figure 1E). Based on our CRISPR counter-screens, we found 8 of the 21 RBP candidates (*RBM39*, *DHX37*, *PABPN1*, *ZFP36L2*, *TRA2B*, *SRSF10*, *HNRNPH1*, and *SUPT6H*) to be selectively required and upregulated in AML (Figures 1C, 1D, and S1D). To further validate these findings, we evaluated four candidates (*RBM39*, *DHX37*, *SUPT6H*, and *HNRNPH1*) by selecting the top sgRNAs for each gene and monitored their ability to inhibit AML cell growth using a competition-based assay. Our results confirmed that all sgRNAs demonstrated robust depletion in two independent AML cell lines, MOLM-13 and THP-1 (MLL-AF9, $NRAS^{mut}$) (Figures 1F and 1G). We also performed gene ontology (GO) analysis focusing on RBPs that were at least 5-fold depleted in our AML screen. Most of the genes significantly clustered into key RNA biological processes that included mRNA export, splicing, and mRNA processing (Figure S1E). Finally, as some of the top-ranking depleted RBPs were also considered core fitness genes (Hart et al., 2015) (Figure S1F), we eliminated them to increase the therapeutic significance of our findings.

RBM39 Is Required to Sustain Leukemia Survival *In Vitro* and *In Vivo*

One of the top-scoring candidates that met the above criteria was *RBM39*, also known as *CAPER-α*, a protein previously characterized as interacting with U2AF65 and SF3B1 (Imai et al., 1993; Loerch et al., 2014; Stepanyuk et al., 2016). Our expression studies were able to show that *RBM39* was significantly more highly expressed in AML patient samples when compared with normal hematopoietic cells (Figure 2A). *RBM39* median expression in AML was highest among all other cancer subtypes in the TCGA. We did not observe any significant differences in *RBM39* expression across molecular or cytogenetics AML risk groups, suggesting a potential unique requirement for RBM39 in this type of leukemia (Figures S2A and S2B). To further evaluate the requirement across different cancers for *RBM39*, we performed competition assays using RNAi or CRISPR/Cas9 to target RBM39 in various cancer cell lines. Our one-by-one validation revealed that AML cell lines, across a variety of molecular subtypes, were the most susceptible to growth inhibition upon loss of RBM39 (Figures 1F, 1G, and S2C–S2E). Additionally, RBM39 suppression in AML led to marked induction of apoptosis, as determined by Annexin V staining (Figure S2F). In contrast, RBM39 was relatively more dispensable for growth in non-AML cell lines (Figures 1C, 1D, S1D, and S2G). Given the evident requirement of RBM39 in AML pathogenesis, we also employed CRISPR/Cas9 "domain scanning" to identify essential RBM39 protein domains for future drug discovery efforts. We designed sgRNAs with low off-target scores to perform a CRISPR scanning of *RBM39* coding exons in two independent human AML cell lines. These include sgRNAs targeting the three RRM domains and the serine/arginine-rich (RS) region of RBM39, which are critical for pre-mRNA splicing (Prigge et al., 2009). Our results revealed strong depletion of sgRNAs that exclusively targeted RRM1 and RRM2 domains, whereas RRM3 and non-domain sgRNAs demonstrated lesser negative selection (Figures 2B and S2H). Overall, these findings demonstrate that RBM39 relies on specific RBDs and supports a critical RBM39 dependency in AML.

We next examined the *in vivo* significance of Rbm39 in AML progression by using a previously established MLL-AF9 $Nras^{G12D}$ mouse model (RN2) (Shi et al., 2015; Zuber et al., 2011a). We designed sgRNAs targeting the RRM domains of *Rbm39* and confirmed knockout of Rbm39 protein in RN2 cells (Figure S2I); subsequently sgRNA-expressing RN2 cells were injected intravenously into sublethally irradiated recipient mice. Strikingly, we observed remarkable delay in leukemia progression in mice receiving Rbm39-deficient RN2 cells, as determined by bioluminescent imaging (Figures 2C and 2D). Moreover, fluorescence-activated cell sorting (FACS) analysis of peripheral blood in late-disease-onset mice detected a low percentage of circulating leukemia cells harboring Rbm39 sgRNAs (Figures 2E and S2J). These findings correlated with prolonged absolute survival (Figure 2F). Indel analysis by next-generation sequencing verified that sgRNA-positive cells mostly lacked CRISPR/Cas9 editing of *Rbm39*, suggesting that the mice succumbed to an outgrowth of unmodified alleles and substitution mutations (Figures S2K–S2L). Taken together, these observations support the notion that Rbm39 is required for AML growth *in vivo*.

Mapping of the RBM39 Proteome Identifies an Essential AML Splicing Network

To identify RBM39 interaction partners in AML cells, we performed immunoprecipitation followed by mass spectrometry (IP-MS) in the AML cell line MOLM-13 (Figures 3A and S3A). We identified a total of 54 RBM39-interacting proteins that were at least 10-fold more enriched over immunoglobulin G (IgG) (Figure 3B and Table S2). Proteomic network analysis identified numerous proteins associated with the spliceosome complex and ribosome biogenesis (Figure 3C). This was in agreement with GO analysis, which identified proteins highly enriched in fundamental processes involved in RNA metabolism

(D) Quantification of bioluminescent imaging in sgRosa-negative control and two independent sgRbm39 at the indicated time points. Box-and-whisker plot, minimum to maximum showing all points, 25^{th}–75^{th} percentiles, and median (horizontal line). Statistical analysis was performed using unpaired Student's t test by GraphPad Prism (**$p < 0.01$, ***$p < 0.001$).

(E) Flow-cytometry analysis of GFP-positive sgRNA-expressing leukemia cells in peripheral blood of MLL-AF9 $Nras^{G12D}$ leukemia recipient mice at indicated time points. Box-and-whisker plot, minimum to maximum showing all points, 25^{th}–75^{th} percentiles, and median (horizontal line). Statistical analysis was performed using unpaired Student's t test by GraphPad Prism (*$p < 0.05$, **$p < 0.01$; ns, not significant).

(F) Kaplan-Meier survival curves of recipient mice transduced with sgRosa-negative control and two independent Rbm39 sgRNAs are plotted. The p values were determined using a log-rank Mantel-Cox test (**$p < 0.01$, ***$p < 0.001$).

See also Figure S2.

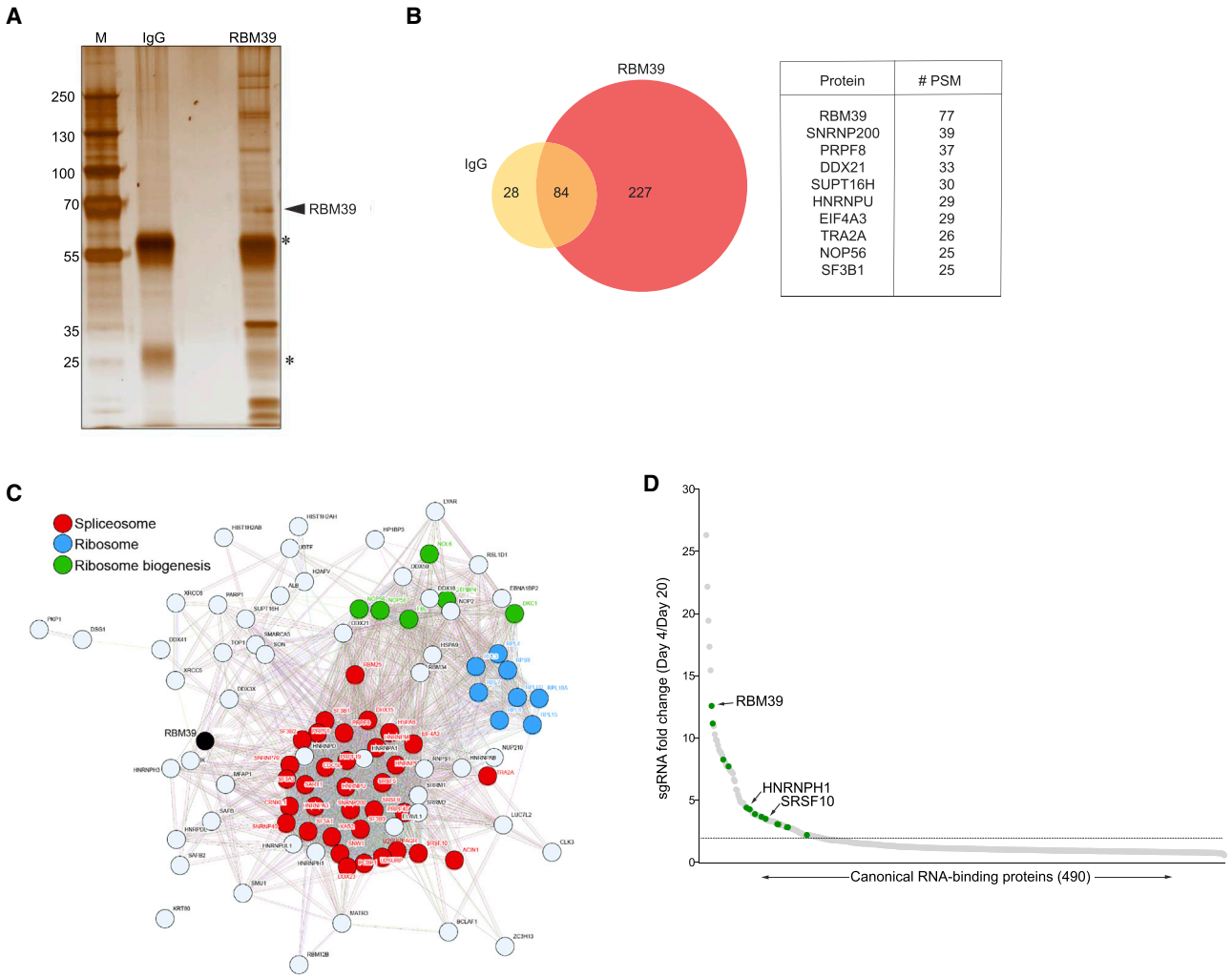


Figure 3. Mass Spectrometry of RBM39 Proteome Identifies an RBP Network Required in AML

(A) Silver staining of endogenous RBM39 immunoprecipitation (IP) in MOLM-13 cells. Arrowhead represents predicted size of RBM39, asterisks denote non-specific IgG heavy-/light-chain signals, and M is protein marker.

(B) Overlay of peptide-spectrum match counts (#PSM) between IgG control and RBM39 IP (left panel). Top RBM39-interacting partners enriched over IgG (right panel).

(C) STRING network analysis of RBM39 protein interactome in AML. Black circle, RBM39; red circles, spliceosome complex; blue circles, ribosome; green circles, ribosome biogenesis.

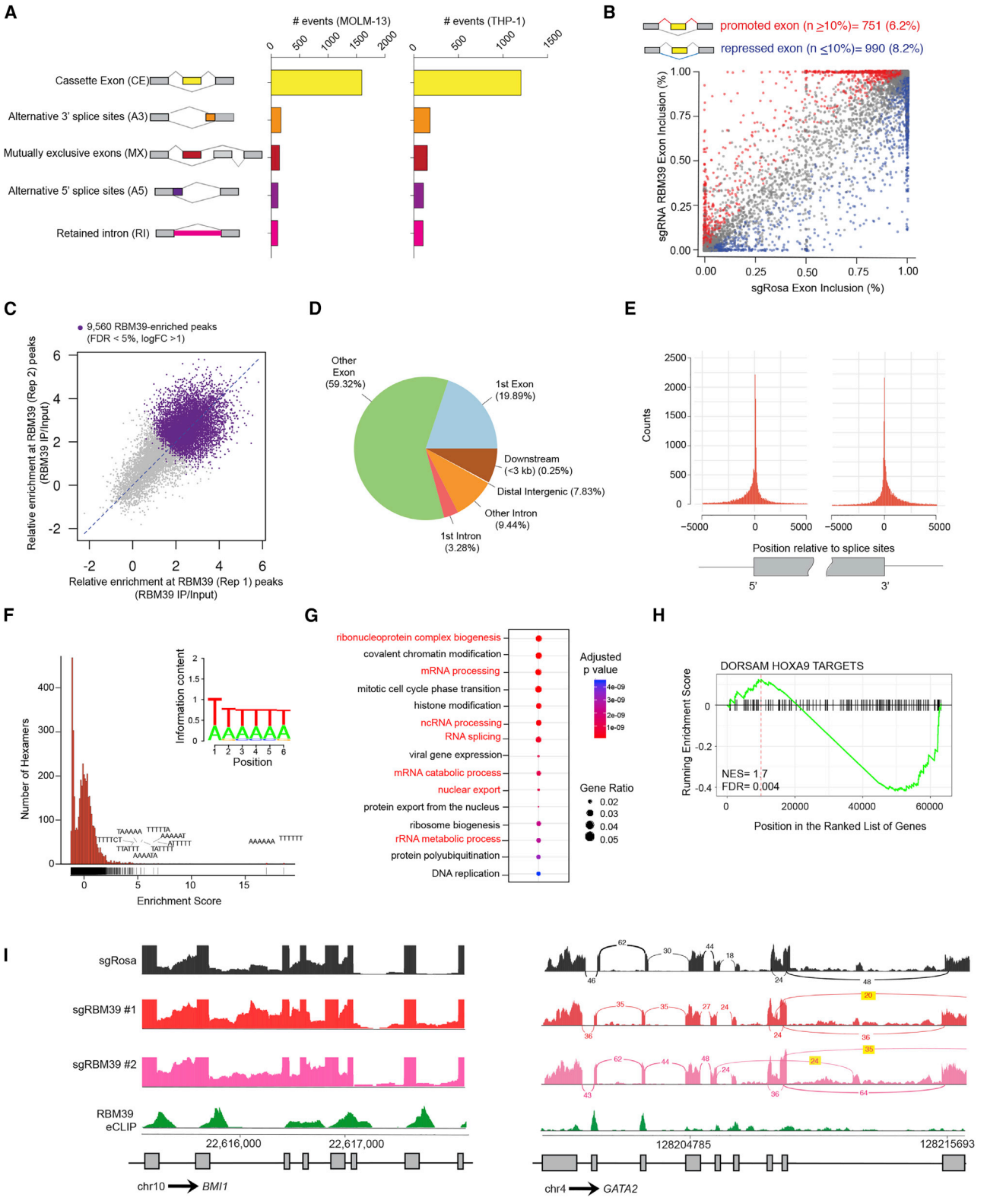
(D) Plot represents data from MOLM-13 AML screen as shown in Figure 1E. Green denotes RBP genes identified in RBM39 IP-MS and overexpressed in TCGA AML patients. Dotted line represents 2-fold depletion.

See also Figure S3 and Table S2.

and ribosomal RNA processing (Figure S3B). In addition, we also found several RBM39-binding partners involved in chromatin and transcriptional regulation (e.g., histone chaperone, SUPT16, and the SWI/SNF-related subunit, SMARCA5), thereby suggesting possible roles for RBM39 in gene transcription. Moreover, our proteomics approach identified 31 RBM39-interacting RBPs that also appeared in our domain CRISPR screens. We found 15 of the 31 RBPs to exhibit strong essentiality (>2-fold depletion) in our genetic screen in AML (Figure 3D). Among these RBM39 interacting partners, the splicing factors *SRSF10* and *HNRNPH1* were identified in our CRISPR screen to be preferentially required in AML and also highly expressed in AML patient samples (Figure 3D). Altogether, these data place RBM39 in an extended RBP network that is essential for leukemia maintenance and suggest that modulation of AML splicing can be therapeutically exploited by targeting the RBM39 interactome.

RBM39 Loss Alters Splicing of mRNAs Essential for AML

Given the physical interaction of RBM39 with key core splicing factors and prior data identifying splicing changes upon RBM39 loss (Han et al., 2017; Uehara et al., 2017), we hypothesized that loss of RBM39 might disrupt splicing and expression of mRNAs preferentially required for leukemia growth. To this end, we initially evaluated changes in splicing by RNA



(legend on next page)

sequencing (RNA-seq) of human AML cells upon RBM39 sgRNA deletion. We measured "percentage spliced in" (ΔPSI) values across five main types of alternative splicing (AS) events (cassette exon [CE], alternative 5′ ss exon [A5E], alternative 3′ ss exon [A3E], mutually exclusive exons [MXE], and retained introns [RI]) in *RBM39* knockout versus control MOLM-13 and THP-1 AML cells. This revealed that the predominant change in RNA splicing upon RBM39 depletion was a change in CE event splicing (68%–74% of differentially spliced events) across both AML cell lines with a false discovery rate (FDR) of <0.1 and ΔPSI >10% (Figure 4A; Tables S3 and S4). Using the same parameters, we also found alterations in MXE (~6%–9%), A5E (5%–6%), A3E (8%–10%), and RI (5%–6%) events, albeit to a much lesser extent compared with CE (Figure 4A). Within CEs, we found approximately equal proportions of exon inclusion and exclusion with *RBM39* genetic loss, which was also observed in other AS groups (Figure 4B).

The direct impact of RBM39-RNA interactions on pre-mRNA splicing has not been explored in cancer, nor is the precise role of RBM39 in splicing well defined. We thus performed anti-RBM39 enhanced UV crosslinking immunoprecipitation (eCLIP) using the MOLM-13 AML cell line to identify genome-wide RNA targets of RBM39 (Van Nostrand et al., 2016) (Figure S4A). eCLIP analysis identified 9,560 significant sequence clusters using an FDR of <0.05 and log(FC) > 1, which corresponded to 4,775 annotated transcripts (Figure 4C and Table S5). GO analysis of the top 200 enriched binding sites showed processes involved in RNA metabolism, cell cycle, and transcriptional regulation (Figure S4B). Approximately 79% of RBM39-binding sites mapped to exons (Figures 4D and S4C), with a preferential occupancy of proximal exonic sequences near 5′ and 3′ splice sites throughout the transcriptome (Figures 4E and S4D). In each instance, we identified that at least 50% of splicing events altered by RBM39 loss in MOLM-13 were also bound by RBM39, suggesting a direct role of RBM39 in splicing regulation of its mRNA targets (Figure S4E). Consistent with prior data suggesting an interaction of RBM39 with U2AF2 (Imai et al., 1993; Loerch et al., 2014; Stepanyuk et al., 2016), motif analysis showed highly pyrimidine-enriched sequences on mapped RBM39-binding sites (Figure 4F). In fact, ten of the top hexamer sequences enriched in RBM39-binding sites were pyrimidine rich, similar to results seen with U2AF2-RNA interactions previously reported (Shao et al., 2014). GO analysis of differentially spliced exons revealed that RBM39 loss was strongly enriched in processes related to RNA splicing, export, and metabolism as well as DNA replication and mitosis (Figure 4G). Moreover, gene set enrichment analysis (GSEA) of differential spliced mRNAs with RBM39 sgRNA treatment identified prominent downregulation of *HOXA9* targets, known to be required for leukemogenesis of these MLL-rearranged AML cell lines (Figure 4H) (Faber et al., 2009). These prominently include aberrant splicing due to intron retention across several introns of *BMI1*, a gene known to be required for leukemogenesis that was also identified as a direct RBM39 mRNA-binding target by our eCLIP sequencing (eCLIP-seq). Additionally, *GATA2*, an essential hematopoietic factor involved in maintaining AML transcriptional homeostasis, was aberrantly spliced upon RBM39 loss (Katsumura et al., 2016; Yang et al., 2017) (Figure 4I). The unannotated splicing changes observed in BMI1 and GATA2 upon RBM39 depletion are predicted to result in mRNAs that would be subjected to nonsense-mediated decay (NMD). These studies suggest that the dependency of RBM39 is, at least in part, due to defective splicing of genes involved in pathways required for AML cell survival.

Potent Anti-leukemic Effects of Pharmacologic RBM39 Degradation

An emerging field in drug discovery has been the characterization of small molecules that modulate E3 ligases to promote targeted protein degradation (Kronke et al., 2015; Lu et al., 2014). Interestingly, a class of compounds known as anti-cancer sulfonamides (including the drugs indisulam [also known as E7070], E7820, and chloroquinoxaline sulfonamide) was recently identified to selectively degrade RBM39 through an interaction with DCAF15, an adapter protein for the CUL4/Ddb1 E3 ubiquitin ligase (Han et al., 2017; Uehara et al., 2017). These compounds demonstrated an excellent safety profile in clinical trials and have previously been shown to exhibit some anti-tumor efficacy; however, overall response rates were low (Owa et al., 1999; Supuran, 2003), most likely because neither the mechanism of action nor potential biomarkers of response were known. To this end, we observed that *DCAF15* was more highly expressed in AML

Figure 4. RBM39 Loss Alters Splicing of mRNAs Essential for AML Cell Growth

(A) Number of differentially spliced events in MOLM-13 and THP-1 cells treated with RBM39 sgRNA versus control sgRNA.

(B) Scatterplot of cassette exons (CEs) promoted (red circles) and repressed (blue circles) in MOLM-13 treated with RBM39 sgRNA versus control sgRNA. Splicing is quantified using a "percent spliced in" value (PSI, or ψ value). Promoted and repressed CEs are defined as those whose inclusion levels are increased or decreased by $\psi \geq 10\%$, respectively. Gray circles represent exons for where $\Delta\psi$ is <10%.

(C) eCLIP analysis of RBM39-binding sites in MOLM-13 cells. Input-normalized peak signals are shown as $\log_2$ fold change. Purple points indicate eCLIP-enriched RBM39 peaks (FDR > 5%, log FC > 1) in biological replicates.

(D) Genomic distribution of RBM39 eCLIP-seq peaks.

(E) Metagene profile of RBM39-binding sites on exons throughout the transcriptome in MOLM-13 cells relative to 5′ and 3′ splice sites. Gray box represents exonic region.

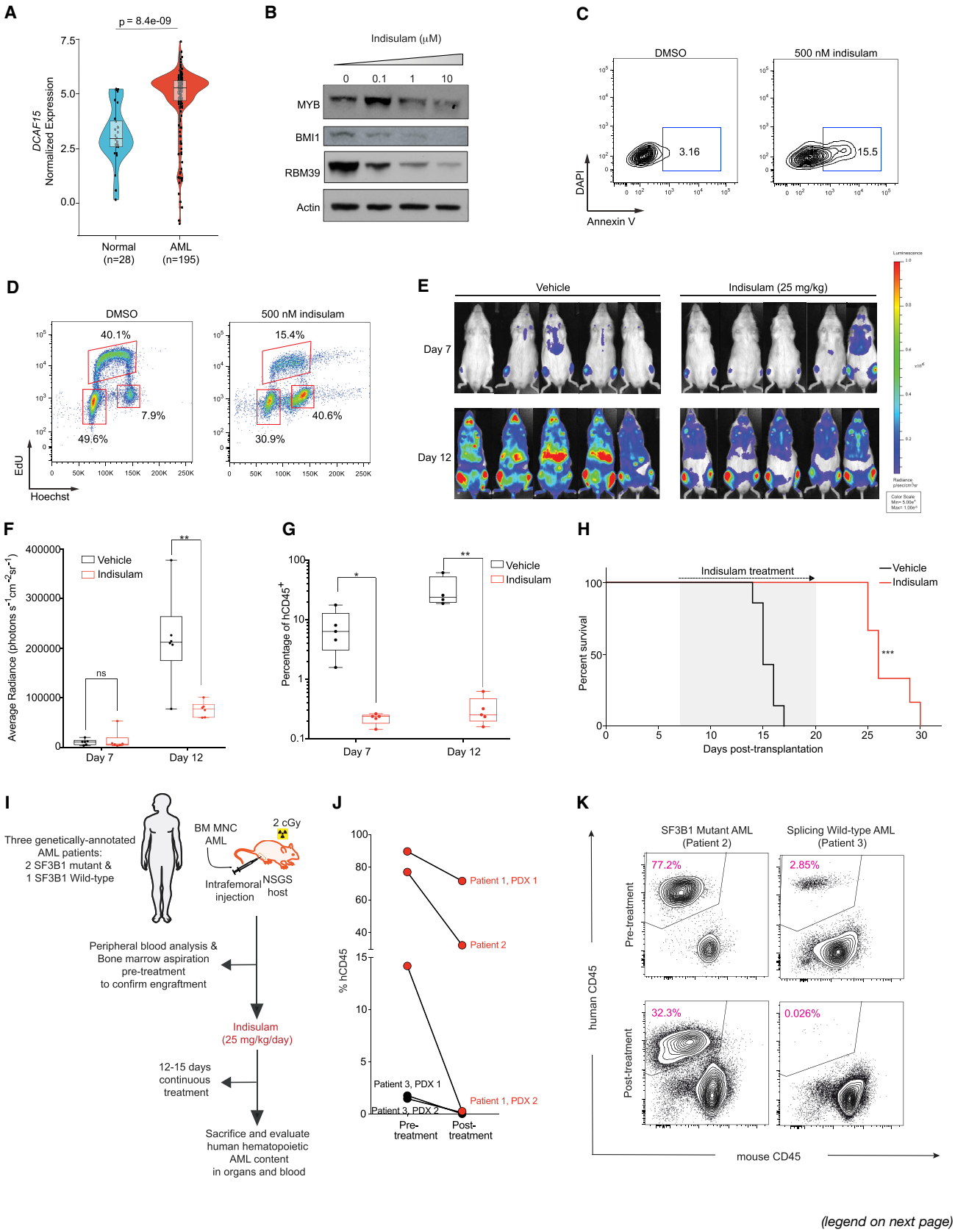
(F) Enriched motifs for RBM39 binding. The top ten motifs are shown. Inset shows consensus sequence, deduced from the top ten motifs.

(G) Gene ontology (GO) enrichment analysis of terms enriched in differentially spliced mRNAs in MOLM-13 cells with RBM39 sgRNA versus control. Terms in red font are related to RNA processing and/or splicing.

(H) Enrichment plots from gene set enrichment analysis (GSEA) of HOXA9 target genes that are differentially spliced upon RBM39 knockout versus control sgRNA (normalized enrichment score [NES] is inferred from permutations of the gene set and the false discovery rate [FDR]).

(I) RNA-seq coverage plot of *BMI1* and Sashimi plots of *GATA2* in MOLM13 cells treated with RBM39 sgRNA or control overlaid with anti-RBM39 eCLIP-seq tracks. Yellow highlighted reads in *GATA2* Sashimi plots highlight exon-skipping events with RBM39 sgRNA.

See also Figure S4; Tables S3, S4, and S5.



(legend on next page)

patient samples than normal hematopoietic progenitors (Figure 5A), and that increasing indisulam concentrations led to a dose-dependent reduction in RBM39 and HOXA9 target genes (BMI1 and MYB) protein levels in human leukemia cells (Figures 5B, S5A, and S5B). Indisulam exposure led to severely impaired G_2/M cell-cycle arrest and increased apoptosis after 48 h of treatment in AML cells *in vitro* (Figures 5C, 5D, and S5C). We additionally confirmed that the cellular effects of indisulam are dependent on *DCAF15* expression by designing sgRNAs to knock out *DCAF15* in AML-sensitive cells. Suppression of DCAF15 had little impact on AML cell viability at baseline. However, DCAF15 loss conferred robust resistance to treatment with indisulam, as demonstrated by the positive selection of AML cells expressing *DCAF15* sgRNAs (Figure S5D). Moreover, we verified RBM39 as a direct target of indisulam by overexpressing a previously described *RBM39* G268V mutation shown to confer resistance to indisulam (Figure S5E).

We next assessed the *in vivo* efficacy of indisulam in AML progression by transplanting AML cells expressing firefly luciferase into immune-deficient recipient mice. To simulate disease burden in patients we waited until onset of disease, and subsequently gave mice 13 consecutive daily treatments of either vehicle or 25 mg/kg indisulam, a previously described dose (Han et al., 2017). In concert with *in vitro* findings, administration of indisulam led to strong anti-leukemic effects *in vivo* using two AML cell line xenograft models, MOLM-13 and OCI-AML3. Bioluminescent quantification and imaging of animals treated with indisulam showed marked delay in AML burden (Figures 5E, 5F, and S5F). Compared with vehicle control, FACS analysis of peripheral blood and bone marrow revealed a substantial reduction in MOLM-13 cells and elevated apoptosis in animals treated with indisulam (Figures 5G and S5G). Immunohistochemistry analysis of RBM39 expression in bone marrow and spleen further confirmed *in vivo* degradation of RBM39 in indisulam-treated mice (Figure S5H). Overall, indisulam administration significantly extended survival in both AML xenograft models (Figures 5H and S5I), with a median survival time of 26 days in MOLM-13 indisulam-treated animals versus 15 days in vehicle control cohorts (Figure 5H). In addition, indisulam treatment to five patient-derived xenografts (PDX) generated from three distinct patients demonstrated reduced leukemia burden with indisulam treatment in each case (Figures 5I–5K and Figure S5J).

In contrast to the robust anti-leukemic effects of RBM39 degradation *in vitro* and *in vivo*, 4 weeks of indisulam treatment in normal C57/B6 mice resulted in no hematologic toxicities while causing on-target degradation of RBM39 (Figures S6A–S6H). Similarly, we noticed no significant effects of pharmacologic degradation of RBM39 on human normal hematopoietic cells *in vivo* using immunocompromised mice transplanted with cord blood $CD34^+$ cells (Figures S6I–S6Q). Overall, these findings demonstrate potent single-agent leukemic effects of indisulam *in vivo* with preferential effects of RBM39 loss on AML over normal human hematopoietic cells.

Preferential Sensitivity of Spliceosomal Mutant AML to RBM39 Degradation

Recent genomic studies of AML and related myeloid malignancies have identified recurrent change-of-function mutations in RNA splicing factors, highlighting a potential role for altered splicing in the pathogenesis of clonal hematopoietic malignancies. These mutations are conspicuously present in splicing factors that physically interact with RBM39 (including SF3B1 and U2AF1) and others (SRSF2) as a series of mutually exclusive, heterozygous point mutations at highly specific residues. Prior work has demonstrated that cancer cells bearing these mutations are genetically dependent on otherwise wild-type splicing catalysis for cell survival (Fei et al., 2016; Lee et al., 2016; Zhou et al., 2015). Similarly, cells bearing these mutations are preferentially sensitive to drugs that modulate core splicing catalysis by binding to SF3B1 (Lee and Abdel-Wahab, 2016; Obeng et al., 2016; Seiler et al., 2018; Shirai et al., 2017). Preferential

Figure 5. Pharmacologic Inhibition of RBM39 Displays Broad Sensitivity across Diverse AMLs

(A) Violin plot of *DCAF15* normalized counts from RNA-seq in AML patients (red) and normal human $CD34^+$ hematopoietic stem and progenitor cells (blue). Horizontal line inside the box represents the mean and 25^{th}–75^{th} percentiles, showing all data points. Statistical analysis was performed using a Wilcoxon rank-sum test.

(B) Western blot of K562 whole-cell lysates treated with escalated doses of indisulam after 48 h.

(C) Representative flow-cytometry plots of Annexin V staining performed on MOLM-13 AML cells treated with 500 nM indisulam or DMSO for 48 h.

(D) Representative flow-cytometry plots showing 5-ethynyl-2′-deoxyuridine cell-cycle analysis of MOLM-13 AML cells exposed to 500 nM indisulam or DMSO for 48 h.

(E) Bioluminescent imaging of mice transplanted with MOLM-13 luciferase cells and treated with vehicle or 25 mg/kg indisulam. Representative images of five mice per group are shown. The same mice are depicted at each time point.

(F) Quantification of bioluminescent signals in vehicle versus indisulam-treated leukemic mice at indicated time points (n = 6/group). Unpaired Student's t test using Prism 7 (GraphPad). Box-and-whisker plot, minimum to maximum showing all points, 25^{th}–75^{th} percentiles, and median (horizontal line). Statistical analysis was performed using unpaired Student's t test by Prism GraphPad (**$p < 0.01$).

(G) Flow-cytometry analysis of human MOLM-13 cells in peripheral blood in DMSO versus indisulam groups. Gating was performed on human $CD45^+$ in DMSO-treated (n = 4–5) and indisulam-treated (n = 5) mice. Box-and-whisker plot, 25^{th}–75^{th} percentiles, minimum to maximum showing all points, and median (horizontal line). Statistical analysis was performed using unpaired Student's t test by Prism GraphPad (*$p < 0.05$, **$p < 0.01$).

(H) Kaplan-Meier overall survival of indisulam-treated AML mice. Gray represents the exposure time (days) to indisulam. The p values were calculated using a log-rank Mantel-Cox test (***$p < 0.001$).

(I) Schematic of patient-derived xenograft (PDX) generation and treatment with indisulam (25 mg/kg/day). Patient 1 (PDX 1) received 3 days of drug treatment.

(J) Percentage of human $CD45^+$ (hCD45) cells among total live bone marrow mononuclear cells by flow-cytometry analysis of bone marrow aspirate pre- and post-treatment with indisulam in five primary AML PDX models from three patients. Red indicates *SF3B1* mutant patients.

(K) Representative flow-cytometry plots of CD45 in patient 2 and patient 3 bone marrow aspirates before and after indisulam treatment. Percentage of $hCD45^+$ cells is indicated.

See also Figures S5 and S6.

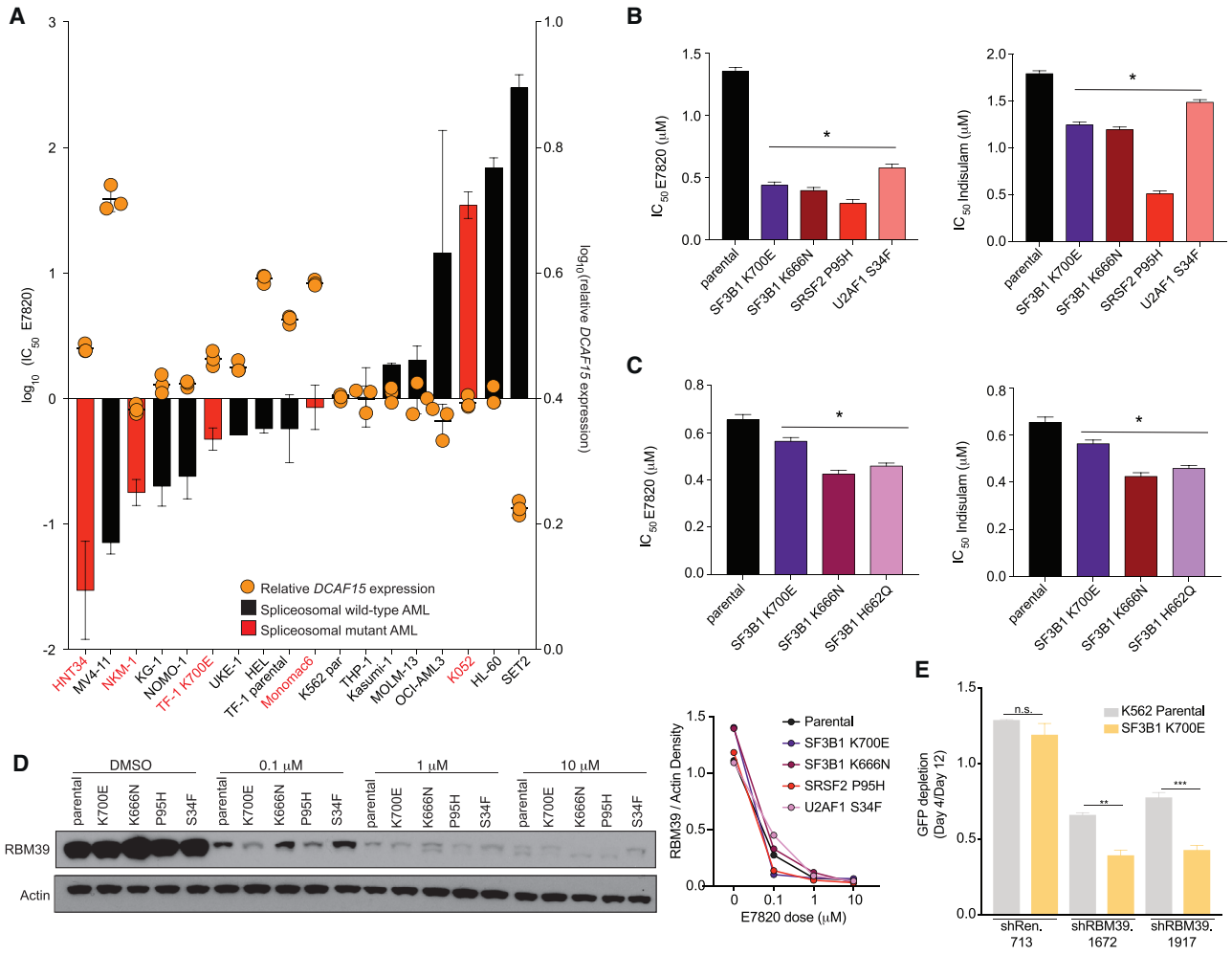


Figure 6. Preferential Effects of Sulfonamides on Spliceosomal Mutant AML

(A) Log_{10} of IC_{50} values of a panel of human AML cell lines to E7820 annotated for spliceosomal gene mutation status (red versus black bars indicate spliceosomal mutant or wild-type cells, respectively) as determined by the CellTiterGlo assay. IC_{50} values are calculated from at least four technical replicates per experiment, each experiment was performed at least twice, and error bars represent SD. Also shown are log_{10} relative *DCAF*15 mRNA expression levels for each cell line.

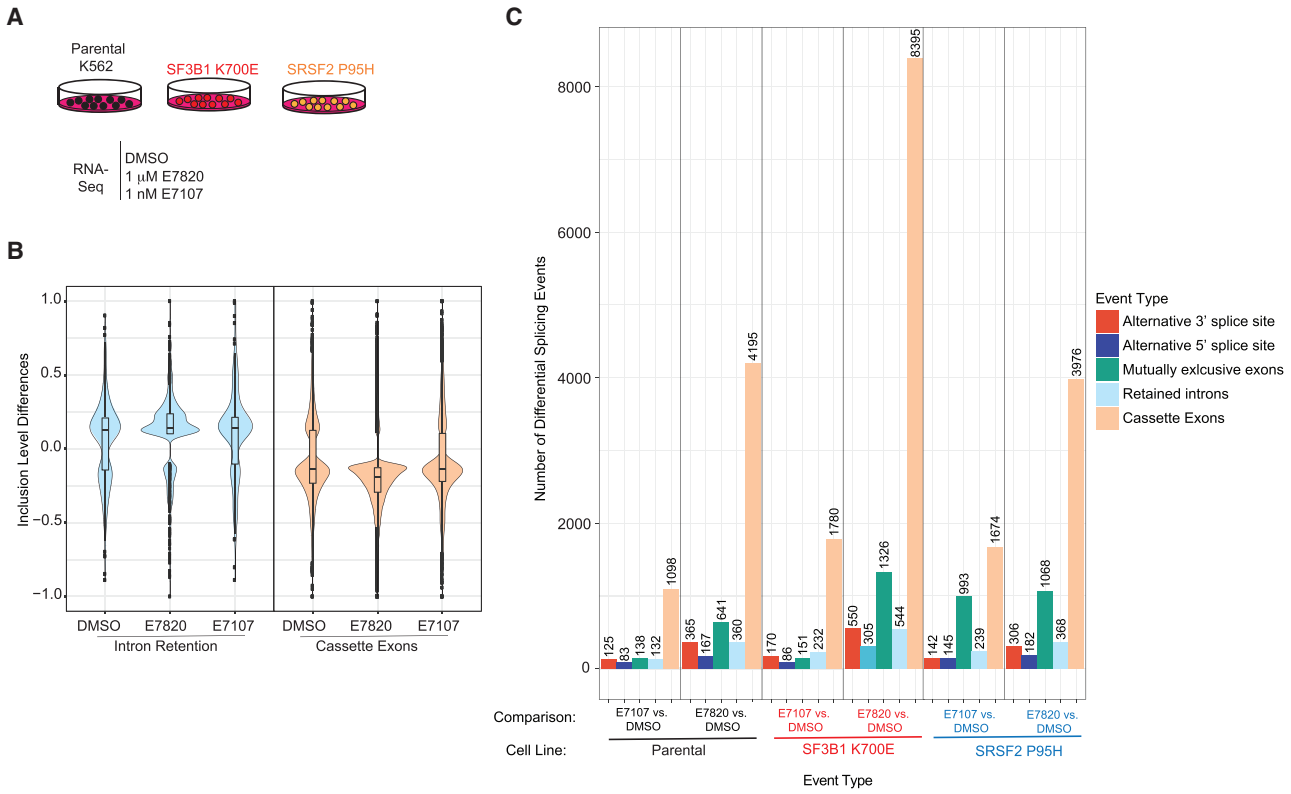
(B) Bar plots of IC_{50} values of isogenic K562 cells with or without knockin of spliceosomal gene mutations to E7820 (left) or indisulam (right). IC_{50} values calculated from four technical replicates per experiment; one representative experiment of three biological replicates is shown; error bars represent SD. Statistical analysis was performed using unpaired Student's t test by Prism GraphPad (*$p < 0.05$).

(C) Bar plots of IC_{50} values of isogenic NALM-6 cells with or without knockin of spliceosomal gene mutations to E7820 (left) or indisulam (right). IC_{50} values calculated from one representative experiment of three biological replicates are shown; error bars represent SD. Statistical analysis was performed using unpaired Student's t test by Prism GraphPad (*$p < 0.05$).

(D) Anti-RBM39 western blot of isogenic K562 cells treated with DMSO or increasing doses of E7820 for 24 h. Densitometry quantification of western blot is shown on the right.

(E) Competition assay of *RBM39* shRNA knockdown in K562 parental versus K562 *SF3B1*^{K700E}. Analysis was determined by unpaired Student's t test. Data with statistical significance of mutant cells relative to control are as indicated (**$p < 0.01$, ***$p < 0.001$; n.s., not significant).

sensitivity of splicing mutant hematopoietic cells over splicing wild-type counterparts has previously been demonstrated for the drugs E7107, sudemycin, and H3B-8800. As mentioned earlier, this has led to an ongoing phase I trial to identify the safety and maximum tolerated dose of H3B-8800 in patients with refractory myeloid leukemia (Lee et al., 2016; Seiler et al., 2018; Shirai et al., 2017). Given our data that RBM39 loss resulted in greatly altered splicing of factors required for RNA splicing, export, and catabolism, we hypothesized that spliceosomal mutant AML might be preferentially sensitive to sulfonamide-induced RBM39 degradation. To this end, we further investigated the *in vitro* selectivity of sulfonamides across a panel of genetically diverse AML cell lines. Measurement of cell viability upon sulfonamide exposure revealed broad anti-leukemic effects with potent inhibitory activity across many AML subtypes, with most cell lines exhibiting submicromolar sensitivity (Figure 6A). We found that leukemia cells bearing mutations in leukemia-associated mutations in splicing factors were among the most sensitive cells to sulfonamides. In addition, a number of AML cell lines without spliceosomal gene mutations



(legend on next page)

also exhibit sensitivity to the sulfonamides. Evaluation of relative *DCAF15* mRNA expression revealed that the highest and lowest relative levels of *DCAF15* mRNA correlated with the greatest and worst response to E7820, respectively (Figure 6A). The preferential effects of sulfonamides on leukemia cells bearing spliceosomal gene mutations was further confirmed in a series of isogenic AML lines (K562 and TF-1) engineered to express hotspot mutations in *SF3B1*, *SRSF2*, and *U2AF1* from their endogenous loci (Figures 6A and 6B) and B cell acute lymphoblastic leukemia (NALM-6 cells; Figure 6C) with exposure to E7820 and indisulam. In each instance, spliceosomal mutant cells were preferentially sensitive to growth inhibition to sulfonamides over spliceosomal wild-type cells. In isogenic cells, E7820 exposure led to similar dose-dependent degradation of RBM39 in leukemia cell lines (Figure 6D). Furthermore, short hairpin RNA (shRNA) knockdown of *RBM39* in K562 cells knockin of SF3B1^{K700E} resulted in greater competitive disadvantage when compared with K562 parental cells, supporting the concept that spliceosomal mutant cells are preferentially sensitive to RBM39 loss (Figure 6E).

RBM39 Degradation Targets an RBP Network Required for AML Survival

To understand the basis for the preferential effects of sulfonamides on spliceosomal mutant cells, we next evaluated RBM39 protein levels across a panel of isogenic AML cells with or without knockin of spliceosomal gene mutations. Degradation of RBM39 occurred in a comparable dose-dependent fashion across cell lines regardless of gene mutation status (Figure 6D), suggesting that a greater dependency on residual wild-type splicing function may be the basis for increased sensitivity of spliceosomal mutant cells to sulfonamides. Given that spliceosomal mutant cells are preferentially sensitive to alterations in splicing over their wild-type counterparts, we next performed RNA-seq of parental K562 and isogenic lines expressing SF3B1^{K700E} and SRSF2^{P95H} mutations treated with 1 mM E7820 (which represents the IC_{50} of parental K562 cells to E7820; Figure 7A). In parallel, we also carried out RNA-seq of the same cell lines treated with E7107, a small molecule that inhibits splicing through impeding binding of SF3B1 to the branchpoint (Cretu et al., 2018; Finci et al., 2018). Treatment with either E7820 or E7107 at the IC_{50} of each drug in parental K562 cells resulted in increased CE skipping and intron retention relative to DMSO treatment regardless of spliceosomal gene mutation status (Figures 7B and 7C). Interestingly, however, at equipotent non-toxic doses, E7820 resulted in greater changes in splicing within each cell type and across each category of splicing versus E7107 at this dose (Figures 7B, 7C, and S7A–S7D). Moreover, a greater number of differential splicing events were identified within each type of splicing in SF3B1^{K700E} cells treated with E7820 versus SF3B1 wild-type counterparts. These data suggest that at least one reason for the preferential effects of sulfonamides on SF3B1-mutant over wild-type cells is a heightened splicing response of SF3B1-mutant cells to RBM39 degradation.

We also noticed that a number of differentially spliced events upon sulfonamide exposure involved mRNAs encoding RBPs identified in the CRISPR screen as upregulated and required for AML cell survival (Figure 1E). These included *SUPT6H*, *HNRNPH1*, and *SRSF10* (Figure 7D), whereby E7820 exposure resulted in intron retention that was most pronounced in spliceosomal mutant cells. In addition and in agreement with our previous findings, RBM39 degradation also resulted in enhanced aberrant splicing of the HOXA9 target genes *BMI1* and *MYB* and a number of RBPs in spliceosomal mutant AML over wild-type counterparts, including aberrant splicing events in *U2AF2* and *RBM3* (Figures 7D and S7A–S7E). Although there was no mis-splicing of *HOXA9* upon indisulam treatment (Figures S7F–S7H), we observed aberrant splicing and reduced expression of *MYB*, *GATA2*, and *BMI1* (Figure S7I). These data are consistent with the prediction that mis-splicing of *MYB*, *GATA2*, and *BMI1* upon RBM39 loss would result in NMD and indicate that RBM39 loss resulted in anti-leukemia effects by altering splicing and expression of HOXA9 target genes without affecting HOXA9 splicing or expression itself. Moreover, GSEA of differentially spliced events in response to E7820 also revealed downregulation of targets of MYC and PI3K-AKT-mTOR signaling as well as mRNAs involved in response to inflammation (Figures S7J–S7K), all known to be important in AML pathogenesis or progression. Overall, these data suggest that the presence of spliceosomal gene mutations as well as *DCAF15* expression may serve as important predictors of response to RBM39 degradation in AML.

DISCUSSION

AML continues to have a dismal survival rate. This can largely be attributed to limited advances in treatment regimens that, for the last decades, have relied on the use of non-targeted cytotoxic drugs. However, recent Food and Drug Administration approvals of the first small-molecule inhibitors targeting recurrent genetic lesions in AML (midostaurin and enasidenib) provides hope for molecularly targeted therapies for this challenging illness. Despite the success in the development of these small-molecule inhibitors, the genetic heterogeneity of AML presents a current therapeutic barrier, whereby available selective inhibitors target only a genetically defined subset of patients. However, it appears that AML cells rely on specific pathways that have not been subjected to genomic alterations for their survival. This

Figure 7. RBM39 Degradation Targets an RBP Network Required for AML Survival

(A) Schema of drug testing (6 h) across isogenic K562 cells with or without heterozygous knockin of the *SF3B1*^{K700E} or *SRSF2*^{P95H} mutation at the endogenous locus for RNA-seq. Each experiment was perform in duplicates.

(B) Violin plot of inclusion level differences of introns or CEs in *SF3B1*$^{K700E/WT}$ cells treated with DMSO versus *SF3B1*^{WT} cells treated with DMSO ("DMSO" column) or *SF3B1*$^{K700E/WT}$ cells treated with E7820 or E7107 versus the same cells treated with DMSO. Horizontal line inside the box represents the mean and 25th–75th percentiles; points are "outliers" (outside the 99th percentile line). Statistical analysis was performed using a Wilcoxon rank-sum test.

(C) Bar plot of number of differential splicing events across parental, *SF3B1*$^{K700E/WT}$, and *SRSF2*$^{P95H/WT}$ K562 cells. The numbers above each bar indicate the number of differentially spliced events.

(D) RNA-seq coverage plots of *SUPT6H*, *HNRNPH1*, *SRSF10*, and *U2AF2* in the isogenic K562 cells from (A).

See also Figures S7.

has led to the discovery of broader targeted therapies that are currently in clinical trials for hematologic malignancies (Kotschy et al., 2016; Zuber et al., 2011b).

From our focused CRISPR/Cas9 screen, we identified several dysregulated RBPs that are essential for AML maintenance. Many of these proteins have not been described previously as leukemia dependencies and warrant future research to elucidate their molecular mechanism in leukemia. Also, the direct comparison to identical CRISPR/Cas9 screens in T-ALL, melanoma, and LUAD gives us the ability to identify RBPs that play specific functions in AML cells. Here, we have focused our efforts on understanding the dependency of RBM39 in AML. Mechanistically, we showed that CRISPR-mediated deletion or pharmacologic degradation of RBM39 causes altered splicing of *HOXA9* target genes, an essential transcriptional network required in AML. This in turn consequently leads to a preferential targeting of AML cells compared with other cancers. Given the mechanistic role of RBM39 in splicing shown here and the requirement of the DCAF15 adapter protein for anti-cancer sulfonamide activity, we identified that both the presence of spliceosomal gene mutations and levels of DCAF15 expression are important predictors of response to sulfonamides. Currently, a regimen that includes indisulam in combination with standard chemotherapy is being investigated in phase II trials in patients with refractory or relapsed myeloid malignancies (Assi et al., 2018). While indisulam is given to patients by intravenous infusion, E7820 is an orally bioavailable drug that has been tested as an improved, highly on-target, second-generation anti-cancer sulfonamide in phase II clinical trials for solid tumors such as colorectal cancer (Milojkovic Kerklaan et al., 2016). The present study supports further clinical investigation of E7820 in patients with myeloid malignancies. These data also provide mechanistic support for expanded use of sulfonamides in clinical trials, as it identifies RBM39 as a key non-oncogenic addiction in AML, describes its mechanism of action, and offers valuable potential biomarkers and genetic predictors of response. The known safety of anti-cancer sulfonamides established through multiple phase I and phase II clinical trials to date provides an additional advantage over drugs targeting splicing by modulating SF3b function whose safety is not yet known. Indeed, we are currently planning a multi-center clinical trial using E7820, targeting AML patients that either carry spliceosomal mutations and/or express high levels of DCAF15, as biomarkers of response.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- CONTACT FOR REAGENT AND RESOURCE SHARING
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Lines and Cell Culture
  - Primary Human Samples
  - Animals
- METHOD DETAILS
  - sgRNA Library Design
  - CRISPR Screen
  - CRISPR/Cas9 Indel Analysis
  - Apoptosis and Cell Cycle
  - RNA-Sequencing Library Preparation and Sequencing
  - eCLIP Library Preparation
  - Western Blotting
  - Sulfonamide Drug Treatment IC_{50} Measurements
  - qRT-PCR Measurement of Gene Expression
  - RBM39 cDNA Overexpression
  - Immunoprecipitation
  - IP-MS
  - Animal Experiments
  - Immunohistochemistry
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Splicing Analysis
  - eCLIP Data Analysis
  - Gene Ontology Analysis
  - AML TCGA Analysis
  - Statistical Analysis
- DATA AND SOFTWARE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information includes seven figures and six tables and can be found with this article online at https://doi.org/10.1016/j.ccell.2019.01.010.

ACKNOWLEDGMENTS

We would like to thank all members of the O.A.-W. and I.A. labs for useful discussions and comments on the manuscript; A. Heguy and the NYU Genome Technology Center (supported in part by National Institutes of Health/National Cancer Institute grant P30CA016087-30) for expertise with sequencing experiments; the NYU Histology Core (5P30CA16087-31) for assistance; C. Loomis and L. Chiriboga for immunohistochemistry experiments. We also thank Dr. Chris Vakoc for RN2-Cas9 cells and MOLM-13 Cas9 cells; We are grateful to Eclipse Bio for performing eCLIP experiments. E.W. is supported by the NIH T32 Genome Integrity grant. S.X.L. is supported by a Conquer Cancer Foundation and ASCO Young Investigator Award, an Aplastic Anemia & Myelodysplastic Syndrome International Foundation research award, as well as an AACR Lymphoma Research Fellowship. O.A.-W. was supported by the US National Institutes of Health (R01 HL128239), the Leukemia & Lymphoma Society, the Pershing Square Sohn Cancer Research Alliance, The Evans MDS Foundation, The Druckenmiller Center for Lung Cancer Research at MSK, and the Department of Defense Bone Marrow Failure Research Program (BM150092 and W81XWH-12-1-0041). O.A.-W. and A.R.K. are supported by grants from the Starr Cancer Consortium (I8-A8-075) and the Taub Foundation. I.A. was supported by the US National Institutes of Health (1R01 CA228135, 1R01 CA216421, 1R01CA194923, 1R01CA169784, 5RO1CA173636), the Leukemia & Lymphoma Society (TRP#6499-17), the Taub Foundation, the Alex's Lemonade Stand Foundation for Childhood Cancer, and the St. Baldrick's Cancer Research Foundation. The work was also supported by the New York State Department of Health (#CO030132, C32587GG, C32563GG).

AUTHOR CONTRIBUTIONS

Conceptualization, E.W., O.A.-W., and I.A.; Methodology, E.W. and I.A.; Formal Analysis, A.P., Y.E.G., and K.-T.L.; Investigation, E.W., A.P., S.X.L., S.C.-W.L., X.C., J.I., Y.E.G., A.Y., L.B., and R.T.; Resources, T.U., T.O., O.A.-W., and A.R.K.; Writing – Original Draft, E.W., O.A.-W., and I.A.; Writing – Review & Editing, E.W., S.X.L., A.Y., O.A.-W., and I.A.; Visualization, E.W.

DECLARATION OF INTEREST

T.H. and T.O. are employees of Eisai Co. Ltd. O.A.-W. has served as a consultant for H3B Biomedicine, Foundation Medicine Inc., Merck, and Janssen. O.A.-W. has received prior research funding from H3B Biomedicine unrelated to the current manuscript. The other authors declare no competing interests.

Received: July 2, 2018
Revised: November 26, 2018
Accepted: January 18, 2019
Published: February 21, 2019

REFERENCES

Anczukow, O., Rosenberg, A.Z., Akerman, M., Das, S., Zhan, L., Karni, R., Muthuswamy, S.K., and Krainer, A.R. (2012). The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. Nat. Struct. Mol. Biol. *19*, 220–228.

Assi, R., Kantarjian, H.M., Kadia, T.M., Pemmaraju, N., Jabbour, E., Jain, N., Daver, N., Estrov, Z., Uehara, T., Owa, T., et al. (2018). Final results of a phase 2, open-label study of indisulam, idarubicin, and cytarabine in patients with relapsed or refractory acute myeloid leukemia and high-risk myelodysplastic syndrome. Cancer *124*, 2758–2765.

Cancer Genome Atlas Research Network, Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Cheng, D.T., Mitchell, T.N., Zehir, A., Shah, R.H., Benayed, R., Syed, A., Chandramohan, R., Liu, Z.Y., Won, H.H., Scott, S.N., et al. (2015). Memorial Sloan Kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J. Mol. Diagn. *17*, 251–264.

Cook, K.B., Kazan, H., Zuberi, K., Morris, Q., and Hughes, T.R. (2011). RBPDB: a database of RNA-binding specificities. Nucleic Acids Res. *39*, D301–D308.

Cooper, T.A., Wan, L., and Dreyfuss, G. (2009). RNA and disease. Cell *136*, 777–793.

Cretu, C., Agrawal, A.A., Cook, A., Will, C.L., Fekkes, P., Smith, P.G., Luhrmann, R., Larsen, N., Buonamici, S., and Pena, V. (2018). Structural basis of splicing modulation by antitumor macrolide compounds. Mol. Cell *70*, 265–273.e8.

Dohner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F.R., Buchner, T., Dombret, H., Ebert, B.L., Fenaux, P., Larson, R.A., et al. (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood *129*, 424–447.

Dohner, H., Weisdorf, D.J., and Bloomfield, C.D. (2015). Acute myeloid leukemia. N. Engl. J. Med. *373*, 1136–1152.

Dvinge, H., Kim, E., Abdel-Wahab, O., and Bradley, R.K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. Nat. Rev. Cancer *16*, 413–430.

Faber, J., Krivtsov, A.V., Stubbs, M.C., Wright, R., Davis, T.N., van den Heuvel-Eibrink, M., Zwaan, C.M., Kung, A.L., and Armstrong, S.A. (2009). HOXA9 is required for survival in human MLL-rearranged acute leukemias. Blood *113*, 2375–2385.

Fei, D.L., Motowski, H., Chatrikhi, R., Prasad, S., Yu, J., Gao, S., Kielkopf, C.L., Bradley, R.K., and Varmus, H. (2016). Wild-type U2AF1 antagonizes the splicing program characteristic of U2AF1-mutant tumors and is required for cell survival. PLoS Genet. *12*, e1006384.

Fiedler, W., Henke, R.P., Ergun, S., Schumacher, U., Gehling, U.M., Vohwinkel, G., Kilic, N., and Hossfeld, D.K. (2008). Derivation of a new hematopoietic cell line with endothelial features from a patient with transformed myeloproliferative syndrome: a case report. Cancer *88*, 344–351.

Finci, L.I., Zhang, X., Huang, X., Zhou, Q., Tsai, J., Teng, T., Agrawal, A., Chan, B., Irwin, S., Karr, C., et al. (2018). The cryo-EM structure of the SF3b spliceosome complex bound to a splicing modulator reveals a pre-mRNA substrate competitive mechanism of action. Genes Dev. *32*, 309–320.

Fox, R.G., Lytle, N.K., Jaquish, D.V., Park, F.D., Ito, T., Bajaj, J., Koechlein, C.S., Zimdahl, B., Yano, M., Kopp, J., et al. (2016). Image-based detection and targeting of therapy resistance in pancreatic adenocarcinoma. Nature *534*, 407–411.

Gerstberger, S., Hafner, M., and Tuschl, T. (2014). A census of human RNA-binding proteins. Nat. Rev. Genet. *15*, 829–845.

Glisovic, T., Bachorik, J.L., Yong, J., and Dreyfuss, G. (2008). RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett. *582*, 1977–1986.

Han, T., Goralski, M., Gaskill, N., Capota, E., Kim, J., Ting, T.C., Xie, Y., Williams, N.S., and Nijhawan, D. (2017). Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. Science *356*, https://doi.org/10.1126/science.aal3755.

Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K.R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., et al. (2015). High-resolution CRISPR screens reveal fitness genes and genotype-specific cancer liabilities. Cell *163*, 1515–1526.

Imai, H., Chan, E.K., Kiyosawa, K., Fu, X.D., and Tan, E.M. (1993). Novel nuclear autoantigen with splicing factor motifs identified with antibody from hepatocellular carcinoma. J. Clin. Invest. *92*, 2419–2426.

Kapeli, K., Martinez, F.J., and Yeo, G.W. (2017). Genetic mutations in RNA-binding proteins and their roles in ALS. Hum. Genet. *136*, 1193–1214.

Katsumura, K.R., Ong, I.M., DeVilbiss, A.W., Sanalkumar, R., and Bresnick, E.H. (2016). GATA factor-dependent positive-feedback circuit in acute myeloid leukemia cells. Cell Rep. *16*, 2428–2441.

Kharas, M.G., Lengner, C.J., Al-Shahrour, F., Bullinger, L., Ball, B., Zaidi, S., Morgan, K., Tam, W., Paktinat, M., Okabe, R., et al. (2010). Musashi-2 regulates normal hematopoiesis and promotes aggressive myeloid leukemia. Nat. Med. *16*, 903–908.

Kotschy, A., Szlavik, Z., Murray, J., Davidson, J., Maragno, A.L., Le Toumelin-Braizat, G., Chanrion, M., Kelly, G.L., Gong, J.N., Moujalled, D.M., et al. (2016). The MCL1 inhibitor S63845 is tolerable and effective in diverse cancer models. Nature *538*, 477–482.

Kronke, J., Fink, E.C., Hollenbach, P.W., MacBeth, K.J., Hurst, S.N., Udeshi, N.D., Chamberlain, P.P., Mani, D.R., Man, H.W., Gandhi, A.K., et al. (2015). Lenalidomide induces ubiquitination and degradation of CK1alpha in del(5q) MDS. Nature *523*, 183–188.

Lavallee, V.P., Baccelli, I., Krosl, J., Wilhelm, B., Barabe, F., Gendron, P., Boucher, G., Lemieux, S., Marinier, A., Meloche, S., et al. (2015a). The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. Nat. Genet. *47*, 1030–1037.

Lavallee, V.P., Gendron, P., Lemieux, S., D'Angelo, G., Hebert, J., and Sauvageau, G. (2015b). EVI1-rearranged acute myeloid leukemias are characterized by distinct molecular alterations. Blood *125*, 140–143.

Lavallee, V.P., Lemieux, S., Boucher, G., Gendron, P., Boivin, I., Armstrong, R.N., Sauvageau, G., and Hebert, J. (2016). RNA-sequencing analysis of core binding factor AML identifies recurrent ZBTB7A mutations and defines RUNX1-CBFA2T3 fusion signature. Blood *127*, 2498–2501.

Lee, S.C., and Abdel-Wahab, O. (2016). Therapeutic targeting of splicing in cancer. Nat. Med. *22*, 976–986.

Lee, S.C., Dvinge, H., Kim, E., Cho, H., Micol, J.B., Chung, Y.R., Durham, B.H., Yoshimi, A., Kim, Y.J., Thomas, M., et al. (2016). Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. Nat. Med. *22*, 672–678.

Loerch, S., Maucuer, A., Manceau, V., Green, M.R., and Kielkopf, C.L. (2014). Cancer-relevant splicing factor CAPERalpha engages the essential splicing factor SF3b155 in a specific ternary complex. J. Biol. Chem. *289*, 17325–17337.

Lu, G., Middleton, R.E., Sun, H., Naniong, M., Ott, C.J., Mitsiades, C.S., Wong, K.K., Bradner, J.E., and Kaelin, W.G., Jr. (2014). The myeloma drug lenalidomide promotes the cereblon-dependent destruction of Ikaros proteins. Science *343*, 305–309.

Lukong, K.E., Chang, K.W., Khandjian, E.W., and Richard, S. (2008). RNA-binding proteins in human genetic disease. Trends Genet. *24*, 416–425.

Lunde, B.M., Moore, C., and Varani, G. (2007). RNA-binding proteins: modular design for efficient function. Nat. Rev. Mol. Cell Biol. *8*, 479–490.

Milojkovic Kerklaan, B., Slater, S., Flynn, M., Greystoke, A., Witteveen, P.O., Megui-Roelvink, M., de Vos, F., Dean, E., Reyderman, L., Ottesen, L., et al. (2016). A phase I, dose escalation, pharmacodynamic, pharmacokinetic, and food-effect study of alpha2 integrin inhibitor E7820 in patients with advanced solid tumors. Invest. New Drugs *34*, 329–337.

Obeng, E.A., Chappell, R.J., Seiler, M., Chen, M.C., Campagna, D.R., Schmidt, P.J., Schneider, R.K., Lord, A.M., Wang, L., Gambe, R.G., et al. (2016). Physiologic expression of Sf3b1(K700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell *30*, 404–417.

Owa, T., Yoshino, H., Okauchi, T., Yoshimatsu, K., Ozawa, Y., Sugi, N.H., Nagasu, T., Koyanagi, N., and Kitoh, K. (1999). Discovery of novel antitumor sulfonamides targeting G1 phase of the cell cycle. J. Med. Chem. *42*, 3789–3799.

Pabst, C., Bergeron, A., Lavallee, V.P., Yeh, J., Gendron, P., Norddahl, G.L., Krosl, J., Boivin, I., Deneault, E., Simard, J., et al. (2016). GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. Blood *127*, 2018–2027.

Palanichamy, J.K., Tran, T.M., Howard, J.M., Contreras, J.R., Fernando, T.R., Sterne-Weiler, T., Katzman, S., Toloue, M., Yan, W., Basso, G., et al. (2016). RNA-binding protein IGF2BP3 targeting of oncogenic transcripts promotes hematopoietic progenitor proliferation. J. Clin. Invest. *126*, 1495–1511.

Papaemmanuil, E., Dohner, H., and Campbell, P.J. (2016). Genomic classification in acute myeloid leukemia. N. Engl. J. Med. *375*, 900–901.

Prigge, J.R., Iverson, S.V., Siders, A.M., and Schmidt, E.E. (2009). Interactome for auxiliary splicing factor U2AF(65) suggests diverse roles. Biochim. Biophys. Acta *1789*, 487–492.

Sebestyen, E., Singh, B., Minana, B., Pages, A., Mateo, F., Pujana, M.A., Valcarcel, J., and Eyras, E. (2016). Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. Genome Res. *26*, 732–744.

Seiler, M., Yoshimi, A., Darman, R., Chan, B., Keaney, G., Thomas, M., Agrawal, A.A., Caleb, B., Csibi, A., Sean, E., et al. (2018). H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nat. Med. *24*, 497–504.

Shalem, O., Sanjana, N.E., Hartenian, E., Shi, X., Scott, D.A., Mikkelson, T., Heckl, D., Ebert, B.L., Root, D.E., Doench, J.G., and Zhang, F. (2014). Genome-scale CRISPR-Cas9 knockout screening in human cells. Science *343*, 84–87.

Shao, C., Yang, B., Wu, T., Huang, J., Tang, P., Zhou, Y., Zhou, J., Qiu, J., Jiang, L., Li, H., et al. (2014). Mechanisms for U2AF to define 3' splice sites and regulate alternative splicing in the human genome. Nat. Struct. Mol. Biol. *21*, 997–1005.

Shi, J., Wang, E., Milazzo, J.P., Wang, Z., Kinney, J.B., and Vakoc, C.R. (2015). Discovery of cancer drug targets by CRISPR-Cas9 screening of protein domains. Nat. Biotechnol. *33*, 661–667.

Shirai, C.L., White, B.S., Tripathi, M., Tapia, R., Ley, J.N., Ndonwi, M., Kim, S., Shao, J., Carver, A., Saez, B., et al. (2017). Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. Nat. Commun. *8*, 14060.

Stepanyuk, G.A., Serrano, P., Peralta, E., Farr, C.L., Axelrod, H.L., Geralt, M., Das, D., Chiu, H.J., Jaroszewski, L., Deacon, A.M., et al. (2016). UHM-ULM interactions in the RBM39-U2AF65 splicing-factor complex. Acta Crystallogr. D Struct. Biol. *72*, 497–511.

Supuran, C.T. (2003). Indisulam: an anticancer sulfonamide in clinical development. Expert Opin. Investig. Drugs *12*, 283–287.

Tsherniak, A., Vazquez, F., Montgomery, P.G., Weir, B.A., Kryukov, G., Cowley, G.S., Gill, S., Harrington, W.F., Pantel, S., Krill-Burger, J.M., et al. (2017). Defining a cancer dependency map. Cell *170*, 564–576.e16.

Uehara, T., Minoshima, Y., Sagane, K., Sugi, N.H., Mitsuhashi, K.O., Yamamoto, N., Kamiyama, H., Takahashi, K., Kotake, Y., Uesugi, M., et al. (2017). Selective degradation of splicing factor CAPERalpha by anticancer sulfonamides. Nat. Chem. Biol. *13*, 675–680.

Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y., Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016). Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). Nat. Methods *13*, 508–514.

Wang, T., Yu, H., Hughes, N.W., Liu, B., Kendirli, A., Klein, K., Chen, W.W., Lander, E.S., and Sabatini, D.M. (2017). Gene essentiality profiling reveals gene networks and synthetic lethal interactions with oncogenic Ras. Cell *168*, 890–903.e15.

Wurth, L., Papasaikas, P., Olmeda, D., Bley, N., Calvo, G.T., Guerrero, S., Cerezo-Wallis, D., Martinez-Useros, J., Garcia-Fernandez, M., Huttelmaier, S., et al. (2016). UNR/CSDE1 drives a post-transcriptional program to promote melanoma invasion and metastasis. Cancer Cell *30*, 694–707.

Yang, L., Sun, H., Cao, Y., Xuan, B., Fan, Y., Sheng, H., and Zhuang, W. (2017). GATA2 inhibition sensitizes acute myeloid leukemia cells to chemotherapy. PLoS One *12*, e0170630.

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.

Zehir, A., Benayed, R., Shah, R.H., Syed, A., Middha, S., Kim, H.R., Srinivasan, P., Gao, J., Chakravarty, D., Devlin, S.M., et al. (2017). Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat. Med. *23*, 703–713.

Zhou, Q., Derti, A., Ruddy, D., Rakiec, D., Kao, I., Lira, M., Gibaja, V., Chan, H., Yang, Y., Min, J., et al. (2015). A chemical genetics approach for the functional assessment of novel cancer genes. Cancer Res. *75*, 1949–1958.

Zuber, J., McJunkin, K., Fellmann, C., Dow, L.E., Taylor, M.J., Hannon, G.J., and Lowe, S.W. (2011a). Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. Nat. Biotechnol. *29*, 79–83.

Zuber, J., Shi, J., Wang, E., Rappaport, A.R., Herrmann, H., Sison, E.A., Magoon, D., Qi, J., Blatt, K., Wunderlich, M., et al. (2011b). RNAi screen identifies Brd4 as a therapeutic target in acute myeloid leukaemia. Nature *478*, 524–528.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
RBM39/Caper-alpha	Bethyl Laboratories	Cat#A300-291A; RRID: AB_263411
RBM39	Atlas Labs (Sigma Aldrich)	Cat#HPA001591-100UL; RRID: AB_1079749
Beta-actin antibody	Sigma Aldrich	Cat#A5441-.5ML; RRID: AB_476744
Anti-MYB antibody	Thermo Fisher Scientific	Cat#05175MI
BMI1 antibody	Abcam	Cat#ab38295; RRID: AB_725719
Beta-Actin Peroxidase antibody	Sigma Aldrich	Cat#A3854-200UL; RRID: AB_262011
Mouse IgG	Santa Cruz	Cat#Sc-2025; RRID: AB_737182
Bacterial and Virus Strains		
MegaX DH10B Electrocomp Cells	Life Technologies	Cat#C640003
Chemicals, Peptides, and Recombinant Proteins		
Indisulam	Sigma Aldrich	Cat#SML1225
E7820	Eisai	N/A
(2-Hydroxypropyl)-β-cyclodextrin powder	Sigma Aldrich	Cat#H107
D-Luciferin, Potassium Salt	Gold Bio Technology	Cat#LUCK-500
Puromycin	Sigma Aldrich	Cat#P7255-25MG
Polyethylenimine (PEI)	Polysciences, Inc	Cat#23966
Polybrene transfection reagent	Fisher Scientific	Cat#TR1003G
Dimethyl sulfoxide (DMSO)	Fisher	Cat#D128-500
Critical Commercial Assays		
Click-iT Plus EdU Alex Fluor 647	Thermo Fisher Scientific	Cat#C10340
APC Annexin V kit	BD bioscience	Cat#BDB559763
NEXTflex Rapid Directional RNA-Seq Kit	Bioo Scientific	Cat#5138-07
QIAamp DNA Mini Kit	Qiagen	Cat#51304
NEBNext Poly(A) mRNA Magnetic Isolation	New England Biolabs	Cat#E7490S
NEXTflex RNA-seq Barcodes-12	Bioo Scientific	Cat#NOVA-512912
RNeasy Plus Mini Kit	Qiagen	Cat#74136
SilverQuest staining	Invitrogen	Cat#LC6070
2x Phusion Master Mix	Thermo Scientific	Cat#F-548
Deposited Data		
RNA-seq Raw data	This paper	GSE114558
RBM39 eCLIP Raw data	This paper	GSE114558
RBM39 splicing analysis	This paper	GSE114558
Experimental Models: Cell Lines		
Human: MOLM-13	DSMZ	ACC 554
Human: K562	ATCC	CCL-243
Human: NALM6	ATCC	CRL-3273
Human: MV4-11	ATCC	CRL-9591
Human: HNT34	DSMZ	ACC 600
Human: NKM-1	JCRB	IFO50476
Human: TF-1	ATCC	CRL-2003
Human: TF-1 *SF3B1*$^{K700E/WT}$ knockin cells	Horizon Discovery	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human: Monomac6	DSMZ	ACC 124
Human: HEL	DSMZ	ACC 11
Human: K052	JCRB	JCRB0123
Human: SET2	DSMZ	ACC 608
Human: UKE-1	Fiedler et al., 2000	N/A
Human: THP-1	ATCC	TIB-202
Human: OCI-AML3	DSMZ	ACC 582
Human: KG-1	ATCC	CCL-246
Human: HL-60	ATCC	CCL-240
Human: KASUMI-1	ATCC	CRL-2724
Human: NOMO-1	DSMZ	ACC 542
Human: A549	Thales Papagiannakopoulos Lab	N/A
Human: CUTLL1	Adolfo Ferrando Lab	N/A
Human: HEK293T	ATCC	CRL-1573
Mouse: MLL-AF9 $Nras^{G12}$ (RN2) Cas9	Zuber et al., 2011a	N/A
Human: JURKAT	ATCC	TIB-152
Human: SK-MEL239	Eva Hernando Lab	N/A
Human: 501MEL	Eva Hernando Lab	N/A
Human: K562	ATCC	CCL-243
Human: K562 *SF3B1*$^{K700E/WT}$, *SF3B1*$^{K666N/WT}$, *U2AF1*$^{S34F/WT/WT}$, *SRSF2*$^{P95H/WT}$ knockin cells	PUBMED ID: 30107174	N/A
Human: NALM6 *SF3B1*$^{K700E/WT}$, *SF3B1*^{K666NWT,} *SF3B1*^{H662QWT} knockin cells	PUBMED ID: 30107174	N/A
Experimental Models: Organisms/Strains		
Mouse: C57BL/6	N/A	N/A
NOD scid gamma	JAX	Cat#005557
Oligonucleotides		
sgRNA sequence for RBP domain library, see Table S6	This paper	N/A
sgRNA sequences for RBP CRISPR scanning, see Table S6	This paper	N/A
In vivo mouse Rbm39 sgRNA, see Table S6	This paper	N/A
sgRNA for DCAF15, see Table S6	This paper	N/A
Human DCAF15 probe set, Hs00384913_m1	Life technologies	Cat#4331182
Human 18S rRNA probe set, Hs99999901_s1	Life technologies	Cat#4448485
Human GAPDH probe set, Hs02786624_g1	Life technologies	Cat#4448485
Recombinant DNA		
MSCV-Cas9 puro	https://www.nature.com/articles/nbt.3235	Addgene: 65655
psPAX2	N/A	Addgene: 12260
pVSVG	N/A	Addgene: 12259
LRG sgRNA vector	https://www.nature.com/articles/nbt.3235	Addgene: 65656
Software and Algorithms		
FlowJo V8.7	TreeStar (BD Biosciences)	https://www.flowjo.com/
Prism 7.0	GraphPad	https://www.graphpad.com
Living Image Software	Perkin Elmer	http://www.perkinelmer.com/product/li-software-for-spectrum-1-seat-add-on-128113
GSEA	Broad Institute	http://software.broadinstitute.org/gsea/

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
rMATS	Xing Lab, UCLA	http://rnaseq-mats.sourceforge.net/
ImageJ	NIH	https://imagej.nih.gov/ij/index.html
IGV	Broad Institute	http://software.broadinstitute.org/software/igv/
Other		
PerkinElmer IVIS Lumina III XR	N/A	N/A
Illumina MiSeq	N/A	N/A
BD LSR FORTESSA	N/A	N/A
Illumina HiSeq 4000	N/A	N/A
QuantStudio 6 Flex Real-Time PCR system	N/A	N/A

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Iannis Aifantis (ioannis.aifantis@nyumc.org).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines and Cell Culture

All human and mouse leukemia cells were cultured in recommended media, typically RPMI medium with 10% FBS and 1% penicillin/streptomycin. The NALM-6 isogenic cell lines (NALM-6 cells engineered to express the single mutations *SF3B1*^{K700E}, *SF3B1*^{K666N}, *SF3B1*^{H662Q} from the endogenous *SF3B1* locus) were cultured in RPMI/10% FBS and 1% penicillin/streptomycin. K562 isogenic cell lines (engineered to express *SF3B1*^{K700E}, *SF3B1*^{K666N}, *SRSF2*^{P95H}, or *U2AF1*^{S34F} mutations from each respective endogenous locus) were cultured in IMDM/10% FBS. TF-1 isogenic cell lines (engineered to express *SF3B1*$^{wild\text{-}type}$ or *SF3B1*^{K700E} from the endogenous *SF3B1* locus by Horizon Discovery Inc. as described previously) (Seiler et al., 2018) were cultured in RPMI/10% FBS and 1% penicillin/streptomycin with 5ng/mL human GM-CSF. Genotyping and variant allele frequency of spliceosomal gene mutations in *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2* in each cell line was performed using the MSKCC IMPACT assay (Cheng et al., 2015; Zehir et al., 2017). The adherent cell lines, A549, 501MEL and HEK293T cells were grown in DMEM medium with 10% FBS and 1% penicillin streptomycin. SK-MEL239 were cultured in RPMI, 10% FBS and 1% penicillin streptomycin. Cell lines transduced with retroviral Cas9 puromycin (Addgene plasmid no. 65655) were selected with puromycin (Sigma Aldrich) 48 hours after transduction. All transfections were performed in HEK293T cells using Polyethylenimine (PEI) reagent at 4:2:3 ratios of sgRNA construct: pVSVG: pPax2 in OPTI-MEM solution. Viral supernatant was collected 48 hrs and 72 hrs post-transfection. Spin infections were performed at room temperature at 1,800 RPM for 30 mins with polybrene reagent (1:2000 dilution) (Fisher Scientific).

Primary Human Samples

Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center and conducted in accordance to the Declaration of Helsinki protocol. Primary human de-identified AML samples derived from whole peripheral blood or BM mononuclear cells were utilized. Mutational genotyping of each sample was performed by the MSKCC IMPACT assay as described previously (Cheng et al., 2015; Zehir et al., 2017). Cord blood was acquired from NY Blood Bank. Informed consent was obtained from all subjects to obtain the patient specimens used in the studies described. Specimens were obtained as part of the Memorial Sloan-Kettering Cancer Center Institutional Review Board approved clinical protocol #06-107 to which all subjects consented. O.A-W is a participating investigator on this protocol.

Animals

8-10 weeks-old C57BL/6 male mice were purchased from Jackson Laboratory. 10 week-old NOD scid gamma as well as NSG-SGM3 (NSGS) female mice were obtained from Jackson Laboratory. Mice were bred and maintained in individual ventilated cages and fed with autoclaved food and water at NYU School of Medicine Animal Facility as well as Memorial Sloan Kettering Cancer Center. All animal experiments were done in accordance with approved protocols from the Institutional Animal Care and Use Committees, according to national and institutional guidelines. Kaplan-Meier survival curves were compared using the Wilcoxon Rank-Sum test via GraphPad Prism. All animal experiments were performed in accordance with protocols approved by the New York University Institutional Animal Care and Use Committee (IACUC).

METHOD DETAILS

sgRNA Library Design

All sgRNAs in this study were designed using http://benchling.com/. Most sgRNAs had an off-target score of 70 or higher. RBP domain sgRNA library was designed to target 490 well-annotated RBDs based on NCBI database annotation. For each RBP gene, we designed on average 6-8 sgRNA for a total of ~2,900 sgRNAs. Library also consisted of positive sgRNAs targeting *RPA3*, *DOT1L*, *BRD4*, and *KMT2D* and the negative controls, sgRosa. Individual sgRNA oligos was synthesized by Twist Bioscience (https://twistbioscience.com/) on a 12K array and amplified using array primers (Table S1). Using a Gibson Assembly master mix (New England Biolabs), we cloned sgRNAs into a lentiviral sgRNA GFP-tagged vector (LRG) (Addgene plasmid no. 65656). Gibson reactions were transformed using DH10B electrocompetent cells (Invitrogen) at 2 kV, 200 Ω, and 25 uF. Bacterial colonies were quantified to obtain ~70X coverage. Subsequently, library was deep-sequenced using Miseq to confirm sgRNA representation. All sgRNA sequences used in this study are provided in Table S1.

CRISPR Screen

Cas9-expressing cell lines were transduced with sgRNA library virus at a low MOI (~0.3). On Day 4 post-transduction, GFP percentage was assessed to determine infection efficiency and sgRNA coverage (~300-500X). Remaining 300-500X cells were placed back into culture after each passage until 20 days' post-transduction. Genomic DNA (gDNA) of cells containing 300-500X coverage were harvested on Day 4 and Day 20 using Qiagen DNA kit based on manufacturer's protocol. For library construction, 200 ng of gDNA was amplified for 20 cycles using Phusion Master Mix. End repair products were generated using T4 DNA polymerase (NEB), DNA polymerase I (NEB) and T4 polynucleotide kinase (NEB). Subsequently, A-tail overhangs were added using Klenow DNA Pol Exo- (NEB). DNA fragments were then ligated with Illumina-compatible barcodes (Bioo Scientific) using T4 Quick Ligase (NEB) and amplified using pre-capture primers (5 cycles). Barcoded libraries were then sequenced using Mi-Seq instrument (150 cycles). For pooled CRISPR screen analysis, individual time-points for all samples were normalized using the formula (sgRNA read count/total read count) x 100,000. Subsequently, normalized reads were then used to calculate $\log_2$ fold change (normalized read count Day 4/normalized read count Day 20).

CRISPR/Cas9 Indel Analysis

To quantify the spectrum of indel mutations in mouse Rbm39 sgRNAs, we transduced RN2 cells with sgRbm39s, followed by cell sorting of GFP+/sgRNA+ populations at day 3 and day 20 post-infection. Cells were then harvested for gDNA and PCR amplicon (ranging from 100-200 bps) were designed to flank the sgRNA recognition sequence. 200 ng of gDNA was amplified using 2x Phusion Master Mix, followed by end repair using T4 DNA polymerase (NEB), DNA polymerase I (NEB) and T4 polynucleotide kinase (NEB). Subsequently, A overhangs were added following end repair using Klenow DNA Pol Exo- (NEB). Adaptor barcodes (Bioo Scientific) were then ligated and products were purified by AMPure beads. Finally, precapture PCR cycles were performed and libraries were sequenced using Mi-Seq instrument (150 cycles). Indel analysis was performed using CRISPResso (http://crispresso.pinellolab.partners.org/).

Apoptosis and Cell Cycle

To assess cell cycle, we used the 5-ethynyl-2'-deoxyuridine (EdU) incorporation using the Click-iT Plus EdU Alexa Fluor 647 kit and performed experiments as described in manufacturer's protocol (Life Technologies). Apoptotic analysis was determined using APC Annexin V (BD bioscience) and performed according to manufacturer's specifications and co-stained with 4',6-Diamidino-2-Phenylindole, Dihydrochloride (DAPI) for DNA content. Both EdU and Annexin V stained cells were analyzed by flow cytometry and FlowJo software.

RNA-Sequencing Library Preparation and Sequencing

CRISPR-mediated RBM39 knockout cells were harvested day 6 post-transduction and wash with 1x PBS. To extract total RNA, we used RNeasy Plus Mini Kit (Qiagen) and QIAshredder (Qiagen), which were performed according to manufacturer's protocol. We isolated poly(A) mRNA using magnetic isolation method (NEB) using ~1 ug of total mRNA. To generate RNA-sequencing libraries, we used NEXTflex Rapid Directional (Bioo Scientific) protocol and NEXTflex RNA-seq barcodes (Bioo Scientific), which were carried out according to manufacturer's protocol. RNA-seq barcoded libraries were then sequenced using Hi-Seq 4000 (150 cycles paired end).

For analysis of sulfonamide vs. E7107 treated K562 cells, RNA sequencing was performed in biological duplicate for each cell line and drug treatment. RNA was extracted using Qiagen RNeasy columns. poly(A)-selected, unstranded Illumina libraries were prepared with a modified TruSeq protocol. 0.5X AMPure XP beads were added to the sample library to select for fragments of 400 bp, followed by 1X beads to select for fragments 4100 bp. These fragments were then amplified with PCR (15 cycles) and separated by gel electrophoresis (2% agarose). 300 bp DNA fragments were isolated and 101 bp reads were sequenced on the Illumina HiSeq 2000 in paired end mode. For analysis of sulfonamide vs. E7107 treated K562 cells, 100 million paired end reads of 100 bp were performed for each sample.

eCLIP Library Preparation

eCLIP studies were performed in duplicates by Eclipse Bioinnovations Inc (San Diego, www.eclipsebio.com) according to the published single-end seCLIP protocol [[can cite pubmed ID 28766298]] with the following modifications. 10 million Molm13 cells were UV crosslinked at 400 mJoules/cm2 with 254 nm radiation, and snap frozen. Cells were then lysed and treated with RNase I to fragment RNA as previously described. RBM39 antibody (A300-291A lot 001, Bethyl) was then pre-coupled to Protein G Dynabeads (Thermo Fisher), added to lysate, and incubated overnight at 4 deg C. Prior to immunoprecipitation, 2% of the sample was taken as the paired input sample, with the remainder magnetically separated and washed with lysis buffer only (as the standard high-salt eCLIP wash buffer gave poor immunoprecipitation yield). eCLIP was performed by excising the area from ~65 kDa to ~140 kDa. RNA adapter ligation, IP-western, reverse transcription, DNA adapter ligation, and PCR amplification were performed as previously described.

Western Blotting

Cell lines were treated with the indicated dose of drug for 24 hours. Lysate protein concentration was measured with the BCA reagent and 10 mcg was loaded per lane onto 4-12% bis-tris protein gels. After transfer, PVDF membranes were probed with anti-RBM39 rabbit polyclonal (Atlas Laboratories) at 1:200, anti-RBM39 (Bethyl Laboratories) at 1:1,000, BMI-1 (Abcam), and MYB (Thermo Fisher Scientific) and visualized by standard methods. Western blot densitometry was performed using ImageJ.

Sulfonamide Drug Treatment IC_{50} Measurements

E7820 (molecular weight: 336.37 grams/mole) was provided by Eisai Pharmaceuticals. For *in vitro* experiments, E7820 was dissolved in DMSO to make a 100 micromolar stock solution, and this was then added to tissue culture media to the appropriate final concentration. Cell lines were plated in 96 well plates and exposed to the indicated sulfonamide compound at concentrations ranging from 1 micromolar to 1 nanomolar with a minimum of four technical replicates per concentration per cell line. Cell viability was measured with the CellTiter Glo reagent (Promega) as per manufacturer's instructions. Absolute viability values were converted to percentage viability versus DMSO control treatment, and then non-linear fit of log(inhibitor) versus response (three parameters) was performed in GraphPad Prism v7.0 to obtain an IC_{50} values. Experiments were performed at least in duplicate or triplicate.

qRT-PCR Measurement of Gene Expression

RNA was extracted from the indicated cell lines and reverse transcribed into cDNA using the Verso cDNA synthesis Kit (ThermoFisher Scientific). Measurement of *DCAF15* (Hs00384913_m1) gene expression was performed using Taqman probes (Life Technologies) with *GAPDH* (Hs02786624_g1) or *18S* ribosomal RNA (Hs99999901_s1) as the housekeeping gene. Assay IDs for other primers: HoxA9 Hs00365956_m1, BMI1 Hs00180411_m1, MYB Hs00920556_m1, GATA2 Hs00231119_m1. Relative expression levels across cell lines were calculated using the Delta-delta Ct method as per standard procedures.

RBM39 cDNA Overexpression

RBM39 was PCR amplified from MOLM-13 cDNA and cloned into MSCV Puro-IRES-GFP construct (Addgene #18751). *RBM39* G268V retroviral overexpression plasmid was a gift from Eisai Pharmaceuticals, G268V CDNA was amplified and subsequently cloned into MSCV Puro-IRES-GFP plasmid. MOLM-13 cells were transduced with either empty vector, *RBM39*, or *RBM39* G268V construct and high GFP+ expressing cells were sorted for *in vitro* indisulam experiments.

Immunoprecipitation

150 million MOLM-13 cells were harvested and washed twice with 1X PBS. Cell pellets were resuspended with Cytosolic Hyptonic Buffer (10 mM HEPES pH 7.9, 10 mM KCl, 1.5 mM MgCl2, and 0.1 mM EDTA) containing protease inhibitor (PI) and phosphatase inhibitors. Subsequently, cell pellets were resuspended in Nuclear Extraction Buffer (20 mM HEPES pH 7.9, 400 mM NaCl, 1.5 mM MgCl2, 0.4% Triton X-100, and 1 mM EDTA) with PI and phosphatase inhibitors. Supernatant was extracted and diluted with Equilibrium Buffer (20 mM HEPES pH 7.9, 10% glycerol and 1 mM EDTA). Cell extracts was incubated with 2 ug of antibody overnight. Next day, cell extract containing antibodies were incubated with protein A magnetic beads and washed with BC-100 (20 mM HEPES pH 7.9, 100 mM NaCl, 10% glycerol, 0.4% Triton X-100, and 1 mM EDTA) and resuspended with loading buffer and boiled for 10 mins at 95°C. Supernatant was taken and western blot analysis was performed. Silver staining was performed with SilverQuest (Invitrogen) according to manufacturer's instruction.

IP-MS

High resolution full MS spectra were acquired with a resolution of 70,000, an AGC target of 1e6, with a maximum ion time of 120 ms, and scan range of 400 to 1500 m/z. Following each full MS twenty data-dependent high resolution HCD MS/MS spectra were acquired. All MS/MS spectra were collected using the following instrument parameters: resolution of 17,500, AGC target of 5e4, maximum ion time of 120ms, one microscan, 2 m/z isolation window, fixed first mass of 150 m/z, and NCE of 27. MS/MS spectra were searched against a Uniprot human + mammalian IgG database using Sequest within Proteome Discoverer.

Animal Experiments

For *in vivo* Cas9 experiments, RN2 Cas9-expressing cells were transduced with sgRosa (negative control) (n=4) or sgRbm39 constructs (n=7/group). At day 2 post-transduction, sgRNA positive cells (GFP+) were sorted by FACS. One million leukemia-sgRNA

expressing cells were intravenously injected into each sub-lethal irradiated (5.5 Gy) B6/SJL recipient mice. For indisulam trials, a 50 mg/ml indisulam (Sigma Aldrich) stock was diluted in 20% (2-Hydroxypropyl)-β-cyclodextrin (Sigma Aldrich) to obtain a final concentration of 25mg/kg. NOD scid mice were intravenously injected with 1 million MOLM-13-expressing luciferase cells. Upon disease onset as measured by bioluminescent imaging, we intraperitoneally injected once daily with either 25 mg/kg indisulam or vehicle (1% DMSO) (n=6/group) for 13 consecutive days. All whole-body bioluminescent imaging was performed by intraperitoneally injection of Luciferin (Goldbio) at a 50 mg/kg concentration and imaging was performed after 5 mins using an IVIS imager. Bioluminescent signals (radiance) were quantified using Living Image software with standard regions of interests (ROI) rectangles. AML patient-derived xenografts were generated from patient peripheral blood and/or bone marrow mononuclear cells and subsequently transplanted intrafemorally into NSGS mice and eventually treated with indisulam (25mg/kg/daily) for 12-15 days. *In vivo* analysis of indisulam toxicity in C57/B6 mice was performed by intraperitoneally injection for 4 weeks (5 days on/2 days off). Additionally, human $CD34^+$ cells were purified by MACS $CD34^+$ from mononuclear cells from cord blood and intrafemorally injected into NSG and eventually treated with two weeks of indisulam (25mg/kg/daily) or vehicle. $hCD34^+$ were also purified from mice following treatment for Western blotting analysis. Indisulam was also obtained from Eisai Pharmaceuticals.

Immunohistochemistry

Immunohistochemistry was performed on bone marrow and spleen mice treated with vehicle (DMSO) or 25 mg/kg indisulam fixed, paraffin-embedded, 5-μm tissue sections collected on plus slides (Fisher Scientific, Cat # 22-042-924) and stored at room temperature. Polyclonal rabbit RBM39 antibody (Bethyl Laboratories Cat# A300-291A) was used for immunohistochemistry. Antigen retrieval conditions were determined followed by serial dilution of the primary antibody to determine optimum dilution/concentration. A nuclear staining pattern in was considered positive. Chromogenic immunohistochemistry was performed on a Ventana Medical Systems Discovery XT instrument using Ventana's reagents and detection kits unless otherwise noted. In Brief, sections were deparaffinized online. Antigen retrieval was performed using Cell Conditioner 2 (Citrate) for 20 minutes. Endogenous peroxidase activity was blocked for all samples. CAPER was diluted 1:500 in Ventana antibody diluent (Ventana catalog number 760-219) and incubated for 3 hours at 37°C. Primary antibody was detected using goat anti-rabbit horseradish peroxidase conjugated multimer incubated for 8 minutes. The complex was visualized with 3,3 diaminobenzidene and enhanced with copper sulfate. Slides were washed in distilled water, counterstained with hematoxylin, dehydrated and mounted with permanent media. Negative controls consisted of diluent only tested with the study sections.

QUANTIFICATION AND STATISTICAL ANALYSIS

Splicing Analysis

Significant changes in alternative splicing events and constitutive spliced exons and intron in response to CRISPR knockout or treatment depletion were quantified by MATS v4.0.2 (ref. MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data https://doi.org/10.1093/nar/gkr1291). Analyses were restricted to events with 10 or more reads. Events were defined as Significant if (i) the FDR corrected p value was smaller than 0.1 and (2) absolute Inclusion Level Difference was larger than 10%. Gene set enrichment was performed using the fgsea R package (1.4.0) using the KEGG, GO and MsigDB specific signatures according to the manual. Specific examples of splicing and mis-splicing events were visualized with IGV (Broad Institute).

eCLIP Data Analysis

The eCLIP data was processed similarly as described previously (Van Nostrand et al., 2016) and is outlined shortly in the following. First, adapter sequences were trimmed from both reads of all read-pairs using cutadapt version 1.14. Then, all remaining reads longer than 16 bases were aligned against the human reference genome sequence hg19/GRCh37 using STAR version 2.5.0c. Only uniquely mapped reads were kept. Read-pair duplicates by position were removed using picard tools version 2.6.0. To identify binding sites, we first ran a custom script to identify clusters of overlapping reads that had a read-depth of at least 10 reads. Then, we calculated significant enrichments for all such identified clusters by comparing IP-samples versus input-samples using edgeR. More specifically, we ran bamutils count version 0.5.7 to counted stranded reads within all identified clusters for all samples. Using this output, we calculated differential coverage between IP-vs-input for each cluster with edgeR after normalizing for total sequencing depth per replicate (resulting in counts per million/CPM per cluster). Final binding sites were called by applying logFC > 2 and FDR < 0.05 thresholds between IP-vs-input.

Gene Ontology Analysis

Gene set enrichment was performed using the fgsea R package (1.4.0) using the KEGG, GO and MsigDB specific signatures according to the manual.

AML TCGA Analysis

Raw RNA-seq reads from 28 Bone marrow progenitor cell populations and 43 AML patient from the Leucegene data set were retrieved from the US National Center for Biotechnology Information (NCBI) sequence read archive (GEO accession numbers

GSE74246, GSE63569, GSE49642). Gene expression quantification was performed using the TCGA pipeline method. Briefly, Reads were aligned with STAR (2.5.3) using the GRCh38.p7 (May 2017) human assembly (https://gdc.cancer.gov/about-data/data-harmonization-and-generation/gdc-reference-files) and gene level expression were quantified using HTSeq (ref. HTSeq—a Python framework to work with high-throughput sequencing data https://doi.org/10.1093/bioinformatics/btu638). Gene expression of 151 AML patients from the TCGA project quantified by HTSeq was retrieved from GDC TCGA-LAML data portal (https://portal.gdc.cancer.gov/projects/TCGA-LAML). Additionally, 43 AML patient samples from Leucegene were obtained and included in our gene expression analysis (Lavallee et al., 2015a, 2015b, 2016; Pabst et al., 2016). Raw counts were transformed and normalized using the voom method (ref. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR doi: 10.12688/f1000research.9005.2). We evaluated *RBM39* expression across known risk groups of AML patients from the TCGA based on (i) the 2017 European LeukemiaNet (ELN) risk stratification system for AML (Dohner et al., 2017) and (ii) standard AML cytogenetic risk groups (Dohner et al., 2017).

Statistical Analysis

Kaplan-Meier survival curve p values were performed using Log rank Mantel-COX test. For statistical comparison, we performed unpaired Student's t test. Statistical analyses were performed using Prism 7 software (GraphPad). Data with statistical significance are as indicated, *p< 0.05, **p< 0.01, ***p< 0.001.

DATA AND SOFTWARE AVAILABILITY

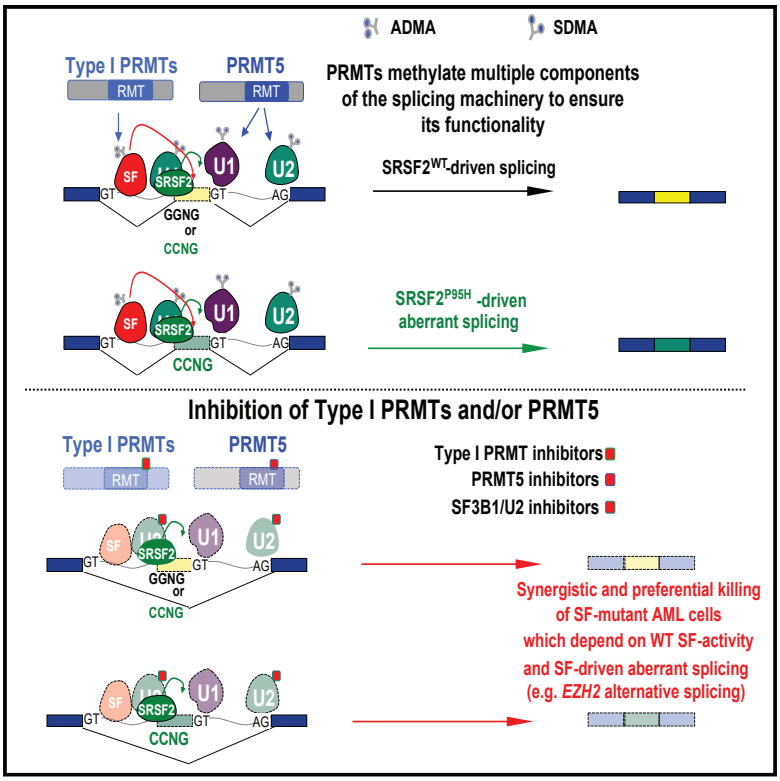
Gene Expression Omnibus: all newly generated RNA-seq data were deposited under accession number GSE114558.

Article

Cancer Cell

Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation

Graphical Abstract



Highlights

- Inhibition of SDMA or ADMA preferentially kills splicing factor (SF)-mutant cells
- Combined inhibition of PRMT5 and type I PRMTs has synergistic effects
- RNA-binding proteins are the most enriched cellular substrates of PRMTs
- Inhibition of RNA splicing underlies the cytotoxic effects of PRMT inhibition

Authors

Jia Yi Fong, Luca Pignata, Pierre-Alexis Goy, ..., Tiziana Bonaldi, Omar Abdel-Wahab, Ernesto Guccione

Correspondence

abdelwao@mskcc.org (O.A.-W.), ernesto.guccione@mssm.edu (E.G.)

In Brief

Fong et al. show that spliceosomal mutant leukemias are preferentially sensitive to inhibition of protein arginine methyltransferases (PRMTs), that RNA-binding proteins are enriched among substrates of PRMT5 and type I PRMTs, and that combined PRMT5 and type I PRMT inhibition synergistically kill these leukemia cells.

Fong et al., 2019, Cancer Cell *36*, 194–209
August 12, 2019 © 2019 Published by Elsevier Inc.
https://doi.org/10.1016/j.ccell.2019.07.003

CellPress

CellPress

Cancer Cell
Article

Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation

Jia Yi Fong,[1,2] Luca Pignata,[1,3] Pierre-Alexis Goy,[1,3] Kimihito Cojin Kawabata,[4] Stanley Chun-Wei Lee,[5] Cheryl M. Koh,[1,16] Daniele Musiani,[6] Enrico Massignani,[6] Andriana G. Kotini,[8] Alex Penson,[5] Cheng Mun Wun,[1] Yudao Shen,[7] Megan Schwarz,[8] Diana HP. Low,[1] Alexander Rialdi,[8] Michelle Ki,[5] Heike Wollmann,[1] Slim Mzoughi,[1] Florence Gay,[1] Christine Thompson,[9] Timothy Hart,[9] Olena Barbash,[9] Genna M. Luciani,[10] Magdalena M. Szewczyk,[10] Bas J. Wouters,[11,12] Ruud Delwel,[12] Eirini P. Papapetrou,[8] Dalia Barsyte-Lovejoy,[10] Cheryl H. Arrowsmith,[10,13] Mark D. Minden,[13] Jian Jin,[7] Ari Melnick,[4] Tiziana Bonaldi,[6] Omar Abdel-Wahab,[5,14,*] and Ernesto Guccione[1,3,8,15,17,*]

[1]Institute of Molecular and Cell Biology, Agency for Science, Technology and Research (A*STAR), Singapore 138673, Singapore
[2]NUS Graduate School for Integrative Sciences and Engineering, National University of Singapore, Singapore 119077, Singapore
[3]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore 119077, Singapore
[4]Departments of Medicine and Pharmacology, Weill Cornell Medicine, New York, NY 10065, USA
[5]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[6]Department of Experimental Oncology, IEO, European Institute of Oncology IRCCS, 20146 Milan, Italy
[7]Mount Sinai Center for Therapeutics Discovery, Departments of Pharmacological Sciences and Oncological Sciences, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[8]Department of Oncological Sciences and Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[9]Epigenetics Research Unit, GlaxoSmithKline, Collegeville, PA 19426, USA
[10]Structural Genomics Consortium, University of Toronto, Toronto, ON M5G 1L7, Canada
[11]Division of Hematology and Medical Oncology, Department of Medicine, Weill Medical College of Cornell University, New York, NY 10065, USA
[12]Department of Hematology, Erasmus University Medical Center, 3015 GD Rotterdam, Netherlands
[13]Ontario Cancer Institute/Princess Margaret Hospital, Toronto, ON M5G 2M9, Canada
[14]Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA
[15]Department of Pharmacological Sciences and Mount Sinai Center for Therapeutics Discovery, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA
[16]Present address: Silicon Therapeutics, Boston, MA, USA
[17]Lead Contact
*Correspondence: abdelwao@mskcc.org (O.A.-W.), ernesto.guccione@mssm.edu (E.G.)
https://doi.org/10.1016/j.ccell.2019.07.003

SUMMARY

Cancer-associated mutations in genes encoding RNA splicing factors (SFs) commonly occur in leukemias, as well as in a variety of solid tumors, and confer dependence on wild-type splicing. These observations have led to clinical efforts to directly inhibit the spliceosome in patients with refractory leukemias. Here, we identify that inhibiting symmetric or asymmetric dimethylation of arginine, mediated by PRMT5 and type I protein arginine methyltransferases (PRMTs), respectively, reduces splicing fidelity and results in preferential killing of SF-mutant leukemias over wild-type counterparts. These data identify genetic subsets of cancer most likely to respond to PRMT inhibition, synergistic effects of combined PRMT5 and type I PRMT inhibition, and a mechanistic basis for the therapeutic efficacy of PRMT inhibition in cancer.

Significance

Pharmacologic suppression of PRMT5 and type I PRMTs is being pursued as a cancer treatment approach, and numerous mechanisms have been proposed for the efficacy of PRMT inhibition. Here we identify that spliceosomal mutant leukemias are preferentially sensitive to PRMT inhibition and that RNA-binding proteins are the most enriched substrates of PRMT5 and/or type I PRMTs. Accordingly, combined PRMT5 and type I PRMT inhibition resulted in synergistic cell killing and pronounced effects on splicing compared with inhibiting either enzymatic activity alone. These data provide a rational combinatorial strategy and a strong basis for ongoing and future clinical trials based on the presence of genetic alterations impacting RNA splicing.

 © 2019 Published by Elsevier Inc.

INTRODUCTION

Recent genomic analyses of cancers have identified numerous means by which splicing is altered in cancer (Dvinge et al., 2016; Kahles et al., 2018; Zhang and Manley, 2013). These include change-of-function mutations in RNA splicing factors (SFs) (Harbour et al., 2013; Martin et al., 2013; Wang et al., 2011; Yoshida et al., 2011), mutations that alter splicing in *cis* (Supek et al., 2014), changes in the expression of splicing regulatory proteins (Anczuków et al., 2012; Karni et al., 2007), and alterations in transcriptional regulators that influence the process of splicing such as c-MYC (Hsu et al., 2015; Koh et al., 2015). In parallel, certain genetic subsets of cancer have been identified to be particularly sensitive to therapeutic inhibition of splicing. These include cells bearing hotspot change-of-function mutations in the RNA SFs SRSF2, SF3B1, and U2AF1 (from hereon collectively referred to as SF-mutants) (Lee et al., 2016; Obeng et al., 2016; Seiler et al., 2018; Shirai et al., 2017), as well as solid tumors and lymphomas driven by *MYC* amplification and overexpression (Hsu et al., 2015; Koh et al., 2015).

Despite the large numbers of proteins involved in splicing, as well as post-translational modifications of splicing proteins and protein-protein, protein-RNA, and RNA-RNA interactions that regulate splicing, clinical efforts to inhibit splicing have largely utilized drugs inhibiting the interaction of the SF3B complex with RNA (Folco et al., 2011; Kaida et al., 2007; Kotake et al., 2007; Seiler et al., 2018). Currently, however, the safety of direct inhibition of RNA splicing catalysis in patients is unknown. More recently, a series of compounds that result in proteasomal degradation of the RNA SF RBM39 have provided an alternative pharmacologic means to perturb splicing (Han et al., 2017; Uehara et al., 2017). These compounds are effective at degrading RBM39 *in vivo* and lead to the deregulation of a splicing network required for acute myeloid leukemia (AML) survival (Wang et al., 2019).

PRMT family members regulate the activity of many proteins and their inhibition may affect splicing activities (Koh et al., 2015; Zhang et al., 2015b). With the observation that SF-mutant leukemias are more sensitive to further perturbation of splicing than their wild-type (WT) counterparts, we seek to identify if inhibitors of protein arginine methyltransferases (PRMTs) can preferentially target these leukemias.

RESULTS

Spliceosomal Interacting Proteins Are Targetable Vulnerabilities in SF-Mutant Cells

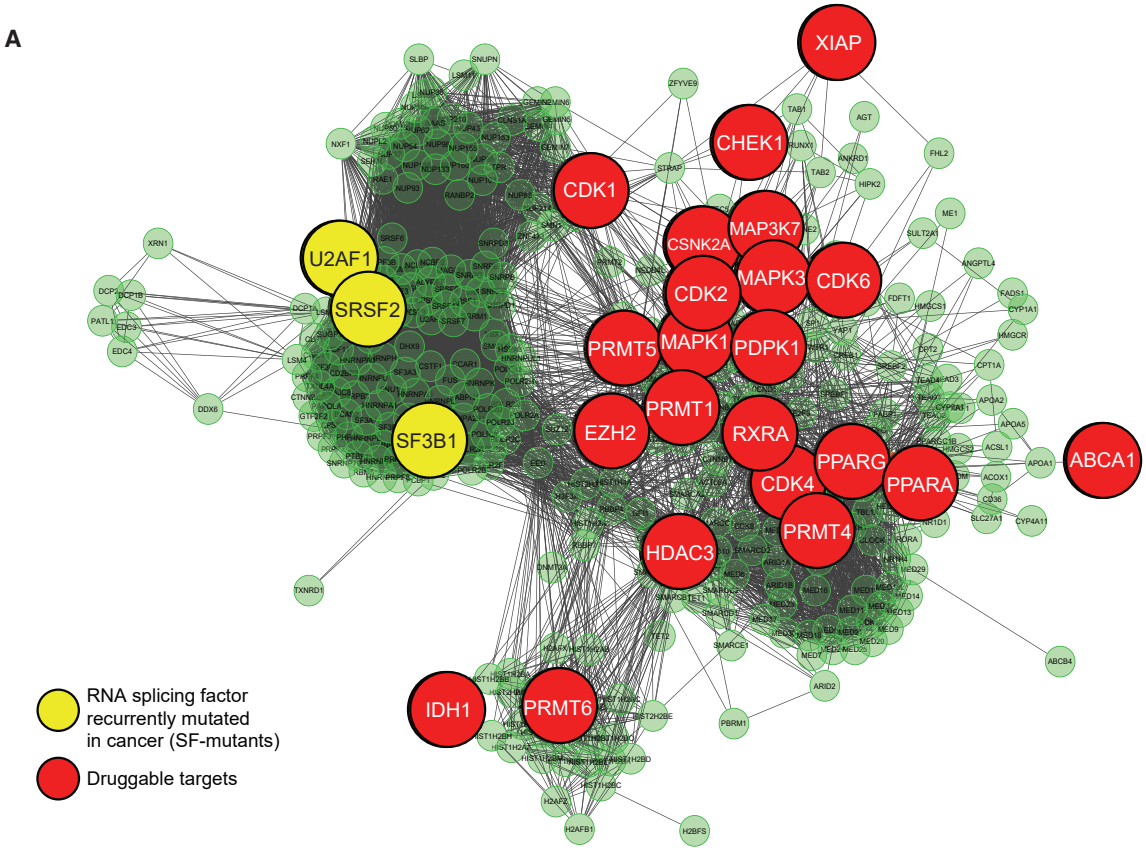
In an effort to identify additional means to therapeutically impact the process of splicing, we sought to identify proteins with functional relationships to components of the core splicing machinery, which might be druggable with available therapies. We built a network based on protein-protein (protein complexes or enzyme/substrate relationships), protein-DNA (transcriptional), and protein-RNA (post-transcriptional) regulation (Shannon et al., 2003), to obtain a list of 312 genes encoding an extended network of proteins with known or putative interactions with the core spliceosome. We then manually selected all druggable targets and sourced available inhibitors for these proteins (Figure 1A; Table S1). Given that most of the genes in the network (71%) were transcriptional regulators, and that splicing is a co-transcriptional process influenced by transcription, we also included small molecules inhibiting epigenetic regulatory proteins (chemical probe collection from the SGC, Toronto) as additional controls.

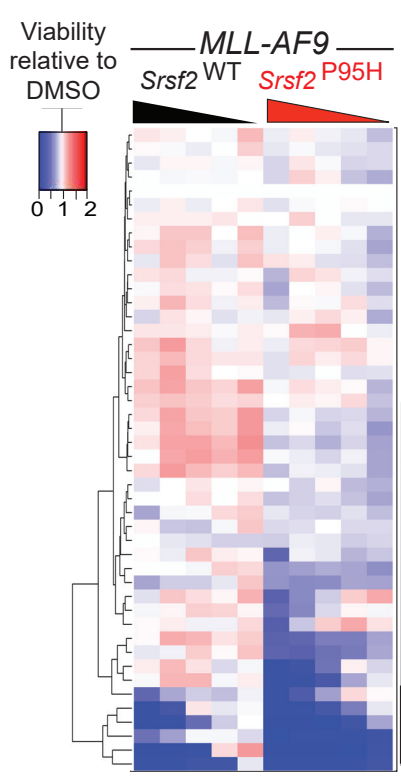
We assembled a panel of 45 compounds based on the above criteria, and performed a drug screen using a range of five concentrations of each drug for 7 days in isogenic murine AML cells driven by the *MLL-AF9* fusion in the absence or presence of mutant *Srsf2* (*MLL-AF9/Vav*-cre *Srsf2*^{WT} or *MLL-AF9/Vav*-cre *Srsf2*^{P95H} cells). The generation of these cells has been described previously in studies identifying the preferential sensitivity of spliceosomal mutant cells to SF3B1 inhibitory compounds, and their growth rate is equivalent both *in vitro* and *in vivo*, regardless of *Srsf2* genotype (Lee et al., 2016). Moreover, this is a genetically relevant model for human AML as *SRSF2* mutations occur in ~10% of adult MLL-rearranged AMLs (Lee et al., 2016). The hallmark global sequence-specific change in RNA splicing characteristic of mutant *SRSF2* is evident in the context of both human MLL-rearranged AML and murine *MLL/AF9* *Srsf2*^{P95H} cells (Lee et al., 2016).

We performed an *in vitro* drug screen and scored cell viability by MTS assay, after 7 days of drug treatment. Values were normalized to DMSO controls. *Srsf2*-mutant cells were more sensitive than WT counterpart cells to several inhibitors targeting components of the extended splicing network, including the SF3B complex inhibitor E7107 (Figure 1B; Table S1). Moreover, distinct inhibitors of PRMTs resulted in preferential killing of *Srsf2*-mutant AML cells over WT counterpart cells. These included GSK3203591 (abbreviated GSK591), a selective inhibitor of PRMT5 (Duncan et al., 2016), and MS023, a pan type I PRMTs inhibitor (Eram et al., 2016). Whereas type I PRMTs include PRMT1, 3, 4, and 6, selective inhibitors for PRMT3, PRMT4, or PRMT6 did not impair proliferation of *Srsf2*-mutant cells, suggesting that PRMT1 was the critical target of MS023 in this context.

PRMT5 and Type I PRMT Inhibition Preferentially Affect SF-Mutant AMLs

We next evaluated the effects of PRMT inhibitors on *Srsf2*^{WT} versus *Srsf2*^{P95H} cells over a range of drug concentrations to verify the results of the screen and to determine the half maximal inhibitory concentration for each compound (Figures 2A and 2B). As calculated by CellTiter Glo cell viability assay, *Srsf2*^{P95H} cells were approximately 10-fold more sensitive than their *Srsf2*^{WT} isogenic counterpart to GSK591 (120 versus 12 nM in *Srsf2*^{WT} versus *Srsf2*^{P95H}-mutant cells, respectively) and MS023 (1,100 versus 120 nM in *Srsf2*^{WT} versus *Srsf2*^{P95H}-mutant cells, respectively). Similar preferential induction of apoptosis was seen for *Srsf2*^{P95H}-mutant versus WT cells for both GSK591- and MS023-treated cells using a Caspase-Glo 3/7 (Figure S1A). The SF3B1 inhibitor E7107 was used in the same assays as a positive control, given previous data demonstrating enhanced sensitivity of *Srsf2*^{P95H} cells to E7107 (Lee et al., 2016) (Figure 2C). Overall, these experiments identified that SF-mutant AML cells were >10 times more sensitive to either PRMT5 or type I PRMT inhibition than WT counterparts. The preferential killing of SF-mutant cells was mirrored by dose-dependent reductions of symmetric dimethylarginine and asymmetric





DRUG	TARGET
I-BRD9	BRD9
SGC-CBP30	CBP, EP300 bromo
BI-9564	BRD9/7
I-CBP112	CBP, EP300 bromo
MS049	PRMT4/6
UNC0638	G9a, EHMT1
SGC707	PRMT3
GSK484	PAD4
IOX2	PHD2
GSK-J4	JMJD3, UTX, Jarid1B
Vorinostat	HDAC3
GSK2334470	PDK1
SGC0946	DOT1L
PD0332991	CDK4/6
CCT244747	CHK1
PF-477736	CHK1
BMS-265246	CDK4/CycD
Irinotecan	Topoisomerase 1
SCH772984	ERK1/2
Alitretinoin	Agonist Retinoic Receptor
CX-4945	CK2 protein kinase
E7107	**SF3B1**
GSK591	**PRMT5**
MS023	**PRMT Type I**
Phenoxodiol	**XIAP1**
5Z-7-oxozeaenol	**ERK2, TAK1 (MEKK7), MKK7, and MEK1**

(legend on next page)

dimethylarginine (SDMA and ADMA), respectively (Figures S1B and S1C), thereby also validating the on-target efficacy of these compounds. Moreover, we validated the increased sensitivity of SF-mutant AMLs to PRMT inhibitors across a panel of 32 genetically annotated human AML samples, half of which contained hotspot mutations in *SF3B1*, *SRSF2*, or *U2AF1* (Figure 2D; Table S2). In this assay, primary AML mononuclear cells were exposed to GSK591 or DMSO for 6 days and evaluated by the number of viable cells at the end of culture. SF-mutant AML cells had a significantly increased sensitivity to GSK591 compared with SF-WT counterparts, an effect seen in *SF3B1*-, *U2AF1*-, and *SRSF2*-mutant samples. Similar effects were additionally observed in drug treatment of nine AML cell lines (four of which had mutations in *SRSF2*, *U2AF1*, or *SF3B1*; Figures S1D and S1E). This latter experiment was performed with the type I PRMT inhibitor GSK3368712, which has very similar potency and selectivity as MS023 (Fedoriw et al., 2019). Importantly these two latter experiments demonstrated that the increased sensitivity of SF-mutant AMLs to PRMT inhibitors extends beyond *SRSF2*-mutant leukemias to leukemias harboring mutations in *SF3B1* and *U2AF1* as well.

PRMT Inhibition Preferentially Affects SF-Mutant AMLs *In Vivo*

We next tested the sensitivity of leukemia cells to either PRMT5 or type I PRMT inhibitors based on SF-mutation status *in vivo*. We used the orally available PRMT5 inhibitor EPZ015666 dosed once a day at 200 mg/kg. Equal numbers of $Srsf2^{WT}$ or $Srsf2^{P95H}$ *MLL-AF9*-transformed leukemia cells were transplanted into secondary recipient mice and treatment was started at 8 days post-transplant to allow tumor engraftment (Figure S2A). EPZ015666 treatment increased the survival of mice transplanted with $Srsf2^{P95H}$ leukemias but not mice transplanted with $Srsf2^{WT}$ leukemias (Figure 3A). Western blot analysis demonstrated downregulation of global SDMA levels in spleen cells from mice treated with EPZ015666 (Figure 3B).

To evaluate the effect of type I PRMT inhibition on leukemias *in vivo*, we first identified the pharmacokinetic (PK) properties of MS023 (Figure S2B) and established an intraperitoneal dose of 80 mg/kg/day as optimal to downregulate ADMA in most organs (Figure S2C). Similar to that observed upon PRMT5 inhibition, treatment with MS023 delayed disease progression in mice transplanted with $Srsf2^{P95H}$ leukemias but not in the case of $Srsf2^{WT}$ AMLs (Figure 3C). Western blot analysis confirmed downregulation of global ADMA in bone marrow and spleen cells from mice treated with MS023. Concurrent upregulation of SDMA was observed, in accordance with previous reports of a scavenging effect by PRMT5 observed upon type I PRMTs inhibition (Dhar et al., 2013) (Figure 3D). Both EPZ015666 and MS023 were well tolerated in the mouse PK, and *in vivo* efficacy studies and no apparent toxicity was observed (Chan-Penebre et al., 2015) (and data not shown).

Synergistic Effects of Combined PRMT5, Type I PRMT, and/or SF3B Inhibition

Given the upregulation of SDMA seen with type I PRMT inhibition and the preferential sensitivity of SF-mutant AMLs to inhibition of symmetric as well as asymmetric arginine dimethylation, we next sought to evaluate the effects of simultaneous inhibition of both axes of PRMT function. To evaluate the synergism between the two inhibitors, we calculated their combination index (CI), which quantitatively defines synergism ($CI < 1$), additive effect ($CI = 1$) and antagonism ($CI > 1$) among two drugs (Chou, 2010). Combined *in vitro* treatment of $Srsf2^{WT}$ or $Srsf2^{P95H}$ *MLL-AF9* leukemia cells with MS023 and GSK591, MS023 and E7107, or GSK591 and E7107, revealed synergistic effects of each combination, regardless of *Srsf2* mutational status (Figures 4A–4C). A similar synergistic effect was seen *in vivo* where combined MS023 (60 mg/kg) and EPZ015666 (150 mg/kg) treatment (given starting at day 8 after engraftment for 10 days) resulted in significantly increased survival in recipient mice engrafted with both $Srsf2^{WT}$ and $Srsf2^{P95H}$ *MLL-AF9* leukemia cells (Figure 4D). Consistent with the extended survival seen with combined MS023 and EPZ015666 treatment, PRMT inhibition *in vivo* resulted in reductions in both ADMA and SDMA levels in splenic tissue (Figure S2D).

We then tested the combinatorial effects of PRMT5 and type I PRMT inhibition in additional models. First, we assessed the effect of combined treatment with GSK591 and MS023 in human AML cell lines (Figures 5A–5C). Importantly, the synergy between the two drugs extent beyond MLL-AF9-driven malignancies and was also observed in those with CALM-AF10 (U937) and BCR-ABL (K562) fusions.

Second, we used induced pluripotent stem cells (iPSCs) with a $SRSF2^{P95L}$ mutation knocked in by CRISPR and isogenic normal iPSCs (Chang et al., 2018). We differentiated the iPSCs into hematopoietic progenitor cells and treated them with PRMT inhibitors. The *SRSF2*-mutant cells were more sensitive than their isogenic WT control (Figure 5D).

Third, we tested three primary patient AML samples, two had $SRSF2^{P95H}$ and one had $SF3B1^{Y765C}$, *in vitro* and found that GSK591/MS023 combination treatment resulted in a synergistic effect of reducing cell viability (Figure 5E; Table S3).

Fourth, we tested the efficacy of drug combination on a spliceosomal WT/*MLL*-rearranged and one $SF3B1^{K700E}$/*EVI1*-rearranged AML PDX models. Treatment with 60 mg/kg MS023 and 150 mg/kg EPZ015666 for 30 days resulted in reduction of human chimerism only in the SF-mutant AML (Figures S3A–S3C).

Figure 1. Spliceosomal Interacting Proteins Are Targetable Vulnerabilities in Spliceosomal Mutant Cells

(A) Molecular interaction network generated by Cytoscape 3.4.0 (Shannon et al., 2003) displaying proteins involved in RNA splicing, snRNP assembly, and/or mutated in acute myeloid leukemia (AML) and their nearest neighbors of a given physical entity (e.g., genes or proteins). Genes and proteins are illustrated by nodes, which are connected by lines to nodes based on physical or functional interaction. Nodes that are RNA SFs mutated in cancer are displayed in yellow while those that are druggable targets are indicated in red.

(B) Heatmap of the relative viability of *MLL-AF9*/*Vav*-cre $Srsf2^{WT}$ and *MLL-AF9*/*Vav*-cre $Srsf2^{P95H}$ to the indicated compounds following 7 days of growth scored by MTS assay and reported as a ratio to control DMSO-treated cells. Blue indicates a reduction, while red indicates an increase in cell viability, relative to DMSO-treated cells. The experiment was conducted in biological triplicate, and each individual run was repeated in technical triplicate.

See also Table S1.

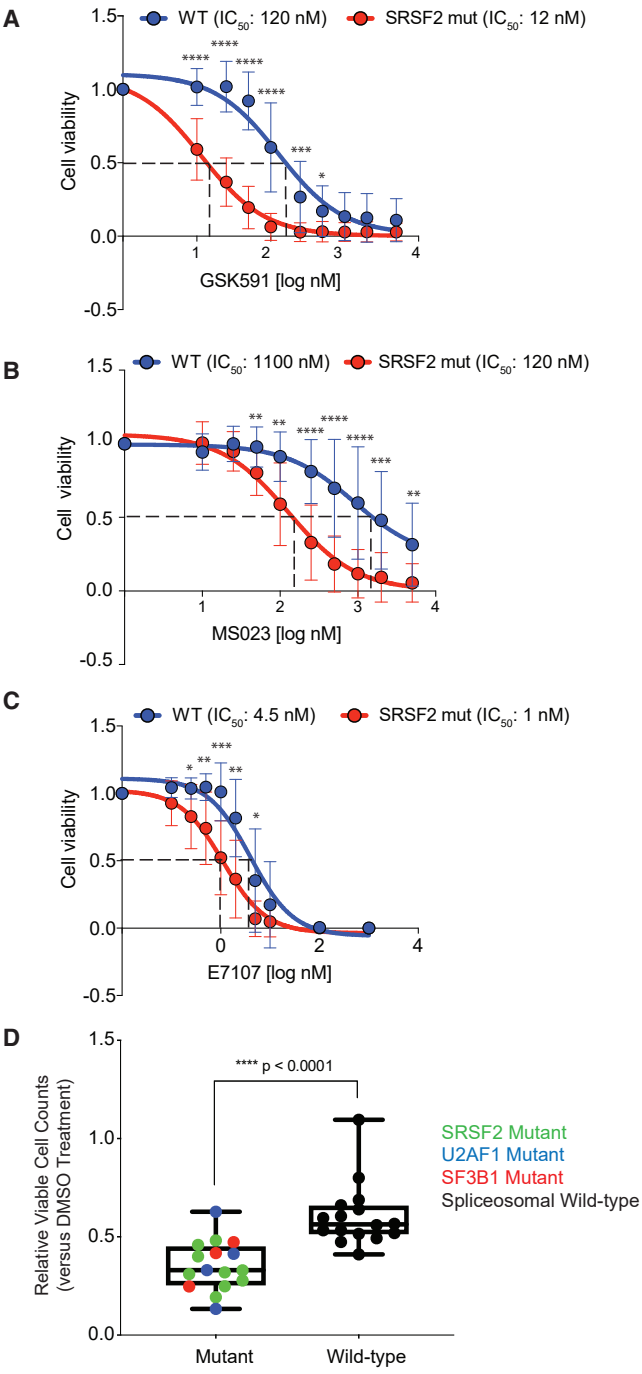


Figure 2. Preferential Effects of PRMT5 or Type I PRMT Inhibition on SF-Mutant AML over WT Counterparts *In Vitro*

(A–C) Relative cell viability of *MLL-AF9/Vav*-cre *Srsf2*^{WT} and *MLL-AF9/Vav*-cre *Srsf2*^{P95H} treated with GSK591 (A), MS023 (B), or E7107 (C), and normalized to control (SGC2096a for PRMT5, MS094 for PRMT1, and DMSO for E7107). Samples were prepared in four to six replicates and averages were calculated, error bars represent SD. Student's t test was used for statistical analysis.

(D) Relative viable cell counts of AML patient samples to GSK591 based on spliceosomal gene mutation status. Primary AML cells with SF-mutations (n = 16) or WT for *SRSF2*, *U2AF1*, and *SF3B1* (n = 16) were incubated with DMSO or GSK591 (0.5 mM) for 6 days. Cells were subjected to flow cytometry to detect 7-AAD-negative, YO-PRO1-negative, viable cells. Relative viable cell numbers were compared with Welch's t test. Boxplot top line of whisker denotes the highest value in dataset and bottom line of whisker denotes the lowest value in dataset, box spans interquartile range and line in box indicates median.

*p = 0.01–0.05, **p = 0.001–0.01, ***p = 0.0001–0.001, ****p < 0.0001. See also Figure S1 and Table S2.

Fifth, we extended our observation to isogenic K562 human leukemia cell lines with knockin mutations of the endogenous *SRSF2* (*SRSF2*^{P95H}) or *SF3B1* (*SF3B1*^{K700E}) (Figures 5F and 5G). In each case, we observed a strong synergistic effect between GSK591 and MS023, which was more pronounced in SF-mutant isogenic cells. The changes in symmetric and asymmetric dimethylation of arginine in K562 cells, from single and combinatorial use of MS023 and GSK591, were validated by western blot (Figure S4).

Finally, given the reported effects of E7107 (Lee et al., 2016), and the synergy observed upon combining MS023 and GSK591 (Figures 4B and 4C), we tested the effect of a triple-drug combination. As expected, triple combination treatment with MS023, GSK591, and E7107 had a strong effect on cell viability and preferentially killed the SRSF2-mutant line over the WT control (Figure 5H). Altogether these experiments confirm the strong synergy between the drugs impacting different aspects of splicing catalysis (type I PRMTs, PRMT5 and SF3B1 inhibitors), which was accentuated in the presence of an SF mutation.

Global Profiling of PRMT Substrates at Single-Site Resolution by Quantitative Liquid Chromatography-Tandem Mass Spectrometry

We next set out to explore the mechanistic basis for the link between inhibition of arginine methylation and preferential effects on SF-mutant leukemias. Although previous studies have suggested a link between PRMT5 and splicing regulation through arginine methylation of spliceosome-associated proteins (Braun et al., 2017; Koh et al., 2015), PRMT5 has numerous substrates (Guo et al., 2010; Hamard et al., 2018; Jansson et al., 2008; Musiani et al., 2019; Sims et al., 2011; Xu et al., 2001). Similarly, type I PRMTs have a large number of cellular substrates. As a consequence, it is currently unclear which PRMT substrate(s) are most relevant to the observed cellular effects of PRMT inhibition. To address this question, we undertook a proteomic approach to identify PRMT5 and PRMT type I substrates in leukemia. In particular, we employed a stable isotope labeling with amino acids in cell culture (SILAC)-mass spectrometry strategy (Musiani et al., 2019) to characterize the arginine methyl proteome in acute promyelocytic leukemia NB4 cells (Figure 6A; Table S4). In brief, in a forward SILAC experiment, cells cultured in the medium complemented with heavy-labeled arginine ("H") and lysine were treated with GSK591 or MS023, while cells cultured with light amino acids ("L") were treated with vehicle. Upon H and L cell harvesting and mixing in equal ratio, protein extraction and in-solution tryptic digestion, peptides bearing mono-methylated arginine (MMA), SDMA, or ADMA, respectively, were enriched using pan-methyl-specific antibodies, before high-resolution nano-liquid chromatography-tandem mass spectrometry analysis for modified peptide identification and quantification (Figure 6A). The SILAC experiments were repeated in biological replicates in SILAC reverse mode,

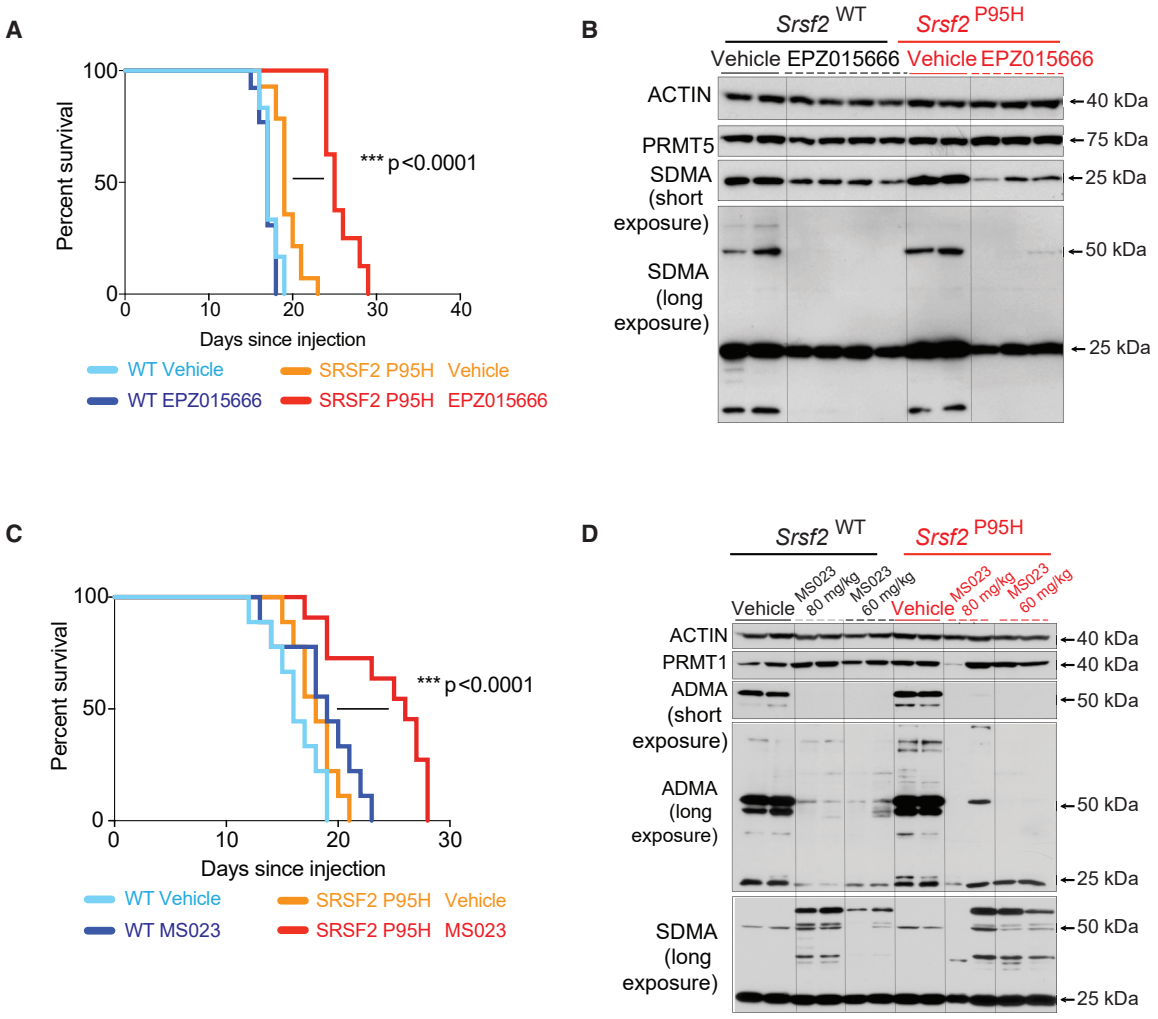


Figure 3. Preferential Effects of PRMT5 or Type I PRMT Inhibition on SF-Mutant AML over WT Counterparts *In Vivo*

(A and B) Kaplan-Meier survival curve of mice treated with vehicle or EPZ015666. Survival comparison by Mantel-Cox log-ranked test (WT vehicle n = 12; WT drug n = 13, *Srsf2*^{P95H} vehicle n = 14, *Srsf2*^{P95H} drug n = 8) (A). Western blot of PRMT5, symmetric dimethyl arginine (SDMA; both a short exposure and a long exposure are shown), and actin in spleens of mice from (A) at time of death. Organs collected were 24 h after the last dose. Each column represents tissue from a distinct individual representative animal (B).

(C and D) Kaplan-Meier survival curve of mice treated with vehicle or MS023. Survival comparison by Mantel-Cox log-ranked test (WT vehicle n = 9, WT drug n = 9, *Srsf2*^{P95H} vehicle n = 9, *Srsf2*^{P95H} drug n = 11). (C) Western blot of PRMT1, asymmetric dimethylarginine (ADMA) (both a low exposure and a high exposure are shown), SDMA, and actin in spleens of mice from (C) at time of death. Organs collected were 24 h after the last dose. Each column represents tissue from a distinct individual representative animal (D). *p = 0.01–0.05, **p = 0.001–0.01, ***p = 0.0001–0.001, ****p < 0.0001.

See also Figure S2.

whereby the amino acid labels were swapped among the two functional states (see the STAR Methods). With this strategy, we were able to quantify a total of 391 and 735 R-methyl-peptides in the GSK591 and MS023 experiments, respectively. Specifically, in the GSK591 treatment experiment, we identified 299 peptides bearing MMA, 40 bearing DMA and 52 bearing both modifications; in the MS023 experiment, we found 433 MMA peptides, 200 DMA peptides, and 102 peptides bearing both modifications (Figure S5A). In total, these peptides carried 1,188 methylation events on arginine, and were distributed on 219 different proteins (Table S4).

Analysis of the methyl-peptide SILAC ratios, normalized by the respective protein ratios, revealed that both GSK591 and MS023 caused a prominent downregulation of methyl sites, with 49 (15%) and 135 (16%) methyl-peptides significantly decreased and only 4 (1%) and 97 (11%) upregulated, respectively (Figure S5B; Table S4). Log analysis of the significantly regulated methyl-peptides showed specific enrichment for glycine at position +1 and −1 around the modified arginine in the MS023- and GSK591-treated cells, respectively (Figures 6B and 6C). Interestingly, PRMT5 and type I PRMTs appear to regulate the methylation of distinct and non-overlapping proteins (Figure 6D). We did not detect histones as differentially arginine methylated, rather, proteins whose methylation level changed upon PRMT inhibition were mainly RNA-binding proteins (RBPs) with established roles in RNA export, regulation of RNA stability, and

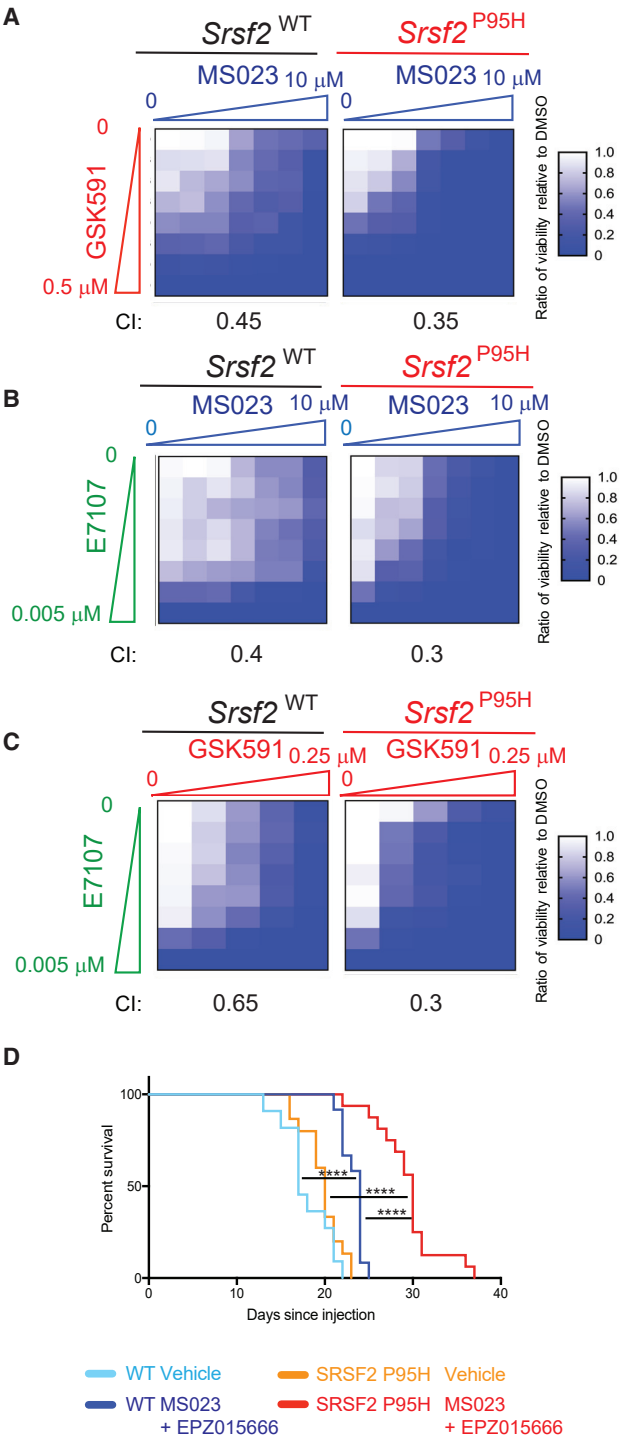


Figure 4. Synergistic Effects of Combined PRMT5, Type I PRMT, and/or SF3B Inhibition on Primary Mouse AMLs

(A–C) Heatmap of half maximal inhibitory concentration values of WT or *Srsf2*^{P95H} *MLL-AF9* cells grown in increasing doses of MS023 plus GSK591 (A), MS023 plus E7107 (B), or GSK591 plus E7107 (C) for 7 days. The combination index (CI) in each cell type for each drug pair is shown.

(D) Kaplan-Meier curve of mice transplanted with MLL-AF9 cells followed by combined treatment with MS023 plus EPZ015666 *in vivo*. Survival comparison by Mantel-Cox log-ranked test (WT vehicle n = 11, WT drug n = 12, *Srsf2*^{P95H} vehicle n = 15, *Srsf2*^{P95H} drug n = 16). *p = 0.01–0.05, **p = 0.001–0.01, ***p = 0.0001–0.001, ****p < 0.0001.

See also Figure S2.

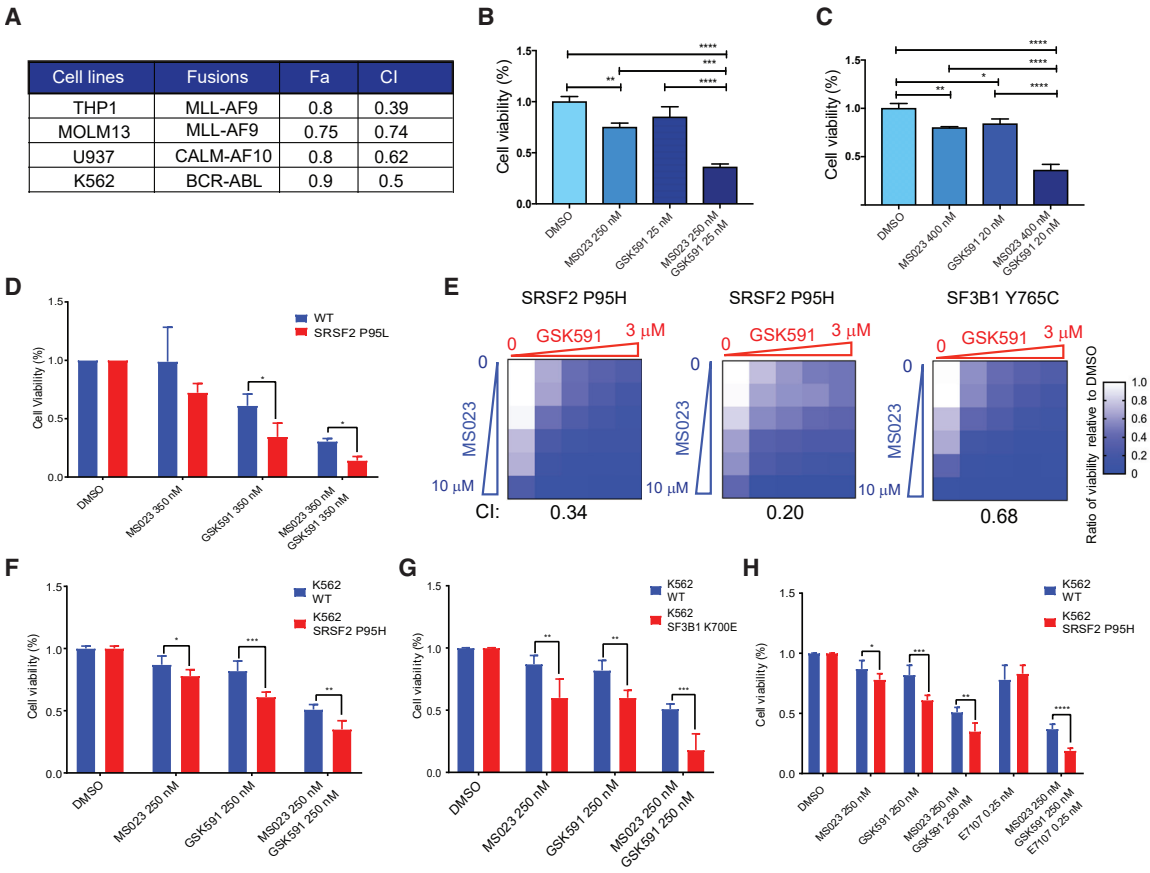
RNA splicing (Figures 6E–6H). In addition, MS023 elicited a specific effect on the methylation of proteins involved in translation (Figures 6E–6G). Most of the methyl-peptides responding to GSK591 (70%) and MS023 (89%) were orthogonally validated through the intersection with a high-quality methyl-proteome dataset annotated through the heavy methyl SILAC labeling strategy (Massignani et al., 2019; Ong et al., 2004) (Figure S5C; Table S4).

Importantly, to exclude a potential bias toward more abundant proteins in our methyl-proteome, we matched the methylated proteins identified with the immuno-enrichment approach to their respective protein abundances, calculated through the Intensity-Based Absolute Quantification (iBAQ) algorithm (Schwanhausser et al., 2011) in the whole proteome dataset, used as input for the methyl-peptide immunoprecipitation (Table S3). This experimental proteome included more than 7,300 proteins, encompassing a dynamic range of abundance of $\sim 10^7$ (Figure S5D, black bars). Within this protein abundance distribution, the 219 experimentally identified methyl-proteins covered a dynamic range of abundance of $>10^5$, spanning from high (e.g., EEF1A1 and HNRNPs) to low (e.g., KMT2C and SVIL) expressed proteins (Figure S5D, red bars). A similar dynamic range was also observed for proteins directly regulated by PRMT5 or type I PRMT inhibitors (Figure S5E).

This result suggests that, despite being a minor proportion of the detectable cellular proteome, the PRMT-dependent methyl-proteome is not biased toward the high abundant subset; hence, the overrepresentation of RBPs is not a mere reflection of the abundance of these proteins, but linked to the cellular function of PRMTs.

Cell-Cycle Deregulation upon PRMT5 and Type I PRMT Inhibition

To further understand the mechanistic underpinning for the synergistic cytotoxic effects of PRMT inhibition on leukemia cells, regardless of the spliceosomal gene mutational status, we performed RNA sequencing analysis of isogenic *SRSF2*^{P95H} K562 leukemia cell lines and its WT counterpart, treated with DMSO, GSK591, and MS023, or a combination of the two. We first analyzed the global changes in gene expression caused by drug treatment. Gene ontology (GO) analysis revealed that there is an upregulation of genes involved in mitosis and cell-cycle regulation for both K562 WT and K562 *SRSF2*^{P95H} cells (Figures 7A and 7B; Table S5). This suggests that drug combination treatment causes cell-cycle deregulation and indeed we could validate this by cell-cycle analysis on cells collected at day 8 of drug treatment. Specifically, inhibition of PRMT5 and type I PRMTs leads to a decrease in proportion of cells in G1 and S phase (Figure 7C). Concomitantly, we observed an increase in an apoptotic sub-G1 population (Figures 7C and 7D). Overall, upon PRMT inhibition, *SRSF2*^{P95H}-mutant cells exited the cell cycle and induced caspase-3/-7 activity more than their WT counterpart.



Cell lines	Fusions	Fa	CI
THP1	MLL-AF9	0.8	0.39
MOLM13	MLL-AF9	0.75	0.74
U937	CALM-AF10	0.8	0.62
K562	BCR-ABL	0.9	0.5

Figure 5. Synergistic Effects of Combined PRMT5, Type I PRMT, and/or SF3B Inhibition on Human AML Lines and Patient Samples

(A) Summary of CI and fraction affected (Fa) values of human AML cell lines treated with MS023 and GSK591.

(B and C) Percentage of viable THP1 (B) and U937 (C) cells after 8 days exposure to MS023, GSK591, or the combination relative to DMSO *in vitro*.

(D) Hematopoietic progenitor cells differentiated from the iPSC lines 5–16 Cre20 (*SRSF2* P95L) and N-2.12 (isogenic normal) were treated as indicated. Cell viability represents cell counts relative to DMSO-treated cells.

(E) Heatmap of cell viability after 6 days of treatment. CI at IC_{50} values of MS023 plus GSK591 is indicated at the bottom.

(F and G) Percentage of viable K562 WT or *SRSF2*^{P95H} knockin mutant cells (F) and K562 WT or *SF3B1*^{K700E} knockin mutant cells (G) after 8 days treatment with MS023, GSK591, or the combination relative to DMSO *in vitro*.

(H) Percentage of viable K562 WT or *SRSF2*^{P95H} knockin mutant cells after 8 days treatment with MS023, GSK591, E7107 or in combination relative to DMSO *in vitro*.

All data are representative of three independent experiments, error bars represent SD. *p = 0.01–0.05, **p = 0.001–0.01, ***p = 0.0001–0.001, ****p < 0.0001; one-way ANOVA was used to analyze THP1 and U937, while unpaired t test was used for K562 WT and *SRSF2*^{P95H}. See also Figures S3 and S4.

PRMT5 and Type I PRMTs Inhibition Leads to Synergistic Changes in Alternative Splicing

We next conducted splicing analysis of the isogenic K562 leukemia cell lines treated with DMSO, GSK591, and MS023, or a combination of the two, to further decipher the possible cause for the overall synergistic killing of PRMT inhibitors, and preferential killing of *SRSF2*^{P95H} over WT cells. As expected, *SRSF2*^{P95H}-mutant and WT cells had a distinct transcriptome, typified by unique changes to RNA splicing, as described previously (Kim et al., 2015; Zhang et al., 2015a). First, the analysis revealed an increase in the number of aberrant splicing events upon treatment with both MS023 and GSK591, compared with control DMSO-treated cells. This was true both in WT and *SRSF2*^{P95H} cells. Interestingly, the number of altered splicing events, particularly in cassette exons, increased upon combined exposure to GSK591 and MS023, both in K562 WT and *SRSF2*^{P95H} cells (Figures 8A and 8B). In addition, drug combination induced a unique pattern of splicing alteration that is not simply the sum of the cassette exon splicing events caused by the individual drugs (Figure S6A).

Second, we compared the overlap of the deregulated events among the two genotypes. Given that *SRSF2*^{P95H} and WT cells have a different transcriptome to begin with, it was not surprising to note a divergence in the inclusion/exclusion of the affected cassette exons. More specifically, about one-third of the cassette exon events were commonly deregulated upon GSK591 + MS023 treatment, while the rest were uniquely spliced in either WT or *SRSF2*^{P95H} cells (Figure S6B).

Third, we confirmed that *SRSF2*^{P95H}-regulated exon inclusion/exclusion events have enrichment for C-rich exonic splicing enhancer sequences over G-rich sequences. Interestingly, treatment with PRMT inhibitors, led to a reduction in events

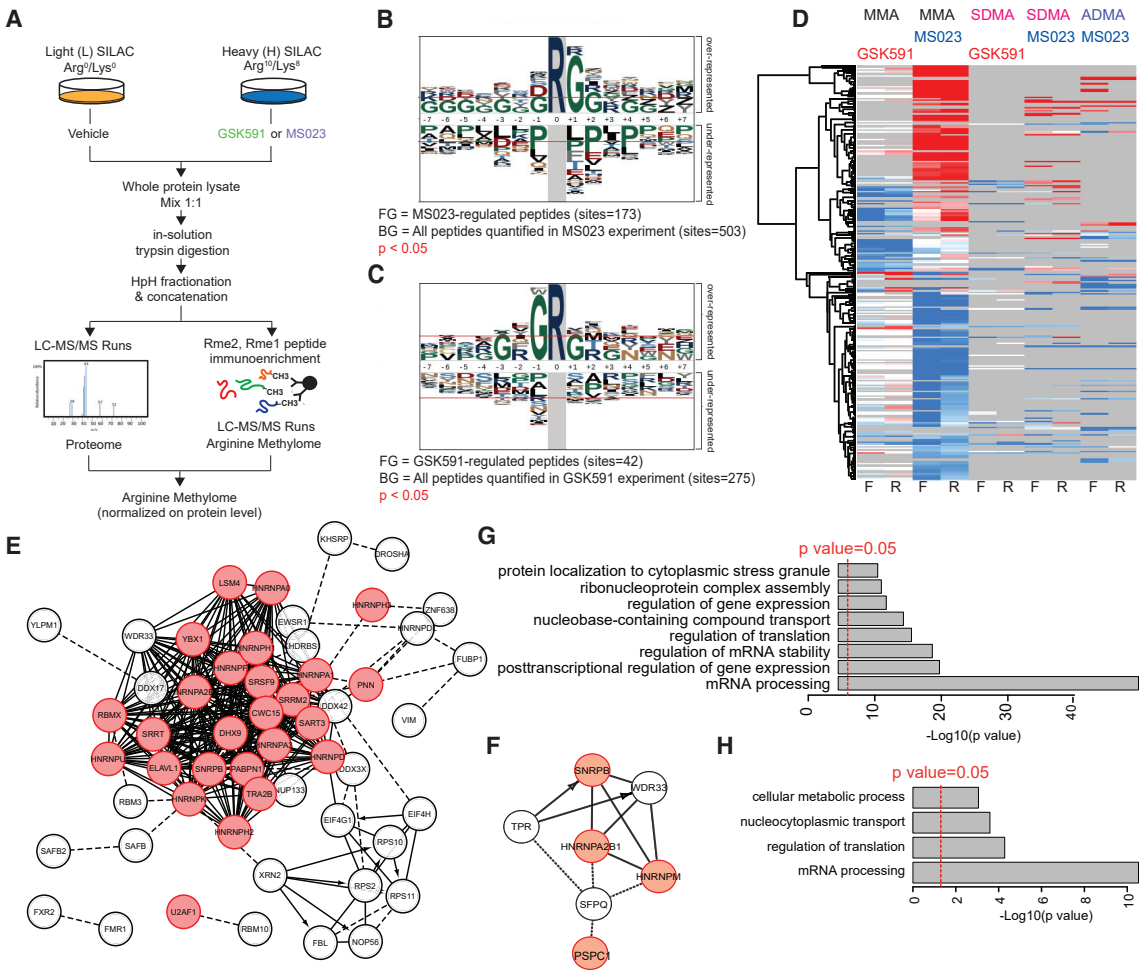


Figure 6. Global Profiling of PRMT Substrates at Single-Site Resolution by Quantitative Liquid Chromatography-Tandem Mass Spectrometry

(A) Workflow of the SILAC methyl-R-proteomic experiments carried out to identify mono- and dimethylarginine substrates regulated by GSK591 and MS023 in NB4 leukemia cells.

(B and C) Sequence motif analysis indicates the consensus sequences significantly enriched in the methyl-peptides regulated by MS023 (B) or GSK591 (C).

(D) Heatmap of $\log_2$ SILAC ratios of each methyl-peptide identified and quantified from the SILAC experiment diagrammed in (A). For each pharmacological treatment two biological replicates, in forward and reverse SILAC mode, were carried out and different degrees of methylation were enriched and profiled. Only peptides reproducibly regulated in at least one pair of forward-reverse experiment are visualized in the heatmap.

(E and F) Network analysis of the proteins displaying methylation changes upon treatment with MS023 (E) or GSK591 (F). RNA-binding proteins are highlighted in red.

(G and H) Functional analysis of changing methylated proteins highlight the biological process terms significantly enriched among proteins displaying methylation changes upon treatment with MS023 (G) and GSK591 (H).

See also Figure S5 and Table S3.

containing a CCNG motif (preferentially regulated by SRSF2^{P95H}) and to an overall depletion of events regulated by WT SRSF2 (containing a GGNG and CGNG motif) (Figure 8C).

We then looked at the GO enrichment of cassette exon events changing upon combination drug treatment in K562 *SRSF2*^{P95H} cells and focused on the top perturbed GO categories, which included microtubule organization, DNA repair, and cell-cycle regulation (Table S5). We generated heatmaps for events in each GO category to depict the change in percentage spliced in (delta PSI) values upon single- and double-drug treatment in both cell lines, relative to K562 WT (Figures 8D and S6C). The rationale for this approach was to identify those splicing changes that could explain the preferential sensitivity of mutant cells to PRMT inhibition. Within these GO categories, we identified cassette exon splicing events that were changing more in K562 *SRSF2*^{P95H} than in their WT counterparts and validated some of these events (Figure S6D). As DNA repair genes are broadly deregulated by PRMT inhibition, we also performed western blot of phospho-H2Ax to check for evidence of DNA damage. Indeed, there is a significant upregulation of γH2Ax, especially in the K562 *SRSF2*^{P95H} cell, upon combination drug treatment (Figure S6E). Overall, although the number of deregulated splicing events is similar between WT and *SRSF2*^{P95H}-mutant cells, they affect different genes and pathways, ultimately resulting in increased DNA damage and cell-cycle arrest.

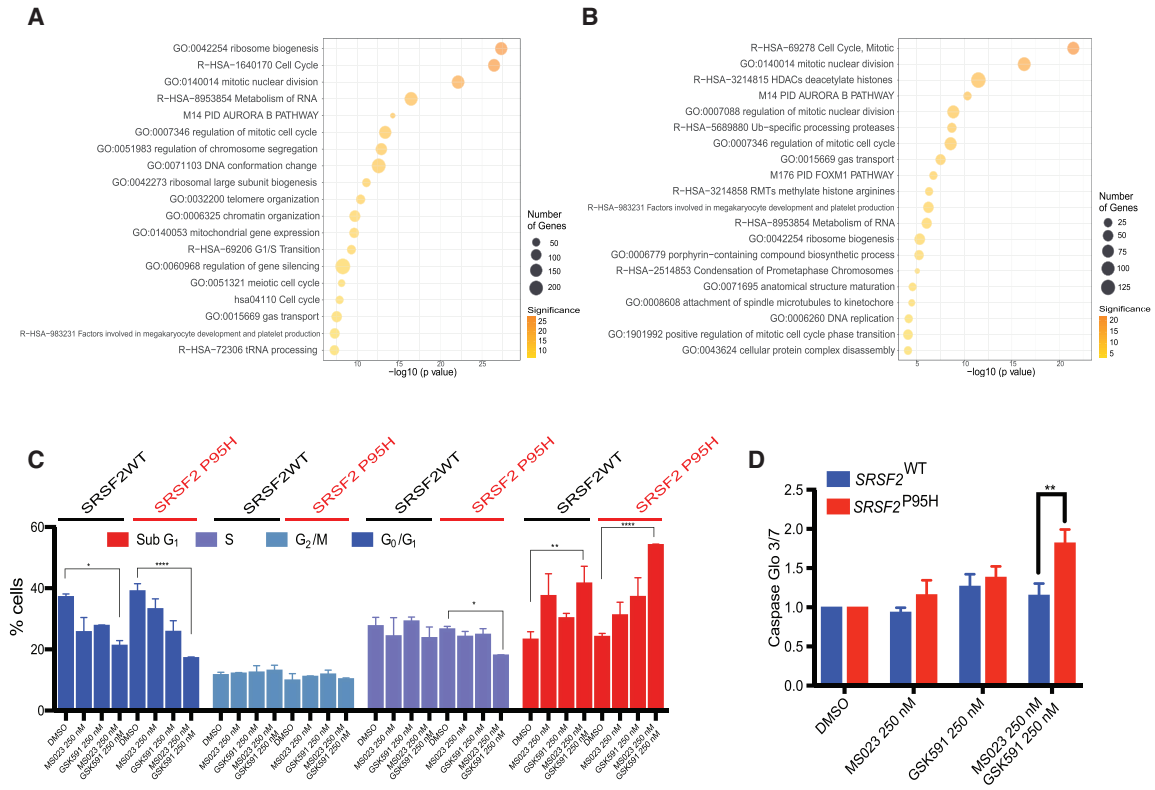


Figure 7. Gene Expression Changes and Cell-Cycle Deregulation with PRMT5 and Type I PRMTs Inhibition

(A and B) GO categories of genes upregulated in K562 WT cells (A) and K562 *SRSF2*^{P95H} cells (B) upon MS023 and GSK591 combination treatment.

(C) Cell-cycle analysis of K562 WT and *SRSF2*^{P95H} cells upon MS023 and GSK591 treatment.

(D) Caspase-Glo 3/7 changes following 8 days treatment with MS023 and GSK591 in K562 WT and *SRSF2*^{P95H} cells.

All data are representative of three independent experiments, error bars represent SD. *p = 0.01–0.05, **p = 0.001–0.01, ***p = 0.0001–0.001, ****p < 0.0001. Student's t test used for statistical analysis.

Among the splicing events in the cell-cycle regulation GO category, we identified an aberrant splicing event within the mRNA encoding *EZH2* (Kim et al., 2015; Lee et al., 2016). This specific event, which is driven by *SRSF2* and not by *SF3B1* mutations (Pellagatti et al., 2018) (Figures S6F and S6G), leads to the inclusion of a poison exon, to the formation of a premature stop codon, and subsequent nonsense-mediated decay (Kim et al., 2015; Lee et al., 2016). Interestingly, similar to that previously observed for E7107 (Lee et al., 2016), we observed decreased poison exon inclusion upon PRMT inhibition (Figures 8E, S6G, and S7A), leading to increased EZH2 protein abundance (Figure 8F). To functionally test the meaning of this event and assess whether toxicity induced by MS023 and GSK591 was at least in part mediated by EZH2 re-expression, we took two approaches. We first designed an antisense oligonucleotide (AON) to specifically mimic the exon-skipping event induced by PRMT inhibitors. The designed AON was able to reduce the inclusion level of the *EZH2* poison exon (Figure S7B), increase EZH2 protein levels (Figure S7C), and reduce cell viability in K562, which was more evident in *SRSF2*^{P95H}-mutant cells than in WT controls (Figure S7D).

To complement the above AON approach, we also used CRISPR/Cas9 to delete the endogenous *EZH2* in K562 cells. We then treated EZH2 null cells or the parental control line with MS023, GSK591, or a combination of the two. EZH2 null cells were significantly more resistant to PRMT inhibition, suggesting that the effect of these small molecules is at least in part mediated by restoration of EZH2 levels in AML cells (Figure S7E).

DISCUSSION

Inhibition of PRMT5 or type I PRMTs has been linked to several downstream effects and proven to impact multiple pathways. Nonetheless, with the exception of the well-described *MTAP* deletion, which renders cells more susceptible to further depletion of PRMT5 (Kryukov et al., 2016; Marjon et al., 2016; Mavrakis et al., 2016) or inhibition of type I PRMTs (Fedoriw et al., 2019; Gao et al., 2019), there has been no clear indication of additional potential vulnerabilities of cancer cells to PRMT deletion/inhibition.

Overall, our data identify a strong impact of inhibiting protein arginine methylation on RNA splicing. PRMT5 or type I PRMTs inhibition resulted in preferential killing of SF-mutant AML, and overall synergy across genetic backgrounds. Mechanistically, this can be explained through inhibition of arginine post-translational modification (SDMA and ADMA) of a large compendium of RBPs involved in splicing regulation.

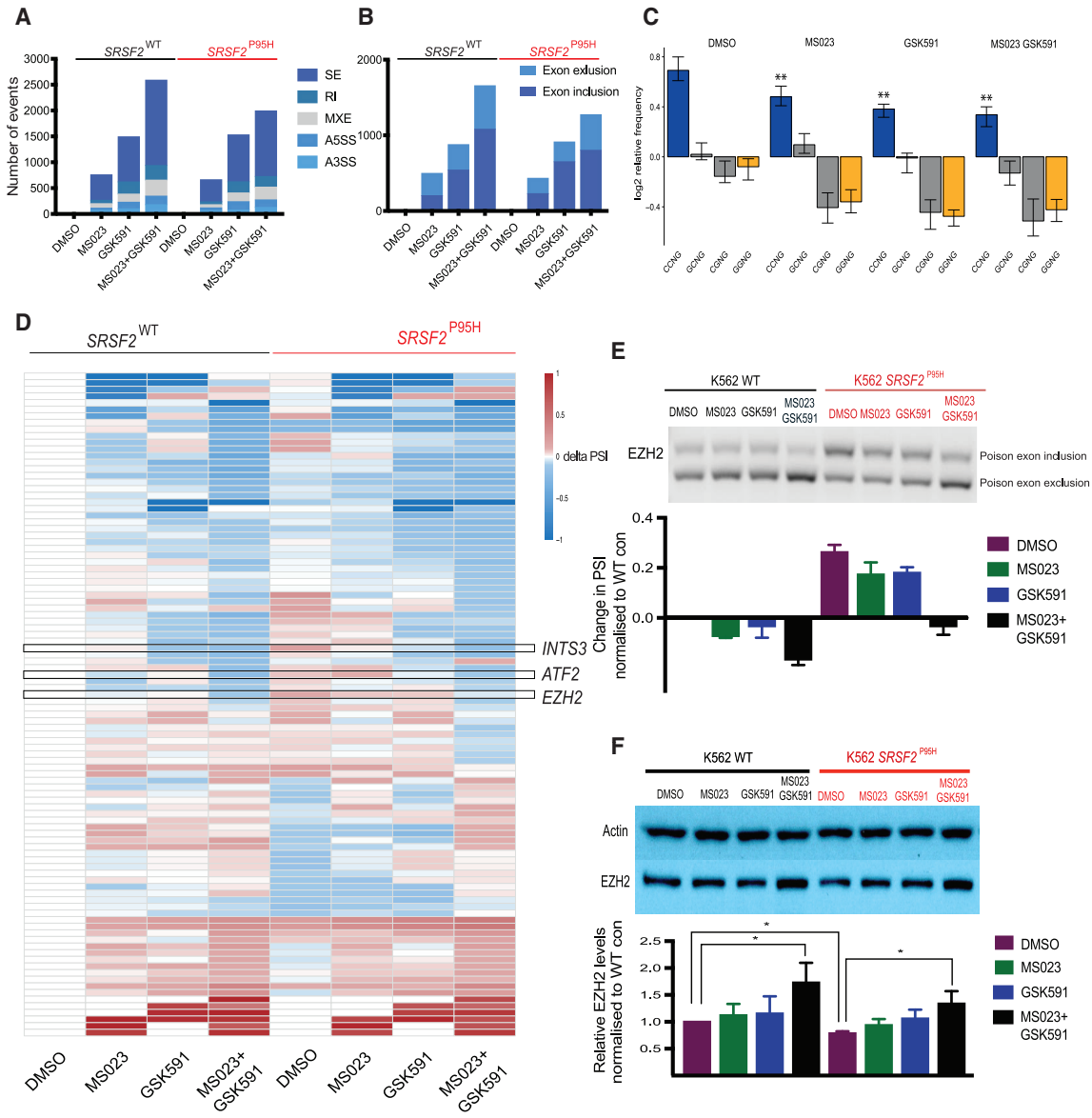


Figure 8. PRMT5 and Type I PRMT Inhibition Leads to Synergistic Changes in Alternative Splicing

(A) Bar graphs enumerating numbers of significantly differentially spliced events in WT or *SRSF2*^{P95H} K562 cells treated with each of the compounds listed versus DMSO control. A3SS, alternative 3′ splice site; A5SS, alternative 5′ splice site; MXE, mutually exclusive exon; RI, retained intron; SE, cassette exon.

(B) Bar graphs enumerating numbers of significantly differentially spliced exons in WT or *Srsf2*^{P95H} K562 cells treated with each of the compounds listed versus DMSO control.

(C) Changes in levels of splicing events with CCNG, GCNG, CGNG, or GGNG motifs upon MS023 and GSK591 treatment in K562 cells.

(D) Heatmap showing change in PSI of exon splicing events normalized to PSI of WT control within the "Cell Cycle" gene ontology category.

(E) RT-PCR data showing changes in PSI levels of *EZH2* poison exon inclusion in K562 cells with or without knockin of *SRSF2*^{P95H} upon MS023 and GSK591 treatment *in vitro*.

(F) Western blot showing changes in EZH2 protein levels upon MS023 and GSK591 treatment *in vitro*.

All data are representative of three independent experiments, error bars represent SD. *p = 0.01–0.05, **p = 0.001–0.01, ***p = 0.0001–0.001, ****p < 0.0001. Student's t test used for statistical analysis. See also Figures S6 and S7 and Table S5.

Currently, there is no standardized way to predict the effects of arginine methylation on protein function. Even less is known on the interplay between ADMA and SDMA occurring at distinct or overlapping sites. However, despite the current limitations on systematic predictions, there are several examples in the literature describing the effect of individual arginine methylation events on the function of PRMT5 and/or type I PRMTs substrates.

As described here, type I PRMTs and PRMT5 methylate different arginine residues on a common set of RBPs involved in RNA post-transcriptional regulation. Both ADMA and SDMA maintain the positive charge of the arginine side chain, but

reduce its hydrogen bonding capabilities. Methylation concurrently increases the hydrophobicity to the arginine side chain, favoring its interaction with aromatic cages. As a consequence, arginine methylation has been shown to alter protein-protein interactions (Erce et al., 2013) and protein-RNA interactions (Denman, 2002; Dolzhanskaya et al., 2006). A well-documented example is the arginine methylation of three of the seven Sm proteins (SNRPB/B′, SNRPD1, and SNRPD3), leading to their binding to the Tudor domain on SMN1 (Bezzi et al., 2013; Friesen et al., 2001; Meister et al., 2001; Meister and Fischer, 2002). This is a critical event in the assembly of spliceosomal snRNPs, and, consistently with previous reports, we detected significant changes in SNRPB methylation in our dataset at arginines 112 and 147, upon PRMT inhibition.

In addition, we identified several arginine methylated residues on SFPQ and PSPC1. Interesting, both proteins are core components of the membrane-less nuclear structures called paraspeckles, which play a central role in cancer and stress response (Adriaens et al., 2016). In particular, SFPQR693me1 and PSPC1R507me1 are reduced in the presence of GSK591, while PSPC1R4me1/me2a are reduced in the presence of MS023. Whereas, in the case of SFPQ, arginine methylation is associated with increased RNA binding (Snijders et al., 2015), there is no report describing the role of PSPC1 methylation, a point that warrants future investigation.

Recently, ADMA methylation by PRMT1 has been shown to weaken the cation π interactions between RGG/RG-rich motifs and low complexity regions of several disease-associated prion-like intrinsically disordered proteins (Hofweber et al., 2018; Qamar et al., 2018; Tsai et al., 2016). We detected changes in methylation of several members of this family of proteins (e.g., FUS, G3BP1, hnRNPA2B1, and hnRNPA1). We hence predict that PRMT5 and PRMT type I inhibitors may have an impact on reversible liquid-liquid phase separation and on the assembly/disassembly of membrane-less organelles, impacting RNA post-transcriptional regulation.

Finally, there are numerous examples of how ADMA/SDMA may regulate localization of proteins including cellular localization of RBPs such as Sam68, Aven, FUS, hnRNPA1, and hnRNPA2 (Cote et al., 2003; Nichols et al., 2000; Thandapani et al., 2015; Tradewell et al., 2012; Wall and Lewis, 2017).

It is well documented that SF genes mutated in cancers are mutually exclusive (Yoshida et al., 2011) and that they have a synthetic lethal relationship when co-mutated in the same cell (Lee et al., 2018). Here we observed a general perturbation of gene expression and, importantly, of alternative splicing (increased number of skipped/included exons and retained introns) upon PRMT inhibition. Given that SF-mutant cells are preferentially sensitive to genetic or pharmacologic perturbations in splicing, compared with spliceosomal WT counterparts (Lee et al., 2016; Obeng et al., 2016; Seiler et al., 2018; Shirai et al., 2017; Wang et al., 2019), an increase in overall splicing perturbation is predicted to lead to their preferential killing. This has been previously observed upon treatment with SF3B1 inhibitors (Lee et al., 2016; Obeng et al., 2016; Seiler et al., 2018; Shirai et al., 2017; Wang et al., 2019) and RBM39 degrading compounds (Wang et al., 2019). Our results reveal how inhibition of both type I and type II PRMTs can lead to a similar preferential killing of SF-mutant AML and myelodysplastic syndrome cells.

At the same time, it is possible that PRMT inhibition causes the aberrant splicing of oncogenic isoforms required for the survival of SF-mutant cells, leading to their preferential killing over WT counterparts. To test this latter point, we focused our analysis on the aberrant splicing of the *EZH2* mRNA. In myeloid malignancies, EZH2 acts as a tumor suppressor, and it is often deleted and/or affected by loss-of-function mutations (Ernst et al., 2010; Nikoloski et al., 2010). Although we acknowledge that multiple genes could mediate the downstream effects of PRMT inhibition, we did observe a partial rescue upon CRISPR-mediated knockout of *EZH2*.

In human *SRSF2*^{P95H}-mutant cells, we detected the previously reported aberrant inclusion of a poison exon in the *EZH2* transcript, which harbors a premature stop codon and targets it for nonsense-mediated decay (Kim et al., 2015; Lee et al., 2016). Combinatorial use of the PRMT inhibitors caused a reduction in poison exon inclusion and increased EZH2 protein levels. The importance of this splicing event as part of the mechanistic basis for increased sensitivity of *SRSF2*^{P95H} cells to PRMT inhibition was further validated by the increased sensitivity of K562 *SRSF2*^{P95H} cells to the use of AON that induced skipping of the poison exon, and with CRISPR/Cas9 deletion of EZH2, which reduced sensitivity of K562 cells to PRMT inhibition.

To conclude, the data provided here have important therapeutic implications for patients with spliceosomal gene mutations, given ongoing clinical trials with GSK3326595 (PRMT5 inhibitor from the same chemical series as EPZ015666 and GSK3203591; NCT03614728) and with a type I inhibitor GSK3368715 (Fedoriw et al., 2019; NCT03666988). Specifically, our data provide a strong basis for patient selection in the use of PRMT5 and type I PRMTs inhibitors in ongoing and future clinical trials of these agents based on the presence of genetic alterations impacting RNA splicing. At the same time, our proteomic data identify that PRMT5 and type I PRMTs largely regulate a distinct set of substrates, and, coincident with this, the simultaneous inhibition of both PRMT axes resulted in strong synergistic effects. A notable example is the increased expression of EZH2 protein in *SRSF2*^{P95H} K562 cells. *EZH2* is recurrently deleted or affected by loss-of-function mutations in myeloid malignancies (Ernst et al., 2010), while it is partially inactivated by aberrant splicing in *SRSF2*-mutant cells. These data extend the therapeutic utility of PRMT inhibitors, identify a combinatorial drug strategy utilizing simultaneous PRMT5 and type I PRMT inhibition, and implicate perturbation of RNA splicing as an important cellular mediator of cell death due to inhibition of PRMTs.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- LEAD CONTACT AND MATERIALS AVAILABILITY
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Cell Lines and Cell Culture
  - Primary Human Samples
  - Animals

- METHOD DETAILS
  - Identification of Druggable Targets in the Extended Splicing Network
  - Animal Studies
  - *In Vitro* Drug Screen with Murine *MLL-AF9* Cells
  - *In Vitro* Drug Titration with Murine *MLL-AF9* Cells
  - *In Vitro* Drug Treatment with Primary Human AML Patient Samples
  - *In Vitro* Combination Drug Titration with Primary Human AML Patient Samples
  - *In Vitro* Cell Viability Assays with Human Cell Lines
  - Treatment of iPSC-derived Hematopoietic Progenitor Cells with PRMT Inhibitors
  - Pharmacokinetic Studies of MS023 *In Vivo*
  - Administration of PRMT Inhibitors *In Vivo*
  - Western Blotting
  - RNA-seq Sample Preparation
  - Arginine Methyl-Peptides Separation and Enrichment Prior to LC-MS/MS Analysis
  - Nano-LC-MS/MS Analysis
  - Validation of RNAseq
  - AON Electroporation
  - CRISPR/Cas9 Knockout of EZH2 in K562 Cells
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Computation of the Combination Index
  - RNA-Sequencing and Bioinformatics Analysis
  - Replicates
  - Data Analysis of SILAC Arginine Methyl-Peptides
  - HmSEEKER: A Perl-Based Pipeline for High-Confidence Assignment of Methyl-Peptides from hmSILAC Data
  - Use of hmLINKER to Intersect the SILAC Methyl-proteome with the hmSILAC Dataset
  - Motif Analysis of R Methyl-Peptides
  - Bioinformatic Analysis of R Methyl-Peptides
  - Calculation of the Protein Abundance Index in the Annotated Proteome
  - Statistical Analyses
- DATA AND CODE AVAILABILITY

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.ccell.2019.07.003.

ACKNOWLEDGMENTS

We thank H3 Biomedicine for providing E7107. We thank the BRC Shared facilities for technical support. We thank F. Raimondi (IEO) for technical support with the proteome to transcriptome data integration analysis. We are grateful to the GIS Genome Sequencing Team for help with the Solexa high-throughput sequencing; to the entire EG lab for critical discussion. This work was supported by AGS fellowship to J.Y.F., AGA-SINGA (Singapore Graduate Award) fellowships to P.A.G. and L.P. E.G. is supported by NMRC/OFIRG/0032/2017 and NRF-CRP17-2017-06 grants. This research was additionally supported by the RNA Biology Center at the Cancer Science Institute of Singapore, NUS, as part of funding under the Singapore Ministry of Education's AcRF Tier 3 grants, grant no. MOE2014-T3-1-006. S.C.-W.L. and O.A.W. are supported by the Leukemia and Lymphoma Society. O.A.W. is supported by grants from NIH/NHLBI (R01 HL128239), NIH/NCI (1 R01 CA201247-01A1), the US Department of Defense Bone Marrow Failure Research Program (W81XWH-12-1-0041), Cycle For Survival, the Edward P. Evans Foundation, the Henry & Marilyn Taub Foundation, the Pershing Square Sohn Foundation, and Starr Foundation grants I8-A8-075 and I9-A9-059. C.H.A. and G.M.L. were supported by the Leukemia and Lymphoma Society of Canada. The Structural Genomics Consortium is a registered charity (1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) (ULTRADD grant no. 115766), Janssen, Merck, Novartis Pharma AG, Ontario Ministry of Economic Development and Innovation, Pfizer, Sao Paulo Research Foundation FAPESP, Takeda, and the Wellcome Trust. Research in the T.B. group was supported by grants from the Italian Ministry of Health (GR-2011-02347880), the Epigen Flagship project (CNR), and the Italian Association for Cancer Research (IG-2018-21834). D.M. was supported by the Fondazione Umberto Veronesi (FUV) fellowship and the Fondazione Istituto Europeo di Oncologia (FIEO) fellowship. A.R. acknowledges support through the NIH (award no. T32 CA078207). J.J. acknowledges support by grant R01GM122749 from the US NIH.

AUTHOR CONTRIBUTIONS

J.Y.F., D.H.PL., L.P., O.A.-W., and E.G. designed the study. J.Y.F., L.P., K.C.K., S.C.-W.L., D.M., E.M., C.M.K., C.M.W., M.S., A.R., S.M., H.W., F.G., M.K., G.M.L., A.G.K., and M.M.S. performed the experiments. P.-A.G., A.P., and D.H.P.L. performed RNA sequencing and computational analysis. J.J. and Y.S. provided MS023 and advice with drug dosing experiments. A.M. and M.D.M. provided access to primary patient samples. D.B.-L., E.P., A.M., T.B., O.A.-W., and E.G. provided general supervision. J.Y.F., L.P., P.-A.G., O.A.-W., and E.G. prepared the manuscript with help from all co-authors.

DECLARATION OF INTERESTS

O.A.-W. has received research funding from H3 Biomedicine, has served as a consultant for H3 Biomedicine and Merck, and has served on advisory board for Foundation Medicine, H3 Biomedicine, and Janssen. E.G. has received research funding from Eli-Lilly and Prelude Therapeutics, has served as a consultant for Prelude Therapeutics, SKBP, and has served on the advisory board for LION TCR and Janssen, he is a co-founder of ImmuNOA Pte. Ltd.. C.M.K. is currently an employee at Mechanistic Biology and Profiling, Discovery Sciences, IMED Biotech Unit, AstraZeneca Pharmaceuticals LP. 35 Gatehouse Dr, Waltham, MA 02451, USA. C.T. and O.B. are Employees and stakeholders of GSK. M.M. receives honorariums from Amgen, Pfizer, and Astellas, and funding from Pfizer, Notable Labs. All SGC Affiliated authors receive funding from the nine SGC-funding companies. C.H.A. and G.M.L. were supported by the Leukemia Lymphoma Society of Canada. The Structural Genomics Consortium is a registered charity (1097737) that receives funds from AbbVie, Bayer Pharma AG, Boehringer Ingelheim, Canada Foundation for Innovation, Eshelman Institute for Innovation, Genome Canada, Innovative Medicines Initiative (EU/EFPIA) (ULTRADD grant no. 115766), Janssen, Merck, Novartis Pharma, Ontario Ministry of Economic Development and Innovation, Pfizer, São Paulo Research Foundation FAPESP, Takeda, and the Wellcome Trust.

Received: February 6, 2019
Revised: June 2, 2019
Accepted: July 8, 2019
Published: August 12, 2019

REFERENCES

Adriaens, C., Standaert, L., Barra, J., Latil, M., Verfaillie, A., Kalev, P., Boeckx, B., Wijnhoven, P.W., Radaelli, E., Vermi, W., et al. (2016). p53 induces formation of NEAT1 lncRNA-containing paraspeckles that modulate replication stress response and chemosensitivity. Nat. Med. *22*, 861–868.

Anczuków, O., Rosenberg, A.Z., Akerman, M., Das, S., Zhan, L., Karni, R., Muthuswamy, S.K., and Krainer, A.R. (2012). The splicing factor SRSF1 regulates apoptosis and proliferation to promote mammary epithelial cell transformation. Nat. Struct. Mol. Biol. *19*, 220–228.

Anders, S., Pyl, P.T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. Bioinformatics *31*, 166–169.

Bezzi, M., Teo, S.X., Muller, J., Mok, W.C., Sahu, S.K., Vardy, L.A., Bonday, Z.Q., and Guccione, E. (2013). Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery. Genes Dev. *27*, 1903–1916.

Braun, C.J., Stanciu, M., Boutz, P.L., Patterson, J.C., Calligaris, D., Higuchi, F., Neupane, R., Fenoglio, S., Cahill, D.P., Wakimoto, H., et al. (2017). Coordinated splicing of regulatory detained introns within oncogenic transcripts creates an exploitable vulnerability in malignant glioma. Cancer Cell *32*, 411–426.e11.

Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E., et al. (2012). The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. *2*, 401–404.

Chan-Penebre, E., Kuplast, K.G., Majer, C.R., Boriack-Sjodin, P.A., Wigle, T.J., Johnston, L.D., Rioux, N., Munchhof, M.J., Jin, L., Jacques, S.L., et al. (2015). A selective inhibitor of PRMT5 with in vivo and in vitro potency in MCL models. Nat. Chem. Biol. *11*, 432–437.

Chang, C.J., Kotini, A.G., Olszewska, M., Georgomanoli, M., Teruya-Feldstein, J., Sperber, H., Sanchez, R., DeVita, R., Martins, T.J., Abdel-Wahab, O., et al. (2018). Dissecting the contributions of cooperating gene mutations to cancer phenotypes and drug responses with patient-derived iPSCs. Stem Cell Reports *10*, 1610–1624.

Cheng, D.T., Mitchell, T.N., Zehir, A., Shah, R.H., Benayed, R., Syed, A., Chandramohan, R., Liu, Z.Y., Won, H.H., Scott, S.N., et al. (2015). Memorial Sloan Kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. J. Mol. Diagn. *17*, 251–264.

Chou, T.C. (2010). Drug combination studies and their synergy quantification using the Chou-Talalay method. Cancer Res. *70*, 440–446.

Cote, J., Boisvert, F.M., Boulanger, M.C., Bedford, M.T., and Richard, S. (2003). Sam68 RNA binding protein is an in vivo substrate for protein arginine N-methyltransferase 1. Mol. Biol. Cell *14*, 274–287.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372.

Denman, R.B. (2002). Methylation of the arginine-glycine-rich region in the fragile X mental retardation protein FMRP differentially affects RNA binding. Cell. Mol. Biol. Lett. *7*, 877–883.

Dhar, S., Vemulapalli, V., Patananan, A.N., Huang, G.L., Di Lorenzo, A., Richard, S., Comb, M.J., Guo, A., Clarke, S.G., and Bedford, M.T. (2013). Loss of the major type I arginine methyltransferase PRMT1 causes substrate scavenging by other PRMTs. Sci. Rep. *3*, 1311.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15–21.

Dolzhanskaya, N., Merz, G., Aletta, J.M., and Denman, R.B. (2006). Methylation regulates the intracellular protein-protein and protein-RNA interactions of FMRP. J. Cell Sci. *119*, 1933–1946.

Duncan, K.W., Rioux, N., Boriack-Sjodin, P.A., Munchhof, M.J., Reiter, L.A., Majer, C.R., Jin, L., Johnston, L.D., Chan-Penebre, E., Kuplast, K.G., et al. (2016). Structure and property guided design in the identification of PRMT5 tool compound EPZ015666. ACS Med. Chem. Lett. *7*, 162–166.

Dvinge, H., Kim, E., Abdel-Wahab, O., and Bradley, R.K. (2016). RNA splicing factors as oncoproteins and tumour suppressors. Nat. Rev. Cancer *16*, 413–430.

Eden, E., Navon, R., Steinfeld, I., Lipson, D., and Yakhini, Z. (2009). GOrilla: a tool for discovery and visualization of enriched GO terms in ranked gene lists. BMC Bioinformatics *10*, 48.

Eram, M.S., Shen, Y., Szewczyk, M., Wu, H., Senisterra, G., Li, F., Butler, K.V., Kaniskan, H.U., Speed, B.A., Dela Sena, C., et al. (2016). A potent, selective, and cell-active inhibitor of human type I protein arginine methyltransferases. ACS Chem. Biol. *11*, 772–781.

Erce, M.A., Abeygunawardena, D., Low, J.K., Hart-Smith, G., and Wilkins, M.R. (2013). Interactions affected by arginine methylation in the yeast protein-protein interaction network. Mol. Cell Proteomics *12*, 3184–3198.

Ernst, T., Chase, A.J., Score, J., Hidalgo-Curtis, C.E., Bryant, C., Jones, A.V., Waghorn, K., Zoi, K., Ross, F.M., Reiter, A., et al. (2010). Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat. Genet. *42*, 722–726.

Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B., et al. (2018). The reactome pathway knowledgebase. Nucleic Acids Res. *46*, D649–D655.

Fedoriw, A., Rajapurkar, S.R., O'Brien, S., Gerhart, S.V., Mitchell, L.H., Adams, N.D., Rioux, N., Lingaraj, T., Ribich, S.A., Pappalardi, M.B., et al. (2019). Anti-tumor activity of the type I PRMT inhibitor, GSK3368715 with PRMT5 inhibition through MTAP loss. Cancer Cell https://doi.org/10.1016/j.ccell.2019.05.014.

Folco, E.G., Coil, K.E., and Reed, R. (2011). The anti-tumor drug E7107 reveals an essential role for SF3b in remodeling U2 snRNP to expose the branch point-binding region. Genes Dev. *25*, 440–444.

Friesen, W.J., Massenet, S., Paushkin, S., Wyce, A., and Dreyfuss, G. (2001). SMN, the product of the spinal muscular atrophy gene, binds preferentially to dimethylarginine-containing protein targets. Mol. Cell *7*, 1111–1117.

Gao, G., Zhang, L., Villarreal, O.D., He, W., Su, D., Bedford, E., Moh, P., Shen, J., Shi, X., Bedford, M.T., et al. (2019). PRMT1 loss sensitizes cells to PRMT5 inhibition. Nucleic Acids Res. *47*, 5038–5048.

Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci. Signal. *6*, pl1.

Garrido-Martín, D., Palumbo, E., Guigó, R., and Breschi, A. (2018). ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. PLoS Comput. Biol. *14*, e1006360.

Guo, Z., Zheng, L., Xu, H., Dai, H., Zhou, M., Pascua, M.R., Chen, Q.M., and Shen, B. (2010). Methylation of FEN1 suppresses nearby phosphorylation and facilitates PCNA binding. Nat. Chem. Biol. *6*, 766–773.

Hamard, P.J., Santiago, G.E., Liu, F., Karl, D.L., Martinez, C., Man, N., Mookhtiar, A.K., Duffort, S., Greenblatt, S., Verdun, R.E., et al. (2018). PRMT5 regulates DNA repair by controlling the alternative splicing of histone-modifying enzymes. Cell Rep. *24*, 2643–2657.

Han, T., Goralski, M., Gaskill, N., Capota, E., Kim, J., Ting, T.C., Xie, Y., Williams, N.S., and Nijhawan, D. (2017). Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. Science *356*, https://doi.org/10.1126/science.aal3755.

Harbour, J.W., Roberson, E.D., Anbunathan, H., Onken, M.D., Worley, L.A., and Bowcock, A.M. (2013). Recurrent mutations at codon 625 of the splicing factor SF3B1 in uveal melanoma. Nat. Genet. *45*, 133–135.

Hofweber, M., Hutten, S., Bourgeois, B., Spreitzer, E., Niedner-Boblenz, A., Schifferer, M., Ruepp, M.D., Simons, M., Niessing, D., Madl, T., et al. (2018). Phase separation of FUS is suppressed by its nuclear import receptor and arginine methylation. Cell *173*, 706–719.e13.

Hsu, T.Y., Simon, L.M., Neill, N.J., Marcotte, R., Sayad, A., Bland, C.S., Echeverria, G.V., Sun, T., Kurley, S.J., Tyagi, S., et al. (2015). The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature *525*, 384–388.

Huber, W., Carey, V.J., Gentleman, R., Anders, S., Carlson, M., Carvalho, B.S., Bravo, H.C., Davis, S., Gatto, L., Girke, T., et al. (2015). Orchestrating high-throughput genomic analysis with Bioconductor. Nat. Methods *12*, 115–121.

Jansson, M., Durant, S.T., Cho, E.C., Sheahan, S., Edelmann, M., Kessler, B., and La Thangue, N.B. (2008). Arginine methylation regulates the p53 response. Nat. Cell Biol. *10*, 1431–1439.

Kahles, A., Lehmann, K.V., Toussaint, N.C., Huser, M., Stark, S.G., Sachsenberg, T., Stegle, O., Kohlbacher, O., Sander, C., Cancer Genome Atlas Research, N., et al. (2018). Comprehensive analysis of alternative splicing across tumors from 8,705 patients. Cancer Cell *34*, 211–224.e6.

Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., et al. (2007).

Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat. Chem. Biol. *3*, 576–583.

Karni, R., de Stanchina, E., Lowe, S.W., Sinha, R., Mu, D., and Krainer, A.R. (2007). The gene encoding the splicing factor SF2/ASF is a proto-oncogene. Nat. Struct. Mol. Biol. *14*, 185–193.

Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C., Ramakrishnan, A., Li, Y., Chung, Y.R., Micol, J.B., Murphy, M.E., et al. (2015). SRSF2 mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. Cancer Cell *27*, 617–630.

Koh, C.M., Bezzi, M., Low, D.H., Ang, W.X., Teo, S.X., Gay, F.P., Al-Haddawi, M., Tan, S.Y., Osato, M., Sabo, A., et al. (2015). MYC regulates the core pre-mRNA splicing machinery as an essential step in lymphomagenesis. Nature *523*, 96–100.

Kotake, Y., Sagane, K., Owa, T., Mimori-Kiyosue, Y., Shimizu, H., Uesugi, M., Ishihama, Y., Iwata, M., and Mizui, Y. (2007). Splicing factor SF3b as a target of the antitumor natural product pladienolide. Nat. Chem. Biol. *3*, 570–575.

Kotini, A.G., Chang, C.J., Boussaad, I., Delrow, J.J., Dolezal, E.K., Nagulapally, A.B., Perna, F., Fishbein, G.A., Klimek, V.M., Hawkins, R.D., et al. (2015). Functional analysis of a chromosomal deletion associated with myelodysplastic syndromes using isogenic human induced pluripotent stem cells. Nat. Biotechnol. *33*, 646–655.

Kotini, A.G., Chang, C.J., Chow, A., Yuan, H., Ho, T.C., Wang, T., Vora, S., Solovyov, A., Husser, C., Olszewska, M., et al. (2017). Stage-specific human induced pluripotent stem cells map the progression of myeloid transformation to transplantable leukemia. Cell Stem Cell *20*, 315–328.e7.

Kryukov, G.V., Wilson, F.H., Ruth, J.R., Paulk, J., Tsherniak, A., Marlow, S.E., Vazquez, F., Weir, B.A., Fitzgerald, M.E., Tanaka, M., et al. (2016). MTAP deletion confers enhanced dependency on the PRMT5 arginine methyltransferase in cancer cells. Science *351*, 1214–1218.

Lee, S.C., Dvinge, H., Kim, E., Cho, H., Micol, J.B., Chung, Y.R., Durham, B.H., Yoshimi, A., Kim, Y.J., Thomas, M., et al. (2016). Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. Nat. Med. *22*, 672–678.

Lee, S.C., North, K., Kim, E., Jang, E., Obeng, E., Lu, S.X., Liu, B., Inoue, D., Yoshimi, A., Ki, M., et al. (2018). Synthetic lethal and convergent biological effects of cancer-associated spliceosomal gene mutations. Cancer Cell *34*, 225–241.e8.

Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N. Engl. J. Med. *368*, 2059–2074.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. *15*, 550.

Marjon, K., Cameron, M.J., Quang, P., Clasquin, M.F., Mandley, E., Kunii, K., McVay, M., Choe, S., Kernytsky, A., Gross, S., et al. (2016). MTAP deletions in cancer create vulnerability to targeting of the MAT2A/PRMT5/RIOK1 axis. Cell Rep. *15*, 574–587.

Martin, M., Masshofer, L., Temming, P., Rahmann, S., Metz, C., Bornfeld, N., van de Nes, J., Klein-Hitpass, L., Hinnebusch, A.G., Horsthemke, B., et al. (2013). Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. Nat. Genet. *45*, 933–936.

Massignani, E., Cuomo, A., Musiani, D., Jammula, S., Pavesi, G., and Bonaldi, T. (2019). hmSEEKER: identification of hmSILAC doublets in maxquant output data. Proteomics *19*, e1800300.

Mavrakis, K.J., McDonald, E.R., 3rd, Schlabach, M.R., Billy, E., Hoffman, G.R., deWeck, A., Ruddy, D.A., Venkatesan, K., Yu, J., McAllister, G., et al. (2016). Disordered methionine metabolism in MTAP/CDKN2A-deleted cancers leads to dependence on PRMT5. Science *351*, 1208–1213.

Meister, G., Eggert, C., Buhler, D., Brahms, H., Kambach, C., and Fischer, U. (2001). Methylation of Sm proteins by a complex containing PRMT5 and the putative U snRNP assembly factor pICln. Curr. Biol. *11*, 1990–1994.

Meister, G., and Fischer, U. (2002). Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. EMBO J. *21*, 5853–5863.

Musiani, D., Bok, J., Massignani, E., Wu, L., Tabaglio, T., Ippolito, M.R., Cuomo, A., Ozbek, U., Zorgati, H., Ghoshdastider, U., et al. (2019). Proteomics profiling of arginine methylation defines PRMT5 substrate specificity. Sci. Signal. *12*, https://doi.org/10.1126/scisignal.aat8388.

Nichols, R.C., Wang, X.W., Tang, J., Hamilton, B.J., High, F.A., Herschman, H.R., and Rigby, W.F. (2000). The RGG domain in hnRNP A2 affects subcellular localization. Exp. Cell Res. *256*, 522–532.

Nikoloski, G., Langemeijer, S.M., Kuiper, R.P., Knops, R., Massop, M., Tonnissen, E.R., van der Heijden, A., Scheele, T.N., Vandenberghe, P., de Witte, T., et al. (2010). Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. Nat. Genet. *42*, 665–667.

O'Shea, J.P., Chou, M.F., Quader, S.A., Ryan, J.K., Church, G.M., and Schwartz, D. (2013). pLogo: a probabilistic approach to visualizing sequence motifs. Nat. Methods *10*, 1211–1212.

Obeng, E.A., Chappell, R.J., Seiler, M., Chen, M.C., Campagna, D.R., Schmidt, P.J., Schneider, R.K., Lord, A.M., Wang, L., Gambe, R.G., et al. (2016). Physiologic expression of Sf3b1(K700E) causes impaired erythropoiesis, aberrant splicing, and sensitivity to therapeutic spliceosome modulation. Cancer Cell *30*, 404–417.

Ong, S.E., Mittler, G., and Mann, M. (2004). Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. Nat. Methods *1*, 119–126.

Pellagatti, A., Armstrong, R.N., Steeples, V., Sharma, E., Repapi, E., Singh, S., Sanchi, A., Radujkovic, A., Horn, P., Dolatshad, H., et al. (2018). Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. Blood *132*, 1225–1240.

Qamar, S., Wang, G., Randle, S.J., Ruggeri, F.S., Varela, J.A., Lin, J.Q., Phillips, E.C., Miyashita, A., Williams, D., Strohl, F., et al. (2018). FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-pi interactions. Cell *173*, 720–734.e15.

Rappsilber, J., Mann, M., and Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat. Protoc. *2*, 1896–1906.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337–342.

Seiler, M., Yoshimi, A., Darman, R., Chan, B., Keaney, G., Thomas, M., Agrawal, A.A., Caleb, B., Csibi, A., Sean, E., et al. (2018). H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nat. Med. *24*, 497–504.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. *13*, 2498–2504.

Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U S A *111*, E5593–E5601.

Shirai, C.L., White, B.S., Tripathi, M., Tapia, R., Ley, J.N., Ndonwi, M., Kim, S., Shao, J., Carver, A., Saez, B., et al. (2017). Mutant U2AF1-expressing cells are sensitive to pharmacological modulation of the spliceosome. Nat. Commun. *8*, 14060.

Sims, R.J., 3rd, Rojas, L.A., Beck, D.B., Bonasio, R., Schuller, R., Drury, W.J., 3rd, Eick, D., and Reinberg, D. (2011). The C-terminal domain of RNA polymerase II is modified by site-specific methylation. Science *332*, 99–103.

Snijders, A.P., Hautbergue, G.M., Bloom, A., Williamson, J.C., Minshull, T.C., Phillips, H.L., Mihaylov, S.R., Gjerde, D.T., Hornby, D.P., Wilson, S.A., et al. (2015). Arginine methylation and citrullination of splicing factor proline- and glutamine-rich (SFPQ/PSF) regulates its association with mRNA. RNA *21*, 347–359.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U S A *102*, 15545–15550.

Supek, F., Bosnjak, M., Skunca, N., and Smuc, T. (2011). REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One *6*, e21800.

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

Thandapani, P., Song, J., Gandin, V., Cai, Y., Rouleau, S.G., Garant, J.M., Boisvert, F.M., Yu, Z., Perreault, J.P., Topisirovic, I., et al. (2015). Aven recognition of RNA G-quadruplexes regulates translation of the mixed lineage leukemia protooncogenes. Elife *4*, https://doi.org/10.7554/eLife.06234.

Tradewell, M.L., Yu, Z., Tibshirani, M., Boulanger, M.C., Durham, H.D., and Richard, S. (2012). Arginine methylation by PRMT1 regulates nuclear-cytoplasmic localization and toxicity of FUS/TLS harbouring ALS-linked mutations. Hum. Mol. Genet. *21*, 136–149.

Tsai, W.C., Gayatri, S., Reineke, L.C., Sbardella, G., Bedford, M.T., and Lloyd, R.E. (2016). Arginine demethylation of G3BP1 promotes stress granule assembly. J. Biol. Chem. *291*, 22671–22685.

Uehara, T., Minoshima, Y., Sagane, K., Sugi, N.H., Mitsuhashi, K.O., Yamamoto, N., Kamiyama, H., Takahashi, K., Kotake, Y., Uesugi, M., et al. (2017). Selective degradation of splicing factor CAPERalpha by anticancer sulfonamides. Nat. Chem. Biol. *13*, 675–680.

Vizcaino, J.A., Csordas, A., del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res. *44*, D447–D456.

Wall, M.L., and Lewis, S.M. (2017). Methylarginines within the RGG-motif region of hnRNP A1 affect its IRES trans-acting factor activity and are required for hnRNP A1 stress granule localization and formation. J. Mol. Biol. *429*, 295–307.

Wang, E., Lu, S.X., Pastore, A., Chen, X., Imig, J., Chun-Wei Lee, S., Hockemeyer, K., Ghebrechristos, Y.E., Yoshimi, A., Inoue, D., et al. (2019). Targeting an RNA-binding protein network in acute myeloid leukemia. Cancer Cell *35*, 369–384.e7.

Wang, L., Lawrence, M.S., Wan, Y., Stojanov, P., Sougnez, C., Stevenson, K., Werner, L., Sivachenko, A., DeLuca, D.S., Zhang, L., et al. (2011). SF3B1 and other novel cancer genes in chronic lymphocytic leukemia. N. Engl. J. Med. *365*, 2497–2506.

Wu, G., Dawson, E., Duong, A., Haw, R., and Stein, L. (2014). ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. F1000Res *3*, 146.

Xu, W., Chen, H., Du, K., Asahara, H., Tini, M., Emerson, B.M., Montminy, M., and Evans, R.M. (2001). A transcriptional switch mediated by cofactor methylation. Science *294*, 2507–2511.

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.

Zehir, A., Benayed, R., Shah, R.H., Syed, A., Middha, S., Kim, H.R., Srinivasan, P., Gao, J., Chakravarty, D., Devlin, S.M., et al. (2017). Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. Nat. Med. *23* (6), 703–713.

Zhang, J., Lieu, Y.K., Ali, A.M., Penson, A., Reggio, K.S., Rabadan, R., Raza, A., Mukherjee, S., and Manley, J.L. (2015a). Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. Proc. Natl. Acad. Sci. U S A *112*, E4726–E4734.

Zhang, L., Tran, N.T., Su, H., Wang, R., Lu, Y., Tang, H., Aoyagi, S., Guo, A., Khodadadi-Jamayran, A., Zhou, D., et al. (2015b). Cross-talk between PRMT1-mediated methylation and ubiquitylation on RBM15 controls RNA splicing. Elife *4*, e07938.

Zhang, J., and Manley, J.L. (2013). Misregulation of pre-mRNA alternative splicing in cancer. Cancer Discov. *3*, 1228–1237.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Anti-Actin	Santa Cruz	sc-47778; RRID: AB_626632
Anti-PRMT5	Abcam	Abcam ab109451; RRID: AB_10863428
Anti-SDMA	Cell signaling	#13222; RRID: AB_2714013
Anti-ADMA	Cell signaling	#13522; RRID: AB_2665370
Anti-PRMT1	Cell signaling	CST #2449; RRID: AB_2237696
Anti-MMA	Cell signaling	#8015; RRID: AB_10891776
Anti-EZH2	Cell signaling	CST #5246; RRID: AB_10694683
Anti- γH2Ax	Cell signalling	CST #9718; RRID: AB_2118009
Biological Samples		
AML primary patient samples	Princess Margaret Hospital, Memorial Sloan Kettering Cancer Center, & Weill Cornell College of Medicine	N/A
Chemicals, Peptides, and Recombinant Proteins		
GSK3203591	SGC Toronto	https://thesgc.org/chemical-probes/GSK591
MS023	SGC Toronto	https://www.thesgc.org/chemical-probes/MS023
EPZ015666	Tocris	Cat. No. 6516
GSK3368712	GSK	N/A
E7107	H3 Biomedicine Inc.	N/A
I-BRD9	SGC Toronto	https://www.thesgc.org/chemical-probes/I-BRD9
SGC-CBP30	SGC Toronto	https://www.thesgc.org/chemical-probes/CBP30
BI-9564	SGC Toronto	https://www.thesgc.org/chemical-probes/BI-9564
I-CBP112	SGC Toronto	https://www.thesgc.org/chemical-probes/I-CBP112
MS049	SGC Toronto	https://www.thesgc.org/chemical-probes/MS049
UNC0638	SGC Toronto	https://thesgc.org/chemical-probes/UNC0638
SGC707	SGC Toronto	https://www.thesgc.org/chemical-probes/SGC707
GSK484	SGC Toronto	https://www.thesgc.org/chemical-probes/GSK484
IOX2	SGC Toronto	https://www.thesgc.org/chemical-probes/IOX2
GSK-J4	SGC Toronto	https://www.thesgc.org/chemical-probes/gskj1
Vorinostat	Cayman chemicals	*CAS № 149647-78-9*
GSK2334470	Cayman chemicals	*CAS № 1227911-45-6*
SGC0946	SGC Toronto	https://www.thesgc.org/chemical-probes/SGC0946
PD0332991	Cayman chemicals	*CAS № 827022-32-2*
CCT244747	Cayman chemicals	*CAS № 1404095-34-6*
PF-477736	Cayman chemicals	*CAS № 952021-60-2*
BMS-265246	Cayman chemicals	*CAS № 582315-72-8*
Irinotecan	Cayman chemicals	*CAS № 100286-90-6*
SCH772984	Cayman chemicals	*CAS № 942183-80-4*
Alitretinoin	Cayman chemicals	*CAS № 5300-03-8*
CX-4945	Cayman chemicals	*CAS № 1009820-21-6*
Phenoxodiol	Sigma	*CAS №81267-65-4*
5Z-7-oxozeaenol	Cayman chemicals	*CAS № 253863-19-3*
Critical Commercial Assays		
Caspase-Glo. 3/7 luminescence assay	Promega	G8090
Cell titer Glo	Promega	G7570

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
MTS assay	Promega	G3582
PTMScan® Asymmetric Di-Methyl Arginine Motif [adme-R] Kit	Cell Signaling Technologies	Kit #13474
PTMScan® Symmetric Di-Methyl Arginine Motif [sdme-RG] Kit	Cell Signaling Technologies	Kit #13563
PTMScan Mono-Methyl Arginine 13 Motif [mme-RG] Kit	Cell Signaling Technologies	Kit #12235
Deposited Data		
RNA-seq data	This paper	GEO: GSE123774
MS-proteomics data	This paper	PRIDE: PXD012007
Experimental Models: Cell Lines		
Human: K562 cells	ATCC	CLL-243
Human: THP1 cells	ATCC	TIB-202
Human: MOLM13 cells	DSMZ	ACC 554
Human: U937 cells	ATCC	CRL-1593.2
Human: K562 *SRSF2*$^{P95H/WT}$ knockin cells	Horizon Discovery	N/A
Human: K562 *SF3B1*$^{K700E/WT}$ knockin cells	Horizon Discovery	N/A
Human: TF1 cells	ATCC	CRL-2003
Human: OCI-AML5 cells	DSMZ	ACC 247
Human: F-36P cells	DSMZ	ACC 543
Human: GDM-1 cells	ATCC	CRL-2627
Human: KO52 cells	JCRB	CVCL-1321
Human: HNT34 cells	DSMZ	ACC 600
Human: MonoMac6 cells	DSMZ	ACC 124
Human: 5-16 Cre20 (SRSF2 P95L)	(Chang et al., 2018)	N/A
Human: N-2.12 (isogenic normal)	(Chang et al., 2018)	N/A
Experimental Models: Organisms/Strains		
Mice: *Srsf2*$^{P95H/+}$	(Kim et al., 2015).	N/A
Mice: *MLL-AF9/Vav-cre Srsf2*$^{WT/WT}$	(Kim et al., 2015).	N/A
Mice: MLL-AF9/Vav-cre Srsf2$^{P95H/WT}$	(Kim et al., 2015).	N/A
Mice: Vav-cre transgenic mice (B6.Cg-Tg(Vav1-icre)A2Kio/J)	The Jackson Laboratory	JAX: 008610
Oligonucleotides		
Human *EZH2* F	(Lee et al., 2016)	TTTCATGCAACACCCAACACT
Human *EZH2* R	(Lee et al., 2016)	CCCTGCTTCCCTATCACTGT
Human *ATF2* F	This paper	AGTTACATGTGAATTCTGCCAGG
Human *ATF2* R	This paper	CTCAAATGGACTCGCCAACTC
Human *INTS3* F	This paper	ATGCCAAGCTGGCTTTGTTTT
Human *INTS3* R	This paper	TCCGACATATGGTTGTCCATCTC
Human *TRPT1* F	This paper	GGCCAACCAGGGCCATT
Human *TRPT1* R	This paper	ATCACCAGCCAAGGAAAGGG
Human *HDAC7* F	This paper	GGAAGAATCCACTGCTCCGA
Human *HDAC7* R	This paper	GACTGGGCAAAGTGGAAGGG
Human *LEF1* F	This paper	CCACCCATCCCGAGAACATC
Human *LEF1* R	This paper	AGGCTTCACGTGCATTAGGT
Anti-sense oligonucleotide SCR	This paper	CGGUGUGUGUAUCAUUCUCUAGUGU
Anti-sense oligonucleotide *EZH2*	This paper	UGAAUCUUCUGUCCAAAAUCCAACAGGCAAUAUA
sgRNA 1 sequence for *EZH2*	This paper	TTATCAGAAGGAAATTTCCG
sgRNA 2 sequence for *EZH2*	This paper	TTATGATGGGAAAGTACACG

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant DNA		
LentiCRISPR v2	N/A	Addgene: 52961
Software and Algorithms		
Cytoscape 3.4.0	(Shannon et al., 2003)	https://cytoscape.org/
Reactome FI	(Wu et al., 2014)	N/A
pLogo web application	(O'Shea et al., 2013)	N/A
hmLINKER	(Massignani et al., 2019)	N/A
hmSEEKER	(Massignani et al., 2019)	N/A
Prism 7.0	Graphpad	https://graphpad.com/
ImageJ	NIH	https://imagej-nih-gov.libproxy1.nus.edu.sg/ij/index.html
STAR	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
HTSeq	(Anders et al., 2015)	N/A
R	R Core Team	https://www.r-project.org/
Bioconductor	(Huber et al., 2015)	N/A
DESeq2	(Love et al., 2014)	10.18129/B9.bioc.DESeq2
rMATS	(Shen et al., 2014)	http://rnaseq-mats.sourceforge.net/
ggsashimi	(Garrido-Martín et al., 2018)	https://github.com/guigolab/ggsashimi

LEAD CONTACT AND MATERIALS AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Ernesto Guccione (eguccione@imcb.a-star.edu.sg; Ernesto.guccione@mssm.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell Lines and Cell Culture

Primary murine *MLL-AF9* leukemia cell lines were generated from bone marrow cells of leukemia-bearing mice and maintained in IMDM/15% FCS supplemented with L-glutamine, murine-SCF (25 ng/mL), murine-IL3 (10 ng/mL) and murine-IL6 (10 ng/mL). K562 isogenic cell lines (engineered to express *SRSF2*^{P95H} *or SF3B1*^{K700E} mutations from each respective endogenous locus from Horizon Discovery Inc.) were cultured in IMDM/10% FBS. THP-1, U937, MOLM13, KO52, GDM-1 and HNT34 cell lines were cultured in RPMI/10% FBS. TF1 cells were cultured in RMPI/10% FBS/ 2 ng/ml recombinant human GM-CSF. OCI-AML5 and F-36p cells were cultured in RPMI/10% FBS/ 10 ng/ml recombinant human GM-CSF. MonoMac 6 cells were cultured in RPMI/10% FBS/10 μg/ml recombinant human insulin. HEK293T cells were grown in DMEM medium with 10% FBS. The iPSC lines 5-16 Cre20 (SRSF2 P95L) and N-2.12 (isogenic normal) (Chang et al., 2018) were differentiated along the hematopoietic lineage as previously described (Kotini et al., 2017). To induce reprogramming, 10,000-300,000 cells were plated on retronectin-coated 24-well dishes and transduced with the OKMS lentiviral vector CMV-fSV2A (Kotini et al., 2015). The cells were harvested one or two days later and plated on mitotically inactivated MEFs in 6-well plates and centrifuged at 500 rpm for 30 min at RT. The following day and every day thereafter, half of the medium was gently changed to hESC medium with 0.5 mM valproic acid (VPA). In the first 10 days, cells contained in the removed medium were collected by centrifugation and placed back in their original wells. After 3-4 weeks, manually picked colonies with hPSC morphology were expanded. iPSCs were cultured on mitotically inactivated MEFs or in feeder-free conditions on Matrigel with hESC media supplemented with 6 ng/ml FGF2. Primary human AML patient sample cells were grown in Iscove's Modified Dulbecco's Medium (IMDM), 20% bovine serum albumin, insulin, and transferrin (BIT) 9500 serum substitute, 16.7 μg/ml human low-density lipoproteins, 55 μM beta-mercaptoethanol with recombinant human (rh) G-CSF (20 ng/ml), rhGM-CSF (20 ng/ml), rh IL3 (20 ng/ml), rh IL6 (20 ng/ml), rh FLT3 ligand (50 ng/ml) and rh SCF (50 ng/mL).

Primary Human Samples

Studies were approved by the Institutional Review Boards of the University Health Network, Weill Cornell College of Medicine, and Memorial Sloan Kettering Cancer Center. Studies were conducted in accordance to the Declaration of Helsinki protocol. De-identified primary human AML samples derived from whole peripheral blood or BM mononuclear cells were utilized. The MSK-IMPACT assay was used to perform mutational genotyping of samples, as described previously (Cheng et al., 2015; Zehir et al., 2017). Cord blood was acquired from NY Blood Bank. Informed consent was obtained from all subjects to use the specimens described

in this study. Specimens were obtained as part of the Memorial Sloan- Kettering Cancer Center Institutional Review Board approved clinical protocol #06-107 to which all subjects consented. O.A-W is a participating investigator on this protocol.

Animals

6-8 weeks old C57BL/6 mice were purchased from InVivos. Mice were maintained in individual ventilated cages and fed with autoclaved food and water in the Biological Resource Center, A*STAR Singapore. 10 week-old NSG-SGM3 (NSGS) female mice were purchased from Jackson Laboratory. Mice were maintained in individual ventilated cages and fed with autoclaved food and water in Memorial Sloan Kettering Cancer Center. All animal experiments were conducted in accordance to approved protocols from Institutional Animal Care and Use Committees of Institute of Molecular and Cell Biology (IMCB), A*STAR and Memorial Sloan Kettering Cancer Center. Kaplan-Meier survival curves were compared using the Mantel-Cox Log-rank test via Graphpad Prism 7.

METHOD DETAILS

Identification of Druggable Targets in the Extended Splicing Network

A list of genes belonging to the Gene Ontology (GO) category GO_SPLICEOSOMAL_ SNRNP_ ASSEMBLY was derived from the Gene Set Enrichment Analysis (GSEA) website (Subramanian et al., 2005). We uploaded this list on the CBIO portal website (Cerami et al., 2012) and selected as a query ''Acute Myeloid Leukemia, TCGA'' study (Cerami et al., 2012; Gao et al., 2013; Ley et al., 2013). The network tab provided a list of neighboring genes (sif file). This network comprises the nearest neighbors of a given physical entity (e.g. gene, protein or small molecule). The following rules govern the construction of the neighborhood: 1) if A is part of a [complex] (A:B), (A:B) is included in the neighborhood, but none of the interactions involving (A:B) are included. 2) if A is a [CONTROLLER] for a [control] interaction, the reaction that is [CONTROLLED] (and all the participants in that reaction) are included in the neighborhood. 3) if A participates in a [conversion] reaction, and this reaction is [CONTROLLED] by another interaction, the [control] interaction (plus its [CONTROLLER]) are included in the neighborhood. The sif file was then imported into Cytoscape 3.4.0(Shannon et al., 2003) to visualize molecular interaction networks and integrate the data with gene expression profiles and other state data. We next created a network of functional interacting genes using the application 'Reactome FI'.(Wu et al., 2014) The list of genes belonging to the extended splicing network was manually curated to identify druggable targets. We used www.drugbank.ca and www.chemicalprobes.org to identify the inhibitors of the druggable genes.

Animal Studies

In leukemia transplant experiments, mice were monitored daily for any sign of distress and leukemia development. The number of mice chosen in each experiment was chosen to give 90% statistical power with a 5% error level given the differences in standard deviation that was observed in the pilot study. Generation and genotyping of the *Srsf2*$^{P95H/+}$ conditional knock-in mice as well as *MLL-AF9/Vav-cre Srsf2*$^{WT/WT}$ (referred to *Srsf2*^{WT} throughout the text) and *MLL-AF9/Vav-cre Srsf2*$^{P95H/WT}$ (referred to *Srsf2*^{P95H} throughout the text), all on C57BL/6 backgrounds, are as previously described (Kim et al., 2015). Briefly, Vav-Cre^{+}Srsf2 $^{WT/WT}$ and Vav-Cre^{+}Srsf2$^{P95H/WT}$ mice were treated with a single dose 5-fluorouracil (150 mg/kg) and bone marrow were harvested 6 days later. The red blood cells from the bone marrow were lysed by ammonium-chloride potassium bicarbonate lysis buffer (ACK lysis buffer). The bone marrow cells were then transduced with viral supernatant containing the murine stem cell virus (MSCV)-driven MLL-AF9 fusion oncogene with an internal ribosomal entry site to express a GFP tag (MSCV-MLL-AF9-IRES-GFP) for 2 days in Iscove's modified Dulbecco's medium (IMDM) containing 15% fetal bovine serum (FCS) and supplemented with mouse stem cell factor (mSCF) (25 ng/ml), mouse interleukin-6 (10 ng/ml) and mouse interleukin-3 (10 ng/ml). Thereafter, lethally irradiated (9.5 Gy) C57BL6/J mice are injected with ~400,000 cells via tail vein injection to generate primary leukemias. All *in vivo* drug experiments were conducted as secondary leukemia transplants whereby 6-8 week old C57BL6/J mice are sublethally irradiated (5 Gy) and injected with 250,000 MLL-AF9 leukemia cells collected from moribund mice with leukemia (Lee et al., 2016).

AML PDXs were generated from patient peripheral blood and/or bone marrow mononuclear cells and subsequently transplanted intrafemorally into 10 week-old NSG-SGM3 (NSGS) female mice (Jackson Laboratory). Once human CD45 chimerism was >25% of the bone marrow, NSGS mice were eventually treated with 60 mg/kg MS023 and 150 mg/kg EPZ015666 for 6 weeks.

In Vitro Drug Screen with Murine *MLL-AF9* Cells

Both *MLL-AF9/Vav-cre Srsf2*^{WT} and *MLL-AF9/Vav-cre Srsf2*^{P95H} cells were seeded at 900 cells/well in 384 well plate. Cells were treated with indicated drugs for 7 days over various concentrations (see Table S1). Cell viability was measured using MTS or Cell TiterGlo assay (Promega) as per manufacturer's instructions and values were normalized to DMSO controls.

In Vitro Drug Titration with Murine *MLL-AF9* Cells

All cell lines were seeded in white flat-well 96 well plates (Costar) at 1000cells/well unless otherwise indicated. For murine *MLL-AF9* cell lines single drug titration experiments, cells were exposed to MS023 from a range of 0-5 μM, to GSK591 from a range of 0-5 μM and to E7107 from a range of 0-1 μM for seven days. For murine *MLL-AF9* cell lines combination drug titration experiments, cells were exposed to MS023 from a range of 0-5 μM, to GSK591 from a range of 0-0.5 μM and to E7107 from a range of 0-0.05 μM for seven days. Cell viability read-out was performed using the Cell-titre glo assay (Promega) as per manufacturer's instructions and normalized to DMSO controls.

In Vitro Drug Treatment with Primary Human AML Patient Samples

Primary AML cells with SF mutation (n=16) or wild-type for SF (n=16) were incubated with DMSO or GSK591 (0.5 μM) for 6 days. Cells were grown in Iscove's Modified Dulbecco's Medium (IMDM), 20% bovine serum albumin, insulin, and transferrin (BIT) 9500 serum substitute, 16.7 μg/ml human low-density lipoproteins, 55 μM beta-mercaptoethanol with recombinant human (rh) G-CSF (20 ng/ml), rhGM-CSF (20 ng/ml), rh IL3 (20 ng/ml), rh IL6 (20 ng/ml), rh FLT3 ligand (50 ng/ml) and rh SCF (50 ng/mL). Following the 6-day incubation period, cells were subjected to flow cytometry to detect 7-AAD negative, and YO-PRO1 negative, viable cells. Samples were prepared in 4-6 replicates and averages were calculated. Relative viable cell numbers were compared with Welch's t test.

In Vitro Combination Drug Titration with Primary Human AML Patient Samples

Patient cells derived from either bone marrow or peripheral blood were thawed and allowed to grow for a minimum of three days on OP9 stroma in 6-well tissue culture plates (Greiner) topping up with fresh media and putting cells onto fresh OP9 as necessary. Patient cells were cultured in IMDM media (Gibco) with 2 mM L-glutamine, 10% FBS (Wisent), 55 μM β-mercaptoethanol (Gibco), 100 μg/mL Primocin (InvivoGen), 100 ng/mL SCF, 50 ng/mL FLT3L, 40 ng/mL THPO, 20 ng/mL IL3, and 20 ng/mL GM-CSF (Shenandoah Biotech or Custom Biologics). Leukemic cells were transferred to GFP labelled OP9 cells in 96-well plates for drug titration at a density of either 25000 or 50000 cells per well in 100 μL media. An additional 100 μL of media with drug was added at day 3-4. Cells were exposed to MS023 from a range of 0-10 μM and to GSK591 from a range of 0-3 μM for six days.

Viable cell number was assessed at day 6. Cells were transferred to 96-well round well suspension plate (Starstedt) along with trypsinized (30uL 0.25% Trypsin-EDTA; Wisent) OP9 stroma and attached leukemic cells. Cells were resuspended in PBS (Wisent) with 2% FBS with 0.2 μM Sytox Blue (Life Technologies) viability dye. Flow cytometry was performed using MACSQuant VYB (Miltenyi), and MACSQuantify software was used to determine viable leukemic cell number (GFP negative, Sytox blue negative).

OP9 mouse stromal cells (ATCC) were grown in α-MEM media without nucleosides that contains GlutaMAX (Gibco), 20% fetal bovine serum (FBS, Wisent), 55 μM β-mercaptoethanol (Gibco) and 100U/100 μg/mL penicillin/streptomycin (Gibco) at 37°C/5% CO2. OP9 cells were transduced with GFP lentivirus and sorted for GFP positive cells to use in assays to test the effect of the drugs on leukemic blasts. GFP positive cells were seeded at 5000-10000 cells/well in 96-well plates (Greiner) and used 1-3 days later for leukemic cell assays.

In Vitro Cell Viability Assays with Human Cell Lines

Cells were treated in duplicated with a 20-point, two-fold dilution series of GSK3203591 (PRMT5 inhibitor) or GSK3368712 (Type I PRMT inhibitor). Cells were treated with drugs for six days and cell growth was measured using Cell-titre glo (Promega).

A plate of untreated cells was read at the time of compound addition to determine the T=0 value representing the starting number of cells. Data were fitted with a four-parameter equation to generate a concentration response curve. Growth inhibition is the percent maximal inhibition and was calculated as 100-((ymin-100)/(ymax-100)*100). Ymin-T0 values were calculated by subtracting the T0 value (100%) from the Ymin value on the curve, and are a measure of net population cell growth or death. Growth Death Index (GDI) is a composite representation of Ymin-T0 and percent maximal inhibition. If Ymin-T0 values are negative, then GDI equals Ymin-T0; otherwise, GDI represents the fraction of cells remaining relative to DMSO control (ymax) and (ymin): (ymin-100)/(ymax-100)*100).

Treatment of iPSC-derived Hematopoietic Progenitor Cells with PRMT Inhibitors

The iPSC lines 5-16 Cre20 (SRSF2 P95L) and N-2.12 (isogenic normal) (Chang et al., 2018) were differentiated along the hematopoietic lineage as previously described (Kotini et al., 2017). 250.000 single cells from day 12 of hematopoietic differentiation were plated in 24-well ultra-low attachment plates. MS023 and GSK591 inhibitors or DMSO were added on days 12 and 17 of hematopoietic differentiation culture at a concentration of 350 nM. Live cells were quantified by counting in a hemocytometer 10 days after the beginning of treatment (day 22 of differentiation culture).

Pharmacokinetic Studies of MS023 *In Vivo*

Male Swiss Albino mice at Sai Life Sciences were administered a single 80 mg/kg intraperitoneal (IP) injection (n=3) or a single oral 150 mg/kg dose of MS023 (n=3). Plasma concentrations of MS023 were then evaluated at 0.083, 0.25, 0.5, 1, 2, 4, 8, and 12 hr post dosing.

Administration of PRMT Inhibitors *In Vivo*

For *in vivo* drug sensitivity studies, 8 mg (80 mg/kg dose) or 6 mg (60 mg/kg dose) of MS023 was dissolved in 50 μL N-Methyl-2-pyrrolidone, 200 μL Captisol, 200 μL polyethylene glycol 400 and 550 μL PBS and administered via IP injection once per day. 20 mg/ml (200 mg/kg dose) or 15 mg/ml (150 mg/kg dose) of EPZ015666 was reconstituted in 0.5% methylcellulose and administered to each mouse via oral gavage route once per day. To generate leukemia in C57BL6 mice (6-10 weeks old), the mice were sublethally irradiated at 5 Gy total body γ irradiation, and tail vein injected with 250,000 primary *MLL-AF9* leukemic cells. (Lee et al., 2016). In single drug administration experiment, 80 mg/kg MS023 or 200 mg/kg EPZ015666 was administered from Day 8 of experiment till mice are moribund. In combinatorial drug administration experiment, 60 mg/kg MS023 and 150 mg/kg EPZ015666 were administered daily to the mice from Day 8 of experiment for 10 days.

Western Blotting

Protein concentrations were determined using RC DC protein assay kit from Bio-Rad. Proteins were run on 8-15% gels and separated with SDS-PAGE. The membranes were blotted with the primary antibodies, anti-actin (Santa Cruz) at 1:1000, anti-PRMT5 (Abcam) at 1:1000, anti-PRMT1 (Cell signaling) at 1:1000, anti-SDMA (Cell signaling) at 1:1000, anti-ADMA (Cell signaling) at 1:1000, anti-MMA (Cell signaling) at 1:1000, anti-EZH2 (Cell signaling) at 1:1000 and anti-γH2Ax (Cell signaling) at 1:1000. Blots were incubated with primary antibodies overnight at 4°C. The next day, blots were washed in TBST, incubated with HRP-conjugated secondary antibodies for 1 hr and visualised on X-ray films with West Pico Chemiluminescent Substrate.

RNA-seq Sample Preparation

RNA was extracted from all human and mouse cell samples using RNeasy mini kit (Qiagen) as per manufacturers instructions. Poly(A)-selected, unstranded Illumina libraries were prepared with a modified TruSeq protocol. 0.5X AMPure XP beads were added to the sample library to select for fragments <400 bp, followed by 1X beads to select for fragments >100 bp. These fragments were then amplified with PCR (15 cycles) and separated by gel electrophoresis (2% agarose). 300 bp DNA fragments were isolated and sequenced on the Illumina HiSeq 2000 (~100 million 101 bp reads per sample).

Arginine Methyl-Peptides Separation and Enrichment Prior to LC-MS/MS Analysis

NB4 cells were treated for 3 days with MS023 (3 μM) or GSK591 (1 μM). Equal numbers of Light and Heavy-labelled NB4 cells differentially treated were mixed in a 1:1 ratio, pelleted and washed twice with PBS. Cell pellets were lysed in urea lysis buffer (9 M urea, 20 mM HEPES pH 8.0), supplemented with 1X Roche proteases and phosphatases inhibitors, sonicated and cleared by ultracentrifugation (20.000 x g for 15 min at 15°C). For in-solution digestion, 50 mg of proteins were reduced by adding 4.5 mM DTT (Sigma-Aldrich) for 30 min at 55°C, alkylated with 5.5 mM iodoacetamide (IAA: 10% v/v for 15 min at room temperature in the dark, Sigma Aldrich) and digested overnight with sequencing-grade trypsin (1:100 w/w, Promega), after a fourfold dilution in 25 mM ammonium bicarbonate solution. Protease digestion was terminated with the addition of trifluoroacetic acid (TFA) to adjust pH < 3. Precipitated material was removed by centrifugation for 15 min at 1780 × g at room temperature. Soluble peptides were purified using reversed-phase Sep-Pak C18 cartridges (Waters, Milford, MA) and eluted off the Sep-Pak with 40% can, followed by a step of acetonitrile removal and concentration through 48 hours of lyophilization. Lyophilized peptides were dissolved in 25 mM ammonium hydroxide (NH4OH) and subsequently off-line fractionated by High-pH (HpH) reversed-phased chromatographic separation using a Phenomenex Jupiter® C12 4 μm Proteo 90 Å, LC column 250 x 4.6 mm, on an ÄKTA-FPLC (fast protein liquid chromatography) system (GE Healthcare) operating at 1 ml/min. Buffer A was 25 mM NH_4OH and Buffer B was 25 mM NH_4OH in 90% ACN. Fractions were collected using a collector in a 96-deep well plate at 1-min intervals. Samples were initially loaded onto the column at 1 ml/min for 3 min, after which the fractionation gradient was as follows: 5% B to 30% B in 60 min, 30% B to 60% in 2 min and ramped to 70% B for 3 min. Then, fraction collection was halted and the gradient was held at 100% B for 5 min, before being ramped back to 5% B, when the column was washed. The 60 fractions collected were concatenated to 14 throughout each experiment. After lyophilisation, each concatenated fraction was dissolved in 250 μl of 1x immuno-Affinity Purification Buffer (IAP buffer, #9993, Cell Signaling Technologies, CST) and subjected to two consecutive steps of methyl-R-peptides enrichment using the SDMA antibody-conjugated beads (PTMScan [sdme-R] Kit #13563, Cell Signaling Technologies) and the MMA antibody-conjugated beads (PTMScan Mono-Methyl Arginine Motif [mme-RG] Kit #12235, Cell Signaling Technologies), following the manufacturer's instructions. For the MS023 experiments, ADMA-bearing peptides were also enriched using the specific PTMScan kit #13474, starting from the same amount of desalted and lyophilized peptides. After peptides incubation with the antibody conjugated beads for 2 hours at 4°C, the immuno-precipitates were washed twice in ice-cold IAP buffer, followed by three washes in water; then, the bound methyl-peptides were eluted with two consecutive runs of elution in 50 μl 0.15% TFA. Peptide eluates were desalted on reversed-phase C18 StageTips, as described previously (Rappsilber et al., 2007) and subjected to a second round of trypsin digestion prior to nano-LC-MS/MS analysis.

Nano-LC-MS/MS Analysis

Peptide mixtures were analysed by online nano-flow liquid chromatography tandem mass spectrometry using an EASY-nLC™ 1000 (Thermo Fisher Scientific, Odense, Denmark) connected to a Q-Exactive instrument (Thermo Fisher Scientific) through a nano-electrospray ion source. The nano-LC system was operated in one column set-up with a 50 cm analytical column (75 μm inner diameter, 350 μm outer diameter) packed with C18 resin (EasySpray PEPMAP RSLC C18 2M 50 cm x 75 M, Thermo Fisher Scientific) configuration. Solvent A was 0.1% formic acid (FA) and solvent B was 0.1% FA in 80% ACN. Samples were injected in an aqueous 0.1% TFA solution at a flow rate of 500 nL/min. SILAC immuno-enriched methyl-peptides were separated with a gradient of 5–40% solvent B over 90 min followed by a gradient of 40–60% for 10 min and 60–80% over 5 min at a flow rate of 250 nL/min in the EASY-nLC 1000 system.

The Q-Exactive was operated in the data-dependent mode (DDA) to automatically switch between full scan MS and MSMS acquisition. Survey full scan MS spectra (from m/z 300-1150) were analysed in the Orbitrap detector with resolution R=35,000 at m/z 400. The ten most intense peptide ions with charge states ≥2 were sequentially isolated to a target value of 3e6 and fragmented by Higher Energy Collision Dissociation (HCD) with a normalized collision energy setting of 25%. The maximum allowed ion accumulation times were 20 ms for full scans and 50ms for MSMS and the target value for MSMS was set to 1e6. The dynamic exclusion time was set to 20 s.

Validation of RNAseq

Primers were designed to flank the exon of interest for validation of cassette exon events. cDNA is synthesized using the Maxima first strand cDNA synthesis kit (ThermoFisher Scientific). The PCR cycling conditions to amplify the alternatively spliced transcripts are as follows: 95°C for 5 min, 28 cycles of 95°C for 45 s, 58°C for 30 s and 72°C for 1 min, followed by 72°C for 5 min.

The primers used to validate the splicing events are as follows:

EZH2
Forward primer: TTTCATGCAACACCCAACACT
Reverse primer: CCCTGCTTCCCTATCACTGT
ATF2
Forward primer: AGTTACATGTGAATTCTGCCAGG
Reverse primer: CTCAAATGGACTCGCCAACTC
INTS3
Forward primer: ATGCCAAGCTGGCTTTGTTTT
Reverse primer: TCCGACATATGGTTGTCCATCTC
HDAC7
Forward primer: GGAAGAATCCACTGCTCCGA
Reverse primer: GACTGGGCAAAGTGGAAGGG
TRPT1
Forward primer: GGCCAACCAGGGCCATT
Reverse primer: ATCACCAGCCAAGGAAAGGG
LEF1
Forward primer: CCACCCATCCCGAGAACATC
Reverse primer: AGGCTTCACGTGCATTAGGT

AON Electroporation

K562 cells were transiently transfected with antisense oligonucleotides using the BTX Gemini X2 electroporation system. Cells were resuspended in cytoporation medium T (BTX #47-0002) at a density of $2x10^7$ cells/ml with 10 μM of AON. Thereafter, cells were electroporated at 250 V x 10 ms X 1 pulse and plated onto fresh medium (3 ml medium per 100 μL electroporated cells). At 48 hours after electroporation, cells were collected for RNA, protein and cell viability readout using Cell-titre Glo (Promega).

AON sequence:
SCR: CGGUGUGUGUAUCAUUCUCUAGUGU
EZH2 (1042): UGAAUCUUCUGUCCAAAAUCCAACAGGCAAUAUA

CRISPR/Cas9 Knockout of EZH2 in K562 Cells

Guide RNA (gRNA) sequences were cloned into a lentiCRISPR V2 vector (Addgene) at the BsmBI restriction site.

Target sequences are
gRNA1: TTATCAGAAGGAAATTTCCG
gRNA 2: TTATGATGGGAAAGTACACG

To generate lentivirus, 293T cells were transfected using Lipofectamine 3000 (Thermo Fisher Scientific) according to the manufacturer's protocol. Virus was collected at 48 and 72 hours post transfection, concentrated with Lenti-X concentrator (Takara), and titered using Lenti-X GoStix (Takara). $1x10^6$ K562 cells were spinoculated with virus (MOI of 3 and 5 μg/ml polybrene) for 1 hour at 2400 rpm and 37°C in non-tissue culture treated plates. Two days post-infection, cells were selected with 5 μg/ml puromycin for four days before starting experiments.

QUANTIFICATION AND STATISTICAL ANALYSIS

Computation of the Combination Index

The presence of synergistic or additive effects was determined following the theorem of Chou-Talalay (Chou, 2010). The resulting combination index (CI) offers quantitative definition for additive effect (CI = 1), synergism (CI < 1), and antagonism (CI > 1) in drug combinations.

RNA-Sequencing and Bioinformatics Analysis

Library preparation was performed following the TruSeq RNA Sample preparation v2 guide (Illumina). In brief, the sequenced reads were mapped to mm9 build of the mouse genome or hg19 build of the human genome using STAR version 2.4.2a. Differential expression analysis was performed using the edgeR package in R. Enriched Gene Ontology terms and KEGG pathways were identified using Metascape. Heatmaps of gene expressions (FPKM) were generated with in-house scripts with R. Alternative splicing analysis was done using rMATS version 4.0.1 with annotation versions Ensembl.NCBIM37v65 for mouse and Ensembl.GRCh37v72 for human. Significant alternative splicing events were defined at FDR=0.05 and inclusion level difference of 10%.

Replicates

RNA-Seq was conducted with 3-5 biological replicates from each group. Genetic phenotyping experiments were replicated three times independently. For *in vivo* experiments, the number of animals was chosen to ensure 90% power with 5% error based on observed standard deviation. Flow cytometric experiments were replicated independently two-three times. Pilot studies were conducted with drug studies and results were replicated in a larger study to achieve enough statistical power. *In vitro* experiments were replicated two-three times, with viability experiments being completed in triplicate.

Data Analysis of SILAC Arginine Methyl-Peptides

Acquired raw data were analysed using the integrated MaxQuant software v1.3.0.5 or v.1.5.2.8, using the Andromeda search engine (Cox and Mann, 2008). In MaxQuant, the estimated false discovery rate (FDR) of all peptide identifications was set to a maximum of 1%. The main search was performed with a mass tolerance of 7 ppm. Enzyme specificity was set to Trypsin/P. A maximum of 3 missed cleavages was permitted, and the minimum peptide length was fixed at 7 amino acids. Carbamidomethylation of Cysteine was set as a fixed modification. The January 2016 version of the Uniprot sequence was used for peptide identification.

To assign and quantify SILAC methyl-peptides, each raw file was analysed with the following set of variable modifications: N-terminal acetylation, Methionine oxidation, mono-methyl-K/R and di-methyl-K/R. The MaxQuant evidence.txt output file was then filtered as follows: potential contaminants and reverse sequences were removed; methyl-peptides not fulfilling the quality criteria of Andromeda score $\geq$ 25 and PTM localization probability $\geq$ 0.50 were also removed. For the methyl-peptides quantified more than once, the median SILAC ratio was calculated. Finally, methyl-peptide SILAC ratios were normalised on the respective protein SILAC ratios extracted from the proteinGroups.txt MaxQuant output file. These were calculated using unmodified peptides in the ''input'' experiment. To define significantly up- or down-regulated methyl-peptides by GSK591, we used mean (μ) and standard deviation (σ), based on the distribution of the unmodified peptide SILAC ratios calculated separately in the forward and reverse experiments and we applied a $\mu \pm 3\sigma$ cut-off to the distributions of the modified peptides of the respective replicate (see Table S4).

HmSEEKER: A Perl-Based Pipeline for High-Confidence Assignment of Methyl-Peptides from hmSILAC Data

To assign hmSILAC peptide sequences, we defined new modifications in MaxQuant with the mass increment and residue specificities corresponding to heavy mono-methylation (mono-methyl4-K/R) and di-methylation (di-methyl4-K/R). Additionally, we defined new modifications for heavy methionine (Met4) and oxidized heavy methionine (OxMet4). To reduce the search complexity, raw data were analysed twice with the following sets of variable modifications: (ii) N-terminal acetylation, Met4, OxMet4, oxidation, mono-methyl-K/R, mono-methyl4-K/R; (ii) N-terminal acetylation, Met4, OxMet4, oxidation, di-methyl-K/R, di-methyl4-K/R.

Identification of high confidence methyl-sites was carried out with an in-house developed, Perl-based pipeline, named hmSEEKER, which identifies doublets of heavy and light hmSILAC peptides from MaxQuant output tables (Massignani et al., 2019). hmSEEKER performs the following steps: methyl-peptides identified in the msms file are first filtered to remove: (i) all contaminants and decoy peptides, (ii) all peptides with single charge and (iii) all peptides bearing simultaneous heavy and light modifications. Then each peptide is associated to its corresponding MS1 peak in the allPeptides file. Finally, the H or L counterpart of each peak is searched among other peaks detected in the same raw data file. Because the pair is searched in msmsScans, hmSEEKER enables the identification of peptide doublets even when one of the two counterparts has not been MS/MS sequenced, thus not appearing in the msms file. We used hmSEEKER to automatically filter the MaxQuant msms.txt file and remove contaminants and reverse sequences, as well as peptides carrying simultaneously light and heavy modifications. To increase the confidence of our findings, remaining peptides were further filtered to remove any peptide with Andromeda score < 25 or Andromeda delta score < 12 and any methylation with a PTM localization probability < 0.75. Heavy and light methyl-peptide pairs were accepted when the difference between calculated and expected mass shift was < 2 ppm and the difference between their retention times was < 30 s.

Use of hmLINKER to Intersect the SILAC Methyl-proteome with the hmSILAC Dataset

Validation of the methylated peptide identified in the SILAC experiments through the hmSILAC identifications was achieved by using hmLINKER, another in-house developed bioinformatic tool that compares the sequences of the peptides in the SILAC dataset to those in the hmSILAC repository (hmSEEKER output). If a match is not found at the sequence level, the peptide is not immediately discarded, but a second round of match-attempt is performed using a 31 amino acids sequence window, centered on each modification site

Motif Analysis of R Methyl-Peptides

Motif analysis of R methyl-sites was performed using the pLogo web application (O'Shea et al., 2013), which allows the visualization of significant enrichment variations between the set of GSK591-regulated sequences and the unchanging methyl-peptides used as a background set. p value threshold was set to 0.05.

Bioinformatic Analysis of R Methyl-Peptides

Motif analysis of changing methyl-sites was performed using the pLogo web application (O'Shea et al., 2013), which allows the visualization of significant enrichment variations between the set of drug-regulated sequences and the unchanging methyl-peptides used

as a background set. P value threshold was set to 0.05. Networks on regulated methyl-proteins were built with the Reactome application of Cytoscape (Fabregat et al., 2018; Shannon et al., 2003) .The Gene Ontology analysis was carried out with GOrilla (Eden et al., 2009) and Revigo (Supek et al., 2011).

Calculation of the Protein Abundance Index in the Annotated Proteome

We matched the methylated proteins identified via the immuno-enrichment approach with their respective protein, annotated in the whole proteome, which was used ad INPUT for the methyl-peptides IP. To calculate the abundance index of each protein in this proteome, we employed the iBAQ algorithm (Intensity-Based Absolute Quantification (Schwanhausser et al., 2011), embedded in the MaxQuant suit.

Statistical Analyses

Statistical significance was determined by unpaired Student's t-test after testing for normal distribution unless indicated otherwise. The Mantel-Cox log-rank test was used to compare survival curves. P values of < 0.05 were considered statistically significant. Data were plotted using GraphPad Prism 7 software as mean values and error bars represent standard deviation. Asterisks indicate * p = 0.01-0.05; ** p = 0.001-0.01, *** p = 0.0001-0.001, **** p <0.0001.

DATA AND CODE AVAILABILITY

The accession number for the RNA-seq data reported in this paper is GEO:GSE123774.

The accession number for the MS-proteomics data have been deposited in the ProteomeXchange Consortium via the PRIDE (Vizcaino et al., 2016) partner repository with the dataset identifier PRIDE: PXD012007.

LETTER

https://doi.org/10.1038/s41586-019-1618-0

Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis

Akihide Yoshimi[1,12], Kuan-Ting Lin[2,12], Daniel H. Wiseman[3,4,12], Mohammad Alinoor Rahman[2], Alessandro Pastore[1], Bo Wang[1], Stanley Chun-Wei Lee[1], Jean-Baptiste Micol[5], Xiao Jing Zhang[1], Stephane de Botton[5], Virginie Penard-Lacronique[5], Eytan M. Stein[6], Hana Cho[1], Rachel E. Miles[1], Daichi Inoue[1], Todd R. Albrecht[7], Tim C. P. Somervaille[3], Kiran Batta[4], Fabio Amaral[3], Fabrizio Simeoni[3], Deepti P. Wilks[8], Catherine Cargo[9], Andrew M. Intlekofer[1], Ross L. Levine[1,6], Heidi Dvinge[10], Robert K. Bradley[11], Eric J. Wagner[7], Adrian R. Krainer[2] & Omar Abdel-Wahab[1,6]*

Transcription and pre-mRNA splicing are key steps in the control of gene expression and mutations in genes regulating each of these processes are common in leukaemia[1,2]. Despite the frequent overlap of mutations affecting epigenetic regulation and splicing in leukaemia, how these processes influence one another to promote leukaemogenesis is not understood and, to our knowledge, there is no functional evidence that mutations in RNA splicing factors initiate leukaemia. Here, through analyses of transcriptomes from 982 patients with acute myeloid leukaemia, we identified frequent overlap of mutations in *IDH2* and *SRSF2* that together promote leukaemogenesis through coordinated effects on the epigenome and RNA splicing. Whereas mutations in either *IDH2* or *SRSF2* imparted distinct splicing changes, co-expression of mutant IDH2 altered the splicing effects of mutant SRSF2 and resulted in more profound splicing changes than either mutation alone. Consistent with this, co-expression of mutant IDH2 and SRSF2 resulted in lethal myelodysplasia with proliferative features in vivo and enhanced self-renewal in a manner not observed with either mutation alone. *IDH2* and *SRSF2* double-mutant cells exhibited aberrant splicing and reduced expression of *INTS3*, a member of the integrator complex[3], concordant with increased stalling of RNA polymerase II (RNAPII). Aberrant *INTS3* splicing contributed to leukaemogenesis in concert with mutant IDH2 and was dependent on mutant SRSF2 binding to *cis* elements in *INTS3* mRNA and increased DNA methylation of *INTS3*. These data identify a pathogenic crosstalk between altered epigenetic state and splicing in a subset of leukaemias, provide functional evidence that mutations in splicing factors drive myeloid malignancy development, and identify spliceosomal changes as a mediator of *IDH2*-mutant leukaemogenesis.

Mutations in RNA splicing factors are common in cancer and impart specific changes to splicing that are identifiable by mRNA sequencing (RNA-seq)[4–6]. Somatic mutations involving the proline 95 residue of the spliceosome component SRSF2 are among the most recurrent in myeloid malignancies and alter SRSF2 binding to RNA in a sequence-specific manner[6,7]. We analysed RNA-seq data in The Cancer Genome Atlas (TCGA)[1] from 179 patients with acute myeloid leukaemia (AML), evaluating them for spliceosomal alterations. Aberrant splicing events characteristic of *SRSF2* mutations, including *EZH2*[6,7] poison exon inclusion, were observed in 19 patients ($P = 1.6 \times 10^{-12}$; Fisher's exact test; Fig. 1a, Extended Data Fig. 1a, b, Supplementary Table 1). Although only one patient with a mutation in *SRSF2* was reported in the TCGA AML publication[1], mutational analysis of RNA-seq data identified *SRSF2* hotspot mutations in each of these 19 patients (11% of the patients with AML). Therefore, these data retrospectively identify *SRSF2* as one of the most commonly mutated genes in the TCGA AML cohort.

Notably, 47% of patients with mutated *SRSF2* also had a mutation in *IDH2* and conversely, 56% of patients with mutated *IDH2* had a mutation in *SRSF2* ($P = 1.7 \times 10^{-6}$; Fisher's exact test; Fig. 1b, Extended Data Fig. 1c, d, Supplementary Table 2). Similar results were seen in RNA-seq data from 498 and 263 patients with AML from the Beat AML[8] and Leucegene[9] studies, respectively (Fig. 1c, d, Extended Data Fig. 1e–j, Supplementary Table 2). Across these datasets, variant allele frequencies of *IDH2* and *SRSF2* mutations were high and significantly correlated (Extended Data Fig. 1k), suggesting their common placement as early events in AML.

Beyond these datasets, combined *IDH2* and *SRSF2* mutations were identified in 5.2–6.2% of 1,643 unselected consecutive patients with AML in clinical practice (Supplementary Table 3). Although not statistically significant, *IDH2* and *SRSF2* double-mutant AML cases had the shortest overall survival across the four studied genotypes (Extended Data Fig. 2a). Whereas patients with *IDH2* and *SRSF2* double mutations had mostly intermediate cytogenetic risk, their prognosis was comparable to those with adverse cytogenetic risk (Extended Data Fig. 2b). The patients with *IDH2* and *SRSF2* double mutations were also significantly older than those with mutations in *IDH2* only, or with wild-type *IDH2* and *SRSF2* (Extended Data Fig. 2d; clinical and genetic features are summarized in Extended Data Fig. 2 and Supplementary Table 3).

Mutations in *IDH2* confer neomorphic enzymatic activity that results in the generation of 2-hydroxyglutarate (2HG)[10], which in turn induces DNA hypermethylation via the competitive inhibition of the α-ketoglutarate-dependent enzymes TET1–TET3. Unsupervised hierarchical clustering of DNA methylation data from the TCGA AML cohort revealed that *IDH2* and *SRSF2* double-mutant AML cases form a distinct cluster with higher DNA methylation than *IDH2* single-mutant AML cases (Extended Data Fig. 1l-o). Collectively, these data identify *IDH2* and *SRSF2* double-mutant leukaemia as a recurrent genetically defined AML subset with a distinct epigenomic profile.

We next sought to understand the basis for co-enrichment of *IDH2* and *SRSF2* mutations. Although mutations in splicing factors are frequently found in leukaemias, there is no functional evidence that they can transform cells in vivo. Overexpression of human $IDH2^{R140Q}$ or $IDH2^{R172K}$ in bone marrow (BM) cells from *Vav-cre* $Srsf2^{P95H/+}$ or *Vav-cre* $Srsf2^{+/+}$ mice revealed a clear collaborative effect between mutant *IDH2* and *Srsf2* (Extended Data Fig. 3a). Four weeks after transplantation, the peripheral blood of recipient mice transplanted with *IDH2* and

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. [3]Leukaemia Biology Laboratory, Cancer Research UK Manchester Institute, The University of Manchester, Manchester, UK. [4]Division of Cancer Sciences, The University of Manchester, Manchester, UK. [5]Gustave Roussy, Université Paris-Saclay, Villejuif, France. [6]Leukemia Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [7]Department of Biochemistry & Molecular Biology, The University of Texas Medical Branch at Galveston, Galveston, Texas, USA. [8]Manchester Cancer Research Centre Biobank, The University of Manchester, Manchester, UK. [9]Haematological Malignancy Diagnostic Service, St James's University Hospital, Leeds, UK. [10]Department of Biomolecular Chemistry, School of Medicine and Public Health, University of Wisconsin-Madison, Madison, WI, USA. [11]Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [12]These authors contributed equally: Akihide Yoshimi, Kuan-Ting Lin, Daniel H. Wiseman. *e-mail: abdelwao@mskcc.org

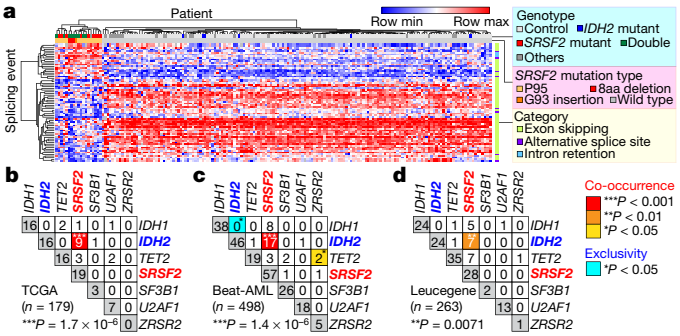


Fig. 1 | Frequently co-occurring *IDH2* and *SRSF2* mutations in AML. **a**, Heat map of per cent-spliced-in values for mutant *SRSF2*-specific splicing events in TCGA AML samples. **b**–**d**, Co-occurrence of mutations in *IDH1*, *IDH2*, *TET2* and RNA-splicing factors in the TCGA (**b**), Beat AML (**c**) and Leucegene (**d**) cohorts. Number of patients indicated; co-occurrence or exclusivity noted by colour coding; two-sided Fisher's exact test. Double refers to *SRSF2* and *IDH2* double mutant throughout.

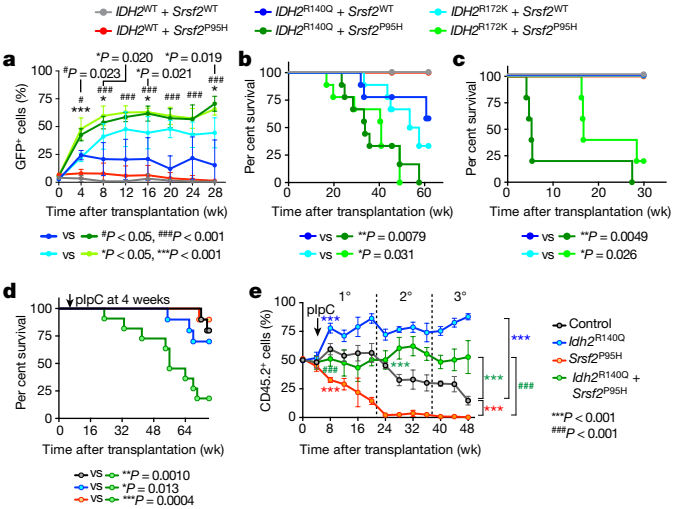


Fig. 2 | Mutant *IDH2* cooperates with mutant *Srsf2* to promote leukaemogenesis. **a**, Chimerism of GFP$^+$ cells in the blood of recipients transplanted with BM cells with indicated genotypes over time ($n = 5$ per group; data at 0 week represent transduction efficiency; mean percentage ± s.d.; two-way analysis of variance (ANOVA) with Tukey's multiple comparison test). **b**–**d**, Kaplan–Meier survival analysis of primary recipients (**b**) ($n = 10$ mice per genotype), recipients of serial transplant (**c**) ($n = 5$) and primary recipients transplanted non-competitively with BM cells from knock-in mice (**d**) ($n = 10$). log-rank Mantel–Cox test (two-sided). **e**, Chimerism of peripheral blood CD45.2$^+$ cells in competitive transplantation. $n = 10$ mice per group; mean ± s.d.; two-way ANOVA with Tukey's multiple comparison test.

Srsf2 double-mutant cells had a substantially higher percentage of GFP$^+$ cells than in an *Srsf2* wild-type background (Fig. 2a, Extended Data Fig. 3b, c). Moreover, these mice exhibited significant myeloid skewing, macrocytic anaemia and thrombocytopaenia of greater magnitude than seen with mutant *IDH2* (Extended Data Fig. 3d–h). *IDH2* and *Srsf2* double mutants showed no difference in plasma 2HG levels from *IDH2* single mutants (Extended Data Fig. 3i, j). Serial replating of BM cells from leukaemic mice revealed markedly enhanced clonogenicity of *IDH2* and *Srsf2* double-mutant cells compared with other genotypes; the *IDH2* and *Srsf2* cells exhibited a blastic morphology and immature immunophenotype (Extended Data Fig. 3k–m). Consistent with these in vitro results, mice transplanted with *IDH2* and *Srsf2* double-mutant cells developed a lethal myelodysplastic syndrome (MDS) characterized by pancytopaenia, macrocytosis, myeloid dysplasia, expansion of immature BM progenitors and splenomegaly (Fig. 2b, Extended Data Fig. 3n–w). The *IDH2* and *Srsf2* double-mutant cells were also serially transplantable in sublethally irradiated recipients (Fig. 2c, Extended Data Fig. 3x). By contrast, *IDH2* single-mutant controls developed leukocytosis, myeloid skewing without clear dysplasia and less pronounced splenomegaly, whereas *Srsf2* single-mutant cells exhibited impaired repopulation capacity. These results provide evidence that spliceosomal gene mutations can promote leukaemogenesis in vivo.

We next sought to verify the effects of mutant *Idh2* and *Srsf2* using models in which both mutants were expressed from endogenous loci. *Mx1-cre Srsf2*$^{P95H/+}$ mice were crossed with *Idh2*$^{R140Q/+}$ mice to generate control, *Idh2*^{R140Q} single-mutant, *Srsf2*^{P95H} single-mutant and *Idh2* and *Srsf2* double knock-in (DKI) mice (Extended Data Fig. 4a). As expected, 2HG levels in peripheral blood mononuclear cells were increased and 5-hydroxymethylcytosine levels in KIT$^+$ BM cells were decreased in *Idh2* single-mutant and DKI primary mice compared with controls (Extended Data Fig. 4b, c). We next performed non-competitive transplantation, in which each mutation was induced alone or together following stable engraftment in recipients. DKI mice showed stable engraftment over time, similar to *Idh2* single-mutant or control mice (Extended Data Fig. 4d). However, DKI mice developed a lethal MDS with proliferative features and significantly shorter survival compared to controls (Fig. 2d). In competitive transplantation, expression of mutant *Idh2*^{R140Q} rescued the impaired self-renewal capacity of *Srsf2* single-mutant cells (Fig. 2e). These observations were supported by an increase in haematopoietic stem–progenitor cells in DKI mice compared with *Srsf2* single-mutant or control mice in primary and serial transplantation (Extended Data Fig. 4e–i). These results confirm cooperativity between mutant IDH2 and SRSF2 in promoting leukaemogenesis in vivo.

On the basis of data identifying 2HG-mediated inhibition of TET2 as a mechanism of *IDH2* mutant leukaemogenesis[11], we also evaluated whether loss of *TET2* might promote transformation of *SRSF2* mutant cells. However, deletion of *Tet2* in an *Srsf2* mutant background was insufficient to rescue the impaired self-renewal capacity of *Srsf2* single-mutant cells (Extended Data Fig. 4j–n). Similarly, restoration of TET2 function did not affect the self-renewal capacity of *Idh2* and *Srsf2* double-mutant cells in vivo (Extended Data Fig. 4o–q). These data indicated that the collaborative effects of mutant *Idh2* and *Srsf2* are not solely dependent on TET2. Consistent with this, combined silencing of *Tet2* and *Tet3* partially rescued the impaired replating capacity of *Srsf2* mutant cells in vitro (Extended Data Fig. 4r, s) and the impaired self-renewal of *Srsf2* mutant cells in vivo (Extended Data Fig. 4t–v). Because FTO and ALKBH5—which have roles in RNA processing as *N*6-methyladenosine (m^{6}A) RNA demethylases[12,13]—are also dependent on α-ketoglutarate, we investigated the effects of their loss on cooperativity with mutant *Srsf2*. However, collaborative effects were not observed between loss of *Fto* or *Alkbh5* and *Srsf2*^{P95H} (Extended Data Fig. 4w, x).

To understand the basis for cooperation between *IDH2* and *SRSF2* mutations, we next analysed RNA-seq data from the TCGA ($n = 179$ patients), Beat AML ($n = 498$ patients) and Leucegene ($n = 263$ patients) cohorts as well as two previously unpublished RNA-seq datasets targeting defined *IDH2* and *SRSF2* genotype combinations ($n = 42$ patients) and the knock-in mouse models. This revealed that cells with mutations in both *IDH2* and *SRSF2* consistently contained more aberrant splicing events than cells with mutations in *SRSF2* only. Moreover, *IDH2* mutations were associated with a small but reproducible change in RNA splicing (Fig. 3a, b, Extended Data Fig. 5a–g, Supplementary Tables 4–20). By contrast, AML cases in which both *TET2* and *SRSF2* were mutated had fewer changes in splicing than those in which *IDH2* and *SRSF2* were mutated (Extended Data Fig. 5h–m, Supplementary Tables 21, 22).

The majority of splicing changes associated with *SRSF2* mutations involved altered cassette-exon splicing, consistent with *SRSF2* mutations promoting inclusion of C-rich RNA sequences[6,7]. The sequence specificity of mutant SRSF2 on splicing was not influenced by concomitant *IDH2* mutations (Extended Data Fig. 5n–q) and a number of these events were validated by PCR with reverse transcription

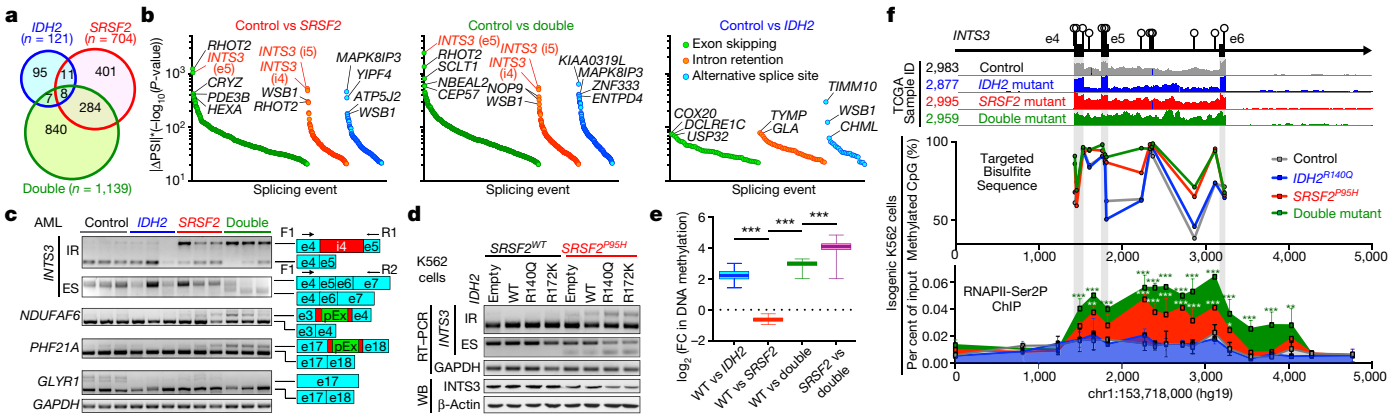


Fig. 3 | Collaborative effects of mutant *IDH2* and *SRSF2* on aberrant splicing. a, Venn diagram showing numbers of differentially spliced events from TCGA AML samples. **b**, Number of differentially spliced events ($|\Delta PSI| > 10\%$ and $P < 0.01$) in indicated genotypes are ranked by $((|\Delta PSI| \times |(-\log(P))$ according to class of event (e5, exon 5; i4 and i5, intron 4 and 5, respectively). PSI and *P* values adjusted for multiple comparisons were calculated using PSI-Sigma[25]. **c**, Representative RT–PCR results of aberrantly spliced transcripts in samples from patients with AML (pEx: exon with premature stop codon; $n = 3$ patients per genotype; three technical replicates with similar results). ES, exon skipping; IR, intron retention; F1, R1 and R2 represent primers used for RT–PCR. **d**, RT–PCR and western blots in isogenic K562 human leukaemia cells (representative images from three biologically independent experiments with similar results). **e**, Mean fold change (expressed in $\log_2$) in DNA cytosine methylation (*y* axis) at regions of genomic DNA encoding mRNA that undergo differential splicing (*x* axis). DNA methylation levels were determined by enhanced reduced representation bisulfite sequencing (eRRBS). $n = 3$ per genotype; the line represents mean, box edges represent the 25th and 75th percentiles and whiskers show 2.5th and 97.5th percentiles; one-way ANOVA with Tukey's multiple comparison test; $^{***}P < 2.2 \times 10^{-16}$. **f**, Genomic locus of *INTS3* around exons 4–6 with CpG dinucleotides indicated (top), representative RNA-seq from four patients with AML (top; $n = 1$ per genotype), results of targeted bisulfite sequencing (middle; $n = 1$ per genotype) and results of ChIP-walking experiments targeting RNAPII phosphorylated on Ser2 (Ser2P) (bottom; $n = 3$; mean ± s.d.; two-way ANOVA with Tukey's multiple comparison test compared with control.). Double mutant, $IDH2^{R140Q}SRSF2^{P95H}$. $^{*}P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$.

(RT–PCR) of primary AML samples from an independent cohort (Fig. 3c). Among the mis-splicing events in AML with mutations in both *IDH2* and *SRSF2* was a complex event in *INTS3* involving intron retention across two contiguous introns and skipping of the intervening exon (Fig. 3b, c, Extended Data Figs. 5e–g, r–y, 6a–c). Aberrant *INTS3* splicing was demonstrated in isogenic and non-isogenic leukaemia cells with or without *IDH2* and/or *SRSF2* mutations (Fig. 3d, Extended Data Fig. 6d–f), and *INTS3* transcripts with both intron retention and exon skipping resulted in nonsense-mediated decay (Extended Data Fig. 6g–j). Consistent with these observations, INTS3 protein expression was reduced in *SRSF2* mutant cells (Fig. 3d, Extended Data Fig. 6e, f, k–n, Supplementary Table 23). Moreover, silencing of INTS3 was associated with reduced protein levels of additional integrator subunits in *SRSF2* mutant AML compared to *SRSF2* wild-type AML. Consistent with these observations, steady-state protein expression levels of integrator subunits were correlated with one another (Extended Data Fig. 6o). Overall, these data indicate that aberrant splicing and consequent loss of INTS3 was a consistent feature of *IDH2* and *SRSF2* double-mutant cells and was associated with reduced expression of multiple integrator subunits.

We next sought to understand how *IDH2* mutations, which affect the epigenome, might influence splicing catalysis. Splice-site choice is influenced by *cis*-regulatory elements engaged by RNA-binding proteins as well as RNAPII elongation, which is regulated by DNA cytosine methylation and histone modifications[14]. We therefore generated a controlled system to dissect the contribution of RNA-binding elements and DNA methylation to *INTS3* intron retention. We constructed a minigene of *INTS3* spanning exons 4 and 5 and the intervening intron 4 (Extended Data Fig. 7a–c). Transfection of this minigene into leukaemia cells containing combinations of *IDH2* and *SRSF2* mutations revealed that retention of *INTS3* intron 4 is driven by mutant SRSF2 and further enhanced in the *IDH2* and *SRSF2* double-mutant setting (Extended Data Fig. 7d). SRSF2 normally binds C- or G-rich motif sequences in RNA equally well to promote splicing[15]. Leukaemia-associated mutations in *SRSF2* promote its avidity for C-rich sequences while reducing the ability to recognize G-rich sequences[6,7]. Of note, exon 4 of *INTS3* contains the highest number of predicted SRSF2-binding motifs over the entire *INTS3* genomic region (Extended Data Fig. 7c). We evaluated the role of putative SRSF2 motifs in regulating *INTS3* splicing by mutating all six CCNG motifs in exon 4 to G-rich sequences. In this G-rich version of the minigene, intron retention no longer occurred (*INTS3*-GGNG) (Extended Data Fig. 7e). Conversely, when all G-rich SRSF2 motifs were converted to C-rich sequences (*INTS3*-CCNG), intron retention became evident (Extended Data Fig. 7f). These results confirmed the sequence-specific activity of mutant SRSF2 in *INTS3* intron retention and identified a role for mutant IDH2 in regulating splicing.

Because *IDH2* mutations promote increased DNA methylation and DNA methylation can affect splicing[14], we generated genome-wide maps of DNA cytosine methylation from patients with AML across four genotypes (Supplementary Table 23). This revealed that differentially spliced events in *IDH2* single-mutant as well as *IDH2* and *SRSF2* double-mutant AML (compared to *IDH2* and *SRSF2* wild-type and *SRSF2* single-mutant AML) contained significant hypermethylation of DNA. Thus regions of differential DNA hypermethylation significantly overlapped with regions of differential RNA splicing (Fig. 3e, Extended Data Fig. 7j).

The above results suggest a strong link between increased DNA methylation mediated by mutant IDH2 and altered RNA splicing by mutant SRSF2. To evaluate this further, we next examined DNA methylation levels around endogenous *INTS3* exons 4–6 by targeted bisulfite sequencing. This revealed increased DNA methylation at all CpG dinucleotides in this region in *IDH2* and *SRSF2* double-mutant cells compared to control or single-mutant cells (Fig. 3f, Extended Data Fig. 7k). A functional role of DNA methylation at these sites was verified by evaluating splicing in versions of the *INTS3* minigene in which each CG dinucleotide was converted to an AT to prevent cytosine methylation. In these CG-to-AT versions of the minigene, *IDH2* mutations no longer promoted mutant-*SRSF2*-mediated intron retention (Extended Data Fig. 7g–i). As further confirmation of the influence of mutant *IDH2* on *INTS3* splicing, cell-permeable 2HG increased *INTS3* intron retention whereas treatment of *IDH2* and *SRSF2* double-mutant cells with the DNA methyltransferase inhibitor 5-aza-2′-deoxycytidine (5-AZA-CdR) inhibited *INTS3* intron retention (Extended Data Fig. 7l, m).

Given that changes in epigenetic state may affect splicing by influencing RNAPII stalling[14,16], we evaluated the abundance of RNAPII

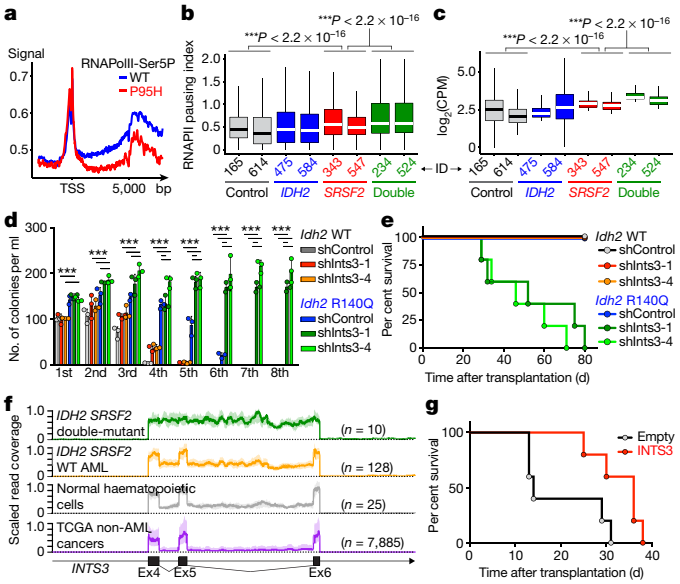


Fig. 4 | RNAPII stalling in *IDH2* and *SRSF2* double-mutant AML and contribution of INTS3 loss to leukaemogenesis. a, Metagene plot of genome-wide RNAPII Ser5 phosphorylation (Ser5P) in isogenic $SRSF2^{WT}$ or $SRSF2^{P95H}$ mutant K562 cells. **b**, **c**, RNAPII pausing index[20] in primary AML samples calculated as the ratio of normalized ChIP–seq reads of RNAPII Ser5P on transcription start sites (TSSs) (± 250 bp) over that of the corresponding bodies (+500 bp to +1000 bp from TSSs) (**b**) and RNAPII abundance over the differentially spliced regions between *SRSF2* single-mutant and *IDH2* and *SRSF2* double-mutant AML determined by RNAPII Ser2P ChIP–seq (*y* axis shows $\log_2$(counts per million (CPM))) (**c**). *x* axis shows patient ID. Box plots were generated from ChIP–seq data from an individual primary AML sample; the line represents mean, box edges represent 25th and 75th percentiles and whiskers show 2.5th and 97.5th percentiles. One-way ANOVA with Tukey's multiple comparison test. **d**, Colony numbers from serial replating assays of either *Mx1-cre* $Idh2^{+/+}$ *or* $Idh2^{R140Q/+}$ BM cells transduced with shRNA targeting *Ints3* (shInts3-1 or shInts3-4) or control shRNA (shControl). $n = 3$ biologically independent experiments; mean ± s.d.; two-way ANOVA with Tukey's multiple comparison test. $^{***}P < 0.004$. **e**, **g**, Kaplan–Meier survival analysis of recipients. $n = 5$ per group; log-rank (Mantel–Cox) test (two-sided). In **e**: shInts3-1 versus shControl, $P = 0.0018$; shInts3-4 versus shControl, $P = 0.0018$. In **g**, $P = 0.034$ (INTS3 versus empty vector control). **f**, RNA-seq read coverage between exons 4–6 of *INTS3* and 1,000 bp either side of *INTS3* is scaled and shown as mean (line) ± s.d. (shaded region) (generated from TCGA datasets; sample list shown in legend for Extended Data Fig. 10e).

using chromatin precipitation with DNA sequencing (ChIP–seq) in isogenic $SRSF2^{WT}$ and $SRSF2^{P95H}$ cells as well as the primary samples from patients with AML. This revealed increased promoter-proximal transcriptional pausing and decreased RNAPII occupancy over gene bodies in *SRSF2* mutant cells, which was further enhanced in *IDH2* and *SRSF2* double-mutant cells (Fig. 4a, b, Extended Data Fig. 7n–q, Supplementary Table 23). Transcriptional pausing was also evident at *INTS5* and *INTS14* in *SRSF2* mutant cells (Extended Data Fig. 7r, s), which—in combination with aberrant splicing of several integrator subunits (Supplementary Table 24)—suggested impaired function of the entire integrator complex in *SRSF2* mutant cells. Similar to DNA cytosine methylation levels, RNAPII was more abundant over differentially spliced regions in *SRSF2* single-mutant AML than in *SRSF2* wild-type AML, and further enhanced over differentially spliced regions in *IDH2* and *SRSF2* double-mutant AML compared with those in *SRSF2* single-mutant AML (Fig. 4c, Extended Data Fig. 7t).

The above data provide further links between increased DNA cytosine methylation and RNAPII stalling with altered RNA splicing in *IDH2* and *SRSF2* double-mutant AML. To further evaluate this model, we performed ChIP for RNAPII across 4,766 bp of the *INTS3* locus in isogenic leukaemia cells (Fig. 3f). This revealed substantial accumulation of RNAPII across this locus in *IDH2* and *SRSF2* double-mutant cells. Treatment with 5-AZA-CdR significantly reduced RNAPII stalling, which was coupled with decreased aberrant *INTS3* splicing (Extended Data Fig. 7k–m). These data reveal that *IDH2* and *SRSF2* mutations coordinately dysregulate splicing through alterations in RNAPII stalling in addition to aberrant sequence recognition of *cis* elements in RNA.

INTS3 encodes a component of the integrator complex that participates in small nuclear RNA (snRNA) processing[3] in addition to RNAPII pause–release[17]. Consistent with this, *SRSF2* single-mutant cells had altered snRNA cleavage similar to those seen with direct INTS3 downregulation, which was exacerbated in *IDH2* and *SRSF2* double-mutant cells (Extended Data Fig. 8a–h). Attenuation of INTS3 expression in *SRSF2* mutant cells caused a blockade of myeloid differentiation, an effect further enhanced in an *IDH2* mutant background (Extended Data Fig. 8i–n). Notably, direct *Ints3* downregulation in the $Idh2^{R140Q/+}$ background resulted in enhanced clonogenic capacity of cells with an immature morphology and immunophenotype (Fig. 4d, Extended Data Fig. 8o–r) and promoted clonal dominance of *Idh2* mutant cells (Extended Data Fig. 9a–d). Moreover, mice transplanted with $Idh2^{R140Q/+}$ BM cells treated with short hairpin RNA (shRNA) targeting *Ints3* exhibited myeloid skewing, anaemia and thrombocytopaenia (Extended Data Fig. 9e–g) and developed a lethal MDS with proliferative features—phenotypes resembling those seen in *IDH2* and *Srsf2* double-mutant mice (Fig. 4e, Extended Data Fig. 9h, i).

The defects in snRNA processing in *SRSF2* single-mutant and *IDH2* and *SRSF2* double-mutant cells were partially rescued by INTS3 cDNA expression (Extended Data Fig. 8s–x). In addition, restoration of INTS3 expression released *SRSF2* single-mutant and *IDH2* and *SRSF2* double-mutant HL-60 cells from differentiation block (Extended Data Fig. 8y, z). Xenografts of *IDH2* and *SRSF2* double-mutant HL-60 cells demonstrated that forced expression of INTS3 induced myeloid differentiation and slowed leukaemia progression in vivo (Extended Data Fig. 9j-s). Collectively, these data suggest that INTS3 loss due to aberrant splicing by mutant IDH2 and SRSF2 contributes to leukaemogenesis.

Although loss of INTS3 resulted in measurable changes in snRNA processing, the degree of snRNA mis-processing did not have a substantial effect on splicing as determined by RNA-seq of $IDH2^{R140Q}$ mutant HL-60 cells with INTS3 silencing. By contrast, INTS3 depletion in these cells significantly affected transcriptional programs associated with myeloid differentiation, multiple oncogenic signalling pathways, RNAPII elongation-linked transcription and DNA repair (Extended Data Fig. 10a–d, Supplementary Table 25). This latter association of INTS3 loss with DNA repair is potentially consistent with previous reports that sensor of single-stranded DNA complexes containing INTS3 participate in DNA damage response[18,19].

These data uncover an important role for RNA splicing alterations in *IDH2* mutant tumorigenesis and identify perturbations in integrator as a driver of transformation of *IDH2* and *SRSF2* mutant cells. However, INTS3 is not known to be recurrently affected by coding-region alterations in leukaemias. We therefore evaluated *INTS3* splicing across 32 additional cancer types as well as normal blood cells to evaluate whether aberrant *INTS3* splicing might be a common mechanism in AML. This revealed that, whereas *INTS3* mis-splicing is most evident in *IDH2* and *SRSF2* double-mutant AML, *INTS3* mis-splicing is also prevalent across other molecular subtypes of AML but is not present in blood cells from healthy subjects or RNA-seq data from more than 7,000 samples from other cancer types (Fig. 4f, Extended Data Fig. 10e, f). To further evaluate the effects of enforced INTS3 expression in myeloid leukaemia with a wild-type splicing phenotype, we used *MLL-AF9;*$Nras^{G12D}$ mouse leukaemia (RN2) cells. INTS3 overexpression reduced colony-forming capacity (Extended Data Fig. 10g, h) and enhanced differentiation of RN2 cells, resulting in decelerated leukaemia progression in vivo (Fig. 4g, Extended Data Fig. 10i–s).

These data highlight a role for loss of INTS3 in broad genetic subtypes of AML. Further efforts to determine how integrator loss

promotes leukaemogenesis and other non-mutational mechanisms mediating *INTS3* aberrant splicing will be critical for understanding and targeting leukaemias with integrator loss. Previous studies have identified that integrator[17,20] and SRSF2[21] have direct roles in modulating transcriptional pause–release. The accumulation of RNAPII at certain mis-spliced loci in this study is consistent with recent data that suggest that mutant SRSF2 is defective in promoting RNAPII pause–release[22]. Identifying how aberrant splicing mediated by mutant SRSF2 is influenced by altered RNAPII pause–release may therefore be informative.

In addition to modifying splicing in *SRSF2* mutant cells, *IDH2* mutations were associated with reproducible changes in splicing in haematopoietic cells. There is a strong correlation between aberrant splicing in *IDH2* and *IDH1* mutant low-grade gliomas ($P = 2.2 \times 10^{-16}$ (binominal proportion test); Extended Data Fig. 10t–w, Supplementary Tables 26–28). A significant number of splicing events that were dysregulated in *IDH2* mutant AML from the TCGA and Leucegene cohorts were differentially spliced in *IDH2* mutants versus *IDH1* and *IDH2* wild-type low-grade gliomas ($P = 1.8 \times 10^{-9}$ and $P = 1.3 \times 10^{-8}$, respectively; binominal proportion test). These data suggest that *IDH1* and *IDH2* mutations impart a consistent effect on splicing regardless of tumour type. Finally, these results have important translational implications given the substantial efforts to pharmacologically inhibit mutant IDH1 and IDH2 as well as mutant splicing factors[23,24]. The frequent coexistence of *IDH2* and *SRSF2* mutations underscores the enormous therapeutic potential for modulation of splicing in the approximately 50% of patients with *IDH2* mutant leukaemia who also have a spliceosomal gene mutation.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1618-0.

Received: 24 May 2018; Accepted: 28 August 2019;
Published online 2 October 2019.

1. The Cancer Genome Atlas Research Network. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. *N. Engl. J. Med.* **368**, 2059–2074 (2013).
2. Papaemmanuil, E. et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616–3627, quiz 3699 (2013).
3. Wu, Y., Albrecht, T. R., Baillat, D., Wagner, E. J. & Tong, L. Molecular basis for the interaction between Integrator subunits IntS9 and IntS11 and its functional importance. *Proc. Natl Acad. Sci. USA* **114**, 4394–4399 (2017).
4. 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).
5. Ilagan, J. O. et al. *U2AF1* mutations alter splice site recognition in hematological malignancies. *Genome Res.* **25**, 14–26 (2015).
6. Kim, E. et al. *SRSF2* mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* **27**, 617–630 (2015).
7. 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).
8. Tyner, J. W. et al. Functional genomic landscape of acute myeloid leukaemia. *Nature* **562**, 526–531 (2018).
9. Lavallée, V. P. et al. The transcriptomic landscape and directed chemical interrogation of *MLL*-rearranged acute myeloid leukemias. *Nat. Genet.* **47**, 1030–1037 (2015).
10. Dang, L. et al. Cancer-associated IDH1 mutations produce 2-hydroxyglutarate. *Nature* **462**, 739–744 (2009).
11. Figueroa, M. E. et al. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553–567 (2010).
12. Jia, G. et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. *Nat. Chem. Biol.* **7**, 885–887 (2011).
13. Zheng, G. et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol. Cell* **49**, 18–29 (2013).
14. Naftelberg, S., Schor, I. E., Ast, G. & Kornblihtt, A. R. Regulation of alternative splicing through coupling with transcription and chromatin structure. *Annu. Rev. Biochem.* **84**, 165–198 (2015).
15. Daubner, G. M., Cléry, A., Jayne, S., Stevenin, J. & Allain, F. H. A syn-anti conformational difference allows SRSF2 to recognize guanines and cytosines equally well. *EMBO J.* **31**, 162–174 (2012).
16. Shukla, S. et al. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* **479**, 74–79 (2011).
17. Gardini, A. et al. Integrator regulates transcriptional initiation and pause release following activation. *Mol. Cell* **56**, 128–139 (2014).
18. Huang, J., Gong, Z., Ghosal, G. & Chen, J. SOSS complexes participate in the maintenance of genomic stability. *Mol. Cell* **35**, 384–393 (2009).
19. Li, Y. et al. HSSB1 and hSSB2 form similar multiprotein complexes that participate in DNA damage response. *J. Biol. Chem.* **284**, 23525–23531 (2009).
20. Stadelmayer, B. et al. Integrator complex regulates NELF-mediated RNA polymerase II pause/release and processivity at coding genes. *Nat. Commun.* **5**, 5531 (2014).
21. Ji, X. et al. SR proteins collaborate with 7SK and promoter-associated nascent RNA to release paused polymerase. *Cell* **153**, 855–868 (2013).
22. Chen, L. et al. The augmented R-loop is a unifying mechanism for myelodysplastic syndromes induced by high-risk splicing factor mutations. *Mol. Cell* **69**, 412–425 (2018).
23. Seiler, M. et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **24**, 497–504 (2018).
24. Stein, E. M. et al. Enasidenib in mutant *IDH2* relapsed or refractory acute myeloid leukemia. *Blood* **130**, 722–731 (2017).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

Data reporting. The number of mice in each experiment was chosen to provide 90% statistical power with a 5% error level. Otherwise, no statistical methods were used to predetermine sample size. The experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

Mice. All mice were housed at Memorial Sloan Kettering Cancer Center (MSK). All mouse procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSK. Six- to eight-week-old week female CD45.1 C57BL/6 mice were purchased from The Jackson Laboratory (Stock No: 002014) . Male and female CD45.2 *Srsf2*$^{P95H/+}$ conditional knock-in mice, *Idh2*$^{R140Q/+}$ conditional knock-in mice, and *Tet2* conditional knockout mice (all on C57BL/6 background) were also analysed and used as bone marrow donors (generation of these mice were as described[6,26,27]). For BM transplantation assays with IDH2 overexpression, *Srsf2*$^{P95H/+}$ and littermate control mice were crossed to *Vav-cre* transgenic mice[28]. CBC analysis was performed on peripheral blood collected from submandibular bleeding, using a Procyte Dx Hematology Analyzer (IDEXX Veterinary Diagnostics). For all mouse experiments, the mice were monitored closely for signs of disease or morbidity daily and were euthanized for visible tumour formation at tumour volume >1 cm^3, failure to thrive, weight loss >10% total body weight, open skin lesions, bleeding, or any signs of infection. In none of the experiments were these limits exceeded.

BM transplantation assays. Freshly dissected femurs and tibias were isolated from *Mx1-cre*, *Mx1-cre/Idh2*$^{R140Q/+}$, *Mx1-cre Srsf2*$^{P95H/+}$, *Mx1-cre Idh2*$^{R140Q/+}$*Srsf2*$^{P95H/+}$, *Mx1-cre Tet2*$^{fl/fl}$, or *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$ CD45.2$^+$ mice. BM was flushed with a 3-cm^3 insulin syringe into cold PBS supplemented with 2% bovine serum albumin to generate single-cell suspensions. BM cells were pelleted by centrifugation at 1,500 r.p.m. for 4 min and red blood cells (RBCs) were lysed in ammonium chloride-potassium bicarbonate lysis (ACK) buffer (150 mM NH_4Cl + 10 mM $KHCO_3$ + 0.1 mM EDTA; Thermo Fisher Scientific) for 3 min on ice. After centrifugation, cells were resuspended in PBS/2% BSA, passed through a 40-μm cell strainer, and counted. For competitive transplantation experiments, 0.5 × 10^6 BM cells from *Mx1-cre*, *Mx1-cre Idh2*$^{R140Q/+}$, *Mx1-cre Srsf2*$^{P95H/+}$, *Mx1-cre Idh2*$^{R140Q/+}$*Srsf2*$^{P95H/+}$, *Mx1-cre Tet2*$^{fl/fl}$ or *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$ CD45.2$^+$ mice were mixed with 0.5 × 10^6 wild-type (WT) CD45.1$^+$ BM and transplanted via tail-vein injection into 8-week-old lethally irradiated (900 cGy) CD45.1$^+$ recipient mice. The CD45.1$^+$:CD45.2$^+$ ratio was confirmed to be approximately 1:1 by flow cytometry analysis pre-transplant. To activate the conditional alleles, mice were treated with 3 doses of polyinosinic:polycytidylic acid (pIpC; 12 mg per kg (body weight) per day; GE Healthcare) every second day via intraperitoneal injection. Peripheral blood chimerism was assessed every four weeks by flow cytometry. For noncompetitive transplantation experiments, 1 × 10^6 total BM cells from *Mx1-cre*, *Mx1-cre Idh2*$^{R140Q/+}$, *Mx1-cre Srsf2*$^{P95H/+}$, *Mx1-cre Idh2*$^{R140Q/+}$*Srsf2*$^{P95H/+}$, *Mx1-cre Tet2*$^{fl/fl}$, or *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$ CD45.2$^+$ mice were injected into lethally irradiated (950 cGy) CD45.1$^+$ recipient mice. Peripheral blood chimerism was assessed as described for competitive transplantation experiments. Additionally, for each bleeding whole blood cell counts were measured on an automated blood analyser. Mice that were lost owing to pIpC toxicity were excluded from analysis.

Retroviral transduction and transplantation of primary haematopoietic cells. *Vav-cre Srsf2*$^{+/+}$ and *Vav-cre Srsf2*$^{P95H/+}$ mice were treated with a single dose of 5-fluoruracil (150 mg kg^{-1}) followed by BM collection from the femurs, tibias and pelvic bones 5 days later. RBCs were removed by ACK lysis buffer, and nucleated BM cells were transduced with viral supernatants containing MSCV-*IDH2*$^{WT/R140Q/R172K}$-IRES-GFP for 2 days in RPMI/20% FCS supplemented with mouse stem cell factor (mSCF, 25 ng ml^{-1}), mouse interleukin-3 (mIL3, 10 ng ml^{-1}) and mIL6 (10 ng ml^{-1}), followed by injection of about 0.5 × 10^6 cells per recipient mouse via tail vein injection into lethally irradiated (950 cGy) CD45.1$^+$ mice. Transplantation of primary BM cells with TET2 catalytic domain cDNA and anti-*Ints3* or *Tet3* shRNAs was similarly performed. For secondary transplantation experiments, 8-week old, lethally (900–950 cGy) or sub-lethally (450–700 cGy) irradiated C57/BL6 recipient mice were injected with unsorted 1 × 10^6 BM cells from the primary transplantation. *IDH2*^{WT} + *Srsf2*^{WT} and *IDH2*^{WT} + *Srsf2*^{P95H} mice were euthanized at day 315 post-transplant to collect BM for the serial transplantation. All cytokines were purchased from R&D Systems.

Flow cytometry analyses and antibodies. Surface-marker staining of haematopoietic cells was performed by first lysing cells with ACK lysis buffer and washing cells with ice-cold PBS. Cells were stained with antibodies in PBS/2% BSA for 30 min on ice. For haematopoietic stem/progenitor staining, cells were stained with the following antibodies: B220-APCCy7 (clone: RA3-6B2; purchased from BioLegend; catalogue no.: 103224; dilution: 1:200); B220-Bv711 (RA3-6B2; BioLegend; 103255; 1:200); CD3-PerCPCy5.5 (17A2; BioLegend; 100208; 1:200); CD3-APC (17A2; BioLegend; 100236; 1:200); CD3-APCCy7 (17A2; BioLegend; 100222; 1:200); Gr1-PECy7 (RB6-8C5; eBioscience; 25-5931-82; 1:500); CD11b-PE (M1/70; eBioscience; 12-0112-85; 1:500); CD11b-APCCy7 (M1/70; BioLegend; 101226; 1:200); CD11c-APCCy7 (N418; BioLegend; 117323; 1:200); NK1.1-APCCy7 (PK136; BioLegend; 108724; 1:200); Ter119-APCCy7 (BioLegend; 116223: 1:200); KIT-APC (2B8; BioLegend; 105812; 1:200); KIT-PerCPCy5.5 (2B8; BioLegend; 105824; 1:100); KIT-Bv605 (ACK2; BioLegend; 135120; 1:200); Sca1-PECy7 (D7; BioLegend; 108102; 1:200); CD16/CD32 (FcγRII/III)-Alexa700 (93; eBioscience; 56-0161-82; 1:200); CD34-FITC (RAM34; BD Biosciences; 553731; 1:200); CD45.1-FITC (A20; BioLegend; 110706; 1:200); CD45.1-PerCPCy5.5 (A20; BioLegend; 110728; 1:200); CD45.1-PE (A20; BioLegend; 110708; 1:200); CD45.1-APC (A20; BioLegend; 110714; 1:200); CD45.2-PE (104; eBioscience; 12-0454-82; 1:200); CD45.2-Alexa700 (104; BioLegend; 109822; 1:200); CD45.2-Bv605 (104; BioLegend; 109841; 1:200); CD48-Bv711 (HM48-1; BioLegend; 103439; 1:200); CD150 (9D1; eBioscience; 12-1501-82; 1:200). DAPI was used to exclude dead cells. For sorting human leukaemia cells, cells were stained with a lineage cocktail including CD34-PerCP (8G12; BD Biosciences; 345803; 1:200); CD117-PECy7 (104D2; eBioscience; 25-1178-42; 1:200); CD33-APC (P67.6; BioLegend; 366606; 1:200); HLA-DR-FITC (L243; BioLegend; 307604; 1:200); CD13-PE (L138; BD Biosciences; 347406; 1:200); CD45-APC-H7 (2D1; BD Biosciences; 560178; 1:200). The composition of mature haematopoietic cell lineages in the BM, spleen and peripheral blood was assessed using a combination of CD11b, Gr1, B220, and CD3. For the haematopoietic stem and progenitor analysis, a combination of CD11b, CD11c, Gr1, B220, CD3, NK1.1 and Ter119 was stained as lineage-positive cells. Fluorescence-activated cell sorting (FACS) was performed on a FACS Aria, and analysis was performed on an LSRII or LSR Fortessa (BD Biosciences). For western blotting, DNA dot blot assays, and ChIP assays, the following antibodies were used: INTS1 (purchased from Bethyl laboratories; catalogue no.: A300-361A; dilution: 1:1,000), INTS2 (Abcam; ab74982; 1:1,000), INTS3 (Bethyl laboratories; A300-427A; 1:1,000, Abcam; ab70451; 1:1,000), INTS4 (Bethyl laboratories; A301-296A; 1:1,000), INTS5 (Abcam; ab74405; 1:1,000), INTS6 (Abcam; ab57069; 1:1,000), INTS7 (Bethyl laboratories; A300-271A; 1:1,000), INTS8 (Bethyl laboratories; A300-269A; 1:1,000), INTS9 (Bethyl laboratories; A300-412A; 1:1,000), INTS11 (Abcam; ab84719; 1:1,000), Flag-M2 (Sigma-Aldrich; F-1084; 1:1,000), Myc-tag (Cell Signaling; 2276S; 1:1,000), β-actin (Sigma-Aldrich; A-5441; 1:2,000), 5-hydroxymehylcytosine (5hmC) (Active motif; 39769), RNAPII CTD repeat YSPTSPS (phospho S2) (Abcam; ab5095), RNAPII CTD repeat YSPTSPS (phospho S5) (Abcam; ab5408), and UPF1 (Abcam; ab109363; 1:1,000).

Minigene assay. We constructed *INTS3*-WT minigene spanning exons 4 to 5 of human *INTS3* into pcDNA3.1(+) vector (Invitrogen) using BamHI and XhoI sites, respectively. Artificial mutations were engineered into *INTS3*-WT minigene using the QuikChange Site-Directed Mutagenesis Kit (Agilent) to generate *INTS3*-GGNG, *INTS3*-CCNG, *INTS3*-WT_CG(−) *INTS3*-GGNG_CG(−), and *INTS3*-CCNG_CG(−) minigenes, respectively, and the sequences of inserts were verified by Sanger sequencing. Plasmids (1 μg) were transfected using Lipofectamine LTX reagent with PLUS reagent (Invitrogen) including 0.2 μg of eGFP and 0.8 μg of *INTS3* minigene, per well of a 6-well plate. Total RNA was extracted 48 h after transfection using TRIzol reagent (Ambion), followed by DNase I treatment (Qiagen). cDNA was synthesized with an oligo-dT primer using ImProm-II reverse transcriptase (Promega). Radioactive PCR was done with ^{32}P-α-dCTP, 1.25 units of AmpliTaq (Invitrogen) and 26 cycles using primer pairs 5′-GCTTGGTACCGAGCTCGGATC-3′ (vector-specific forward primer) and 5′-CAGTTCCCGTACCAACCACAC-3′ (reverse primer for *INTS3* versions of minigene), or 5′-CAGTTCCATTACCAACCACAC-3′ (reverse primer for *INTS3*_CG(-) versions of minigene). Products were run on a 5% PAGE and the bands were quantified using a Typhoon FLA 7000 (GE Healthcare). eGFP was used as a control for transfection efficiency and exogenous eGFP was amplified using a vector specific forward primer and reverse primer on eGFP. eGFP products were loaded after we ran the INTS3 products for 20–30 min. Percentages of intron 4 retention were normalized against exogenous eGFP.

Cell culture. K562 (human chronic myeloid/erythroleukaemia cell line) and HL-60 (human promyelocytic leukaemia cell line) leukaemia cells, K052 (human multilineage leukaemia cell line) leukaemia cells, TF1 (human erythroleukaemia cell line) leukaemia cells, *MLL-AF9/Nras*^{G12D} murine leukaemia (RN2) cells[29], and Ba/F3 (murine pro-B cell line) cells were cultured in RPMI/10% fetal calf serum (FCS, heat inactivated), RPMI/20% FCS, RPMI/10% FCS + human granulocyte-macrophage colony-stimulating factor (GM-CSF, R&D Systems; 5 ng ml^{-1}), and RPMI/10% FCS + mIL3 (R&D Systems; 1 ng ml^{-1}), respectively. None of the cell lines above were listed in the database of commonly misidentified cell lines maintained by ICLAC and NCBI Biosample.

MSCV-*IDH2*$^{WT/R140Q/R172K}$-IRES-GFP, MSCV-3×Flag-*INTS3*-puro, MSCV-IRES-3×Flag-*INTS3*-mCherry, MSCV-IRES-TET2 catalytic domain cDNA-mCherry ('TET2CD'), and empty vectors of these constructs were used for retroviral overexpression studies and pRRLSIN.cPPT.PGK-mCherry.WPRE-*SRSF2*$^{WT/P95H}$ constructs were used for lentiviral overexpression studies. TET2CD

cDNA fragment with Myc tag was generated by PCR amplification using pCMVT$_N$T- TET2CD[30] as a template and inserted in the BglII restriction sites of MSCV-IRES-mCherry. Retroviral supernatants were produced by transfecting 293 GPII cells with cDNA constructs and the packaging plasmid VSV.G using XtremeGene9 (Roche) or polyethylenimine hydrochloride (Polysciences). Lentiviral supernatants were produced by similarly transfecting HEK293T cells with cDNA constructs and the packaging plasmid VSV.G and psPAX2. Virus supernatants were used for transduction in the presence of polybrene (5 μg ml^{-1}). GFP$^+$mCherry$^+$ double-positive HL-60 cells and mCherry$^+$ K562 cells were FACS-sorted to obtain cells expressing wild-type or mutant IDH2 and SRSF2 in various combinations. Isogenic HL-60 cells transduced with 3×Flag-tagged INTS3 or empty vector were obtained by puromycin selection (1 μg ml^{-1}). To let the cells fully establish epigenetic changes, they were analysed after culture for more than 30 days.

For in vitro colony-forming assays, a single-cell suspension was prepared and 15,000 cells per 1.5 ml were plated in triplicates in cytokine-supplemented methylcellulose medium (MethoCult GF M3434; StemCell Technologies), and colonies were enumerated every week. For the colony-forming assays shown in Extended Data Fig. 3k, *IDH2*^{WT} + *Srsf2*^{WT} and *IDH2*^{WT} + *Srsf2*^{P95H} mice were euthanized at day 315 post-transplant to collect BM as controls.

shRNA-mediated silencing. shRNAs against human *INTS3* (hINTS3), mouse *Ints3* (mInts3), and mouse *Tet3* (mTet3) were cloned into MLS-E-Cherry and/or MLS-E-GFP vector and those against human *UPF1* (hUPF1), mouse *Fto* (mFto), and mouse *Alkbh5* (mAlkbh5) were cloned into LT3GEPIR (pRRL) Lenti-GFP-Puro-Tet-ON all-in-one vector. The antisense sequences were: hINTS3-1: TTTTCGAAACATAACCAGGTTA; hINTS3-2: TAAATATTAGGTACAGAGGCTT; mInts3-1: TTAAAAACAATTTAAAACTCGA; mInts3-2: TACAAATGCAGACTGACAGGAA; mInts3-3: TTCTTATCCTGAAAGGAGGGGA; mInts3-4: TTTAAAACTCGATTATCTTTGC; mInts3-5: TAATCTTACAAGGTCCCGGCCA; mTet3-1: TTATTAAGACCAAACCTGGCTA; mTet3-2: TTAAATGAAGTGTAGGCCATGC; mTet3-3: TTAAATGGAATTTTAAAACTAC; mTet3-4: GCCTGTTAGGCAGATTGTTCT; mTet3-5: GCTCCAACGAGAAGCTATTTG; hUPF1-1: TGGTATTACAGTAAACCACGCA; hUPF1-2: TTGTGATTTAAACTCGTCACCA; mFto-1: TTCTAAGATATAATCCAAGGTG; mFto-2: TCTGGTTTCTGCTGTACTGGTA; mAlkbh5-1: TTGAACTGGAACTTGCAGCCGA; mAlkbh5-2: TTCATCAGCAGCATACCCACTG. mCherry$^+$ or GFP$^+$ cells with shRNAs against hINTS3, mInts3, or mTet3 were FACS-sorted.

Semi-quantitative and quantitative RT–PCR and mRNA stability assay. Total RNA was isolated using TRIzol reagent (Life Sciences) with standard RNA extraction protocol for snRNA quantification or using an RNeasy Mini or Micro kit (Qiagen) with DNase I treatment (Qiagen). For cDNA synthesis, total RNA was reverse transcribed with EcoDry kits (Random Hexamer or Oligo dT kits; Clontech), SuperScript (Invitrogen), RNA-Quant cDNA synthesis Kit (System Biosciences), or Verso cDNA Synthesis Kit (Thermo Fisher Scientific). Primers used in reverse-transcriptase polymerase chain reactions (RT–PCR) were: *INTS3* forward1: TGAGTCGTGATGGCATGAAT (exon 4), reverse1: TCTTCACCAGTTCCCGTACC (exon 5; for detection of intron 4 retention), reverse2: CTGCTCTTCAGGACCCACTC (exon 7; for detection of exon 5 skipping); *NDUFAF6* forward: GCCTGTGGCCATTGAACTAT, reverse: ACAATGCCTTGTGCTTTTCC; *PHF21A* forward: TCCATGGCCTGGAACTTTAG, reverse: GCCAGGATGGTGTTCTTCAT; *GLYR1* forward: AGGTCAGGCCCAGTTCTCTT, reverse: TCACGTCTAAGCGTCCAGTG; *GAPDH* forward: GCAAATTCCATGGCACCGTC, reverse: TCGCCCCACTTGATTTTGG.

The PCR cycling conditions (33 cycles) chosen were as follows: (1) 30 s at 95 °C (2) 30 s at 60 °C (3) 30 s at 72 °C with a final 5-min extension at 72 °C. Reaction products were analysed on 2% agarose gels. The bands were visualized by ethidium bromide staining.

Quantitative real-time reverse transcriptase PCR (qPCR) analyses were performed on an Applied Biosystems QuantStudio 6 Flex cycler using SYBR Green Master Mix (Roche). The following primers were used: *hINTS3*: forward2: CTGCAGGATACCTGCCGTA (exon 4), reverse3: CTTTCCCGTTCCTGACAGAG (intron 5; for specific quantification of transcript with intron 4 retention); forward1: TGAGTCGTGATGGCATGAAT (exon 4), reverse4: GGCTGTAACATCTCCACCTGA (exon 4–6; for specific quantification of transcript with exon 5 skipping); forward3: GGGCAATGCTGAGAGAGAAG (exon 14), reverse5: TGCCTCTGCATTGTCATAGC (exon 15); *mInts3*: forward: GTGGCTGTTATTGACTCTGCAC, reverse: CAGGTTCCCCATCATCACAT; *mFto*: forward: CACTTGGCTTCCTTACCTGACCCCC, reverse: GGTATGCTGCCGGCCTCTCGG; *mAlkbh5*: forward: CGGCCTCAGGACATTAAGGA, reverse: TCGCGGTGCATCTAATCTTG; Total U2snRNA: forward: CTTCTCGGCCTTTTGGCTAAGAT, reverse: GTACTGCAATACCAGGTCGATGC; uncleaved U2snRNA: forward: ACGTCCTCTATCCG+AGGACAATA, reverse: GCAGGTGCTACCGTCTCTCAC; total U4snRNA: forward: GCAGTATCGTAGCCAATGAGGTCTA, reverse: CCAGTGCCGACTATATTGCAAGTC.

Uncleaved U4snRNA: forward: CGTAGCCAATGAGGTCTATCCG, reverse: CCTCTGTTGTTCAACTGCAAGAAA; *hGAPDH*: forward: GCAAATTCCATGGCACCGTC, reverse: TCGCCCCACTTGATTTTGG; *mGapdh*: forward: TGGAGAAACCTGCCAAGTATG, reverse: GGAGACAACCTGGTCCTCAG.

All samples, including the template controls were assayed in triplicate. The relative number of target transcripts was normalized to the housekeeping gene found in the same sample. The relative quantification of target gene expression was performed with the standard curve or comparative cycle threshold (C_T) method.

mRNA stability assay was performed as previously described[6]. In brief, anti-*UPF1* shRNA- or control shRNA lentivirus-infected K562 *SRSF2*^{P95H} knock-in cells were generated by puromycin selection (1 μg ml^{-1}) for 7 days and shRNAs against *UPF1* were expressed by doxycycline (2 μg ml^{-1}) for 2 days. GFP (shRNA)-positive cells were FACS-sorted, treated with 2.5 μg ml^{-1} actinomycin D (Life Technologies), and collected at 0, 2, 4, 8, and 12 h.

ChIP assays. Cells were crosslinked and collected. Chromatin was broken down into 200–1,000-bp fragments using an E220 Focused-ultrasonicator. An antibody was added into the lysate and incubated overnight at 4 °C. Twenty microlitres of ChIP–grade Protein A/G Dynabeeds was added into each IP tube and incubated for 2 h. IP samples were washed and crosslinks reversed by adding proteinase K and incubating overnight at 65 °C. DNA was purified with AMPureXP beads and eluted DNA was subjected to qPCR to measure the enrichment. RNAPII antibody (05-623; EMD Millipore) was used in this study. Primer sequences used for ChIP-PCR were as follows: Intron 3-1 forward: atacccggcccttgctatac, reverse: gcaacttccttagcctgctg; Intron 3-1 forward: atacccggcccttgctatac, reverse: gcaacttcctta gcctgctg; Intron 3-2 forward: ctggcaggtgaaaagcagat, reverse: ggcaggggagagaaaagc; Intron 3-3 forward: agcaggcttttctgcctcat, reverse: tttctttccacaggggtcct; Exon 4 forward: cgggacttagctctggtgag, reverse: cctgagtacggcaggtatcc; Intron 4 forward: ctct gtcaggaacgggaaag, reverse: tgtgagtttgagaagggagcta; Exon 5 forward: acgggaactggtgaa gagtg, reverse: ctgggctctcctcctttctt; Intron 5-1 forward: ctccaccccattatctgaa, reverse: aaatgtcagggtctgttctgtg; Intron 5-2 forward: tcggtgacatctgtctgagc, reverse: cagtgggctaa tggtgaggt; Intron 5-3 forward: aacactgatgctcctgttttga, reverse: actatgccttgccccaggt; Intron 5-4 forward: gctgttgtcagccacctgta, reverse: tttggcccttgaaaatgaac; Intron 5-5 forward: tgtgttaattctgccccaca, reverse: ggatgtcctgagtcctgcac; Intron 5-6 forward: gtaatgggatggcagtcagg, reverse: cctgatttcaaaaggggaaa; Exon 6 forward: agcaaaggtagc atccacca, reverse: cttgcctccccctctctaac; Intron 6-1 forward: tttgatccagacctccttgg, reverse: gcaggggagaaaaggatacc; Intron 6-2 forward: gggggtacatattgggcttt, reverse: gaaagcctcacctccaaaca; Intron 6-3-CTCF binding site forward: ctcctcccaacgttcacact, reverse: atccgtgcccagagcacta; Intron 6-4 forward: agggggcctttcaactctt, reverse: atggggacaggacgtatttg; Intron 6-5 forward: ttccctgccttccaacag, reverse: tcccagttgctt taaaaggagt.

ChIP–seq libraries were prepared as previously described[31] and sequenced by the Integrated Genomics Operation (IGO) at MSK with 50 bp paired-end reads.

ChIP–seq of primary human AML samples. ChIP was performed as previously described[32] using the following antibodies: RNAPolII-Ser2P antibody ChIP Grade (Abcam ab5095), RNAPII-Ser5P antibody [4H8] (Abcam ab5408), and anti-HP1γ antibody, clone 42s2 (05-690 from Merck Millipore). Libraries were size-selected with AMPure beads (Beckman Coulter) for 200–800-bp size range and quantified by qPCR using a KAPA Library Quantification Kit. ChIP–seq data were generated using the NextSeq platform from Illumina with 2 × 75 bp Hi Output (all samples pooled, and sequenced on four consecutive runs before merger of FASTQ files).

Histological analyses. Mice were euthanized and autopsied, and dissected tissue samples were fixed in 4% paraformaldehyde, dehydrated, and embedded in paraffin. Paraffin blocks were sectioned at 4 μm and stained with haematoxylin and eosin (H&E). Images were acquired using an Axio Observer A1 microscope (Carl Zeiss) or scanned using a MIRAX Scanner (Zeiss).

Patient samples. Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center (under MSK IRB protocol 06-107), Université Paris-Saclay (under declaration DC-200-725 and authorization AC-2013-1884), and the University of Manchester (institution project approval 12-TISO-04), and conducted in accordance with the Declaration of Helsinki protocol. Written informed consent was obtained from all participants. Manchester samples were retrieved from the Manchester Cancer Research Centre Haematological Malignancy Tissue Biobank, which receives sample donations from all consenting patients with leukaemia presenting to The Christie Hospital (REC Reference 07/H1003/161+5; HTA license 30004; instituted with approval of the South Manchester Research Ethics Committee). Patient samples were anonymized by the Hematologic Oncology Tissue Bank of MSK, Biobank of Gustave Roussy, and the Manchester Cancer Research Centre Haematological Malignancy Tissue Biobank.

Mutational analysis of patient samples. Genomic DNA is routinely extracted from mononuclear cell samples submitted to the Manchester Cancer Research Centre Haematological Tissue Biobank. Targeted sequencing for recurrent myeloid mutations, using either: (a) a 54 gene panel (TruSight Myeloid; Illumina),

pooling 96 samples with 5% PhiX onto a single NextSeq high output, 2 × 151-bp sequencing run; variant call format (VCF) files were analysed using Illumina's Variant Studio software; (b) a 40 gene panel (Oncomine Myeloid Research Assay; ThermoFisher), processing 8 samples per Ion 530 chip on the IonTorrent platform; data analysis performed using the Ion Reporter software; (c) a 27 gene custom panel (48 × 48 Access Array; Fluidigm) sequenced by Leeds HMDS on the MiSeq platform (300v2); or (d) MSK HemePACT[33] targeting all coding regions of 585 genes known to be recurrently mutated in leukaemias, lymphomas, and solid tumours. All panels provide sufficient coverage to detect minimum variant allele fraction 5% for all genes, except for the Access Array panel and *SRSF2*; all samples genotyped by this approach underwent manual Sanger sequencing of *SRSF2* exon 1 using the following primers (tagged with Fluidigm Access Array sequencing adaptors CS1/CS2): forward: acactgacgacatggttctacacccgtttacctgcggctc, reverse: tacggtagc agagacttggtctccttcgttcgctttcacgacaa.

Statistics and reproducibility. Statistical significance was determined by (1) unpaired two-sided Student's *t*-test after testing for normal distribution, (2) one-way or two-way ANOVA followed by Tukey's, Sidak's, or Dunnett's multiple comparison test, or (3) Kruskal–Wallis tests with uncorrected Dunn's test where multiple comparisons should be adjusted (unless otherwise indicated). Data were plotted using GraphPad Prism 7 software as mean values, with error bars representing standard deviation. For categorical variables, statistical analysis was done using Fisher's exact test or χ^2-test (two-sided). Representative western blot and PCR results are shown from three or more than three biologically independent experiments. Representative flow cytometry results and cytomorphology are shown from biological replicates ($n \geq 3$). $^*P < 0.05$, $^{**}P < 0.01$ and $^{***}P < 0.001$, respectively, unless otherwise specified.

mRNA isolation, sequencing, and analysis. RNA was extracted as shown above. Poly(A)-selected, unstranded Illumina libraries were prepared with a modified TruSeq protocol. 0.5× AMPure XP beads were added to the sample library to select for fragments <400 bp, followed by 1× beads to select for fragments >100 bp. These fragments were then amplified with PCR (15 cycles) and separated by gel electrophoresis (2% agarose). DNA fragments 300 bp in length were isolated and sequenced on an Illumina HiSeq 2000 (about 100 × 10^6 101-bp reads per sample).

Primary samples from the Manchester Cancer Research Centre Haematological Malignancies Biobank with known *IDH2/SRSF2* mutation genotype were FACS-sorted to enrich for blasts on a FACS Aria III sorter using a panel including the following antibodies (all mouse anti-human): CD34-PerCP (8G12, BD); CD117-PECy7 (104D2, eBioscience); CD33-APC (P67.6, BioLegend); HLA-DR-FITC (L243, BioLegend); CD13-PE (L138, BD); CD45-APC-H7 (2D1, BD). RNA was extracted immediately using a Qiagen Micro RNeasy kit. All RNA samples had RIN values > 8. Poly(A)-selected, strand-specific SureSelect (Agilent) mRNA libraries were prepared using 200 ng RNA according to the manufacturer's protocol. Libraries were pooled and sequenced (2 × 101 bp paired end) to >100 million reads per sample on two HiSeq 2500 high throughput runs before retrospective merger of FASTQ files for downstream alignment and splicing analysis as described below. Transcriptional analysis was done using gene set enrichment analysis (GSEA)[34].

Publicly available RNA-seq data. Unprocessed RNA-seq reads of TCGA and Leucegene datasets (patients with AML) were downloaded from NCI's Genomic Data Commons Data Portal (GDC Legacy Archive; TCGA-LAML dataset) and NCBI's Sequence Read Archive (SRA; accession number SRP056295). The TCGA dataset consists of paired-end 2 × 50-bp libraries, with an average read count of 76.92 M. The Leucegene dataset consists of paired-end 2 × 100-bp libraries, with an average read count of 50.40 M per sample. The RNA-seq samples in the Leucegene dataset have 1–3 sequencing runs (about 50 M each run), and only one run was used to represent each RNA-seq sample.

Genome and splice junction annotations. Human assembly hg38 (GRCh38) and Ensembl database (human release 87) were used as the reference genome and gene annotation, respectively. RNA-seq reads were aligned by using 2-pass STAR 2.5.2a[35]. Known splice junctions from the gene annotation and new junctions identified from the alignments of the TCGA dataset were combined to create the database of alternative splicing events for splicing analysis.

Mutational analysis for the RNA-seq data. Samtools (1.3.1) was used to generate VCF files for seven target genes: *IDH1*, *IDH2*, *TET2*, *SF3B1*, *SRSF2*, *U2AF1*, and *ZRSR2* with mpileup parameters (-Bvu). The VCF files were further processed by our in-house scripts to filter out mutations with a variant allele frequency (VAF) lower than 15%. The filtered VCF files were used for variant effect predictor (v.89.4) to annotate the consequences of the mutations. We defined control patient samples as those without mutations in the seven target genes, *IDH2* mutated samples as those with only *IDH2* mutations but no mutations in the other six target genes, *SRSF2* mutated samples as those with only *SRSF2* mutations but no mutations in the other six target genes, Double-mutant samples as those with both *IDH2* and *SRSF2* mutations but no mutations in the other five target genes, and 'others' as those with mutations in *IDH1*, *TET2*, *SF3B1*, *U2AF1*, and *ZRSR2*.

Identification and quantification of differential splicing. The inclusion ratios of alternative exons or introns were estimated by using PSI-Sigma[25]. In brief, the new PSI index considers all isoforms in a specific gene region and can report the PSI value of individual exons in a multiple-exon-skipping or more complex splicing event. The database of splicing events was constructed based on both gene annotation and the alignments of RNA-seq reads. A new splicing event not known to the gene annotation is labelled as 'novel' and a splicing event with a reference transcript that is known to induce nonsense-mediated decay is labelled as 'NMD' in Supplementary Tables. The inclusion ratio of an intron retention isoform is estimated based on the median of 5 counts of intronic reads at the 1st, 25th, 50th, 75th and 99th percentiles in the intron. A splicing event is reported when both sample-size and statistical criteria are satisfied. The sample-size criterion requires a splicing event to have more than 20 supporting reads in more than 75% of the 2 populations in the comparison. For example, for a comparison of 130 control versus 6 *IDH2* mutant samples, a splicing event would be reported only when having more than 98 controls and 5 *IDH2* mutant samples with more than 20 supporting reads. In addition, a splicing event is reported only when it has more than 10% PSI change in the comparison and has a *P* value lower than 0.01.

To generate Fig. 4f, RNA-seq reads were mapped and PSI values were calculated using junction-spanning reads as previously described[36,37]. All reads mapping to the INTS3 introns (chr1:153,718,433-153,722,231; hg19) were extracted from the .bam files and the per-nucleotide coverage was calculated. Data from normal peripheral blood and BM mononuclear cells and $CD34^+$ cord blood cells are combined and shown as normal haematopoietic cells.

Motif enrichment and distribution. Motif analysis was done by using MEME SUITE[38]. In brief, the sequences of alternative exons of exon-skipping events were extracted from a given strand of the reference genome. The sequences were used as the input for MEME SUITE to search for motifs. One occurrence per sequence was set to be the expected site distribution. The width of motif was set to 5. The top 1 motif was selected on the basis of the ranking of *E*-value.

Heat map and sample clustering (differential splicing). The heat maps and sample clustering were done by using MORPHEUS (https://software.broadinstitute.org/morpheus/). The individual values in the matrix for the analysis were PSI values of a splicing event from a given RNA-seq sample. Splicing events were selected based on three criteria: (1) present in both TCGA and Leucegene datasets; (2) more than 15% PSI changes; and (3) false discovery rate smaller than 0.01. Unsupervised hierarchical clustering was based on one minus Pearson's correlation (complete linkage).

Correlation between global changes in splicing and DNA methylation. DNA methylation levels were determined by eRRBS and differentially spliced events were obtained from RNA-seq data. In Fig. 3e, Overlaps of differentially methylated regions of DNA with differential splicing was obtained by evaluating differential cytosine methylation in 500-bp segments of DNA at genomic coordinates at which differential RNA splicing were observed comparing AML with distinct *IDH2/SRSF2* genotypes shown (WT represents patients without mutations in *IDH1*/*IDH2*/spliceosomal genes).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

RNA-seq, ChIP–seq and eRRBS data have been deposited in the NCBI Sequence Read Archive under accession number SRP133673. Gel source data are shown in Supplementary Fig. 1. Other data that support the findings of this study are available from the authors upon reasonable request.

25. Lin, K. T. & Krainer, A. R. PSI-Sigma: a comprehensive splicing-detection method for short-read and long-read RNA-seq analysis. *Bioinformatics* btz438 (2019).
26. Moran-Crusio, K. et al. Tet2 loss leads to increased hematopoietic stem cell self-renewal and myeloid transformation. *Cancer Cell* **20**, 11–24 (2011).
27. Shih, A. H. et al. Combination targeted therapy to disrupt aberrant oncogenic signaling and reverse epigenetic dysfunction in *IDH2*- and *TET2*-mutant acute myeloid leukemia. *Cancer Discov.* **7**, 494–505 (2017).
28. Georgiades, P. et al. *VavCre* transgenic mice: a tool for mutagenesis in hematopoietic and endothelial lineages. *Genesis* **34**, 251–256 (2002).
29. Zuber, J. et al. Toolkit for evaluating genes required for proliferation and survival using tetracycline-regulated RNAi. *Nat. Biotechnol.* **29**, 79–83 (2011).
30. Lee, M. et al. Engineered split-TET2 enzyme for inducible epigenetic remodeling. *J. Am. Chem. Soc.* **139**, 4659–4662 (2017).
31. Kleppe, M. et al. Dual targeting of oncogenic activation and inflammatory signaling increases therapeutic efficacy in myeloproliferative neoplasms. *Cancer Cell* **33**, 29–43 (2018).
32. Maiques-Diaz, A. et al. Enhancer activation by pharmacologic displacement of LSD1 from GFI1 induces differentiation in acute myeloid leukemia. *Cell Rep.* **22**, 3641–3659 (2018).
33. Cheng, D. T. et al. Memorial Sloan Kettering–integrated mutation profiling of actionable cancer targets (MSK–IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J. Mol. Diagn.* **17**, 251–264 (2015).

34. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).
35. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
36. Dvinge, H. & Bradley, R. K. Widespread intron retention diversifies most cancer transcriptomes. *Genome Med.* **7**, 45 (2015).
37. Hubert, C. G. et al. Genome-wide RNAi screens in human brain tumor isolates reveal a novel viability requirement for PHF5A. *Genes Dev.* **27**, 1032–1045 (2013).
38. Bailey, T. L. et al. MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, W202–W208 (2009).
39. Robinson, J. T. et al. Integrative genomics viewer. *Nat. Biotechnol.* **29**, 24–26 (2011).
40. Intlekofer, A. M. et al. Hypoxia induces production of l-2-hydroxyglutarate. *Cell Metab.* **22**, 304–311 (2015).
41. Dvinge, H. et al. Sample processing obscures cancer-specific alterations in leukemic transcriptomes. *Proc. Natl Acad. Sci. USA* **111**, 16802–16807 (2014).
42. Macrae, T. et al. RNA-seq reveals spliceosome and proteasome genes as most consistent transcripts in human cancer cells. *PLoS ONE* **8**, e72884 (2013).

Acknowledgements We thank D. L. Fei, Y. Huang, E. Wang, I. Aifantis, M. Patel, A. S. Shih, A. Penson, E. Kim, Y. R. Chung, B. H. Durham and H. Kunimoto for technical support, J. Wilusz for sharing recent data on integrator and B. J. Druker for sharing the Beat AML RNA-seq data. A.Y. is supported by grants from the Aplastic Anemia and MDS International Foundation (AA&MDSIF) and the Lauri Strauss Leukemia Foundation. A.Y. is a Special Fellow of The Leukemia and Lymphoma Society. A.Y., S.C.-W.L. and D.I. are supported by the Leukemia and Lymphoma Society Special Fellow Award. A.Y. and D.I. are supported by JSPS Overseas Research Fellowships. D.H.W. is supported by a Bloodwise Clinician Scientist Fellowship (15030). D.H.W. and K.B. are supported by fellowships from The Oglesby Charitable Trust. S.C.-W.L. is supported by the NIH/NCI (K99 CA218896) and the ASH Scholar Award. T.C.P.S. is supported by Cancer Research UK grant number C5759/A20971. E.J.W. is supported by grants from the CPRIT (RP140800) and the Welch Foundation (H-1889-20150801). R.K.B. and O.A.-W. are supported by grants from NIH/NHLBI (R01 HL128239) and the Department of Defense Bone Marrow Failure Research Program (W81XWH-16-1-0059). A.R.K. and O.A.-W. are supported by grants from the Starr Foundation (I8-A8-075) and the Henry & Marilyn Taub Foundation. O.A.-W. is supported by grants from the Edward P. Evans Foundation, the Josie Robertson Investigator Program, the Leukemia and Lymphoma Society and the Pershing Square Sohn Cancer Research Alliance.

Author contributions A.Y., K.-T.L., A.R.K. and O.A.-W. designed the study. A.Y., B.W., S.C.-W.L., J.-B.M., X.J.Z., H.C., R.E.M., D.I., T.R.A., K.B., F.S. and E.J.W. performed mouse experiments. K.-T.L. and M.A.R. performed RNA-seq analyses and minigene assays, respectively, under the supervision of A.R.K. A.P. performed DNA methylation and ChIP–seq analyses. T.R.A. and E.J.W. provided antibodies to detect integrator components and assays for snRNA cleavage. H.D., R.K.B. and F.A. performed RNA-seq analyses. D.H.W., T.C.P.S., D.P.W., S.d.B., V.P.-L., E.M.S. and R.L.L. provided clinical samples. D.H.W. and C.C. provided clinical correlative data for primary datasets. D.H.W. performed ChIP–seq experiments under the supervision of T.C.P.S. A.M.I. provided *Idh2*^{R140Q} knock-in mice. R.L.L. provided *Tet2* knockout mice. A.Y., K.-T.L., D.H.W. and O.A.-W. prepared the manuscript with help from all co-authors.

Competing interests A.M.I. has served as a consultant and advisory board member for Foundation Medicine. E.M.S. has served on advisory boards for Astellas Pharma, Daiichi Sankyo, Bayer, Novartis, Syros, Pfizer, PTC Therapeutics, AbbVie, Agios and Celgene and has received research support from Agios, Celgene, Syros and Bayer. R.L.L. is on the Supervisory Board of Qiagen and the Scientific Advisory Board of Loxo, reports receiving commercial research grants from Celgene, Roche and Prelude, has received honoraria from the speakers bureaus of Gilead and Lilly, has ownership interest (including stock and patents) in Qiagen and Loxo, and is a consultant and/or advisory board member for Novartis, Roche, Janssen, Celgene and Incyte. A.R.K. is a founder, director, advisor, stockholder and chair of the Scientific Advisory Board of Stoke Therapeutics and receives compensation from the company. A.R.K. is a paid consultant for Biogen; he is a member of the SABs of Skyhawk Therapeutics, Envisagenics BioAnalytics and Autoimmunity Biologic Solutions, and has received compensation from these companies in the form of stock. A.R.K. is a research collaborator of Ionis Pharmaceuticals and has received royalty income from Ionis through his employer, Cold Spring Harbor Laboratory. O.A.-W. has served as a consultant for H3 Biomedicine, Foundation Medicine, Merck and Janssen. O.A.-W. has received personal speaking fees from Daiichi Sankyo. O.A.-W. has received previous research funding from H3 Biomedicine unrelated to the current manuscript. D.I., R.K.B. and O.A.-W. are inventors on a provisional patent application (patent number FHCC.P0044US.P) applied for by Fred Hutchinson Cancer Research Center on the role of reactivating BRD9 expression in cancer by modulating aberrant *BRD9* splicing in *SF3B1* mutant cells.

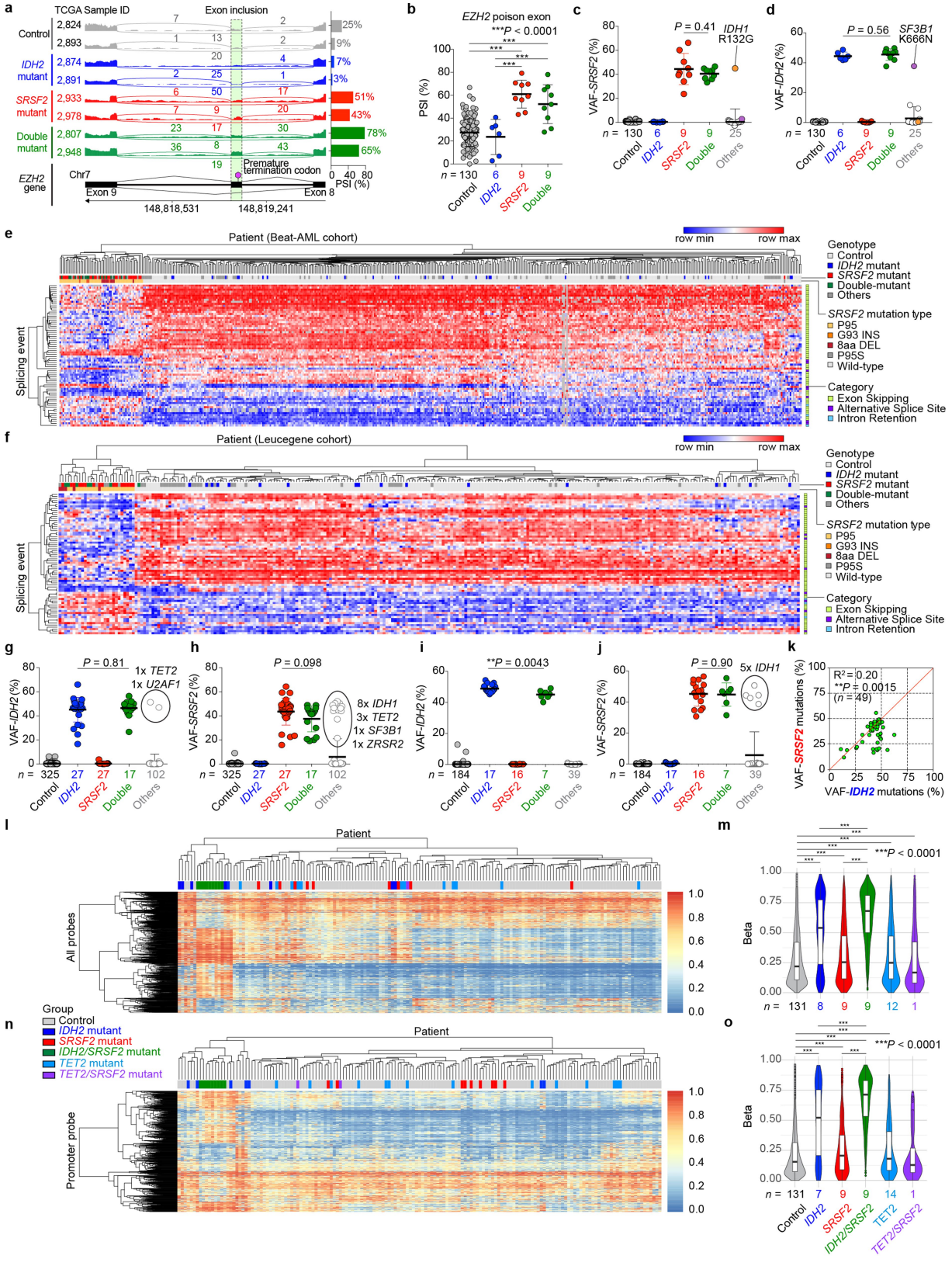
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1618-0.

Correspondence and requests for materials should be addressed to O.A.-W.

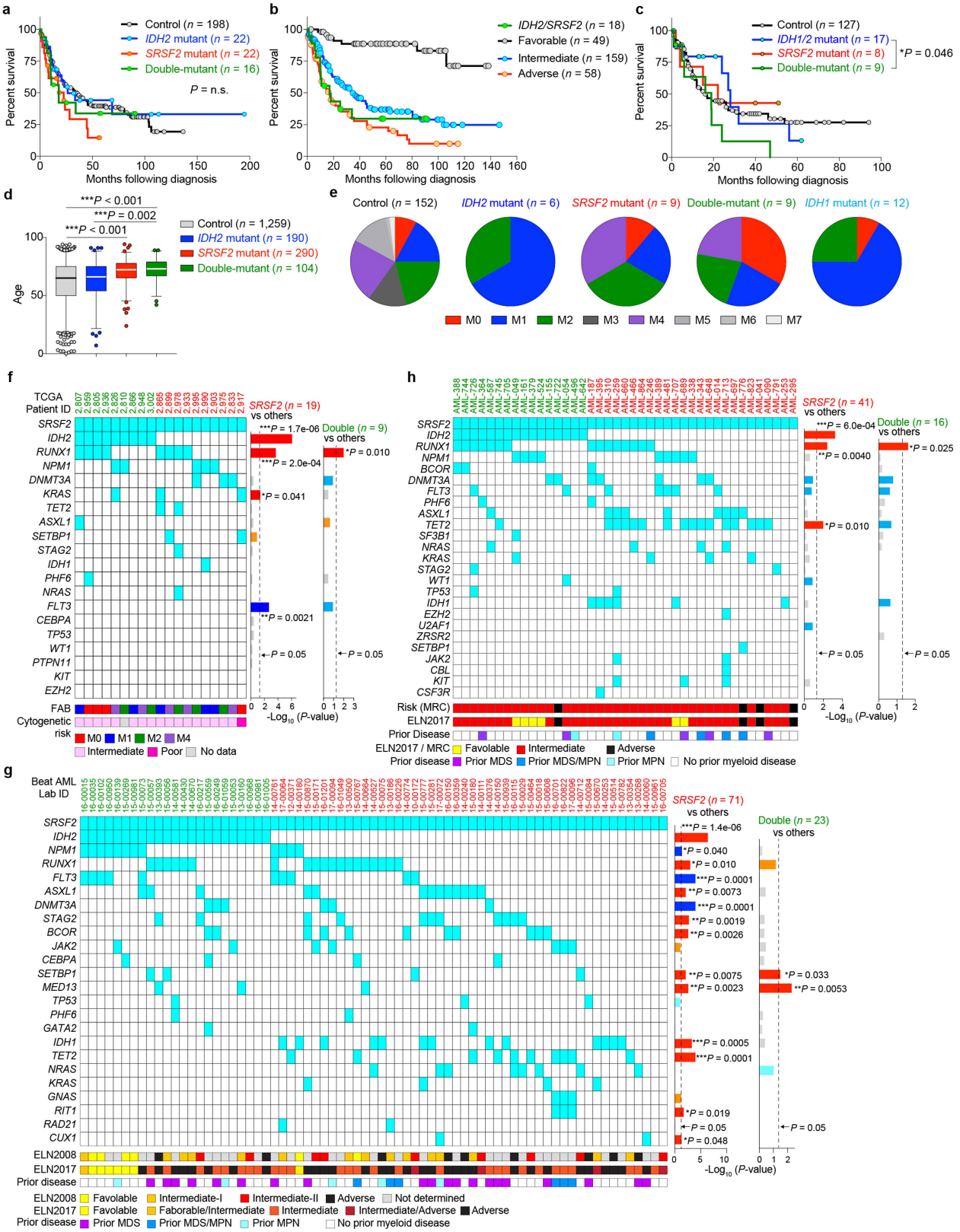
Peer review information *Nature* thanks Rotem Karni and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints.



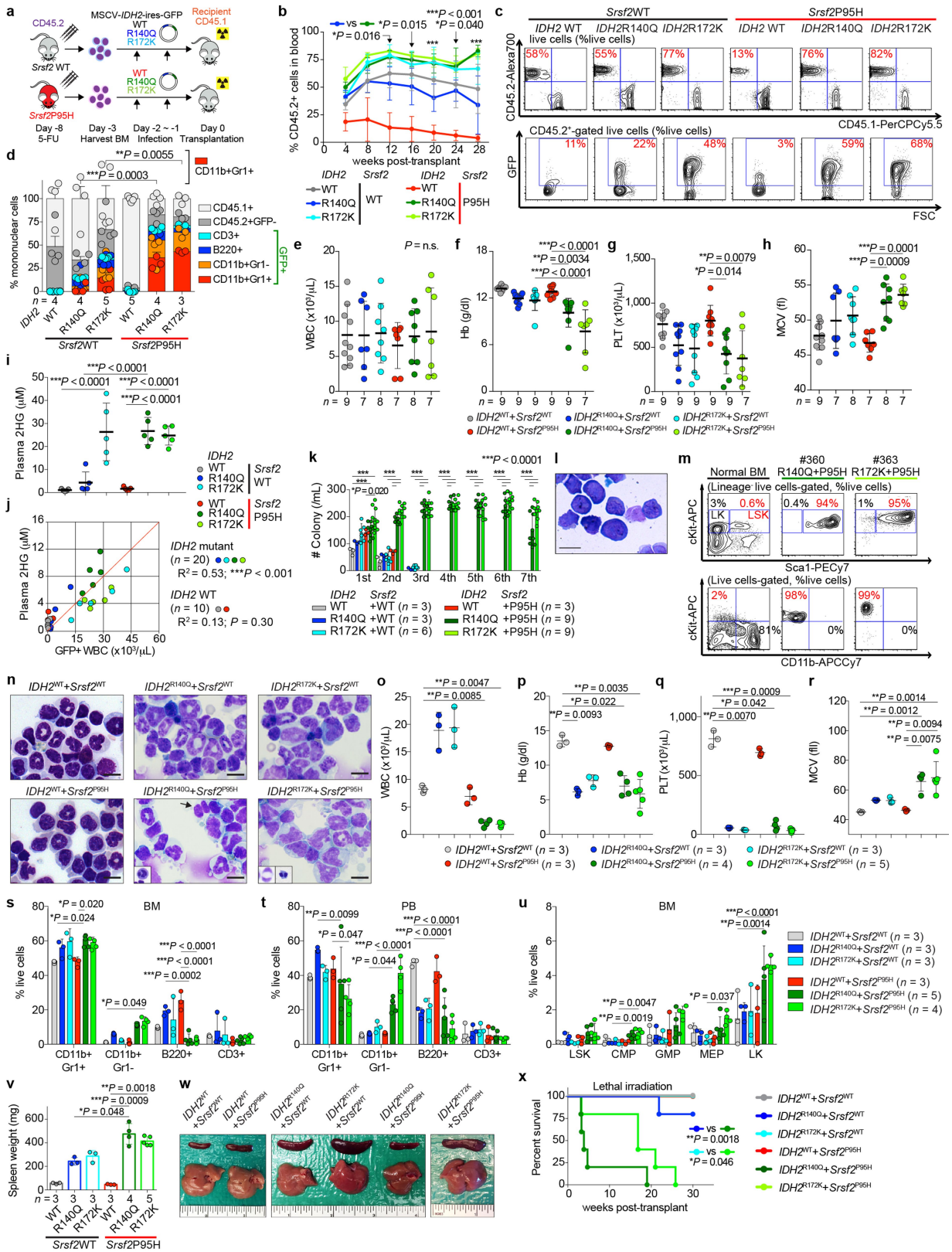
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Mutant *SRSF2*-mediated splicing events in acute myeloid leukaemia (AML). a, Representative Sashimi plots of RNA-seq data from the TCGA showing the poison exon inclusion event in *EZH2* ('Control' represents samples that are wild type (WT) for the following seven genes: *IDH1*, *IDH2*, *TET2*, *SRSF2*, *SF3B1*, *U2AF1*, and *ZRSR2*; '*IDH2* mutant' refers to patients with an *IDH2* mutation and no mutation in the other six genes; '*SRSF2* mutant' refers to patients with an *SRSF2* mutation and no mutation in the other six genes; 'double-mutant' refers to patients with an *IDH2* and *SRSF2* mutation and no mutation in the other five genes; 'others' refers to patients with mutations in *IDH1*, *TET2*, *SF3B1*, *U2AF1* or *ZRSR2*; figure made using Integrative Genomics Viewer (IGV 2.3)[39]). **b**, PSI values of *EZH2* poison exon inclusion (the number of analysed patients is indicated; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). Note that patients classified as 'others' include one patient with an $SRSF2^{P95L}$ mutation with a coexisting $IDH1^{R132G}$ mutation (TCGA ID: 2990) and one patient with an $IDH2^{R140Q}$ mutation also having an $SF3B1^{K666N}$ mutation (TCGA ID: 2973), which were excluded from the analyses shown above. **c**, **d**, **g**–**j**, VAFs of *SRSF2* mutations affecting the proline 95 residue (**c**, **h**, **j**) and *IDH2* mutations affecting *IDH2* arginine 140 or 172 (**d**, **g**, **i**) in TCGA (**c**, **d**), Beat AML (**g**, **h**) and Leucegene (**i**, **j**) datasets (mean ± s.d.; two-sided Student's *t*-test). **e**, **f**, Heat map based on the ΔPSI of mutant *SRSF2*-specific splicing events in AML from Beat AML (**e**) and Leucegene (**f**) cohorts. '8aa DEL' represents samples with 8 amino acid deletions in SRSF2 starting from proline 95, which has similar effects on splicing as point mutations affecting SRSF2 P95. Detailed information of splicing events shown is available in Supplementary Table 1. **k**, VAFs of *IDH2* (*x* axis) and *SRSF2* mutations (*y* axis) in *IDH2* and *SRSF2* double-mutant AML determined by RNA-seq data from the TCGA, Beat AML, Leucegene and our previously unpublished cohorts (Pearson correlation coefficient; *P* value (two-tailed) was calculated by Prism7). **l**, **n**, Unsupervised hierarchical clustering of DNA methylation levels of all probes (**l**) or at the promoter probes (**n**) in the TCGA AML cohort based on *IDH2*, *SRSF2* and *TET2* genotypes. **m**, **o**, DNA methylation levels of AML samples from each genotype are quantified and visualized from **l** and **n** as violin plots (the line represents mean, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; one-way ANOVA with Tukey's multiple comparison test). $^{**}P < 0.01$, $^{***}P < 0.001$.



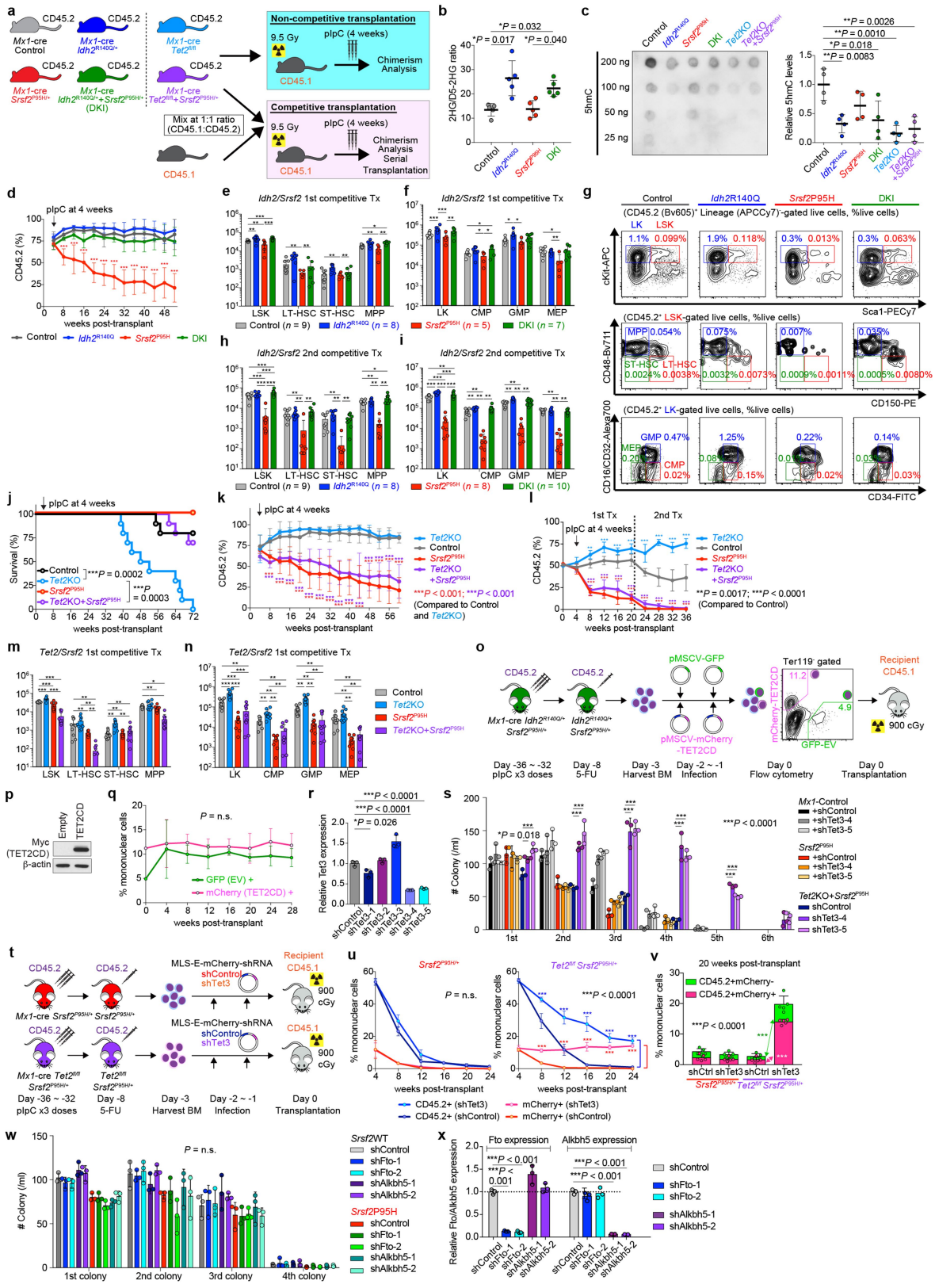
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Clinical relevance of coexisting *IDH2* and *SRSF2* mutations in AML. a–c, Kaplan–Meier survival analysis of patients with AML from the Manchester/Christie Biobank dataset (**a**: based on *IDH2* and *SRSF2* genotype ($n = 258$); **b**: based on cytogenetic risk ($n = 284$)) and the TCGA (**c**; $n = 161$; based on *IDH1*, *IDH2* and *SRSF2* genotypes (log-rank (Mantel–Cox) test (two-sided)). **d**, Age at diagnosis of patients from the TCGA, Beat AML, and Manchester/Christie Biobank cohorts combined (the line represents mean, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; samples below 2.5th percentile and above 97.5th percentile are shown as dots; one-way ANOVA with Tukey's multiple comparison test). **e**, Distribution of French–American–British (FAB) classification of patients with AML with the indicated genotypes from the TCGA cohort. **f–h**, Mutations coexisting with *IDH2* and *SRSF2* double-mutant and *SRSF2* single-mutant AML from the TCGA (**f**), Beat AML (**g**), and Manchester/Christie Biobank (**h**) cohorts are shown with FAB classification, cytogenetic risk, prior history of myeloid disorders, and genetic risk stratification based on European LeukaemiaNet (ELN) 2008 and ELN2017 guidelines (the number of patients is indicated; *P* values on the right represent statistical significance of co-occurrence (red and orange) or mutual exclusivity (blue and light blue) of each gene mutation with *SRSF2* (including those in *IDH2* and *SRSF2* double-mutant AML) or coexisting *IDH2* and *SRSF2* mutations; Fisher's exact test (two-sided)). $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.



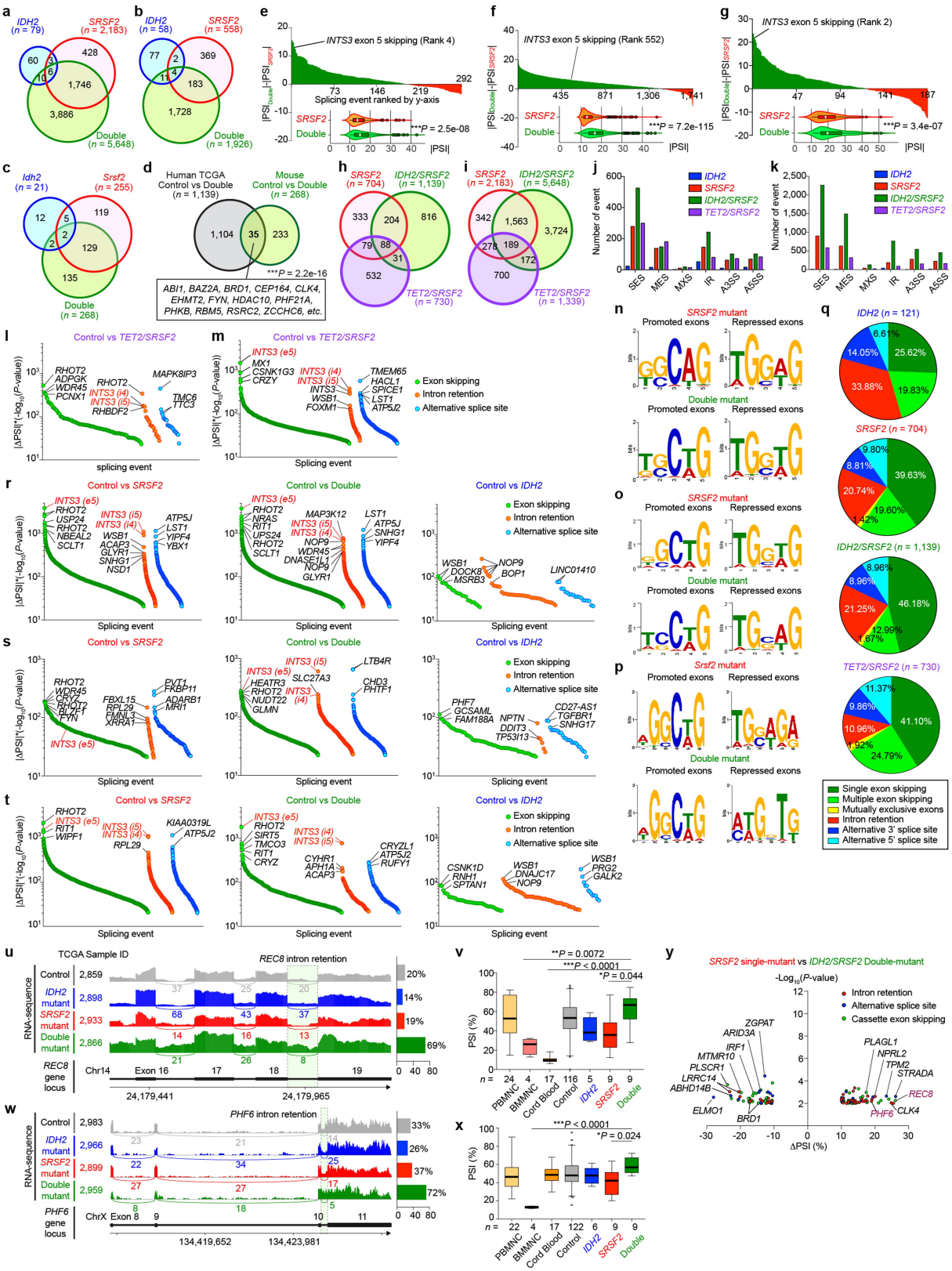
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Mutant *IDH2* cooperates with mutant *Srsf2* to generate lethal MDS with proliferative features in vivo. a, Schematic of BM transplantation model. **b**, **c**, Chimerism of CD45.2$^+$ cells in the peripheral blood of recipient mice over time (**b**) ($n = 5$ per group at 4 weeks; mean percentage ± s.d.; two-way ANOVA with Tukey's multiple comparison test) and representative flow cytometry data showing the chimerism of CD45.2$^+$ versus CD45.1$^+$ (top) or GFP$^+$ (bottom) cells in peripheral blood at 16 weeks post-transplant (**c**) (representative results from five recipient mice; the percentages listed represent the percent of cells within live cells). **d**, Composition of peripheral blood mononuclear cells (PBMNCs) at 28 weeks post-transplant (the number of analysed mice is indicated; mean + s.d.; two-way ANOVA with Tukey's multiple comparison tests; statistical significances were detected in percentage of CD11b$^+$Gr1$^+$ cells in *IDH2*^{R140Q} + *Srsf2*^{WT} versus *IDH2*^{R140Q} + *Srsf2*^{P95H} and in *IDH2*^{R172K} + *Srsf2*^{WT} versus *IDH2*^{R172K} + *Srsf2*^{P95H}). **e**–**h**, Blood counts at 20 weeks post-transplant (white blood cells (WBC) (**e**); haemoglobin (Hb) (**f**); platelets (PLT) (**g**); mean corpuscular volume (MCV) (**h**); the number of analysed mice is indicated; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison tests). **i**, Plasma 2HG levels at 20 weeks post-transplant (2HG levels were quantified as previously described[40]; $n = 5$ per group were randomly selected; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **j**, Correlations between plasma 2HG levels and number of GFP$^+$ cells in peripheral blood at 24 weeks post-transplant ($n = 5$ per group; the Pearson correlation coefficient (R^2) and *P* values (two-tailed) were calculated using PRISM 7). **k**, Colony numbers from serial replating assays of BM cells collected from end-stage mice from Fig. 2b are shown (mean value ± s.d. represented by lines above the box; the number of analysed mice is indicated; two-way ANOVA with Tukey's multiple comparison test). **l**, Giemsa staining of *IDH2*^{R140Q} *Srsf2*^{P95H} double-mutant cells from the sixth plating (scale bar, 10 μm; original magnification, ×400; representative result from 9 biologically independent experiments). **m**, Immunophenotype of colony cells at the sixth plating. Normal BM cells were used as a control (the percentage listed represent the percent of cells within live cells; representative result from nine recipient mice). **n**, Cytomorphology of BM mononuclear cells (BMMNCs) from recipient mice at end stage. BM cells from *IDH2* single-mutant and *IDH2* and *Srsf2* double-mutant groups have increased granulocytes. In addition, *IDH2* and *Srsf2* double-mutant groups had proliferation of monoblastic and monocytic cells as well as dysplastic features such as abnormally segmented neutrophils (black arrow and inset) and binucleated erythroid precursors with irregular nuclear contours (insets) (scale bar, 10 μm; original magnification, ×400; representative results from 3 controls and 9 recipients are shown; number of mice indicated in **o**–**r**). **o**–**r**, Blood counts at end-stage (WBC (**o**); Hb (**p**); PLT (**q**); MCV (**r**); the number of analysed mice is indicated; mean ± s.d.; Kruskal–Wallis tests with uncorrected Dunn's test). **s**–**u**, Results from flow cytometry analysis of BM (**s**) and peripheral blood (**t**) mature lineages as well as BM haematopoietic stem/progenitor cells (HSPC) from two tibias, two femurs, and two pelvic bones (**u**) are quantified (LSK, Lineage$^-$SCA1$^+$KIT$^+$; LT-HSC, long-term haematopoietic stem cell (HSC); ST-HSC, short-term HSC; MPP, multi-potent progenitor; LK: Lineage$^-$SCA1$^-$KIT$^+$; CMP, common myeloid progenitor; GMP, granulocyte-monocyte progenitor; MEP, megakaryocyte-erythroid progenitor; the number of analysed mice is indicated; mean + s.d.; two-way ANOVA with Tukey's multiple comparison test). **v**, **w**, Spleen weight at end stage (**v**; the number of analysed mice is indicated; mean ± s.d.; two-way ANOVA with Tukey's multiple comparison test) and representative photographs of spleens from recipient mice from **v** (**w**; each photograph was taken with an inch ruler). **x**, Kaplan–Meier survival analysis of serially transplanted recipient mice that were lethally irradiated ($n = 5$ per group; log-rank (Mantel–Cox) test (two-sided)). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



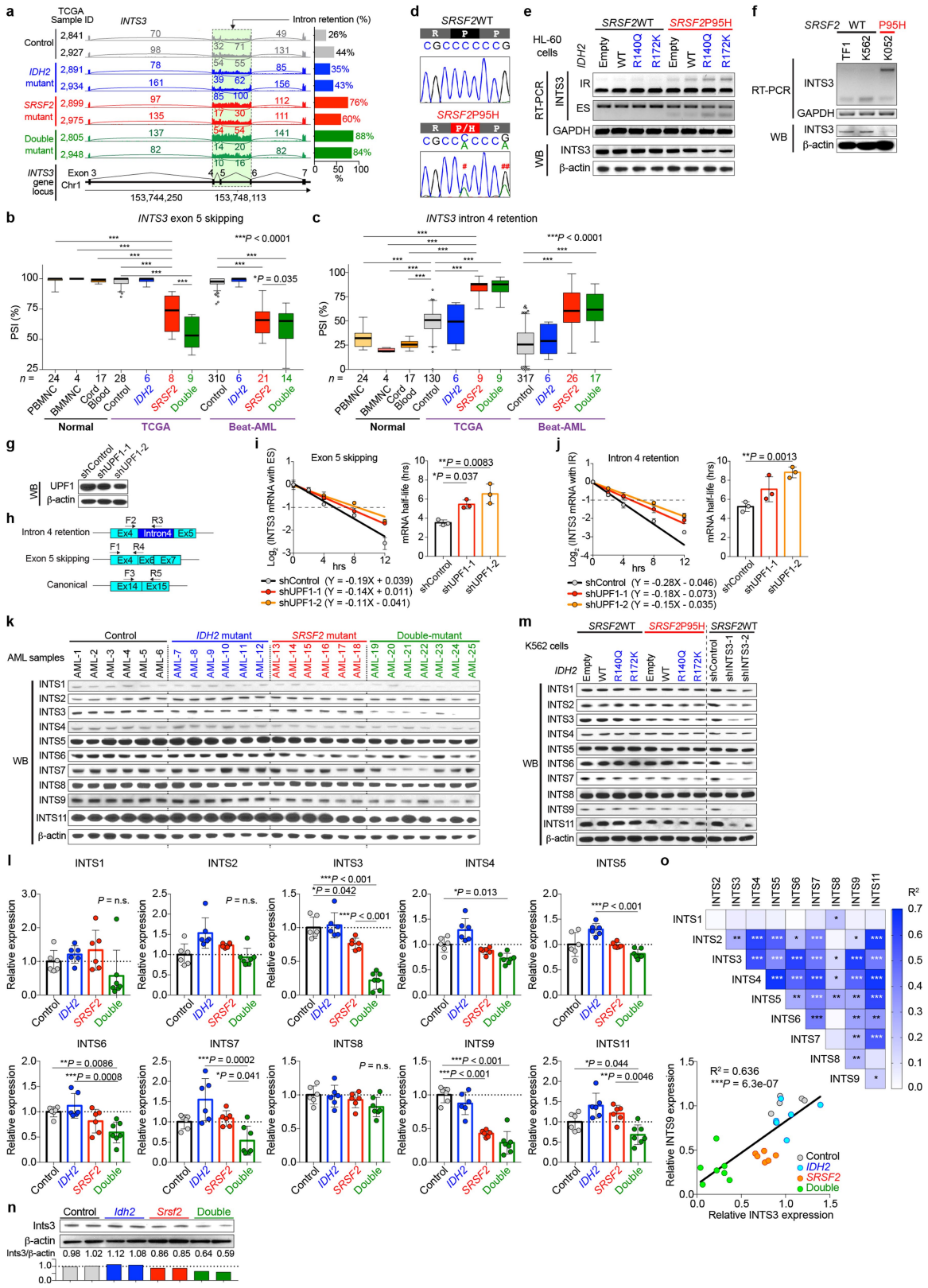
Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Collaborative effects of mutant *Idh2* and mutant *Srsf2* are not dependent on *Tet2* loss alone. a, Schematic of competitive and non-competitive transplantation assays of CD45.2$^+$ *Mx1-cre* control, *Mx1-cre Idh2*$^{R140Q/+}$, *Mx1-cre Srsf2*$^{P95H/+}$, *Mx1-cre Idh2*$^{R140Q/+}$*Srsf2*$^{P95H/+}$ mice, *Mx1-cre Tet2*$^{fl/fl}$, *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$ mice into CD45.1$^+$ recipient mice. **b**, 2HG levels of bulk PBMNCs from primary *Mx1-cre* mice were measured at three months post-pIpC (polyinosinic:polycytidylic acid) and normalized to internal standard (D-2-hydroxyglutaric-2,3,3,4,4-*d5* acid; D5-2HG) (2HG and D5-2HG levels were quantified as described[40]; $n = 5$ per group; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **c**, DNA extracted from sorted KIT$^+$ BM cells from primary *Mx1-cre* mice at one month post-pIpC was probed with antibodies specific for 5-hydroxymethylcytosine (5hmC) (left). Relative intensity of each dot was measured by ImageJ and divided by input DNA amount for comparison (right; $n = 4$; intensity of each dot divided by amount of input DNA was combined per genotype; representative results from 3 biologically independent experiments with similar results; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **d**, Chimerism of peripheral blood CD45.2$^+$ cells in non-competitive transplantation (pIpC was injected at 4 weeks post-transplant; mean ± s.d.; $n = 10$ (control and *Idh2*^{R140Q}), $n = 8$ (*Srsf2*^{P95H}), and $n = 9$ (DKI) at 0 week; two-way ANOVA with Tukey's multiple comparison test; *P* values from comparison between *Srsf2*^{P95H} and each of other groups are shown). **e**–**i**, Absolute number of BM HSPCs from two tibias, two femurs, and two pelvic bones were measured in the primary (**e**, **f**) and serial (**h**, **i**) competitive transplant of *Idh2* and *Srs2* mutant cells, and representative flow cytometry of BM HSPCs from the primary competitive transplant of *Idh2* and *Srsf2* mutant cells from **e**, **f** (the percentage listed represents the percent of cells within live cells) (the number of analysed mice is indicated; mean + s.d.; two-way ANOVA with Tukey's multiple comparison test). **j**, Kaplan–Meier survival analysis of CD45.1$^+$ recipient mice transplanted non-competitively with BM cells from CD45.2$^+$ *Mx1-cre* control, *Mx1-cre Tet2*$^{fl/fl}$, *Mx1-cre Srsf2*$^{P95H/+}$, and *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$ mice (pIpC was injected at 4 weeks post-transplant; $n = 10$ per genotype; log-rank (Mantel–Cox) test (two-sided)). **k**, **l**, Chimerism of peripheral blood CD45.2$^+$ cells in non-competitive (**k**) ($n = 10$ (control and *Tet2* knockout (*Tet2*KO)), $n = 8$ (*Srsf2*^{P95H}), and $n = 5$ (*Tet2*KO + *Srsf2*^{P95H}) at 0 weeks) or competitive (**l**) ($n = 9$ (control), $n = 10$ (*Tet2*KO), $n = 8$ (*Srsf2*^{P95H}), and $n = 10$ (*Tet2*KO + *Srsf2*^{P95H}) at 0 weeks) transplantation (pIpC was injected at 4 weeks post-transplant; percentages of CD45.2$^+$ cells at pre-transplant are also shown as data at 0 weeks in **l**; mean ± s.d.; two-way ANOVA with Tukey's multiple comparison test). **m**, **n**, Absolute number of BM HSPCs from two tibias, two femurs, and two pelvic bones were measured in the primary competitive transplant of *Tet2* and *Srsf2* mutant cells ($n = 10$ per genotype; mean + s.d.; two-way ANOVA with Tukey's multiple comparison test). **o**, Schematic of TET2 catalytic domain (CD: catalytic domain; EV: empty vector) retroviral BM transplantation model. **p**, Western blot analysis confirming the expression of Myc-tagged TET2 CD in Ba/F3 cells transduced with or without TET2 CD (representative images from two biologically independent experiments with similar results). **q**, Chimerism of mCherry–TET2 CD$^+$ and GFP–EV$^+$ cells in peripheral blood of recipient mice over time ($n = 10$; mean percentage ± s.d.; two-way ANOVA with Sidak's multiple comparison test). **r**, qPCR of *Tet3* in the first colony cells from **s** ($n = 3$; mean ± s.d.; a two-sided Student's *t*-test). **s**, Colony numbers from serial replating assays of BM cells from *Mx1-cre* control, *Mx1-cre Srsf2*$^{P95H/+}$, and *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$ mice transduced with shRNAs targeting *Tet3* (shTet3) ($n = 3$; mean + s.d.; two-way ANOVA with Tukey's multiple comparison test). **t**, Schematic of shTet3 retroviral BM transplantation model. **u**, **v**, Chimerism of mCherry$^+$ cells in CD45.2$^+$ donor cells in peripheral blood of recipient mice over time (**u**; left, *Mx1-cre Srsf2*$^{P95H/+}$; right, *Mx1-cre Tet2*$^{fl/fl}$*Srsf2*$^{P95H/+}$; $n = 5$ per group) and at 20 weeks post-transplant (**v**) (mean percentage ± s.d.; two-way ANOVA with Sidak's multiple comparison test). **w**, Colony numbers from serial replating assays of either *Mx1-cre Srsf2*$^{+/+}$ or *Srsf2*$^{P95H/+}$ BM cells transduced with an shRNA against *Fto* or *Alkbh5*. BM cells were collected at one month post-pIpC ($n = 3$; mean value ± s.d.; two-way ANOVA with Tukey's multiple comparison test). **x**, qPCR of *Fto* or *Alkbh5* in Ba/F3 cells transduced with shRNAs targeting mouse *Fto* or *Alkbh5* ($n = 3$; mean value ± s.d.; one-way ANOVA with Tukey's multiple comparison test). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



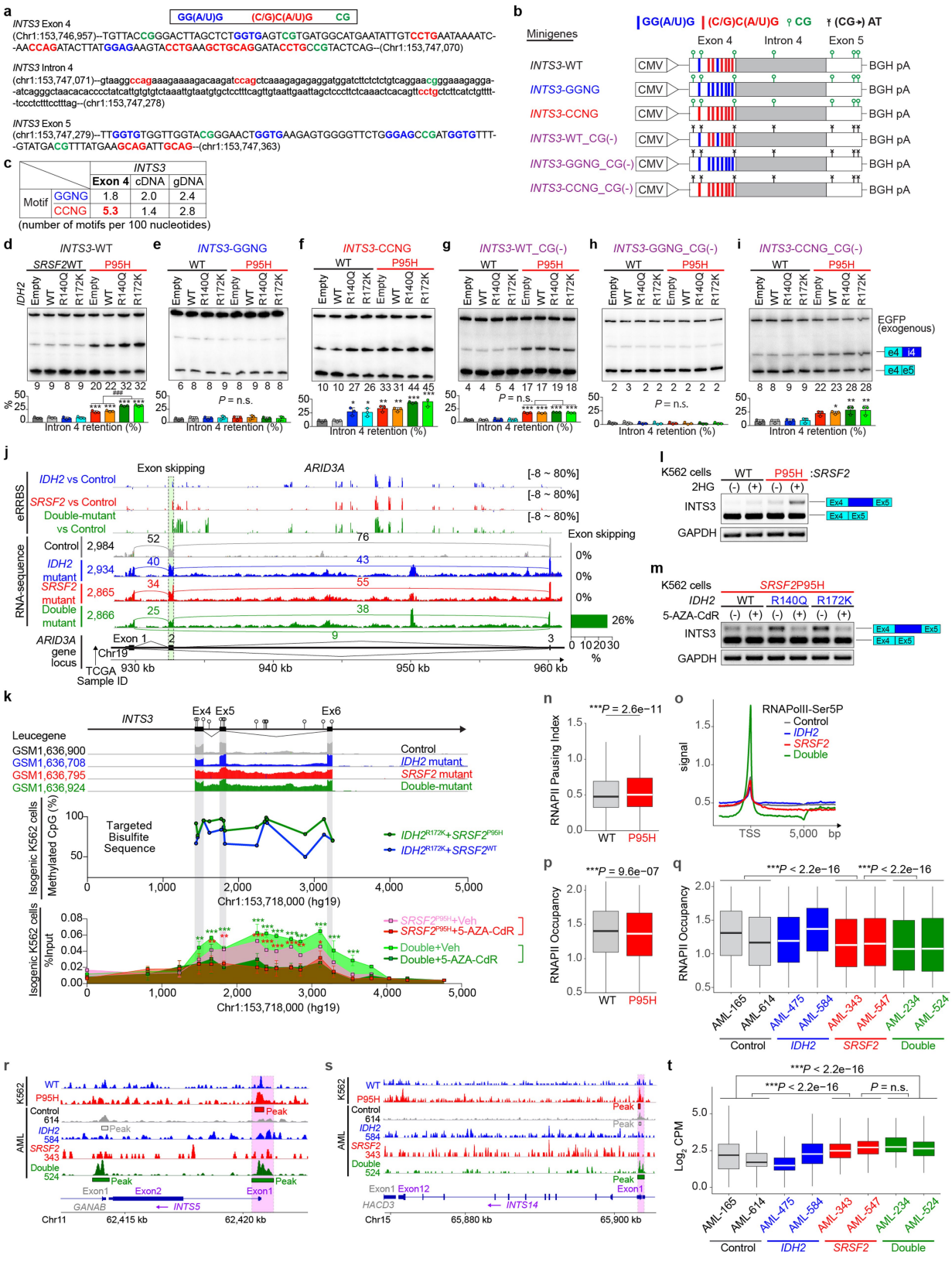
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | *IDH2* mutations augment the RNA splicing defects of *SRSF2* mutant leukaemia. a–c, Venn diagram showing numbers of differentially spliced events from the Beat AML cohort (**a**), unpublished collaborative cohort 2 (**b**) and mouse $Lin^{-}KIT^{+}$ bone marrow cells at 12 weeks post-pIpC (**c**) based on *IDH2* and *SRSF2* mutant genotypes. **d**, Venn diagram showing the numbers of overlapping alternatively spliced events between *IDH2* and *SRSF2* double-mutant AMLs and mouse models ($^{***}P = 2.2 \times 10^{-16}$; binominal test). **e–g**, $\Delta|\text{PSI}|$ ($\Delta|\text{PSI}| = |\text{PSI}|_{\text{Double}} - |\text{PSI}|_{SRSF2}$) values for each overlapping mis-spliced event in *SRSF2* single-mutant and *IDH2* and *SRSF2* double-mutant AML from the TCGA (**e**), Beat AML cohort (**f**) and unpublished collaborative cohort 2 (**g**) are plotted along the *y* axis. Spliced events shown in green and red represent events that are more robust in *IDH2* and *SRSF2* double-mutant and *SRSF2* single-mutant AML, respectively, in terms of |PSI| values. The mean |PSI| value of each event was visualized as violin plots on the bottom ($n = 292$, $n = 1{,}741$, and $n = 187$, respectively; PSI values were calculated using PSI-Sigma; the line represents mean, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; samples below 2.5th percentile and above 97.5th percentile are shown as dots; paired two-tailed Student *t*-test). **h**, **i**, Venn diagram of numbers of differentially spliced events from the TCGA (**h**) and Beat AML (**i**) datasets based on *IDH2*, *TET2* and *SRSF2* genotypes. **j**, **k**, Absolute numbers of each class of alternative splicing event from TCGA (**j**) and Beat AML (**k**) datasets are shown. SES, single-exon skipping; MES, multiple-exon skipping; MXS, mutually-exclusive splicing; A5SS, alternative 5′ splice site; A3SS, alternative 3′ splice site. **l**, **m**, Differentially spliced events ($|\Delta\text{PSI}| > 10\%$ and $P < 0.01$ were used as thresholds) in indicated genotype from the TCGA (**l**) ($n = 730$ differentially spliced events) and Beat AML (**m**) ($n = 1{,}339$ differentially spliced events) cohorts are ranked by *y* axis and class of event (PSI and *P* values adjusted for multiple comparisons were calculated using PSI-Sigma). **n–p**, Sequence logos of nucleotide motifs of exons preferentially promoted or repressed in splicing in *SRSF2* single-mutant (top) or *IDH2* and *SRSF2* double-mutant (bottom) AML from the TCGA cohort (**n**), Beat AML cohort (**o**) and mouse models (**p**). **q**, Percentage of each class of alternative splicing event in indicated genotype from TCGA cohort is shown in a pie-chart. **r–t**, Differentially spliced events ($|\Delta\text{PSI}| > 10\%$ and $P < 0.01$ were used as thresholds) in indicated genotype from the Beat AML (**r**) ($n = 2{,}183$, 5,648, and 79 differentially spliced events, respectively), unpublished collaborative cohort 2 (**s**) ($n = 558$, 1,926, and 94 differentially spliced events, respectively) and Leucegene cohort (**t**) ($n = 2{,}571$, 787, and 122 differentially spliced events, respectively) are ranked by *y*-axis and class of event (PSI and *P* values adjusted for multiple comparisons were calculated using PSI-Sigma). **u**, **w**, Representative Sashimi plots of RNA-seq data showing the intron retention events in *REC8* (**u**) and *PHF6* (**q**) from the TCGA dataset. **v**, **x**, PSI values for intron retention events in *REC8* (**v**) and *PHF6* (**x**) in normal PBMNCs (GSE58335[41]), BMMNCs (GSE61410[41]), cord blood $CD34^{+}$ cells (GSE48846[42]), and AML samples with indicated genotypes (the line represents the median, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; samples below 2.5th percentile and above 97.5th percentile are shown as dots; PSI and *P* values adjusted for multiple comparisons were calculated using PSI-Sigma; one-way ANOVA with Tukey's multiple comparison test; $^{*}P < 0.05$; $^{**}P < 0.01$; $^{***}P < 0.001$). **y**, Volcano plots of aberrant splicing events in TCGA AML data comparing *SRSF2* single-mutant and *IDH2*/*SRSF2* double-mutant AML ($n = 122$ differentially spliced events; PSI and *P* values adjusted for multiple comparisons were calculated using PSI-Sigma; $|\Delta\text{PSI}| > 10\%$ and $P < 0.01$ were used as thresholds).



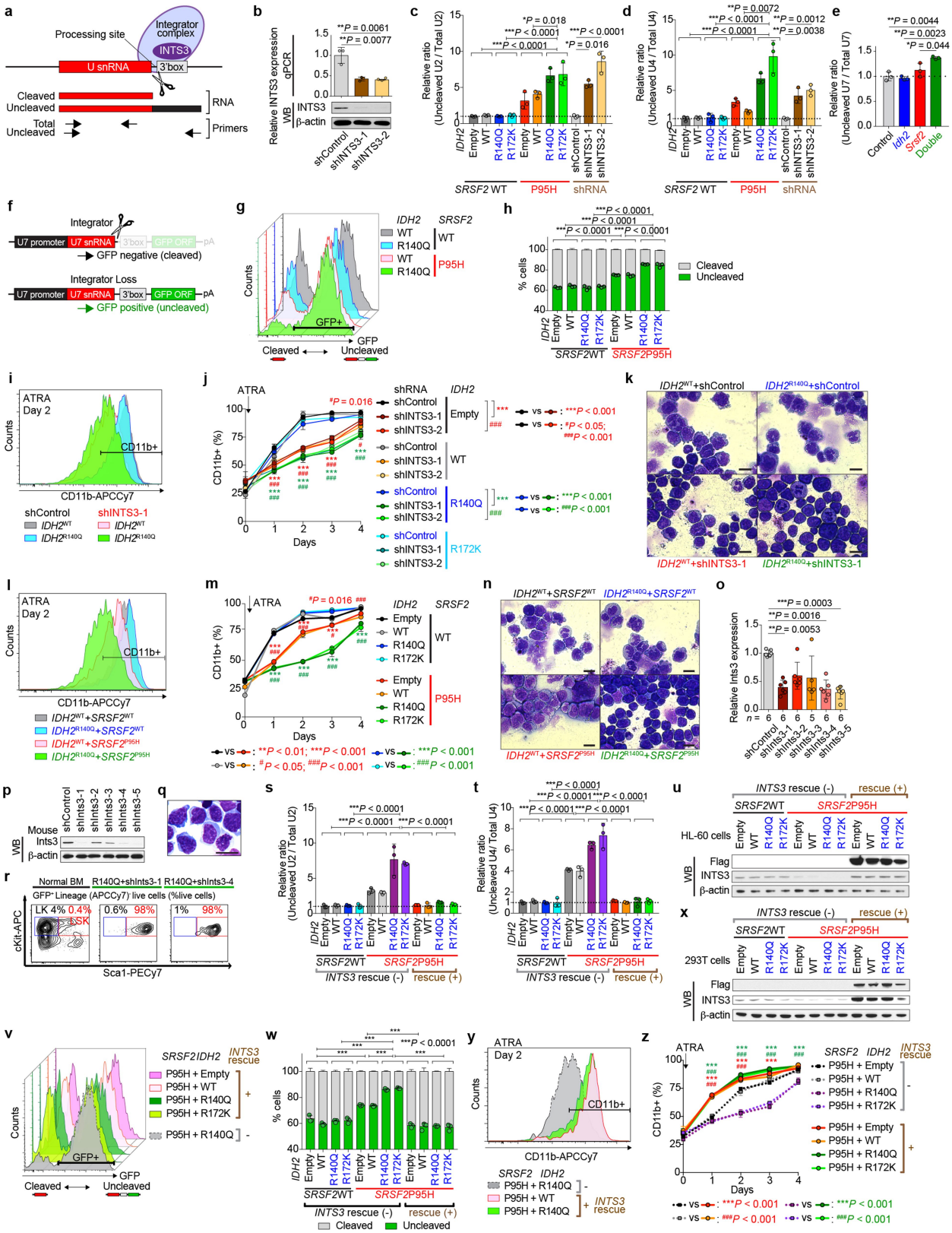
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Aberrant *INTS3* transcripts undergo nonsense-mediated decay, and effect of INTS3 loss extends to other members of the integrator complex. **a**, Representative Sashimi plots of RNA-seq data from the TCGA showing intron retention in *INTS3*. **b**, **c**, PSI values for *INTS3* exon 5 skipping (**b**) and intron 4 retention (**c**) in normal PBMNC (GSE58335[41]), BMMNC (GSE61410[41]), cord blood $CD34^+$ cells (GSE48846[42]) and AML samples with indicated genotypes (the number of RNA-seq samples analysed is indicated; PSI and *P* values adjusted for multiple comparisons were calculated using PSI-Sigma; the line represents mean, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; samples below 2.5th percentile and above 97.5th percentile are shown as dots; one-way ANOVA with Tukey's multiple comparison test). **d**, Sanger sequencing of cDNA showing wild-type or mutant *SRSF2* expression in isogenic K562 knock-in cells. $^{\#}$a nonsynonymous mutation that alters P95; $^{\#\#}$a synonymous mutation that does not change the amino acid. **e**, RT–PCR and western blot analysis of INTS3 in isogeneic HL-60 cells with various combinations of *IDH2* and *SRSF2* mutations. IR: intron retention; ES: exon skipping. Representative results from three biologically independent experiments with similar results. **f**, RT–PCR and western blot of INTS3 in non-isogenic myeloid leukaemia cell lines. *SRSF2* genotypes are shown together (representative results from three independent experiments with similar results). **g**, Western blot analysis of K562 $SRSF2^{P95H}$ knock-in cells transduced with shRNAs against *UPF1* (representative results from three biologically independent experiments with similar results). **h**, Primers used to specifically measure *INTS3* isoform with intron 4 retention and exon 5 skipping, and those for the normal *INTS3* isoform. **i**, **j**, Half-life of *INTS3* transcripts with exon 5 skipping (**i**) and intron 4 retention (**j**) were measured by qPCR ($n = 3$; mean ± s.d.; a two-sided Student's *t*-test). **k**, **l**, Western blot analysis of protein lysates of samples from patients with AML with the indicated *IDH2* and *SRSF2* genotypes (**k**). Expression level of each integrator subunit was quantified using ImageJ and relative expression levels are shown in **l**, in which the mean expression levels of control samples were set as 1 ($n = 6$ for control, *IDH2* single-mutant, and *SRSF2* single-mutant AML, and $n = 7$ for *IDH2* and *SRSF2* double-mutant AML; detailed information of the primary patient samples used for this analysis is provided in Supplementary Table 23; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **m**, Western blot analysis of protein lysates from isogenic K562 cells with indicated *IDH2* and *SRSF2* genotypes (left) or with INTS3 knockdown (right) (representative results from three biologically independent experiments are shown). **n**, Western blot analysis of murine Lin^-KIT^+ BM cells at 12 weeks post-pIpC based on *Idh2* and *Srsf2* mutant genotypes. Expression level of INTS3 was quantified using ImageJ and relative expression levels are shown below; $n = 2$ mice per genotype were analysed. **o**, Correlation among indicated Integrator subunits and *P* value were calculated in Excel and R^2 values are visualized as a heat map generated by Prism 7 (top). Correlation between INTS3 and INTS9 protein expression is shown (bottom) ($n = 25$ from **k**; the Pearson correlation coefficient (R^2) and *P* values (two-tailed) were calculated in Excel). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



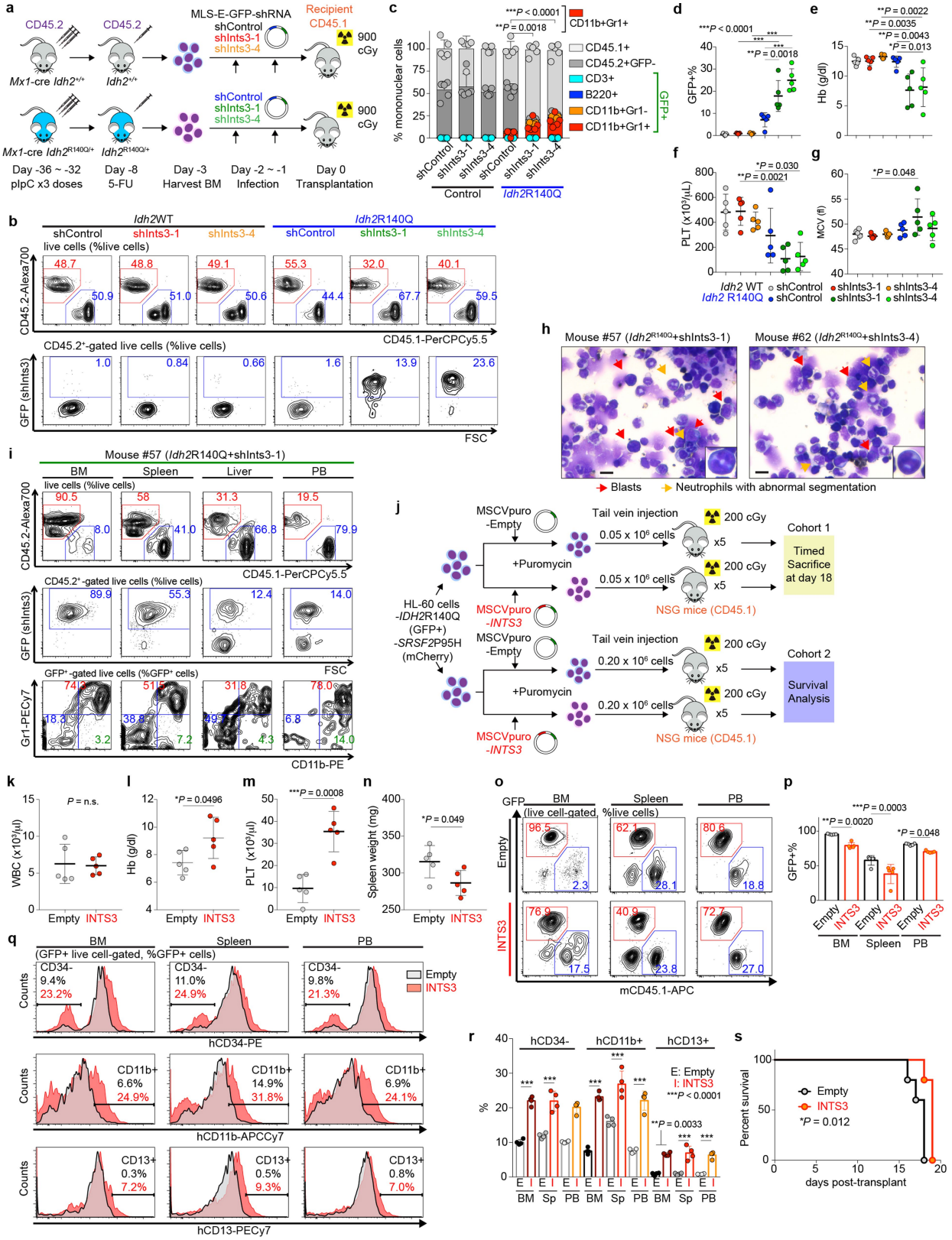
Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | DNA hypermethylation at *INTS3* enhances *INTS3* mis-splicing, which is associated with RNAPII stalling. **a**, Sequence of human *INTS3* exon 4, intron 4 and exon 5, and schematic of *INTS3* minigene constructs. GG(A/U)G motifs, (C/G)C(A/U)G motifs, and CG dinucleotides are highlighted in blue, red, and green, respectively. **b**, Schematic of *INTS3* minigene constructs. **c**, Table revealing the number of GGNG or CCNG motifs in exon 4, entire cDNA of *INTS3* or entire genomic DNA (gDNA) of *INTS3* per 100 nucleotides. **d**–**i**, Radioactive RT–PCR results of *INTS3* minigene assays using indicated versions of the minigene in isogenic K562 cells. Percentage of intron 4 retention were normalized against exogenous eGFP ($n = 3$; mean percentage ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **j**, Mean percentage of methylated CpGs at *ARID3A* in samples from patients with AML with indicated genotypes determined by eRRBS ($n = 3$ patients per genotype), followed by IGV plots of RNA-seq data of *ARID3A* from the TCGA. **k**, Results of eRRBS ($n = 1$ per genotype) and RNAPII-Ser2P ChIP-walking experiments are represented as shown in Fig. 3f ($n = 3$; mean percentage ± s.d.; two-way ANOVA with Tukey's multiple comparison test). **l**, **m**, RT–PCR results detecting *INTS3* intron retention in isogenic K562 cells containing various combinations of *IDH2* and *SRSF2* mutations that were treated with cell-permeable 2HG at 0.5 μM (**l**) or 5-AZA-CdR at 5 μM (**m**) for 8 days (representative results from three biologically independent experiments with similar results). **n**, RNAII pausing index in isogenic $SRSF2^{WT}$ or $SRSF2^{P95H}$ mutant K562 cells was calculated as previously described[20] as a ratio of normalized ChIP–seq reads of RNAPII-Ser5P on TSSs (±250 bp) over that of the corresponding bodies (+500 to +1,000 from TSSs) (the line represents the median, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; each box plot was made by analysing ChIP–seq data from one cell line; two-sided Student's *t*-test). **o**, Metagene plots showing genome-wide RNAPII-Ser5P occupancy in primary samples from patients with AML with indicated genotypes (patient samples used for this analysis are described in Supplementary Table 23). **p**, **q**, RNAPII occupancy representing ChIP–seq reads of RNAPII-Ser2P over gene bodies was calculated for isogenic K562 cells (**p**) and AML samples (**q**) (the line represents the median, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; each box plot was made by analysing ChIP–seq data from one cell line (**p**) or one primary AML sample (**q**); two-sided Student's *t*-test (**p**) and one-way ANOVA with Tukey's multiple comparison test (**q**)). **r**, **s**, Genome browser view of ChIP–seq signal for RNAPII Ser5P at *INTS5* (**r**) and *INTS14* (**s**) in isogenic K562 cells with or without *SRSF2* mutation ($n = 1$) and primary AML samples with indicated genotype (results generated from $n = 2$ primary AML samples are shown). **t**, RNAPII abundance over the differentially spliced regions between *IDH2* and *SRSF2* wild-type control and *SRSF2* single-mutant AML determined by RNAPII-Ser2P ChIP–seq (*y* axis, $\log_2$(counts per million); the line represents the median, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; each box plot was made by analysing ChIP–seq data from one primary AML sample; one-way ANOVA with Tukey's multiple comparison test). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



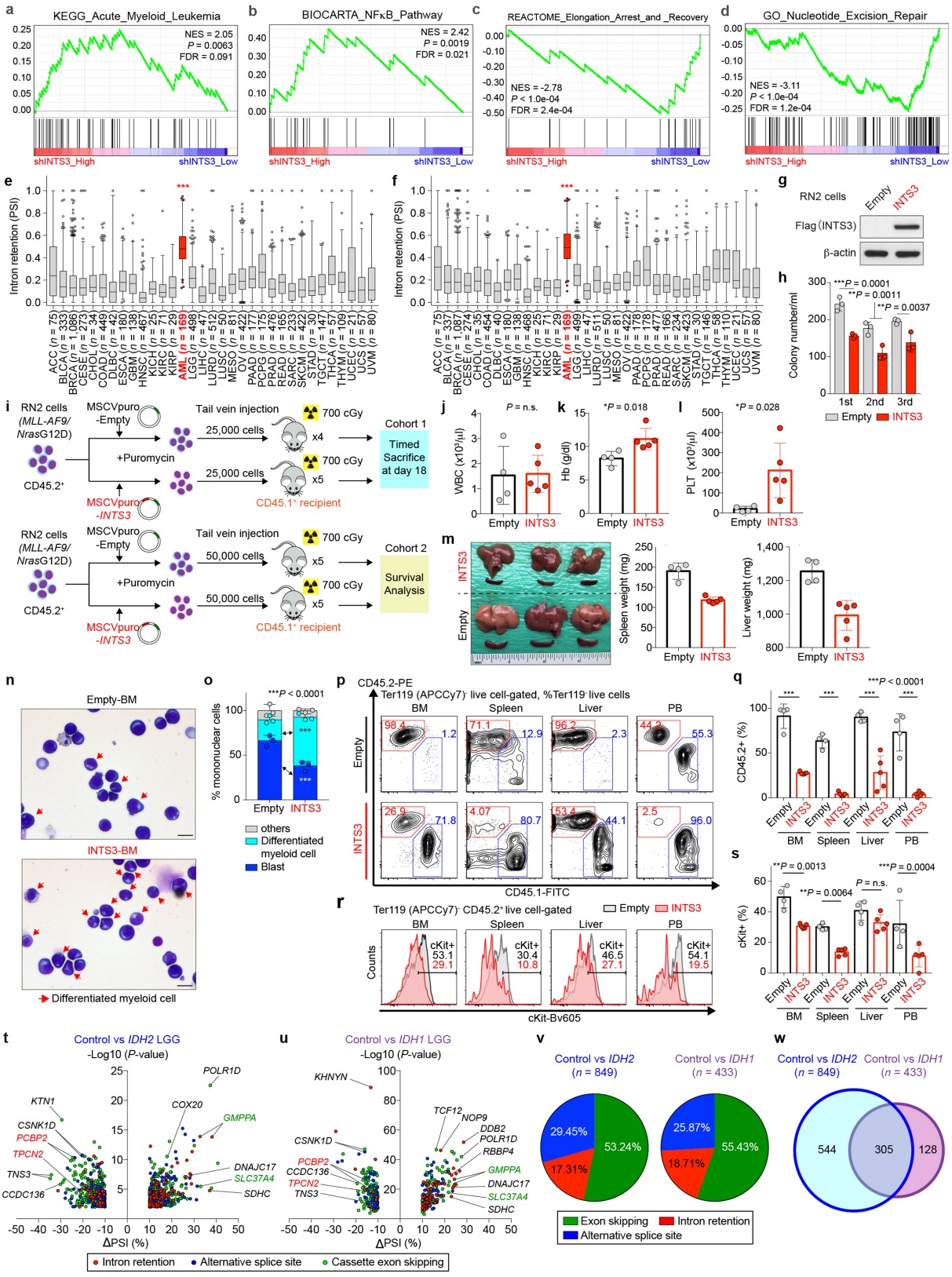
Extended Data Fig. 8 | See next page for caption.

Extended Data Fig. 8 | Loss of INTS3 impairs uridine-rich small nuclear RNA processing and blocks myeloid differentiation. **a**, Schematic of snRNA processing site and qPCR primers for detecting cleaved or uncleaved snRNA. **b**, qPCR (top; $n = 3$; mean ± s.d.; a two-sided Student's *t*-test) and representative western blot of INTS3 in HL-60 cells transduced with shRNAs targeting human *INTS3* (bottom, representative results from three biologically independent experiments). **c**–**e**, **s**, **t**, qPCR results of U2 (**c**, **s**) and U4 (**d**, **t**) snRNAs in isogenic HL-60 cells and U7 snRNA in murine cells from Extended Data Fig. 6n (**e**). Ratio of uncleaved/total snRNAs expression was compared ($n = 3$, mean ratio ± s.d.; one-way ANOVA with Tukey's multiple comparison test; the largest *P* values calculated among 2 × 2 comparisons of two components from different groups are shown. For example, *P* values were calculated from the following four comparisons; bars 1 versus 3, 2 versus 3, 1 versus 4, 2 versus 4). **f**, Schematic of the U7 snRNA–GFP reporter. **g**, **v**, Flow cytometry analysis of 293T cells transduced with U7 snRNA-GFP reporter and *IDH2*, *SRSF2* and *INTS3* constructs as labelled on the right (representative results from three biologically independent experiments are shown). **h**, **w**, Quantification of per cent GFP^- and GFP^+ 293T cells ($n = 3$ biologically independent experiments, mean percentage ± s.d.; one-way ANOVA with Tukey's multiple comparison test; *P* values are shown as in **c**). **i**, **l**, **y**, Flow cytometry analysis of CD11b expression in isogenic HL-60 cells after ATRA treatment for two days (representative results from three biologically independent experiments are shown). **j**, **m**, **z**, Quantification of percentages of $CD11b^+$ HL-60 cells over time ($n = 3$; mean percentage ± s.d.; two-way ANOVA with Tukey's multiple comparison test). **k**, **n**, Cytomorphology of isogenic HL-60 cells after ATRA treatment for two days (Giemsa staining; scale bar, 10 μm; original magnification, ×400; representative results from three biologically independent experiments are shown). **o**, **p**, qPCR of *Ints3* (**o**) (mean ± s.d.; Kruskal–Wallis tests with uncorrected Dunn's test) and western blot of INTS3 (**p**) in Ba/F3 cells transduced with shRNAs targeting mouse *Ints3*. **q**, **r**, Representative cytomorphology (**q**) and immunophenotype (**r**) of colony cells at the sixth colony. Normal BMMNCs were used as a control (the percentage listed represent the percent of cells within live cells; representative results from three biologically independent experiments are shown). **u**, **x**, Western blot of proteins extracted from HL-60 cells (**u**) assayed in **s**-**t** and **y**-**z** and 293T cells (**x**) assayed in **v** and **w** (representative results from three biologically independent experiments). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{\#}P < 0.05$, $^{\#\#}P < 0.01$, $^{\#\#\#}P < 0.001$.



Extended Data Fig. 9 | See next page for caption.

Extended Data Fig. 9 | Mutant Idh2 cooperates with Ints3 loss to generate a lethal myeloid neoplasm in vivo. a, Schematic of shRNA targeting *Ints3* (shInts3) retroviral BM transplantation model. **b**, Flow cytometry data showing the chimerism of $CD45.2^+$ versus $CD45.1^+$ (top) or GFP^+ (bottom) cells in peripheral blood at four weeks post-transplant (the percentages listed represent the percent of cells within live cells; representative results from five recipient mice). **c**, Composition of PBMNCs at four weeks post-transplant ($n = 5$ per group; mean + s.d.; represented by lines above the box. statistical significance was detected in percentage of $CD11b^+Gr1^+$ cells; by two-way ANOVA with Tukey's multiple comparison test). **d**–**g**, Chimerism of GFP^+ cells in peripheral blood (**d**) and blood counts of recipients at four weeks post-transplant (Hb (**e**); PLT (**f**); MCV (**g**); $n = 5$ per group; mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **h**, Giemsa staining of BMMNCs from moribund mice with indicated genotypes (red and yellow arrows represent blastic cells and dysplastic neutrophils, respectively; inset, representative neutrophils with abnormal segmentation; scale bar, 10 μm; original magnification, ×400; representative results from five mice per genotype). **i**, Flow cytometry data of BM, spleen, liver, and peripheral blood from $Idh2^{R140Q}$ mice treated with shInts3 (representative results from five mice). **j**, Schematic of HL-60 xenograft model in which recipient mice from cohort 1 were euthanized at day 18 post-transplant and mice from cohort 2 were observed for survival analysis until end stage. **k**–**n**, Blood counts (WBC (**k**); Hb (**l**); PLT (**m**)) and spleen weight (**n**) of mice from cohort 1 at day 18 post-transplant (mean ± s.d.; $n = 5$ per group; a two-sided Student's *t*-test). **o**, **p**, Representative flow cytometry data of BM, spleen, and peripheral blood from the recipient mice from cohort 1 (**o**) (the percentage represents the percent of cells within live cells) and the mean percentage of GFP^+ cells (**p**) ($n = 5$ per group; mean + s.d.; two-way ANOVA with Sidak's multiple comparison test). **q**, **r**, Representative flow cytometry data of BM, spleen and peripheral blood from cohort 1 (**q**) (the percentage represents the percent of cells within GFP^+ live cells) and the mean percentage of $hCD34^-$, $hCD11b^+$ and $hCD13^+$ cells (**r**) ($n = 4$ per group; mean + s.d.; two-way ANOVA with Sidak's multiple comparison test). **s**, Kaplan–Meier survival analysis of recipient mice from cohort 2 ($n = 5$ per group; log-rank (Mantel-Cox) test (two-sided)). $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$.



Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Gene expression and biological consequences of INTS3 loss, and effect of *IDH1* and *IDH2* mutations on splicing in low-grade glioma. a–**d**, GSEA based on RNA-seq data generated from isogenic $IDH2^{R140Q}$ mutant HL-60 cells with or without INTS3 depletion. Representative results from gene sets associated with leukaemogenesis and myeloid differentiation (**a**), oncogenic signalling pathways (**b**), RNAPII elongation-linked transcription (**c**) and DNA damage response (**d**) with statistical significance ($P < 0.01$) are shown (*y* axis; enrichment score; NES: normalized enrichment score; FDR: false discovery rate; RNA-seq data generated from isogenic HL-60 cells in duplicate were analysed using GSEA[34]). **e**, **f**, PSI values for *INTS3* intron 4 (**e**) and 5 (**f**) retention events across 33 cancer cell types (the same datasets were analysed in Fig. 4f). ACC,: adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; CHOL, cholangiocarcinoma; DLBC, diffuse large B-cell lymphoma; ESCA, oesophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; KICH, kidney chromophobe; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LGG, low-grade glioma; LIHC, liver hepatocellular carcinoma; LUSC, lung squamous cell carcinoma; MESO, mesothelioma; OV, ovarian serous cystadenocarcinoma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; TGCT, testicular germ cell tumours; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma. The line represents the median, box edges show 25th and 75th percentiles and whiskers represent 2.5th and 97.5th percentiles; samples below 2.5th percentile and above 97.5th percentile are shown as dots; one-way ANOVA with Dunnett's multiple comparison test; $***P < 0.001$ represents the *P* values from all the comparisons between AML and any of other 32 non-AML cancer type. **g**, Western blot analysis confirming overexpression of 3× Flag-tagged INTS3 in RN2 (*MLL-AF9* $Nras^{G12D}$) leukaemia cells (representative results from three biologically independent experiments). **h**, Colony numbers from serial replating assays of RN2 cells with or without INTS3 overexpression ($n = 3$; mean + s.d. represented by lines above the box; two-way ANOVA with Sidak's multiple comparison test). **i**, Schematic of INTS3 retroviral BM transplantation models in which recipient mice from cohort 1 were euthanized at day 18 post-transplant and mice from cohort 2 were observed for survival analysis until end-stage. **j**–**l**, Blood counts (WBC (**j**); Hb (**k**); PLT (**l**)) of mice from cohort 1 at day 18 post-transplant (mean ± s.d.; $n = 4$ ('empty' group); $n = 5$ ('INTS3' group) recipient mice; a two-sided Student's *t*-test). **m**, Representative photograph of spleens and livers from cohort 1 with an inch scale (left), and spleen (middle) and liver weight (right) ($n = 4$ (empty); $n = 5$ (INTS3); mean ± s.d.; two-sided Student's *t*-test). **n**, **o**, Representative Giemsa staining (**n**) (red arrows represent differentiated cells; scale bar, 10 μm; original magnification, ×400) and percentages of blasts, differentiated myeloid cells, and other cells in BMMNCs (**o**) from moribund mice from cohort 2 ($n = 3$ per genotype; 100 cells per mouse were classified; mean percentage + s.d.; two-way ANOVA with Sidak's multiple comparison test). **p**, **q**, Representative flow cytometry analysis of BM, spleen, liver, and peripheral blood (**p**) and percentages of $CD45.2^+$ cells in $Ter119^-$ live cells (**q**) in recipient from cohort 1 ($n = 4$ (empty); $n = 5$ (INTS3); mean ± s.d.; two-way ANOVA with Tukey's multiple comparison test). **r**, **s**, Representative flow cytometry analysis showing KIT expression in RN2 cells with or without INTS3 overexpression (**r**) and quantification of KIT^+ cells (**s**) from cohort 1 ($n = 4$ (Empty); $n = 5$ (INTS3); mean ± s.d.; one-way ANOVA with Tukey's multiple comparison test). **t**, **u**, Volcano plots of aberrant splicing events in the LGG TCGA dataset based on *IDH2* (**t**) or *IDH1* (**u**) mutant genotypes. $|\Delta PSI| > 10\%$ and $P < 0.01$ were used as thresholds ($n = 849$ and $n = 433$ differentially spliced events, respectively; RNA-seq data were analysed using PSI-Sigma). **v**, Percentage of each class of alternative splicing event in *IDH2* (left) and *IDH1* (right) mutant LGG is shown in pie-chart. **w**, Venn diagram of numbers of alternatively spliced events from the LGG TCGA dataset based on *IDH1* and *IDH2* mutant genotypes. 'Control' represents LGG with wild-type *IDH1* and *IDH2*. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$.

nature research

Corresponding author(s): Omar Abdel-Wahab

Last updated by author(s): Jul 17, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No code was used for data collection.
Data analysis	The inclusion ratios of alternative exons or introns were estimated by using PSI-Sigma. GraphPad Prism 7 software was used to analyze the data and to make figures. RNA-seq reads were aligned by using 2-pass STAR 2.5.2a. Samtools (1.3.1) were used to generate variant call format (VCF) files for 6 target genes: IDH1, IDH2, SF3B1, SRSF2, U2AF1, and ZRSR2 with mpileup parameters (-Bvu). The VCF files were further processed by our in-house scripts to filter out mutations whose VAF was lower than 15%. The filtered VCF files were used for variant effect predictor (version 89.4) to annotate the consequences of the mutations. Motif analysis was done by using MEME SUITE. The heatmaps and sample clustering were done by using MORPHEUS (software.broadinstitute.org/morpheus/). VCF files from the TruSightTM Myeloid 54 gene panel from Illumina were analyzed using Illumina's Variant Studio software while those from a 40 gene panel (Oncomine Myeloid Research Assay; ThermoFisher), processing eight samples per Ion 530 chip on the IonTorrent platform were analyzed using the Ion Reporter software. ChIP-seq reads were mapped to the genome by calling Bowtie v1.0.048 with the arguments '-v 2 -k 1 -m 1 --best --strata'. Peaks were called using MACS2 v2.1.1.2016030952 against input control libraries with P < 1e-5 and subsequently filtered to remove peaks contained within ENCODE blacklisted regions and the mitochondrial genome. Subsequent data analysis was performed with Bioconductor in the R programming environment. Consensus peaks between samples were called using the soGGI package v1.14.0. Peaks were annotated using the ChIPseeker package v1.18.0.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. The RNA sequencing data have been deposited in NCBI Sequencing Read Archive (SRA) under accession number SRP133673.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	For the in vivo experiments, the number of mice in each experiment was chosen to provide 90% statistical power with a 5% error level. For the RNA-seq experiments, maximal sample sizes available were used to obtain statistical power as much as possible to detect significant splicing alterations.
Data exclusions	No data were excluded from the analyses.
Replication	The experiments were repeated at a minimum of 3 times for all the in vitro experiments. All attempts at replication were successful.
Randomization	Animals were assigned to experimental group based on genotype and there was no drug treatment groups, therefore randomization was not utilized.
Blinding	For survival and blood count analyses of mice, actual measurements were carried out by a member of the lab who did not have knowledge of which alleles were expected to alter survival or blood count parameters. All other experiments were not blinded and it was not necessary to be as they were less subjective.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Flow cytometry antibodies: B220-APCCy7 (clone: RA3-6B2; purchased from BioLegend; catalog #: 103224; dilution: 1:200); B220-Bv711 (RA3-6B2; BioLegend; 103255; 1:200); CD3-PerCPCy5.5 (17A2; BioLegend; 100208; 1:200); CD3-APC (17A2; BioLegend; 100236; 1:200); CD3-APCCy7 (17A2; BioLegend; 100222; 1:200); Gr1-PECy7 (RB6-8C5; eBioscience; 25-5931-82; 1:500); CD11b-PE (M1/70; eBioscience; 12-0112-85; 1:500); CD11b-APCCy7 (M1/70; BioLegend; 101226; 1:200); CD11c-APCCy7 (N418; BioLegend; 117323; 1:200); NK1.1-APCCy7 (PK136; BioLegend; 108724; 1:200); Ter119-APCCy7 (BioLegend; 116223: 1:200); cKit-APC (2B8; BioLegend; 105812; 1:200); cKit-PerCPCy5.5 (2B8; BioLegend; 105824; 1:100); cKit-Bv605 (ACK2; BioLegend; 135120; 1:200); Sca1-PECy7 (D7; BioLegend; 108102; 1:200); CD16/CD32 (FcγRII/III)-Alexa700 (93; eBioscience; 56-0161-82; 1:200); CD34-FITC (RAM34; BD Biosciences; 553731; 1:200); CD45.1-FITC (A20; BioLegend; 110706; 1:200); CD45.1-PerCPCy5.5 (A20; BioLegend; 110728; 1:200); CD45.1-PE (A20; BioLegend; 110708; 1:200); CD45.1-APC (A20; BioLegend; 110714; 1:200); CD45.2-

PE (104; eBioscience; 12-0454-82; 1:200); CD45.2-Alexa700 (104; BioLegend; 109822; 1:200); CD45.2-Bv605 (104; BioLegend; 109841; 1:200); CD48-Bv711 (HM48-1; BioLegend; 103439; 1:200); CD150 (9D1; eBioscience; 12-1501-82; 1:200); CD34-PerCP (8G12; BD Biosciences; 345803; 1:200); CD117-PECy7 (104D2; eBioscience; 25-1178-42; 1:200); CD33-APC (P67.6; BioLegend; 366606; 1:200); HLA-DR-FITC (L243; BioLegend; 307604; 1:200); CD13-PE (L138; BD Biosciences; 347406; 1:200); CD45-APC-H7 (2D1; BD Biosciences; 560178; 1:200).
Western blotting, DNA dot blot assays, and ChIP: INTS1 (purchased from Bethyl laboratories; catalog #: A300-361A; dilution: 1:1,000), INTS2 (Abcam; ab74982; 1:1,000), INTS3 (Bethyl laboratories; A300-427A; 1:1,000, Abcam; ab70451; 1:1,000), INTS4 (Bethyl laboratories; A301-296A; 1:1,000), INTS5 (Abcam; ab74405; 1:1,000), INTS6 (Abcam; ab57069; 1:1,000), INTS7 (Bethyl laboratories; A300-271A; 1:1,000), INTS8 (Bethyl laboratories; A300-269A; 1:1,000), INTS9 (Bethyl laboratories; A300-412A; 1:1,000), INTS11 (Abcam; ab84719; 1:1,000), Flag-M2 (Sigma-Aldrich; F-1084; 1:1,000), Myc-tag (Cell Signaling; 2276S; 1:1,000), β-actin (Sigma-Aldrich; A-5441; 1:2,000), 5-Hydroxymehylcytosine (5hmC) (Active motif; 39769), RNA polymerase II CTD repeat YSPTSPS (phospho S2) (Abcam; ab5095), RNA polymerase II CTD repeat YSPTSPS (phospho S5) (Abcam; ab5408), and UPF1 (Abcam; ab109363; 1:1,000).

Validation

All antibodies were validated by the supplier for human samples, and were checked in the lab by Western blotting on cell lysate and by comparing to the manufacturer's or in-house results.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)

K562, HL-60, TF1, and HEK293T cells were obtained from the American Type Culture Collection (ATCC) . K052 cells were obtained from JCRB Cell Bank. Ba/F3 cells were obtained from DSMZ. 293 GPII cells were purchased from Clontech. The isogenic K562 cell lines with or without SRSF2 P95H were generated at Horizon Discovery. MLL-AF9/NrasG12D murine leukemia (RN2) cells were obtained from Dr. Iannis Aifantis (NYU School of Medicine).

Authentication

An aliquot of each cell lines were authenticated using ATCC/JCRB/DSMZ DNA fingerprinting.

Mycoplasma contamination

All cell lines are frequently tested for mycoplasma contamination. Cell lines used in this study were verified to be mycoplasma negative before undertaking any experiments with them.

Commonly misidentified lines
(See ICLAC register)

No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals

6-8 week female CD45.1 C57BL/6 mice were purchased from The Jackson Laboratory (Stock No: 002014) . Male and female CD45.2 Srsf2P95H/+ conditional knock-in mice, Idh2R140Q/+ conditional knock-in mice, and Tet2 conditional knockout mice (all on C57BL/6 background) were also analyzed and used as bone marrow donors.

Wild animals

The study did not involve wild animals.

Field-collected samples

The study did not involve samples collected from the field.

Ethics oversight

All animal procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics

The covariate-relevant population characteristics of the human research participants of Memorial Sloan Kettering Cancer Center (MSKCC), Université Paris-Saclay, and the University of Manchester are provided below: Samples were obtained from acute myeloid leukemia (AML) patients treated at MSKCC, Université Paris-Saclay, and the University of Manchester. All samples were viably frozen and used to extract DNA, RNA, and protein. All subjects with AML were eligible for inclusion regardless of age, sex, or race.

Recruitment

All the participants were recruited without knowing their genotypes. Samples were genotyped and classified based on IDH2/SRSF2 genotypes. Samples that had mutations in IDH1, SF3B1, U2AF1, or ZRSR2 were excluded from RNA-seq and targeted RNA and protein analyses. Then RNA-seq was performed to analyze the splicing alterations. Therefore, there was no self-selection bias and it is unlikely that bias, if any, impacted the splicing analysis.

Ethics oversight

Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center (under MSK IRB protocol 06-107), Université Paris-Saclay (under declaration DC-200-725 and authorization AC-2013-1884), and the University of Manchester (institution project approval 12-TISO-04), and conducted in accordance with the Declaration of Helsinki protocol. Written informed consent was obtained from all participants. Manchester samples were retrieved from the Manchester Cancer Research Centre Haematological Malignancy Tissue Biobank, which receives sample donations from all consenting leukemia patients presenting to The Christie Hospital (REC Reference 07/H1003/161+5; HTA license 30004; instituted with approval of the

South Manchester Research Ethics Committee). Patient samples were anonymized by the Hematologic Oncology Tissue Bank of MSK, Biobank of Gustave Roussy, and the Manchester Cancer Research Centre Haematological Malignancy Tissue Biobank.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	The ChIP-seq data (and RNA-seq data and eRRBS data) have been deposited in NCBI Sequence Read Archive (SRA) under accession number SRP133673.
Files in database submission	Sample_WT-WT-614, Sample_WT-WT-247, Sample_RQ-WT-584, Sample_RQ-WT-475, Sample_WT-PH-547, Sample_WT-PH-343, Sample_RQ-PH-524, Sample_RQ-PH-475, k562_WT_input, k562_WT_si_PolII-Ser2, k562_WT_si_PolII-Ser5, k562_p95_input, k562_p95H_si_PolII-Ser2, k562_p95H_si_PolII-Ser5
Genome browser session (e.g. UCSC)	The ChIP-seq data (and RNA-seq data and eRRBS data) have been deposited in NCBI Sequence Read Archive (SRA) under accession number SRP133673.

Methodology

Replicates	Two primary AML patient samples per IDH2/SRSF2 genotype (IDH2/SRSF2 WT/WT, Mutant/WT, WT/Mutant. and Mutant/Mutant) were used for the ChIP-seq experiments.
Sequencing depth	An average of 75 million paired reads was generated per sample (125 bp single-end).
Antibodies	Antibodies used for ChIP were as follows: RNA polymerase II CTD repeat YSPTSPS (phospho S2) (Abcam; ab5095) RNA polymerase II CTD repeat YSPTSPS (phospho S5) (Abcam; ab5408)
Peak calling parameters	Narrow peaks were called using the callpeak function from MACS2 v2.1.1.20160309 against matched input samples, using default parameters and a P-value cutoff of 1e-5, according to the ENCODE Histone ChIP-seq Data Standards and Processing Pipeline (https://www.encodeproject.org/chip-seq/histone/).
Data quality	For all samples, a P-value cutoff of 1e-5 against input was used. All peaks were called at a q-value of < 0.017. For each sample, the number of peaks with a fold-change > 5, and the average total number of peaks called is 19,200.
Software	ChIP-seq reads were mapped to the genome by calling Bowtie v1.0.048 with the arguments '-v 2 -k 1 -m 1 --best --strata'. Peaks were called using MACS2 v2.1.1.2016030952 against input control libraries with P < 1e-5 and subsequently filtered to remove peaks contained within ENCODE blacklisted regions and the mitochondrial genome. Subsequent data analysis was performed with Bioconductor in the R programming environment. Consensus peaks between samples were called using the soGGI package v1.14.0. Peaks were annotated using the ChIPseeker package v1.18.0.

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	Surface-marker staining of hematopoietic cells was performed by first lysing cells with ACK lysis buffer and washing cells with ice-cold PBS. Cells were stained with antibodies in PBS/2% BSA for 30 minutes on ice. For hematopoietic stem/progenitor staining, cells were stained with the following antibodies: B220-APCCy7 (clone: RA3-6B2; purchased from BioLegend; catalog #: 103224; dilution: 1:200); B220-Bv711 (RA3-6B2; BioLegend; 103255; 1:200); CD3-PerCPCy5.5 (17A2; BioLegend; 100208; 1:200); CD3-APC (17A2; BioLegend; 100236; 1:200); CD3-APCCy7 (17A2; BioLegend; 100222; 1:200); Gr1-PECy7 (RB6-8C5; eBioscience; 25-5931-82; 1:500); CD11b-PE (M1/70; eBioscience; 12-0112-85; 1:500); CD11b-APCCy7 (M1/70; BioLegend; 101226; 1:200); CD11c-APCCy7 (N418; BioLegend; 117323; 1:200); NK1.1-APCCy7 (PK136; BioLegend; 108724; 1:200); Ter119-APCCy7 (BioLegend; 116223: 1:200); cKit-APC (2B8; BioLegend; 105812; 1:200); cKit-PerCPCy5.5 (2B8; BioLegend; 105824; 1:100); cKit-Bv605 (ACK2; BioLegend; 135120; 1:200); Sca1-PECy7 (D7; BioLegend; 108102; 1:200); CD16/CD32 (FcγRII/III)-Alexa700 (93; eBioscience; 56-0161-82; 1:200); CD34-FITC (RAM34; BD Biosciences; 553731; 1:200); CD45.1-FITC (A20; BioLegend; 110706;

	1:200); CD45.1-PerCPCy5.5 (A20; BioLegend; 110728; 1:200); CD45.1-PE (A20; BioLegend; 110708; 1:200); CD45.1-APC (A20; BioLegend; 110714; 1:200); CD45.2-PE (104; eBioscience; 12-0454-82; 1:200); CD45.2-Alexa700 (104; BioLegend; 109822; 1:200); CD45.2-Bv605 (104; BioLegend; 109841; 1:200); CD48-Bv711 (HM48-1; BioLegend; 103439; 1:200); CD150 (9D1; eBioscience; 12-1501-82; 1:200). DAPI was used to exclude dead cells. For sorting human leukemia cells, cells were stained with a lineage cocktail including CD34-PerCP (8G12; BD Biosciences; 345803; 1:200); CD117-PECy7 (104D2; eBioscience; 25-1178-42; 1:200); CD33-APC (P67.6; BioLegend; 366606; 1:200); HLA-DR-FITC (L243; BioLegend; 307604; 1:200); CD13-PE (L138; BD Biosciences; 347406; 1:200); CD45-APC-H7 (2D1; BD Biosciences; 560178; 1:200). The composition of mature hematopoietic cell lineages in the BM, spleen and peripheral blood was assessed using a combination of CD11b, Gr1, B220, and CD3. For the hematopoietic stem and progenitor analysis, a combination of CD11b, CD11c, Gr1, B220, CD3, NK1.1, and Ter119 was stained as lineage-positive cells.
Instrument	All the FACS sorting was performed on FACS Aria, and analysis was performed on an LSRII or LSR Fortessa (BD Biosciences).
Software	FlowJo Ver.9 was used for analysis of flow cytometry data.
Cell population abundance	To check the purity of GFP and/or mCherry positivity in post-sort samples, the sorted samples were analyzed for GFP and/or mCherry by FACS Aria (BD Biosciences), and samples with >99% purity were used for analyses.
Gating strategy	The FSC/SSC gates of the starting cell population was set in order to include all the lineages of mouse hematopoietic cells such as granulocytes, monocytes, and lymphocytes. Then doublet cells were excluded by SSC-H vs SSC-W and FSC-H vs FSC-W gating. The boundaries between “positive” and “negative” staining cell population were defined by using unstained and single color-stained controls that were prepared by staining the whole BM mononuclear cells from B6 mice at 8-12 weeks with antibodies against mouse CD11b. “Positive” staining cell population was defined as CD11b+ population and “negative” staining cell population was defined by unstained control. The boundary for each fluorescence was set between these “positive” and “negative” staining cell populations.

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

LETTER

https://doi.org/10.1038/s41586-019-1646-9

Spliceosomal disruption of the non-canonical BAF complex in cancer

Daichi Inoue[1,2,10], Guo-Liang Chew[3,4,10], Bo Liu[1], Brittany C. Michel[5,6,7], Joseph Pangallo[3,4], Andrew R. D'Avino[6,7], Tyler Hitchman[1], Khrystyna North[3,4,8], Stanley Chun-Wei Lee[1], Lillian Bitner[1], Ariele Block[1], Amanda R. Moore[1], Akihide Yoshimi[1], Luisa Escobar-Hoyos[1], Hana Cho[1], Alex Penson[1], Sydney X. Lu[1], Justin Taylor[1], Yu Chen[1,9], Cigall Kadoch[6,7], Omar Abdel-Wahab[1,9]* & Robert K. Bradley[3,4,8]*

SF3B1 is the most commonly mutated RNA splicing factor in cancer[1–4], but the mechanisms by which *SF3B1* mutations promote malignancy are poorly understood. Here we integrated pan-cancer splicing analyses with a positive-enrichment CRISPR screen to prioritize splicing alterations that promote tumorigenesis. We report that diverse *SF3B1* mutations converge on repression of BRD9, which is a core component of the recently described non-canonical BAF chromatin-remodelling complex that also contains GLTSCR1 and GLTSCR1L[5–7]. Mutant SF3B1 recognizes an aberrant, deep intronic branchpoint within *BRD9* and thereby induces the inclusion of a poison exon that is derived from an endogenous retroviral element and subsequent degradation of *BRD9* mRNA. Depletion of BRD9 causes the loss of non-canonical BAF at CTCF-associated loci and promotes melanomagenesis. BRD9 is a potent tumour suppressor in uveal melanoma, such that correcting mis-splicing of *BRD9* in *SF3B1*-mutant cells using antisense oligonucleotides or CRISPR-directed mutagenesis suppresses tumour growth. Our results implicate the disruption of non-canonical BAF in the diverse cancer types that carry *SF3B1* mutations and suggest a mechanism-based therapeutic approach for treating these malignancies.

SF3B1 is subject to recurrent missense mutations at specific residues in myeloid leukaemia[1,2] and lymphoid leukaemia[3,8] as well as in solid tumours, at rates of up to 14–29% for uveal melanoma (UVM)[9–12] and 65–83% for myelodysplastic syndromes with ring sideroblasts[1,2]. Consistent with the critical role of SF3B1 in the recognition of 3′ splice sites[13], several previous studies have reported that *SF3B1* mutations induce widespread usage of abnormal 3′ splice sites[10,14,15]. Although many mis-spliced genes have been identified in *SF3B1*-mutant samples, few of these have been functionally implicated in driving disease.

We hypothesized that effectors of the pro-tumorigenic consequences of *SF3B1* mutations might appear as pan-cancer targets of mutant SF3B1. We accordingly identified mis-spliced events that were shared between erythroleukaemic (K562) and UVM (MEL270) cells that expressed wild-type *SF3B1* or the most-common *SF3B1* mutation, $SF3B1^{K700E}$. A compact set of 40 events exhibited concordant splicing changes, and was sufficient to infer *SF3B1* mutational status across 249 samples from patients with chronic lymphocytic leukaemia, myelodysplastic syndromes and UVM (Fig. 1a, Extended Data Fig. 1a, Supplementary Tables 1–3).

We designed a single-guide RNA (sgRNA) library that targeted both pan-cancer and cancer-type-specific targets of mutant SF3B1, focusing on genes for which *SF3B1* mutations are predicted to cause mis-splicing that triggers nonsense-mediated RNA decay (NMD) (Fig. 1b, Supplementary Table 4). We tested whether the knockout of any of these genes promoted the transformation of Ba/F3 cells (a mouse cell line with a wild-type spliceosome, with a requirement for IL-3 that can be overcome by oncogenic lesions) (Fig. 1c). In addition to the positive control *Pten*, our screen revealed that the loss of *Brd9* promoted the transformation of Ba/F3 cells (Fig. 1d, Extended Data Fig. 1b–d, Supplementary Tables 5, 6). *Brd9* was a notable hit because *BRD9* exhibited notable mis-splicing in all cohorts of patients with *SF3B1*-mutant cancer (Fig. 1e). *Brd9* knockout conferred cytokine independence to mouse 32Dcl3 cells, and growth advantage to human cancer cells with a wild-type spliceosome derived from UVM, cutaneous melanoma, and pancreatic cancer (Extended Data Fig. 1d–f). By contrast, acute myeloid leukaemia cells with rearranged *MLL* (also known as *KMT2A*) required BRD9 for growth (Extended Data Fig. 1g), as previously reported[16].

SF3B1 mutations cause the exonization of a *BRD9* intronic sequence, which results in the inclusion of a poison exon that interrupts the open reading frame of *BRD9*. This *BRD9* poison exon is derived from a primate-specific endogenous retroviral element, explaining its absence from mice (Extended Data Fig. 1h, i). We confirmed that the inclusion of the poison exon was induced by the expression of endogenous or ectopic mutant SF3B1 in K562 and NALM-6 cells, whereas *SF3B1* knockdown in *SF3B1* wild-type cells had no effect (Extended Data Fig. 1j–m). The poison exon was included in an *SF3B1*-mutation-dependent manner in diverse cell lines and in samples of chronic lymphocytic leukaemia, myelodysplastic syndromes and UVM that bear 19 different *SF3B1* mutations—but not in healthy tissues (Extended Data Fig. 1m–p, Supplementary Table 7).

The inclusion of the *BRD9* poison exon triggered NMD and reduced the half-life of *BRD9* mRNA and steady-state levels of full-length BRD9 protein (Extended Data Fig. 1q–w). Patients with *SF3B1* mutations exhibited reduced total levels of *BRD9* mRNA relative to patients with wild-type *SF3B1* (Extended Data Fig. 1x). We tested whether the inclusion of the poison exon could result in the production of C-terminally truncated BRD9 by knocking an N-terminal haemagglutinin tag into the *BRD9* locus in MEL270 and K562 cells that transgenically express wild-type or mutant SF3B1 (Extended Data Fig. 2a–c). Mutant SF3B1 suppressed levels of full-length BRD9 protein, without generating a truncated BRD9 protein (Fig. 1f).

SF3B1 mutations promote the use of cryptic 3′ splice sites[10,14,15], probably by altering the normal role of SF3B1 in branchpoint recognition[17]. We therefore mapped the *BRD9* branchpoints used in K562, MEL270 and T47D (breast cancer) cells that express mutant SF3B1 (Fig. 2a, Extended Data Fig. 2d–f). The inclusion of the poison exon was associated with an unusually close branchpoint (close branchpoints are rare and normally inefficiently recognized[18]). Mutating the aberrant branchpoint abolished poison exon recognition (Fig. 2b, Extended Data

[1]Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [2]Department of Hematology-Oncology, Institute of Biomedical Research and Innovation, Foundation for Biomedical Research and Innovation at Kobe, Kobe, Japan. [3]Computational Biology Program, Public Health Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [4]Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [5]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston, MA, USA. [6]Broad Institute of MIT and Harvard, Cambridge, MA, USA. [7]Biomedical and Biological Sciences Program, Harvard Medical School, Boston, MA, USA. [8]Department of Genome Sciences, University of Washington, Seattle, WA, USA. [9]Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA. [10]These authors contributed equally: Daichi Inoue, Guo-Liang Chew. *e-mail: abdelwao@mskcc.org; rbradley@fredhutch.org

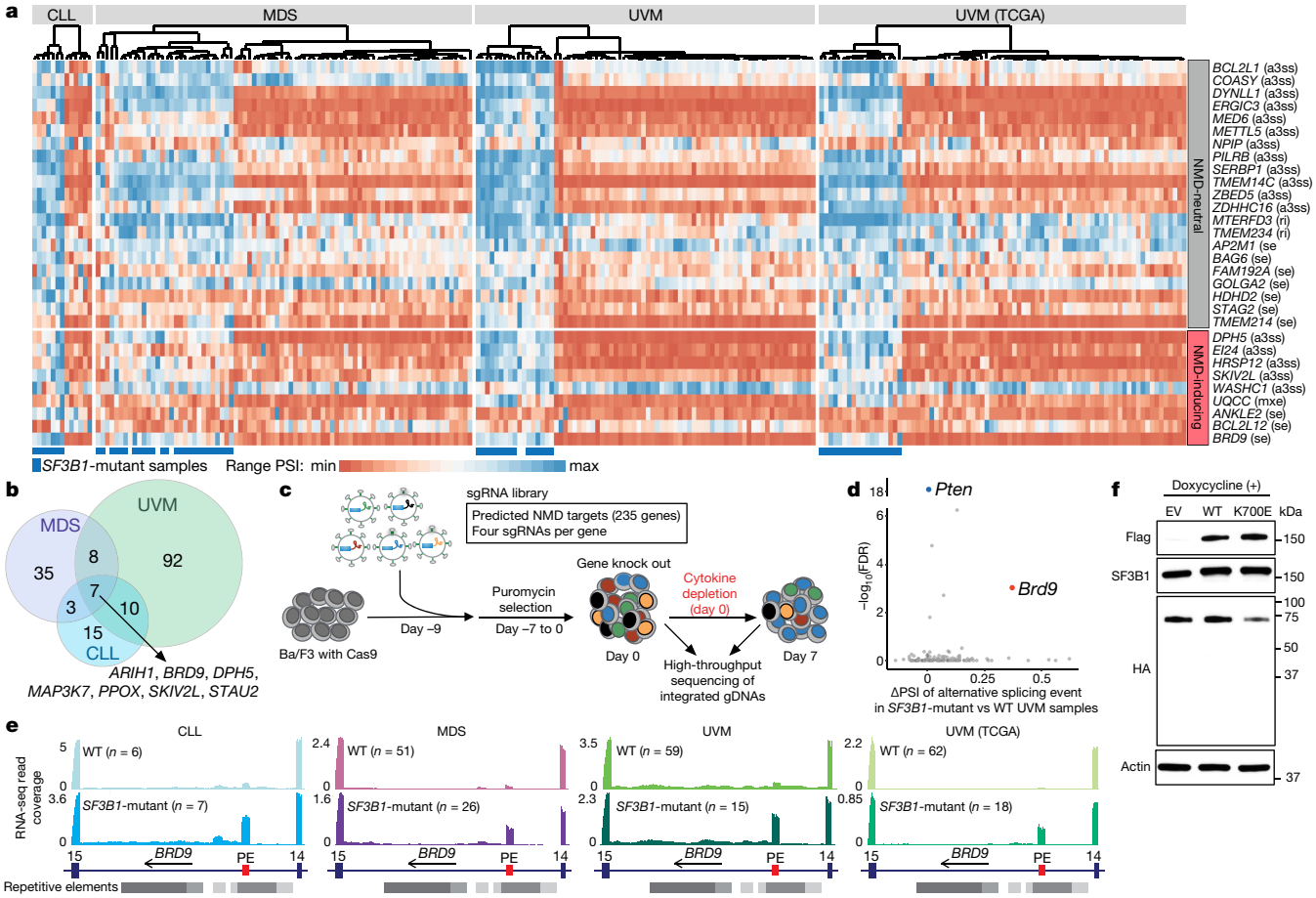


Fig. 1 | *BRD9* mis-splicing causes BRD9 loss and proliferative advantage in *SF3B1*-mutated cancers. a, Unsupervised clustering of patient samples on the basis of events that are differentially spliced in UVM (MEL270) and myeloid leukaemia (K562) cells that express $SF3B1^{K700E}$ versus wild-type (WT) *SF3B1*. a3ss, alternative 3′ splice site; CLL, chronic lymphocytic leukaemia (data are from ref. [15]); MDS, myelodysplastic syndromes (data are from ref. [27]); mxe, mutually exclusive exons; PSI, percentage spliced in (fraction of mRNA that corresponds to the mutant SF3B1-promoted isoform), with per-event and per-cohort range normalization; ri, retained intron; se, skipped exon; TCGA, The Cancer Genome Atlas. Data for UVM are from ref. [10] (middle) or the TCGA (right). *FAM192A*, *HRSP12*, *MTERFD3*, *NPIP* and *UQCC* are also known as *PSME3IP1*, *RIDA*, *MTERF2*, *NPIPA1* and *UQCC1*, respectively. **b**, Genes for which mutant SF3B1 promotes an isoform predicted to trigger NMD (alternative 3′ splice site and skipped exon events only) in one or more cohorts. **c**, CRISPR–Cas9-based positive-selection screen targeting genes for which mutant SF3B1 promotes an isoform predicted to trigger NMD. **d**, Per-gene scatter plot comparing CRISPR screen enrichment (*y* axis) to differential splicing in TCGA cohort of patients with UVM (*x* axis). *Pten* was used as a positive control. $n = 6$ biologically independent experiments. Per-gene significance computed with two-sided correlation-adjusted mean rank gene set (CAMERA) test. The false-discovery rate (FDR) was computed using the Benjamini–Hochberg method. **e**, *BRD9* RNA sequencing (RNA-seq) read coverage in patient samples. *n*, number of patients. PE, *BRD9* poison exon; 14 and 15, flanking constitutive exons. Repetitive elements from RepeatMasker[28]. **f**, Western blot for N-terminally haemagglutinin (HA)-tagged endogenous BRD9 in MEL270 cells transduced with empty vector (EV) or doxycycline-inducible Flag–SF3B1(WT) or Flag–SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments.

Fig. 2g). Consistent with the lack of an obvious polypyrimidine tract upstream of the poison exon, neither *U2AF1* nor *U2AF2* knockdown compromised poison exon recognition, whereas introducing a poly(Y) tract resulted in robust inclusion of the poison exon even in wild-type cells (Fig. 2b, Extended Data Fig. 2h–j). Finally, we identified a putative exonic splicing enhancer that was essential for inclusion of the poison exon (Fig. 2c, Extended Data Fig. 2k). We confirmed that the aberrant branchpoint, lack of a polypyrimidine tract, and exonic splicing enhancer were essential for poison exon recognition in the context of $SF3B1^{R625H}$, the most common mutation of *SF3B1* in UVM (Extended Data Fig. 2l–n). Disrupting the 3′ splice site and/or exonic splicing enhancer of the poison exon with CRISPR-directed mutagenesis markedly increased the levels of BRD9 protein in UVM cells with mutated *SF3B1* (Fig. 2d, Extended Data Fig. 2o, Supplementary Table 8), but had no effect on *BRD9* splicing or expression in cells with wild-type *SF3B1* (Extended Data Fig. 2p–r).

Several studies have recently described BRD9 as part of a non-canonical (nc) BAF complex, which is biochemically distinct from canonical BAF and polybromo-associated BAF[5–7] (Fig. 3a). Although ncBAF is not recurrently mutated in cancer—unlike canonical BAF and polybromo-associated BAF (Extended Data Fig. 3a)—our data suggested that ncBAF is nonetheless frequently disrupted via *SF3B1* mutations.

We investigated the consequences of BRD9 loss by *SF3B1* mutations for ncBAF function. Immunoprecipitation and mass spectrometry to identify the chromatin-associated interaction partners of BRD9 in K562 cells specifically recovered ncBAF components (Extended Data Fig. 3b, c, Supplementary Table 9). We confirmed these results by immunoblotting against shared and complex-specific components of canonical BAF, polybromo-associated BAF and ncBAF in K562 and UVM cells (Fig. 3b, Extended Data Fig. 3d). Expression of mutant, but not wild-type, SF3B1 reduced the levels of BRD9 protein and abolished interactions between BRG1 and GLTSCR1 while leaving interactions between BRG1 and BAF155 intact, which indicates that *SF3B1* mutations specifically perturb ncBAF rather than disrupting all BAF complexes (Fig. 3c, Extended Data Fig. 3e). Chemical degradation of BRD9[19] or *BRD9* knockout similarly reduced the BRG1–GLTSCR1 interaction (Fig. 3c, Extended Data Fig. 3f). We next identified the BRD9 domains that are necessary for ncBAF formation by generating 3×Flag–BRD9 deletion mutants

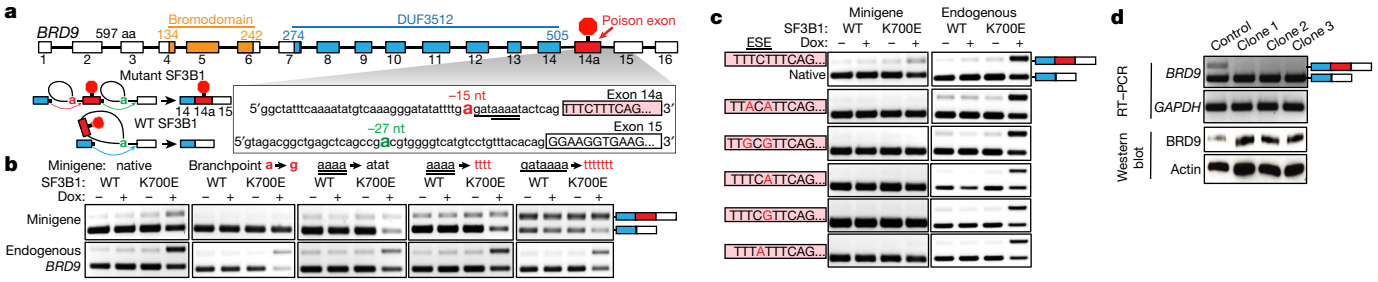


Fig. 2 | Mutant SF3B1 recognizes an aberrant, deep intronic branchpoint within *BRD9*. **a**, *BRD9* gene structure and protein domains. Inset illustrates the branchpoints used when the poison exon is included (top) or excluded (bottom). Single and double underlining indicates sequence motifs that were subsequently mutated. aa, amino acid. **b**, PCR with reverse transcription (RT–PCR) analysis of inclusion of the *BRD9* poison exon in a minigene (top) or endogenous (bottom) context, following transfection of minigenes with the illustrated mutations into MEL270 cells with doxycycline (dox)-inducible Flag–SF3B1(WT) or Flag–SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. Native, no mutations. **c**, As in **b**, but for minigene mutations (shown in red) at the 5′ end of the poison exon. ESE, exonic splicing enhancer. **d**, RT–PCR (top) illustrating the loss of inclusion of the *BRD9* poison exon, and corresponding western blot (bottom) in MEL202 ($SF3B1^{R625G}$) clones following CRISPR–Cas9 targeting of the poison exon. Indels are illustrated in Extended Data Fig. 2o. Control, unedited cells. Representative images from $n = 2$ (RT–PCR) and $n = 3$ (western blot) biologically independent experiments.

and testing for interactions with GLTSCR1 and GLTSCR1L. These experiments revealed that the DUF3512 domain of BRD9 mediates its interactions with GLTSCR1 and GLTSCR1L (Extended Data Fig. 3g–h).

We next determined how *SF3B1* mutations altered ncBAF localization to chromatin. We mapped the genome-wide binding of the pan-BAF component BRG1, and the ncBAF-specific components BRD9 and GLTSCR1, in MEL270 cells that express wild-type or mutant SF3B1. We additionally performed the same chromatin immunoprecipitation with sequencing (ChIP–seq) experiments after treatment with dimethylsulfoxide (DMSO) or a BRD9 degrader to identify BRD9-dependent effects. BRD9 and GLTSCR1 exhibited substantial co-localization, consistent with their mutual requirement for ncBAF formation, and were found at a subset of the loci bound by BRG1 (Fig. 3d). BRD9 and GLTSCR1 bound to promoters, gene bodies, and probable enhancers, with focal binding at promoters relative to BRG1 (Fig. 3e, Extended Data Fig. 4a). CTCF motifs exhibited notable co-localization with GLTSCR1, but only modest co-localization with BRG1 (Fig. 3f, Extended Data Fig. 4b).

We then tested how the depletion of BRD9, induced by $SF3B1^{K700E}$ or by chemical degradation of BRD9, altered ncBAF localization. We defined the genomic loci bound by GLTSCR1 in all samples as constitutive sites. Conversely, we defined genomic loci bound by GLTSCR1 in both control (wild-type SF3B1 or DMSO) but not BRD9-depleted

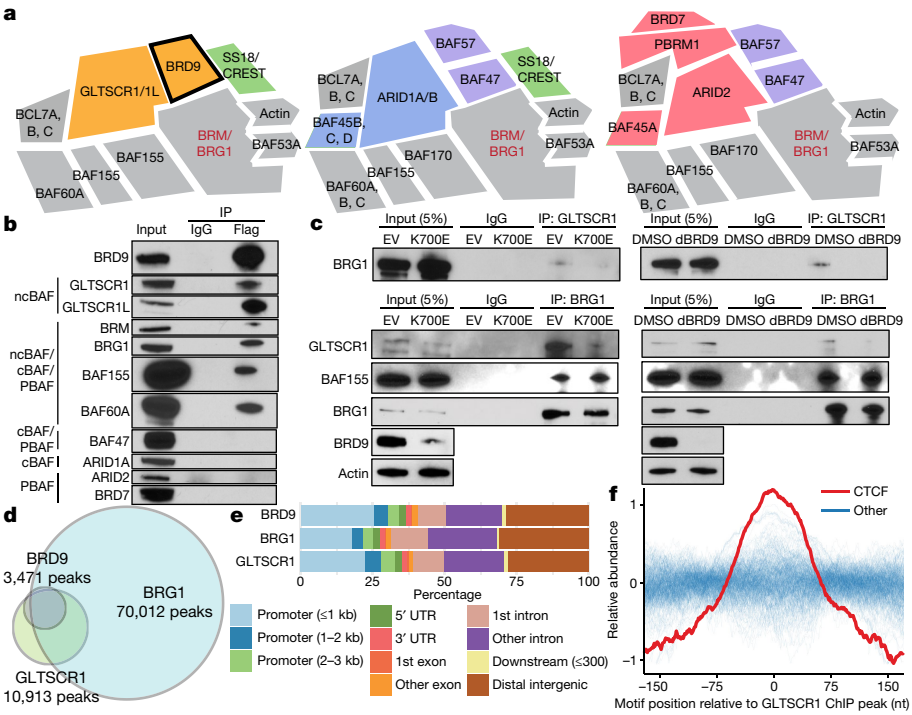


Fig. 3 | BRD9 loss perturbs the formation and localization of the ncBAF complex. **a**, Schematic of non-canonical BAF (ncBAF; left), canonical BAF (cBAF; middle) and polybromo-associated BAF (PBAF; right) complexes[5,6]. Solidus denotes one of the proteins is present; comma denotes one or more of the proteins are present or that mutually exclusive inclusion of proteins may occur. **b**, Cross-linking and immunoprecipitation (IP) with IgG or Flag followed by immunoblotting in K562 cells that express 3×Flag–BRD9. Representative images from $n = 3$ biologically independent experiments. **c**, Immunoprecipitation with GLTSCR1 or BRG1 antibody followed by immunoblotting in MEL270 cells that express exogenous SF3B1(K700E) (left) or were treated with BRD9 degrader[19] (dBRD9) (right). Representative images from $n = 3$ biologically independent experiments. **d**, Overlap of consensus BRD9, BRG1 and GLTSCR1 ChIP–seq peaks called in both MEL270 control samples (DMSO and ectopic expression of wild-type SF3B1). **e**, Genomic localization of BRD9-, BRG1-, and GLTSCR1-bound loci in **d**. UTR, untranslated region. **f**, Distributions of transcription factor-binding motifs at GLTSCR1-bound loci (20 nucleotide (nt) rolling mean). $n = 401$ transcription factors analysed.

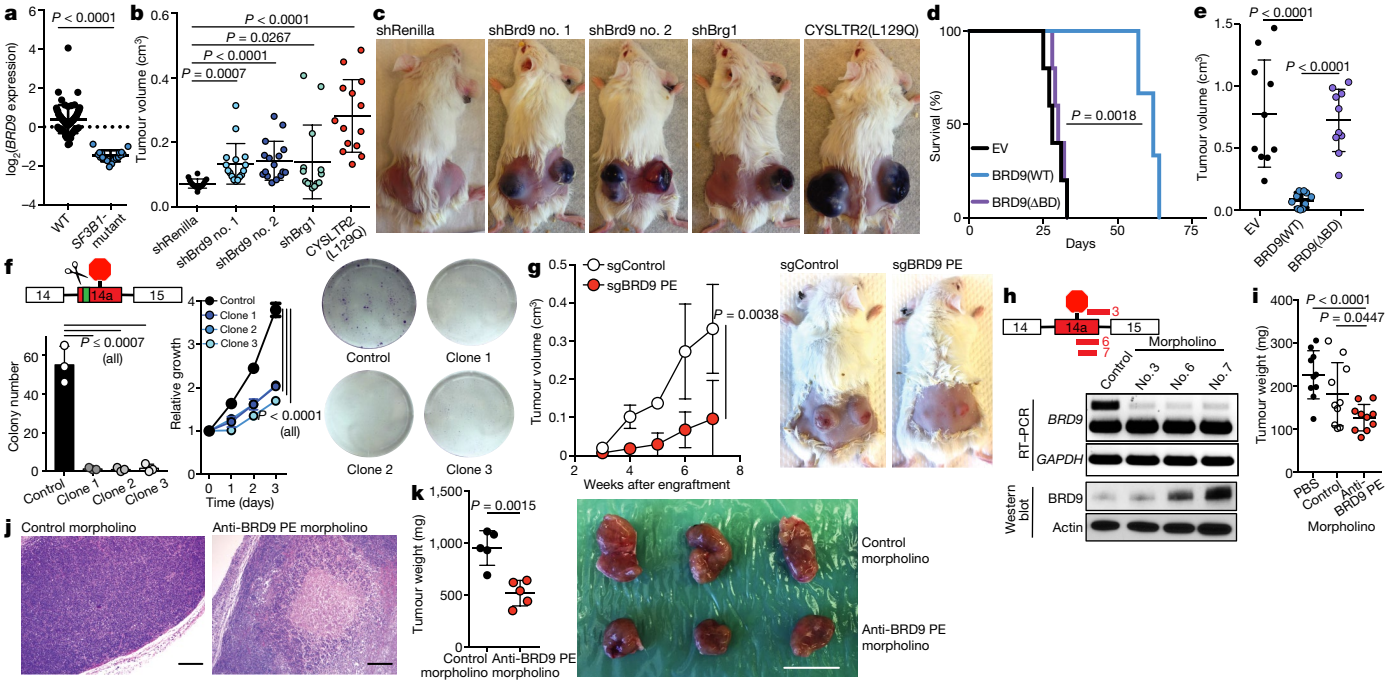


Fig. 4 | BRD9 is a therapeutically targetable tumour suppressor in melanoma. a, *BRD9* expression (*z*-score normalized) in TCGA UVM samples with ($n = 18$) or without ($n = 62$) *SF3B1* mutations. *P* value calculated by two-sided *t*-test. **b**, Tumour volume 49 days after subcutaneous engraftment of Melan-a cells transduced with the indicated shRNAs into SCID mice. $n = 16$, 16, 16, 14 and 14 tumours per group (left to right). Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test. **c**, Representative mice from **b** at day 63. **d**, Survival of SCID mice engrafted with MEL270 cells that express empty vector, full-length wild-type BRD9 or a BRD9 bromodomain-deletion mutant (ΔBD). $n = 5$ mice per group. *P* value calculated by log-rank test. **e**, Tumour volume from experiments shown in **d**, 21 days after engraftment. $n = 10$ tumours per group. Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test. **f**, Colony number (left) and representative images (right) of MEL202 cells (*SF3B1*^{R625G}) without (control) or with (clone 1, clone 2 and clone 3) CRISPR–Cas9-induced disruption of the *BRD9* poison exon. Indels are illustrated in Extended Data Fig. 2o. $n = 3$ biologically independent experiments. Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test at day 3 (middle). **g**, Tumour volume (left) and representative images (right) of mice engrafted with control or clone 1 cells from **f**. $n = 6$ tumours per group. Error bars, mean ± s.d. *P* value calculated by two-sided *t*-test at week 7. **h**, ASO design (top), RT–PCR (middle) and western blot (bottom) for BRD9. MEL202 cells (*SF3B1*^{R625G}) were treated with a non-targeting (control) or targeting morpholino at 10 μM for 24 h. Representative images from $n = 3$ biologically independent experiments. **i**, Tumour weight following 16 days of in vivo treatment of MEL202-derived xenografts (*SF3B1*^{R625G}) with PBS or a non-targeting (control) or poison-exon-targeting (no. 6) morpholino (12.5 mg kg^{-1}, every other day to a total of 8 intratumoral injections). $n = 10$ tumours per group. Error bars, mean ± s.d. *P* values calculated by two-sided *t*-test. **j**, Haematoxylin and eosin (H&E) images of tumours from **i**. Scale bars, 200 μm. Representative images from $n = 3$ biologically independent histological analyses. **k**, Tumour weight (left) and representative images (right) following in vivo morpholino treatment of a patient-derived rectal melanoma xenograft (*SF3B1*^{R625C}). Scale bar, 1 cm. $n = 5$ tumours per group. *P* value calculated by two-sided *t*-test. Error bars, mean ± s.d.

(mutant SF3B1 or BRD9 degradation) samples as BRD9-sensitive sites (Extended Data Fig. 4c). GLTSCR1 peaks were more sensitive to BRD9 loss than were BRG1 peaks, and CTCF motifs were uniquely enriched in BRD9-sensitive loci ($P < 1 \times 10^{-8}$) versus constitutive GLTSCR1-bound loci (Extended Data Fig. 4d). CTCF was similarly highly enriched at BRG1-bound loci that were BRD9-sensitive ($P < 1 \times 10^{-55}$) (Extended Data Fig. 4e, f). We conclude that the BRD9 loss induced by *SF3B1* mutations causes specific loss of ncBAF at CTCF-associated loci.

We identified genes with BRD9-sensitive ncBAF binding in their promoters or enhancers and found that BRD9 loss in UVM preferentially affects genes involved in apoptosis and cell growth, adhesion and migration (Extended Data Fig. 4g). To understand how BRD9 loss altered gene expression, we identified genes with promoters that exhibited BRD9-sensitive ncBAF binding and that were differentially expressed in patients with UVM with mutant versus wild-type *SF3B1*. Loss of ncBAF binding was associated with promotion as well as repression of gene expression, suggesting that ncBAF—similar to other SWI/SNF complexes—has both activating and repressive roles[20] (Extended Data Fig. 4h–j).

Several recent studies have reported that BRD9 is required for the survival of some cancer types, particularly cancers with mutations that affect polybromo-associated BAF and canonical BAF[6,16,21]. Because *BRD9* loss conferred a proliferative advantage to types of cancer with recurrent *SF3B1* mutations (Fig. 1d, Extended Data Fig. 1), we wondered whether normalizing the levels of BRD9 might suppress the growth of *SF3B1*-mutant cells.

As *SF3B1* is recurrently mutated in uveal (Fig. 4a), mucosal and cutaneous melanomas, we first tested whether *BRD9* loss induced melanomagenesis in vivo. We transduced non-tumorigenic mouse melanocytes (Melan-a cells), which require oncoprotein expression for sustained growth, with a non-targeting short hairpin RNA (shRNA) (Extended Data Fig. 5a), doxycycline-inducible shRNAs targeting *Brd9* or *Brg1* (also known as *Smarca4*) or a cDNA encoding the oncoprotein CYSLTR2(L129Q) (as a positive control[22]). Knockdown of either *Brd9* or *Brg1* resulted in potent tumour growth, augmented melanocyte pigmentation, and expression of melanocyte-lineage-specific genes in vivo (Fig. 4b, c, Extended Data Fig. 5b–g).

We next tested whether *Brd9* expression influences metastasis. *Brd9* knockdown significantly increased the number of pulmonary metastatic foci following intravenous injection of cells from a mouse model of melanoma (B16) or of human UVM (92.1) cells into mice (Extended Data Fig. 6a–f). By contrast, restoring *Brd9* expression in established tumours in vivo, by withdrawing doxycycline, suppressed tumour growth (Extended Data Fig. 6g, h). Similarly, ectopic expression of full-length BRD9, but not the bromodomain- or DUF3512-deletion mutants, suppressed the growth of UVM cell lines and xenografts (Fig. 4d, e, Extended Data Fig. 6i–k). These data demonstrate that loss of *Brd9* promotes cell transformation, tumour maintenance, and metastatic progression, and that the bromodomain and DUF3512 domain of BRD9 are essential for its anti-proliferative effects.

We sought to understand how BRD9 loss promotes melanoma tumorigenesis. We identified BRD9-bound genes that exhibited dysregulated expression in samples from patients with UVM with mutated versus wild-type *SF3B1*, and in isogenic UVM cells with or without mutant SF3B1 and with or without forced loss of BRD9. *HTRA1*, a known tumour suppressor in melanoma[23,24], was the most downregulated gene in UVM (Extended Data Fig. 7a–c). HTRA1 was suppressed by mutant SF3B1 expression and BRD9-degradation treatment of UVM cells with wild-type *SF3B1*, and mutagenesis of the *BRD9* poison exon increased levels of HTRA1 in UVM cells with mutated *SF3B1* (Extended Data Fig. 7d, e). *HTRA1* is bound by ncBAF in UVM, and this binding is reduced by mutant SF3B1 (Extended Data Fig. 7f). *HTRA1* knockdown promoted the growth of UVM cells with wild-type *SF3B1*, and ectopic expression of *HTRA1* suppressed the growth of UVM cells with mutated *SF3B1* (Extended Data Fig. 7g–k). These data suggest that perturbation of ncBAF-dependent regulation of *HTRA1* contributes to the pro-tumorigenic effects of BRD9 loss.

We next tested whether correcting *BRD9* mis-splicing suppressed tumorigenesis. CRISPR-based mutagenesis of the poison exon markedly slowed the growth of cells with mutated *SF3B1*, but not of wild-type cells, both in vitro and in vivo (Figs. 2d, 4f, g, Extended Data Figs. 2o–r, 8). We then designed antisense oligonucleotides (ASOs) to block the inclusion of the *BRD9* poison exon (Fig. 4h, Extended Data Fig. 9a). We treated *SF3B1*-mutated cells with a non-targeting (control) or poison-exon-targeting ASO, and measured *BRD9* splicing, BRD9 protein levels, and cell growth. Each targeting ASO prevented the inclusion of the poison exon, increased the level of BRD9 protein, and suppressed cell growth relative to the control ASO (Fig. 4h, Extended Data Fig. 9b). The relative abilities of each ASO to restore BRD9 protein levels and suppress cell growth were strongly correlated, consistent with on-target effects.

We therefore tested whether ASO treatment slowed tumour growth in vivo. We treated *SF3B1*-mutated xenografts (derived from MEL202 cells) with each ASO via intratumoral injection for 16 days. Treatment with the poison-exon-targeting ASO—but not with the non-targeting ASO—corrected *BRD9* mis-splicing, significantly reduced tumour growth, and induced tumour necrosis (Fig. 4i, j, Extended Data Fig. 9c–f). We observed a similar ASO efficacy in a patient-derived xenograft model of rectal melanoma with the $SF3B1^{R625C}$ mutation (Fig. 4k, Extended Data Fig. 9g–i, Supplementary Table 7). By contrast, when we performed an identical experiment with a patient-derived xenograft model of UVM that lacked an *SF3B1* mutation, treatment with the poison-exon-targeting ASO had no effect (Extended Data Fig. 9j–l, Supplementary Table 7). We conclude that correcting *BRD9* mis-splicing restores the tumour suppressor activity of BRD9 in cancers with *SF3B1* mutations.

Although recognition of the *BRD9* poison exon requires mutant SF3B1, *BRD9* mis-splicing and ncBAF disruption may also have roles in cancers with wild-type *SF3B1*. We identified significant pan-cancer expression correlations between *BRD9* and many genes that encode RNA-binding proteins, as well as six additional *BRD9* isoforms that are predicted to trigger NMD that are expressed in cancers with wild-type *SF3B1* and are predictive of *BRD9* expression (Extended Data Fig. 10, Supplementary Table 10). A recent study has also identified a promoter polymorphism associated with decreased *GLTSCR1* (also known as *BICRA*) expression as a common risk allele for acute myeloid leukaemia[25].

As we observed *BRD9* mis-splicing in a range of cancer types that carry distinct *SF3B1* mutations, targeting *BRD9* mis-splicing could be a productive pan-cancer therapy. Although ASO treatment merely restored *BRD9* mRNA and BRD9 protein to normal levels, we nonetheless observed a strong suppression of tumour growth. The functional effect of correcting *BRD9* mis-splicing was particularly notable given that the UVM models used here contain hundreds of other mis-splicing events and multiple pro-tumorigenic mutations. Given recent clinical successes with treating spinal muscular atrophy and other diseases with ASOs[26], the tumour-suppressive effects of correcting *BRD9* mis-splicing suggest that oligonucleotide-based therapy may prove similarly promising for treating cancers with spliceosomal mutations.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-019-1646-9.

Received: 12 January 2019; Accepted: 30 August 2019;
Published online 9 October 2019.

1. Yoshida, K. et al. Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* **478**, 64–69 (2011).
2. Papaemmanuil, E. et al. Somatic *SF3B1* mutation in myelodysplasia with ring sideroblasts. *N. Engl. J. Med.* **365**, 1384–1395 (2011).
3. Wang, L. et al. *SF3B1* and other novel cancer genes in chronic lymphocytic leukemia. *N. Engl. J. Med.* **365**, 2497–2506 (2011).
4. Seiler, M., et al. Somatic mutational landscape of splicing factor genes and their functional consequences across 33 cancer types. *Cell Rep.* **23**, 282–296.e284 (2018).
5. Alpsoy, A. & Dykhuizen, E. C. Glioma tumor suppressor candidate region gene 1 (*GLTSCR1*) and its paralog GLTSCR1-like form SWI/SNF chromatin remodeling subcomplexes. *J. Biol. Chem.* **293**, 3892–3903 (2018).
6. Michel, B. C. et al. A non-canonical SWI/SNF complex is a synthetic lethal target in cancers driven by BAF complex perturbation. *Nat. Cell Biol.* **20**, 1410–1420 (2018).
7. Gatchalian, J. et al. A non-canonical BRD9-containing BAF chromatin remodeling complex regulates naive pluripotency in mouse embryonic stem cells. *Nat. Commun.* **9**, 5139 (2018).
8. Quesada, V. et al. Exome sequencing identifies recurrent mutations of the splicing factor *SF3B1* gene in chronic lymphocytic leukemia. *Nat. Genet.* **44**, 47–52 (2011).
9. Furney, S. J. et al. *SF3B1* mutations are associated with alternative splicing in uveal melanoma. *Cancer Discov.* **3**, 1122–1129 (2013).
10. Alsafadi, S. et al. Cancer-associated *SF3B1* mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat. Commun.* **7**, 10615 (2016).
11. Harbour, J. W. et al. Recurrent mutations at codon 625 of the splicing factor *SF3B1* in uveal melanoma. *Nat. Genet.* **45**, 133–135 (2013).
12. Martin, M. et al. Exome sequencing identifies recurrent somatic mutations in *EIF1AX* and *SF3B1* in uveal melanoma with disomy 3. *Nat. Genet.* **45**, 933–936 (2013).
13. Cretu, C. et al. Molecular architecture of SF3b and structural consequences of its cancer-related mutations. *Mol. Cell* **64**, 307–319 (2016).
14. 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).
15. 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).
16. Hohmann, A. F. et al. Sensitivity and engineered resistance of myeloid leukemia cells to BRD9 inhibition. *Nat. Chem. Biol.* **12**, 672–679 (2016).
17. Gozani, O., Potashkin, J. & Reed, R. A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell. Biol.* **18**, 4752–4760 (1998).
18. Pineda, J. M. B. & Bradley, R. K. Most human introns are recognized via multiple and tissue-specific branchpoints. *Genes Dev.* **32**, 577–591 (2018).
19. Remillard, D. et al. Degradation of the BAF complex factor BRD9 by heterobifunctional ligands. *Angew. Chem. Int. Edn Engl.* **56**, 5738–5743 (2017).
20. Martens, J. A. & Winston, F. Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr. Opin. Genet. Dev.* **13**, 136–142 (2003).
21. Brien, G. L. et al. Targeted degradation of BRD9 reverses oncogenic gene expression in synovial sarcoma. *eLife* **7**, e41305 (2018).
22. Moore, A. R. et al. Recurrent activating mutations of G-protein-coupled receptor CYSLTR2 in uveal melanoma. *Nat. Genet.* **48**, 675–680 (2016).
23. Baldi, A. et al. The HtrA1 serine protease is down-regulated during human melanoma progression and represses growth of metastatic melanoma cells. *Oncogene* **21**, 6684–6688 (2002).
24. Chien, J. et al. Serine protease HtrA1 associates with microtubules and inhibits cell migration. *Mol. Cell. Biol.* **29**, 4177–4187 (2009).
25. Walker, C. J. et al. Genome-wide association study identifies an acute myeloid leukemia susceptibility locus near *BICRA*. *Leukemia* **33**, 771–775 (2019).
26. Stein, C. A. & Castanotto, D. FDA-approved oligonucleotide therapies in 2017. *Mol. Ther.* **25**, 1069–1075 (2017).
27. Pellagatti, A. et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood* **132**, 1225–1240 (2018).
28. Smit, A., Hubley, R. & Green, P. RepeatMasker Open-4.0, http://www.repeatmasker.org (2013–2015).

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

© The Author(s), under exclusive licence to Springer Nature Limited 2019

METHODS

Sample sizes for xenograft experiments were chosen on the basis of published studies of known oncogenic drivers of relevant models (for example, expression of the oncoprotein CYSLTR2(L129Q) in Melan-a cells). Mice were randomly assigned to experimental groups. The data presented did not require the use of blinding.

Cell lines and tissue culture. All cell lines underwent short-tandem repeat testing (ATCC) and Memorial Sloan Kettering integrated mutation profiling of actionable cancer targets (MSK IMPACT) genetic analysis[29] to evaluate for spliceosome-gene mutational status and status of recurrently mutated genes in cancer. HEK293T cells were grown in DMEM with 10% FCS. Ba/F3 cells and Melan-a cells (provided by D. Bennett) were grown in RPMI with 10% FCS with 1 ng/ml IL-3 (PeproTech, 213-13) and 200 nM TPA (Sigma-Aldrich), respectively, unless noted otherwise. The K562 and NALM-6 isogenic cell lines (engineered to express *SF3B1*^{K700E}, *SF3B1*^{K666N} or *SF3B1*^{K700K} (wild-type control for genome engineering) from the endogenous *SF3B1* locus) were cultured in RPMI with 10% FCS and their generation has previously been described[30]. MEL270, MEL285 and RN2 cell lines were cultured in RPMI with 10% FCS. MEL202, 92-1 and SK-MEL30 cells were grown in RPMI with 10% FCS and 1% GlutaMAX (Gibco). UPMD1 and UPMD2 cells were grown in Ham F-12 with 10% FCS. CFPAC1 cells were cultured in IMDM with 10% FCS. KPC, Miapaca2 and B16 cells were cultured in DMEM with 10% FCS. Panc 05.04 cells were grown in RPMI with 20% FCS and 20 units per millilitre human recombinant insulin. T47D cells were cultured in RPMI1640 supplemented with 10% fetal bovine serum (Corning), 100 μg/ml penicillin, 100 mg/ml streptomycin (Corning), and 4 mM glutamine. All of the cell culture media included penicillin (100 U/ml) and streptomycin (100 μg/ml).

Primary human samples and human patient-derived xenograft models. Studies were approved by the Institutional Review Boards (IRBs) of Memorial Sloan Kettering Cancer Center (MSK), informed consent was obtained from all subjects (under MSK IRB protocol 06-107) and studies were conducted in accordance to the Declaration of Helsinki protocol. Patients provided samples after their informed consent, and samples of primary human de-identified chronic lymphocytic leukaemia derived from whole peripheral blood or bone marrow mononuclear cells were used. Patient-derived xenograft models were performed using tumour biopsies from de-identified patients under MSK IRB protocol 14-191. Genomic alterations in melanoma tumour biopsies and chronic lymphocytic leukaemia cells were analysed using the MSK IMPACT[29] assay or FoundationOne Heme[31] assay, both as previously described. Patient samples were anonymized by the Hematologic Oncology Tissue Bank of MSK (for chronic lymphocytic leukaemia samples) and the MSK Antitumour Assessment Core Facility (for patient-derived xenograft samples).

Mice. All mice were housed at Memorial Sloan Kettering Cancer Center (MSKCC). All mouse procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC. All mouse experiments were performed in accordance with a protocol approved by the MSKCC Institutional Animal Care and Use Committee (11-12-029). SCID mice (Jackson Laboratories stock no. 001303) were used for all human cell line xenografts and NSG mice (Jackson Laboratories stock no. 005557) were used for patient-derived xenografts. For all mouse experiments, the mice were monitored closely for signs of disease or morbidity daily and were killed when they showed a volume of the visible tumour formation above 1 cm^3, failed to thrive, experienced weight loss >10% total body weight or showed open skin lesions, bleeding or any signs of infection. These limits were not exceeded in any experiments.

HA-tag knock-in into endogenous *BRD9*. The following guide RNA sequence targeting the *BRD9* transcriptional start site was selected using the optimized CRISPR design tool (http://crispr.mit.edu): CGAGTGGCGCTCGTCCTACG. DNA oligonucleotides were purchased from IDT and cloned into the px458-GFP vector. For homologous recombination, we purchased a custom IDT Ultramer 197-bp repair template (single-stranded donor DNA) with the HA sequence (TACCCATACGATGTTCCAGATTACGCT) directly following the *BRD9* start codon. This 197-bp fragment contained two silent mutations, one to remove the PAM site (AGG > AAG) and another to introduce an XhoI restriction enzyme (CTCGGG > CTCGAG) site upstream of the HA tag. The 197-bp fragment also contained 83 bp of homology to the *BRD9* 5′ UTR upstream of the HA tag and 87 bp of homology to the *BRD9* exon 1 downstream of the start codon. Five micrograms of the targeting construct and 500 nM of the repair template were nucleofected into K562 *SF3B1*^{K700E} cells and MEL270 cells using the Lonza Nucleofector V kit and Program T-003 on the nucleofector device. Nucleofected cells were single-cell-sorted on the basis of GFP positivity 48 h after nucleofection. Clones were screened for the presence of successful HA insertion by *BRD9* exon 1 PCR and subsequent restriction enzyme digestion with XhoI and direct Sanger sequencing. A single positive clone containing the HA coding sequencing was selected to carry out further studies. The sgRNAs were CACCCGAGTGGCGCTCGTCCTACG (top) and AAACCGTAGGACGAGCGCCACTCG (bottom); the single-strand donor DNA was CCAGGGGGCGGTGGCGGCCAAGGTCCGACCGGGTGC CAGCTGTTCCCAGCCCCCGCCTCGAGCCCGCCGCCGGCGCCGCCATGT ACCCATACGATGTTCCAGATTACGCTGGCAAGAAGCACAAGAAGCACA AGGCCGAGTGGCGCTCGTCCTACGAAGGTGAGGCGGCGGCGCTTTGTG ACGCGCGGCGGGCGGGG; the PCR primers were forward (fwd) AGCG AGCTCGGCAACCTCG and reverse (rev) CTTCAGGACTAGCTTTAGAGGC; the Sanger sequence primer was rev TGCAGCCTCGAACCCAGAAC

Overexpression of *SF3B1* cDNA in K562 cells. Two micrograms of PiggyBac Transposase construct (CMV-PB-Transposase-IRES-TK-HSV) and 6 μg of wild-type *SF3B1* (ITR-CAG-Flag-SF3B1^{WT}-IRES-Puro-ITR) or *SF3B1*^{K700E} (ITR-CAG-Flag-SF3B1^{K700E}-IRES-Puro-ITR) cDNA constructs were electroporated into 2 × 10^6 cells (in 200 μl volume) using the Amaxa Nucleofector Protocol (Program T-003) according to manufacturer instructions (Lonza). Puromycin selection (1 μg/ml) was initiated 4 days after electroporation to select for cells that successfully incorporated the constructs. Sanger sequencing was performed to confirm successful integration of the cDNA plasmid using the following primers: fwd, TCCAATCAAAGATCTTCTTCCAA and rev, GAGCAGGTTTCTGCAACGAT.

RT–PCR and quantitative RT–PCR. Total RNA was isolated using RNeasy Mini or Micro kit (Qiagen). For cDNA synthesis, total RNA was reverse-transcribed to cDNA with SuperScript VILO cDNA synthesis kit (Life Technologies). The resulting cDNA was diluted 10–20 fold before use. Quantitative RT–PCR (qRT–PCR) was performed in 10-μl reactions with SYBR Green PCR Master Mix. All qRT–PCR analysis was performed on an Applied Biosystems QuantStudio 6 Flex Cycler (ThermoFisher Scientific). Relative gene expression levels were calculated using the comparative C_T method. Primers used in RT–PCR reactions were: *BRD9* (human) fwd, GCAATGACATACAATAGGCCAGA and rev, GAGCTGCCTGTTTGCTCATCA; *Brd9* (mouse) fwd, TTGGAG ATGGAAGTCTGCTCT and rev, GCAACTTGCTAGACAGTGAACT; *BRD9* poison exon (human) fwd, AGCTCTGTTCTGGAGTTCATG and rev, CTGAAG AAACTCATAGGGGTCGTG; *Brd9* poison exon (mouse) fwd, GGCCCT GTTCTGGACTTCATG and rev, CTGAAGGAATTCATAAGGGTCGTG; *BRD9* poison exon inclusion for small interfering RNA (siRNA) experiment (human) fwd, CAGCAGCTCTGTTCTGGAGT and rev: CCTGAAAGAAACCAGAGAGCTG; *BRD9* poison exon exclusion for siRNA experiment (human) fwd, CAGCAGCTCTGTTCTGGAGT and rev, TCACCTTCCCCAGAGAGCTG; *EPB49* (also known as *DMTN*) cassette exon inclusion (human) fwd, GCCTGCAGAACGGAGAGG and rev: ACCACTAGCATTTCATAGGGATAGATCT; *EPB49* cassette exon exclusion (human) fwd, GCCTGCAGATCTATCCCTATGAAAT and rev, CTCAAGCCGCATCCGATCC; *BRD9* poison exon for mRNA half-life experiment (human) fwd, GTTGGGGACACCCTAGGAGA, rev (exclusion-specific), CTTCACCTTCCCCAGAGAGC and rev (inclusion-specific), CCCTGAAAGAAACCAGAGAGC; 18S rRNA (human) fwd CTAC CACATCCAAGGAAGCA and rev, TTTTTCGTCACTACCTCCCCG; *Mitf* (mouse) fwd, CCAACAGCCCTATGGCTATGC and rev: CTGGGCA CTCACTCTCTGC; *Dct* (mouse) fwd, GTCCTCCACTCTTTTACAGACG and rev, ATTCGGTTGTGACCAATGGGT; *Pmel* (mouse) fwd, GAGCTTCCTTCCCGTGCTT and rev, TGCCTGTTCCAGGTTTTAGTTAC; *Tyrp1* (mouse) fwd, CCCCTAGCCTATATCTCCCTTTT and rev, TACCATCGT GGGGATAATGGC; *GAPDH* (human) fwd, GGAGCGAGATCCCTCCAAAAT and rev, GGCTGTTGTCATACTTCTCATGG; and *Gapdh* (mouse) fwd, AGGTCGGTGTGAACGGATTTG and rev, TGTAGACCATGTAGTTGAGGTCA.

mRNA stability assay. For mRNA half-life measurement using qRT–PCR, K562 and NALM-6 cells with isogenic *SF3B1*^{K700E} mutations were infected with anti-*UPF1* shRNAs or control shRNA, and treated with 2.5 μg/ml actinomycin D (Life Technologies) and collected at 0, 2, 4, 6 and 8 h (using methods as previously described[32]). *BRD9* poison exon inclusion or exclusion and 18S rRNA mRNA levels were measured by qRT–PCR.

Western blotting. For western blotting, the following antibodies to the following proteins were used: BRD9 (Bethyl Laboratories A303-781A and Active Motif 61538), SF3B1/Sap-155 (MBL D221-3), Flag-M2 (Sigma-Aldrich F-1084), β-actin (Sigma-Aldrich A-5441), GLTSCR1 (Santa Cruz Biotechnology sc-515086), GLTSCR1L (Thermo Fisher Scientific PA5-56126), BRM (Bethyl Laboratories A303-015A), BRG1 (Santa Cruz Biotechnology sc-17796), BAF155 (Santa Cruz Biotechnology sc-48350), BAF60A (Santa Cruz Biotechnology sc-135843), BAF47 (Santa Cruz Biotechnology sc-166165), ARID1A (Santa Cruz Biotechnology sc-373784), ARID2 (Santa Cruz Biotechnology sc-166117), BRD7 (Thermo Fisher Scientific PA5-49379), U2AF2 (Bethyl Laboratories A303-665A), U2AF1 (Bethyl Laboratories A302-080A), histone H3 (Abcam ab1791), HTRA1 (R&D Systems MAB2916-SP). All primary antibodies for western blotting were diluted to a final concentration of 1:500 to 1,000, in either 5% BSA (Sigma-Aldrich) in 0.05% TBS–Tween 20 (TBS-T) or 5% skim milk in 0.05% TBS-T. Nuclear extracts were quantified using BCA and 1 mg protein (1 mg ml^{-1} in immunoprecipitation buffer supplemented with protease inhibitors) was used per immunoprecipitation.

Proteins were incubated for 3 h with 2–5 μg of antibody or with protein A/G PLUS-Agarose (Santa Cruz Biotechnology sc-2003) with rotation at 4 °C. After washing three times with Pierce IP Lysis Buffer (Thermo Fisher Scientific 87787), immunoprecipitated proteins were eluted with Pierce Lane Marker Reducing Sample Buffer (Thermo Fisher Scientific 39000) and loaded onto 4–12% Bis-Tris NuPAGE Gels (Life Technologies).

ChIP. For ChIP–seq studies in MEL270 cells, antibodies to endogenous BRG1 (Abcam EPNCIR111A, lot no. GR3208604-8), GLTSCR1 (Santa Cruz SC-240516, lot no. A2313) and BRD9 (Abcam ab137245) were used, and ChIP was performed as previously described in detail[6]. MEL270 cells transduced with empty vector, doxycycline-inducible wild-type *SF3B1* cDNA, or doxycycline-inducible *SF3B1*^{K700E} cDNA in the backbone of pInducer20, were treated with doxycycline (1 μg/ml) plus BRD9 degrader[19] (250 nM) or DMSO for 72 h before crosslinking.

Mass spectrometry. For anti-Flag–BRD9 ChIP followed by mass spectrometry, K562 cells transduced with empty vector or 3×Flag-tagged BRD9 were grown in RPMI with 10% FCS. Ten million cells were crosslinked according to the manufacturer's instruction (Active Motif) and as previously described[33]. Cells were fixed with 1% methanol-free formaldehyde (Sigma, F-8775) for 8 min and quenched with 0.125 M glycine (Sigma, G-7403). Chromatin was isolated by the addition of lysis buffer, followed by disruption with a Dounce homogenizer. Lysates were sonicated and the DNA sheared to an average length of 300–500 bp. Genomic DNA (input) was prepared by treating aliquots of chromatin with RNase, proteinase K and heat for reverse-crosslinking, followed by ethanol precipitation. Pellets were resuspended and the resulting DNA was quantified on a NanoDrop spectrophotometer. Extrapolation to the original chromatin volume allowed quantification of the total chromatin yield. An aliquot of chromatin (150 μg) was precleared with protein G agarose beads (Invitrogen). Proteins of interest were immunoprecipitated using 15 μg of antibody against Flag and protein G magnetic beads. Protein complexes were washed and trypsin was used to remove the immunoprecipitate from beads and digest the protein sample. Protein digests were separated from the beads and purified using a C18 spin column (Harvard Apparatus). The peptides were vacuum-dried using a SpeedVac. Digested peptides were analysed by liquid chromatography and tandem mass spectrometry on a Thermo Scientific Q Exactive Orbitrap mass spectrometer in conjunction with a Proxeon Easy-nLC II HPLC (Thermo Scientific) and Proxeon nanospray source.

Protein identifications were accepted if they contained at least one identified peptide. Proteins that contained similar peptides and could not be differentiated on the basis of tandem mass spectrometry analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Protein peptide evidence is specified in Supplementary Table 9. The final list was generated by taking all proteins with a spectral count of five and above from each replicate reaction and comparing them in a Venn diagram against IgG control replicates. Proteins unique to both experimental replicates were then applied to the PANTHER database for protein ontology results.

shRNA experiments. Cells were transduced with a doxycycline-inducible LT3GEPIR lentiviral vector, T3G-GFP-mirE-PGK-Puro-IRES-rtTA3[34], expressing shRNAs against *BRD9* or a non-targeting shRNA against *Renilla*. shRNAs were induced with the addition of doxycycline (2.0 μg/ml, Sigma Aldrich). All shRNAs were designed using the SplashRNA algorithm[35]. The shRNA sequences are: *BRD9* shRNA no. 1 (human, shBRD9_352): TTTATTATCATTGAATATCCAG; *BRD9* shRNA no. 2 (human, shBRD9_353): TTTTATTATCATTGAATATCCA; *Brd9* shRNA no. 1 (mouse, shBrd9_511): TTTATTATCATTGAATACCCAG; *Brd9* shRNA no. 2 (mouse, shBrd9_512): TTTTATTATCATTGAATACCCA; *HTRA1* shRNA no. 1 (human, shHTRA1_1192): TTTTTAATCTTATCAGATGGGA; *HTRA1* shRNA no. 2 (human, shHTRA1_1669): TGAACAAACAAAATGGCAGTCA; and *HTRA1* shRNA no. 3 (human, shHTRA1_1898): TTCTATCTACGCATTGTATCGA.

siRNA transfections. K562 cells were transfected with a non-targeting control siRNA (Dharmacon, D-001810-01, target sequence: UGGUUUACAUGUCGACUAA), an siRNA pool against *U2AF1* (Dharmacon ON-TARGETplus SMARTpool, L-012325-01) or an siRNA pool against *U2AF2* (Dharmacon ON-TARGETplus SMARTpool, L-012380-02) using the Nucleofector II device from Lonza with the Cell Line Nucleofector Kit V (program T16). RNA and protein were extracted 48 h after transfection. cDNA was produced using 1 μg of RNA and the Superscript III first strand synthesis system (Thermo Fisher, 18080051).

In vitro competition assay. For the competition assay to evaluate the cellular effect of sgRNA against *BRD9* or *Brd9*, cell lines were transduced with LentiCas9-Blast (Addgene no. 52962) and then single-cell-sorted into 96-well plates. Among these clones, we used single clones with strong Cas9 expression, which was confirmed by western blotting. Cas9-expressing cells were lentivirally transduced with iLenti-guide-GFP vectors subcloned with the gRNA sequences, in which sgRNA expression was linked to GFP expression. The percentage of GFP-expressing cells was then measured over time after infection using BD LSRFortessa. The GFP-positive rates in living cells at each point compared to that of day 2 were calculated. Similarly, to evaluate the cellular effect of *BRD9* fragment cDNA, *HTRA1* cDNA or shRNAs against *BRD9* or *HTRA1*, GFP-positive rates were measured after transducing pMIGII-backbone plasmids (cDNA) and LT3GEPIR plasmids (shRNA), respectively, into melanoma cell lines. The sgRNAs were as follows: *BRD9* sgRNA no. 1 (human): ACTCCAGTTACTATGATGAC, *BRD9* sgRNA no. 2 (human): AGAGAGGGAGCACTGTGACA, *BRD9* sgRNA no. 3 (human): AGATACCGTGTACTACAAGT, *Brd9* sgRNA no. 1 (mouse): ATTAACCGGTTTCTCCCGGG, *Brd9* sgRNA no. 2 (mouse): GGAACACT GCGACTCAGAGG, *Brd9* sgRNA no. 3 (mouse): ACTTGCTAGACAGTGAACTC, and control (scrambled sgRNA): ACGGAGGCTAAGCGTCGCAA.

CRISPR enrichment screening for NMD targets. First, Ba/F3 cells were transduced with LentiCas9-Blast (Addgene no. 52962) and single-cell-sorted into 96-well plates. Among these clones, we used a single clone with strong Cas9 expression. The sgRNA library of NMD targets in SF3B1-mutant cells were amplified and packaged as lentivirus. The library includes 4 sgRNAs against each target gene (a total of 274 genes), 100 control sgRNAs and positive-control sgRNAs against *Pten*. Ba/F3 cells were transduced with the lentivirus-carrying sgRNA library produced by 293FT cells, and puromycin selection (2 μg/ml) was performed in IL-3-containing medium for 7 days. Then, we washed out IL-3 (on day 0) and the surviving cells were collected 7 days after IL-3 depletion (day 7). Cell pellets were lysed and genomic DNA was extracted (Qiagen), and quantified by Qubit (Thermo Scientific). A quantity of gDNA covering a 1,000× representation of gRNA was PCR-amplified using Q5 high-fidelity polymerase (NEB cat. no. M0491) to add Illumina adapters and multiplexing barcodes. Amplicons were quantified by Qubit and Bioanalyzer (Agilent) and sequenced on Illumina HiSeq 2500. Sequencing reads were aligned to the screened library; counts were computed for each gRNA and counts for each sgRNA were compared between days 0 and 7 after cytokine depletion. For the probe-level analysis, we fitted a negative binomial generalized log-linear model and performed a likelihood ratio test with glmFit and glmLRT in the Bioconductor edgeR package. For the gene-level analysis, we used the CAMERA test as implemented in edgeR[36–38]. FDR values were computed using the Benjamini–Hochberg method.

CRISPR-directed mutations. Cas9-expressing MEL202 and MEL270 cells were transduced with iLenti-guide-Puro vector targeting the 5′ end of the *BRD9* poison exon. The sgRNA used to induce mutations was: AAAATACTCAGTTTCTTTCA. Single-cell sorting was performed into a 96-well plate using a BD FACSAria III cell sorter after puromycin selection (2.0 μg/ml for 7 days). The mutations caused by Cas9 and the sgRNA were confirmed by PCR amplification, followed by Sanger sequencing. The PCR primers were fwd, TGTTGGGTCAGGAAGAGACTTG and rev, CCATGGACTGAACGGATTCC; and the Sanger sequencing primer was fwd, TGTTGGGTCAGGAAGAGACTTG.

Colony-forming assays. Single-cell suspensions were prepared from MEL202 and MEL270 cells with or without CRISPR-mediated insertions and deletions (indels), 3,000 cells from each cell line were plated in triplicates in 6-well treated plates and colonies were enumerated 10 days later. After 10 days, colonies were fixed with 3.7% paraformaldehyde for 5 min and stained in a solution of 0.05% crystal violet for 30 min at room temperature and washed in PBS and tap water.

In vitro cell viability assays. Cells were seeded in white flat-well 96-well plates (Costar) at a density of 1,000 cells per well. ATP luminescence readings were taken every 24 h after seeding, using Cell Titer Glo (Promega) according to the manufacturer's instructions.

***BRD9* minigene assay.** We identified putative exonic splicing enhancers with SpliceAid 2[39]. We used SF3B1 wild-type MEL270 and T47D cells transduced with Flag–SF3B1(WT) or Flag–SF3B1(K700E) (in the backbone of pInducer20 (Addgene no. 44012)). After drug selection with neomycin (Thermo Fisher Scientific 10131027), the selected cells were treated with 1 μg /ml doxycycline (Sigma D9891). The *BRD9* minigene construct was generated by inserting the DNA fragment containing the *BRD9* genomic sequence from exon 14 to exon 15 in between the BamHI and AgeI restriction sites in the FRE5 plasmid (Addgene 62377) via Gibson assembly. *BRD9* minigene mutagenesis was performed with the Agilent QuikChange II site-directed mutagenesis kit with specific primers according to the manufacturer's directions. For transient transfection experiments, cells were seeded into a 24-well plate one day before transfection of *BRD9* minigene constructs in the presence of X-tremeGENE HP DNA transfection reagent (Roche) according to the manufacturer's directions. Forty-eight hours after transfection, cells were collected and RNA was extracted using Qiagen RNeasy mini kit. Minigene-derived and endogenous BRD9 transcripts were analysed by RT–PCR using specific primers. Primers and oligonucleotides used in RT–PCR reactions were: cloning fwd GGACCCAGTACCAGGATCCGTTGGGGACACCCTAGGAG and rev CTTGGAGGAGCGCACACCGGTCAGGTGGTGCTGGTCCCTC; RT–PCR fwd (minigene) GGATTACAAGGATGACGATGAC, fwd (endogenous) CATGAAGCCTCCAGATGAAG and rev (common) CTTCGTCGT CTCATCCAAGTTC; mutagenesis (AAAA to TTTT) fwd CAAAGGGATATA

TTTTGAGATTTTTTACTCAGTTTCTTTCAGG and rev CCTGAAAG AAACTGAGTAAAAAATCTCAAAATATATCCCTTTG; mutagenesis (AAAA to ATAT) fwd CAAAGGGATATATTTTGAGATATATTACTCAGTTTCTTTCAGG and rev CCTGAAAGAAACTGAGTAATATATCTCAAAATATATCCCTTTG; mutagenesis (GATAAAA to TTTTTTT) fwd GTCAAAGGGATATATTTT GATTTTTTTTACTCAGTTTCTTTCAGG and rev ; mutagenesis (TTTCT to TTACA at exon 14a) fwd GAGATAAAATACTCAGTTACATTCAG GGCCTGCCATCTATC and rev GATAGATGGCAGGCCCTGAATG TAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTGCG at exon 14a) fwd GAGATAAAATACTCAGTTGCGTTCAGGGCCTGCCATCTATC, and rev GATAGATGGCAGGCCCTGAACGCAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTTCA at exon 14a) fwd GAGATAAAATAC TCAGTTTCATTCAGGGCCTGCCATCTATC and rev GATAGATGGCAG GCCCTGAATGAAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTTCG at exon 14a) fwd GAGATAAAATACTCAGTTTCGTTCAGG GCCTGCCATCTATC, and rev GATAGATGGCAGGCCCTGAAC GAAACTGAGTATTTTATCTC; mutagenesis (TTTCT to TTTAT at exon 14a) fwd GAGATAAAATACTCAGTTTATTTCAGGGCCTGCCATCTATC and rev GATAGATGGCAGGCCCTGAAATAAACTGAGTATTTTATCTC; and mutagenesis (branchpoint A to G) fwd GTCAAAGGGATATATTTTGG GATAAAATACTCAGTTTCTTTC and rev GAAAGAAACTGAGTAT TTTATCCCAAAATATATCCCTTTGAC.

Lariat sequencing. To map the branchpoints that were used when the *BRD9* poison exon was included or excluded, RT–PCR was performed to amplify branchpoint-spanning fragments from lariat RNAs arising during normal (poison exon exclusion) or aberrant (poison exon inclusion) splicing of *BRD9* pre-mRNA. In brief, SuperScript III reverse transcriptase (Invitrogen) and a primer complementary to the intronic sequences downstream of the 5′ splice sites were used to generate cDNA from lariat RNAs. Branchpoint-spanning fragments were then amplified from lariat RNAs by nested PCR with pairs of outer primers (with the RT primer being the reverse primer) and inner primers. The forward primers were complementary to sequences about 200–300 nucleotides upstream of the 3′ splice sites and the reverse primers were complementary to sequences downstream of the 5′ splice sites (Supplementary Table 11). The PCR products were cloned into the pGEM-T vector (Promega) and sequenced by Sanger sequencing.

Melanoma transplant model. We resuspended stably transduced Melan-a, MEL270 and MEL202 cells (1M cells) with doxycycline-inducible shRNAs, cDNAs or sgRNAs in 100 μl of a 1:1 mix of medium and Matrigel (BD Biosciences), and subcutaneously and bilaterally injected the mix into the flanks of 7-week-old female CB17-SCID mice (Taconic). For doxycycline-regulated shRNA induction, we used doxycycline-containing diets (625 mg/kg diet, Envigo). To assess tumour growth, at least five mice per group were injected for a total of ten tumours per group. No randomization or blinding was used in the analysis of tumour growth. Tumours were measured with callipers every seven days. Growth curves were visualized with Prism GraphPad 8.0. Tumour volume was calculated using the formula; Volume = π(length)(width)(height)/6.

In vitro morpholino transfection. To deliver morpholinos into cultured cell lines, we followed the manufacturer's instruction (GeneTools). In brief, we used 6 μM Endo-Porter after adding morpholinos (final concentration of 10 μM). RNA and proteins are collected 48 h after delivery. Morpholino target sequence no. 3 was TAATGAGGCAAGTCCAGTCCCGCTT; no. 6 was AAAGAGGGGAT AATGAGGCAAGTCC; and no. 7 was GGGATAATGAGGCAAGTCCAGTCCC.

In vivo morpholino treatment. Treatment with morpholinos was started when the tumour volume in mice reached 100–200 mm^3. Cohorts were treated intratumorally with 12.5 mg/kg scrambled or poison-exon-targeting Vivo-Morpholinos (AAAGAGGGGATAATGAGGCAAGTCC, GeneTools) dissolved in 50 μl PBS, every 2 days for 8 doses in total. The mice were dissected 24 h after the final treatment. For the patient-derived xenograft model, patient-derived rectal melanoma cells (*SF3B1*^{R625C}) and UVM cells (*SF3B1* wild type) were serially transplanted into SCID mice and treated similarly.

In vivo metastasis model. For lung experimental metastasis, B16 and 92.1 melanoma cells retrovirally transduced with shRNAs targeting *Renilla*, *BRD9* (no. 1 and no. 1) or *Brd9* (no. 1 or no. 2) in MLS-E vector (sorted using GFP) were trypsinized, resuspended in PBS and then 0.4 M cells in 0.2 ml PBS were injected via the lateral tail vein using a 27-gauge needle. Mice were killed 14 days after injection and tissues were isolated and fixed in 4% paraformaldehyde (Thermo Fisher Scientific). For evaluation of metastatic colonization of the lung using 92.1 human UVM cells, the burden of metastatic cells was evaluated using GFP expression by flow cytometry as well as anti-GFP immunohistochemistry 14 days following tail-vein injection of 0.4 M cells into NOD–SCID *Il2rg*$^{-/-}$ mice.

Histological analysis. Tissues were fixed in 4% paraformaldehyde, processed routinely in alcohol and xylene, embedded in paraffin, sectioned at 5-μm thickness and stained with H&E. Immunohistochemistry was performed on a Leica Bond RX automated stainer (Leica Biosystems). Following heat-induced epitope retrieval at pH 6.0, the primary antibody against Ki67 (Vector VP-K451) was applied, followed by application of a polymer detection system (DS9800, Novocastra Bond Polymer Refine Detection, Leica Biosystems) in which the chromogen was 3,3-diaminobenzidine tetrachloride (DAB) and the counterstain was haematoxylin. Photomicrograph examination of all H&E and immunohistochemistry slides were performed using a Zeiss Axioskop imaging.

***BRD9* expression correlates.** The cor.test (in R) was used to calculate Spearman's ρ and the P value associated with the correlation of *BRD9* expression with the expression of each coding gene across all samples within each cohort from the TCGA. Analysis was restricted to coding genes that are not on the same chromosome arm as *BRD9* (chromosome 5p) to remove potential confounding effects of local correlations. Coding genes with $P < 0.01$ in at least 10 cancer types were ranked by their absolute mean value of ρ (computed across all TCGA cohorts) and classified as RNA-binding if they were annotated with the 'RNA-binding' Gene Ontology term (GO: 0003723).

***BRD9* alternative splicing.** Potential NMD-targeted isoforms of *BRD9* were identified as follows: we queried the MISO v.2.0 alternative splicing annotation[40] for exon skipping and competing splice site events within the *BRD9* gene locus, restricted to those events with evidence of alternative splicing based on spliced junction reads (described in 'Genome annotation, RNA-seq read mapping, and estimation of gene and isoform expression'), assigned open reading frames for the isoforms resulting from each alternative splicing event based on the *BRD9* isoform with the longest open reading frame, and classified isoforms as predicted NMD substrates if they contained a termination codon >50 nt upstream of a splice junction.

Robust linear modelling of *BRD9* expression on the basis of the identified alternatively spliced isoforms of *BRD9* that are predicted NMD substrates was performed for each TCGA cohort with the rlm function in the MASS package in R. Relative expression of each isoform in each sample was estimated from RNA-seq data across all TCGA cohorts as described in 'Genome annotation, RNA-seq read mapping, and estimation of gene and isoform expression'. A z-score normalization was performed across all samples for each isoform in each cohort before model fitting. The resulting coefficients from the fitted models were subsequently used to predict *BRD9* expression from *BRD9* NMD-targeted isoform expression.

RNA-seq library preparation. RNA-seq libraries were prepared from TRIzol-isolated (Thermo Fisher cat. no. 15596026) RNA using the Illumina TruSeq RNA Library Prep Kit v.2 (Illumina cat. no. RS-122-2001/2). K562 libraries were sequenced at MSKCC with 101-bp single-end reads. MEL270 libraries were sequenced by the FHCRC Genomics Shared Resource with 2 × 51-bp paired-end reads.

ChIP–seq library preparation. ChIP–seq libraries were prepared and sequenced as previously described[6] by the Molecular Biology Core Facilities at the Dana-Farber Cancer Institute with 75-bp single-end reads.

Genomic analysis of SWI–SNF complex members from TCGA. Mutational analysis of genes encoding members of the SWI–SNF complex was performed as previously described[41].

Genome annotation, RNA-seq read mapping, and estimation of gene and isoform expression. RNA-seq reads were processed for gene expression and isoform ratio quantification as previously described[42]. In brief, RNA-seq reads were aligned to the hg19/GRCh37 assembly of the human genome using a gene annotation created by merging the UCSC knownGene gene annotation[43], Ensembl v.71.1 gene annotation[44] and MISO v.2.0 isoform annotation[40]. Read alignment and expression estimation were performed with RSEM v.1.2.4[45], Bowtie v.1.0.0[46] and TopHat v.2.1.1[47]. Isoform ratios were quantified with MISO v.2.0[40]. Gene expression estimates were normalized by applying the trimmed mean of M values method[48] to coding genes. Statistical tests for differential gene and isoform expression were performed for single-sample comparisons with Wagenmakers' Bayesian framework[49] and for sample group comparisons with the Mann–Whitney U-test. RNA-seq read-coverage plots (for example, Fig. 1e) represent reads normalized by the number of reads mapping to all coding genes in each sample (per million).

RNA-seq coverage plots. RNA-seq coverage plots were made using the UCSC Genome Browser[50] and/or the ggplot2 package in R[51]. Repetitive elements were annotated by RepeatMasker[28].

Cluster analysis. Unsupervised clustering of chronic lymphocytic leukaemia, myelodysplastic syndrome and UVM samples (Fig. 1a) was based on the 40 events that were differentially spliced in isogenic UVM (MEL270 cells) as well as myeloid leukaemia (K562 cells) cells expressing *SF3B1*^{K700E} versus wild-type *SF3B1*, restricted to the 30 of these events that had sufficient read coverage in all cohorts for clustering.

ChIP–seq data analysis. ChIP–seq reads were mapped to the genome by calling Bowtie v.1.0.0[46] with the arguments '-v 2 -k 1 -m 1–best–strata'. Peaks were called using MACS2 v.2.1.1.20160309[52] against input control libraries with $P < 10^{-5}$ and subsequently filtered to remove peaks contained within ENCODE blacklisted regions[53] and the mitochondrial genome. Subsequent data analysis was performed with Bioconductor in the R programming environment[54]. Consensus

peaks between samples were called using the soGGI package v.1.14.0[55]. Peaks were annotated using the ChIPseeker package v.1.18.0[56]. Potential transcription factor binding in a 300-nucleotide region around the centre of consensus peaks was scored using the TFBSTools package v.1.20.0[57], with models taken from the HOCOMOCO v.11 human core collection[58] and applied with a threshold of $P < 10^{-4}$. The highest scores for each consensus peak region were collated for each transcription factor. A two-sided Mann–Whitney *U*-test was used to assess the significance of the difference in scores between constitutive and sensitive peaks for each transcription factor.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

RNA-seq and ChIP–seq data generated as part of this study were deposited in the Gene Expression Omnibus (accession number GSE124720). RNA-seq data from published studies were downloaded from CGHub (TCGA UVM[59]), EMBL-EBI ArrayExpress (Illumina Human BodyMap 2.0: E-MTAB-513), the Gene Expression Omnibus (accession numbers GSE72790 and GSE114922 for chronic lymphocytic leukaemia[15] and myelodysplastic syndromes[27], respectively), or directly obtained from the authors (for UVM[10]). Gel source data can be found in Supplementary Fig. 1. Other data that support the findings of this study are available from the corresponding authors upon reasonable request.

29. Cheng, D. T. et al. Memorial Sloan Kettering-integrated mutation profiling of actionable cancer targets (MSK-IMPACT): a hybridization capture-based next-generation sequencing clinical assay for solid tumor molecular oncology. *J. Mol. Diagn.* **17**, 251–264 (2015).
30. Seiler, M. et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* **24**, 497–504 (2018).
31. Leeksma, A. C. et al. Clonal diversity predicts adverse outcome in chronic lymphocytic leukemia. *Leukemia* **33**, 390–402 (2019).
32. Kim, E. et al. *SRSF2* mutations contribute to myelodysplasia by mutant-specific effects on exon recognition. *Cancer Cell* **27**, 617–630 (2015).
33. Mohammed, H. et al. Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. *Cell Rep.* **3**, 342–349 (2013).
34. Fellmann, C. et al. An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
35. Pelossof, R. et al. Prediction of potent shRNAs with a sequential classification algorithm. *Nat. Biotechnol.* **35**, 350–353 (2017).
36. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
37. McCarthy, D. J., Chen, Y. & Smyth, G. K. Differential expression analysis of multifactor RNA-seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).
38. Wu, D. & Smyth, G. K. Camera: a competitive gene set test accounting for inter-gene correlation. *Nucleic Acids Res.* **40**, e133 (2012).
39. Piva, F., Giulietti, M., Burini, A. B. & Principato, G. SpliceAid 2: a database of human splicing factors expression data and RNA target motifs. *Hum. Mutat.* **33**, 81–85 (2012).
40. Katz, Y., Wang, E. T., Airoldi, E. M. & Burge, C. B. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat. Methods* **7**, 1009–1015 (2010).
41. Mashtalir, N., et al. Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288.e1220 (2018).
42. Dvinge, H. et al. Sample processing obscures cancer-specific alterations in leukemic transcriptomes. *Proc. Natl Acad. Sci. USA* **111**, 16802–16807 (2014).
43. Meyer, L. R. et al. The UCSC genome browser database: extensions and updates 2013. *Nucleic Acids Res.* **41**, D64–D69 (2013).
44. Flicek, P. et al. Ensembl 2013. *Nucleic Acids Res.* **41**, D48–D55 (2013).
45. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
46. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
47. Trapnell, C., Pachter, L. & Salzberg, S. L. TopHat: discovering splice junctions with RNA-seq. *Bioinformatics* **25**, 1105–1111 (2009).
48. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
49. Wagenmakers, E. J., Lodewyckx, T., Kuriyal, H. & Grasman, R. Bayesian hypothesis testing for psychologists: a tutorial on the Savage-Dickey method. *Cognit. Psychol.* **60**, 158–189 (2010).
50. Kent, W. J. et al. The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
51. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2016).
52. Zhang, Y. et al. Model-based analysis of ChIP–seq (MACS). *Genome Biol.* **9**, R137 (2008).
53. The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
54. Huber, W. et al. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat. Methods* **12**, 115–121 (2015).
55. Dharmalingam G. & Carroll, T. soGGi: Visualise ChIP–seq, MNase-seq and motif occurrence as aggregate plots summarised over grouped genomic intervals. R package version 1.14.0, https://rdrr.io/bioc/soGGi/ (2018).
56. Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).
57. Tan, G. & Lenhard, B. TFBSTools: an R/bioconductor package for transcription factor binding site analysis. *Bioinformatics* **32**, 1555–1556 (2016).
58. Kulakovskiy, I. V. et al. HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. *Nucleic Acids Res.* **46**, D252–D259 (2018).
59. Robertson, A.G., et al. Integrative analysis identifies four molecular and clinical subsets in uveal melanoma. *Cancer cell* **32**, 204–220.e215 (2017).
60. Pollard, K. S., Hubisz, M. J., Rosenbloom, K. R. & Siepel, A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res.* **20**, 110–121 (2010).
61. McLean, C. Y. et al. GREAT improves functional interpretation of *cis*-regulatory regions. *Nat. Biotechnol.* **28**, 495–501 (2010).

Acknowledgements We are grateful for the support of the MSK RNAi core facility for help with the CRISPR screens performed in the study, the MSK anti-tumor assessment core facility for help with patient-derived xenograft experiments and the Genomics Shared Resource of the Fred Hutchinson/University of Washington Cancer Consortium (P30 CA015704). D.I., S.C.-W.L., A.Y. and O.A.-W. are supported by the Leukemia & Lymphoma Society. D.I. is supported by grants from Lydia O'Leary Memorial Pias Dermatological Foundation and Kobayashi Foundation for Cancer Research. A.Y. is supported by grants from the Aplastic Anemia and MDS International Foundation (AA&MDSIF) and the Lauri Strauss Leukemia Foundation. S.X.L. is supported by a Conquer Cancer Foundation and ASCO Young Investigator Award, an Aplastic Anemia & Myelodysplastic Syndrome International Foundation research award, as well as an AACR Lymphoma Research Fellowship. G.-L.C. is a Mahan Fellow. O.A.-W. is supported by the Pershing Square Sohn Cancer Research Alliance, the Henry & Marilyn Taub Foundation and the Starr Cancer Consortium. R.K.B. is a Scholar of The Leukemia and Lymphoma Society (1344-18) and is supported in part by the US National Institutes of Health (R01 DK103854). O.A.-W. and R.K.B. are supported by the Evans MDS Foundation, the US National Institutes of Health (R01 HL128239) and the Department of Defense Bone Marrow Failure Research Program (BM150092 and W81XWH-12-1-0041). The results shown here are in part based upon data generated by the TCGA research network (https://cancergenome.nih.gov/).

Author contributions D.I., G.-L.C., B.L., B.C.M., C.K., O.A.-W. and R.K.B. designed the study. D.I., B.L. and J.P. performed minigene assays. D.I., B.C.M. and C.K. performed ChIP–seq experiments. A.R.D., G.-L.C. and C.K. performed ChIP–seq analyses. G.-L.C., K.N., A.P. and R.K.B. performed computational analyses of the CRISPR screen, RNA splicing and gene expression data. D.I. and S.X.L. performed the CRISPR screen. D.I. and B.L. generated and validated the anti-BRD9 PE morpholinos in in vitro assays. D.I., L.B., A.B., S.C.-W.L., A.Y. and H.C. performed the mouse experiments. A.R.M., D.I. and L.E.-H. performed experiments using pancreatic cancer cell lines. T.H. and Y.C. provided melanoma cell lines. D.I., G.-L.C., O.A.-W. and R.K.B. prepared the manuscript with help from all co-authors. J.T. and O.A.-W. provided clinically annotated samples from the MSK Hematology/Oncology Tissue Bank and from the MSK Antitumor Assessment Core Facility. O.A.-W. and R.K.B. provided funding and study supervision.

Competing interests C.K. is a Scientific Founder, fiduciary Board of Directors member, Scientific Advisory Board member, consultant and shareholder of Foghorn Therapeutics, none of which are related to the current manuscript. O.A.-W. has served as a consultant for H3 Biomedicine, Foundation Medicine, Merck and Janssen; O.A.-W. has received personal speaking fees from Daiichi Sankyo. O.A.-W. has received previous research funding from H3 Biomedicine unrelated to the current manuscript. D.I., O.A.-W. and R.K.B. are inventors on a provisional patent application submitted by the Fred Hutchinson Cancer Research Center that covers BRD9 activation in cancer.

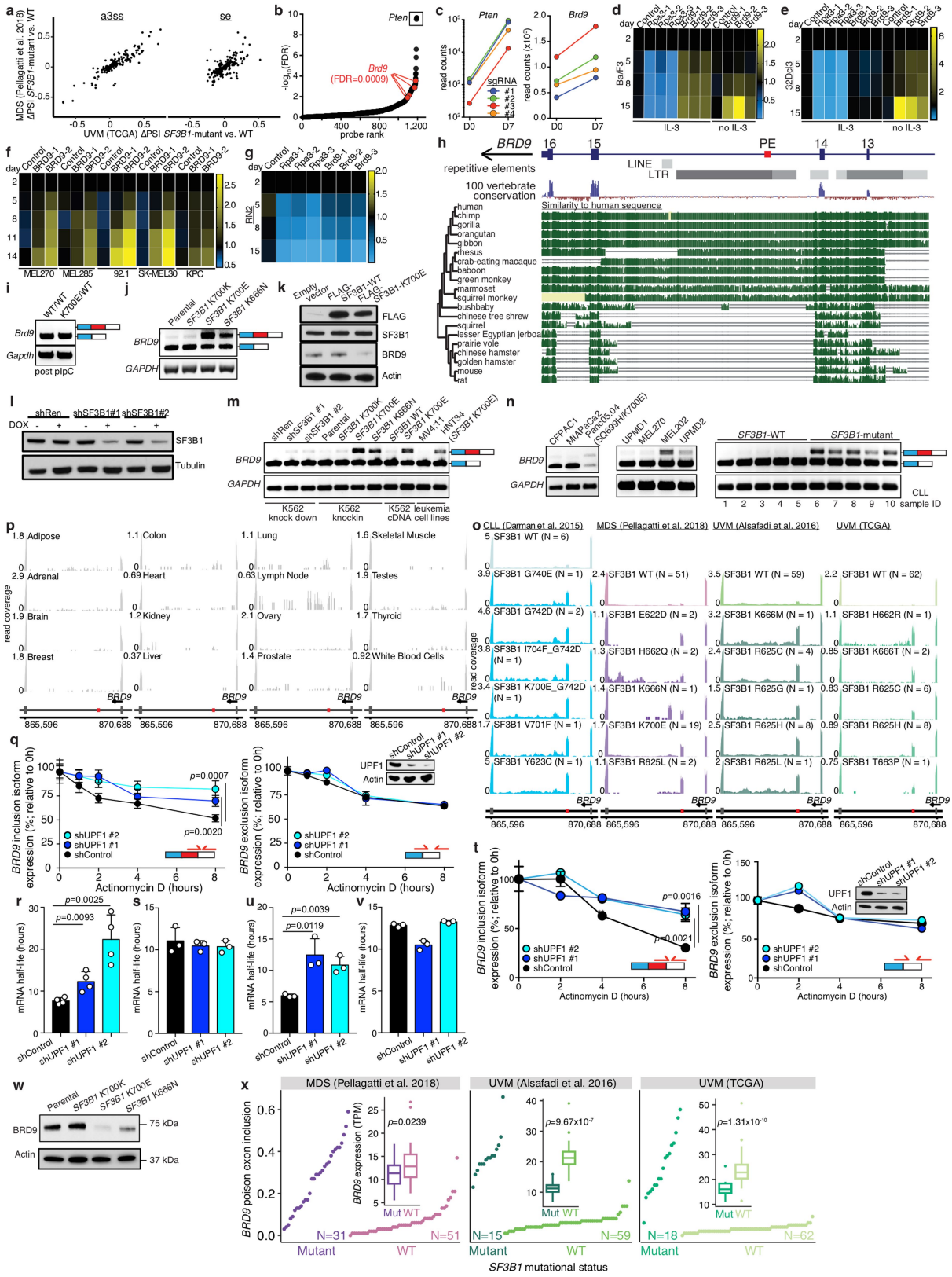
Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1646-9.

Correspondence and requests for materials should be addressed to O.A. or R.K.B.

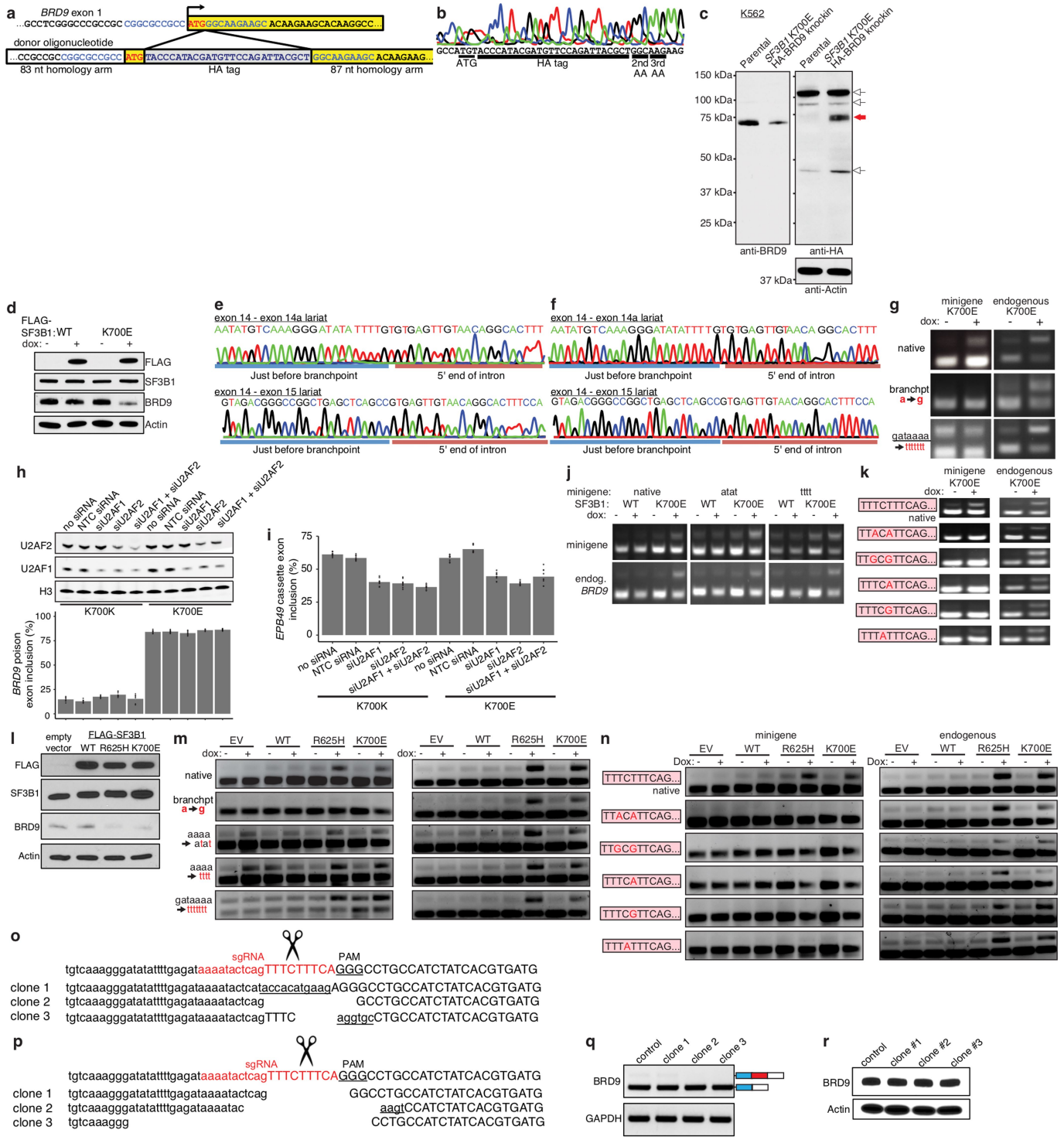
Peer review information *Nature* thanks Boris Bastian, Rotem Karni and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at http://www.nature.com/reprints.



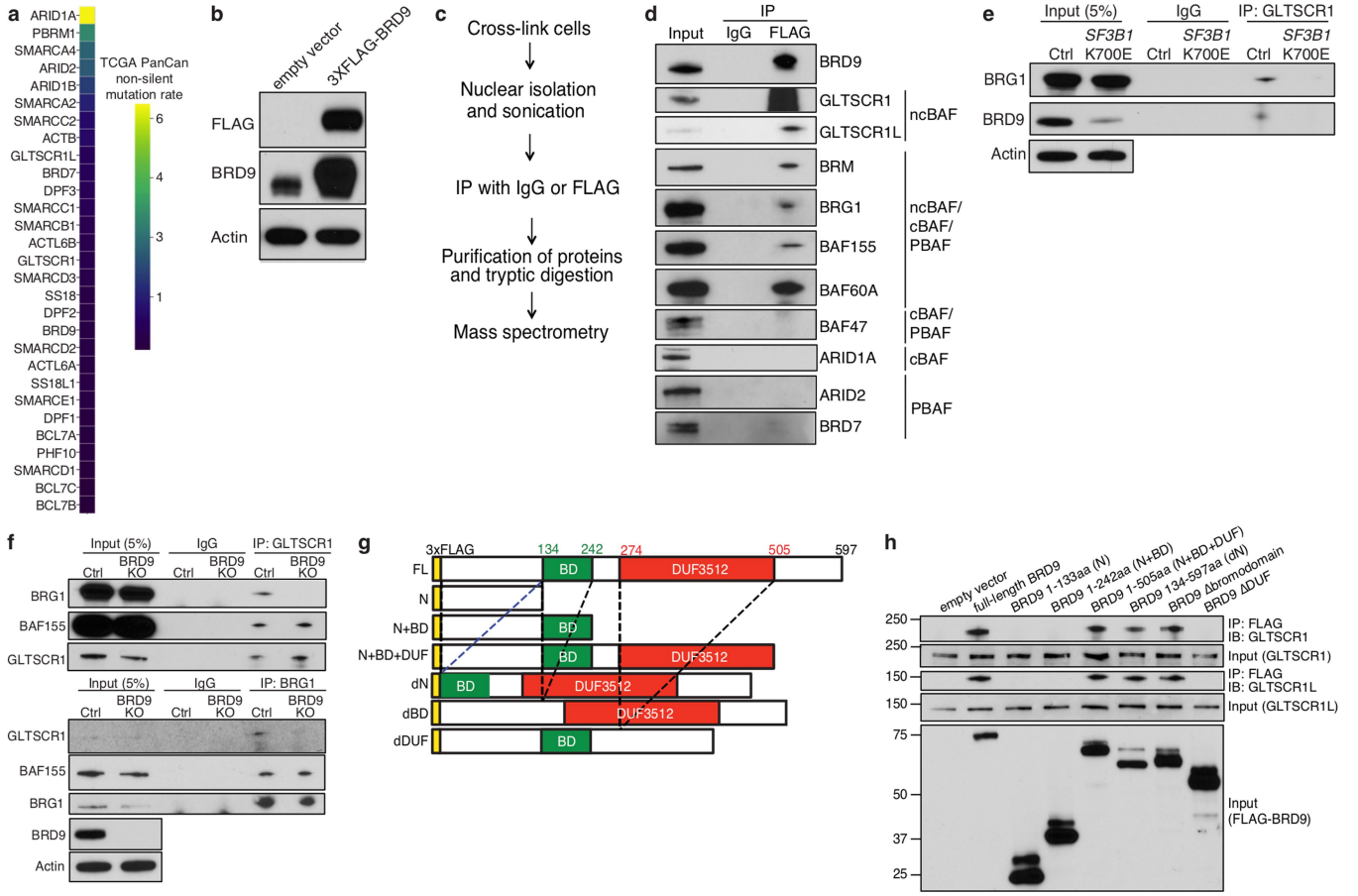
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | *BRD9* is mis-spliced in *SF3B1*-mutated human cells, and BRD9 loss confers a proliferative advantage. a, Scatter plots comparing differential splicing (ΔPSI) between patients in the TCGA UVM cohort with *SF3B1* mutations or wild-type *SF3B1*[59] (*x* axis) and patients from a myelodysplastic syndromes cohort with *SF3B1* mutations or wild-type *SF3B1*[27] (*y* axis). Events were classified as alternative 3′ splice sites or skipped exons. **b**, Rank plot for the $-\log_{10}$(FDR) associated with each sgRNA in our CRISPR–Cas9 positive-selection screen. sgRNAs targeting the positive control (*Pten*) and *Brd9* are highlighted. For the probe-level (per-sgRNA) analysis, we fitted a negative binomial generalized log-linear model and performed a likelihood ratio test. FDR values were computed using the Benjamini–Hochberg method. **c**, Read counts for sgRNAs targeting the positive control (*Pten*) or *Brd9*. D0 and D7 indicate days following withdrawal of IL-3. **d**, Heat map summarizing the results of a competition assay to measure the effect of each indicated sgRNA on the growth of Cas9-expressing Ba/F3 cells. Cell growth was computed with respect to cells treated with a non-targeting (control) sgRNA and the percentages of GFP^+ cells on day 14 were normalized to the percentages on day 2. The illustrated values correspond to the mean computed value over $n = 3$ biological replicates. *Rpa3* sgRNAs were used as a negative control. **e**, As in **d**, but for 32Dcl3 cells. **f**, As in **d**, but for the indicated melanoma and pancreatic ductal adenocarcinoma cells. **g**, As in **d**, but for RN2 cells. **h**, Sequence conservation of the *BRD9* poison exon locus as estimated by phyloP[60]. Conservation and repetitive element annotation is from the UCSC Genome Browser[43]. **i**, RT–PCR analysis of *Brd9* poison exon inclusion using whole bone marrow cells from *Mx1-cre Sf3b1*$^{WT/WT}$ (WT/WT) and *Mx1-cre Sf3b1*$^{K700E/WT}$ (K700E/WT) mice. Three weeks after pIpC treatment, RT–PCR was performed with mouse primers corresponding to those used to assay *BRD9* poison exon inclusion in human cells. Representative images from $n = 2$ technically independent replicates. **j**, RT–PCR analysis to confirm mutant-SF3B1-dependent inclusion of the *BRD9* poison exon in isogenic NALM-6 cell lines engineered to contain the indicated mutations. *SF3B1*^{K700K} is a wild-type control for genome engineering. Representative images from $n = 2$ technically independent replicates. **k**, Western blot for Flag, SF3B1 and BRD9 in K562 cells overexpressing N-terminally Flag-tagged wild-type *SF3B1* or *SF3B1*^{K700E} cDNAs, or an empty vector; this panel corresponds to the cells evaluated in **m**. Representative images from $n = 2$ biologically independent replicates. **l**, Western blot for SF3B1 in K562 cells treated with doxycycline-inducible *SF3B1*-targeting shRNAs or a non-targeting control shRNA (shRen); this panel corresponds to cells evaluated in **m**. Representative images from $n = 2$ technically independent replicates. **m**, RT–PCR illustrating the specificity of *BRD9* poison exon inclusion for *SF3B1*-mutated cells in the indicated cell lines. These include K562 cells treated with control shRNA (shRen) or *SF3B1*-targeting shRNAs (the columns labelled 'K562 knock-down'); knock-in of the *SF3B1*^{K700K}, *SF3B1*^{K700E} or *SF3B1*^{K666N} mutation into the endogenous locus of *SF3B1* (the columns labelled 'K562 knock-in'); or overexpression of wild-type *SF3B1* or *SF3B1*^{K700E} cDNA (the columns labelled 'K562 cDNA'). The two right-most lanes show acute myeloid leukaemia cell lines with wild-type *SF3B1* (MV4;11) or a naturally occurring endogenous *SF3B1*^{K700E} mutation (HNT34 cells; the columns labelled 'leukaemia cell lines'). Representative images from $n = 3$ biologically independent experiments. **n**, As in **m**, but for the indicated pancreatic ductal adenocarcinoma cell lines (left), UVM cell lines (centre) and a cohort of patients with chronic lymphocytic leukaemia (right). CFPAC1 and MIA PaCa2 cells lack *SF3B1* mutations; Panc05;04 cells carry *SF3B1*$^{Q699H/K700E}$; UPMD1 and MEL270 cells lack *SF3B1* mutations; MEL202 and UPMD2 cells carry *SF3B1*^{R625G} and *SF3B1*^{Y623H} mutations, respectively. Sample identifiers for patients with chronic lymphocytic leukaemia correspond to the genotypes shown in Supplementary Table 7. Representative images from $n = 2$ technically independent experiments (left and centre) and $n = 3$ biologically independent experiments (right). **o**, RNA-seq read coverage plots of the *BRD9* poison exon locus from patient samples with the indicated *SF3B1* genotypes. All *SF3B1*-mutated samples exhibit *BRD9* poison exon inclusion. **p**, As in **o**, but for the indicated tissues from healthy donors (from BodyMap 2.0). **q**, qRT–PCR measurement of the half-lives of the poison exon inclusion (left) and exclusion (right) isoforms in isogenic K562 *SF3B1*^{K700E} cells treated with the indicated shRNAs and actinomycin D to inhibit transcription. NMD inhibition via *UPF1* knockdown stabilizes the inclusion, but not exclusion, isoform. Red arrows indicate primers used to specifically detect the two isoforms. $n = 2$ biologically independent experiments and $n = 2$ technically independent experiments for the inclusion isoform; $n = 3$ technically independent experiments for the exclusion isoform. *P* value was calculated by two-sided *t*-test at 8 h. **r**, Bar graph illustrating the estimated poison exon inclusion isoform half-life in the indicated conditions from the data in **q**. Error bars, mean + s.d. $n = 2$ biologically independent experiments and $n = 2$ technically independent experiments. *P* value was calculated by two-sided *t*-test. **s**, As in **r**, but for the exclusion isoform. Error bars, mean + s.d. $n = 3$ technically independent experiments. *P* value was calculated by two-sided *t*-test. **t**, As in **q**, but for NALM-6 *SF3B1*^{K700E} cells. $n = 3$ technically independent experiments for the inclusion isoform and the exclusion isoform. *P* value was calculated by two-sided *t*-test at 8 h. **u**, As in **r**, but for NALM-6 *SF3B1*^{K700E} cells. $n = 3$ technically independent experiments. Error bars, mean + s.d. *P* value was calculated by two-sided *t*-test. **v**, As in **s**, but for NALM-6 *SF3B1*^{K700E} cells. Error bars, mean + s.d. $n = 3$ technically independent experiments. *P* value was calculated by two-sided *t*-test. **w**, Western blot for BRD9 in NALM-6 cells with or without knock-in of an *SF3B1* mutation. Actin, loading control. Representative images from $n = 3$ biologically independent experiments. **x**, Rank plot of *BRD9* poison exon inclusion (scale of 0 to 1; top) and box plot of gene expression (inset) for patients stratified by *SF3B1* mutational status (data are from cohorts of patients with myelodysplastic syndromes or UVM, as in Fig. 1a). *SF3B1* mutations were strongly associated with high poison exon inclusion and low *BRD9* expression. Boxes illustrate 1st and 3rd quartiles, with whiskers extending to 1.5× interquartile range. *P* value computed with one-sided Mann–Whitney *U*-test.

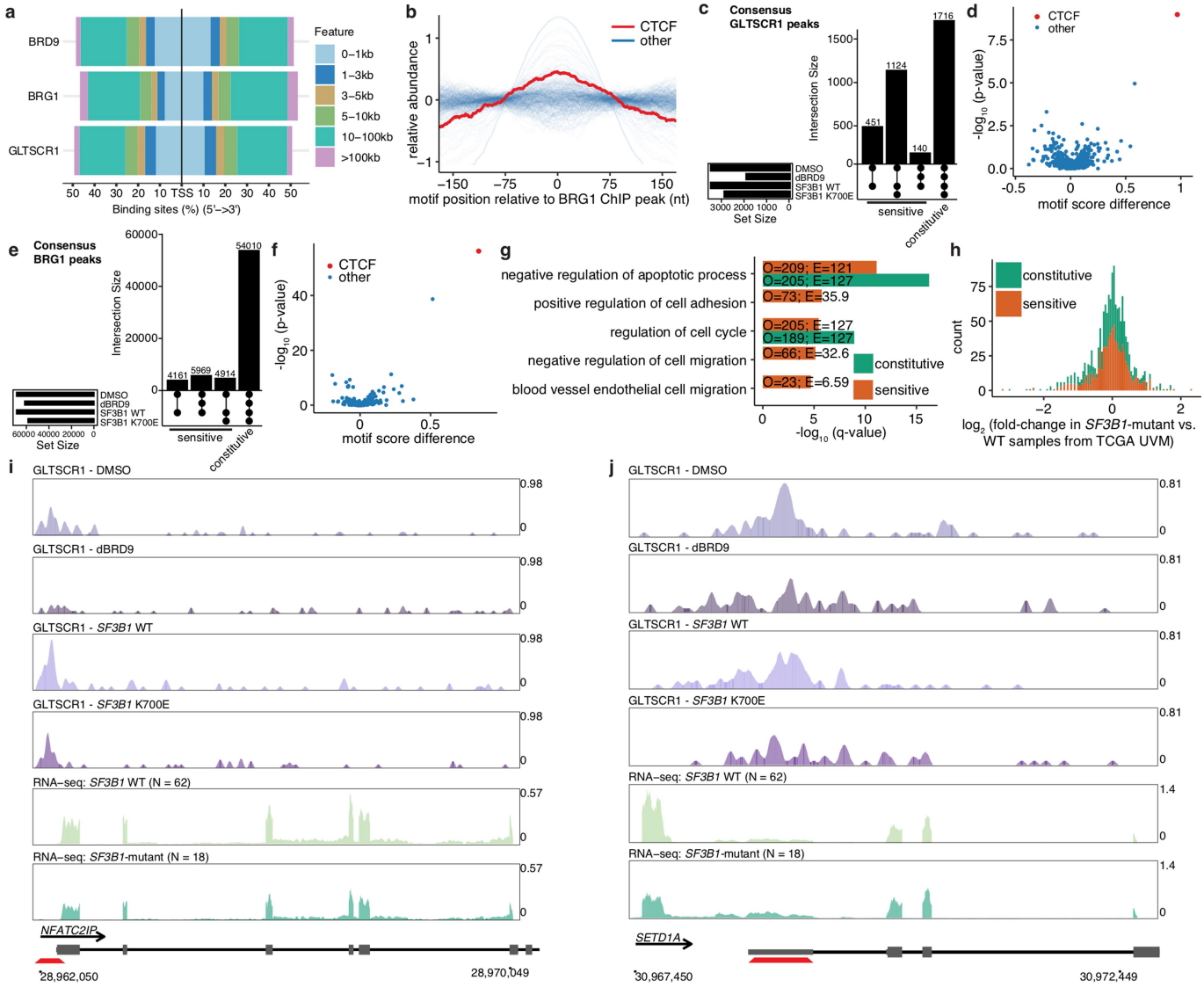


Extended Data Fig. 2 | See next page for caption.

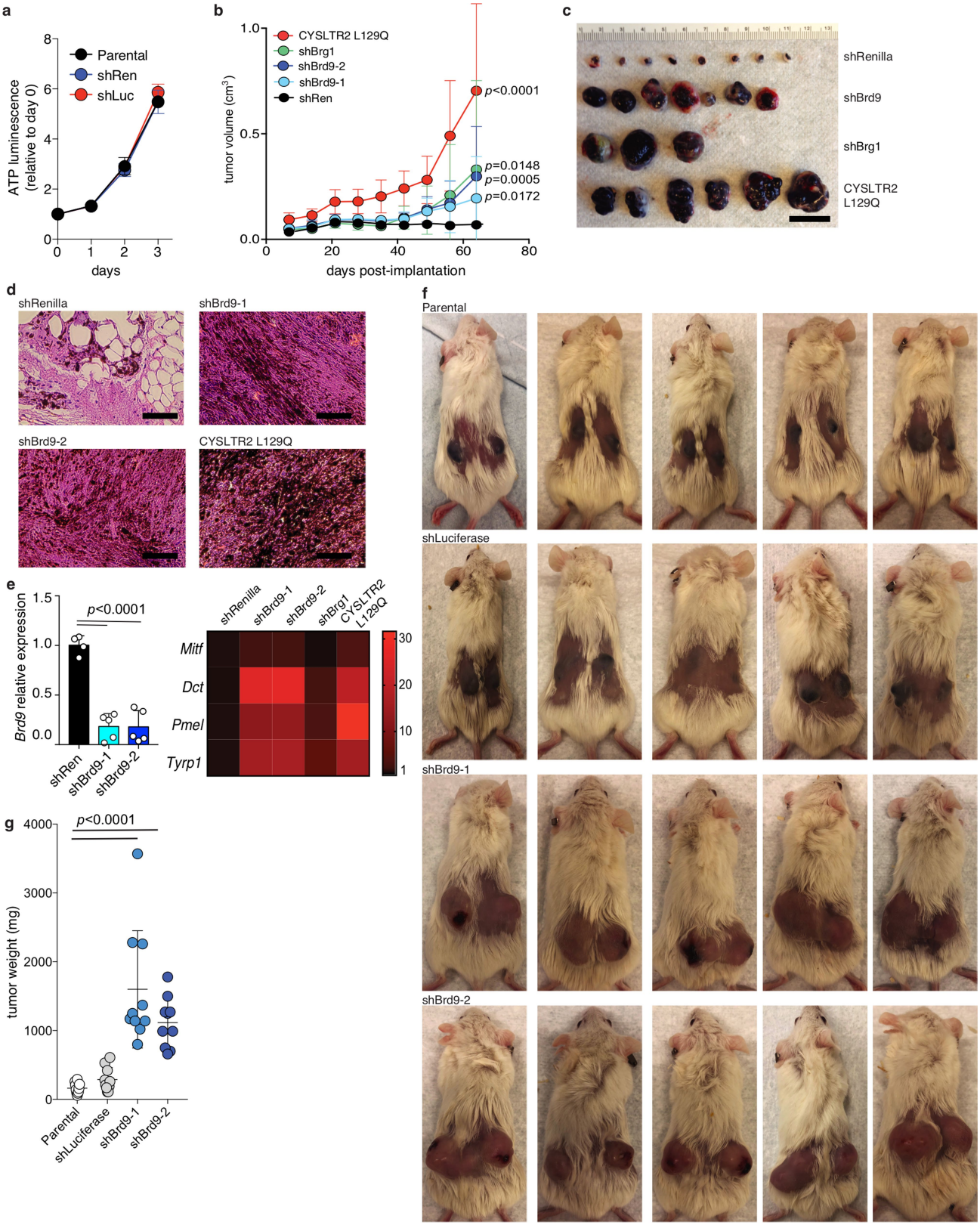
Extended Data Fig. 2 | Mutant SF3B1 recognizes an aberrant branchpoint within *BRD9* to promote poison exon inclusion, causing loss of full-length BRD9 protein. a, Schematic illustrating the strategy for knock-in of an HA tag into the endogenous *BRD9* locus. The single-stranded donor DNA contained a 197-nt fragment, including 83 nt homologous to the *BRD9* 5′ UTR (upstream of the HA tag) and 87 nt homologous to *BRD9* exon 1 (downstream of the start codon). **b**, Sanger sequencing of genomic DNA validating successful HA tag knock-in in K562 *SF3B1*^{K700E} cells. Representative images from $n = 2$ biologically independent experiments. **c**, Western blot with anti-BRD9 (left), anti-HA (right, top) or anti-actin (right, bottom) used to probe K562 *SF3B1*^{K700E} cells carrying an endogenously N-terminally HA-tagged BRD9. White arrows, non-specific bands. Red arrow, expected size of BRD9 protein. Representative images from $n = 3$ biologically independent experiments. **d**, Western blot for Flag, SF3B1 and endogenous BRD9 protein in MEL270 cells with doxycycline-inducible Flag-tagged wild-type SF3B1 or Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **e**, Sanger sequencing of cDNA arising from reverse transcription of lariats arising from inclusion (top) (exon 14–exon 14a splicing) or exclusion (bottom) (exon 14–exon 15 splicing) of the *BRD9* poison exon in MEL270 cells with doxycycline-inducible Flag-tagged wild-type SF3B1 (bottom) or Flag-tagged SF3B1(K700E) (top). The branchpoints are illustrated in Fig. 2a. Representative images from $n = 3$ biologically independent experiments. **f**, As in **e**, but for T47D cells. Representative images from $n = 3$ biologically independent experiments. **g**, As in Fig. 2b, but for the indicated minigene mutagenesis in T47D cells with doxycycline-inducible Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **h**, Western blot of U2AF2, U2AF1 and histone H3 in K562 cells transfected with siRNAs against *U2AF1* and/or *U2AF2* (top) and bar plot illustrating mean *BRD9* poison exon inclusion as measured by quantitative PCR (qPCR) following siRNA knockdown of *U2AF1* and/or *U2AF2* (bottom). Experiment performed with $n = 1$ biologically independent replicate for siRNA transfection, $n = 1$ technically independent replicate for western blot and $n = 3$ technically independent replicates for RT–PCR. Poison exon inclusion was computed over all $n = 3 \times 3$ (9) combinations of technical replicates for RT–PCR for the inclusion and exclusion isoforms. Bars illustrate mean inclusion. **i**, *EPB49* cassette exon inclusion as measured by qPCR following siRNA knockdown of *U2AF1* and/or *U2AF2*. As the *EPB49* cassette exon is U2AF-dependent, this experiment serves as a positive control for the functional efficacy of *U2AF1* and *U2AF2* knockdown. $n = 3$ technically independent experiments. Cassette exon inclusion was computed over all $n = 3 \times 3$ (9) combinations of technical replicates for RT–PCR for the inclusion and exclusion isoforms. Bars illustrate mean inclusion. **j**, As in Fig. 2b, but for the indicated minigene mutagenesis in T47D cells with doxycycline-inducible Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **k**, As in Fig. 2c, but for the indicated minigene mutagenesis in T47D cells with doxycycline-inducible Flag-tagged SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **l**, Western blot for Flag, SF3B1, BRD9 and actin in MEL270 cells expressing an empty vector or N-terminally Flag-tagged wild-type *SF3B1*, *SF3B1*^{R625H} or *SF3B1*^{K700E} cDNA. Representative images from $n = 3$ biologically independent experiments. **m**, RT–PCR analysis of *BRD9* splicing in MEL270 cells expressing doxycycline-inducible empty vector, wild-type SF3B1, SF3B1(R625H) or SF3B1(K700E). The left column illustrates minigene splicing and the right column illustrates endogenous *BRD9* splicing. Representative images from $n = 3$ biologically independent experiments. **n**, As in **m**, but for the illustrated minigene mutations at the 5′ end of the poison exon. Representative images from $n = 3$ biologically independent experiments. **o**, Mutations generated at the 5′ end of the *BRD9* poison exon by CRISPR–Cas9-mediated indels in MEL202 cells (*SF3B1*^{R625G}). The PAM sequence is illustrated with uppercase, underlined nucleotides. Red nucleotides hybridize to the sgRNA. Substitutions are illustrated with lowercase, underlined nucleotides. **p**, As in **o**, but for MEL270 cells. Representative images from $n = 3$ biologically independent experiments. **q**, As in Fig. 2d top, but for MEL270 cells. Representative images from $n = 3$ biologically independent experiments. **r**, As in Fig. 2d bottom, but for MEL270 cells. Representative images from $n = 3$ biologically independent experiments.



Extended Data Fig. 3 | BRD9 loss impairs ncBAF complex formation. **a**, Mutation rate observed across TCGA cohorts for canonical BAF, polybromo-associated BAF and ncBAF components. **b**, Western blot confirming Flag-tagged BRD9 protein expression in 3×Flag–BRD9-expressing K562 cells. Representative images from $n = 3$ biologically independent experiments. **c**, Experimental workflow for using rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME)[33] for purification and identification of the chromatin-associated interactions partners of BRD9. **d**, Cross-linking and immunoprecipitation with IgG or Flag followed by probing with the indicated antibodies. Data from 3×Flag–BRD9-expressing MEL270 cells. Representative images from $n = 3$ biologically independent experiments. **e**, Immunoprecipitation of GLTSCR1 followed by western blotting with the indicated antibodies in *SF3B1*^{K700E} knock-in NALM-6 cells. Representative images from $n = 3$ biologically independent experiments. **f**, Immunoprecipitation of GLTSCR1 (top) or BRG1 (bottom) followed by blotting with the indicated antibodies in K562 cells with CRISPR-mediated knockout (KO) of BRD9. Representative images from $n = 3$ biologically independent experiments. **g**, Schematic of the BRD9 full-length (FL) protein and the deletion mutants that we constructed. BD, bromodomain; DUF, domain of unknown function; N, amino acids 1–133 of BRD9; N + BD, amino acids 1–242 of BRD9; N + BD + DUF, amino acids 1–505 of BRD9; dN, amino acids 134–597 of BRD9; dBD, bromodomain-deletion mutant of BRD9; dDUF, DUF-domain deletion mutant of BRD9. **h**, Immunoprecipitation with Flag following by probing (immunoblot, IB) for GLTSCR1 or GLTSCR1L in 293T cells expressing 3×Flag-tagged versions of the indicated deletion mutants. Deletion mutants illustrated in **g**. Representative images from $n = 3$ biologically independent experiments.

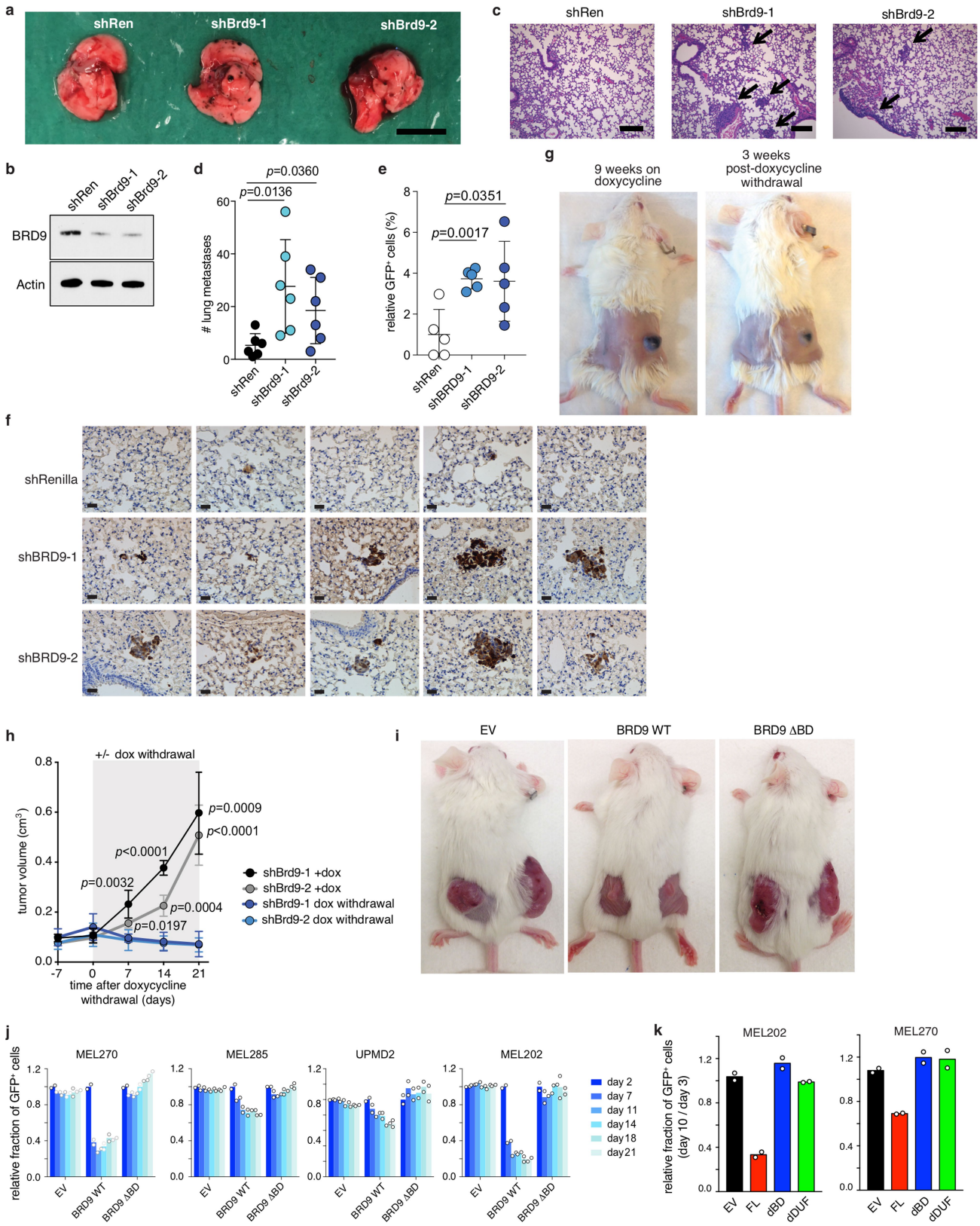


Extended Data Fig. 4 | BRD9 loss drives relocalization of GLTSCR1 away from CTCF-associated loci. a, As in Fig. 3e, but illustrating relative positions with respect to transcription start sites (TSSs). **b**, As in Fig. 3f, but for motifs at BRG1-bound loci. $n = 401$ transcription factors analysed. **c**, UpSet plots depicting the overlap of consensus GLTSCR1-bound loci in MEL270 cells with the indicated treatments. **d**, Volcano plot illustrating the difference in the mean motif scores at BRD9-sensitive versus constitutive GLTSCR1-bound loci for the transcription factors in Fig. 3f, as well as associated statistical significance. $n = 401$ transcription factors analysed. *P* values computed with a two-sided Mann–Whitney *U*-test. **e**, As in **c**, but for BRG1-bound loci. **f**, As in **d**, but for BRG1-bound loci. $n = 401$ transcription factors analysed. **g**, Selected enriched annotation terms from a Genomic Regions Enrichment of Annotations Tool (GREAT) analysis[61] of genes near BRD9-sensitive and constitutive GLTSCR1-bound loci. Plot illustrates $-\log_{10}$(FDR), computed with a one-sided binomial test and corrected for multiple testing using the Benjamini–Hochberg procedure. O and E, numbers of genes that were observed and expected, respectively. **h**, Differences in gene expression in *SF3B1*-mutated versus wild-type samples in the TCGA UVM cohort for genes with GLTSCR1-bound promoters identified in MEL270 cells. Colours indicate the responsiveness of peaks to BRD9 loss. **i**, Read coverage from GLTSCR1 ChIP–seq (MEL270 cells) and RNA-seq (TCGA UVM cohort) around *NFATC2IP*. Red trapezoid indicates GLTSCR1 binding in the promoter, with reduced binding upon treatment with BRD9 degrader or expression of $SF3B1^{K700E}$. *NFATC2IP* was significantly differentially expressed in UVM samples with *SF3B1* mutations relative to wild-type samples. Vertical axis scales were rendered comparable by normalizing ChIP–seq read coverage to mapped library size and RNA-seq read coverage to mapped library size, restricted to coding genes. ChIP–seq experiment performed for $n = 1$ biologically independent replicate. **j**, As in **i**, but for *SETD1A*. *SETD1A* was significantly differentially expressed in UVM samples with *SF3B1* mutations relative to wild-type samples.



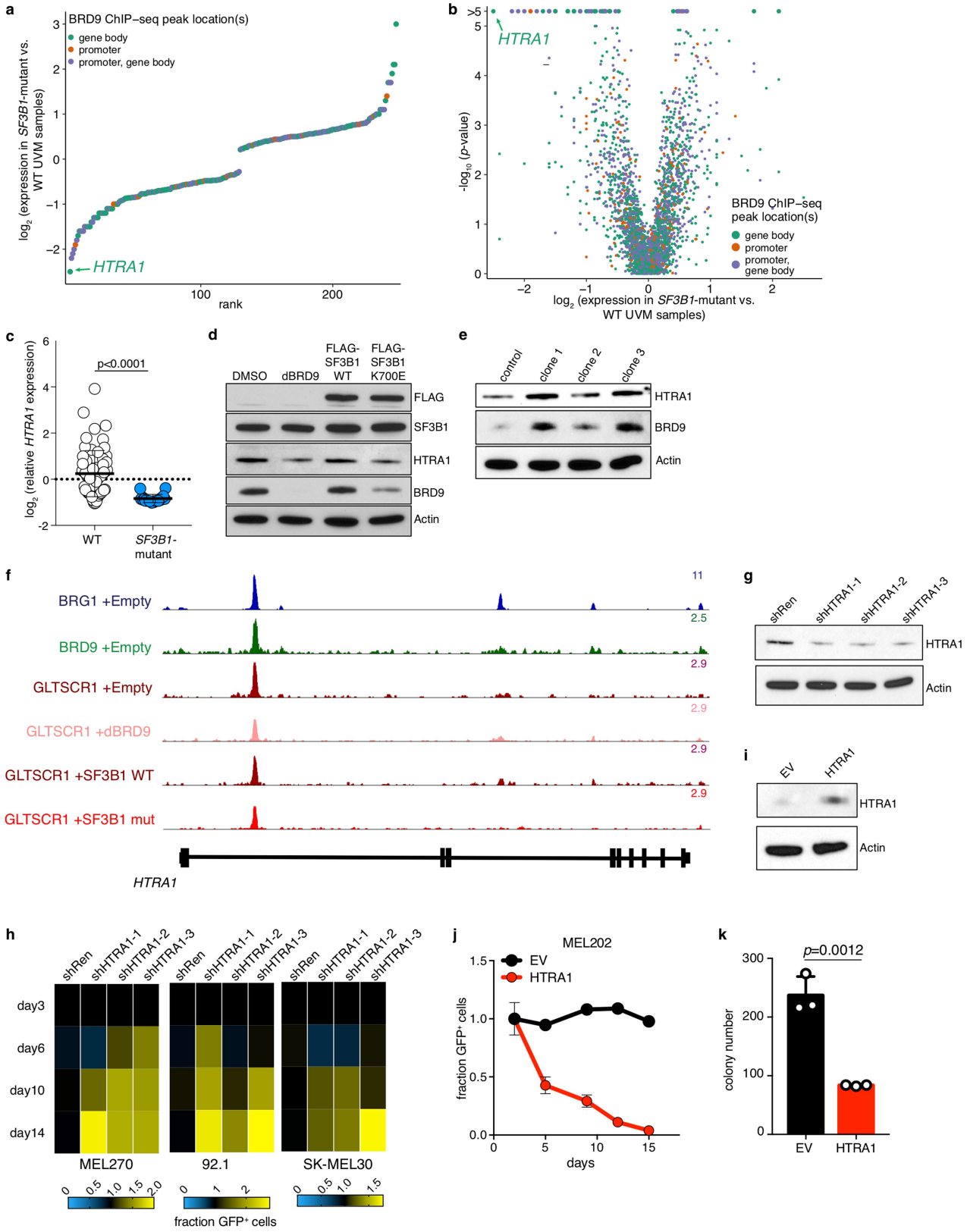
Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | BRD9 is a potent tumour suppressor in UVM. a, In vitro growth curves of Melan-a cells treated with two non-targeting shRNAs (against *Renilla* (shRen) and luciferase (shLuc)) versus parental, un-manipulated Melan-a cells. $n = 3$ biologically independent experiments per group. Data are presented as mean ± s.d. **b**, Tumour volume in SCID mice subcutaneously injected with Melan-a cells expressing a control shRNA (against *Renilla*), shRNA against *Brd9* (*Brd9* shRNA no. 1 and no. 2), shRNA against *Brg1* or cDNA encoding CYSLTR2(L129Q) ($n = 8$ mice per group). Data are presented as mean ± s.d. *P* value at day 64 was calculated compared to *Renilla* shRNA group with a two-sided *t*-test. **c**, Representative images of the dissected melanomas from **b**. **d**, H&E images of melanomas from **b**. Scale bars, 100 μm. Representative images from $n = 3$ biologically independent experiments. **e**, qRT–PCR measuring expression of *Brd9* (left) and melanoma-associated genes (*Mitf*, *Dct*, *Pmel* and *Tyrp1*) of melanomas from **a**. $n = 4$ (*Renilla* shRNA) and $n = 5$ (*Brd9* shRNA no. 1 and no. 2) biologically independent experiments. Data are presented as mean ± s.d. *P* value was calculated by two-sided *t*-test. **f**, Images of mice transplanted with parental, un-manipulated Melan-a cells or Melan-a cells transduced with a non-targeting shRNA or *Brd9*-targeting shRNA. Cells were subcutaneously engrafted into SCID mice and tumour volume was estimated 36 days after transplant. **g**, Volumes of tumours from **f** at day 36. Data are presented as mean ± s.d. $n = 10$ tumours per group. *P* value was calculated relative to the parental group by a two-sided *t*-test.



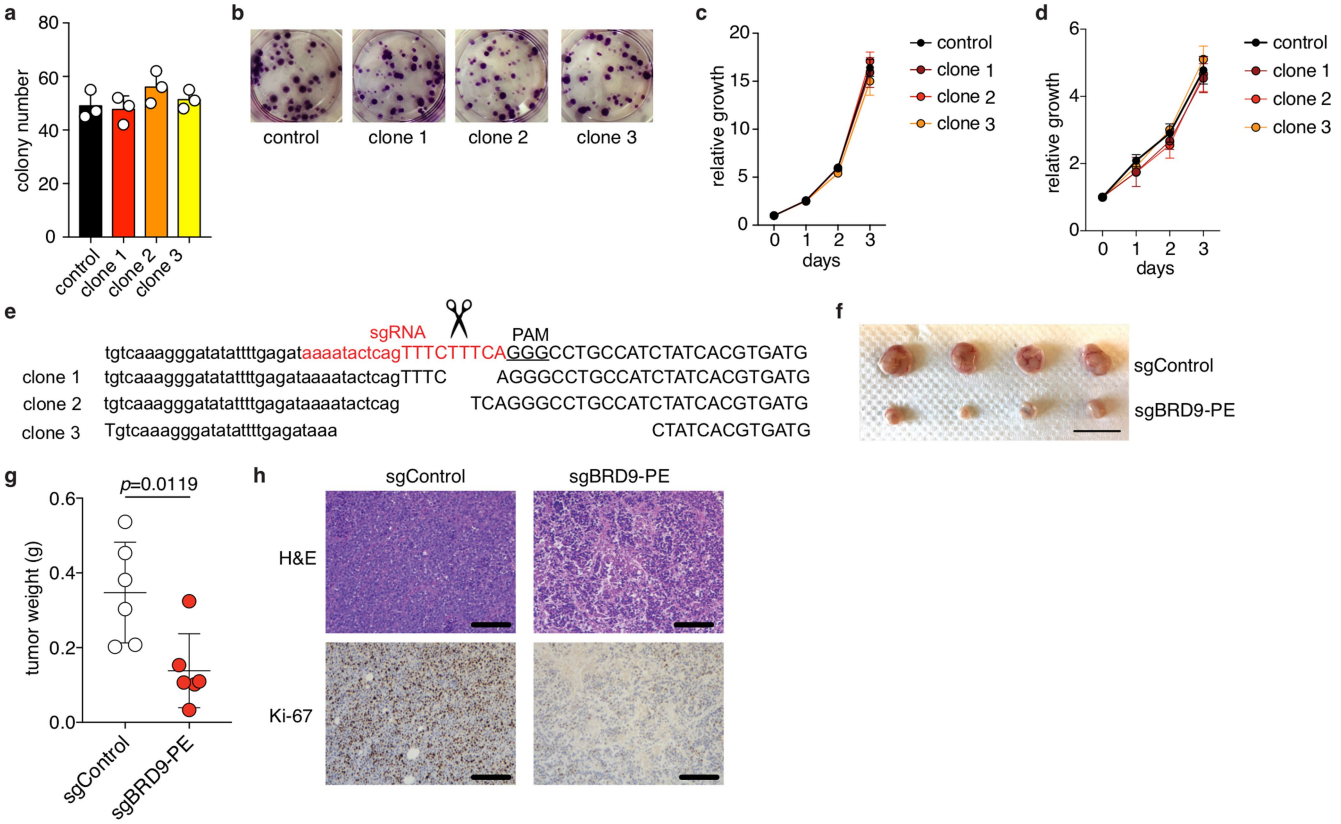
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | BRD9 is a potent tumour suppressor in UVM. **a**, Representative images of pulmonary metastatic foci produced 14 days after intravenous injection of B16 cells with or without *Brd9* shRNA (MLS-E vector). Scale bar, 5 mm. **b**, Western blot of endogenous BRD9 in B16 cells immediately before injection. Actin, loading control. The experiment was repeated three times with similar results. **c**, H&E sections of lung metastases. Arrows indicate metastatic foci. Scale bar, 100 μm. The experiment was repeated three times with similar results. **d**, Numbers of pulmonary B16 metastases identified in the experiments from **a**. $n = 6$ mice per group. *P* value was calculated relative to the *Renilla* shRNA group by a two-sided *t*-test. **e**, Relative percentages of GFP^+ 92.1 cells with or without *BRD9* shRNA (MLS-E vector), assessed by flow cytometric analysis of lung tissue in recipient NSG (NOD–SCID $Il2rg^{-/-}$) mice 14 days after intravenous injection by tail vein. The signal was normalized by dividing by the average percentage of GFP^+ cells in the *Renilla* shRNA (control) group. $n = 5$ biologically independent experiments per group. *P* value was calculated relative to the *Renilla* shRNA group by a two-sided *t*-test. **f**, Anti-GFP immunohistochemistry for sections of lung metastases from the experiment in **e**. Scale bar, 200 μm. The experiment was repeated three times with similar results. **g**, Representative images of tumours derived from transplantation of Melan-a cells transduced with doxycycline-inducible *Brd9* shRNA. Doxycycline was administered for nine weeks (left) and followed by doxycycline withdrawal for three weeks (right). **h**, Tumour volume for the experiment in **g**. $n = 4$ mice per group. The experiment was repeated twice with similar results. *P* value was calculated relative to the parental group by a two-sided *t*-test at day 7, day 14 and day 21. **i**, Representative images of recipient mice engrafted with MEL270 cells transduced with empty vector, full-length BRD9 (BRD9 WT) or a bromodomain-deletion mutant of BRD9 (BRD9 ΔBD) at day 12. $n = 5$ mice per group. **j**, Results of a competition assay to measure the effects of expression of the indicated cDNAs on growth of the indicated melanoma cells. Transduced cells were identified by co-expression of GFP (pMIGII vector). The percentage of GFP^+ cells was tracked over 21 days and normalized to the GFP percentage on day 2. Data are presented as mean ± s.d. $n = 2$ biologically independent experiments per group. **k**, Results of a competition assay to measure the effects of expression of the indicated cDNAs on growth of the indicated melanoma cells. Transduced cells were identified by the co-expression of GFP (pMIGII vector). The percentage of GFP^+ cells was tracked over 10 days and normalized to the GFP percentage on day 3. $n = 2$ biologically independent experiments per group.

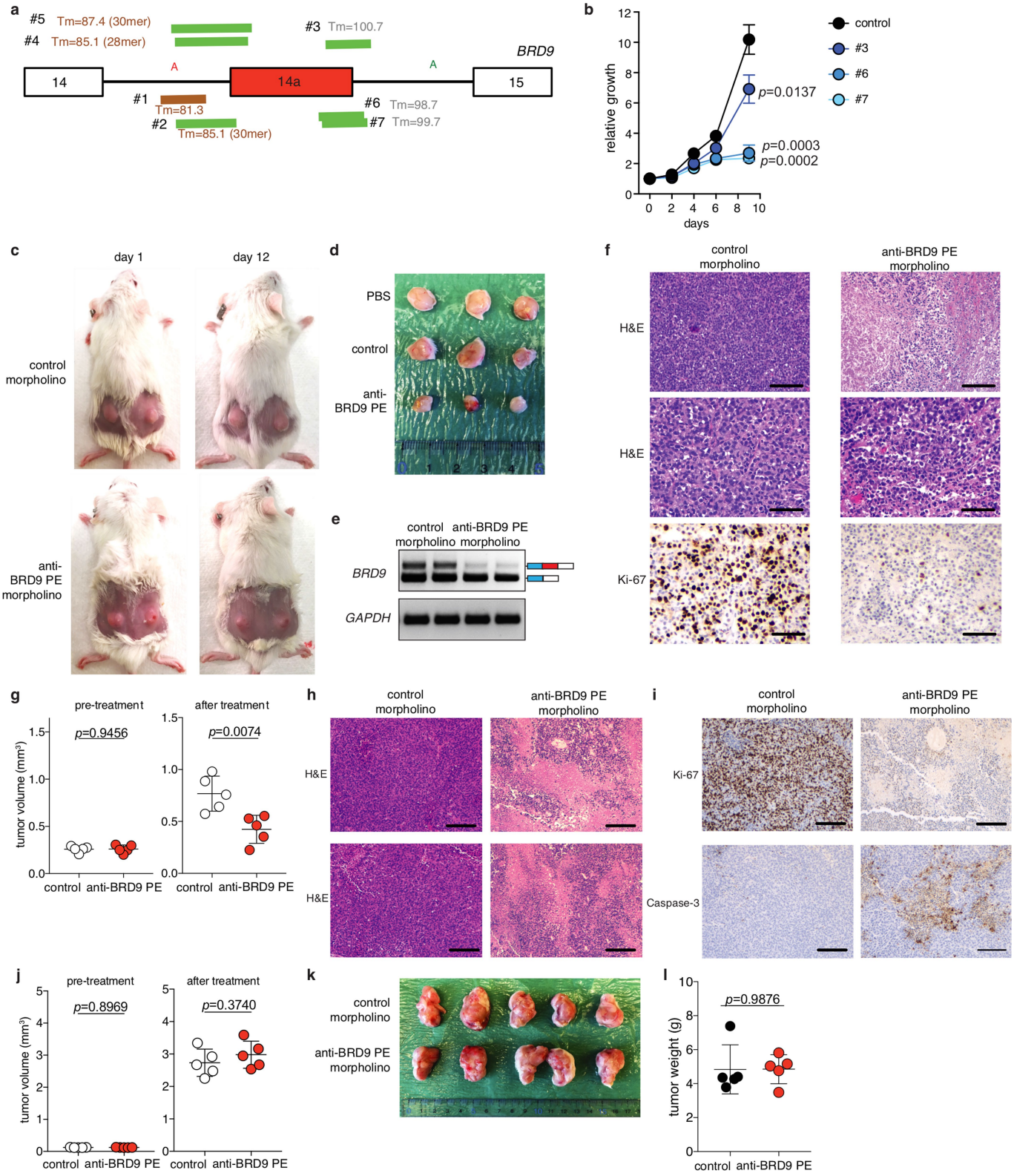


Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | BRD9 regulates *HTRA1* expression to promote UVM tumorigenesis. a, Rank plot illustrating fold change (expressed in $\log_2$) of each significantly differentially expressed gene identified by comparing samples from patients in the TCGA UVM cohort with mutated versus wild-type *SF3B1*. The plot is restricted to genes with BRD9 ChIP–seq peaks within their promoters or gene bodies in the absence of perturbations to BRD9 (MEL270 cells treated with DMSO or following ectopic expression of wild-type SF3B1). $n = 3{,}122$ genes analysed, of which $n = 248$ met the significance ($P < 0.001$) and expression (median expression in both wild-type and mutant samples > 2 transcripts per million) thresholds, and so are illustrated here. *P* value was computed using a two-sided Mann–Whitney *U*-test. **b**, As in **a**, but a volcano plot additionally illustrating the *P* value associated with the comparison between *SF3B1*-mutated and wild-type *SF3B1* samples. $n = 3{,}122$ genes analysed and illustrated. *P* value computed with two-sided Mann–Whitney *U*-test. **c**, *HTRA1* expression in samples from patients in the TCGA UVM cohort with ($n = 18$) or without ($n = 62$) *SF3B1* mutations. Expression is *z*-score normalized across all samples. Data are presented as mean ± s.d. *P* value computed with two-sided *t*-test. **d**, Western blot for Flag, SF3B1, HTRA1, BRD9 and actin in MEL270 cells (wild-type *SF3B1*), treated with DMSO, BRD9 degrader, Flag–SF3B1(WT) or Flag–SF3B1(K700E). Representative images from $n = 3$ biologically independent experiments. **e**, Western blot for HTRA1, BRD9 and actin in MEL202 cells ($SF3B1^{R625G}$) following CRISPR–Cas9-mediated mutagenesis of the *BRD9* poison exon (as shown in Extended Data Fig. 2o). Representative images from $n = 3$ technically independent experiments. **f**, Read coverage for BRG1, BRD9 and GLTSCR1 ChIP–seq at the *HTRA1* locus in MEL270 cells (shown in **d**) treated with an empty vector, BRD9 degrader, or $SF3B1^{WT}$ or $SF3B1^{K700E}$ cDNAs ($n = 1$ ChIP–seq experiment performed for each condition). **g**, Western blot for HTRA1 and actin in MEL270 cells treated with shRNAs against *HTRA1* or with a non-targeting control shRNA (against *Renilla*). Representative images from $n = 3$ biologically independent experiments. **h**, Heat map summarizing the results of a competition assay to measure the effect of each indicated shRNA on the growth of Cas9-expressing UVM cell lines with wild-type *SF3B1*. Cell growth was computed with respect to cells treated with a non-targeting control shRNA (against *Renilla*) and the percentage of GFP^+ cells at day 14 was normalized to that at day 2. The illustrated values correspond to the mean values computed over $n = 3$ biologically independent experiments. **i**, Western blot for HTRA1 and actin in MEL202 cells ($SF3B1^{R625G}$) following stable overexpression of an empty vector or HTRA1 (both with an MSCV-IRES-GFP vector). Representative images from $n = 3$ biologically independent experiments. **j**, Ratio of GFP^+ to GFP^- MEL202 cells ($SF3B1^{R625G}$) from a competition experiment in which GFP^+ cells from **i** were seeded at an initial 1:1 ratio with GFP^- control cells. Data are presented as mean of $n = 2$ biologically independent experiments. **k**, Colony number of MEL202 cells expressing empty vector or *HTRA1* cDNA from **i**, following 10 days of growth in soft agar. Data are presented as mean of $n = 3$ biologically independent experiments.

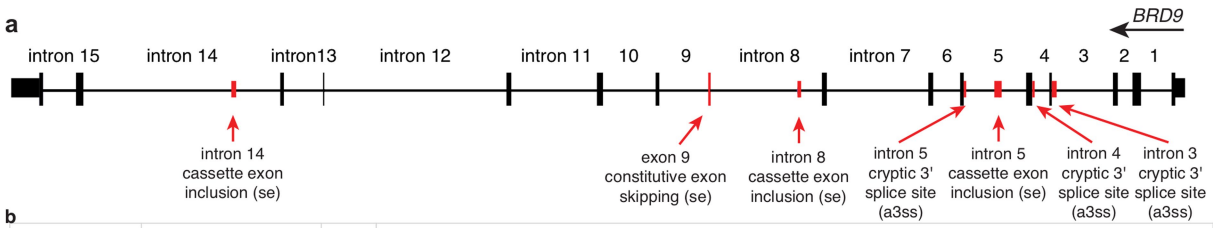


Extended Data Fig. 8 | CRISPR–Cas9-mediated mutagenesis of the *BRD9* poison exon corrects *BRD9* aberrant splicing and abrogates growth of *SF3B1*-mutated melanoma. a, Colony number for MEL270 cells (wild-type *SF3B1*) without (control) or with (clone 1, clone 2 and clone 3) CRISPR–Cas9-induced indels that disrupted *BRD9* poison exon recognition. Data presented as mean ± s.d. $n = 3$ biologically independent experiments per group. **b**, Representative images from **a**. **c**, Proliferation of the clones described in **a**. $n = 3$ biologically independent experiments per group. **d**, Proliferation of MEL285 cells (wild-type *SF3B1*) without (control) or with (clones 1, 2 and 3) CRISPR–Cas9-induced indels that disrupted *BRD9* poison exon recognition. $n = 3$ biologically independent experiments per group. **e**, Mutations generated at the 5′ end of the *BRD9* poison exon by CRISPR–Cas9-mediated indels in clones 1, 2 and 3 of MEL285 cells from **d**. The PAM sequence is illustrated with uppercase, underlined nucleotides. Red nucleotides hybridize to the sgRNA. **f**, Representative images of dissected tumours from recipient mice transplanted with CRISPR–Cas9-modified MEL202 clones. **g**, Tumour weight for the tumours illustrated in **f**. Data presented as mean ± s.d. $n = 6$ biologically independent experiments per group. *P* value was calculated relative to the control shRNA group by a two-sided *t*-test. **h**, H&E staining, as well as Ki-67 immunohistochemistry images, for the tumours illustrated in **f**. Representative images from $n = 3$ independent histological analyses.



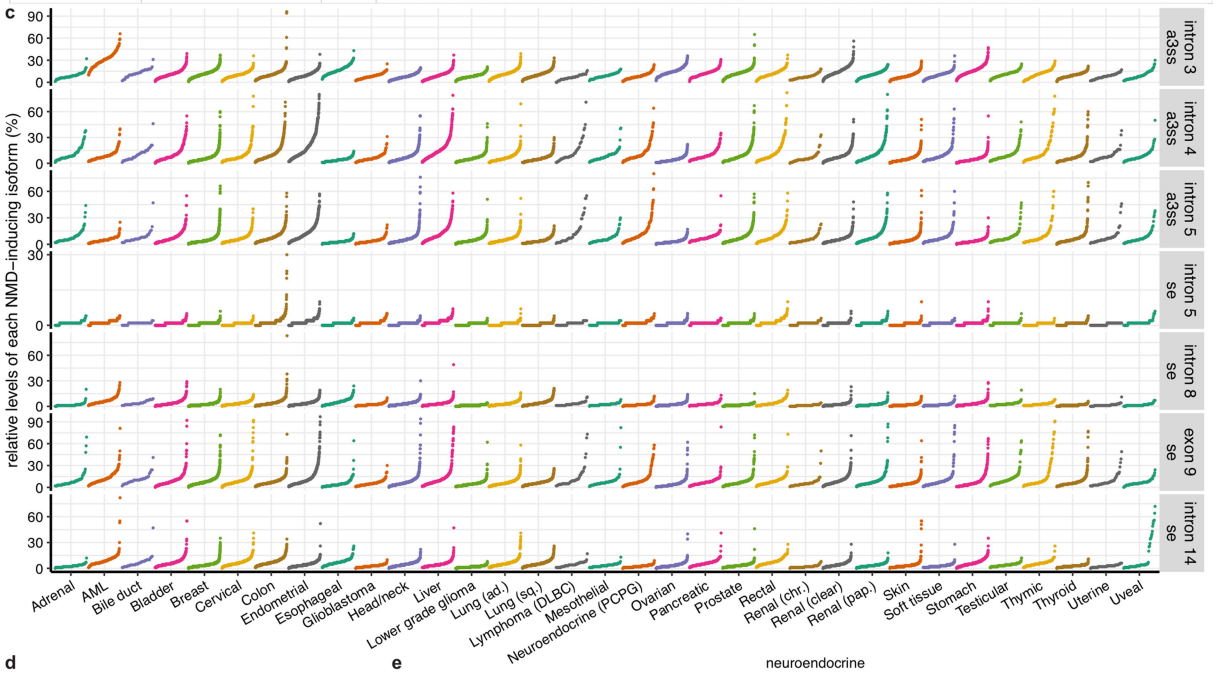
Extended Data Fig. 9 | See next page for caption.

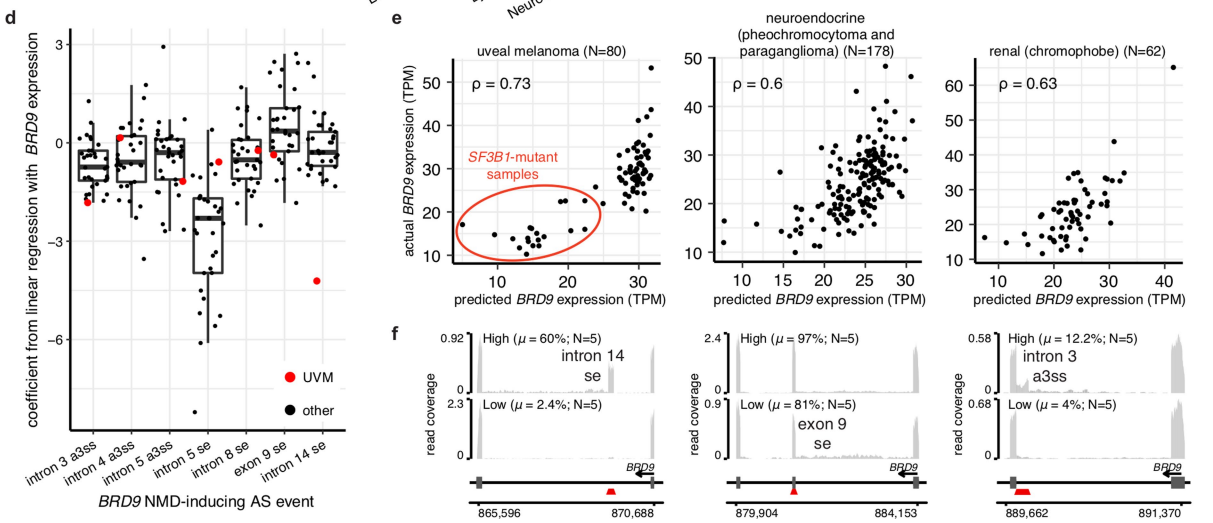
Extended Data Fig. 9 | Correcting *BRD9* mis-splicing in *SF3B1*-mutated xenografts with ASOs suppresses tumour growth. a, Cartoon representation of the *BRD9* loci targeted by each designed morpholino. Melting temperature (T_m) is shown. Lengths of target sequences are indicated in parentheses; if these are not indicated, then the length is 25 nt. **b**, Growth of MEL202 cells (*SF3B1*^{R625G}) treated with 10 μM of control non-targeting (control) or *BRD9* poison-exon-targeting morpholinos (no. 3, no. 6 and no. 7). $n = 3$ biologically independent experiments per group. *P* values at day 9 were calculated relative to the control group by a two-sided *t*-test. **c**, Representative images of recipient mice xenografted with MEL202 cells and treated with PBS or morpholinos in vivo. Each tumour was analysed after in vivo treatment with PBS, control morpholino or morpholino against the *BRD9* poison exon (no. 6) (12.5 mg kg^{-1}, every other day to a total of 8 intratumoral injections). $n = 10$ tumours per group. **d**, Representative images of dissected tumours from the experiment described in **c**. **e**, RT–PCR results of tumours from **c** to evaluate *BRD9* splicing. The experiment was repeated three times with similar results. **f**, Representative H&E and Ki-67 staining images of tumours from **c**. Scale bars, 100 μm (top), 50 μm (middle and bottom). The experiment was repeated three times with similar results. **g**, Estimated tumour volume for recipient mice transplanted with a patient-derived xenograft model of *SF3B1*^{R625C} rectal melanoma and treated with in vivo morpholinos (control or morpholino against *BRD9* poison exon, no. 6) (12.5 mg kg^{-1}, every other day to a total of 8 intratumoral injections). $n = 5$ mice per group. Estimated tumour volumes before and after treatment are shown. Data are presented as mean ± s.d. *P* values were calculated relative to the control group by a two-sided *t*-test. **h**, Representative H&E staining images of tumours from **g**. The experiment was repeated three times with similar results. **i**, Representative Ki-67 staining images of tumours from **g**. The experiment was repeated three times with similar results. **j**, Estimated tumour volume for recipient mice transplanted with a patient-derived xenograft model of UVM (wild-type *SF3B1*), treated with in vivo morpholinos (control or morpholino against *BRD9* poison exon, no. 6) (12.5 mg kg^{-1}, every other day to a total of 8 intratumoural injections). $n = 5$ mice per group. Estimated tumour volumes before and after treatment are shown. Data are presented as mean ± s.d. *P* values weres calculated relative to the control group by a two-sided *t*-test. **k**, Representative images of dissected tumours from **j**. **l**, Tumour weight for tumours from **k**. $n = 5$ mice per group. *P* value was calculated relative to the control group by a two-sided *t*-test.



b

intron 3 a3ss	chr5:889763-889881	prox.	TTCTCGGTGGAACTTAAAATGCTGTGAGACACCAGACAGACAGATACTGTGAACTTGGAGCTCTCTAATGAAGGGATACCAAAGTCTTGTATTCAATTTTTTTTTCCTTAAATTGTCAG
intron 4 a3ss	chr5:889276-889280	dist.	AAAAG
intron 5 se	chr5:888350-888522	inc.	GGTTTGCTGACTCACAGAAACACAACAGAACAGACTTAAAGGAGAGTGGGAGAAAAGGAAACTCGATGAGAGGATCAGGCAGCGCCTGAGGGTAGATGTTAAGAGCACTCCATGCCCGAAAGCCCTCTGTATTCCCTGCTTCGGGCCCCAGTCTGGCCATTCCTAGCAACAAA
intron 5 a3ss	chr5:887583-887586	dist.	GCAG
intron 8 se	chr5:883466-883565	inc.	GAGACGTTCTAGACCCAGTGTCTTCAGTGGAAACAGAGCACAGGCTGTCTGCAGTCTCAGGCTTCCTAGCAGCCCTGATAAGAGAAGTAGAGCAATGCAG
exon 9 se	chr5:881222-881297	exc.	ATGGGCTATCTGAAGAGGAACGGGGACGGGAGCCTGCTCTACAGCGTGGTCAACACGGCCGAGCCGGACGCTGATG
intron 14 se	chr5:869360-869519	inc.	TTTCTTTCAGGGCCTGCCATCTATCACGTGATGCTGTACTAAAGTCGGATCGGAATTTGGTGTCTTATTACCACGGAGTCTGCTTTGTCAACCTTCAGATTTCTGTTTTAAAGTTAATGCTGGTCAGCTGTGCCTGAATTCCAAAAGAGGGGATAATGAG





Extended Data Fig. 10 | See next page for caption.

Extended Data Fig. 10 | Use of multiple, distinct NMD isoforms of *BRD9*. **a**, *BRD9* gene structure illustrating constitutive *BRD9* exons and alternative splicing events that are predicted to induce NMD. *SF3B1* mutations promote inclusion of the BRD9 poison exon in intron 14. **b**, Genomic coordinates (hg19/GRCh37 assembly) of each NMD-inducing event illustrated in **a**, as well as genomic sequence of each alternatively spliced region highlighted in red in **a**. The third column indicates the specific isoform that is a predicted NMD substrate. Prox, intron-proximal competing 3′ splice site; dist, intron-distal competing 3′ splice site; inc, exon inclusion; exc, exon exclusion. **c**, Rank plot illustrating levels of each NMD-inducing isoform relative to total *BRD9* mRNA levels for each sample in each indicated TCGA cohort. Boxes illustrate first and third quartiles, with whiskers extending to 1.5× interquartile range. **d**, Box plot illustrating the distribution of coefficients estimated by fitting a linear model to predict *BRD9* gene expression on the basis of relative levels of each NMD-inducing isoform. The relative levels of NMD-inducing isoforms illustrated in **c**, as well as *BRD9* gene expression estimates for each sample, were used to construct an independent linear model with robust regression for each TCGA cohort. The coefficients resulting from this model fitting procedure are illustrated in the box plot, in which each dot corresponds to the coefficient associated with the corresponding NMD-inducing event for a single TCGA cohort. Coefficients for the TCGA UVM cohort are highlighted in red. The coefficients are typically negative (as expected for NMD-inducing isoforms), with the exception of constitutive exon 9 skipping, for which the coefficients are generally positive—as expected for an event in which NMD is induced when a constitutive exon is excluded. The *SF3B1*-mutation-responsive poison exon in intron 14 dominates the fit for UVM, as expected. $n = 33$ TCGA cohorts analysed and illustrated. **e**, Scatter plots comparing actual (y axis) and predicted (x axis) *BRD9* expression levels for three TCGA cohorts. Each dot corresponds to a single sample. ρ, Spearman's correlation between actual and predicted values. **f**, RNA-seq read coverage plots for patient samples from the TCGA cohorts illustrated in **e** for representative alternative splicing events illustrated in **a**. Each coverage plot illustrates data averaged over the $n = 5$ patient samples from the vertically matched cohort in **e** that exhibit the lowest or highest relative expression of the NMD-inducing isoform. μ, mean relative expression of the illustrated NMD-inducing isoform, computed over each group of samples.

nature research

Corresponding author(s): Omar Abdel-Wahab
Robert K. Bradley

Last updated by author(s): Jul 6, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☐	☒	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection.
Data analysis	Publicly available software was used in this study. Specific programs are RSEM (v1.2.4), Bowtie (v1.0.0), TopHat (v2.0.8b), MISO (v2.0), and Bioconductor (v3.7) within the R (v3.5.1) programming environment.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

RNA-seq and ChIP-seq data generated as part of this study were deposited in the Gene Expression Omnibus (accession number GSE124720). RNA-seq data from published studies were downloaded from CGHub (TCGA UVM57), EMBL-EBI ArrayExpress (Illumina Human BodyMap 2.0: E-MTAB-513), the Gene Expression Omnibus (accession numbers GSE72790, GSE114922) (CLL: Darman et al. 201515, MDS: Pellagati et al. 201858), or directly obtained from the authors (UVM: Alsafadi et al. 201610). Gel source data can be found in Supplementary Fig. 1. Other data that support this study's findings are available from the authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample sizes for xenograft experiments were chosen based on published studies of known oncogenic drivers of relevant models (e.g., expression of the oncoprotein CYSLTR2 L129Q in Melan-a cells).
Data exclusions	No data were excluded.
Replication	Attempts at replication were successful.
Randomization	Animals were randomly assigned to experimental groups.
Blinding	The data presented did not require the use of blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☐	☒ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	For Western blotting, the following antibodies to the following proteins were used: BRD9 (Bethyl Laboratories; A303-781A and Active Motif; 61538), SF3B1/Sap-155 (MBL; D221-3), Flag-M2 (Sigma-Aldrich; F-1084), β-actin (Sigma-Aldrich; A-5441), GLTSCR1 (Santa Cruz Biotechnology; sc-515086), GLTSCR1L (Thermo Fisher Scientific; PA5-56126), BRM (Bethyl Laboratories; A303-015A), BRG1 (Santa Cruz Biotechnology; sc-17796), BAF155 (Santa Cruz Biotechnology; sc-48350), BAF60A (Santa Cruz Biotechnology; sc-135843), BAF47 (Santa Cruz Biotechnology; sc-166165), ARID1A (Santa Cruz Biotechnology; sc-373784), ARID2 (Santa Cruz Biotechnology; sc-166117), BRD7 (Thermo Fisher Scientific; PA5-49379), U2AF2 (Bethyl Laboratories; A303-665A), U2AF1 (Bethyl Laboratories; A302-080A), Histone H3 (Abcam; ab1791), HTRA1 (R&D systems; MAB2916-SP). For ChIP-seq studies in MEL270 cells, antibodies to endogenous BRG1 (Abcam EPNCIR111A, Lot # GR3208604-8), GLTSCR1 (Santa Cruz SC-240516, Lot # A2313), and BRD9 (Abcam, ab137245) were used.
Validation	All antibodies were validated by the supplier for human samples, and were checked in the lab by Western blotting on cell lysate and by comparing to the manufacturer's or in-house results.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Ba/F3, K562, NALM6, TF1, HEK293T, MCF10A, Miapaca2, CFPAC1, B16, and HEK293T cells were obtained from the American Type Culture Collection (ATCC) . The isogenic K562 and NALM6 cell lines with or without SF3B1 mutations were generated at Horizon Discovery. Melan-a cells were provided by Dr. Dorothy Bennett (PMID: 3102392); MEL202, MEL270, UPMD1, UPMD2, and 92-1 were obtained from Dr. Boris Bastian (PMID: 22236444). KPC cells were obtained from Dr. Ben Stanger (PMID: 21436454). SK-MEL30 cells and RN2 cells were provided by Dr. David B. Solit (PMID: 21725359) and Dr. Scott W. Lowe (PMID: 21131983), respectively.

Authentication	An aliquot of each cell lines were authenticated using ATCC fingerprinting. Otherwise, the cells were submitted for short tandem repeat (STR) profiling and MSK-IMPACT (integration mutation profiling of actionable cancer targets) for mutational status at MSKCC to confirm their authenticity.
Mycoplasma contamination	All cell lines are frequently tested for mycoplasma contamination. Cell lines used in this study were verified to be mycoplasma negative before undertaking any experiments with them.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	6-8 week SCID and NGS mice were purchased from The Jackson Laboratory (stock #001303 and stock #005557, respectively).
Wild animals	The study did not involve wild animals.
Field-collected samples	The study did not involve samples collected from the field.
Ethics oversight	All animals were housed at Memorial Sloan Kettering Cancer Center (MSKCC). All animal procedures were completed in accordance with the Guidelines for the Care and Use of Laboratory Animals and were approved by the Institutional Animal Care and Use Committees at MSKCC. All mouse experiments were performed in accordance with a protocol approved by the MSKCC Institutional Animal Care and Use Committee (11-12-029).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Human research participants

Policy information about studies involving human research participants

Population characteristics	Patients with chronic lymphocytic leukemia (CLL) and uveal and mucosal melanoma seen at Memorial Sloan Kettering Cancer Center who provided anonymized peripheral blood mononuclear cells (for CLL) and tissue biopsies (for melanoma).
Recruitment	Patients with CLL and uveal and mucosal melanoma seen at Memorial Sloan Kettering Cancer Center (MSK) who consented to MSK IRB protocol 06-107 (for CLL) and IRB 14-191 (for melanoma samples) were eligible for inclusion regardless of race, gender, ethnicity or other characteristics.
Ethics oversight	Studies were approved by the Institutional Review Boards of Memorial Sloan Kettering Cancer Center (MSK), informed consent was obtained from all subjects (under MSK IRB protocol 06-107) and conducted in accordance to the Declaration of Helsinki protocol. Patients provided samples after their informed consent and primary human de-identified CLL samples derived from whole peripheral blood or BM mononuclear cells were utilized. PDX models were performed using tumor biopsies from de-identified patients under MSK IRB protocol 14-191. Genomic alterations in melanoma tumor biopsies and CLL cells were analyzed using MSK IMPACT28 assay or FoundationOne Heme30 assay, both as previously described. Patient samples were anonymized by the Hematologic Oncology Tissue Bank of MSK (for CLL samples) and the MSK Antitumor Assessment Core Facility (for PDX samples).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

ChIP-seq

Data deposition

☒ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☒ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	The data (including related RNA-seq data) has been deposited in the Gene Expression Omnibus under accession number GSE124720 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE124720)
Files in database submission	The above accession includes related RNA-seq data. The relevant ChIP-seq files (.fastq, .bigwig and associated peak calls) can be obtain at the link above: GSM3544057 MEL270.dBRD9.BRG1 GSM3544058 MEL270.dBRD9.GLTSCR1 GSM3544059 MEL270.dBRD9.input GSM3544061 MEL270.DMSO.BRD9 GSM3544062 MEL270.DMSO.BRG1 GSM3544063 MEL270.DMSO.GLTSCR1 GSM3544064 MEL270.DMSO.input GSM3544066 MEL270.SF3B1_K700E.BRG1 GSM3544067 MEL270.SF3B1_K700E.GLTSCR1 GSM3544068 MEL270.SF3B1_K700E.input

GSM3544070 MEL270.SF3B1_WT.BRD9
GSM3544071 MEL270.SF3B1_WT.BRG1
GSM3544072 MEL270.SF3B1_WT.GLTSCR1
GSM3544073 MEL270.SF3B1_WT.input

Genome browser session
(e.g. UCSC)

https://genome.ucsc.edu/s/chewguoliang/SF3B1%20BRD9%20MEL270%20ChIP

Methodology

Replicates

The ncBAF complex was pulled-down with two distinct antibodies against BRD9 and GLTSCR1 in two separate control and ncBAF perturbation experiments that used orthogonal methods to perturb ncBAF formation. Analyses used the intersection of data from these distinct experiments.

Sequencing depth

Sequencing depth, as well as mapping rates are described below:

Sample Reads_total Reads_mapped Reads_unmapped Reads_multimapped
MEL270.dBRD9.GLTSCR1 21663452 17327309 (79.98%) 3317111 (15.31%) 1019032 (4.70%)
MEL270.dBRD9.input 18584217 16915436 (91.02%) 761948 (4.10%) 906833 (4.88%)
MEL270.dBRD9.BRG1 20392804 18016161 (88.35%) 1654723 (8.11%) 721920 (3.54%)
MEL270.DMSO.BRD9 19785290 15734102 (79.52%) 2934586 (14.83%) 1116602 (5.64%)
MEL270.DMSO.GLTSCR1 19281052 15911493 (82.52%) 2501078 (12.97%) 868481 (4.50%)
MEL270.DMSO.input 19743934 17876186 (90.54%) 899111 (4.55%) 968637 (4.91%)
MEL270.DMSO.BRG1 21373496 18958611 (88.70%) 1645389 (7.70%) 769496 (3.60%)
MEL270.SF3B1_K700E.GLTSCR1 20098772 15780486 (78.51%) 3347543 (16.66%) 970743 (4.83%)
MEL270.SF3B1_K700E.input 18866831 17097903 (90.62%) 832672 (4.41%) 936256 (4.96%)
MEL270.SF3B1_K700E.BRG1 19879373 17670132 (88.89%) 1491006 (7.50%) 718235 (3.61%)
MEL270.SF3B1_WT.BRD9 15700266 12196453 (77.68%) 2546295 (16.22%) 957518 (6.10%)
MEL270.SF3B1_WT.GLTSCR1 19755923 16270285 (82.36%) 2565754 (12.99%) 919884 (4.66%)
MEL270.SF3B1_WT.input 19401125 17523685 (90.32%) 895478 (4.62%) 981962 (5.06%)
MEL270.SF3B1_WT.BRG1 19408244 17112467 (88.17%) 1589600 (8.19%) 706177 (3.64%)

Antibodies

BRG1 (Abcam EPNCIR111A, Lot # GR3208604-8)
GLTSCR1 (Santa Cruz SC-240516, Lot # A2313)
BRD9 (Abcam, ab137245)

Peak calling parameters

Narrow peaks were called using the callpeak function from MACS2 v2.1.1.20160309 against matched input samples, using default parameters and a p-value cutoff of 1e-5.

Data quality

For all samples, a p-value cutoff of 1e-5 against input was used. All peaks were called at a q-value of < 0.017. For each sample, the number of peaks with a fold-change > 5, and the total number of peaks called is as follows:
MEL270.dBRD9.BRG1 68229 94156
MEL270.dBRD9.GLTSCR1 2982 4504
MEL270.DMSO.BRD9 13390 20591
MEL270.DMSO.BRG1 73624 101272
MEL270.DMSO.GLTSCR1 4453 6564
MEL270.SF3B1_K700E.BRG1 60362 85463
MEL270.SF3B1_K700E.GLTSCR1 6939 11118
MEL270.SF3B1_WT.BRD9 17434 25208
MEL270.SF3B1_WT.BRG1 64209 89486
MEL270.SF3B1_WT.GLTSCR1 5794 8678

Software

ChIP-seq reads were mapped to the genome by calling Bowtie v1.0.048 with the arguments '-v 2 -k 1 -m 1 --best --strata'. Peaks were called using MACS2 v2.1.1.2016030952 against input control libraries with p < 10-5 and subsequently filtered to remove peaks contained within ENCODE blacklisted regions and the mitochondrial genome. Subsequent data analysis was performed with Bioconductor in the R programming environment. Consensus peaks between samples were called using the soGGI package v1.14.0. Peaks were annotated using the ChIPseeker package v1.18.0.