Splicing Deregulation and Splicing Modulation in Acute Myeloid Leukemia

van der Werf, Inge Margaretha

2022

document version
Publisher's PDF, also known as Version of record

Splicing Deregulation and Splicing Modulation in Acute Myeloid Leukemia: DNA is not Destiny

I.M. van der Werf

Splicing Deregulation and Splicing Modulation in Acute Myeloid Leukemia: DNA is not Destiny

Inge van der Werf

Cover and layout: Willem Thomassen and Inge van der Werf
Printing: Proefschriftmaken.nl

Financial support by the OOA Diamond Program, for the work within this thesis is gratefully acknowledged.

The printing of this thesis was additionally kindly supported by: Servier Nederland Farma and H3 Biomedicine, Inc

VRIJE UNIVERSITEIT

Splicing Deregulation and Splicing Modulation in Acute Myeloid Leukemia: DNA is not Destiny

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr C.M. van Praag,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op dinsdag 1 februari 2022 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door
Inge Margaretha van der Werf
geboren te Voorschoten

promotoren:	prof.dr. G.J.L. Kaspers
	prof.dr. J. Cloos
copromotoren:	dr. A.M. Wojtuszkiewicz
	dr. R. Groen
promotiecomissie:	prof.dr. S. Zweegman
	prof.dr. C.M. Zwaan
	prof.dr. A.W. Griffioen
	prof.dr. O.T. Heidenreich
	dr. M. Spaargaren

TABLE OF CONTENTS

CHAPTER 1

Introduction

ACUTE MYELOID LEUKEMIA

Acute myeloid leukemia, a cancer of the bone marrow, is a disease characterized by a differentiation block of myeloid precursors resulting in an abnormal accumulation of leukemic blast cells which expand at the expense of normal bone marrow cells.[1] AML predominantly affects adults over the age of 60. However, despite it is rare, AML also develops in younger adults and children. It is the most common type of acute leukemia in adults, and though less frequent, it is the acute leukemia in children with the worst outcome.[2,3]

To date, AML treatment includes two cycles of induction chemotherapy with typically cytarabine, idarubicin, daunorubicin or mitoxantrone.[4] Ultimately, the goal of therapy is to eradicate all leukemic cells. Therefore, after achievement of complete remission upon induction treatment, consolidation therapy is administered to maintain a durable response.[5] Consolidation therapy can consist either of further cycles of chemotherapy or stem cell transplantation. While most AML patients achieve complete remission, the five-year event-free survival rates reach only 35-40% in adults and 70-75% in children.[2,3,6] Treatment failure is hypothesized to occur in many cases due to incomplete eradication of a subpopulation of leukemia stem-like cells that are thought to be responsible for the outgrowth of a clonal cell population.[7] As a result, many studies have focused on the characterization of leukemic cells and therapy resistant stem cells specifically.[8-12] These efforts have resulted in improvements of diagnostics, revisions in risk stratification and the development of targeted therapies for AML patients.[13,14] Yet, as a rare hematologic malignancy, pediatric AML has been less extensively studied. Nonetheless, initial studies have hinted at substantial differences between the genomic landscapes of adult and pediatric AML.[3,15]

Recent studies have revealed that in adult AML, mutations are much more common in comparison to structural alterations (i.e. chromosomal abnormalities) while the opposite has been found in pediatric patients.[3,15] In addition, DNA mutations that are frequently present in pediatric AML have found to be occurring at a different frequency in adult AML. For example, a higher frequency of *NRAS* and *KRAS* mutations has been identified in pediatric AML compared to adult AML.[3] In contrast, mutations in *TP53*, *RUNX1*, *DNMT3A* and splicing factor genes are frequent in patients with adult AML while these mutations have found to be nearly absent in pediatric AML patients (Figure 1).[3]

In addition, cellular processes can be disturbed via different mechanisms and do not always involve genetic mutations. For instance, many studies have reported deregulation of histone modifications, DNA or RNA methylation, RNA editing or the process of splicing in cancer in the absence of genomic mutations in splicing factors.[16-27] The process of RNA splicing is of great importance in gene regulation and alterations in this pathway have been implicated in many human cancers, including AML.[28,29] Actually, the highest frequency of differential splicing have been displayed in AML as compared to matched normal tissue, while liver cancers were characterized by the lowest rate of alternative splicing.[30] Accordingly, it is of great interest to study the impact of splicing deregulation in AML which is the focus of this thesis.

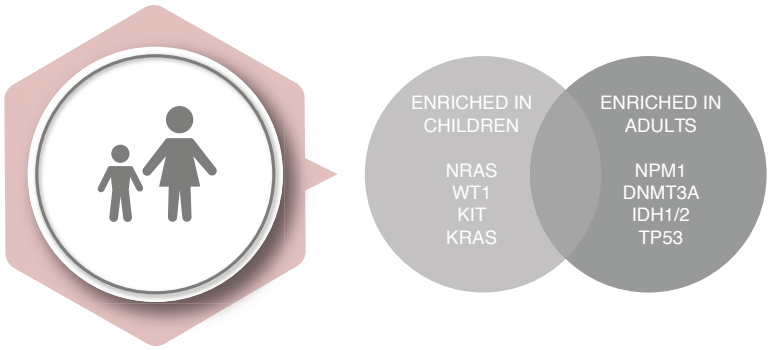


Figure 1. Adult AML and Pediatric AML are distinct on the genomic level. Notable differences between adult and pediatric cases include distinct frequencies of specific genetic mutations.

RNA SPLICING

In 1977, Richard J. Roberts and Phillip A. Sharp were awarded the Nobel prize for discovering that our DNA is not co-linear with our RNA as the coding sequences of our genes were found to be interrupted by non-coding regions.[31,32] These non-coding segments, later termed introns, are removed from the precursor messenger RNA (pre-mRNA) transcript, to obtain the mature mRNA, which contains the coding sequences, known as exons (Figure 2). To date, the process of splicing has been described as intron removal and exon joining to form a sequence that serves as template for protein translation.[33] Frequently, splicing takes place using alternative splice sites. Exon skipping occurs when an exon is spliced out together with its flanking introns and this constitutes the most prevalent type of event, responsible for approximately 40% of all alternative splicing events.[34,35] Alternative 3' or 5' splice site selection designate recognition of an alternative splice site as compared to the canonical site within an exon, while intron retention is called as an entire intron has failed to be spliced out. As multiple regions of a single gene can undergo alternative splicing, this process results in the production of numerous distinct mRNAs from a single DNA sequence, each encoding for different protein isoforms with unique structural and functional properties. In extreme cases, hundreds of isoforms can arise from the same gene. In addition, nearly all genes present in our genome are subject to alternative splicing. Consequently, alternative splicing greatly expands the coding capacity of the genome, or in other words, genes are modular which adds to the complexity of life.[36,37]

The process of splicing is a tightly controlled phenomenon regulated by a dynamic complex, the spliceosome, in concert with multiple auxiliary factors.[37,38] This includes numerous (ribonucleo) proteins that interact with each other and recognize specific nucleotide sequences embedded in pre-mRNA. In short, splicing factors (i.e., SR proteins) drive splice site selection and recruit the spliceosome machinery which subsequently coordinates and executes the process of splicing. This process is affected by many factors influencing its outcome including a characteristic class of splicing regulators which repress exon inclusion (i.e., hnRNPs).[37] However, more recent studies show that both SR proteins and hnRNPs work in a context dependent manner, which adds another layer of complexity that we are only beginning to understand.[39-41]

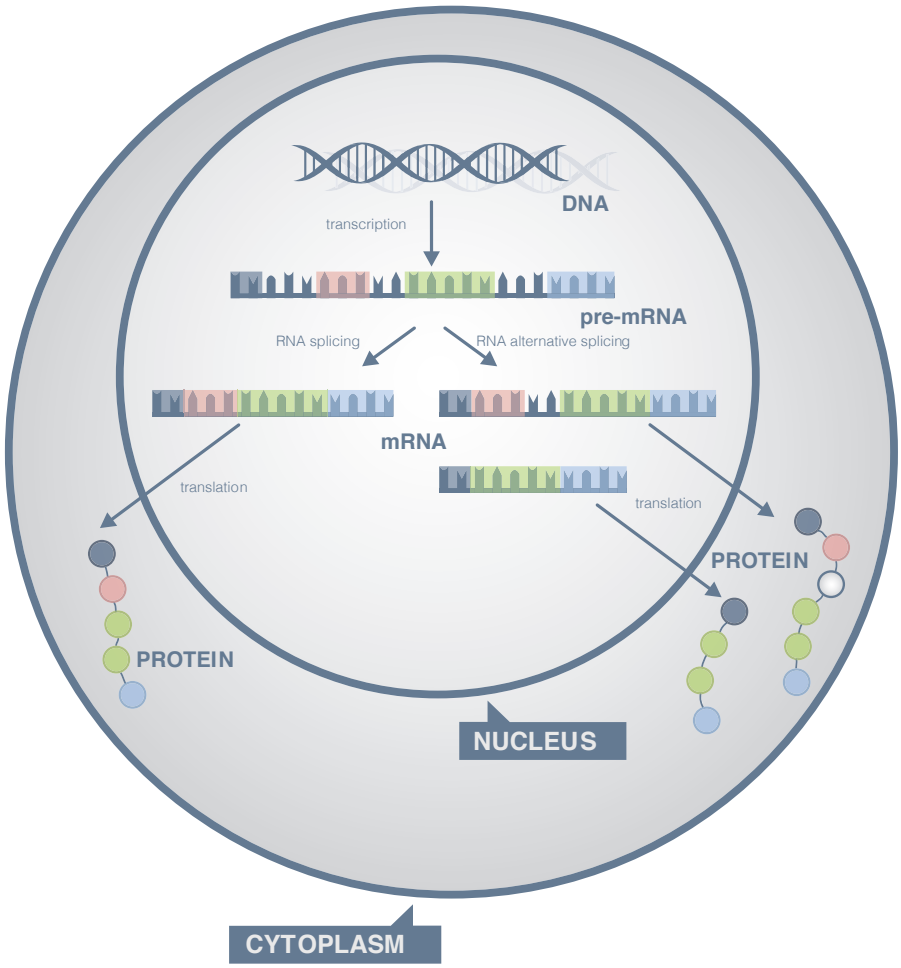


Figure 2. The mechanism of alternative splicing. Alternative splicing is the process by which splice sites in precursor messenger RNAs (pre-mRNAs) are differentially selected to produce multiple mature mRNAs and protein isoforms with distinct structural and functional properties.

RNA SPLICING IN STEM CELLS

To date, splicing is the most extensively studied mRNA modification that influences protein composition, which is essential for many cellular processes including proliferation, apoptosis, and differentiation.[33] Upon differentiation hundreds of cell types arise, most of which contain identical DNA copies of the nearly 21.000 genes that comprise the human genome. Thus, cellular identity is not driven by differences in genomic content but rather by epigenomic, transcriptomic, and proteomic heterogeneity.[42] Accordingly, studies revealed that human embryonic stem cells (hESCs) present distinct splicing profiles compared to differentiated cells. Actually, splice isoform diversity is highest in hESCs and decreases upon differentiation.[43] This phenomenon, referred to as isoform specialization, is mediated in part by a number of splicing factors and other RNA binding proteins that are differentially expressed during development.

Hence, alternative splicing can influence gene expression and either promote or impair stem cell function. For example, depletion of splicing factor SRSF2 was shown to decrease the expression of pluripotency factors and disrupt self-renewal of hESCs.[44] In contrast, muscle blind like splicing factors (MBNLs) were shown to negatively regulate stem cell self-renewal.[45,46] In addition, alternative splicing can also directly affect the function of genes important for pluripotency. This way, *OCT4A*, a splice variant of *OCT4*, is specifically expressed in pluripotent stem cells and is necessary for their self-renewal, whereas *OCT4B*, another splice variant of *OCT4*, is expressed by both stem cells and somatic cells and has no apparent role in regulating stem cell function.[47] As AML is often described as a disease initiated by stem cells, it might not be surprising that the process of alternative splicing has been implicated in its disease pathogenesis.

SPLICING DEREGULATION IN AML

In hematological malignancies, genome wide sequencing studies found the genes involved in the process of splicing to be surprisingly often disrupted. Mutations in genes encoding splicing regulators were first reported in hematological malignancies.[48-50] In 2016, these studies resulted in an updated version of the World Health Organization (WHO) classification for myeloid neoplasms and acute leukemia and incorporation of *SF3B1* mutations as a diagnostic criterion for MDS with ring sideroblasts.[13,14,51] More recently, splicing regulators, such as *SRSF2*, *U2AF1*, *ZRSF2* and *SF3B1*, were recurrently found to be mutated in AML.[12,52-54]

All splicing factors found to be recurrently mutated, act during the early stage of spliceosome assembly. Though, recent work has demonstrated that these mutations alter the function of the spliceosome via distinct mechanisms. For instance, *SF3B1* mutations cause changes in 3' splice site selection, while mutations in *SRSF2* alter the RNA-binding preferences of the spliceosome in a sequence-dependent manner resulting in changes in the efficiency of exon inclusion.[48,55] Regardless of the specific mechanism, mutations in these splicing factors result in widespread changes in the transcriptome which have been associated with a leukemic phenotype. While the role of splicing deregulation has not been clearly elucidated in pediatric AML, and recurrent mutations characteristic of adult AML are not detected in a pediatric setting, the splicing regulator MBNL1 was recently found to be deleted in a small proportion of pediatric AML patients, suggesting a functional role for splicing deregulation in pediatric leukemogenesis.[15]

To date, much attention has been dedicated to characterization of alternative splicing in AML cells carrying mutations in splicing factors. Yet, disruption of splicing appears to be a global phenomenon in hematological malignancies regardless of the presence of these aberrations.[17] Recent results suggest that hematopoietic stem and progenitor AML cells in general harbor unique mRNA splicing profiles characterized by intron retention and exon skipping.[56] Thus, the process of splicing is deregulated in both AML cells with and without mutations in splicing factors. As a consequence, both AML with and without mutations in splicing factors are of interest when studying splicing dysregulation.

SPLICING MODULATION

Given that aberrant splicing is a characteristic feature of AML cells, and many other cancer types, it is not surprising that significant advances have been made in the development of approaches to manipulate splicing for therapeutic purposes.[38,57,58] Most prominent splicing modulators target SF3B1 resulting in the formation of a defective spliceosome. Thus far, many studies have focused on spliceosome mutant AML cells. Concurrently, several studies have uncovered that these spliceosome mutated cells are hypersensitive to splicing modulation by both E7107 and H3B-8800.[57-59] Interestingly, the splicing modulator 17S-FD-895 was shown to reverse pro-survival splicing patterns acquired by leukemic cells of patients with secondary AML, independent of the spliceosome mutational status.[56] In addition, this modulator impaired the maintenance of leukemic stem cells in mouse xenograft models while sparing normal hematopoietic stem cells. Recently, oncogenic activation of MYC has been associated with hypersensitivity to splicing modulation by upregulating the small nuclear ribonucleoprotein particles, including PRMT5.[60] In addition, cells with either high *MCL1* or *BCL2A1* expression were shown to be preferentially sensitive to E7107. Thus, as disruption of splicing in AML in general has become more evident, investigation of splicing modulation-based therapies is warranted in AML cells independent of their splicing factor mutation status.[61]

AIM OF THE THESIS

Cancer is thought to arise from accumulation of DNA mutations.[62] However, AML is a subtype with a relatively low mutational load compared to other cancer types. In addition, in children, DNA is far less damaged by environmental exposure or age.[63-65] Thus, AML might not only be driven by differences in genomic content but also by transcriptomic heterogeneity. Since human AML constitutes a pool of cells organized as a hierarchy that originates from a primitive hematopoietic (stem) cell, Chapter 2 of this thesis compared the vital role of RNA splicing in stem cell identity and function in both healthy and diseased stem cells. This adapted version of our literature review aimed to discuss how post-transcriptional mechanisms affect transcriptomic heterogeneity and enable pluripotent somatic stem cells to maintain homeostasis and respond to developmental cues and environmental stressors by rapidly shaping the content of their proteome. Next to RNA splicing, this chapter aimed to introduce other post-transcriptional mechanisms including RNA editing and RNA methylation. As the role of splicing in leukemia initiating- and therapy resistant stem cells has been of great interest in hematological malignancies, this thesis aimed to highlight the importance of splicing deregulation in AML specifically.

Accordingly, we first aimed to characterize the prognostic impact of splicing factor mutations which have recently been found to be recurrently mutated in AML. To gain more insight into the characteristics of adult AML patients carrying splicing factor mutations specifically, we studied their association with clinical features, cytogenetic and molecular abnormalities as well as clinical outcome in Chapter 3.

Furthermore, we aimed to identify additional patient subgroups (other than splicing factor mutated) which will benefit from the emerging splicing modulators. While splicing modulation in splicing factor mutated AML cells was already extensively studied, in Chapter 4, we demonstrated that leukemia cells of adult AML patients carrying *FLT3*/ITD, show increased sensitivity to splicing modulation. Subsequently, we explored alternative splicing in in *FLT3*/ITD and *NPM1* mutated AML in Chapter 5. In Chapter 6 we explored the potential of splicing modulation as novel therapeutic option in pediatric AML. Additionally, we generated a comprehensive transcriptome expression map of purified pediatric hematopoietic stem and hematopoietic progenitor cells specifically and evaluated genome wide alternative splicing events. This analysis identified distinctive splicing alterations in both hematopoietic stem and progenitor cells compared to non-leukemic counterparts. Subsequently, in Chapter 7, we studied the underlying mechanism of splicing deregulation in pediatric AML in more detail.

Finally, in Chapter 8 we did put our findings in the context of current literature and discussed the implications that these results may have for future AML risk stratification and treatment.

REFERENCES

1. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. N Engl J Med. 2015;373(12):1136-1152.
2. Reedijk AMJ, Klein K, Coebergh JWW, et al. Improved survival for children and young adolescents with acute myeloid leukemia: a Dutch study on incidence, survival and mortality. Leukemia. 2019;33(6):1349-1359.
3. Brunner AM, Graubert TA. Genomics in childhood acute myeloid leukemia comes of age. Nature Medicine. 2018;24(1):7-9.
4. Walter RB, Appelbaum FR, Tallman MS, Weiss NS, Larson RA, Estey EH. Shortcomings in the clinical evaluation of new drugs: acute myeloid leukemia as paradigm. Blood. 2010;116(14):2420-2428.
5. Derman BA, Larson RA. Post-remission therapy in acute myeloid leukemia: Are we ready for an individualized approach? Best Pract Res Clin Haematol. 2019;32(4):101102.
6. Major CK, Kantarjian H, Sasaki K, et al. Survivorship in AML - a landmark analysis on the outcomes of acute myelogenous leukemia patients after maintaining complete remission for at least 3 years. Leuk Lymphoma. 2020;61(13):3120-3127.
7. Bachas C, Schuurhuis GJ, Assaraf YG, et al. The role of minor subpopulations within the leukemic blast compartment of AML patients at initial diagnosis in the development of relapse. Leukemia. 2012;26(6):1313-1320.
8. Zeijlemaker W, Grob T, Meijer R, et al. CD34(+)CD38(-) leukemic stem cell frequency to predict outcome in acute myeloid leukemia. Leukemia. 2019;33(5):1102-1112.
9. Hanekamp D, Denys B, Kaspers GJL, et al. Leukaemic stem cell load at diagnosis predicts the development of relapse in young acute myeloid leukaemia patients. British Journal of Haematology. 2018;183(3):512-516.
10. Hanekamp D, Cloos J, Schuurhuis GJ. Leukemic stem cells: identification and clinical application. Int J Hematol. 2017;105(5):549-557.
11. Vetrie D, Helgason GV, Copland M. The leukaemia stem cell: similarities, differences and clinical prospects in CML and AML. Nature Reviews Cancer. 2020;20(3):158-173.
12. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. New England Journal of Medicine. 2016;374(23):2209-2221.
13. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-447.
14. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-2405.
15. Bolouri H, Farrar JE, Triche T, et al. The molecular landscape of pediatric acute myeloid leukemia reveals recurrent structural alterations and age-specific mutational interactions. Nature Medicine. 2018;24(1):103-112.
16. Zhang Y, Qian J, Gu C, Yang Y. Alternative splicing and cancer: a systematic review. Signal Transduction and Targeted Therapy. 2021;6(1):78.
17. Adamia S, Haibe-Kains B, Pilarski PM, et al. A genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. Clin Cancer Res. 2014;20(5):1135-1145.
18. Cooper TA, Wan L, Dreyfuss G. RNA and disease. Cell. 2009;136(4):777-793.
19. Climente-González H, Porta-Pardo E, Godzik A, Eyras E. The Functional Impact of Alternative Splicing in Cancer. Cell Rep. 2017;20(9):2215-2226.
20. Jung H, Lee D, Lee J, et al. Intron retention is a widespread mechanism of tumor-suppressor inactivation. Nat Genet. 2015;47(11):1242-1248.
21. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer. 2016;16(7):413-430.
22. Goldman SL, Hassan C, Khunte M, et al. Epigenetic Modifications in Acute Myeloid Leukemia: Prognosis, Treatment, and Heterogeneity. Front Genet. 2019;10:133.
23. Ehrlich M. DNA methylation in cancer: too much, but also too little. Oncogene. 2002;21(35):5400-5413.

24. Zipeto MA, Court AC, Sadarangani A, et al. ADAR1 Activation Drives Leukemia Stem Cell Self-Renewal by Impairing Let-7 Biogenesis. Cell Stem Cell. 2016;19(2):177-191.
25. Jiang Q, Crews LA, Barrett CL, et al. ADAR1 promotes malignant progenitor reprogramming in chronic myeloid leukemia. Proceedings of the National Academy of Sciences. 2013;110(3):1041-1046.
26. Han L, Diao L, Yu S, et al. The Genomic Landscape and Clinical Relevance of A-to-I RNA Editing in Human Cancers. Cancer Cell. 2015;28(4):515-528.
27. Chen X-Y, Zhang J, Zhu J-S. The role of m6A RNA methylation in human cancer. Molecular Cancer. 2019;18(1):103.
28. Sebestyén E, Singh B, Miñana B, et al. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. Genome Res. 2016;26(6):732-744.
29. Seiler M, Peng S, Agrawal AA, et al. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. Cell Rep. 2018;23(1):282-296.e284.
30. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. Genome Med. 2015;7(1):45.
31. Berget SM, Moore C, Sharp PA. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proceedings of the National Academy of Sciences. 1977;74(8):3171.
32. Chow LT, Gelinas RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell. 1977;12(1):1-8.
33. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol. 2009;10(11):741-754.
34. Alekseyenko AV, Kim N, Lee CJ. Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes. Rna. 2007;13(5):661-670.
35. Mehmood A, Laiho A, Venäläinen MS, McGlinchey AJ, Wang N, Elo LL. Systematic evaluation of differential splicing tools for RNA-seq studies. Briefings in Bioinformatics. 2019;21(6):2052-2065.
36. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010;463(7280):457-463.
37. Bonnal SC, López-Oreja I, Valcárcel J. Roles and mechanisms of alternative splicing in cancer — implications for care. Nature Reviews Clinical Oncology. 2020;17(8):457-474.
38. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nature Medicine. 2016;22(9):976-986.
39. Fu X-D, Ares M. Context-dependent control of alternative splicing by RNA-binding proteins. Nature Reviews Genetics. 2014;15(10):689-701.
40. Van Nostrand EL, Freese P, Pratt GA, et al. A large-scale binding and functional map of human RNA-binding proteins. Nature. 2020;583(7818):711-719.
41. Carazo F, Gimeno M, Ferrer-Bonsoms JA, Rubio A. Integration of CLIP experiments of RNA-binding proteins: a novel approach to predict context-dependent splicing factors from transcriptomic data. BMC Genomics. 2019;20(1):521.
42. Chua BA, Van Der Werf I, Jamieson C, Signer RAJ. Post-Transcriptional Regulation of Homeostatic, Stressed, and Malignant Stem Cells. Cell Stem Cell. 2020;26(2):138-159.
43. Chen K, Dai X, Wu J. Alternative splicing: An important mechanism in stem cell biology. World J Stem Cells. 2015;7(1):1-10.
44. Lu Y, Loh YH, Li H, et al. Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. Cell Stem Cell. 2014;15(1):92-101.
45. Han H, Irimia M, Ross PJ, et al. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature. 2013;498(7453):241-245.
46. Holm F, Hellqvist E, Mason CN, et al. Reversion to an embryonic alternative splicing program enhances leukemia stem cell self-renewal. Proceedings of the National Academy of Sciences. 2015;112(50):15444-15449.
47. Atlasi Y, Mowla SJ, Ziaee SA, Gokhale PJ, Andrews PW. OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells. Stem Cells. 2008;26(12):3068-3074.

48. Darman RB, Seiler M, Agrawal AA, et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. Cell Rep. 2015;13(5):1033-1045.
49. Rossi D, Bruscaggin A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011;118(26):6904-6908.
50. Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med. 2011;365(15):1384-1395.
51. Nikpour M, Scharenberg C, Liu A, et al. The transporter ABCB7 is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts. Leukemia. 2013;27(4):889-896.
52. Malcovati L, Papaemmanuil E, Bowen DT, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood. 2011;118(24):6239-6246.
53. Damm F, Thol F, Kosmider O, et al. SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. Leukemia. 2012;26(5):1137-1140.
54. Hou HA, Liu CY, Kuo YY, et al. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. Oncotarget. 2016;7(8):9084-9101.
55. Kim E, Ilagan JO, Liang Y, et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell. 2015;27(5):617-630.
56. Crews LA, Balaian L, Delos Santos NP, et al. RNA Splicing Modulation Selectively Impairs Leukemia Stem Cell Maintenance in Secondary Human AML. Cell Stem Cell. 2016;19(5):599-612.
57. Seiler M, Yoshimi A, Darman R, et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nature Medicine. 2018;24(4):497-504.
58. Buonamici S, Yoshimi A, Thomas M, et al. H3B-8800, an Orally Bioavailable Modulator of the SF3b Complex, Shows Efficacy in Spliceosome-Mutant Myeloid Malignancies. Blood. 2016;128(22):966-966.
59. Lee SC, North K, Kim E, et al. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018;34(2):225-241.e228.
60. Hsu TY, Simon LM, Neill NJ, et al. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature. 2015;525(7569):384-388.
61. Aird D, Teng T, Huang C-L, et al. Sensitivity to splicing modulation of BCL2 family genes defines cancer therapeutic strategies for splicing modulators. Nature Communications. 2019;10(1).
62. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science. 2015;347(6217):78-81.
63. Blokzijl F, De Ligt J, Jager M, et al. Tissue-specific mutation accumulation in human adult stem cells during life. Nature. 2016;538(7624):260-264.
64. Osorio FG, Rosendahl Huber A, Oka R, et al. Somatic Mutations Reveal Lineage Relationships and Age-Related Mutagenesis in Human Hematopoiesis. Cell Rep. 2018;25(9):2308-2316.e2304.
65. Welch John S, Ley Timothy J, Link Daniel C, et al. The Origin and Evolution of Mutations in Acute Myeloid Leukemia. Cell. 2012;150(2):264-278.

CHAPTER 2

Post-Transcriptional Regulation of Homeostatic, Stressed and Malignant Stem Cells

Inge Van Der Werf[1,2*]
Bernadette A. Chua[1*]
Catriona Jamieson[1,2]
Robert A.J. Signer[1]

1. Division of Regenerative Medicine, Department of Medicine Moores Cancer Center, University of California San Diego, La Jolla, CA, 92093 USA; 2. Sanford Stem Cell Clinical Center, La Jolla, CA 92037, USA

*These authors contributed equally to this work

Derived from Post-Transcriptional Regulation of Normal, Stressed and Diseased Stem Cells. Cell Stem Cell. 2020. 6;26(2):138-159

Because of the expanded use of single cell nucleic acid sequencing technology, cell identity is being increasingly defined by transcriptional profiles. Transcriptional networks play a central role in governing stem cell function and fate. This is best exemplified by pluripotent stem cells, in which four transcription factors, MYC, OCT4, SOX2 and NANOG, are essential for driving the genetic programs that support pluripotency and self-renewal and are sufficient for reprogramming somatic cells into induced pluripotent stem cells.[1-5] A central question that consequently arose from these studies was whether cell-autonomous mechanisms shape cellular identity or vice versa. While initial work focused on epigenetic mechanisms, we now appreciate that transcriptional events do not entirely determine cellular identity. Recent studies have revealed that diverse post-transcriptional mechanisms influence the functional output of genetic programs (i.e. proteome content) required by stem cells.[6-8]

In this Review, we will discuss how transcript sequence, stability and translational efficiency are regulated, at least in part, by a variety of biochemical modifications to influence stem cell identity and function. In addition, we will discuss how defects in these post-transcriptional mechanisms deregulate tissue-specific stem cells and progenitors in human disease and stress conditions and examine their potential as both diagnostic and therapeutic targets.

RNA PROCESSING

Post-transcriptional regulation begins with extensive processing and modification of RNA. Precursor messenger RNAs (pre-mRNAs) transcribed from coding genes may be capped, spliced, cleaved and polyadenylated to make them competent for translation into functional proteins. In addition to these processing events, coding and non-coding RNAs can be biochemically modified via methylation, pseudouridylation or editing. These epitranscriptomic modifications can alter RNA coding sequences, localization, stability and translational efficiency. Therefore, RNA processing and epitranscriptomic alterations play a key role in regulating proteome content and diversity.[9,10]

RNA SPLICING

To date, alternative pre-mRNA splicing is the most extensively studied mRNA modification that influences protein composition.[11] Splicing is the process of intron removal and exon joining that is necessary for converting intron containing pre-mRNAs into mRNAs that are competent for translation into functional proteins. In many cases, splicing takes place using alternative splice sites that can result in the production of distinct mRNAs that code for different protein isoforms that exhibit unique structural and functional properties. Many human splicing events are not conserved in mice thereby suggesting that splicing is essential for fine-tuning human gene regulation.[12,13]

Splice isoform diversity is highest in human embryonic stem cells (hESCs) and decreases upon differentiation. This phenomenon, referred to as isoform specialization, is mediated in part by a number of splicing factors and other RNA binding proteins that are differentially expressed during development.[14] Splice isoform expression patterns also distinguish human stem and progenitor cell fate, aging and malignant transforming potential.[15,16]

Alternative splicing can influence gene expression to either promote or impair stem cell function.[14,17] For example, depletion of SRSF2, a member of the serine/arginine-rich pre-mRNA splicing factor family, was shown to decrease the expression of pluripotency factors *OCT4* and *NANOG*, and disrupt self-renewal of hESCs.[18] Interestingly, OCT4 itself can bind to the *SRSF2* promoter, and depletion of OCT4 reduces *SRSF2* expression, suggesting that there is reciprocal regulation of splicing and pluripotency factors.[18] In contrast to SRSF2, muscleblind like splicing factors (MBNLs) negatively regulate stem cell self-renewal. MBNL proteins are more highly expressed by differentiated cells than hESCs and have been shown to repress stem cell specific splicing patterns.[19,20] One key splice variant whose production is repressed by MBNL proteins encodes a specific isoform of FOXP1, a transcription factor that is expressed by hESCs but absent during differentiation. This stem cell specific FOXP1 variant arises from inclusion of a stem cell specific exon that alters its DNA binding specificity to promote the expression of pluripotency factors, including *OCT4*, *SOX2* and *NANOG*, and suppress the expression of differentiation factors.[21] Consistent with their role in suppressing pluripotency, knockdown of MBNL proteins enhances the reprogramming of somatic cells into induced pluripotent stem cells (iPSCs).[22] In addition, reprogramming has been associated with re-acquisition of a pluripotent alternative splicing profile.[23] Thus, alternative splicing has important roles in pluripotency as well as cellular reprogramming.

Alternative splicing can also directly affect the function of genes important for pluripotency, including *OCT4*, *TCF3*, *MBD2* and *SALL4*.[18, 24-26] The expression of *OCT4A*, a splice variant of *OCT4*, is specifically expressed in pluripotent stem cells and is necessary for their self-renewal.[24] In contrast, *OCT4B*, another alternatively spliced isoform of *OCT4*, is expressed by both stem cells and somatic cells, but has no apparent role in regulating stem cell function.[24] A third splice isoform, *OCT4B1*, is also specifically expressed in hESCs, and its expression is increased in response to various cellular stresses.[27] Indeed, mRNA splicing efficiency is thought to be altered by stress, and may provide stem cells with a mechanism to rapidly alter the content of their proteome in response to environmental cues without requiring epigenetic and transcriptional changes.[28] Overall, these studies indicate that alternative splicing is important for regulating pluripotency and cell fate specification.

In addition to generating functional variants, splicing can alter transcript stability and translational efficiency to influence stem cell development and function. Alternative splicing can introduce premature stop codons that can trigger nonsense mediated mRNA decay and can cause untranslated region (UTR) variation, which can affect translation efficiency, mRNA stability and subcellular localization.[28-30] One recent study identified temporal splicing changes in the 3' UTR of *HMGA2* during human hematopoietic stem cell (HSC) ontogeny.[15] Human fetal liver and cord blood HSCs were shown to express distinct isoforms of *HMGA2*. Fetal HSCs express a longer isoform of *HMGA2* that contains a distinct terminal exon and a 3-fold longer 3'-UTR as compared to the short isoform which is highly expressed by neonatal HSCs. Although the function of both isoforms is similar, the shorter 3'UTR present in the short isoform of *HMGA2* enables it to more effectively escape repression mediated by various microRNAs (miRNAs). This short isoform enables neonatal HSCs to sustain *HMGA2* expression and self-renewal potential despite the increased presence of let-7 family miRNAs that can suppress *HMGA2*.[15]

Given the importance of RNA splicing in regulating both transcriptome and proteome diversity in normal stem cells, recent attention has focused on cancer stem cells and their capacity to hijack splicing to support malignant growth. Cancer stem cells exhibit splicing patterns reminiscent of undifferentiated stem cells, and these splicing patterns are at least partly mediated by MBNL1.[16,31] Furthermore, splicing factors, such as *SRSF2*, *U2AF1* and *SF3B1*, have been found to be mutated or epigenetically modified in pre-leukemic and leukemic disorders.[32,33] This dysregulation is coupled with observations that cancer stem cells express both stem cell regulatory and pro-survival splice variants of a number of genes. Chronic myeloid leukemia blast crisis stem cells have been shown to express high levels *CD44v3*, an isoform of CD44 that is typically expressed by hESCs.[20] CD44 isoform switching also occurs in breast cancer stem cells.[34] Mis-splicing of *GSK3β* and concomitant activation of β-catenin was shown to be important for leukemia stem cell self-renewal, while pro-apoptotic splice variants of the *BCL2* family promote leukemia stem cell survival.[35,36] Overall, dozens of dysregulated spliceosome components, splice variants and splicing patterns have been identified in cancer stem cells, and much work remains to uncover the functional significance of these changes. However, these cancer stem cell specific splicing events are already revealing new opportunities to improve diagnostic and prognostic tools as well as to develop new targeted therapies.[16] Thus, it is likely that we have only begun to uncover how cell-type and context specific differences in splicing enable stem cells to remodel their proteome for optimal function in response to developmental signals and environmental cues.

MICRO-RNA

MicroRNAs (miRNAs) are small non-coding single stranded RNA molecules that repress gene expression through translational inhibition or by promoting degradation of mRNA. miRNA biogenesis is a stepwise process that starts with transcription of primary miRNAs (pri-miRNAs) in the nucleus. Pri-miRNAs are subsequently processed into stem-loop precursor miRNA (pre-miRNA) by a complex composed of DGCR8 and other factors.[37] Mature miRNAs are then cleaved by Dicer1 and incorporated into the RNA induced silencing complex (RISC).[38] Base pairing between the RISC-bound miRNA and 3' UTR of target mRNA triggers mRNA decay or translational repression.[39]

MicroRNA-mediated gene silencing is a mechanism that regulates stem cell pluripotency. Mouse ESCs deficient in *Dgcr8* lose the ability to differentiate and express the pluripotency genes *Oct4*, *Sox2*, and *Nanog* at high levels.[40,41] Additionally, *Dicer1*-null mouse ESCs have diminished expression of differentiation markers *in vitro* and *in vivo*.[42] Consistent with these observations, miR-134, miR-296, and miR-470 mediate mouse ESC differentiation by disrupting *Oct4*, *Sox2*, and *Nanog* expression.[43] MicroRNAs similarly repress hESC pluripotency by targeting *OCT4*, *SOX2*, and *KLF4* transcripts.[44] Studies on somatic cell reprogramming also point to the role of miRNA in regulating pluripotency. miRNAs can increase the reprogramming efficiency of mouse embryonic fibroblasts into iPSCs and human skin cancer cells to a pluripotent state.[45,46]

miRNAs can also regulate the activation, proliferation, and differentiation of somatic stem cells. miR-128 and miR-181 maintain hematopoietic stem and progenitor cells by inhibiting their dif-

ferentiation into mature hematopoietic lineages.[47] In muscle stem cells Pax3, which controls stem cell activation, is subject to repression by miR-206.[48] miRNAs can also promote myogenesis by targeting repressors of muscle-related transcription factors or can enhance myoblast proliferation by targeting transcripts essential for differentiation.[49] Lastly, neuronal lineages and astrocytes differentially express miRNA species that influence lineage specification.[50,51]

Disruptions in the miRNA pathway have been implicated in several types of cancer. Germline and somatic mutations in *DICER1* can predispose individuals to cancer and impaired DICER1 function can promote colon cancer and endometrial cancer stemness.[52-54] Many cancer types show miRNA signatures characterized by a defect in miRNA biogenesis and global downregulation of miRNA production.[55-57] Downregulation of miR-34a has been observed in breast cancer, colon cancer, pancreatic cancer, neuroblastoma, hepatocellular carcinoma, and non-small-cell lung cancer suggesting that it may function as a tumor suppressor.[58] miR-34a also inhibits the proliferation of breast cancer stem cells and prostate cancer stem cells by suppressing CD44 expression.[59,60] Lastly, miRNAs have also been implicated in stem cell-related signal transduction pathways including Wnt, Notch, and Hedgehog.[58] Given their role in cancer pathogenesis, miRNAs show potential as diagnostic and prognostic biomarkers and provide a new avenue for treating cancer.

RNA METHYLATION

RNA can undergo a variety of biochemical modifications that collectively are referred to as the epitranscriptome. The most prevalent mRNA modification is methylation of adenosine at the nitrogen-6 position (N6-methyl adenosine [m^6A]).[61] m^6A deposition is catalyzed by the m^6A methyltransferase ('writer') complex that consists of methyltransferase like protein 3 (METTL3) or METTL14 along with WTAP and VIRMA (KIAA1429).[62-65] Although RNA methylation has been known about for decades, only recently was it shown to be reversible through the discovery of m6A demethylases ('erasers') such as FTO and ALKBH5.[66,67] The dynamic nature of m^6A modifications has sparked tremendous interest in its biological function, which is mediated by RNA binding proteins that recognize and bind m^6A modified RNA ('readers').

m^6A modifications can functionally alter mRNAs, pre-mRNAs, miRNAs and non-coding RNAs, such as rRNA and tRNA. The major effects of m^6A on mRNA are mediated by the reader proteins YTHDF1 and YTHDF2. YTHDF1 can promote cap dependent translation of m^6A modified mRNAs by enhancing interaction with translation initiation factors. In contrast, YTHDF2 typically promotes mRNA decay, thereby suppressing translation.[68,69] m^6A modification of pre-mRNAs can alter mRNA export or induce structural changes that promote interaction with different RNA binding proteins that in turn alter splicing or editing patterns.[70-74] m^6A in pri-miRNAs promotes processing and miRNA biogenesis.[75] Hence, m^6A can influence the transcriptome and proteome through the regulation of diverse post-transcriptional mechanisms.

Patterns of m^6A modifications vary dramatically in a temporal, tissue and cell type specific manner. Although these patterns are mediated in part by differential expression of writers, erasers and readers, how these context specific patterns of m^6A are established remains largely unknown. *In vivo*, germline deletion of *Mettl3* results in early (E5.5-7.5) embryonic lethality

associated with impaired induction of cellular differentiation.[71] In zebrafish embryos, morpholino-mediated knockdown of either *mettl3* or *wtap* also cause widespread differentiation defects.[76] Germline deletion of the erasers *Fto* or *Alkbh5* in mice are not lethal, but the former causes severe growth defects and the latter impairs male fertility.[67,77] Thus, m^{6}A methylation is critical for normal development.

By regulating mRNA stability, m^{6}A modifications have a striking impact on ESC self-renewal and differentiation. A wide range of transcripts, including the core pluripotency transcription factors *SOX2* and *NANOG*, are marked by m^{6}A. Genetic inactivation of *Mettl3* in naïve mouse ESCs results in widespread loss of m^{6}A modifications that enhance self-renewal and impair differentiation *in vitro* and *in vivo*.[78] Since m^{6}A modifications can promote mRNA degradation, the loss of m^{6}A in ESCs stabilizes pluripotency-promoting transcripts such as *NANOG*.[71] Conversely, overexpression of METTL3 enhances reprograming efficiency of human fibroblasts into iPSCs by stabilizing pluripotency factors.[14] Interestingly, *Mettl3* or *Mettl14* knockdown within primed mouse ESCs also reduces m^{6}A modifications, but has dichotomous effects on self-renewal and differentiation as compared to naïve ESCs. Knockdown of *Mettl3* or *Mettl14* impairs self-renewal and promotes differentiation of primed mouse ESCs.[79] This difference can be partially explained by the rebalancing of self-renewal and differentiation transcripts that occurs in primed but not naïve ESCs. In naive ESCs, the loss of m^{6}A enhances the stability of highly expressed self-renewal genes, while in primed ESCs the loss of m^{6}A enhances the stability of highly expressed differentiation genes. This example demonstrates the importance and precision with which post-transcriptional mechanisms of gene regulation influence stem cell identity.

m^{6}A also regulates the emergence, self-renewal and differentiation of somatic stem cells. *Mettl3*-deficient zebrafish embryos do not undergo the endothelial to hematopoietic transition and fail to produce early hematopoietic stem and progenitor cells.[80] This occurs in part because in the absence of m^{6}A, YTHDF2, a reader protein that promotes mRNA decay, is delayed in binding to the arterial endothelial mRNAs for *notch1a* and *rhoca*. This results in sustained Notch signaling in endothelial cells, which suppresses hematopoietic specification.[80] Knockdown of either *METTL3* or *METTL14* in human cord blood derived hematopoietic stem and progenitor cells modestly impairs stem cell proliferation and promotes myeloid differentiation.[81,82] However, conditional deletion of *Mettl3* from adult mouse HSCs leads to HSC accumulation, reduced reconstituting activity and impaired differentiation due in part to a loss of m^{6}A-mediated translation of c-Myc.[83] Conditional deletion of *Mettl14* from adult mouse HSCs also reduces long-term multilineage reconstituting activity in transplantation assays.[82] In addition, conditional deletion of *Mettl3* from mouse skeletal stem cells impairs osteogenic differentiation and bone development by regulating the translational efficiency of parathyroid hormone receptor 1 (*Pthr1*).[84] Thus, m^{6}A exhibits exquisite context dependent regulation of gene expression that can contribute to a divergent transcriptome and proteome. Overall, our understanding of how m^{6}A influences proteome content and cellular function is still in its infancy, but it clearly plays a key role in stem cell regulation and cell fate determination.

m^{6}A methylation also influences cancer stem cells. Breast cancer stem cells exhibit reduced m^{6}A methylation of NANOG and KLF4, which contributes to elevated expression of both plu-

ripotency factors.[85] In both cervical cancer and acute myeloid leukemia, high *FTO* expression has been reported to be important for cell survival.[86,87] High *METTL3* expression has also been reported in acute myeloid leukemia, with subsequent methylation of *MYC*, *MYB*, *PTEN*, and *BCL2*, which could support leukemia stem cell survival.[81,82,88] The finding that both m^6A writers and erasers are highly expressed in cancer highlights the importance of determining the role of m^6A methylation and subsequent binding of reader proteins on mRNA stability and translation efficiency in a temporal, tissue and cell type specific manner.

In addition to m^6A, adenosines can be methylated at the nitrogen-1 position (m^1A).[89-92] m^1A modifications were traditionally thought to regulate the stability of tRNAs and rRNAs. However, recent advances in sequencing technology have revealed tissue specific methylation of mRNAs, in the 5' UTRs within the mRNA cap that enhances translational efficiency. To date, m^1A modifications have been associated with both increased and suppressed protein synthesis.[89-91] Moreover, m^1A deposition occurs in a tissue and cell type specific manner and can be dynamically regulated in response to environmental stress. Because m^1A methylation was shown to be highly conserved in mice, it is likely to be essential for gene regulation. However, the functional importance of m^1A, particularly in stem cells, remains largely unknown.

RNA can also be methylated at the carbon-5 position of cytosine (m^5C).[93] The m^5C modification most commonly occurs on tRNAs and rRNAs.[93,94] However, NSUN2, one of seven known cytosine-5 methylases, was recently shown to deposit m^5C on some mRNAs as well.[95,96] There are at least six other enzymes capable of methylating cytosine 5, including NSUN1, NSUN3, NSUN4, NSUN5, NSUN6 and DNMT2.[97]

Dynamic changes in m^5C deposition in rRNAs and tRNAs can impact ribosome biogenesis, polysome assembly, translation fidelity and tRNA stability.[98-102] This widespread influence on the translational apparatus enables m^5C levels to modulate global protein synthesis and regulate specific translational programs.[103] Loss of m^5C is associated with suppression of global protein synthesis. Deletion or loss of function of NSUN2 leads to widespread loss of m^5C in most tRNAs, leading to cleavage and the accumulation of tRNA-derived small non-coding RNAs, which can impair translation elongation and reduce protein synthesis.[98] In addition to dampening global protein synthesis, loss of m^5C also increases translation of stress response genes, as well as genes regulating cell motility, morphogenesis and apoptosis.[104-107]

Loss of m^5C in tRNAs associated with *Nsun2* and/or *Dnmt2* deficiency impairs differentiation in multiple murine tissues, including the brain, blood, skin, testis, liver and fat.[102,105, 107-109] The specialized translational program associated with m^5C loss is sufficient to maintain epidermal stem cells in their undifferentiated state, but does not enable normal differentiation. Increased m^5C is required for epidermal stem cells to increase protein synthesis in response to cytotoxic stress.[98] These studies suggest that dynamic control of the epitranscriptome is required for stem cells to appropriately survive and promote regeneration in response to stress.

PSEUDOURIDYLATION

Pseudouridine (Ψ, 5-ribosyluracil) is the most widespread RNA modification.[110,111] Pseudouridine is present within mRNAs and non-coding RNAs such as rRNAs, tRNAs, splicesomal small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs) and telomerase RNA. Because pseudouridine contains an extra hydrogen bond donor, it promotes base stacking interactions that typically make RNA backbones more rigid.[110] This modification thus primarily influences RNA structure, which can in turn influence interactions with other biomolecules. Pseudouridylation plays an important role in regulating protein synthesis by enhancing tRNA stability, influencing base pairing within the ribosome decoding center and altering translation termination.[112] While pseudouridylation is thought to promote translation, its effects are diverse thereby making it a complex modification to understand.[113]

There are at least 13 pseudouridine synthases (PUSs). These PUSs can catalyze pseudouridylation in a guide RNA independent manner.[114] Pseudouridylation can also be catalyzed in a guide RNA dependent manner by Dyskerin (DKC1), in a process that depends upon target sequence complementarity to a box H/ACA snoRNA in complex with several other proteins.[114] Currently, pseudouridylation is thought to be irreversible as no readers or erasers have yet been identified.

The effects of pseudouridylation on stem cells were initially observed in the context of *DKC1* mutations. *DKC1* is mutated in the X-linked form of the human disease dyskeratosis congenita, which is a disorder characterized by short telomeres.[115] TERC contains a highly conserved pseudouridylation site in a key region required for TERT binding, raising the possibility that defects in pseudouridylation impair telomerase activity. DKC1 promotes telomere elongation in iPSCs, and also regulates the expression of *OCT4* and *SOX2*.[115,116] In agreement with a potential role in promoting pluripotency, *DKC1* mutant fibroblasts exhibit impaired iPSC reprogramming.[117] *DKC1* mutations are also associated with widespread loss of rRNA modifications, and the catalytic activity of Dkc1 is required for normal HSC differentiation.[118]

Guide independent pseudouridylation has been shown to be important for stem cells as well. *PUS7* deficient hESCs exhibit impaired activation of tRNA derived small fragments that are required for translational control.[111] *PUS7* deficiency is associated with increased protein synthesis and impaired germ layer specification.[111] Dysregulation of the Pus7 mediated translational program is also required for HSC commitment.[111] Pseudouridylation is thus required for translational control and normal stem cell function.

RNA EDITING

Another mechanism of RNA sequence modification that contributes to transcriptomic diversity is RNA editing. RNA editing is the most common post-transcriptional modification detected by whole transcriptome RNA sequencing in human cells.[119] The most frequent type of RNA editing event in mammals involves deamination of adenosine into inosine (A-to-I).[120-122] When A-to-I editing occurs within protein coding exons, inosine bases are read as guanosines by the translational apparatus.[121-123] Although editing events can occur within coding regions, most RNA editing sites are located within non-coding regions, such as introns and UTRs.[119] In humans,

approximately 90% of these editing sites are located within primate-specific Alu sequences, which are transposable elements that represent approximately 11% of the human genome.[124] Editing of Alu sequences within non-coding regions can cause the introduction of new splice sites (splicing machinery also recognizes inosines as guanosines), promote RNA degradation, and induce sequestration of RNAs in discrete nuclear compartments.[103, 125-127] Also, RNA editing can modulate gene expression by impairing miRNA biogenesis.[122,128]

Editing of RNA is catalyzed by members of the adenosine deaminase associated with RNA (ADAR) family. To date, three members of the ADAR family have been identified in vertebrate animals: *ADAR* (ADAR1), *ADARB1* (ADAR2) and *ADARB2* (ADAR3).[121,123] *ADAR1* is ubiquitously expressed and is essential for embryonic development.[120] *Adar1* deficiency causes embryonic lethality in mice as a consequence of defective erythropoiesis and is associated with hyperactive interferon signaling and widespread apoptosis.[120] Germline *ADAR1* mutations in humans are associated with Aicardi–Goutières syndrome and dyschromatosis symmetrica hereditaria.[129,130] *ADAR2* is also widely expressed, but is not required for embryonic development.[131] However, *Adar2* deficiency is associated with neuronal death and seizures that cause postnatal lethality in the first few weeks of life.[132,133] Aberrant RNA editing profiles are accordingly associated with a number of human neurological and psychiatric disorders.[134] *ADAR3* expression is largely restricted to the brain and has not yet been shown to exhibit RNA editing activity.[135] Rather, ADAR3 has been shown to inhibit RNA editing.[136]

hESCs exhibit high levels of RNA editing.[137] RNA editing in hESCs is enriched within non-coding regions of double stranded RNA marked by inverted Alu repeats.[137] The global abundance of transcript editing is reduced during differentiation, particularly in the neural lineage.[121,137] Knockdown of *ADAR1* is associated with increased expression of genes associated with differentiation and developmental processes and may be dispensable for hESCs.[137,138] Human fibroblasts reprogrammed into iPSCs exhibit RNA editing profiles that more closely resemble hESCs than mature fibroblasts, suggesting that there is reprogramming of the RNA editome. Furthermore, modulating *ADAR1* expression influences the efficiency of reprogramming.[139, 140] Overall, these studies suggest that ADAR1-mediated RNA editing contributes to the establishment of pluripotency and cell fate determination.

RNA editing also regulates somatic stem and progenitor cell populations. Conditional deletion of *Adar1* impairs the multi-lineage reconstituting activity of mouse HSCs.[141,142] Adar1 deficiency also increases hematopoietic progenitor cell apoptosis. This phenotype depends upon the RNA editing domain of Adar1 and is associated with upregulation of interferon signaling.[142] Based on these phenotypes, it remains unclear whether Adar1 deficiency directly impairs HSC function or whether reconstitution is impaired because of defects in progenitor cells. In addition to regulating cell death, ADAR1 can also regulate quiescence and cell cycle entry of human hematopoietic stem and progenitor cells. Lentiviral overexpression of ADAR1 within human cord blood derived hematopoietic stem and progenitor cells increased expression of specific cell cycle and self-renewal regulatory transcripts that enhance the expansion of these cells *in vitro*.[128] At least some of the effects of ADAR1 on stem and progenitor cell expansion occur through an RNA editing dependent mechanism. ADAR1 mediated RNA editing of pri-miR-26a at the

Drosha cleavage site impaired maturation of miR-26a. Subsequent reduction of miR-26a led to a cascade of gene expression changes that enhanced cell cycle transit. Expression of *EZH2*, a direct target of miR-26a, was increased in hematopoietic stem and progenitor cells following ADAR1 overexpression. Moreover, EZH2 subsequently repressed the expression of *CDKN1A*, a negative regulator of cell cycle entry that can promote HSC quiescence.[128]

ADAR1 also promotes intestinal homeostasis and stem cell maintenance. *Adar1* is highly expressed by Lgr5+ cells in the intestine of adult mice. Conditional deletion of *Adar1* results in rapid apoptosis of Lgr5+ stem cells in both the small and large intestine. In contrast to the Lgr5+ cells, *Adar1* deficiency caused expansion of intestinal progenitors and Paneth cells, although enterocytes, goblet cells and enteroendocrine cells were all depleted.[143] Similar to the hematopoietic system, *Adar1* deficiency in the intestine was associated with increased interferon signaling, but was also marked by endoplasmic reticulum stress and activation of the unfolded protein response that at least partially contributed to crypt apoptosis.[143] Together, these studies support an essential role of ADAR1 and RNA editing in both tissue homeostasis and stem cell maintenance.

Though only a handful of editing sites have been identified in normal cells, editing events in cancer have been characterized in more detail. Editing of mRNAs encoding *GLI1*, *GSK3β*, *AZIN1* and *APOBEC3D* have been identified and were found to be required for survival of leukemia stem and progenitor cells.[144] ADAR1-mediated editing of the *MDM2* 3'UTR has also been shown to reduce the binding of mir-155 as well as other negative regulatory miRNAs.[128] As a consequence of hyper-editing, *MDM2* mRNA is stabilized within leukemia stem cells, resulting in increased MDM2 protein and enhanced p53 degradation. In acute myeloid leukemia, editing of *PTPN6* was found to abrogate splicing and is thought to be important for leukemogenesis.[145]

While only a small number of editing sites have been characterized, bioinformatics analyses have predicted A-to-I changes to be far more abundant. To date, A-to-I editing profiles of more than 6000 patient samples of 17 cancer types revealed a surprising increase in RNA editing events in tumor tissue relative to normal tissue.[19] Furthermore, it has been suggested that RNA editing could also affect the therapeutic response to immunotherapy. Loss of function of ADAR1 was found to improve response to PD-1 checkpoint blockade, and PD-L1 expression is under significant translational control.[146,147] Thus, posttranscriptional regulation appears to play a major role in resistance to immunotherapy. Overall, targets of RNA editing vary dramatically across species, tissues and cell types.[121] Recent parallel analysis of RNA secondary structure sequencing (PARS-seq) reveals that ADAR1 regulates RNA topology and ribosomal occupancy resulting in cell type and context specific changes in protein turnover rates.[148] Consistent with these context dependent effects on gene expression, RNA editing exhibits distinct functional effects on stem and progenitor cells. Future studies must focus on identifying edited transcripts that further explain the emerging role of RNA editing in normal and malignant stem cell biology.

CONCLUSION

Cumulative advances in whole genome, whole transcriptome and single cell RNA sequencing, in addition to murine and humanized model systems, have helped to shape our understanding

of stem cell hierarchies in homeostatic, stressed and diseased states particularly in pre-malignancy and during malignant transformation. More recently, post-transcriptional (epitranscriptomic) RNA processing alterations, such as RNA editing and methylation, have been shown to alter transcript splicing, stability, and microRNA targeting that can cause recoding of transcripts or alter translational efficiency. The observation that many of these epitranscriptomic mechanisms are deregulated in cancer underscores the importance of determining their role in regulating proteome content in a cell type and tissue specific manner. Overall, transcriptomic and epitranscriptomic events that dictate alterations in translation can now be interrogated to predict stem cell function in both benign and malignant settings and can be developed as predictive biomarkers of stem cell fitness and may ultimately add new layers in the advancement and complexity of precision medicine.

Ultimately, RNA processing and the epitranscriptome, but also translational control and protein degradation, are key regulators of proteome complexity and play pivotal roles in regulating stem cell identity and function and when dysregulated contribute to cancer stem cell propagation.[149] The intersection of these nascent fields provides a fulcrum for developing clinically tractable methods to track stem cell fitness and cancer stem cell propagation for the benefit of patients with stem cell driven degenerative diseases and cancer.

REFERENCES

1. Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. Cell 122, 947-956.
2. Park, I.H., Zhao, R., West, J.A., Yabuuchi, A., Huo, H., Ince, T.A., Lerou, P.H., Lensch, M.W., and Daley, G.Q. (2008). Reprogramming of human somatic cells to pluripotency with defined factors. Nature 451, 141-146.
3. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of pluripotent stem cells from adult human fibroblasts by defined factors. Cell 131, 861-872.
4. Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126, 663-676.
5. Yu, J., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., et al. (2007). Induced pluripotent stem cell lines derived from human somatic cells. Science 318, 1917-1920.
6. Nilsen, T.W., and Graveley, B.R. (2010). Expansion of the eukaryotic proteome by alternative splicing. Nature 463, 457-463.
7. van den Berg, P.R., Budnik, B., Slavov, N., and Semrau, S. (2017). Dynamic post-transcriptional regulation during embryonic stem cell differentiation. bioRxiv, 123497.
8. Williamson, A.J., Smith, D.L., Blinco, D., Unwin, R.D., Pearson, S., Wilson, C., Miller, C., Lancashire, L., Lacaud, G., Kouskoff, V., et al. (2008). Quantitative proteomics analysis demonstrates post-transcriptional regulation of embryonic stem cell differentiation to hematopoiesis. Mol Cell Proteomics 7, 459-472.
9. Kim, E., Goren, A., and Ast, G. (2008). Alternative splicing: current perspectives. Bioessays 30, 38-47.
10. Kwon, S.C., Yi, H., Eichelbaum, K., Fohr, S., Fischer, B., You, K.T., Castello, A., Krijgsveld, J., Hentze, M.W., and Kim, V.N. (2013). The RNA-binding protein repertoire of embryonic stem cells. Nat Struct Mol Biol 20, 1122-1130.
11. Chen, M., and Manley, J.L. (2009). Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol 10, 741-754.
12. Thanaraj, T.A., Clark, F., and Muilu, J. (2003). Conservation of human alternative splice events in mouse. Nucleic Acids Res 31, 2544-2552.
13. Yeo, G.W., Van Nostrand, E., Holste, D., Poggio, T., and Burge, C.B. (2005). Identification and analysis of alternative splicing events conserved in human and mouse. Proc Natl Acad Sci U S A 102, 2850-2855.
14. Chen, K., Dai, X., and Wu, J. (2015). Alternative splicing: An important mechanism in stem cell biology. World J Stem Cells 7, 1-10.
15. Cesana, M., Guo, M.H., Cacchiarelli, D., Wahlster, L., Barragan, J., Doulatov, S., Vo, L.T., Salvatori, B., Trapnell, C., Clement, K., et al. (2018). A CLK3-HMGA2 Alternative Splicing Axis Impacts Human Hematopoietic Stem Cell Molecular Identity throughout Development. Cell Stem Cell 22, 575-588.e577.
16. Crews, L.A., Balaian, L., Delos Santos, N.P., Leu, H.S., Court, A.C., Lazzari, E., Sadarangani, A., Zipeto, M.A., La Clair, J.J., Villa, R., et al. (2016). RNA Splicing Modulation Selectively Impairs Leukemia Stem Cell Maintenance in Secondary Human AML. Cell Stem Cell 19, 599-612.
17. Aaronson, Y., and Meshorer, E. (2013). Stem cells: Regulation by alternative splicing. Nature 498, 176-177.
18. Lu, Y., Loh, Y.H., Li, H., Cesana, M., Ficarro, S.B., Parikh, J.R., Salomonis, N., Toh, C.X., Andreadis, S.T., Luckey, C.J., et al. (2014). Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. Cell Stem Cell 15, 92-101.
19. Han, L., Diao, L., Yu, S., Xu, X., Li, J., Zhang, R., Yang, Y., Werner, H.M.J., Eterovic, A.K., Yuan, Y., et al. (2015). The Genomic Landscape and Clinical Relevance of A-to-I RNA Editing in Human Cancers. Cancer Cell 28, 515-528.
20. Holm, F., Hellqvist, E., Mason, C.N., Ali, S.A., Delos-Santos, N., Barrett, C.L., Chun, H.J., Minden, M.D., Moore, R.A., Marra, M.A., et al. (2015). Reversion to an embryonic alternative splicing program enhances leukemia stem cell self-renewal. Proc Natl Acad Sci U S A 112, 15444-15449.
21. Gabut, M., Samavarchi-Tehrani, P., Wang, X., Slobodeniuc, V., O'Hanlon, D., Sung, H.K., Alvarez, M., Talukder, S., Pan, Q., Mazzoni, E.O., et al. (2011). An alternative splicing switch regulates embryonic stem cell pluripotency and reprogramming. Cell 147, 132-146.

22. Han, H., Irimia, M., Ross, P.J., Sung, H.K., Alipanahi, B., David, L., Golipour, A., Gabut, M., Michael, I.P., Nachman, E.N., et al. (2013). MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature 498, 241-245.
23. Ohta, S., Nishida, E., Yamanaka, S., and Yamamoto, T. (2013). Global splicing pattern reversion during somatic cell reprogramming. Cell Rep 5, 357-366.
24. Atlasi, Y., Mowla, S.J., Ziaee, S.A., Gokhale, P.J., and Andrews, P.W. (2008). OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells. Stem Cells 26, 3068-3074.
25. Yamazaki, T., Liu, L., Lazarev, D., Al-Zain, A., Fomin, V., Yeung, P.L., Chambers, S.M., Lu, C.W., Studer, L., and Manley, J.L. (2018). TCF3 alternative splicing controlled by hnRNP H/F regulates E-cadherin expression and hESC pluripotency. Genes Dev 32, 1161-1174.
26. Yeo, G.W., Xu, X., Liang, T.Y., Muotri, A.R., Carson, C.T., Coufal, N.G., and Gage, F.H. (2007). Alternative splicing events identified in human embryonic stem cells and neural progenitors. PLoS Comput Biol 3, 1951-1967.
27. Farashahi Yazd, E., Rafiee, M.R., Soleimani, M., Tavallaei, M., Salmani, M.K., and Mowla, S.J. (2011). OCT4B1, a novel spliced variant of OCT4, generates a stable truncated protein with a potential role in stress response. Cancer Lett 309, 170-175.
28. Kalsotra, A., and Cooper, T.A. (2011). Functional consequences of developmentally regulated alternative splicing. Nat Rev Genet 12, 715-729.
29. Licatalosi, D.D., Yano, M., Fak, J.J., Mele, A., Grabinski, S.E., Zhang, C., and Darnell, R.B. (2012). Ptbp2 represses adult-specific splicing to regulate the generation of neuronal precursors in the embryonic brain. Genes Dev 26, 1626-1642.
30. Lou, C.H., Shao, A., Shum, E.Y., Espinoza, J.L., Huang, L., Karam, R., and Wilkinson, M.F. (2014). Posttranscriptional control of the stem cell and neurogenic programs by the nonsense-mediated RNA decay pathway. Cell Rep 6, 748-764.
31. 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.
32. Ogawa, S. (2014). Splicing factor mutations in AML. Blood 123, 3216-3217.
33. Yoshida, K., and Ogawa, S. (2014). Splicing factor mutations and cancer. Wiley Interdiscip Rev RNA 5, 445-459.
34. Zhang, H., Brown, R.L., Wei, Y., Zhao, P., Liu, S., Liu, X., Deng, Y., Hu, X., Zhang, J., Gao, X.D., et al. (2019). CD44 splice isoform switching determines breast cancer stem cell state. Genes Dev 33, 166-179.
35. Abrahamsson, A.E., Geron, I., Gotlib, J., Dao, K.H., Barroga, C.F., Newton, I.G., Giles, F.J., Durocher, J., Creusot, R.S., Karimi, M., et al. (2009). Glycogen synthase kinase 3beta missplicing contributes to leukemia stem cell generation. Proc Natl Acad Sci U S A 106, 3925-3929.
36. Goff, D.J., Court Recart, A., Sadarangani, A., Chun, H.J., Barrett, C.L., Krajewska, M., Leu, H., Low-Marchelli, J., Ma, W., Shih, A.Y., et al. (2013). A Pan-BCL2 inhibitor renders bone-marrow-resident human leukemia stem cells sensitive to tyrosine kinase inhibition. Cell Stem Cell 12, 316-328
37. Gregory, R.I., Yan, K.P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature 432, 235-240.
38. Hammond, S.M. (2005). Dicing and slicing: the core machinery of the RNA interference pathway. FEBS Lett 579, 5822-5829.
39. Gregory, R.I., Chendrimada, T.P., Cooch, N., and Shiekhattar, R. (2005). Human RISC couples microRNA biogenesis and posttranscriptional gene silencing. Cell 123, 631-640.
40. Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blelloch, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. Nat Genet 39, 380-385.
41. Wang, X.W., Hao, J., Guo, W.T., Liao, L.Q., Huang, S.Y., Guo, X., Bao, X., Esteban, M.A., and Wang, Y. (2017c). A DGCR8-Independent Stable MicroRNA Expression Strategy Reveals Important Functions of miR-290 and miR-183-182 Families in Mouse Embryonic Stem Cells. Stem Cell Reports 9, 1618-1629.
42. Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. Genes Dev 19, 489-501.

43. Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 455, 1124-1128.
44. Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J.A., and Kosik, K.S. (2009b). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell 137, 647-658.
45. Li, Z., Yang, C.S., Nakashima, K., and Rana, T.M. (2011). Small RNA-mediated regulation of iPS cell generation. Embo j 30, 823-834.
46. Lin, S.L., Chang, D.C., Chang-Lin, S., Lin, C.H., Wu, D.T., Chen, D.T., and Ying, S.Y. (2008). Mir-302 reprograms human skin cancer cells into a pluripotent ES-cell-like state. Rna 14, 2115-2124.
47. Georgantas, R.W., 3rd, Hildreth, R., Morisot, S., Alder, J., Liu, C.G., Heimfeld, S., Calin, G.A., Croce, C.M., and Civin, C.I. (2007). CD34+ hematopoietic stem-progenitor cell microRNA expression and function: a circuit diagram of differentiation control. Proc Natl Acad Sci U S A 104, 2750-2755.
48. de Morree, A., Klein, J.D.D., Gan, Q., Farup, J., Urtasun, A., Kanugovi, A., Bilen, B., van Velthoven, C.T.J., Quarta, M., and Rando, T.A. (2019). Alternative polyadenylation of Pax3 controls muscle stem cell fate and muscle function. Science 366, 734-738.
49. Chen, J.F., Mandel, E.M., Thomson, J.M., Wu, Q., Callis, T.E., Hammond, S.M., Conlon, F.L., and Wang, D.Z. (2006). The role of microRNA-1 and microRNA-133 in skeletal muscle proliferation and differentiation. Nat Genet 38, 228-233.
50. Smirnova, L., Grafe, A., Seiler, A., Schumacher, S., Nitsch, R., and Wulczyn, F.G. (2005). Regulation of miRNA expression during neural cell specification. Eur J Neurosci 21, 1469-1477.
51. Visvanathan, J., Lee, S., Lee, B., Lee, J.W., and Lee, S.K. (2007). The microRNA miR-124 antagonizes the anti-neural REST/SCP1 pathway during embryonic CNS development. Genes Dev 21, 744-749.
52. Foulkes, W.D., Priest, J.R., and Duchaine, T.F. (2014). DICER1: mutations, microRNAs and mechanisms. Nat Rev Cancer 14, 662-672.
53. Iliou, M.S., da Silva-Diz, V., Carmona, F.J., Ramalho-Carvalho, J., Heyn, H., Villanueva, A., Munoz, P., and Esteller, M. (2014). Impaired DICER1 function promotes stemness and metastasis in colon cancer. Oncogene 33, 4003-4015.
54. Wang, X.J., Jiang, F.Z., Tong, H., Ke, J.Q., Li, Y.R., Zhang, H.L., Yan, X.F., Wang, F.Y., and Wan, X.P. (2017b). Dicer1 dysfunction promotes stemness and aggression in endometrial carcinoma. Tumour Biol 39, 1010428317695967.
55. Calin, G.A., and Croce, C.M. (2006). MicroRNA signatures in human cancers. Nat Rev Cancer 6, 857-866.
56. Gaur, A., Jewell, D.A., Liang, Y., Ridzon, D., Moore, J.H., Chen, C., Ambros, V.R., and Israel, M.A. (2007). Characterization of microRNA expression levels and their biological correlates in human cancer cell lines. Cancer Res 67, 2456-2468.
57. Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., et al. (2005). MicroRNA expression profiles classify human cancers. Nature 435, 834-838.
58. Asadzadeh, Z., Mansoori, B., Mohammadi, A., Aghajani, M., Haji-Asgarzadeh, K., Safarzadeh, E., Mokhtarzadeh, A., Duijf, P.H.G., and Baradaran, B. (2019). microRNAs in cancer stem cells: Biology, pathways, and therapeutic opportunities. J Cell Physiol 234, 10002-10017.
59. Ma, W., Xiao, G.G., Mao, J., Lu, Y., Song, B., Wang, L., Fan, S., Fan, P., Hou, Z., Li, J., et al. (2015). Dysregulation of the miR-34a-SIRT1 axis inhibits breast cancer stemness. Oncotarget 6, 10432-10444.
60. Liu, C., Kelnar, K., Liu, B., Chen, X., Calhoun-Davis, T., Li, H., Patrawala, L., Yan, H., Jeter, C., Honorio, S., et al. (2011). The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. Nat Med 17, 211-215.
61. Dominissini, D., Moshitch-Moshkovitz, S., Salmon-Divon, M., Amariglio, N., and Rechavi, G. (2013). Transcriptome-wide mapping of N(6)-methyladenosine by m(6)A-seq based on immunocapturing and massively parallel sequencing. Nat Protoc 8, 176-189.
62. Balacco, D.L., and Soller, M. (2019). The m(6)A Writer: Rise of a Machine for Growing Tasks. Biochemistry 58, 363-378.
63. Bokar, J.A., Shambaugh, M.E., Polayes, D., Matera, A.G., and Rottman, F.M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. Rna 3, 1233-1247.
64. Liu, J., Yue, Y., Han, D., Wang, X., Fu, Y., Zhang, L., Jia, G., Yu, M., Lu, Z., Deng, X., et al. (2014). A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol 10, 93-95.

65. Yue, Y., Liu, J., Cui, X., Cao, J., Luo, G., Zhang, Z., Cheng, T., Gao, M., Shu, X., Ma, H., et al. (2018). VIRMA mediates preferential m(6)A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. Cell Discov 4, 10.

66. Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., Yang, Y., Yi, C., Lindahl, T., Pan, T., Yang, Y.G., et al. (2011). N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7, 885-887.

67. Zheng, G., Dahl, J.A., Niu, Y., Fedorcsak, P., Huang, C.M., Li, C.J., Vagbo, C.B., Shi, Y., Wang, W.L., Song, S.H., et al. (2013). ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell 49, 18-29.

68. Wang, X., Lu, Z., Gomez, A., Hon, G.C., Yue, Y., Han, D., Fu, Y., Parisien, M., Dai, Q., Jia, G., et al. (2014b). N6-methyladenosine-dependent regulation of messenger RNA stability. Nature 505, 117-120.

69. Wang, X., Zhao, B.S., Roundtree, I.A., Lu, Z., Han, D., Ma, H., Weng, X., Chen, K., Shi, H., and He, C. (2015b). N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. Cell 161, 1388-1399.

70. Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., Osenberg, S., Cesarkas, K., Jacob-Hirsch, J., Amariglio, N., Kupiec, M., et al. (2012). Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485, 201-206.

71. Geula, S., Moshitch-Moshkovitz, S., Dominissini, D., Mansour, A.A., Kol, N., Salmon-Divon, M., Hershkovitz, V., Peer, E., Mor, N., Manor, Y.S., et al. (2015). Stem cells. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. Science 347, 1002-1006.

72. Liu, N., Dai, Q., Zheng, G., He, C., Parisien, M., and Pan, T. (2015). N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. Nature 518, 560-564.

73. Peer, E., Moshitch-Moshkovitz, S., Rechavi, G., and Dominissini, D. (2018). The Epitranscriptome in Translation Regulation. Cold Spring Harb Perspect Biol.

74. Xiao, W., Adhikari, S., Dahal, U., Chen, Y.S., Hao, Y.J., Sun, B.F., Sun, H.Y., Li, A., Ping, X.L., Lai, W.Y., et al. (2016). Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. Mol Cell 61, 507-519.

75. Alarcon, C.R., Lee, H., Goodarzi, H., Halberg, N., and Tavazoie, S.F. (2015). N6-methyladenosine marks primary microRNAs for processing. Nature 519, 482-485.

76. Ping, X.L., Sun, B.F., Wang, L., Xiao, W., Yang, X., Wang, W.J., Adhikari, S., Shi, Y., Lv, Y., Chen, Y.S., et al. (2014). Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res 24, 177-189.

77. Boissel, S., Reish, O., Proulx, K., Kawagoe-Takaki, H., Sedgwick, B., Yeo, G.S., Meyre, D., Golzio, C., Molinari, F., Kadhom, N., et al. (2009). Loss-of-function mutation in the dioxygenase-encoding FTO gene causes severe growth retardation and multiple malformations. Am J Hum Genet 85, 106-111.

78. Batista, P.J., Molinie, B., Wang, J., Qu, K., Zhang, J., Li, L., Bouley, D.M., Lujan, E., Haddad, B., Daneshvar, K., et al. (2014). m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15, 707-719.

79. Wang, Y., Li, Y., Toth, J.I., Petroski, M.D., Zhang, Z., and Zhao, J.C. (2014c). N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat Cell Biol 16, 191-198.

80. Zhang, C., Chen, Y., Sun, B., Wang, L., Yang, Y., Ma, D., Lv, J., Heng, J., Ding, Y., Xue, Y., et al. (2017). m(6)A modulates haematopoietic stem and progenitor cell specification. Nature 549, 273-276.

81. Vu, L.P., Pickering, B.F., Cheng, Y., Zaccara, S., Nguyen, D., Minuesa, G., Chou, T., Chow, A., Saletore, Y., MacKay, M., et al. (2017). The N(6)-methyladenosine (m(6)A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and leukemia cells. Nat Med 23, 1369-1376.

82. Weng, H., Huang, H., Wu, H., Qin, X., Zhao, B.S., Dong, L., Shi, H., Skibbe, J., Shen, C., Hu, C., et al. (2018). METTL14 Inhibits Hematopoietic Stem/Progenitor Differentiation and Promotes Leukemogenesis via mRNA m(6)A Modification. Cell Stem Cell 22, 191-205.e199.

83. Lee, H., Bao, S., Qian, Y., Geula, S., Leslie, J., Zhang, C., Hanna, J.H., and Ding, L. (2019). Stage-specific requirement for Mettl3-dependent m(6)A mRNA methylation during haematopoietic stem cell differentiation. Nat Cell Biol 21, 700-709.

84. Wu, Y., Xie, L., Wang, M., Xiong, Q., Guo, Y., Liang, Y., Li, J., Sheng, R., Deng, P., Wang, Y., et al. (2018). Mettl3-mediated m(6)A RNA methylation regulates the fate of bone marrow mesenchymal stem cells and osteoporosis. Nat Commun 9, 4772.

85. Zhang, C., Zhi, W.I., Lu, H., Samanta, D., Chen, I., Gabrielson, E., and Semenza, G.L. (2016a). Hypoxia-inducible factors regulate pluripotency factor expression by ZNF217- and ALKBH5-mediated modulation of RNA methylation in breast cancer cells. Oncotarget 7, 64527-64542.

86. Li, Z., Weng, H., Su, R., Weng, X., Zuo, Z., Li, C., Huang, H., Nachtergaele, S., Dong, L., Hu, C., et al. (2017b). FTO Plays an Oncogenic Role in Acute Myeloid Leukemia as a N(6)-Methyladenosine RNA Demethylase. Cancer Cell 31, 127-141.

87. Wang, X., Li, Z., Kong, B., Song, C., Cong, J., Hou, J., and Wang, S. (2017a). Reduced m(6)A mRNA methylation is correlated with the progression of human cervical cancer. Oncotarget 8, 98918-98930.

88. Barbieri, I., Tzelepis, K., Pandolfini, L., Shi, J., Millan-Zambrano, G., Robson, S.C., Aspris, D., Migliori, V., Bannister, A.J., Han, N., et al. (2017). Promoter-bound METTL3 maintains myeloid leukaemia by m(6)A-dependent translation control. Nature 552, 126-131.

89. Dominissini, D., Nachtergaele, S., Moshitch-Moshkovitz, S., Peer, E., Kol, N., Ben-Haim, M.S., Dai, Q., Di Segni, A., Salmon-Divon, M., Clark, W.C., et al. (2016). The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature 530, 441-446.

90. Li, X., Xiong, X., Zhang, M., Wang, K., Chen, Y., Zhou, J., Mao, Y., Lv, J., Yi, D., Chen, X.W., et al. (2017a). Base-Resolution Mapping Reveals Distinct m(1)A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. Mol Cell 68, 993-1005.e1009.

91. Safra, M., Sas-Chen, A., Nir, R., Winkler, R., Nachshon, A., Bar-Yaacov, D., Erlacher, M., Rossmanith, W., Stern-Ginossar, N., and Schwartz, S. (2017). The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature 551, 251-255.

92. Zhang, C., and Jia, G. (2018). Reversible RNA Modification N(1)-methyladenosine (m(1)A) in mRNA and tRNA. Genomics Proteomics Bioinformatics 16, 155-161.

93. Schaefer, M., Pollex, T., Hanna, K., and Lyko, F. (2009). RNA cytosine methylation analysis by bisulfite sequencing. Nucleic Acids Res 37, e12.

94. Agris, P.F. (2008). Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications. EMBO Rep 9, 629-635.

95. Khoddami, V., and Cairns, B.R. (2013). Identification of direct targets and modified bases of RNA cytosine methyltransferases. Nat Biotechnol 31, 458-464.

96. Squires, J.E., Patel, H.R., Nousch, M., Sibbritt, T., Humphreys, D.T., Parker, B.J., Suter, C.M., and Preiss, T. (2012). Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. Nucleic Acids Res 40, 5023-5033.

97. Yang, X., Yang, Y., Sun, B.F., Chen, Y.S., Xu, J.W., Lai, W.Y., Li, A., Wang, X., Bhattarai, D.P., Xiao, W., et al. (2017). 5-methylcytosine promotes mRNA export - NSUN2 as the methyltransferase and ALYREF as an m(5)C reader. Cell Res 27, 606-625.

98. Blanco, S., Bandiera, R., Popis, M., Hussain, S., Lombard, P., Aleksic, J., Sajini, A., Tanna, H., Cortes-Garrido, R., Gkatza, N., et al. (2016). Stem cell function and stress response are controlled by protein synthesis. Nature 534, 335-340.

99. Gigova, A., Duggimpudi, S., Pollex, T., Schaefer, M., and Kos, M. (2014). A cluster of methylations in the domain IV of 25S rRNA is required for ribosome stability. Rna 20, 1632-1644.

100. Schosserer, M., Minois, N., Angerer, T.B., Amring, M., Dellago, H., Harreither, E., Calle-Perez, A., Pircher, A., Gerstl, M.P., Pfeifenberger, S., et al. (2015). Methylation of ribosomal RNA by NSUN5 is a conserved mechanism modulating organismal lifespan. Nat Commun 6, 6158.

101. Sharma, S., Yang, J., Watzinger, P., Kotter, P., and Entian, K.D. (2013). Yeast Nop2 and Rcm1 methylate C2870 and C2278 of the 25S rRNA, respectively. Nucleic Acids Res 41, 9062-9076.

102. Tuorto, F., Herbst, F., Alerasool, N., Bender, S., Popp, O., Federico, G., Reitter, S., Liebers, R., Stoecklin, G., Grone, H.J., et al. (2015). The tRNA methyltransferase Dnmt2 is required for accurate polypeptide synthesis during haematopoiesis. Embo j 34, 2350-2362.

103. Mellis, I.A., Gupte, R., Raj, A., and Rouhanifard, S.H. (2017). Visualizing adenosine-to-inosine RNA editing in single mammalian cells. Nat Methods 14, 801-804.

104. Chan, C.T., Dyavaiah, M., DeMott, M.S., Taghizadeh, K., Dedon, P.C., and Begley, T.J. (2010). A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. PLoS Genet 6, e1001247.

105. Flores, J.V., Cordero-Espinoza, L., Oeztuerk-Winder, F., Andersson-Rolf, A., Selmi, T., Blanco, S., Tailor, J., Dietmann, S., and Frye, M. (2017). Cytosine-5 RNA Methylation Regulates Neural Stem Cell Differentiation and Motility. Stem Cell Reports 8, 112-124.

106. Blanco, S., Dietmann, S., Flores, J.V., Hussain, S., Kutter, C., Humphreys, P., Lukk, M., Lombard, P., Treps, L., Popis, M., et al. (2014). Aberrant methylation of tRNAs links cellular stress to neuro-developmental disorders. Embo j 33, 2020-2039.

107. Hussain, S., Tuorto, F., Menon, S., Blanco, S., Cox, C., Flores, J.V., Watt, S., Kudo, N.R., Lyko, F., and Frye, M. (2013). The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation. Mol Cell Biol 33, 1561-1570.

108. Blanco, S., Kurowski, A., Nichols, J., Watt, F.M., Benitah, S.A., and Frye, M. (2011). The RNA-methyltransferase Misu (NSun2) poises epidermal stem cells to differentiate. PLoS Genet 7, e1002403.

109. Rai, K., Chidester, S., Zavala, C.V., Manos, E.J., James, S.R., Karpf, A.R., Jones, D.A., and Cairns, B.R. (2007). Dnmt2 functions in the cytoplasm to promote liver, brain, and retina development in zebrafish. Genes Dev 21, 261-266.

110. Charette, M., and Gray, M.W. (2000). Pseudouridine in RNA: what, where, how, and why. IUBMB Life 49, 341-351.

111. Guzzi, N., Ciesla, M., Ngoc, P.C.T., Lang, S., Arora, S., Dimitriou, M., Pimkova, K., Sommarin, M.N.E., Munita, R., Lubas, M., et al. (2018). Pseudouridylation of tRNA-Derived Fragments Steers Translational Control in Stem Cells. Cell 173, 1204-1216.e1226.

112. De Zoysa, M.D., and Yu, Y.T. (2017). Posttranscriptional RNA Pseudouridylation. Enzymes 41, 151-167.

113. Roundtree, I.A., Evans, M.E., Pan, T., and He, C. (2017). Dynamic RNA Modifications in Gene Expression Regulation. Cell 169, 1187-1200.

114. Hamma, T., and Ferre-D'Amare, A.R. (2006). Pseudouridine synthases. Chem Biol 13, 1125-1135.

115. Batista, L.F., Pech, M.F., Zhong, F.L., Nguyen, H.N., Xie, K.T., Zaug, A.J., Crary, S.M., Choi, J., Sebastiano, V., Cherry, A., et al. (2011). Telomere shortening and loss of self-renewal in dyskeratosis congenita induced pluripotent stem cells. Nature 474, 399-402.

116. Fong, Y.W., Ho, J.J., Inouye, C., and Tjian, R. (2014). The dyskerin ribonucleoprotein complex as an OCT4/SOX2 coactivator in embryonic stem cells. Elife 3.

117. Agarwal, S., Loh, Y.H., McLoughlin, E.M., Huang, J., Park, I.H., Miller, J.D., Huo, H., Okuka, M., Dos Reis, R.M., Loewer, S., et al. (2010). Telomere elongation in induced pluripotent stem cells from dyskeratosis congenita patients. Nature 464, 292-296.

118. Bellodi, C., McMahon, M., Contreras, A., Juliano, D., Kopmar, N., Nakamura, T., Maltby, D., Burlingame, A., Savage, S.A., Shimamura, A., et al. (2013). H/ACA small RNA dysfunctions in disease reveal key roles for noncoding RNA modifications in hematopoietic stem cell differentiation. Cell Rep 3, 1493-1502.

119. Peng, Z., Cheng, Y., Tan, B.C., Kang, L., Tian, Z., Zhu, Y., Zhang, W., Liang, Y., Hu, X., Tan, X., et al. (2012). Comprehensive analysis of RNA-Seq data reveals extensive RNA editing in a human transcriptome. Nat Biotechnol 30, 253-260.

120. Hartner, J.C., Walkley, C.R., Lu, J., and Orkin, S.H. (2009). ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling. Nat Immunol 10, 109-115.

121. Tan, M.H., Li, Q., Shanmugam, R., Piskol, R., Kohler, J., Young, A.N., Liu, K.I., Zhang, R., Ramaswami, G., Ariyoshi, K., et al. (2017). Dynamic landscape and regulation of RNA editing in mammals. Nature 550, 249-254.

122. Zipeto, M.A., Court, A.C., Sadarangani, A., Delos Santos, N.P., Balaian, L., Chun, H.J., Pineda, G., Morris, S.R., Mason, C.N., Geron, I., et al. (2016). ADAR1 Activation Drives Leukemia Stem Cell Self-Renewal by Impairing Let-7 Biogenesis. Cell Stem Cell 19, 177-191.

123. Zipeto, M.A., Jiang, Q., Melese, E., and Jamieson, C.H. (2015). RNA rewriting, recoding, and rewiring in human disease. Trends Mol Med 21, 549-559.

124. Batzer, M.A., and Deininger, P.L. (2002). Alu repeats and human genomic diversity. Nat Rev Genet 3, 370-379.

125. Hsiao, Y.E., Bahn, J.H., Yang, Y., Lin, X., Tran, S., Yang, E.W., Quinones-Valdez, G., and Xiao, X. (2018). RNA editing in nascent RNA affects pre-mRNA splicing. Genome Res 28, 812-823.

126. Rueter, S.M., Dawson, T.R., and Emeson, R.B. (1999). Regulation of alternative splicing by RNA editing. Nature 399, 75-80.

127. Solomon, O., Oren, S., Safran, M., Deshet-Unger, N., Akiva, P., Jacob-Hirsch, J., Cesarkas, K., Kabesa, R., Amariglio, N., Unger, R., et al. (2013). Global regulation of alternative splicing by adenosine deaminase acting on RNA (ADAR). Rna 19, 591-604.

128. Jiang, Q., Isquith, J., Zipeto, M.A., Diep, R.H., Pham, J., Delos Santos, N., Reynoso, E., Chau, J., Leu, H., Lazzari, E., et al. (2019). Hyper-Editing of Cell-Cycle Regulatory and Tumor Suppressor RNA Promotes Malignant Progenitor Propagation. Cancer Cell 35, 81-94.e87.

129. Rice, G.I., Kasher, P.R., Forte, G.M., Mannion, N.M., Greenwood, S.M., Szynkiewicz, M., Dickerson, J.E., Bhaskar, S.S., Zampini, M., Briggs, T.A., et al. (2012). Mutations in ADAR1 cause Aicardi- Goutières syndrome associated with a type I interferon signature. Nat Genet 44, 1243-1248.

130. Suzuki, N., Suzuki, T., Inagaki, K., Ito, S., Kono, M., Fukai, K., Takama, H., Sato, K., Ishikawa, O., Abe, M., et al. (2005). Mutation analysis of the ADAR1 gene in dyschromatosis symmetrica hereditaria and genetic differentiation from both dyschromatosis universalis hereditaria and acropigmentatio reticularis. J Invest Dermatol 124, 1186-1192.

131. Jacobs, M.M., Fogg, R.L., Emeson, R.B., and Stanwood, G.D. (2009). ADAR1 and ADAR2 expression and editing activity during forebrain development. Dev Neurosci 31, 223-237.

132. Higuchi, M., Maas, S., Single, F.N., Hartner, J., Rozov, A., Burnashev, N., Feldmeyer, D., Sprengel, R., and Seeburg, P.H. (2000). Point mutation in an AMPA receptor gene rescues lethality in mice deficient in the RNA-editing enzyme ADAR2. Nature 406, 78-81.

133. Yamashita, T., and Kwak, S. (2019). Cell death cascade and molecular therapy in ADAR2-deficient motor neurons of ALS. Neurosci Res 144, 4-13.

134. Slotkin, W., and Nishikura, K. (2013). Adenosine-to-inosine RNA editing and human disease. Genome Med 5, 105.

135. Whitney, N.P., Peng, H., Erdmann, N.B., Tian, C., Monaghan, D.T., and Zheng, J.C. (2008). Calcium-permeable AMPA receptors containing Q/R-unedited GluR2 direct human neural progenitor cell differentiation to neurons. Faseb j 22, 2888-2900.

136. Oakes, E., Anderson, A., Cohen-Gadol, A., and Hundley, H.A. (2017). Adenosine Deaminase That Acts on RNA 3 (ADAR3) Binding to Glutamate Receptor Subunit B Pre-mRNA Inhibits RNA Editing in Glioblastoma. J Biol Chem 292, 4326-4335.

137. Osenberg, S., Dominissini, D., Rechavi, G., and Eisenberg, E. (2009). Widespread cleavage of A-to-I hyperedited substrates. Rna 15, 1632-1639.

138. Chung, H., Calis, J.J.A., Wu, X., Sun, T., Yu, Y., Sarbanes, S.L., Dao Thi, V.L., Shilvock, A.R., Hoffmann, H.H., Rosenberg, B.R., et al. (2018). Human ADAR1 Prevents Endogenous RNA from Triggering Translational Shutdown. Cell 172, 811-824.e814.

139. Shtrichman, R., Germanguz, I., and Itskovitz-Eldor, J. (2013). Induced pluripotent stem cells (iPSCs) derived from different cell sources and their potential for regenerative and personalized medicine. Curr Mol Med 13, 792-805.

140. Germanguz, I., Shtrichman, R., Osenberg, S., Ziskind, A., Novak, A., Domev, H., Laevsky, I., Jacob-Hirsch, J., Feiler, Y., Rechavi, G., et al. (2014). ADAR1 is involved in the regulation of reprogramming human fibroblasts to induced pluripotent stem cells. Stem Cells Dev 23, 443-456.

141. Orkin, S.H., and Zon, L.I. (2008). Hematopoiesis: an evolving paradigm for stem cell biology. Cell 132, 631-644.

142. XuFeng, R., Boyer, M.J., Shen, H., Li, Y., Yu, H., Gao, Y., Yang, Q., Wang, Q., and Cheng, T. (2009). ADAR1 is required for hematopoietic progenitor cell survival via RNA editing. Proc Natl Acad Sci U S A 106, 17763-17768.

143. Qiu, W., Wang, X., Buchanan, M., He, K., Sharma, R., Zhang, L., Wang, Q., and Yu, J. (2013). ADAR1 is essential for intestinal homeostasis and stem cell maintenance. Cell Death Dis 4, e599.

144. Crews, L.A., Jiang, Q., Zipeto, M.A., Lazzari, E., Court, A.C., Ali, S., Barrett, C.L., Frazer, K.A., and Jamieson, C.H. (2015). An RNA editing fingerprint of cancer stem cell reprogramming. J Transl Med 13, 52.

145. Beghini, A., Ripamonti, C.B., Peterlongo, P., Roversi, G., Cairoli, R., Morra, E., and Larizza, L. (2000). RNA hyperediting and alternative splicing of hematopoietic cell phosphatase (PTPN6) gene in acute myeloid leukemia. Hum Mol Genet 9, 2297-2304.

146. Ishizuka, J.J., Manguso, R.T., Cheruiyot, C.K., Bi, K., Panda, A., Iracheta-Vellve, A., Miller, B.C., Du, P.P., Yates, K.B., Dubrot, J., et al. (2019). Loss of ADAR1 in tumours overcomes resistance to immune checkpoint blockade. Nature 565, 43-48.

147. Xu, Y., Poggio, M., Jin, H.Y., Shi, Z., Forester, C.M., Wang, Y., Stumpf, C.R., Xue, L., Devericks, E., So, L., et al. (2019). Translation control of the immune checkpoint in cancer and its therapeutic targeting. Nat Med 25, 301-311.

148. Solomon, O., Di Segni, A., Cesarkas, K., Porath, H.T., Marcu-Malina, V., Mizrahi, O., Stern-Ginossar, N., Kol, N., Farage-Barhom, S., Glick-Saar, E., et al. (2017). RNA editing by ADAR1 leads to context-dependent transcriptome-wide changes in RNA secondary structure. Nat Commun 8, 1440.

149. Chua BA, Van Der Werf I, Jamieson C, Signer RAJ. Post-Transcriptional Regulation of Homeostatic, Stressed, and Malignant Stem Cells. Cell Stem Cell. 2020;26(2):138-159.

CHAPTER 3

Splicing Factor Gene Mutations in Acute Myeloid Leukemia Offer Additive Value If Incorporated in Current Risk Classification

Inge van der Werf[1*]
Anna Wojtuszkiewicz[1*]
Manja Meggendorfer[2]
Stephan Hutter[2]
Constance Baer[2]
Martijn Heymans[3]
Peter J. M. Valk[4]
Wolfgang Kern[2]
Claudia Haferlach[2]
Jeroen Janssen[1]
Gert J. Ossenkoppele[1]
Jacqueline Cloos[1#]
Torsten Haferlach[2#]

1. Dept. of Hematology, Amsterdam University Medical Center, VU University Medical Center, Cancer Center Amsterdam, Amsterdam, The Netherlands; 2. MLL Munich Leukemia Laboratory, Munich, Germany; 3. Dept. of Epidemiology & Biostatistics, Amsterdam University Medical Center, VU University Medical Center, Cancer Center Amsterdam; 4. Department of Hematology, Erasmus University Medical Center Rotterdam.

*These authors contributed equally to this work
#Shared senior authorship

Blood Advances. 2021. 14;5(17):3254-3265

INTRODUCTION

Recurrent mutations in genes regulating splicing (splicing factors [SF]) were first discovered in hematological malignancies.[1–3] The most commonly mutated genes from this novel class include Splicing Factor 3B Subunit 1 (*SF3B1*), Serine and Arginine Rich Splicing Factor 2 (*SRSF2*), U2 small nuclear RNA Auxiliary Factor 1 (*U2AF1*) and Zinc Finger, CCCH type, RNA-binding Motif and Serine and Arginine Rich 2 (*ZRSR2*), all of which are thought to act during the early stages of spliceosome assembly.[4] Several studies documented the contribution of SF mutations to pathogenesis of myeloid malignancies. Mutations in *SRSF2*, *SF3B1* and *U2AF1* were demonstrated to result in widespread changes in the transcriptome accompanied by altered hematopoiesis[5] and they are generally considered to be early leukemogenic events.[1,5]

Interestingly, the prognostic impact of SF mutations was shown to differ between hematological malignancies.[6] Thus far, the association of *SF3B1* mutations with a better prognosis in the context of low-risk MDS is already well established.[2] In 2016, an updated version of the World Health Organization (WHO) classification for myeloid neoplasms and acute leukemia incorporated *SF3B1* mutations as a diagnostic criterion for MDS with ring sideroblasts.[7] In contrast, two recent whole genome investigations as well as a small study focused on SF mutations suggest that SF mutations are associated with inferior treatment outcome in AML and, in particular *SRSF2* mutations, could be considered for incorporation into prognostic guidelines.[8-10] However, an in-depth study focused specifically on SF mutations in a large well-annotated cohort is currently lacking.

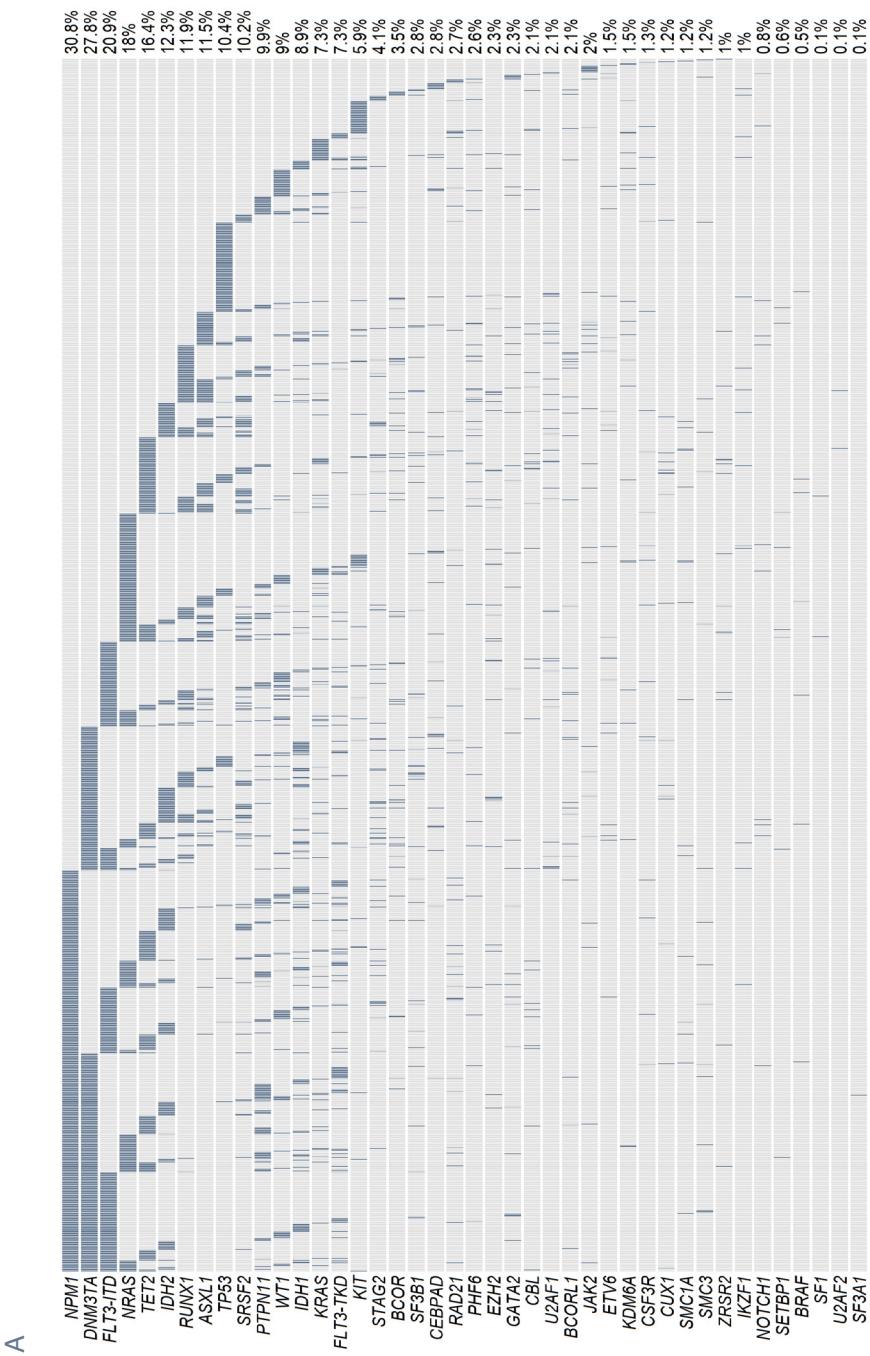
To date, the most widely accepted classification and prognostic schemes for AML include cytogenetic lesions together with multiple genetic mutations, including those in *NPM1*, *FLT3* and *CEBPA*.[7] Importantly, European Leukemia Net (ELN) endorsed mutations in *TP53*, *RUNX1* and *ASXL1* as adverse risk factors, while AML with mutations in *RUNX1* are considered a new provisional entity according to WHO.[11] Interestingly, mutations in both *RUNX1* and *ASXL1* were previously shown to co-occur with SF mutations.[8,9] The high frequency of SF mutations in AML together with the pattern of their co-occurrence with *RUNX1* and *ASXL1* mutations suggest that the prognostic impact of each of these factors could be examined in more detail.

Here, we report a comprehensive study of mutations in SF genes in a large well-annotated cohort of AML patients (N=1447). This cohort provides the unique opportunity to explore the significance of mutations in SF genes individually and cumulatively as an entity as well as to investigate clinical and biological features of AML with SF mutations. In the present study, we assessed the frequency, genetic background and prognostic value of SF mutations. To the best of our knowledge this is the largest study which specifically focusses on this emerging class of molecular aberrations.

Table 1. Characteristics of patients treated with intensive chemotherapy.

	Total Cohort	SF Mutated Patients
Age (y)		
Median	54	60
Range	17.8 - 85.9	18.0 - 83.8
Gender, n (%)		
Male	664 (54.3)	109 (68.1)
Female	559 (45.7)	51 (31.9)
Entity, n (%)		
De Novo AML	1047 (85.6)	126 (78.7)
Secondary AML	61 (5.0)	16 (10.0)
MDS	72 (5.9)	15 (9.4)
Treatment-related AML	43 (3.5)	3 (1.9)
WBC Count, x10^9/L, n (%)		
≤ 100	1047 (85.6)	142 (88.7)
> 100	110 (9.0)	10 (6.3)
Missing	66 (5.4)	8 (5.0)
ELN Classification, n (%)		
Favorable	483 (39.4)	33 (20.6)
Intermediate	337 (27.6)	33 (20.6)
Adverse	403 (33.0)	94 (58.8)
Induction Response, n (%)		
Complete Remission	1045 (85.4)	122
Partial Response	35 (2.9)	8 (5.0)
Early Death	65 (5.3)	17 (10.6)
Refractory Disease	62 (5.1)	13 (8.1)
Missing	16 (1.3)	-
Disease Course, n (%)		
Relapse	411 (33.6)	52 (32.5)
Death	629 (51.4)	107 (66.9)
Stem Cell Transplantation, n (%)		
None	541 (44.2)	90 (56.3)
Autologous Transplantation	122 (10.0)	9 (5.6)
Allogenic Transplantation	560 (45.8)	61 (38.1)

AML, Acute Myeloid Leukemia; MDS, Myelodysplastic Syndrome; ELN, European Leukemia Net (2017 classification); WBC – white blood cell count.


A
NPM1 30.8%
DNM3TA 27.8%
FLT3-ITD 20.9%
NRAS 18%
TET2 16.4%
IDH2 12.3%
RUNX1 11.9%
ASXL1 11.5%
TP53 10.4%
SRSF2 10.2%
PTPN11 9.9%
WT1 9%
IDH1 8.9%
KRAS 7.3%
FLT3-TKD 7.3%
KIT 5.9%
STAG2 4.1%
BCOR 3.5%
SF3B1 2.8%
CEBPAD 2.8%
RAD21 2.7%
PHF6 2.6%
EZH2 2.3%
GATA2 2.3%
CBL 2.1%
U2AF1 2.1%
BCORL1 2.1%
JAK2 2%
ETV6 1.5%
KDM6A 1.5%
CSF3R 1.3%
CUX1 1.2%
SMC1A 1.2%
SMC3 1.2%
ZRSR2 1%
IKZF1 1%
NOTCH1 0.8%
SETBP1 0.6%
BRAF 0.5%
SF1 0.1%
U2AF2 0.1%
SF3A1 0.1%

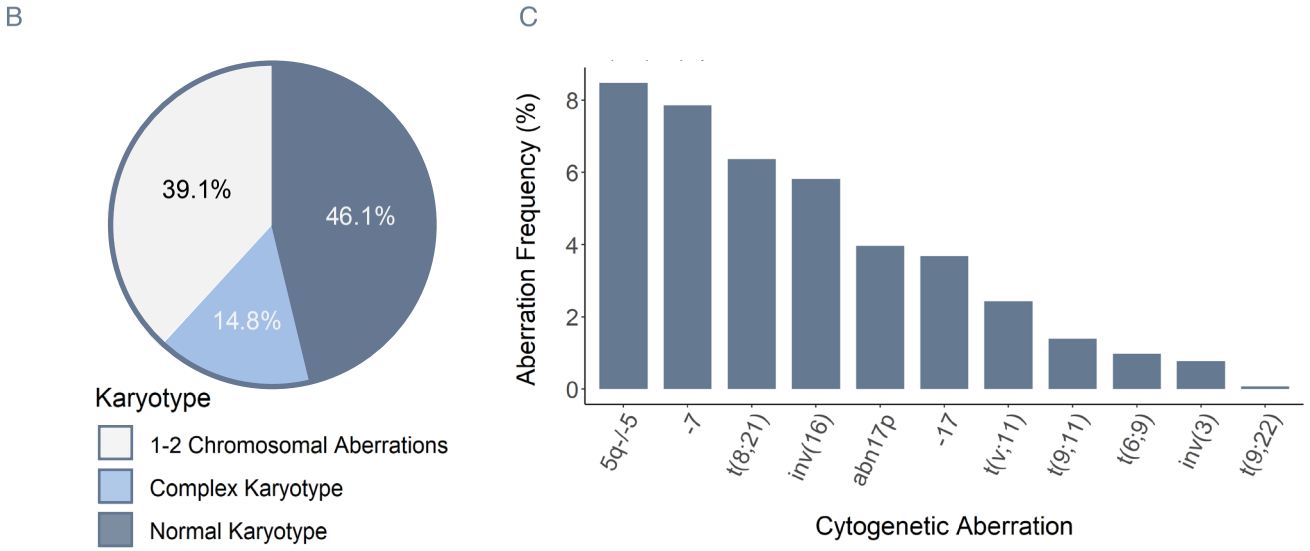


Figure 1. Genetic make-up of the total AML patient cohort. A - Frequency of recurrent mutations. B, C - Proportion of patients depending on the number of karyotypic aberrations (B) and frequency of recurring cytogenetic abnormalities (C). The assignment of patients to particular cytogenetic subgroups was not hierarchical (therefore a single patient could be assigned to multiple subgroups when appropriate, i.e. when more than one cytogenetic aberration was found).

METHODS

PATIENTS

This study included a total of 1447 samples taken at the time of diagnosis from AML/RAEB patients, including 1253 *de novo* AML, 70 secondary AML (sAML), 72 MDS (RAEB) and 52 therapy-related AML (tAML) patients. The patients were either included in clinical trials of the Dutch-Belgian Cooperative Trial Group for Hematology-Oncology (HOVON, N=889) and treated according to one of the 3 clinical protocols: HOVON-42A (N=133), -92 (N=43) or -102 (N=713),[12,13] or treated according to standard protocols in Germany (collected by the Munich Leukemia Laboratory, MLL between 2005 - 2016, N=558). Despite the heterogenous origin of our cohort no incremental differences were found in the genetic landscape and no significant differences in overall survival were found between HOVON and intensively treated MLL patients (Figure S9A). More details regarding treatment protocols and study design can be found in the Supplemental files. In total, the cohort included 1223 patients treated with intensive chemotherapy (1047 de novo AML, 61 sAML, 72 MDS and 43 tAML patients; Table 1, Table S1). All patients have signed a written informed consent. The study was approved by the internal review board of the MLL and local ethics committee of Amsterdam UMC and was conducted in accordance with the Declaration of Helsinki.

GENETIC PROFILE

The mutational profile of each patient from the HOVON dataset was defined based on molecular diagnostics as described previously.[12,13] In addition, for 430 patients (all of whom reached CR) additional data on mutations in 54 genes based on targeted sequencing using Illumina TruSight Myeloid Panel (Illumina, San Diego, CA, USA) were previously generated as described by Jongen-Lavrencic et al.[14] The mutational profiles of patients from the MLL dataset were based on routine molecular diagnostics (including a combination of gene scan analysis, melting curve analysis, Sanger sequencing and next-generation amplicon sequencing as described previously)[15-18] complemented by whole genome sequencing (Supplemental Methods). The cytogenetics were determined as described previously for patients included in the HOVON trials[12,13] and in the MLL dataset following ISCN guidelines (2013).[16,19,20]

STATISTICAL ANALYSIS

The analyses of associations between SF mutations and other genetic abnormalities as well as categorical clinical and biological variables were performed using Fisher's exact test with Benjamini-Hochberg correction for multiple testing. Associations of continuous variables with the presence of SF mutations were assessed using Mann-Whitney U test. Survival analyses were performed in a subset of AML patients treated with an intensive chemotherapy regimen (N = 1223). Samples for which data on the mutational status for a particular SF was missing were excluded from the analyses (for SFmut4: patients for whom data was missing for one of the four major SF factors were excluded from the analysis). These analyses included univariable testing (Kaplan-Meier analysis with log-rank test) as well as multivariable Cox proportional hazards model of the association of SF mutations with primary endpoints: overall survival (OS) and event-free survival (EFS). Further details of all statistical analyses used can be found in the Supplemental Methods.

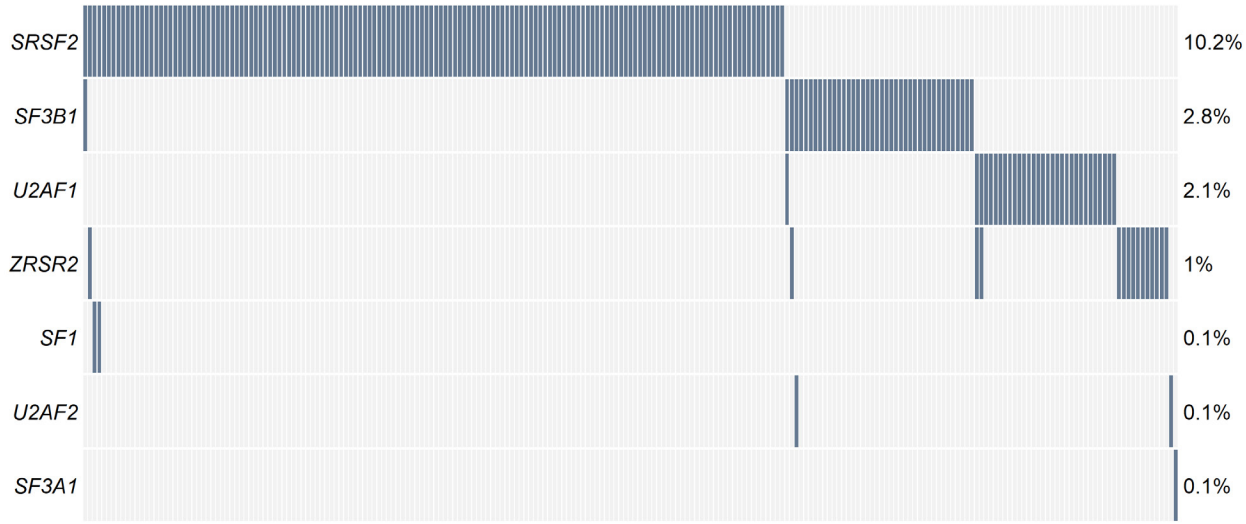


Figure 2. The frequency of SF mutations in the total AML cohort. The figure depicts the percentage of patients with SF mutations among the total number of AML cases in this study (N = 1447). For number of cases evaluated for each gene see Table S2. Cumulatively, SF mutations were identified in 22% of evaluated cases, including 148 individuals with a *SRSF2* mutation, 41 patients with a *SF3B1* mutation, 31 individuals with an *U2AF1* mutations and 15 patients with a *IntermediateAdverseFavorable* mutations. Furthermore, two cases of *SF1* mutation and a single case of *SF3A1* and *U2AF2* mutation were found.

RESULTS

RECURRENT SPLICING FACTOR MUTATIONS IN AML

Recurrent SF mutations were assessed in a total of 1447 patients, including data gathered in multicenter clinical trials of the HOVON group (N=889) as well as data collected by the MLL (N=558). While all three HOVON studies included patients aged 18 to 65, the MLL dataset contains data of patients 18 years and older without an upper age restriction. The combination of the different cohorts of AML patients, resulted in a large and potentially heterogeneous population of interest. However, comprehensive analysis of genetic and clinical features as well as treatment outcome revealed comparability of characteristics that allowed combined analyses (Study Design; Supplemental Files). Accordingly, the genetic landscape of our cohort was typical of AML with the frequencies of cytogenetic and molecular aberrations being consistent with those reported in previous studies (Figure 1).[8,9]

The data regarding the mutational status of seven splicing regulators was available for a variable number of cases: *SF3B1* (N=1447), *SRSF2* (N=1447), *U2AF1* (N=988), *ZRSR2* (N=988), *SF1*, *SF3A1* and *U2AF2* (each N=558). The four most common SF mutations (*SRSF2*, *SF3B1*, *U2AF1* and *ZRSR2*) were identified in a total of 22% of all the assessed cases (Figure 2, see Table S2 for percentages of mutated cases per gene and Table S3 for mutated cases divided per disease type). The majority of identified mutations were hotspot mutations (Table S4). Although SF mutations were mostly mutually exclusive, nine out of 231 patients had coinciding mutations in two different SF genes at diagnosis (Table S5 and Supplemental Results).

THE GENETIC LANDSCAPE AND CLINICAL FEATURES OF SF MUTATED AML

To gain more insight into the characteristics of AML patients carrying SF mutations, we studied their associations with other recurrent gene mutations, cytogenetics and clinical characteristics (Figure 3). Mutations in any single SF gene are relatively infrequent and patients with mutations in SFs have not been previously elaborately studied as a subgroup. Therefore, associations were examined for mutations in each individual SF gene separately, as well as collectively. The analysis combining multiple SFs included two variables: mutations in at least one of the four most commonly mutated SF (*SF3B1*, *SRSF2*, *U2AF1* and *ZRSR2*; SFmut4) and mutations in at least one of all the evaluated SF genes, including sporadic mutations in *SF1*, *SF3A1* and *U2AF2* (SFmut7). Since the latter variable included a limited number of cases (as the sporadic SF mutations were only assessed in 558 patients), we primarily focused our analyses on SFmut4.

When associations with other commonly mutated genes were investigated, SF mutations frequently coincided with *RUNX1*, *ASXL1*, *IDH2* and *TET2* mutations and were mutually exclusive with *KIT* and *NPM1* mutations as well as *FLT3-ITD* (Figure 3A). Furthermore, while we did not find any cytogenetic aberrations or chromosomal abnormalities co-occurring with SF mutations, we did find the core-binding factor aberrations and del5q (regardless of other molecular features), to be mutually exclusive with SF mutations (Figure 3B).

With respect to clinical characteristics, SFmut4 was correlated to older age (median 64.5 vs

56.0 years, p<0.001) and lower white-blood-cell counts (WBC, median; range: 10.8; 0.3-264.3 vs 16.3; 0.3-351.0 x10^9/L, p<0.001), both of which are common features of AML with antecedent myeloid disorders (Figure 3C). Accordingly, SFmut4 was more common in the sAML/MDS/tAML subgroup (34.2% of cases compared to 20.5% of *de novo* AML samples). Furthermore, eighty-three percent of all SF mutations affected patients classified as adverse or intermediate risk according to the ELN 2017 criteria. We also examined the relation between SF mutations and differentiation status of the leukemia cells as reflected by the French-American-British (FAB) classification, even though it is no longer used in risk stratification. We observed a strong association between SFmut4 and M0 (immature) FAB subtype, which indicates undifferentiated state of cells carrying these mutations. Finally, SF mutations were associated with male gender (SFmut4: 28.8% of males' vs 14.3% of females, Figure 3C). The majority of the associations with clinical characteristics, cytogenetics and recurrent gene mutations held true when examined in *de novo* AML excluding all sAML/MDS/tAML cases (Table S6).

SF MUTATIONS ARE ASSOCIATED WITH INFERIOR SURVIVAL

All analyses of OS and EFS were performed in a subset of the cohort including patients treated with an intensive chemotherapy regimen only (N=1223, Table 1, Table S1 also see methods section). Inferior survival was observed when the relation of *SF3B1*, *SRSF2*, *U2AF1* and *ZRSR2* mutations with OS and EFS was analyzed individually for each gene in the total cohort (Figure S1 and S2). In addition, we found that SFmut4 was associated with shorter EFS as well as OS in Kaplan-Meier analysis (Figure 4A). The same associations were apparent within *de novo* AML as well as sAML/MDS/tAML patient subset (data not shown). Furthermore, comparable trends were observed when the association of SFmut4 with OS and EFS was analyzed in separate ELN 2017 risk groups (Figure S3). Notably, mutations in *SF3B1* were related to significantly inferior EFS and OS in the intermediate risk group while borderline significant for OS (P=0.058) in favorable risk patients (Figure 4B-C, Table S7). Survival rates of *SF3B1* mutated patients in the intermediate risk group were comparable to that of adverse risk patients (5-year OS: 20.0% vs 23.8% respectively; 5-year EFS: 10.0% vs 17.0% respectively). In the favorable risk group 5-year OS of SF3B1 mutated patients was similar to adverse risk patients (28.6% vs 23.8% respectively), while 5-year EFS was resembling that of intermediate risk patients (28.6% vs 36.4% respectively). *SRSF2* mutations revealed a slightly shorter EFS only within adverse risk group (Figure S3B). *SRSF2* mutated patients were classified into this subgroup predominantly due to co-occurrence with *RUNX1* or *ASXL1* mutations and had in majority normal karyotype (data not shown). This suggests that co-occurrence of *SRSF2* mutations with *RUNX1* or *ASXL1* mutations could be associated with particularly adverse outcome.

Finally, in order to assess the independent prognostic value of SF mutations, we performed a multivariable Cox regression analysis accounting for known risk factors and potential confounders (which were also significantly associated with OS and EFS in univariable analysis, Table S8). These included age, gender, WBC, disease entity, type of stem cell transplantation (none vs autologous or allogeneic stem cell transplantation) and ELN 2017 risk classification. We found that adding SFmut4 does not improve the baseline model of ES and OS when the current ELN 2017 classification is included (Figure 4D; Figure S4A). Since SF mutations often co-occur with both *RUNX1* and *ASXL1* mutations, which are currently part of the ELN 2017 classification,

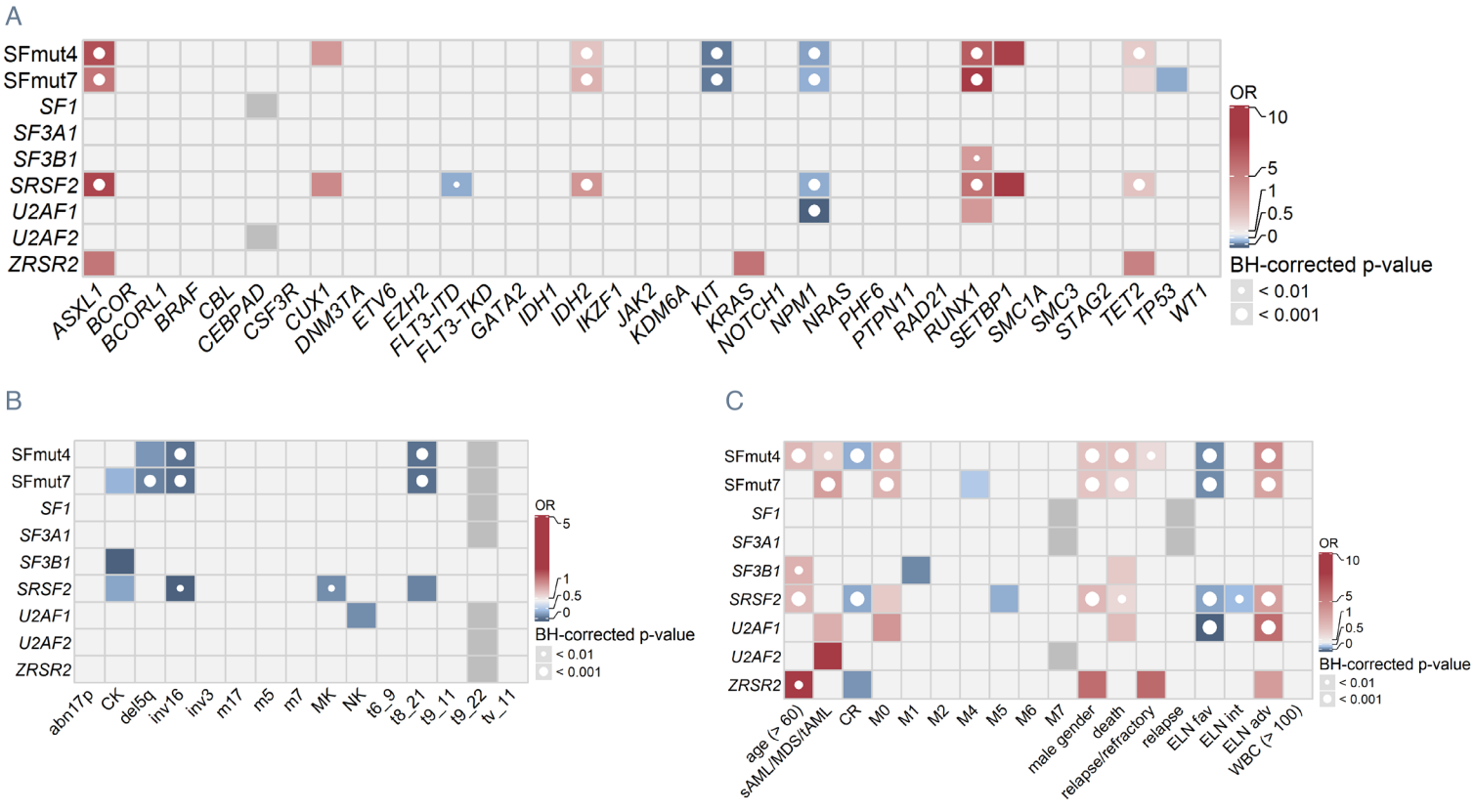


Figure 3. Associations between SF mutations and other recurrent gene mutations, cytogenetics, as well as clinical variables. Associations were determined using the Fisher's exact test with Benjamini-Hochberg correction for multiple testing. The odds ratios for the significant associations are color coded. Red colors indicate positive association (co-occurrence) while blue colors indicate negative association (mutual exclusivity). Grey indicates missing data. The significance levels are indicated by the presence and size of the circle in each cell. A - Co-occurrence and mutual exclusivity of SF mutations and other gene mutations. B - Associations between SF mutations and cytogenetics. C - Associations between gene mutations and clinical patient characteristics.

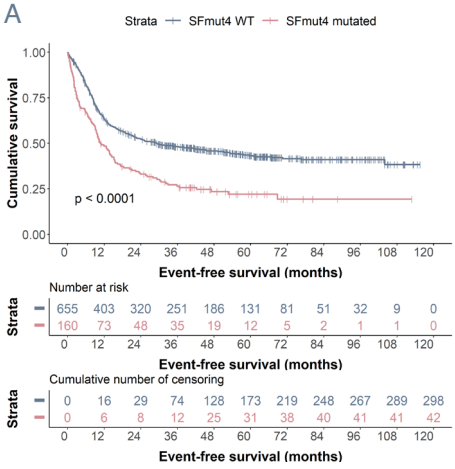
we created a modified ELN 2017 classification by excluding *RUNX1* and *ASXL1* mutations (Figure 4E; Figure S4B). In the model constructed with this modified ELN classification, SFmut4 was independently associated with OS and EFS. Furthermore, addition of SFmut4 improved the model fit (AIC 4368.7 for baseline model vs 4365.7 for SF-including model of EFS, AIC 3626.1 for baseline model vs 3622.6 for SF-including model of OS, also (Table S9-13). In addition, the AIC values for models including SFmut4 and modified ELN 2017 classification were superior to AIC values of the baseline models with the current ELN 2017 classification (AIC 4376.2 for EFS and AIC 3634.5 for OS). The effects of SFmut4 were stronger in comparison to mutations in individual SFs separately, none of which showed significant independent predictive value. Strikingly, in contrast to SFmut4, addition of *RUNX1* or *ASXL1* mutations to the baseline model with modified ELN 2017 classification did not improve its predictive value (Figure S9). Thus, both *RUNX1* and *ASXL1* mutations on its own, are not strong independent predictors of the treatment outcome. Taken together, these findings highlight the importance of SF mutations for prognostication of AML treatment outcome.

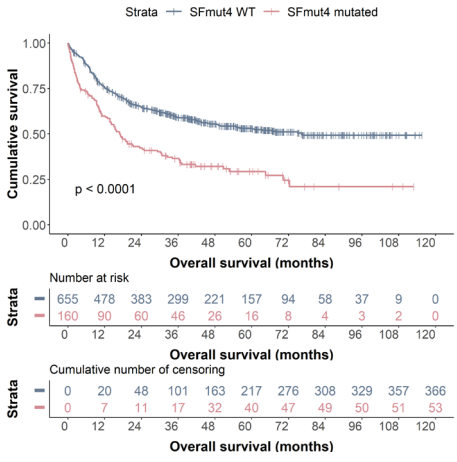
INFLUENCE OF GENE INTERACTIONS ON SURVIVAL

Next, we further explored the prognostic value of the interactions between SFmut4 and *RUNX1* as well as *ASXL1* mutations. Interestingly, Kaplan-Meier analysis of EFS and OS showed that the co-occurrence of SFmut4 with *RUNX1* mutations was associated with particularly adverse outcome (Figure 5A). This effect was largely driven by the interaction of *SRSF2* mutations with RUNX1 mutations (Figure 5B). Moreover, the presence of *RUNX1* mutations without SFmut4 did not indicate inferior survival anymore (Figure 5A-B; Figure S6A), which suggests that mutated *RUNX1* is not the driver behind the independent prognostic value of SFmut4 in the context of modified ELN 2017 classification. In addition, *RUNX1* mutations were previously shown to be significantly associated with adverse outcome in AML patients younger than 60 years[21], therefore we analyzed the interaction between SFmut4 and *RUNX1* mutations also in this group separately and confirmed our findings from the total cohort (Supplemental Results). The lack of adverse prognostic value of *RUNX1* mutations without co-occurring SFmut4 was further supported by the analysis within the adverse ELN 2017 risk group, where patients with a mutation in *RUNX1* without any of these four most common SF mutations (*SRSF2*, *SF3B1*, *U2AF1* and *ZRSR2*) had relatively good outcome (Figure S10). The OS as well as EFS of these patients was comparable to the intermediate risk group: 5-year OS of 43.6% (as compared to 45.6% for intermediate risk patients) and 5-year EFS of 38.0% (36.4% for intermediate risk patients). In addition, double-mutated patients showed worse OS and EFS within the chromatin-spliceosome subgroup defined by Papaemanuil et al (data not shown). Similar associations, although somewhat less pronounced, were found for the combination of SFmut4 with *ASXL1* mutations (Figure S6B).

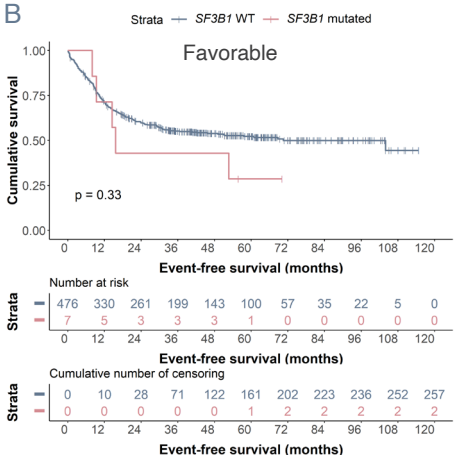
Our multivariable analysis including the modified ELN 2017 risk classification further supports the observation that co-occurrence of *RUNX1* mutations with SFmut4 and in particular co-occurrence of *RUNX1* mutations with *SRSF2* mutations, is independently associated with a particularly poor EFS and OS (Figure 5C-D; Figure S8A). In the absence of any of the four most recurrent SF mutations, *RUNX1* was not indicative of adverse risk in the multivariable model. The interactions of SFmut4 with *ASXL1* mutations (Figure S8B) was not indicative of inferior EFS or OS when examined in the multivariable model (Table S9-12).

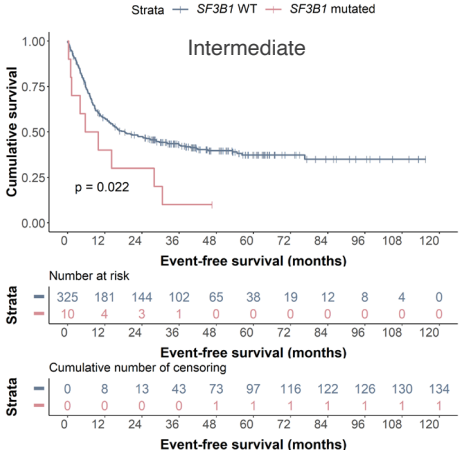
A


Strata SFmut4 WT SFmut4 mutated
Cumulative survival
p < 0.0001
Event-free survival (months)
Number at risk
655 403 320 251 186 131 81 51 32 9 0
160 73 48 35 19 12 5 2 1 1 0
Event-free survival (months)
Cumulative number of censoring
0 16 29 74 128 173 219 248 267 289 298
0 6 8 12 25 31 38 40 41 41 42
Event-free survival (months)

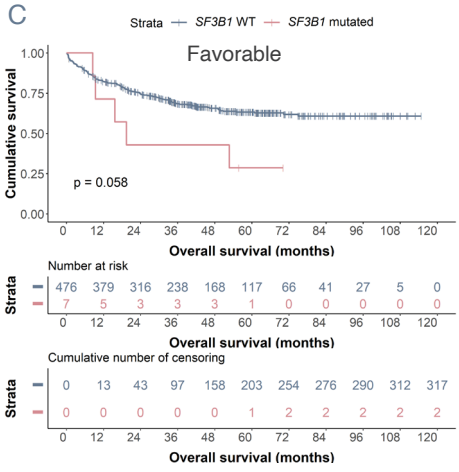

Strata SFmut4 WT SFmut4 mutated
Cumulative survival
p < 0.0001
Overall survival (months)
Number at risk
655 478 383 299 221 157 94 58 37 9 0
160 90 60 46 26 16 8 4 3 2 0
Overall survival (months)
Cumulative number of censoring
0 20 48 101 163 217 276 308 329 357 366
0 7 11 17 32 40 47 49 50 51 53
Overall survival (months)

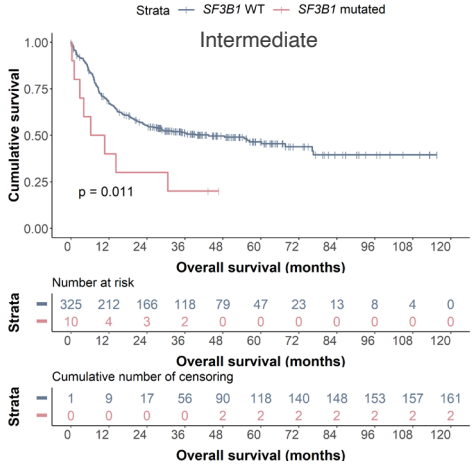
B


Strata SF3B1 WT SF3B1 mutated
Favorable
Cumulative survival
p = 0.33
Event-free survival (months)
Number at risk
476 330 261 199 143 100 57 35 22 5 0
7 5 3 3 3 1 0 0 0 0 0
Event-free survival (months)
Cumulative number of censoring
0 10 28 71 122 161 202 223 236 252 257
0 0 0 0 0 1 2 2 2 2 2
Event-free survival (months)

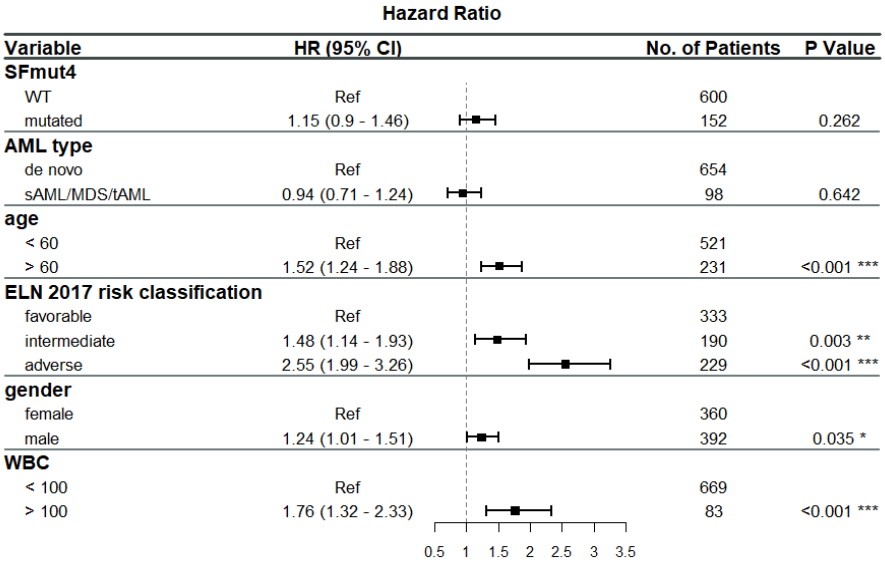

Strata SF3B1 WT SF3B1 mutated
Intermediate
Cumulative survival
p = 0.022
Event-free survival (months)
Number at risk
325 181 144 102 65 38 19 12 8 4 0
10 4 3 1 0 0 0 0 0 0 0
Event-free survival (months)
Cumulative number of censoring
0 8 13 43 73 97 116 122 126 130 134
0 0 0 0 1 1 1 1 1 1 1
Event-free survival (months)

C


Strata SF3B1 WT SF3B1 mutated
Favorable
Cumulative survival
p = 0.058
Overall survival (months)
Number at risk
476 379 316 238 168 117 66 41 27 5 0
7 5 3 3 3 1 0 0 0 0 0
Overall survival (months)
Cumulative number of censoring
0 13 43 97 158 203 254 276 290 312 317
0 0 0 0 0 1 2 2 2 2 2
Overall survival (months)


Strata SF3B1 WT SF3B1 mutated
Intermediate
Cumulative survival
p = 0.011
Overall survival (months)
Number at risk
325 212 166 118 79 47 23 13 8 4 0
10 4 3 2 0 0 0 0 0 0 0
Overall survival (months)
Cumulative number of censoring
1 9 17 56 90 118 140 148 153 157 161
0 0 0 0 2 2 2 2 2 2 2
Overall survival (months)

D



E

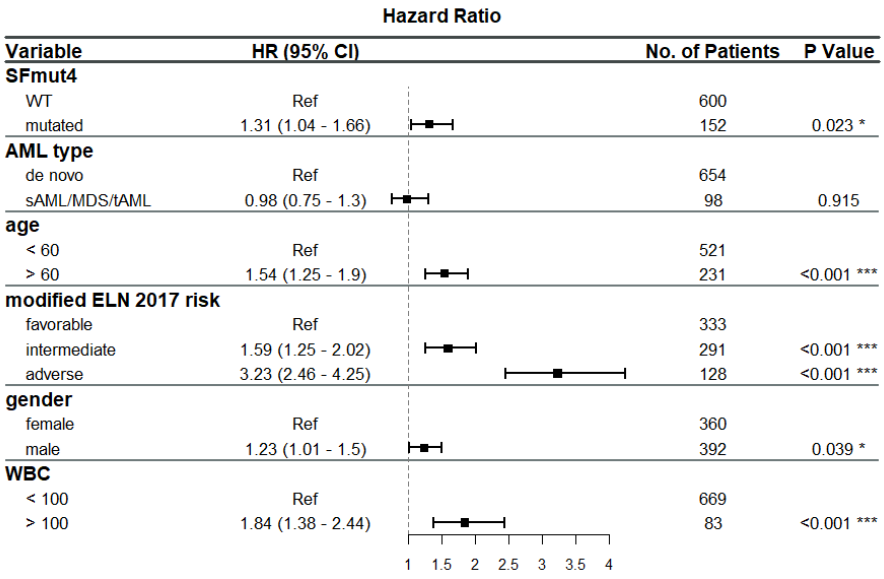


Figure 4. Survival of AML patients in relation to the presence of SF mutations. A - The figure depicts Kaplan-Meier analyses of event-free survival (left) and overall survival (right) in the total AML cohort in relation to SFmut4. B, C - Kaplan-Meier curves for event-free survival (B) and overall survival (C) in relation to the mutation status of *SF3B1* within the favorable or intermediate risk groups as defined by the ELN 2017 classification. D,E - Multivariable Cox regression analysis of event-free survival in relation to SFmut4 with complete (D) and modified (E) ELN 2017 classification. In the modified ELN 2017 classification *RUNX1* and *ASXL1* mutations were excluded, so that patients carrying *RUNX1* or *ASXL1* mutations were re-classified based on the presence of the rest of aberrations in this classification system. Type of stem cell transplantation violated the proportional hazard assumption and therefore it was used as strata variable in all the multivariable Cox regression models (for variables used in the model as strata the statistics are not calculated and therefore, they do not appear in the results).

DISCUSSION

In this study, we focussed specifically on characterization of genetic background and clinical relevance of recurrent SF mutations in a large cohort of AML patients. For the purpose of our analyses, we combined several independent AML patient cohorts with confirmed compatibility (and limited heterogeneity) with respect to genetic and clinical features. The heterogeneity between cohorts was further limited by inclusion in our survival analyses of only intensively treated patients.

We demonstrated that SF mutations as combined groups affect 22% or 35% of AML patients for SFmut4 and SFmut7 respectively. In addition, the vast majority of SF mutations were found to be mutually exclusive.[5] The presence of SF mutations in AML and mutual exclusivity of these mutations together with the effectivity of recently developed splicing modulators indicate that splicing deregulation in AML is key and cooperates with other lesions to promote leukemogenesis.[22] Therefore, it is essential to explore the involvement of SF mutations in relation to other mutations and their (combined) relevance for patient outcome.

In this respect, the two most commonly co-mutated genes together with SFs include two transcription factors (*RUNX1* and *ASXL1*). As splicing occurs co-transcriptionally it is conceivable that mutations in this class of gene expression regulators work together to initiate leukemia.[23] Accordingly, mutations in *RUNX1* were previously reported to be associated with increased intron retention in AML.[24] In line with these observations, a recent study demonstrated that mutated *RUNX1* knockout alone resulted in splicing alterations, which were further broadened in the concomitant presence of *SRSF2* mutation.[9,10]

Exploration of associations between SF mutations and specific AML features showed that any of the four common SF mutations (also coinciding with *RUNX1* mutations) were strongly associated with a very immature cell state (FAB-M0). In addition, an intriguing association was noted between SF mutations (particularly *SRSF2* and *ZRSR2*) and male gender. Several previous studies observed the same association in MDS and AML, however in some cases it did not reach statistical significance due to low sample numbers.[25,26] Furthermore, it was suggested that gender has impact on splicing profiles of various tissues and the degree of global aberrant splicing was reported to be higher in men with MDS as compared to women with the same disorder.[8,9] Future studies better designed to examine gender-specific genetic background and clinical associations with SF mutations, including treatment outcome, should be conducted.

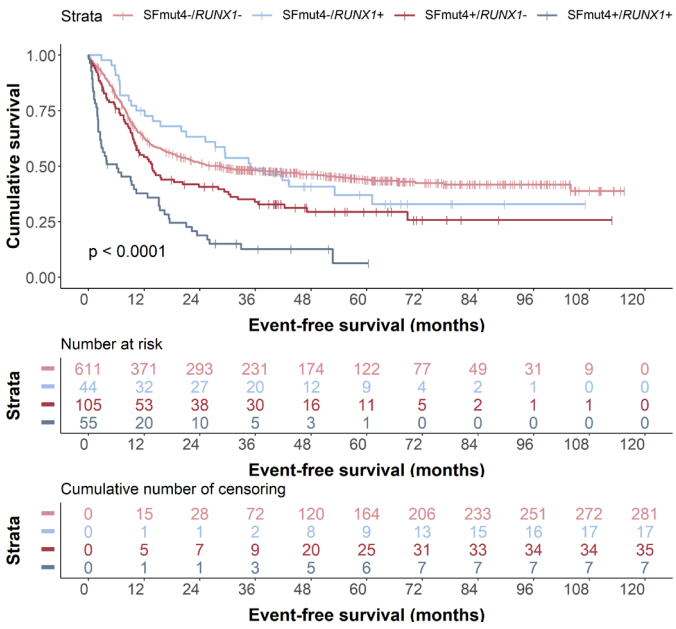
We found that SF mutations were associated with older age, lower WBC and secondary AML, which is in agreement with previous observations.[27] In parallel to these clinical features, SF mutations in our cohort were correlated with shorter OS and EFS, which can be explained by inferior response to therapy reflected in lower rates of CR, higher rates of disease relapse or refractoriness and death. Recently, Lindsey et al. postulated that mutations in the chromatin-spliceosome subtype, including mutations in *SRSF2*, *SF3B1*, *U2AF1* and *ZRSR2*, are characteristic to secondary AML, and in elderly *de novo* AML they identify a subtype with poor prognosis.[28] While this study also highlights the importance of cohesin mutations in this con-

text, we did not identify significant co-occurrence of this type of aberrations with SF mutations. In addition, another recent report suggests that the chromatin-spliceosome mutational signature in *de novo* AML identifies a disease subtype with a poor treatment outcome resembling that of secondary AML.[9,29] This is further supported by the gene interactions observed in our study. SF mutations frequently co-occurred with mutations in *ASXL1*, *RUNX1* and *TET2*, which constitute aberrations commonly found in older patients and secondary AML/MDS.[9,29] At the same time SF mutations were mutual exclusive with CBF translocations which typically show lower incidence in older patients.[9] Taken together, these results support the notion that SF mutations identify a subtype of *de novo* AML with more secondary AML-like features. However, even in our subgroup analysis patients with sAML/MDS/tAML, SF mutations were associated with worse OS and EFS (data not shown).

Overall SFmut4 was associated with shorter OS and EFS in univariable and multivariable models. Just as reported by Papaemanuil et al, this suggests that splicing deregulation in general, via mutation in any splicing factor, confers poor prognosis. In our study, we identified two facets that could drive further improvements in prognostication of AML. Firstly, *SF3B1* mutations among patients classified as favourable and in particular intermediate risk marked individuals with worse OS and EFS, suggesting that these patients would benefit from more intensive treatment or innovative therapies such as splicing modulation. Secondly, we found that patients carrying concomitant SFmut4 and *RUNX1* mutations had a particularly poor prognosis in both univariable and multivariable analysis. This effect was largely due to the particularly strong interaction of *SRSF2* with *RUNX1* mutations. Strikingly, patients with *RUNX1* mutation without any of the four common SF mutations had longer OS and EFS as compared to the rest of patients within the adverse risk group according to the ELN 2017 classification. Hence, *RUNX1* mutations may not be relevant for risk assessment without the co-occurrence with any of the four most common SF mutations. Importantly, exclusion of patients carrying any of the four most common SF mutations from *RUNX1* mutated subgroup had stronger effect than exclusion of *SRSF2* mutated patients only, highlighting the importance of considering mutations in any SF as a subgroup. Further underscoring the prognostic value of SFmut4, the multivariable model including SFmut4 and modified ELN 2017 classification, excluding *ASXL1* and *RUNX1* mutations, was superior at predicting survival than a multivariable model without SFmut4 but including the current ELN 2017 classification.

Finally, it was previously suggested that *RUNX1* mutations are predictive of inferior outcome only in AML patients younger than 60 years.[8] Strikingly, the association of concomitant SFmut4 and *RUNX1* mutations with inferior survival was even more pronounced in AML patients within this age group. The fact that this association was weaker in AML patients older than 60 years seems to, at least partly, stem from the fact that the survival of older patients is in general relatively short. As a result, the impact of genetic lesions is thought to be less pronounced in this subgroup. The interaction of SFmut4 with mutations in *ASXL1* showed similar but less pronounced associations with survival in Kaplan-Meier analysis and was not associated with survival in multivariable analysis, although a previous study suggested that co-occurrence of any two genes from within the chromatin-spliceosome subgroup typically confers particularly short survival.[14] Accordingly, the prognostic value of *ASXL1* mutations without SF mutations should be further explored.

A



B

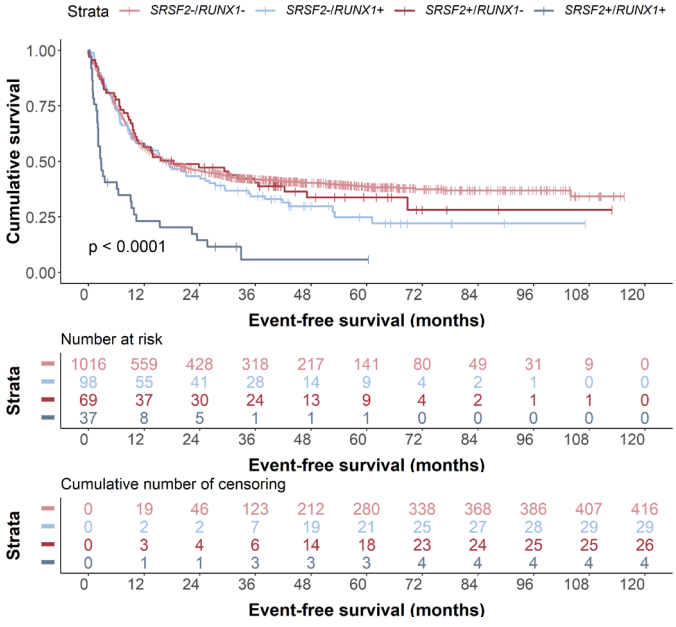
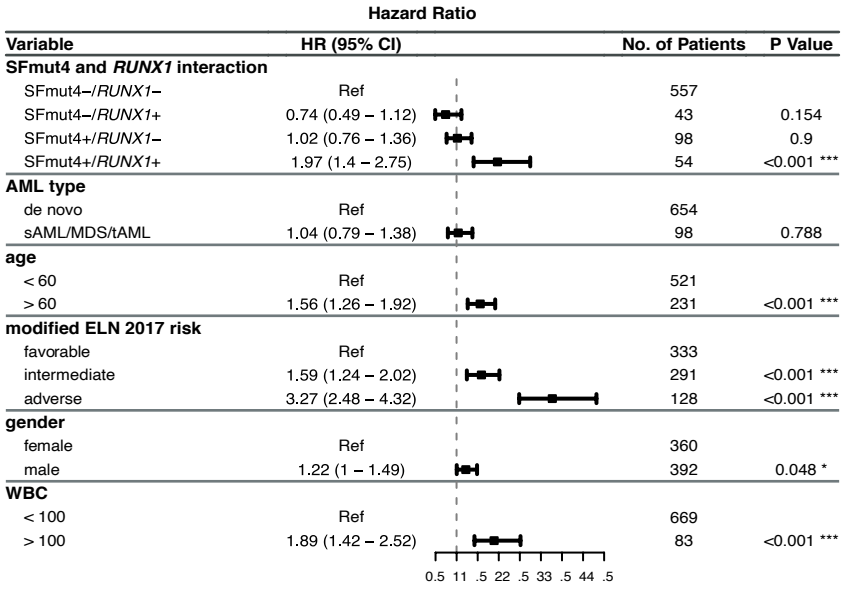


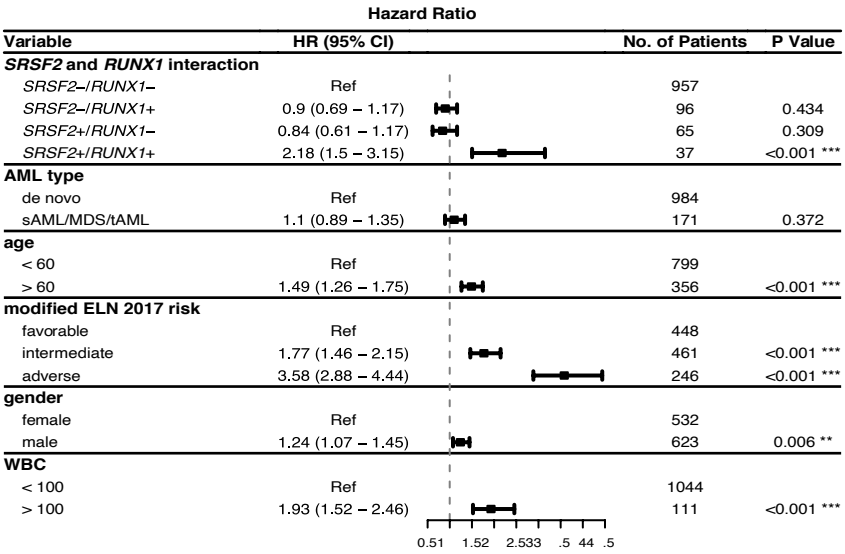
Figure 5. Influence of SF mutations on survival of *RUNX1* and *ASXL1* mutated patients. A - Kaplan Meier curves for event-free survival in relation to SFmut4 in combination with the mutation status of *RUNX1*. B - Kaplan Meier curves for event-free survival in relation to the mutation status of *SRSF2* in combination with *RUNX1*.

C



Variable	HR (95% CI)	No. of Patients	P Value
SFmut4 and *RUNX1* interaction			
SFmut4–/*RUNX1*–	Ref	557	
SFmut4–/*RUNX1*+	0.74 (0.49 – 1.12)	43	0.154
SFmut4+/*RUNX1*–	1.02 (0.76 – 1.36)	98	0.9
SFmut4+/*RUNX1*+	1.97 (1.4 – 2.75)	54	<0.001 ***
AML type			
de novo	Ref	654	
sAML/MDS/tAML	1.04 (0.79 – 1.38)	98	0.788
age			
< 60	Ref	521	
> 60	1.56 (1.26 – 1.92)	231	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	333	
intermediate	1.59 (1.24 – 2.02)	291	<0.001 ***
adverse	3.27 (2.48 – 4.32)	128	<0.001 ***
gender			
female	Ref	360	
male	1.22 (1 – 1.49)	392	0.048 *
WBC			
< 100	Ref	669	
> 100	1.89 (1.42 – 2.52)	83	<0.001 ***

D



Variable	HR (95% CI)	No. of Patients	P Value
***SRSF2* and *RUNX1* interaction**			
SRSF2–/*RUNX1*–	Ref	957	
SRSF2–/*RUNX1*+	0.9 (0.69 – 1.17)	96	0.434
SRSF2+/*RUNX1*–	0.84 (0.61 – 1.17)	65	0.309
SRSF2+/*RUNX1*+	2.18 (1.5 – 3.15)	37	<0.001 ***
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.1 (0.89 – 1.35)	171	0.372
age			
< 60	Ref	799	
> 60	1.49 (1.26 – 1.75)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	1.77 (1.46 – 2.15)	461	<0.001 ***
adverse	3.58 (2.88 – 4.44)	246	<0.001 ***
gender			
female	Ref	532	
male	1.24 (1.07 – 1.45)	623	0.006 **
WBC			
< 100	Ref	1044	
> 100	1.93 (1.52 – 2.46)	111	<0.001 ***

Figure 5. Influence of SF mutations on survival of *RUNX1* and *ASXL1* mutated patients (continued). C, D - Multivariable Cox regression analysis of event free survival in relation to the mutation status of SRSF2 (C) or SFmut4 (D) with *RUNX1* mutations and modified ELN 2017 classification. In the modified ELN 2017 classification *RUNX1* and *ASXL1* mutations were excluded, so that patients carrying *RUNX1* or *ASXL1* mutations were re-classified based on the presence of the rest of aberrations in this classification system. Type of stem cell transplantation violated the proportional hazard assumption and therefore it was used as strata variable in all the multivariable Cox regression models (for variables used in the model as strata the statistics are not calculated and therefore, they do not appear in the results). Since the number of cases for which data on all 4 most common SF mutations (included in the SFmut4 variable) were available was lower than the number of cases for which *SRSF2* alone was available, the patient numbers in analyses in panel B and C differ. WBC – white blood cell count; sAML – secondary AML, treatment-related AML and MDS.

In conclusion, we propose that SF mutations should be considered for incorporation in diagnostic and prognostic guidelines, since the adverse prognostic value of *RUNX1* mutations may not be valid without co-occurrence of any of the four most common SF mutations. Furthermore, mutations in *SF3B1* should be considered as a prognostic marker in individuals currently classified as favourable and intermediate risk. *SF3B1* mutated patients in those subgroups currently have relatively short OS and EFS suggesting that these individuals could benefit from more intensive treatment or innovative therapies such as splicing modulation. While our study already included several independent cohorts combined, future studies should validate our proposed refinements and extend the analysis of SF mutations in relation to minimal residual disease (MRD).

REFERENCES

1. Yoshida K, Sanada M, Shiraishi Y, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011;478(7367):64-69. doi:10.1038/nature10496
2. Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. New England Journal of Medicine. 2011;365(15):1384-1395. Accessed July 2, 2014. http://www.nejm.org/doi/full/10.1056/nejmoa1103283
3. Rossi D, Bruscaggin A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011;118(26):6904-6908. doi:10.1182/blood-2011-08-373159
4. Dvinge H, Kim E, Abdel-wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nature Reviews Cancer. 2016;16(7):413-430. doi:10.1038/nrc.2016.51
5. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nature Medicine. 2016;22(9):976-986. doi:10.1038/nm.4165
6. Chen M, Manley J. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nature Reviews Molecular Cell Biology. 2009;10(11):741-754. doi:10.1038/nrm2777
7. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-2405. doi:10.1182/blood-2016-03-643544
8. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. New England Journal of Medicine. 2016;374(23):2209-2221. doi:10.1056/NEJMoa1516192
9. Metzeler KH, Herold T, Rothenberg-thurley M, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood. 2016;128(5):686-699. doi:10.1182/blood-2016-01-693879.Presented
10. Hou H, Liu C, Kuo Y, Chou W. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. Oncotarget. Published online 2016:Published online Jan 24.
11. Döhner H, Estey E, Grimwade D, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-448. doi:10.1182/blood-2016-08-733196.424
12. Pabst T, Vellenga E, van Putten W, et al. Favorable effect of priming with granulocyte colony-stimulating factor in remission induction of acute myeloid leukemia restricted to dose escalation of cytarabine. Blood. 2012;119(23):5367-5373. doi:10.1182/blood-2011-11-389841
13. Löwenberg B, Pabst T, Maertens J, et al. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. Blood. 2017;129(12):1636-1645. doi:10.1182/blood-2016-10-740613
14. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Minimal residual disease in acute myeloid leukemia. New England Journal of Medicine. 2018;378(13):1189-1199. doi:10.1007/978-3-319-97746-1_7
15. Höllein A, Jeromin S, Meggendorfer M, et al. Minimal residual disease (MRD) monitoring and mutational landscape in AML with RUNX1-RUNX1T1: a study on 134 patients. Leukemia. 2018;32(10):2270-2274. doi:10.1038/s41375-018-0086-0
16. Stengel A, Kern W, Haferlach T, Meggendorfer M, Fasan A, Haferlach C. The impact of TP53 mutations and TP53 deletions on survival varies between AML, ALL, MDS and CLL: An analysis of 3307 cases. Leukemia. 2017;31(3):705-711. doi:10.1038/leu.2016.263
17. Cappelli LV, Meggendorfer M, Dicker F, et al. DNMT3A mutations are over-represented in young adults with NPM1 mutated AML and prompt a distinct co-mutational pattern. Leukemia. 2019;33(11):2741-2746. doi:10.1038/s41375-019-0502-0
18. Baer C, Pohlkamp C, Haferlach C, Kern W, Haferlach T. Molecular patterns in cytopenia patients with or without evidence of myeloid neoplasm—a comparison of 756 cases. Leukemia. 2018;32(10):2295-2298. doi:10.1038/s41375-018-0119-8
19. Schoch C, Schnittger S, Bursch S, et al. Comparison of chromosome banding analysis, interphase- and hypermetaphase-FISH, qualitative and quantitative PCR for diagnosis and for follow-up in chronic myeloid leukemia: A study on 350 cases. Leukemia. 2002;16(1):53-59. doi:10.1038/sj.leu.2402329
20. McGowan-Jordan, J., Simon, A., Schmid M. An International System for Human Cytogenetic Nomenclature.; 2016.
21. Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T. Epigenetics in alternative pre-mRNA splicing. Cell. 2011;144(1):16-26. doi:10.1016/j.cell.2010.11.056

22. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. Genome Medicine. 2015;7(1):1-13. doi:10.1186/s13073-015-0168-9
23. Huang Y-J, Yan M, Chen J-Y, et al. RUNX1 Deficiency Cooperates with SRSF2 Mutation to Further Disrupt RNA Splicing and Exacerbate Myelodysplastic Syndromes in Mouse Models. Blood. 2018;132:(Supplement 1): 1796.
24. Karlebach G, Veiga DFT, Mays AD, et al. The impact of sex on alternative splicing. bioRxiv. 2018;1:490904.
25. Yang Y, Chiu Y, Kao C, et al. The prognostic signi fi cance of global aberrant alternative splicing in patients with myelodysplastic syndrome. Blood Cancer Journal. Published online 2018. doi:10.1038/s41408-018-0115-2
26. Lindsley RC, Mar BG, Mazzola E, et al. Acute myeloid leukemia ontogeny is de fi ned by distinct somatic mutations. Blood. 2015;125(9):1367-1377. doi:10.1182/blood-2014-11-610543.The
27. Caprioli C, Lussana F, Salmoiraghi S, et al. Identification of a Chromatin-Splicing Mutational Signature to Define Secondary Acute Myeloid Leukemia: A Report from the Northern Italy Leukemia Group (NILG) Prospective Trial 02/06. Blood. 2019;134(Supplement_1):1443. doi:10.1182/blood-2019-124497
28. Sperling AS, Gibson CJ, Ebert BL. The genetics of myelodysplastic syndrome: From clonal haematopoiesis to secondary leukaemia. Nature Reviews Cancer. 2017;17(1):5-19. doi:10.1038/nrc.2016.112

SUPPLEMENTAL METHODS

PATIENTS AND TREATMENT PROTOCOL

The detailed treatment protocols and eligibility criteria for HOVON trials were described previously.[1,2] All HOVON patients received intensive chemotherapy induction followed by risk-based assignment to either chemotherapy or stem cell transplantation as consolidation treatment. Patients from the MLL cohort were treated either with intensive chemotherapy, non-intensive chemotherapy or were given supportive care, all according to the standard treatment protocols in Germany. For survival analyses only MLL patients treated with intensive chemotherapy were included, in combination with all HOVON patients. All three HOVON protocols (H42A, H92 and H102) included the same therapeutic backbone with high dose chemotherapeutics and investigational drugs. Specifically, H102 study, included patients 18 to 65 years of age, investigated Clofarabine, H92 randomly assigned Laromustine as investigational drug while in H42A, patients 18 to 60 years of age, received similar induction chemotherapy but were randomly administered G-SCF. The MLL patients were treated with the same 7+3 backbone as the HOVON patients without the addition of an investigational drugs. No significant differences in overall survival were found between HOVON (N=889) and intensively treated MLL patients (N=334; Figure S9A), which may be expected since, the investigational drugs did not render survival benefits. In addition, we included an analysis of patients who achieved complete remission only since for both H42A (N=133) and H92 (N=43) only patients with complete remission were included in the current study (due to lack of availability of data on the presence of SF mutations in patients who did not reach CR from H42A and H92 studies). It should be noted that the majority of patients in our study were treated on the H102 protocol (N=889, including both CR and non-CR patients) and therefore the relatively small number of patients from H42A and H92 result in only a minor overrepresentation of CR patients in our cohort, which is not expected to substantially affect the conclusions. Accordingly, all patients with complete remission presented comparable outcome for both HOVON and MLL patients (Figure S9B).

A specific subgroup of patients was defined as secondary AML and included patients with antecedent myeloid disorder (N=61). Since this is supposed to be a subgroup with substantially different characteristics, they were analysed separately together with MDS (N=72) and treatment related AML (N= 43; Supplemental Table S6).

SAMPLE PREPARATION

Mononuclear cell fraction was isolated from bone marrow (BM) or peripheral blood (PB) by Ficoll-Hypaque (Nygaard, Oslo, Norway) density centrifugation and cryopreserved for further processing. For targeted NGS (see below), following thawing, cells were lysed in RLT buffer with the addition of DTT (Qiagen, Venlo, The Netherlands). Subsequently DNA was isolated as described previously.[3]

MUTATIONAL PROFILE

The mutational status was determined for 41 commonly mutated genes in AML, including 7 splicing factors (ASXL1, BCOR, BCORL1, BRAF, CBL, CEBPA, CSF3R, CUX1, DNMT3A, ETV6, EZH2, FLT3-ITD, FLT3-TKD, GATA2, IDH1, IDH2, IKZF1, JAK2, KDM6A, KIT, KRAS, NOTCH1,

NPM1, NRAS, PHF6, PTPN11, RAD21, RUNX1, SETBP1, SF1, SF3A1, SF3B1, SMC1A, SMC3, SRSF2, STAG2, TET2, TP53, U2AF1, U2AF2, WT1, ZRSR2).

WHOLE GENOME SEQUENCING

Whole genome sequencing (WGS) libraries were generated from 1μg of DNA extracted from bone marrow or peripheral blood samples using the TruSeq PCR free library prep kit, following the manufacturer's recommendations (Illumina, San Diego, CA, USA) and sequenced on a NovaSeq6000 or HiSeqX Illumina instruments following a 2x150bp paired-end reads standard protocol at a mean depth of coverage of >100x. Bioinformatic analysis of WGS data was performed using Illumina's BaseSpace Sequence Hub and in-house pipelines. Reads were aligned against human genome build GRCh37/hg19 with the tool Isaac3.[4] Variant calling was performed using Strelka2[5] and additional variant annotation was performed using Ensembl VEP.[6] Only exonic (non-synonymous single nucleotide variants (SNVs) and small insertions/deletions (indels)) and variants at splicing acceptor/donor sites were considered in this study. Because matched normal samples were not available, tumor-unmatched normal variant calling was performed using a pool gender-matched DNA (Promega, Madison, WI, USA). In order to remove germline and benign variants as well as technical artifacts from the dataset a strict filtering strategy was applied as described below.

FILTERING WHOLE GENOME SEQUENCING DATA

Upon variant calling the following in-house filtering strategy[7] was applied to remove sequencing artifacts as well as germline and likely benign variants, resulting in a dataset consisting of reliable and likely somatic, pathogenic variants.

Sequencing artifacts

Repetitive regions and regions potentially troublesome for variant calling were annotated based on Ensembl repeat database and the Global Alliance for Genomics and Health (GA4GH) database. Firstly, variants located in regions of established low confidence variant calling (as specified in Genome in a Bottle Consortium using sample NA12878/HG001, https://github.com/genome-in-a-bottle). Subsequently, we discarded variants with low frequency (VAF<15%) located in low complexity regions (homopolymers), tandem repeats (i.e. microsatellites), segmental duplications or repetitive regions interspersed throughout the genome (i.e. transposable elements). Finally, variants supported by ≤ 3 reads for the alternative allele as well as with the total depth of coverage ≤ 5 were removed from downstream analysis.

Germline variants

First, variants with the global population frequency ≥ 0.001 (based on the genome aggregation database; gnomAD) were discarded. Next, variants labelled as germline in ClinVar or COSMIC databases were filtered out. Finally, variants not well annotated in hematological malignancies with a VAF between 40-60% or > 90% in all samples were eliminated. This strategy was applied in order to remove as many of the germline variants as possible although it should be noted that due to the lack of matched germline control some residual germline variants are likely to be retained in our dataset. However, since this study focuses on genes recurrently mutated in AML these residual germline variants are not expected to affect the analysis.

Likely benign variants

First, variants listed as benign/likely benign in ClinVar database were removed. Subsequently, we used an in-house developed tool HePPy (Hematological Predictor of Pathogenicity)[8] to remove missense variants with HePPy score < 0.75 (indicating low pathogenicity).

STATISTICAL ANALYSIS

In our study, OS was defined as time from the initial diagnosis to death or last follow-up. EFS was defined as time from initial diagnosis to an event (refractoriness, relapse, death or last follow-up, whichever occurred first). In the multivariable Cox proportional hazards model, the prognostic value of SF mutations was evaluated in the context of demographic and clinical variables (including age, gender, type of AML, white blood cell count, type of stem cell transplantation administered) as well as the ELN 2017 risk classification or modified ELN 2017 classification (all of which were also significantly associated with OS and EFS in univariable Cox model, Supplemental Table S8). In order to assess the individual and combined prognostic value of *RUNX1* and *ASXL1* with SF mutations, a modified ELN 2017 classification was generated by excluding *RUNX1* and *ASXL1* from the list of genetic criteria due to their frequent co-occurrence with SF mutations (hence patients carrying these mutations were re-classified based on the remaining criteria). The proportional hazard assumption was evaluated for each variable using the Schoenfeld test and upon examination of the plots of Schoenfeld residuals. Type of stem cell transplantation violated the proportional hazard assumption and therefore it was used as strata variable in all the multivariable Cox regression models (for variables used in the model as strata the statistics are not calculated and therefore, they do not appear in the results). In addition, the interaction between SF mutations and *RUNX1* variables violated the proportional hazard assumption and was additionally evaluated as time-dependent co-variate. The Akaike Information Criterion (AIC) was used to compare the fit of the baseline models with the fit of models containing variables of interest (SF mutations or their interaction with other genetic mutations). This AIC criterion informs not only about the goodness of fit of the model but also penalizes on the number of variables in the model which makes it an attractive method to compare models (the lower the AIC value the better, regardless of the magnitude of difference). To further substantiate the differences in the fit of the models, we have additionally performed ANOVA analyses to compare the fit of the baseline and new model for each tested variable.

In all analyses p-value < 0.05 was considered statistically significant (for Fisher's test the 0.05 cut-off was applied to BH-corrected p-values). All statistical analyses were carried out in R version 3.6.3/R studio version 1.2.5, including the following packages: survminer (version 0.4.6)[9], survival (version 3.1-11, https://CRAN.R-project.org/package=survival)[10], ggplot2 (version 3.2.1)[11], ComplexHeatmap (version 2.2.0)[12]. The patient numbers in particular analyses vary depending on the amount of available data for specific variables.

REFERENCES

1. Pabst, T. et al. Favorable effect of priming with granulocyte colony-stimulating factor in remission induction of acute myeloid leukemia restricted to dose escalation of cytarabine. Blood 119, 5367–5373 (2012).
2. Lowenberg, B. et al. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. Blood 129, 1636-1645 (2017).
3. Jongen-Lavrencic, M. et al. Minimal residual disease in acute myeloid leukemia. New England Journal of Medicine 378:1189–1199, 2018
4. Raczy C.et al: Isaac: Ultra-fast whole-genome secondary analysis on Illumina sequencing platforms. Bioinformatics 29:2041–2043, 2013
5. Kim S. et al: Strelka2: fast and accurate calling of germline and somatic variants. Nature Methods 15:591–594, 2018
6. McLaren, W. et al. The Ensembl Variant Effect Predictor. Genome Biology 6: 122, 2016
7. Palomo L. et al. Molecular landscape and clonal architecture of adult myelodysplastic/myeloproliferative neoplasms. Blood Jun 23:blood.2019004229, 2020
8. Hutter S. et al: A Novel Machine Learning Based in silico Pathogenicity Predictor for Missense Variants in a Hematological Setting. Blood 134:2090–2090, 2019
9. Kassambara, A., Kosinski, M. & Biecek, P. Survminer: Drawing Survival Curves Using “Ggplot2”. Version 0.4.6. URL: https://CRAN.R-project.org/package=survminer (2019) doi:.1037//0033-2909.I26.1.78.
10. Therneau, T. A Package for Survival Analysis in R. (2020).
11. Wickham, H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York (2016) doi:10.18637/jss.v035.b01.
12. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849 (2016).

SUPPLEMENTAL RESULTS

STUDY DESIGN

In this study, two independent cohorts of AML patients were included (the HOVON and MLL cohort, see Materials and Methods in the main paper). Acknowledging the heterogenous origin of our cohort, the genetic landscape as well as clinical features and treatment outcome of patients of both cohorts were carefully compared to assure that the inter-cohort heterogeneity will not bias the analysis.

The mutational profiles of patients included in both cohorts were primarily defined based on routine molecular diagnostics. In case of MLL patients, molecular diagnostics were additionally complemented by whole genome sequencing (WGS). The routine diagnostics data for 430 patients of the HOVON dataset were complemented by targeted sequencing using Illumina TruSight Myeloid Panel. The complete genetic landscape of both HOVON as well as MLL patients was found to be consistent and typical of AML with frequencies of cytogenetic and molecular aberrations being consistent with those reported in previous studies (data not shown).[1,2] In agreement, no significant differences in overall survival were found between HOVON (with all patients treated with intensive chemotherapy) and intensively treated MLL patients (Figure S9A). The experimental drugs included in HOVON studies (Clofarabine, Laromustine and G-SCF) did not improve treatment outcome, corresponding previously reported results (more details regarding treatment protocols can be found in the methods section).[3-5]

Altogether, this cohort of 1447 patients, of which 1223 were treated with intense chemotherapy, allowed us to address our research questions, including the assessment of effects of relatively low frequency events such as SF mutations in subgroups of patients, and maximized the benefits to be derived from existing cohort studies.

CO-OCCURING SF MUTATIONS

Nine patients in the analyzed cohort had coinciding mutations in two different SF genes (Supplemental Table S5). In almost all cases a common (recurrent) allele for a particular SF was paired with a less common allele of the second SF. Furthermore, these patients were almost exclusively over 65 years of age and in majority presented with mutations in *ASXL1*. Interestingly, a recent study uncovered that less common alleles with reduced effects on alternative splicing were enriched in patients with double SF mutations as compared to single mutants.[2]

THE INFLUENCE OF GENE INTERACTIONS ON SURVIVAL IN YOUNGER AML PATIENTS

It was previously shown that in AML patients younger than 60 years harboring *RUNX1* mutations had an independent negative prognostic value.[6] Therefore, we assessed the association of the interaction between SFmut4 (as well as *SRSF2* mutations) and *RUNX1* mutations within this subgroup of patients (Figure S10). Again, the co-occurrence of *SRSF2* or SFmut4 with *RUNX1* mutations was associated with adverse outcome, while the presence of *RUNX1* mutations without SFmut4 or *SRSF2* mutations did not indicate inferior survival. In this subset of patients (with age below 60 years) the observed associations were even stronger than in the total cohort.

REFERENCES

1. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. New England Journal of Medicine. 2016;374(23):2209-2221. doi:10.1056/NEJMoa1516192
2. Metzeler KH, Herold T, Rothenberg-thurley M, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood. 2016;128(5):686-699.
3. Lowenberg, B. et al. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. Blood. 2017;129(12): 1636-1645.
4. Lowenberg, B. et al. Effect of priming with granulocyte colony-stimulating factor on the outcome of chemotherapy for acute myeloid leukemia. New England Journal of Medicine. 2003;349(8):743-52.
5. Jongen-Lavrencic M, Grob T, Hanekamp D, et al. Molecular Minimal Residual Disease in Acute Myeloid Leukemia. New England Journal of Medicine. 2018;378(13):1189-1199.
6. Taylor, J. et al. Single-cell genomics reveals the genetic and molecular bases for escape from mutational epistasis in myeloid neoplasms. Blood. 2020;doi:10.118.

SUPPLEMENTAL TABLES

Table S1. Additional characteristics of patients treated with intensive chemotherapy.

Prior Disease, n (%)	
None	1079 (88.2)
Hematological Disorder	68 (5.6)
Other Cancer	53 (4.3)
Other Disease	1 (0.1)
Missing	22 (1.8)
Prior Treatment, n (%)	
None	1139 (93.1)
Chemo- or Radiotherapy	60 (4.9)
Missing	24 (2.0)

FAB, French American British Classification; RAEB, Refractory Anemia with Excess Blasts.

Table S2. The frequencies of SF mutations.

Gene	Evaluated cases	Positive cases	Negative Cases	Percentage Positive (among evaluated)	Percentage Positive (among total cases)
SF1	558	2	556	0.36	0.14
SF3A1	558	1	557	0.18	0.07
SF3B1	1447	41	1406	2.83	2.83
SRSF2	1447	148	1299	10.23	10.23
U2AF1	988	31	957	3.14	2.14
U2AF2	558	2	556	0.36	0.14
ZRSR2	988	15	973	1.52	1.04
SFmut7	661	231	430	34.90	15.96
SFmut4	1039	229	810	22.00	15.80

Table S3 and S4 can be found online at https://ashpublications.org/bloodadvances/article/5/17/3254/476647/Splicing-factor-gene-mutations-in-acute-myeloid

Table S5. Characteristics of patients with co-occurring SF mutations at diagnosis.

Sample ID	Gender	Age	SF Mutations (VAF)	Other Mutations (VAF)	Cytogenetics and Karyotype
#1	M	67,9	*SRSF2* P95H (0.4693) *SF3B1* T871I(0.3964)	*IDH2* (0.3885), *RUNX1* (0.1513), *SMC3* (0.4646)	Trisomy 11
#2	M	75,1	*SRSF2* P95H (0.5407) *SF1* G497S (0.4834)	*ASXL1* (0.5192), *NRAS* (0.1724), *TET2* (0.9794)	Normal Karyotype
#3	F	84	*SRSF2* P95H (0.5229) *SF1* K341N (0.4877)	*ASXL1* (0.2857), *TET2* (0.4154), *WT1* (0.3875)	Trisomy 8
#4	M	78,3	*SRSF2* P95H (0.1053) *ZRSR2* R169X (0.2449)	*CUX1* (0.092), *IDH1**, *TET2* (0.381)	Normal Karyotype
#5	F	63,3	*SF3B1* K666N (0.4567) *ZRSR2* G438R442dup (0.4713)	*NRAS* (VAF 0.0989)	Monosomy 7
#6	M	80,3	*SF3B1* D781Q (0.3364) *U2AF2* - (0.1132)	*ASXL1* (0.4302), *RUNX1* (0.2804)	Normal Karyotype
#7	M	80,1	*U2AF1* Q157P (0.2595) *ZRSR2* Q120RfsX10 (0.7391)	*ASXL1* (0.4145), *KRAS* (0.0693)	Del7q
#8	M	64,6	*U2AF1* Q157P (0.4142) *ZRSR2* S445R448dup (1)	*ASXL1* (0.4057), *BCOR* (0.0506), *EZH2* (0.4828), *FLT3*^{ITD} (AR >0.5), *RUNX1* (0.4706), *SMC1A* (0.8116), WT1 (0.5263)	Trisomy 8
#9*	F	45	*SF3B1* - *U2AF1* -	*RUNX1, NRAS, STAG2*	Trisomy 8

Note: Copy number variation data was not included in the analysis and therefore VAFs should be interpreted with caution.

* - no VAF data available; Male; F, Female; VAF, Variant Allele Frequency.

Table S6-13 can be found online at https://ashpublications.org/bloodadvances/article/5/17/3254/476647/Splicing-factor-gene-mutations-in-acute-myeloid

SUPPLEMENTAL FIGURES

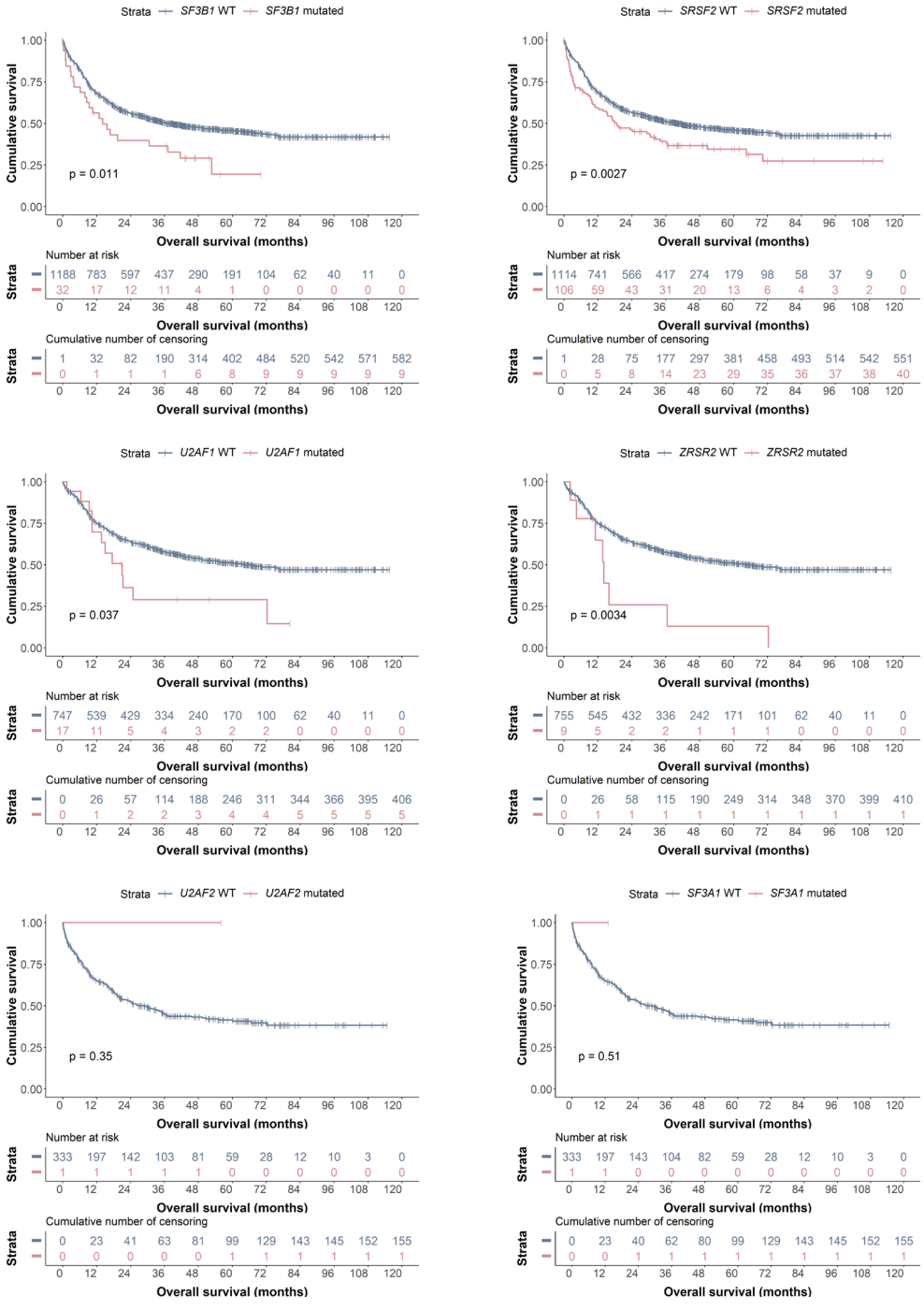


Figure S1. Overall survival of AML patients in the relation to the presence of individual SF mutations. Kaplan-Meier curves for overall survival in relation to the mutation status of *SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *SF3A1* or *U2AF2* are depicted.

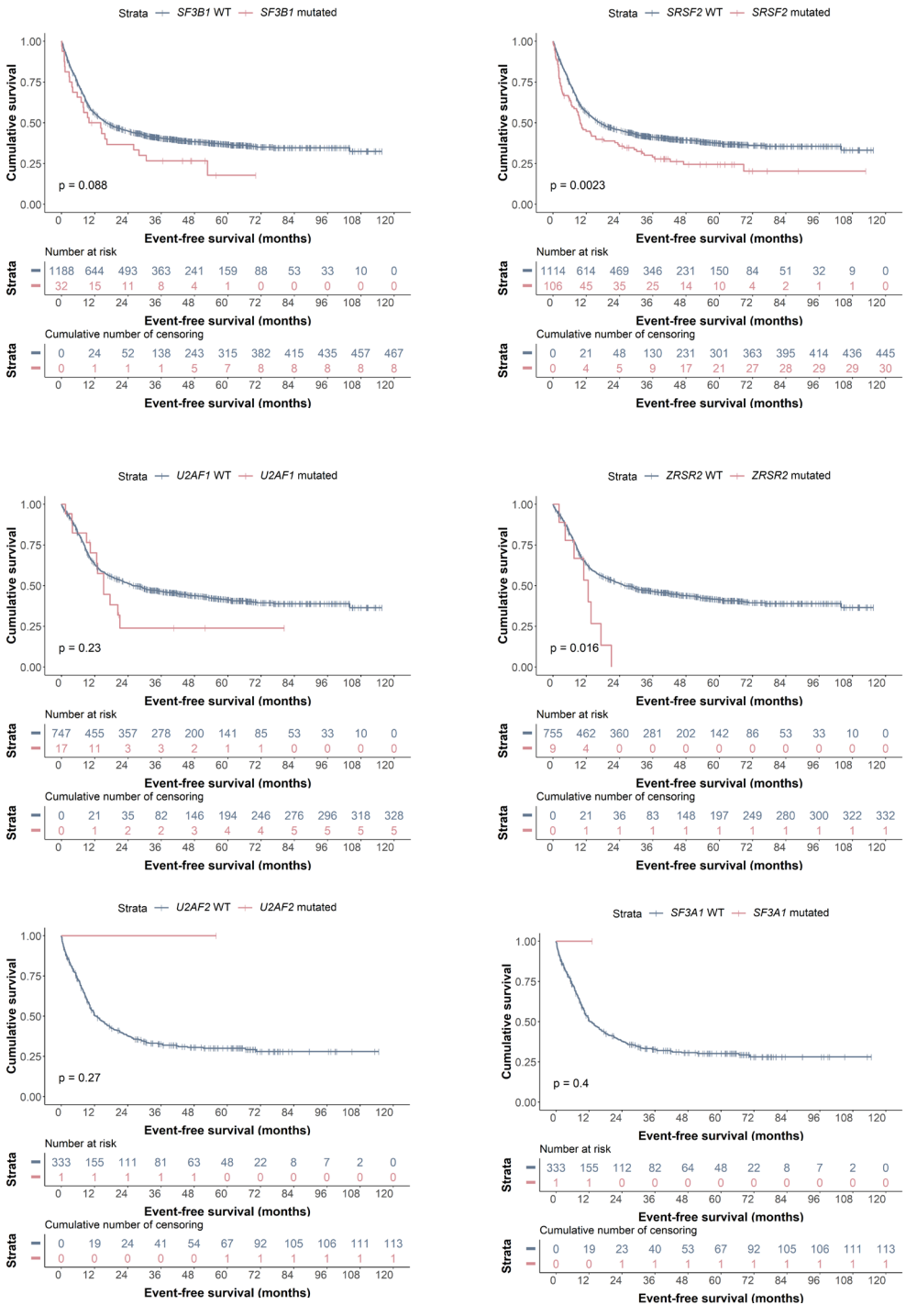
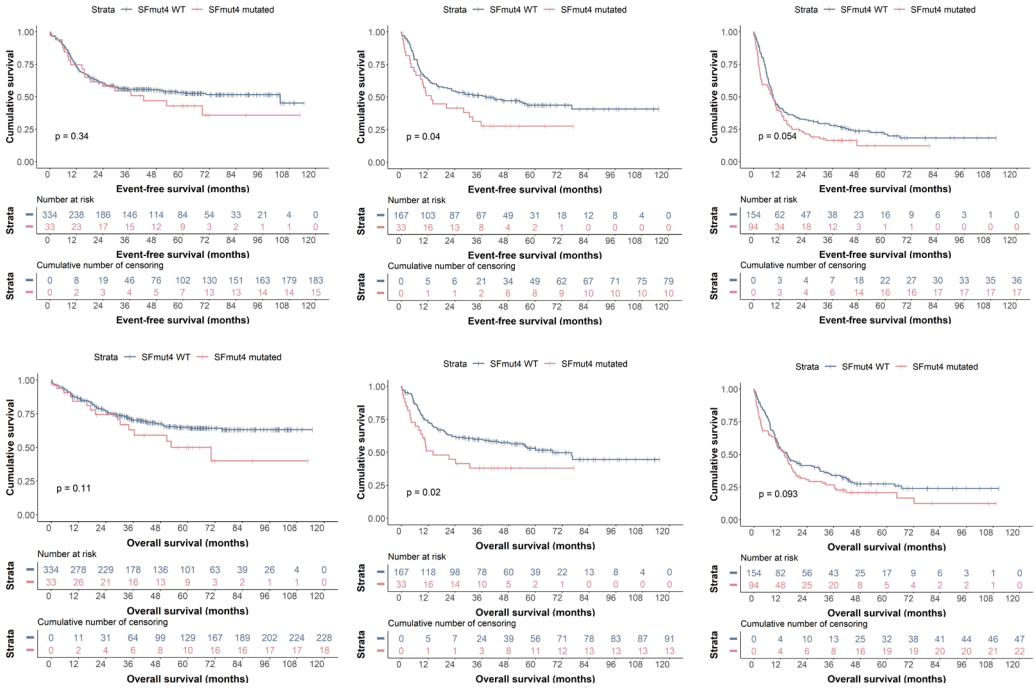
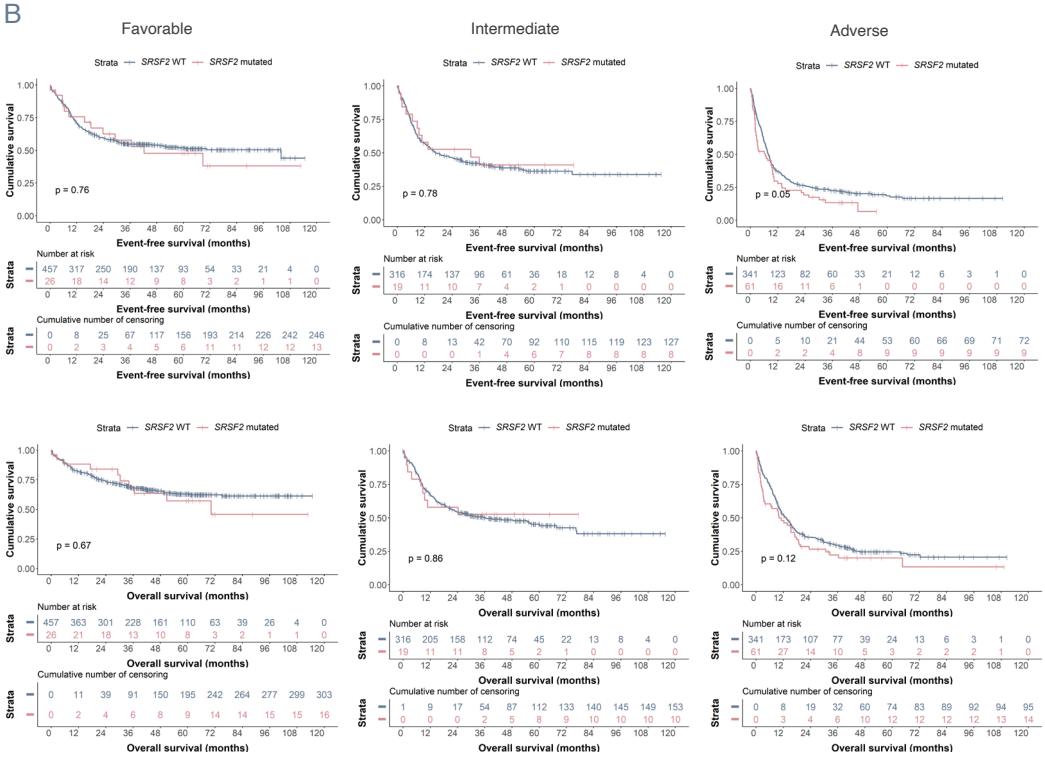


Figure S2. Event-free survival of AML patients in the relation to the presence of individual SF mutations. Kaplan-Meier curves for event-free survival in relation to the mutation status of *SRSF2*, *SF3B1*, *U2AF1*, *ZRSR2*, *SF3A1* or *U2AF2* are depicted.

A


Favorable
Intermediate
Adverse
Strata SFmut4 WT SFmut4 mutated
Cumulative survival
p = 0.34
p = 0.04
p = 0.054
Event-free survival (months)
Number at risk
Cumulative number of censoring
p = 0.11
p = 0.02
p = 0.093
Overall survival (months)

B


Favorable
Intermediate
Adverse
Strata SRSF2 WT SRSF2 mutated
Cumulative survival
p = 0.76
p = 0.78
p = 0.05
Event-free survival (months)
Number at risk
Cumulative number of censoring
p = 0.67
p = 0.86
p = 0.12
Overall survival (months)

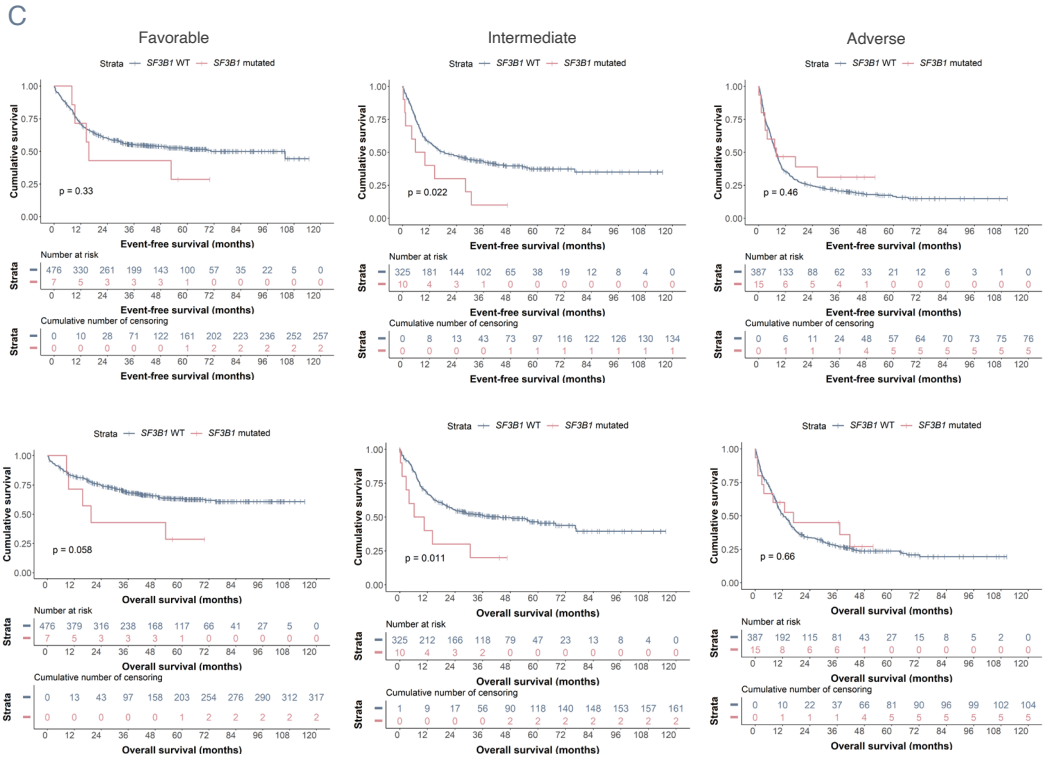
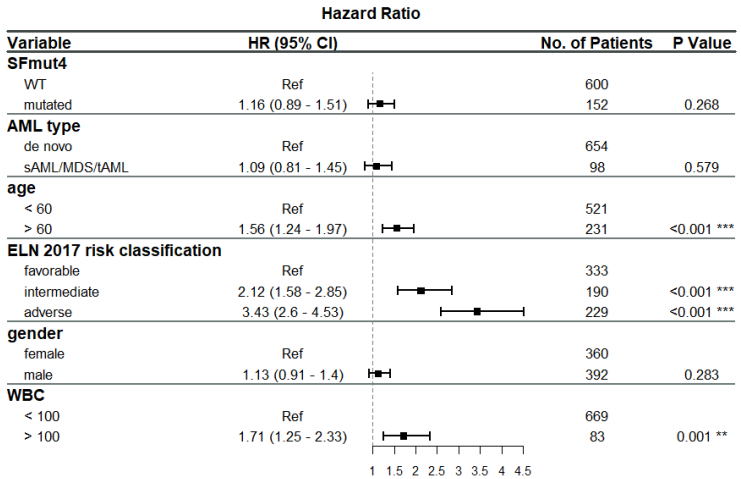


Figure S3. The influence of SF mutations on survival within ELN risk groups. The figure depicts Kaplan-Meier curves for event-free survival and overall survival in relation to SFmut4 status (A), mutation status of *SRSF2* (B) or mutation status of *SF3B1* (C) within the favorable, intermediate or adverse risk groups as defined by the ELN 2017 classification.

A



B

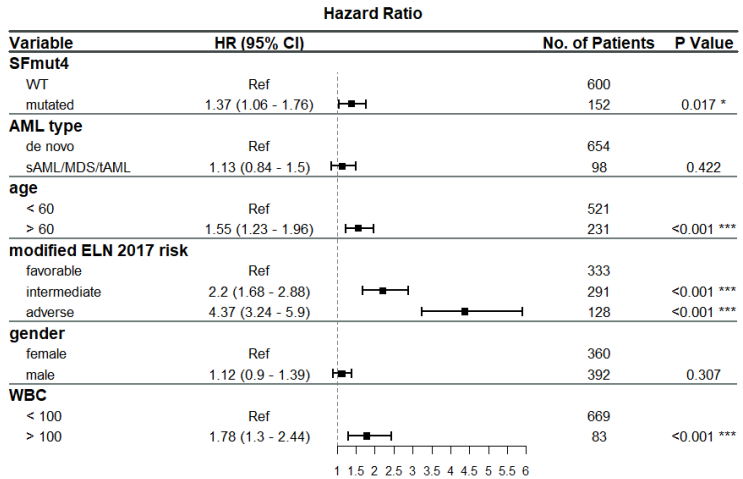


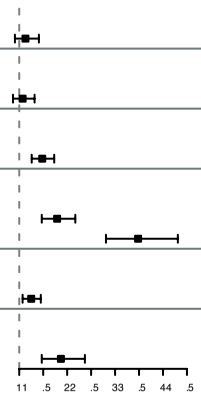
Figure S4. Multivariable analysis of overall survival of AML patients in relation to the presence of SF mutations. A,B. - Multivariable Cox regression analysis of overall survival in relation to SFmut4 with complete (A) and modified (B) ELN 2017 classification. In the modified ELN 2017 classification *RUNX1* and *ASXL1* mutations were excluded, so that patients carrying *RUNX1* or *ASXL1* mutations were re-classified based on the presence of the rest of aberrations in this classification system. Type of stem cell transplantation violated the proportional hazard assumption and therefore it was used as strata variable in all the multivariable Cox regression models (for variables used in the model as strata the statistics are not calculated and therefore, they do not appear in the results). WBC – white blood cell count; sAML – secondary AML, tAML - treatment-related AML.

Figure S5. Survival of AML patients in relation to the presence of a mutation in *RUNX1* or *ASXL1* (see next page). A - Multivariable Cox regression analysis of event-free (left) or overall survival (right) in relation to mutations in *RUNX1* including modified ELN 2017 classification. B - Multivariable Cox regression analysis of event-free (left) or overall survival (right) in relation to mutations in *ASXL1* including modified ELN 2017 classification. In the modified ELN 2017 classification *RUNX1* and *ASXL1* mutations were excluded, so that patients carrying *RUNX1* or mutations were re-classified based on the presence of the rest of aberrations in this classification system. In both models type of stem cell transplantation was included as strata. WBC – white blood cell count; sAML – secondary AML, tAML - treatment-related AML.

A

Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
RUNX1			
WT	Ref	1022	
mutated	1.13 (0.91 – 1.41)	133	0.264
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.07 (0.87 – 1.32)	171	0.497
age			
< 60	Ref	799	
> 60	1.48 (1.26 – 1.73)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	1.79 (1.47 – 2.17)	461	<0.001 ***
adverse	3.48 (2.81 – 4.31)	246	<0.001 ***
gender			
female	Ref	532	
male	1.25 (1.07 – 1.45)	623	0.004 **
WBC			
< 100	Ref	1044	
> 100	1.87 (1.47 – 2.37)	111	<0.001 ***



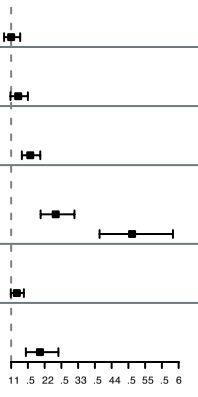
Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
ASXL1			
WT	Ref	1040	
mutated	1.02 (0.8 – 1.29)	115	0.895
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.08 (0.88 – 1.33)	171	0.473
age			
< 60	Ref	799	
> 60	1.48 (1.26 – 1.75)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	1.82 (1.5 – 2.21)	461	<0.001 ***
adverse	3.53 (2.85 – 4.37)	246	<0.001 ***
gender			
female	Ref	532	
male	1.25 (1.07 – 1.45)	623	0.005 **
WBC			
< 100	Ref	1044	
> 100	1.86 (1.47 – 2.37)	111	<0.001 ***

B

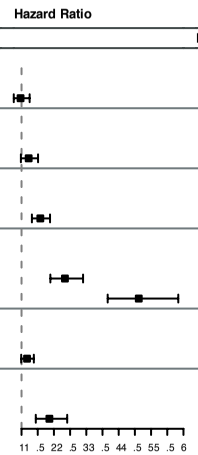
Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
RUNX1			
WT	Ref	1022	
mutated	1 (0.79 – 1.27)	133	0.98
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.21 (0.98 – 1.5)	171	0.079
age			
< 60	Ref	799	
> 60	1.57 (1.32 – 1.87)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	2.33 (1.88 – 2.89)	461	<0.001 ***
adverse	4.61 (3.64 – 5.83)	246	<0.001 ***
gender			
female	Ref	532	
male	1.17 (0.99 – 1.38)	623	0.063
WBC			
< 100	Ref	1044	
> 100	1.86 (1.44 – 2.41)	111	<0.001 ***



Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
ASXL1			
WT	Ref	1040	
mutated	0.97 (0.76 – 1.25)	115	0.823
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.22 (0.98 – 1.51)	171	0.076
age			
< 60	Ref	799	
> 60	1.58 (1.32 – 1.88)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	2.34 (1.89 – 2.9)	461	<0.001 ***
adverse	4.62 (3.66 – 5.84)	246	<0.001 ***
gender			
female	Ref	532	
male	1.17 (0.99 – 1.38)	623	0.061
WBC			
< 100	Ref	1044	
> 100	1.86 (1.44 – 2.41)	111	<0.001 ***



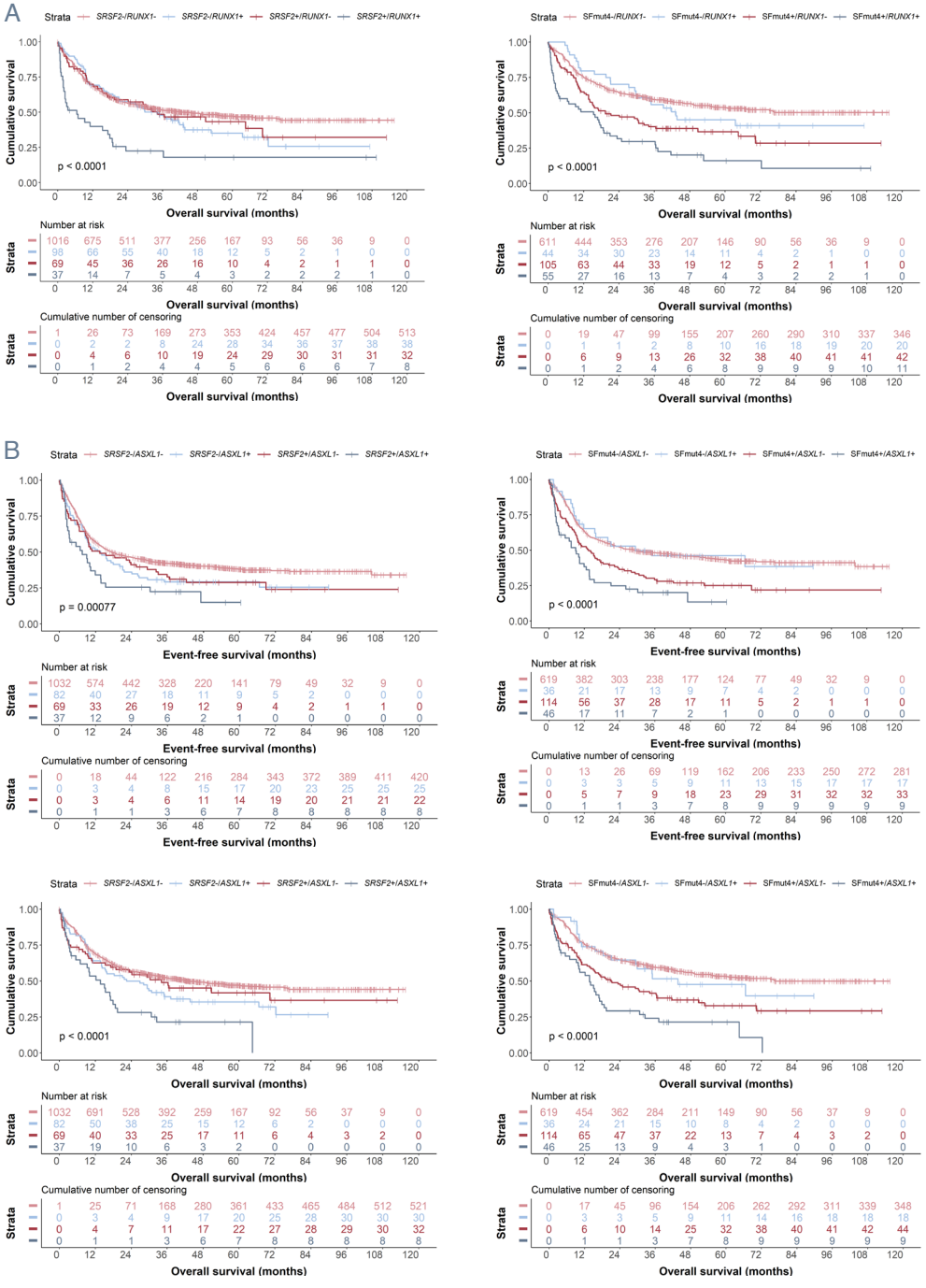


Figure S6. Influence of interactions between SF mutations and *RUNX1* as well as *ASXL1* mutations on overall survival. A - Kaplan Meier curves for overall survival in relation to the mutation status of *SRSF2* (left) or SFmut4 (right) in combination with *RUNX1* mutations. B - Kaplan Meier curves for event-free and overall survival based on the mutation status of *SRSF2* (left) or SFmut4 (right) in combination with *ASXL1* mutations.

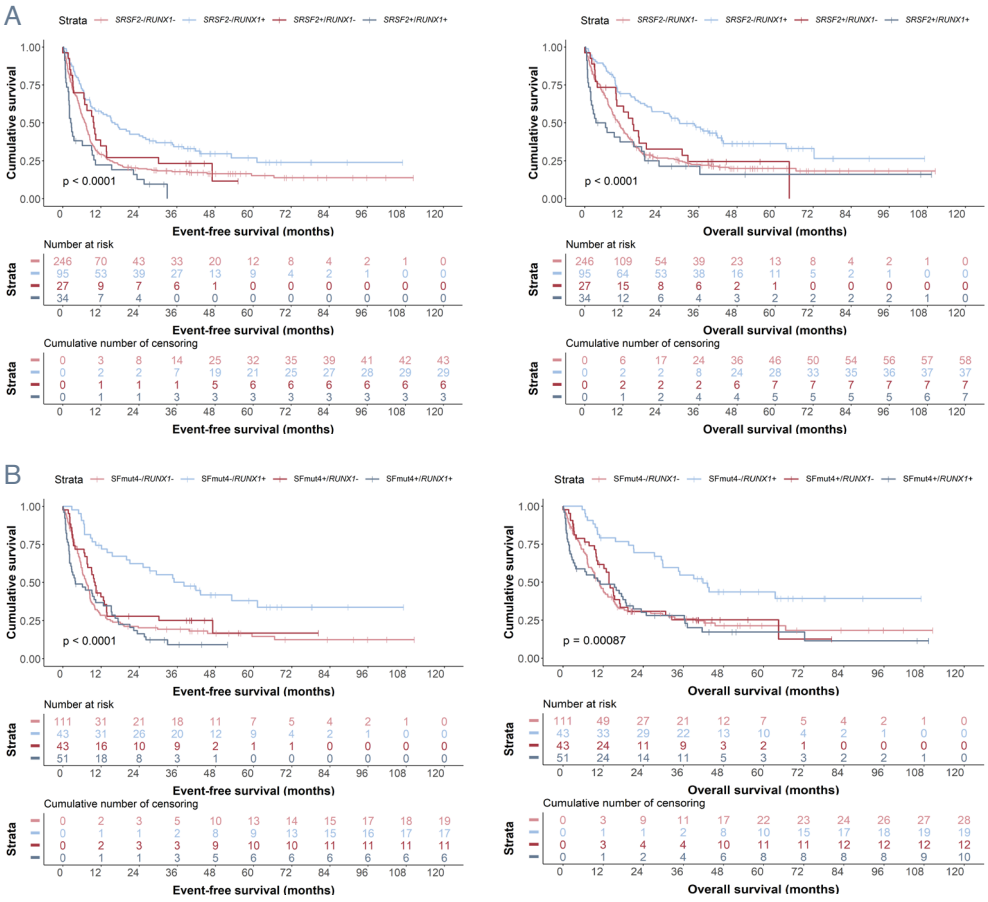
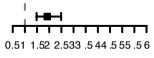


Figure S7. Analysis of survival in relation to interactions of SF mutations with *RUNX1* and *ASXL1* mutations in adverse risk category according to ELN 2017 classification. A - Kaplan Meier curves for event-free survival (left) and overall survival (right) in relation to the mutation status of *SRSF2* and *RUNX1* within the adverse ELN 2017 risk group. B - Kaplan Meier curves for event-free survival (left) and overall survival (right) in relation to the SFmut4 and *RUNX1* mutations.

A

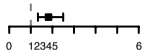
Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
***SRSF2* and *RUNX1* interaction**			
SRSF2–/*RUNX1*–	Ref	957	
SRSF2–/*RUNX1*+	0.81 (0.61 – 1.07)	96	0.14
SRSF2+/*RUNX1*–	0.83 (0.58 – 1.19)	65	0.302
SRSF2+/*RUNX1*+	1.75 (1.18 – 2.6)	37	0.005 **
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.24 (1 – 1.53)	171	0.05
age			
< 60	Ref	799	
> 60	1.59 (1.33 – 1.9)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	2.32 (1.86 – 2.88)	461	<0.001 ***
adverse	4.74 (3.74 – 6)	246	<0.001 ***
gender			
female	Ref	532	
male	1.17 (0.99 – 1.38)	623	0.07
WBC			
< 100	Ref	1044	
> 100	1.92 (1.48 – 2.49)	111	<0.001 ***



Hazard Ratio

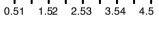
Variable	HR (95% CI)	No. of Patients	P Value
SFmut4 and *RUNX1* interaction			
SFmut4–/*RUNX1*–	Ref	557	
SFmut4–/*RUNX1*+	0.76 (0.49 – 1.18)	43	0.217
SFmut4+/*RUNX1*–	1.14 (0.83 – 1.55)	98	0.419
SFmut4+/*RUNX1*+	1.72 (1.2 – 2.46)	54	0.003 **
AML type			
de novo	Ref	654	
sAML/MDS/tAML	1.18 (0.88 – 1.58)	98	0.265
age			
< 60	Ref	521	
> 60	1.58 (1.25 – 1.99)	231	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	333	
intermediate	2.21 (1.68 – 2.9)	291	<0.001 ***
adverse	4.44 (3.28 – 6.01)	128	<0.001 ***
gender			
female	Ref	360	
male	1.11 (0.89 – 1.38)	392	0.359
WBC			
< 100	Ref	669	
> 100	1.82 (1.33 – 2.5)	83	<0.001 ***



B

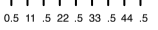
Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
***SRSF2* and *ASXL1* interaction**			
SRSF2–/*ASXL1*–	Ref	975	
SRSF2–/*ASXL1*+	0.97 (0.73 – 1.29)	78	0.824
SRSF2+/*ASXL1*–	1.17 (0.85 – 1.59)	65	0.337
SRSF2+/*ASXL1*+	1.18 (0.8 – 1.75)	37	0.408
AML type			
de novo	Ref	984	
sAML/MDS/tAML	1.08 (0.88 – 1.33)	171	0.455
age			
< 60	Ref	799	
> 60	1.47 (1.25 – 1.73)	356	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	448	
intermediate	1.8 (1.48 – 2.18)	461	<0.001 ***
adverse	3.56 (2.87 – 4.41)	246	<0.001 ***
gender			
female	Ref	532	
male	1.23 (1.05 – 1.43)	623	0.009 **
WBC			
< 100	Ref	1044	
> 100	1.86 (1.47 – 2.36)	111	<0.001 ***



Hazard Ratio

Variable	HR (95% CI)	No. of Patients	P Value
SFmut4 and *ASXL1* interaction			
SFmut4–/*ASXL1*–	Ref	568	
SFmut4–/*ASXL1*+	0.73 (0.43 – 1.24)	32	0.244
SFmut4+/*ASXL1*–	1.26 (0.96 – 1.64)	106	0.091
SFmut4+/*ASXL1*+	1.36 (0.92 – 1.99)	46	0.12
AML type			
de novo	Ref	654	
sAML/MDS/tAML	0.99 (0.75 – 1.32)	98	0.966
age			
< 60	Ref	521	
> 60	1.55 (1.25 – 1.91)	231	<0.001 ***
modified ELN 2017 risk			
favorable	Ref	333	
intermediate	1.61 (1.26 – 2.05)	291	<0.001 ***
adverse	3.23 (2.46 – 4.26)	128	<0.001 ***
gender			
female	Ref	360	
male	1.22 (1 – 1.5)	392	0.048 *
WBC			
< 100	Ref	669	
> 100	1.83 (1.38 – 2.44)	83	<0.001 ***



C

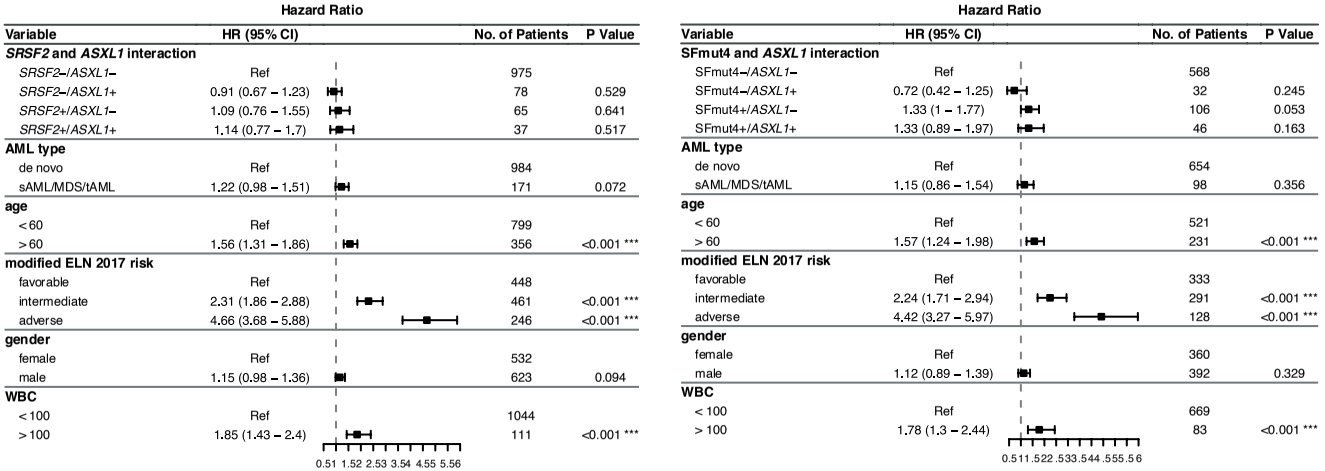


Figure S8. Multivariable analysis of survival of AML patients in relation to the interaction of SF mutations with mutations in *RUNX1* or *ASXL1*. A - Multivariable Cox regression analysis of overall survival in relation to the mutation status of *SRSF2* or SFmut4 and *RUNX1* including modified ELN 2017 classification. B,C – Multivariable Cox regression analysis of event-free (B) and overall survival (C) in relation to the mutation status of *SRSF2* or SFmut4 and *ASXL1* including modified ELN 2017 classification. In the modified ELN 2017 classification *RUNX1* and *ASXL1* mutations were excluded, so that patients carrying *RUNX1* or *ASXL1* mutations were re-classified based on the presence of the rest of aberrations in this classification system. Type of stem cell transplantation violated the proportional hazard assumption and therefore it was used as strata variable in all the multivariable Cox regression models (for variables used in the model as strata the statistics are not calculated and therefore, they do not appear in the results). WBC – white blood cell count; sAML – secondary AML, tAML - treatment-related AML.

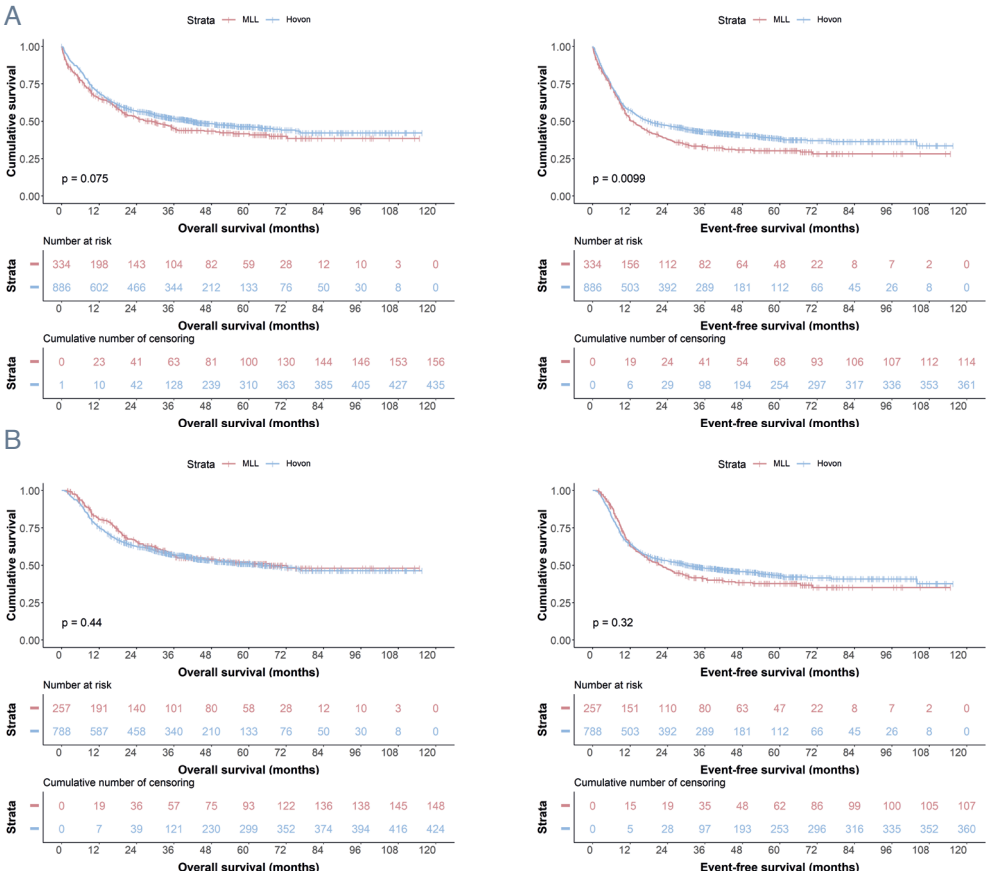
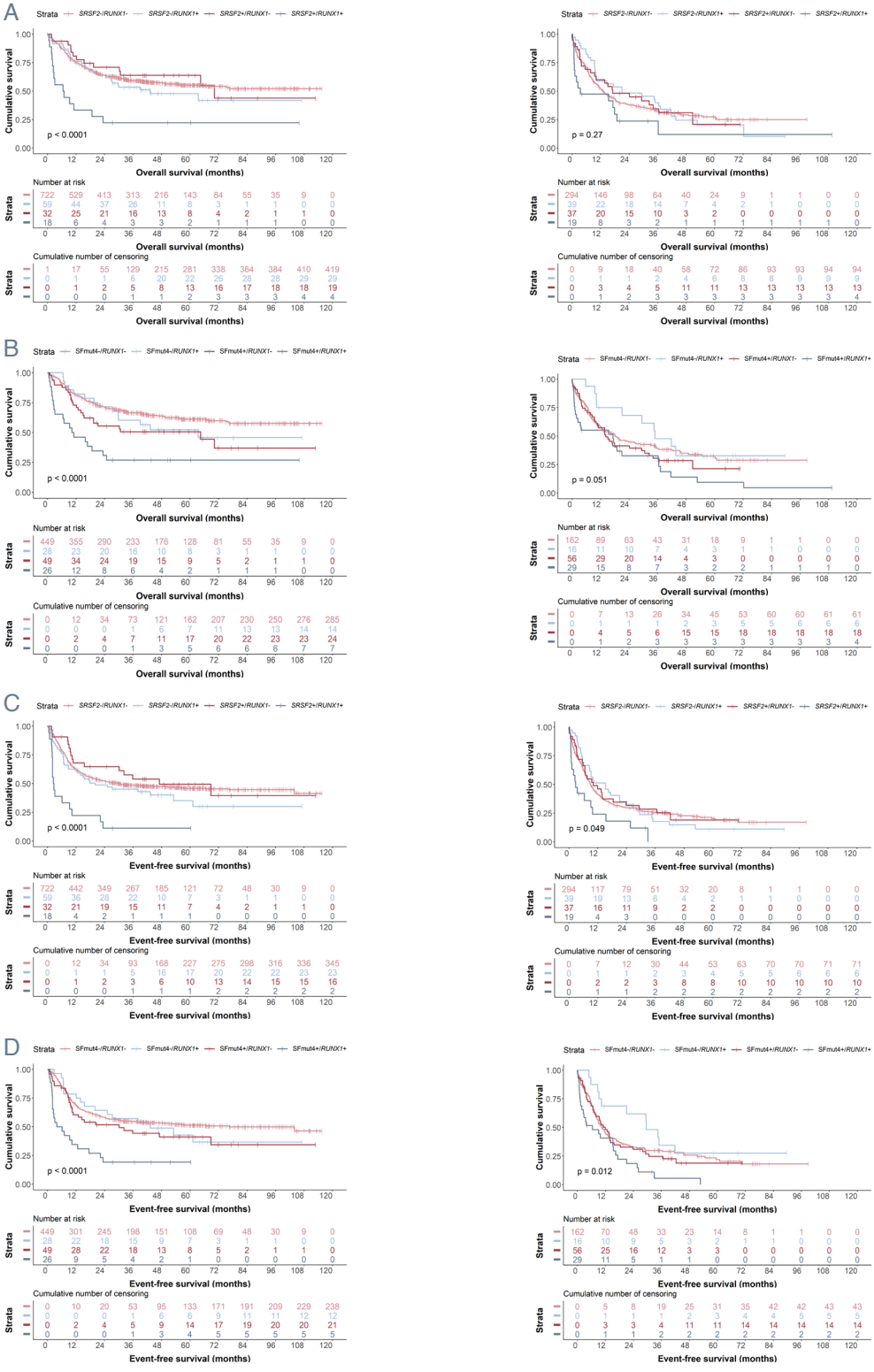


Figure S9. Survival analysis of our combined cohort. The figure depicts Kaplan-Meier curves for overall and event free survival in all HOVON and MLL patients (A) and patients with complete remission only (B).

Figure S10. Survival of young and old AML patients in relation to the presence of SF mutations and *RUNX1* mutations (see next page). A - Kaplan Meier curves for overall survival in relation to the mutation status of *SRSF2* in combination with *RUNX1* mutations in AML patients younger than 60 years (left), or 60 years and older (right). B - Kaplan Meier curves for overall survival in relation to the presence of SFmut4 in combination with *RUNX1* mutations in AML patients younger than 60 years (left), or 60 years and older (right). C - Kaplan Meier curves for event-free survival in relation to the mutation status of *SRSF2* in combination with *RUNX1* mutations in AML patients younger than 60 years (left), or 60 years and older (right). D - Kaplan Meier curves for event-free survival in relation to the presence of SFmut4 in combination with *RUNX1* mutations in AML patients younger than 60 years (left), or 60 years and older (right).


A
B
C
D
Strata
SRSF2-/RUNX1-
SRSF2-/RUNX1+
SRSF2+/RUNX1-
SRSF2+/RUNX1+
SFmut4-/RUNX1-
SFmut4-/RUNX1+
SFmut4+/RUNX1-
SFmut4+/RUNX1+
Cumulative survival
p < 0.0001
p = 0.27
p = 0.051
p = 0.049
p = 0.012
Overall survival (months)
Event-free survival (months)
Number at risk
Cumulative number of censoring

CHAPTER 4

SF3B1 as Therapeutic Target in *FLT3*/ITD Positive Acute Myeloid Leukemia

Inge van der Werf[1]
Anna Wojtuszkiewicz[1]
Huilan Yao[2]
Rocco Sciarrillo[1]
Manja Meggendorfer[3]
Stephan Hutter[3]
Wencke Walter[3]
Jeroen Janssen[1]
Wolfgang Kern[3]
Claudia Haferlach[3]
Torsten Haferlach[3]
Gerrit Jansen[4]
Gertjan J. L. Kaspers[5,6]
Richard Groen[1]
Gert Ossenkoppele[1]
Jacqueline Cloos[1]

1. Dept. of Hematology, Amsterdam University Medical Center, VU University Medical Center, Cancer Center Amsterdam, Amsterdam, The Netherlands; 2. H3 Biomedicine, Inc. Boston, MA, USA; 3. MLL Munich Leukemia Laboratory, Munich, Germany; 4. Amsterdam Rheumatology and Immunology Center, Amsterdam University Medical Center, VU University Medical Center, Amsterdam; 5. Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands 6. Department of Pediatric Oncology, Emma Children's Hospital Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Leukemia. 2021. 35(9):2735

TO THE EDITOR

Recently, significant advances have been made in the development of splicing modulators for therapeutic purposes. In this respect, several studies demonstrated that acute myeloid leukemia (AML) cells carrying spliceosome mutations are preferentially sensitive to Splicing Factor 3B subunit 1 (SF3B1) modulation.[1-3] Whereas approximately 15% of AML patients have mutations in this class of genes, disruption of splicing appears to be a global phenomenon in hematological malignancies.[4] Therefore, we aim to identify additional patient subgroups which will benefit from these emerging modulators.

Towards this goal, we assessed the response to splicing modulation in a set of AML samples with different molecular aberrations. This included both cell lines as well as cryopreserved mononuclear cells isolated from diagnostic bone marrow samples. Cells were exposed to the SF3B1 modulator E7107, a stable analogue of Pladienolide B, and the newly developed H3B-8800. Subsequently, *in vitro* and *ex vivo* cytotoxicity measurements, functional assays and differential expression analyses were performed (supplementary materials and methods).

While growth of all cell lines was inhibited at nanomolar concentrations, cell lines carrying an internal tandem duplication (ITD) in the FMS-like tyrosine kinase 3 (*FLT3*) gene were particularly sensitive (Supplemental Figure S1A). This effect in *FLT3*/ITD^{pos} cells was related to induction of cell cycle arrest (Supplemental Figure S1B) with a concomitant shift in the splicing pattern of *MCL1* towards its pro-apoptotic variant (Supplemental Figure S1C-D) upon relatively low drug concentrations compared to *FLT3*/ITD^{neg} cells. *FLT3*/ITD is a recurrent aberration in AML, which results in activation of downstream signaling pathways involved in proliferation, differentiation and apoptosis.[5] To date, patients with a high allelic ratio (AR; mutated over wild type) have been classified as high risk and are eligible for FLT3 inhibitors such as midostaurin. However, this drug only modestly improved survival in *FLT3*/ITD^{pos} AML patients[6], hence, there is still an unmet need for new drugs to treat this aggressive type of leukemia.

Our hypothesis of preferential sensitivity to SF3B1 modulation in *FLT3*/ITD^{pos} samples was further tested in an *ex vivo* setting. For this purpose, cells taken at diagnosis were selected based on their *FLT3* mutation status and the absence of other aberrations based on our diagnostic panel (Supplemental Table S1). Corroborating *in vitro* results, the total population of white blood cells (WBC), and in particular the CD34+ progenitor population, within the leukemic bone marrow was more affected in *FLT3*/ITD^{pos} samples compared to *FLT3*/ITD^{neg} samples upon incubation with either E7107 or H3B-8800 (Figure 1A and 1B, Supplemental Figure S2A-B). Beside cell count, clonogenicity was found to be affected upon splicing modulation in *FLT3*/ITD^{pos} cells as well (Figure 1C). To verify the selectivity of our findings, cells were simultaneously treated with the proteasome inhibitor Bortezomib, a compound whose mechanistic effects are not splicing dependent. Accordingly, no differences in response between *FLT3*/ITD^{pos} and *FLT3*/ITD^{neg} cells were observed with this compound (Figure 1D).

Next, the response to splicing modulation was reflected in the dose-dependent splicing perturbation of a selected panel of mRNA transcripts (Figure 1E-F, Supplemental Figure S2C; Nanostring). Overall, decreased production of mature mRNA transcripts and concomitant ac-

cumulation of pre-mRNA transcripts seems to be more pronounced in *FLT3*/ITD^{pos} samples suggesting increased target engagement.[7] In addition, splicing of mature mRNA as well as pre-mRNA transcripts correlated with *ex vivo* response confirming enhanced splicing perturbation in *FLT3*/ITD^{pos} cells (Supplemental Figure S2D-E). Importantly, this effect seems to be irrespective of other molecular aberrations (Table S1) due to their absence in *FLT3*/ITD^{pos} samples. Furthermore, response rates to both splicing modulators were highly correlated (Figure 1G) confirming the general high activity of splicing modulators in *FLT3*/ITD^{pos} cells.

In order to further explore the role of the *FLT3*/ITD in the sensitivity to E7107, we evaluated samples with a low (<0.5) or high (>0.5) AR based on the ELN2017 classification.[8,9] CD34+ cells within the leukemic bone marrow of *FLT3*/ITD^{pos} samples with a high AR responded significantly better to splicing modulation as compared to *FLT3*/ITD^{neg} specimens (p=0.03; Figure 2A, Supplemental Figure S3A, Supplemental Table S1). In addition, we found a gradual increase of response with increased AR which suggests a direct link of response to this aberration. The size of the ITD was also demonstrated to have prognostic significance in AML patients.[10] Concordantly, our data showed that the size of the ITD was associated with response (p=0.02; Figure 2B, Supplemental Figure S3B, Supplemental Table 1). Thus, both the AR and ITD length seem to be important determinants of the efficacy of splicing modulation within this specific subgroup of AML patients.

To determine the therapeutic index, we examined total WBC counts within healthy bone marrow upon splicing modulation. These were not affected upon treatment with either E7107 or H3B-8800 (Figure 2C; Supplemental Figure S3C). Accordingly, lymphocytes count within all AML patient samples were not impacted by splicing modulation (Figure 2D; Supplemental Figure S3D). However, the subpopulation of CD34+ cells derived from healthy bone marrow was negatively affected by both splicing modulators (Figure 2E; Supplemental Figure S3E). This finding highlights the importance of selecting particularly sensitive cells, such as *FLT3*/ITD positive leukemia with high AR and/or long ITD length, together with assessing the right dose to avoid toxicity to non-malignant cells.

Having established high sensitivity in *FLT3*/ITD^{pos} patients for SF3B1 modulation, we next hypothesized that the downstream effectors of this aberration could explain the identified differences in sensitivity. Previously, oncogenic MYC activation was shown to confer stress on splicing machinery via upregulation of *PRMT5*, which makes the spliceosome an attractive target in MYC-driven cancers.[11] In *FLT3*/ITD^{pos} cells enrichment of c-MYC gene sets has been reported.[12] In addition, cells with either high *MCL1* or *BCL2A1* expression were shown to be preferentially sensitive to E7107 and elevated *MCL1* levels were demonstrated in *FLT3*/ITD^{pos} cells.[13,14] To assess expression of these genes in *FLT3*/ITD^{pos} and *FLT3*/ITD^{neg} samples we used both qPCR as well as RNAseq data. Increased expression of none of the gene candidates was confirmed in *FLT3*/ITD^{pos} samples in our dataset (Supplemental Figure S4, S5A-B). In fact, we found significantly decreased expression of *MYC* in *FLT3*/ITD^{pos} patients. In addition, we did not identify differential expression of MYC target genes which indicates there are no differences in post-transcriptional regulation either (data not shown). Functionally, we did identify differences in apoptosis upon treatment. We found that *FLT3*/ITD^{pos} cells presented increased markers

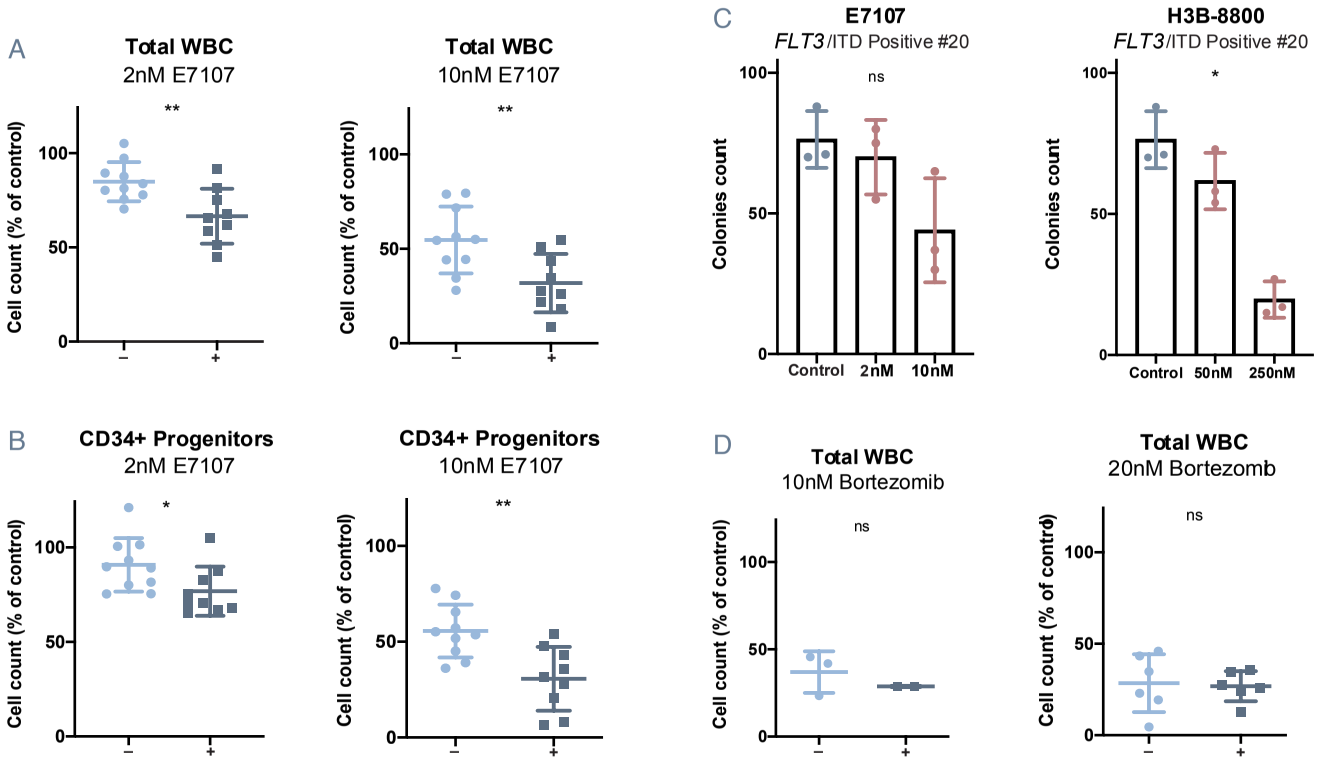


Figure 1. Response of primary AML cells with or without *FLT3*/ITD to splicing modulation. *Ex vivo* AML patients' cells taken at diagnosis were selected based on their *FLT3* mutation status and the absence of any other aberrations based on our diagnostic panel. Light blue dots (-) represent *FLT3*/ITD^{neg} samples whereas dark blue dots (+) represent *FLT3*/ITD^{pos} specimens. A, B - Cells were treated for 48h with indicated doses of E7107. Cell count of the total white blood cell populations (WBC) was determined using Flow Cytometry. Progenitors were identified based on CD34 expression. Cell counts are depicted as per-centage of untreated cells. C - Clonogenic capacity of *FLT3*/ITD^{pos} cells derived from a primary bone marrow sample upon treatment with E7107 or H3B-8800. P-values are based on Kruskas Wallis test. D - Cells were treated for 48h with indicated doses of Bortezomib. Cell count of the total white blood cell populations (WBC) was determined using Flow Cytometry. Cell counts are depicted as percentage of untreated cells. P-values are based on Mann-Whitney U test.

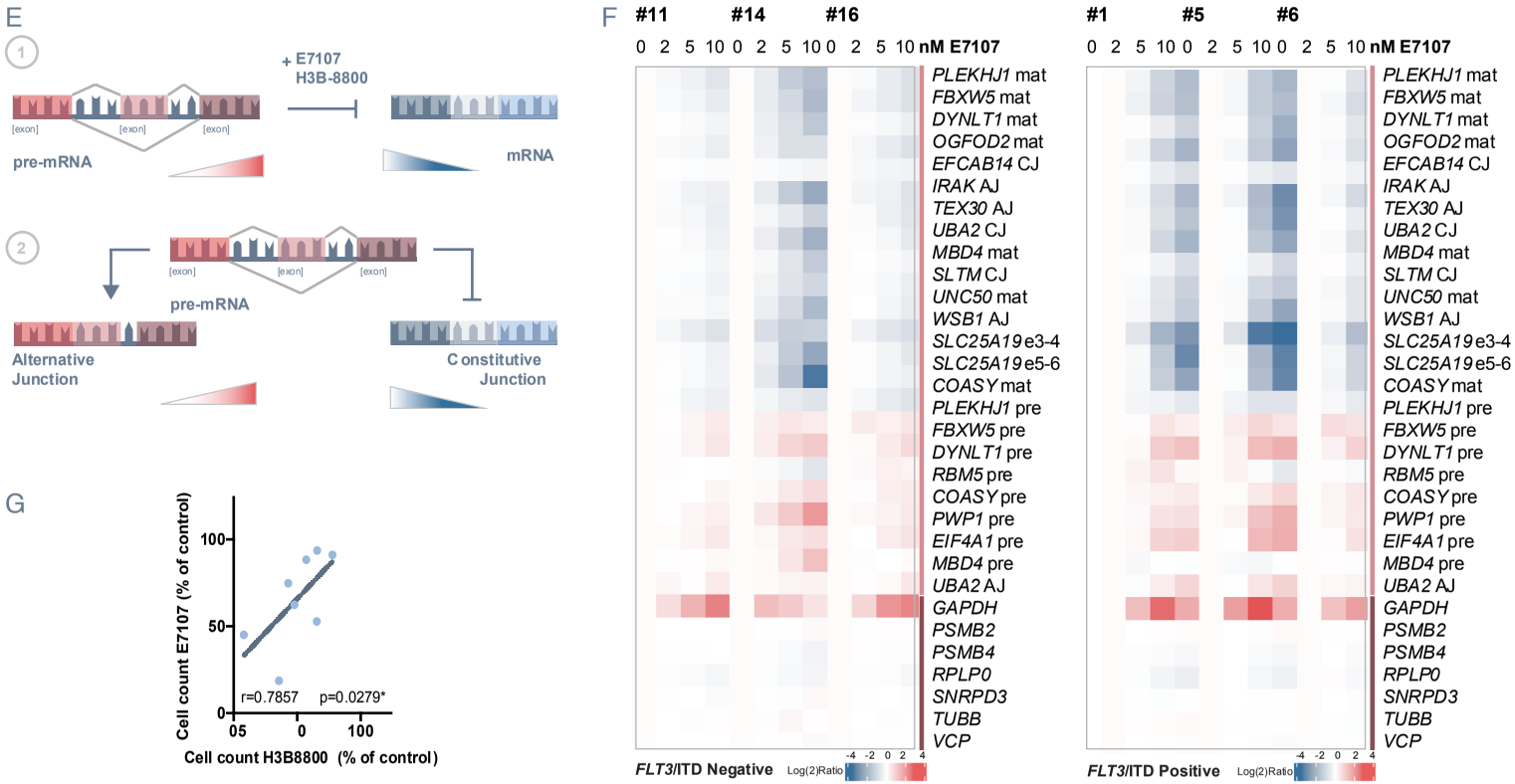


Figure 1. Response of primary AML cells with or without *FLT3*/ITD to splicing modulation (continued). E - Schematic of the effect of splicing modulation on both pre-mRNA as well as mature mRNA levels presented in G. In short, splicing modulation via E7107 or H3B-8800 results in decreased production of mature mRNA transcripts (mat; blue) and concomitant accumulation of pre-mRNA transcripts (pre; red). In addition, alternative junctions (AJ; red) of certain genes are preferred upon splicing modulation resulting in decreased expression of constitutive spliced mRNAs (CJ; blue). F - Representative heatmap of 3 patients samples depicting dose-dependent modulation for both mature as well as pre-mature mRNA markers in *FLT3*/ITD^{neg} (left) and *FLT3*/ITD^{pos} (right) cells. G - Correlation between the response of AML patients' cells to 10nM E7107 or 250nM H3B-8800. A strong association between both splicing modulators as indicated by high Spearman's rho correlation coefficient (R).

of apoptosis upon treatment as well as subtle increase of pro-apoptotic *MCL1-S*, which was correlated with response (Supplemental Figure S5C-E). Yet, this finding seems to be particularly pronounced when using relatively high concentrations of E7107. Therefore, apoptosis induction does not seem to reflect the cause, but rather the consequence of hypersensitivity of *FLT3*/ITD^{pos} cells to splicing modulation.

Notably, we did find significantly increased expression of *SF3B1* in patients with high AR or long ITD length in our dataset (Supplemental Figure S6A), although, *FLT3*/ITD^{pos} patients could not be clustered based on expression of genes involved in splicing regulation (Supplemental Figure S6B). In addition, response rates to SF3B1 modulation were not correlated with *SF3B1* expression levels (Supplemental Figure S6C).

Interestingly, we did observe a drastic decrease of FLT3 RNA and protein levels upon 24h of incubation with E7107 (Supplemental Figure S7). However, downregulation does not seem to be derived from aberrant splicing since we did not identify changes in mRNA length suggesting lack of altered splicing (data not shown). Splicing modulation is known to occur rapidly following treatment[3] while FLT3 downregulation was only detected after 24h, which suggest that the effect on FLT3 expression to be indirect. These findings, together with the that observation that cells with long ITD length and high AR are particularly sensitive to splicing modulation, suggests that hypersensitivity of *FLT3*/ITD^{pos} samples to splicing modulation could rely on FLT3 gene expression regulation and subsequent downstream signaling events. This hypothesis is supported by the finding that long ITD length was associated with higher FLT3 kinase activity.[10] While kinase activity was not assessed in our this study, as downregulation of FLT3 is bound to abrogate FLT3 kinase activity, *FLT3*/ITD positivity was previously linked with increased *ex vivo* response to FLT3 inhibition by gilteritinib supporting the idea of dependency of these cells on FLT3 signaling.[9]

Altogether, this study provides several lines of evidence that splicing modulation holds potential as novel therapeutic option for AML patients carrying *FLT3*/ITD with high AR and/or long ITD length, who currently suffer from poor treatment outcome. Future studies should further validate our findings in isogenic models and uncover the precise role of splicing in *FLT3*/ITD^{pos} patients. Especially SF3B1 expression levels in the context of high AR or long ITD length seem to be of interest.

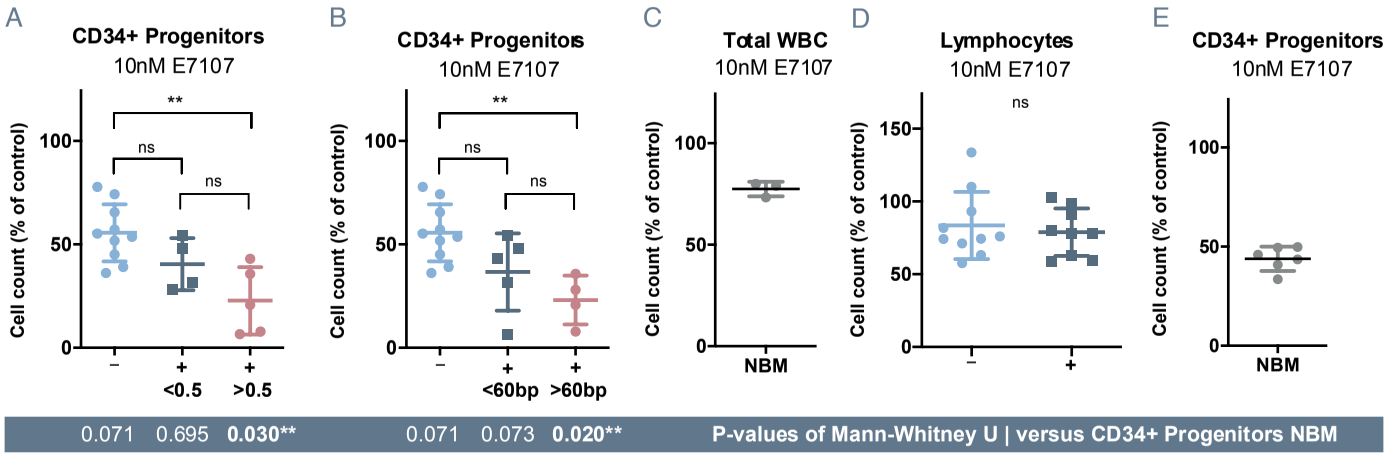


Figure 2. Preferential sensitivity to splicing modulation of *FLT3*/ITD^{pos} AML patients with high AR or long ITD length. A - Cell counts of CD34 positive cells within bone marrows of patient samples grouped based on their *FLT3*/ITD allelic ratio according to the ELN 2017 following treatment with 10nM E7107. B - Cell counts of CD34 positive cells within bone marrows of patient samples grouped on their ITD size. Cut off was based on mean size of ITD length of selected patients. C - Cell count of the total white blood cell population within healthy bone marrow was assessed upon treatment. D - Cell counts of lymphocytes within AML bone marrow samples were identified based on CD45, CD34 and CD7 expression. E - Cell counts of CD34 positive cells within healthy bone marrow samples was determined upon treatment with 10nM E7107 for 48h. Cell counts were plotted as a percentage of untreated cells. The Mann-Whitney U test was used to compare response levels between different groups of patients or normal bone marrow.

REFERENCES

1. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nature Medicine. 2016;22(9):976-986.
2. Buonamici S, Yoshimi A, Thomas M, Seiler M, Chan B, Caleb B, et al. H3B-8800, an Orally Bioavailable Modulator of the SF3b Complex, Shows Efficacy in Spliceosome-Mutant Myeloid Malignancies. Blood. 2016;128(22):966-966.
3. Seiler M, Yoshimi A, Darman R, Chan B, Keaney G, Thomas M, et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nature medicine. 2018;24(4):497-504.
4. Adamia S, Haibe-Kains B, Pilarski PM, Bar-Natan M, Pevzner S, Avet-Loiseau H, et al. A genome-wide aberrant RNA splicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. Clin Cancer Res. 2014;20(5):1135-1145.
5. Stirewalt DL, Radich JP. The role of FLT3 in haematopoietic malignancies. Nat Rev Cancer. 2003;3(9):650-665.
6. Stone RM, Mandrekar SJ, Sanford BL, Laumann K, Geyer S, Bloomfield CD, et al. Midostaurin plus Chemotherapy for Acute Myeloid Leukemia with a FLT3 Mutation. New England Journal of Medicine. 2017;377(5):454-464.
7. Steensma DP, Wermke M, Klimek VM, Greenberg PL, Font P, Komrokji RS, et al. Results of a Clinical Trial of H3B-8800, a Splicing Modulator, in Patients with Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML) or Chronic Myelomonocytic Leukemia (CMML). Blood. 2019;134(Supplement_1):673-673.
8. Döhner H, Estey E, Grimwade D, Amadori S, Appelbaum FR, Büchner T, et al. Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. Blood. 2017;129(4):424-447.
9. Cucchi DGJ, Denys B, Kaspers GJL, Janssen JJWM, Ossenkoppele GJ, de Haas V, et al. RNA-based FLT3-ITD allelic ratio is associated with outcome and ex vivo response to FLT3 inhibitors in pediatric AML. Blood. 2018;131(22):2485-2489.
10. Liu S-B, Dong H-J, Bao X-B, Qiu Q-C, Li H-Z, Shen H-J, et al. Impact of FLT3-ITD length on prognosis of acute myeloid leukemia. Haematologica. 2019;104(1):e9-e12.
11. Hsu TY, Simon LM, Neill NJ, Marcotte R, Sayad A, Bland CS, et al. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature. 2015;525(7569):384-388.
12. Li L, Osdal T, Ho Y, Chun S, McDonald T, Agarwal P, et al. SIRT1 activation by a c-MYC oncogenic network promotes the maintenance and drug resistance of human FLT3-ITD acute myeloid leukemia stem cells. Cell Stem Cell. 2014;15(4):431-446.
13. Yoshimoto G, Miyamoto T, Jabbarzadeh-Tabrizi S, Iino T, Rocnik JL, Kikushige Y, et al. FLT3-ITD up-regulates MCL-1 to promote survival of stem cells in acute myeloid leukemia via FLT3-ITD-specific STAT5 activation. Blood. 2009;114(24):5034-5043.
14. Aird D, Teng T, Huang C-L, Pazolli E, Banka D, Cheung-Ong K, et al. Sensitivity to splicing modulation of BCL2 family genes defines cancer therapeutic strategies for splicing modulators. Nature Communications. 2019;10(1).

SUPPLEMENTAL METHODS

LEUKEMIC CELLS

The human acute myeloid leukemia cell lines Molm13, Kasumi-1 and MM6 were obtained from DSMZ (Braunschweig, Germany) while MV4-11, KG1 and THP1 were obtained from ATCC (Manassas, Virginia, USA). All cell lines were maintained in RPMI-1640 medium (Gibco, Carlsbad, CA, USA) with either 10% or 20% fetal calf serum (FCS; Greiner Bio-One, Frickenhausen, Germany), 100 units/mL penicillin G (Gibco) and 100 ug/mL streptomycin sulphate (Gibco). Cryopreserved mononuclear cells isolated from diagnostic bone marrow samples were obtained from patients in clinical trials of the Dutch-Belgian Cooperation Trial Group for Hematology-Oncology (HOVON) in accordance with the declaration of Helsinki. Mononuclear cells were thawed in RPMI-medium supplemented with 20% FCS and incubated for 30 minutes with 100 µg/ml DNAse (Roche, Basel, Switzerland) and 10 mM $MgCl^2$. Subsequently, cells were washed and resuspended in RPMI-medium supplemented with 10% FCS, H-SCF (100ng/mL; Peprotech, Rocky Hill, NJ, USA), FLT3-L (100ng/mL; Peprotech) and IL3 (20ng/mL; Peprotech). Both cell lines as well as mononuclear fractions were cultured at 37°C incubator in a 5% CO^2 humidified atmosphere. Cells were repeatedly authenticated using STR profiling and tested for mycoplasma contamination.

EXPOSURE TO SPLICEOSOME MODULATORS

Both E7107 and H3B-8800 were provided by H3 Biomedicine (Boston, MA, USA). Growth inhibitory effect of splicing modulators in AML cells lines was determined using CellTiter-Blue® Cell Viability Assay (Promega, Madison, WI, USA) upon drug incubations for 96 hours. In addition, exponentially growing cells were exposed to E7107 and H3B-8800 for 24 hours followed by flow cytometry-based cell cycle or apoptosis assessment and for 6 hours followed by RNA extraction (see below).

FLT3 MUTATION ANALYSIS

The status of *FLT3*/ITD and other gene mutations was determined according to the method described by Stone et al (reference 6 main text). Data for patient samples treated with splicing modulation can be found in Supplemental Table S1.

RNA ISOLATION AND PCR ANALYSIS

Total RNA was extracted from leukemic cells using the RNA easy minikit (Qiagen, Venlo, The Netherlands). Reverse transcription was performed using 1ug of the obtained RNA and Moloney Murine Leukemia Virus reverse transcriptase (M-MLV; Invitrogen, Carlsbad, CA, USA) in a reaction buffer containing random hexamers (Roche) DNTPs (Roche) and ribonuclease inhibitor RNasin (Promega). Subsequently, splicing patterns of *MCL1* were examined using PCR with 2x Reddy Mix PCR Master (Thermo Scientific, Waltham, MA, USA) following the manufactures instructions. The PCR products were resolved on 2% agarose gels with Ethidium Bromide. Also, real-time PCR was used to asses splicing patterns of *MCL1-S*, *MCL1-L* and expression levels of *C-MYC*, *N-MYC* and *SF3B1*. The Lightcycler 480II Sybr Green 1 Master mix (Roche) were used for analysis according to manufacturer's protocol. Primer sequences are listed in Table S2.

FLOW CYTOMETRY

Samples were treated as described above. For cell cycle analysis, cells were permeabilized in 70% ethanol followed by 30 minutes incubation with RNAse (100ug/mL; Qiagen) and staining with Propidium Iodide (Thermo Fisher Scientific). Subsequently, cell count and Propidum Iodide (PI) staining were plotted to assess DNA content and distinguish G1, S, G2 and M phases. Apoptosis induction was analyzed using the Apoptest-FITC kit (VPC Diagnostics, Hoeven, The Netherlands) and 7-AAD (BD Via-ProbeTM, BD Bioscience, San Jose, CA, USA). Furthermore, flow cytometry (BD FACSCelesta) was used for cell count assessment and phenotypic analysis of mononuclear cell fractions. First, the quality of cells was assessed upon 24 hours drug incubation using the Apoptest-FITC kit and 7-AAD. Samples with less than 50% viability were rejected at this time point. Upon 48 hours of incubation, cells were stained with 7-AAD. Absolute cell counts were determined based on the percentage of 7-AAD negative cells in a total volume of 70uL. Within the CD45dim cell populations progenitor cells were identified as CD34+ cells. The analysis of lymphocytes was performed upon staining with CD45-KO, CD34-BV421 and CD7-PE (BD Biosciences) and characterized by high CD45, high CD7 and low side scatter. Expression levels of FLT3 (CD135) were determined using CD135-PE (BD Bioscience) antibodies. Cells were incubated with CD-135 PE antibodies, followed by the measurement of mean fluorescence intensity (MFI) of 7-AAD negative cells.

CFU ASSAY

Samples were thawed as described above. Subsequently, cells were washed and resuspended in RPMI plus 20% FCS (Gibco) and let to rest overnight at 37°C and 5% CO2. Subsequently, drug dilutions were prepared in Iscove's Modified Dulbecco's Media (IMDM; Gibco). Next, cells were washed and resuspended in IMDM containing E7107 or H3B-8800. Then cells were plated in methylcellulose (1 mL/well; Cat. H4435; STEMCELL Technologies Inc., Vancouver, BC, Canada) in concentrations varying from 10,000 to 50,000 cells/well, depending on colony forming capacity that was assessed in previous experiments. For each condition cells were plated in triplicate. The number of colonies was scored after two weeks of culture.

NANOSTRING PANEL

Custom Nanostring panel was designed by H3 Biomedicine through Nanostring and validated at H3 Biomedicine based on splicing modulation using H3B-8800. nCounter Elements TagSet (84) and other consumables were ordered from Nanostring (Seattle, WA). Customized oligonucleotide probe pools were ordered through Integrated DNA Technologies (Coralville, IA). Experimental procedure was done according to nCounter Elements XT Reagents User Manual and nCounter Analysis System User Manual for MAX. Briefly, hybridization master mix was generated using different oligonucleotide probe pools and TagSet84. For each reaction, 350ng of RNA was mixed with this mix, and incubated in a thermal cycler (67°C) for 18hrs followed by incubation at 4°C. Next, samples were run and analyzed. Raw counts were normalized using negative probes plus two times the standard deviation. In addition, positive control normalization was done using geometric mean of all the positive controls, with minimum threshold of 0.3 and maximum threshold of 3. In the end, content normalization was done using geometric mean of housekeeping genes, with minimum threshold of 0.1 and maximum threshold of 10.

Subsequently, normalized gene counts were used for further data analysis.

RNA-Sequencing Analysis

The TruSeq Total Stranded RNA kit was used, starting with 250ng of total RNA, to generate RNA libraries following the manufacturer's recommendations (Illumina, San Diego, CA, USA). 2x100bp paired-end reads were sequenced on the NovaSeq 6000 with a median of 50 mio. reads per sample (Illumina, San Diego, CA, USA). Using BaseSpace's RNA-seq Alignment app (v2.0.1) with default parameters, reads were mapped with the STAR aligner (v2.5.0a,) to the human reference genome hg19 (RefSeq annotation). For gene expression analysis estimated gene counts were normalized applying Trimmed mean of M-values (TMM) normalization method of the edgeR package (PMID: 19910308). The resulting log2 counts per million (CPMs) were used as a proxy of gene expression. Genes with a CPM < 1 were filtered out.

STATISTICS

All statistical analyses were performed using Prism 8 Software as well as R version 3.6.3/R studio version 1.2.5, including ggplot2 (version 3.2.1) and ComplexHeatmap (version 2.2.0) packages.[1,2] The Mann-Whitney U test was used to compare cell counts between different subgroups of treated AML patient samples, differences in apoptosis induction, as well as mRNA expression levels. Associations between response rates, and response rates to expression levels were analyzed using Spearman's Rho test. In all analyses p-values at below 0.05 were considered statistically significant.

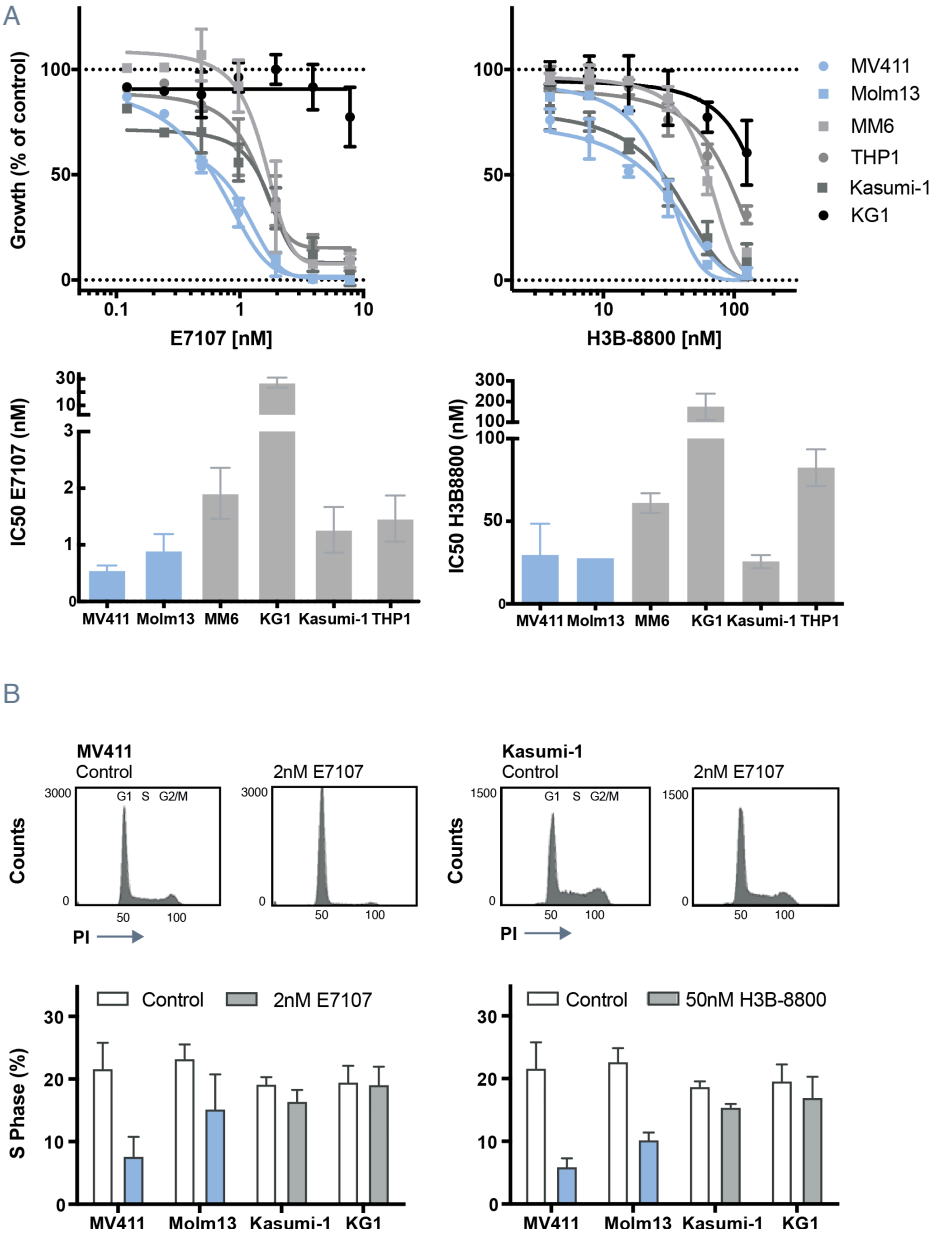
REFERENCES

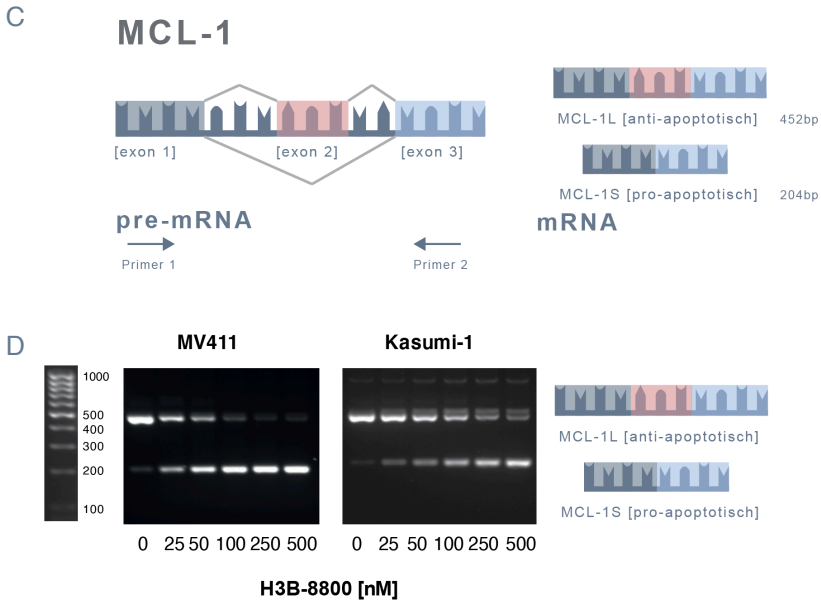
1. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847-2849, doi:10.1093/bioinformatics/btw313 (2016).
2. Valero-Mora, P. M. ggplot2: Elegant Graphics for Data Analysis. 2010 35, 3, doi:10.18637/jss.v035.b01 (2010).

SUPPLEMENTAL TABLES

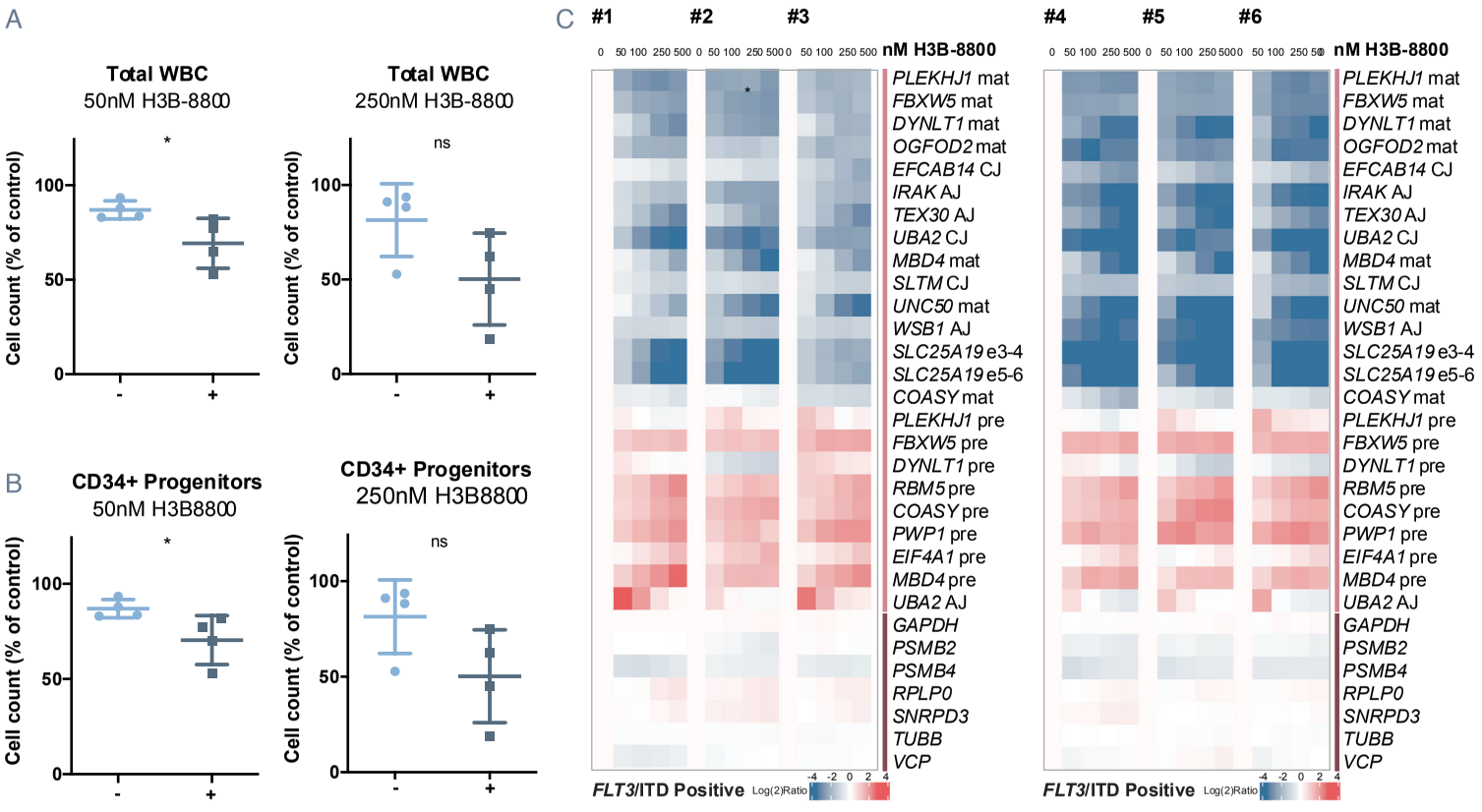
Table S1 and S2 can be found online at https://www.nature.com/articles/s41375-021-01273-7

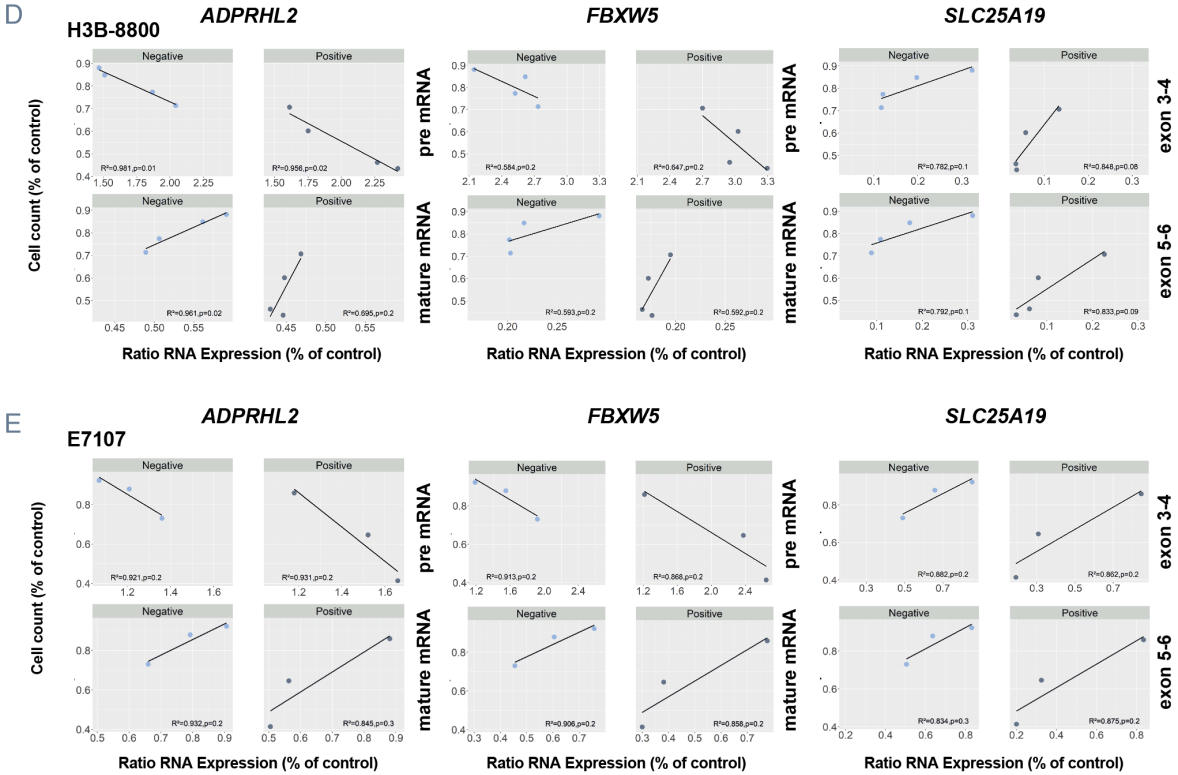
SUPPLEMENTAL FIGURES



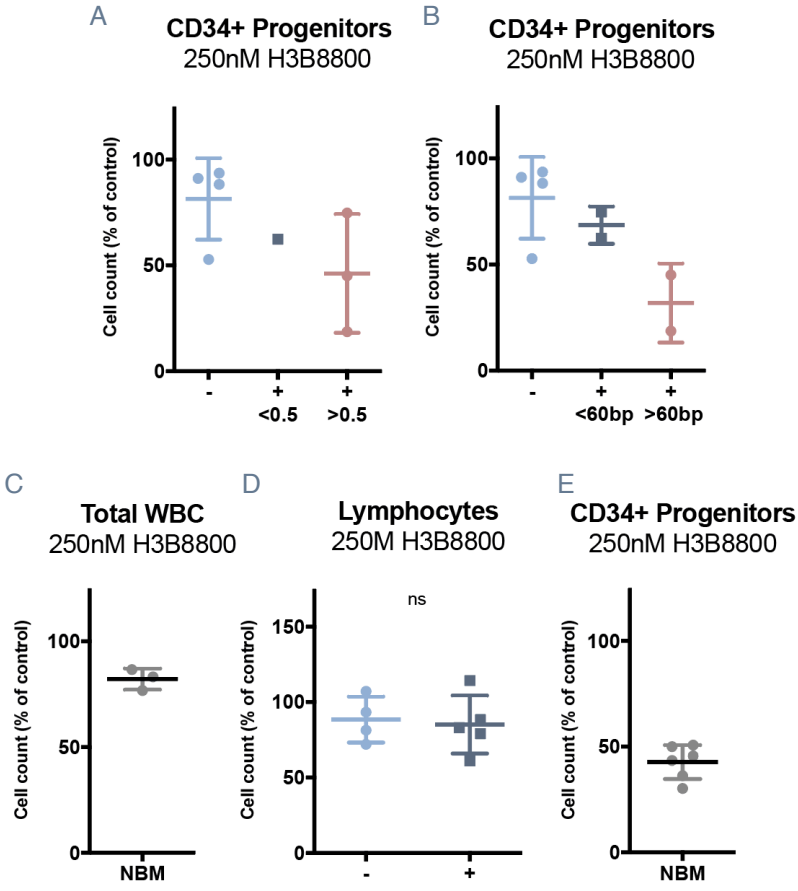


Supplemental Figure S1. Sensitivity of AML cell lines to splicing modulation. Both MV4-11 and Molm13 represent *FLT3*/ITD^{pos} cell lines - MV4-11 (homozygous for *FLT3*/ITD) and Molm13 (heterozygous for *FLT3*/ITD). A - The response of cell lines incubated with a range of E7107 (left) or H3B-8800 (right) concentrations for 96h (Top) and IC50 calculations based on dose response curves (Bottom; N=3). B - Cells were incubated with 2nM E7107 (left) or 50nM H3B-8800 (right). Cell cycle phases are visualized in a histogram plot of the PI stain upon flow cytometric analysis, in addition the fraction of cells (%) in the S phase of the cell cycle with or without treatment is indicated (N=3). C - Schematic representation of *MCL-1* splicing. D - *MCL1* expression after incubation with a range of H3B-8800 concentrations. Pro- (*MCL-S*) and anti-apoptotic (*MCL1-L*) splice variants are indicated.

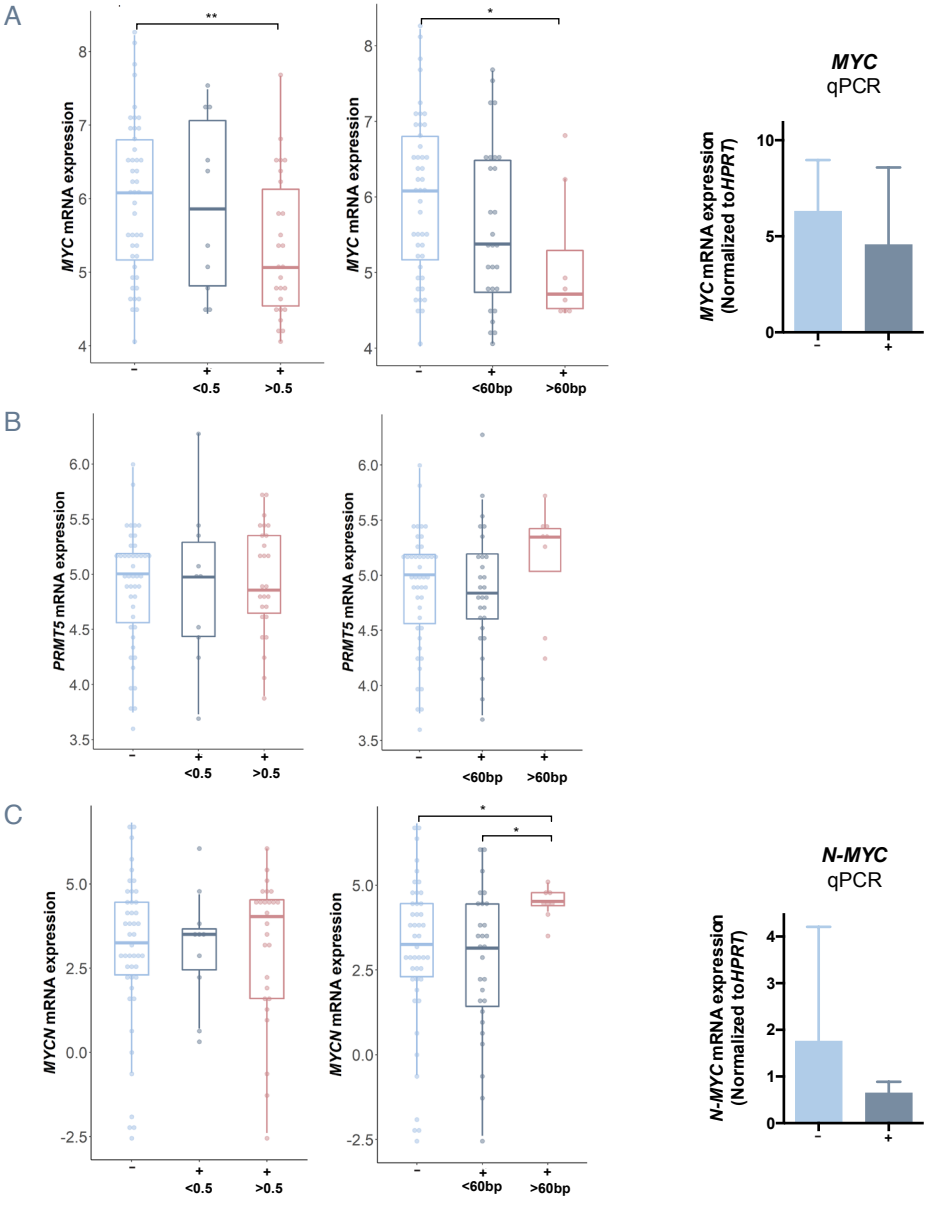

A
Total WBC
50nM H3B-8800
Cell count (% of control)
100
50
0
*
-
+
Total WBC
250nM H3B-8800
ns
B
CD34+ Progenitors
50nM H3B8800
CD34+ Progenitors
250nM H3B8800
C
#1
#2
#3
#4
#5
#6
0 50 100 250 500
nM H3B-8800
PLEKHJ1 mat
FBXW5 mat
DYNLT1 mat
OGFOD2 mat
EFCAB14 CJ
IRAK AJ
TEX30 AJ
UBA2 CJ
MBD4 mat
SLTM CJ
UNC50 mat
WSB1 AJ
SLC25A19 e3-4
SLC25A19 e5-6
COASY mat
PLEKHJ1 pre
FBXW5 pre
DYNLT1 pre
RBM5 pre
COASY pre
PWP1 pre
EIF4A1 pre
MBD4 pre
UBA2 AJ
GAPDH
PSMB2
PSMB4
RPLP0
SNRPD3
TUBB
VCP
FLT3/ITD Positive
Log(2)Ratio
-4 -2 0 2 4



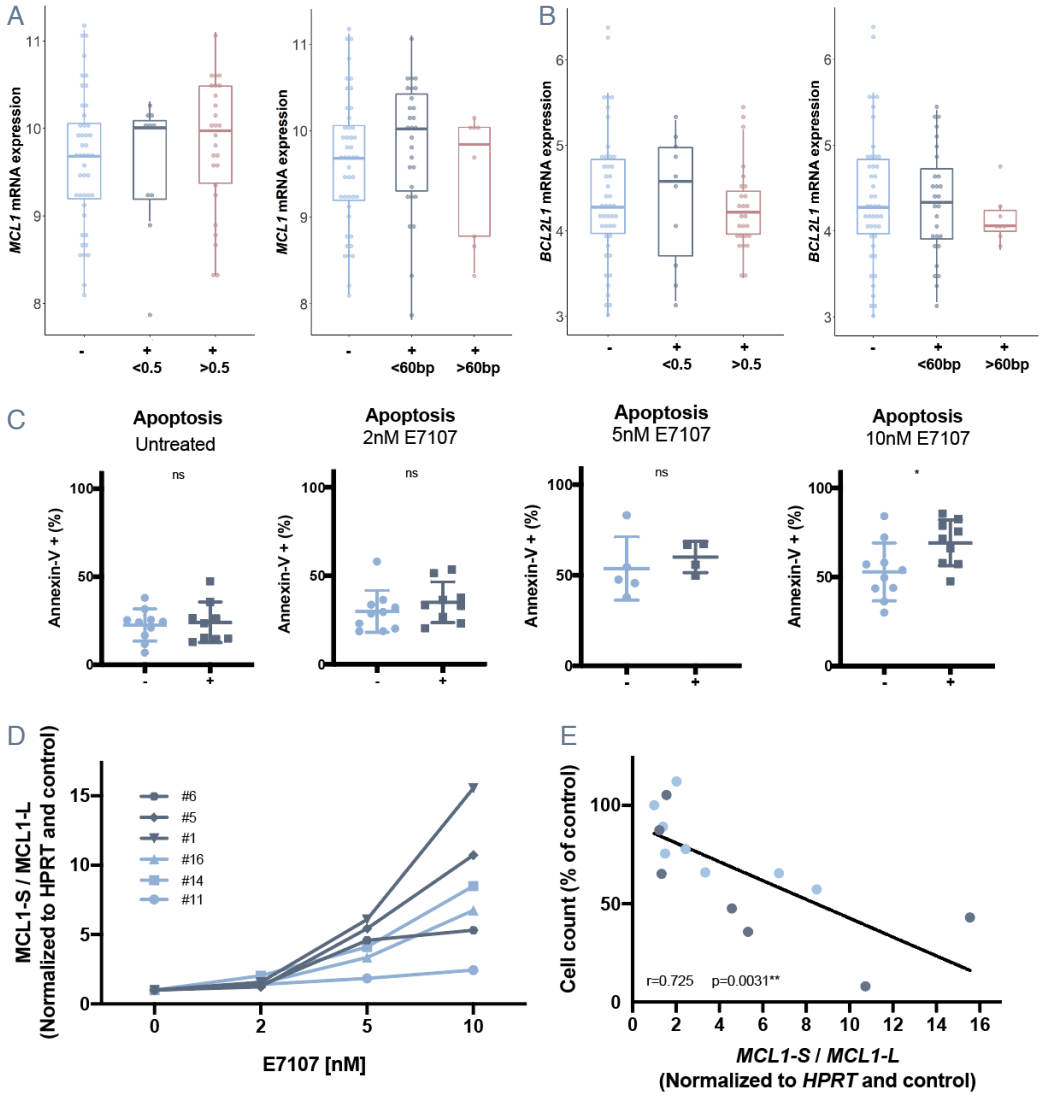
Supplemental Figure S2. Response of primary AML cells with or without *FLT3*/ITD to splicing modulation. A, B - Cells were treated with different doses H3B-8800 for 48h. Cell count of the total white blood cell populations (WBC) was determined using Flow Cytometry. Progenitors were identified based on CD34 expression. Cell counts are plotted as a percentage of untreated cells. P-values are based on Mann-Whitney U test. C - Splicing modulation via E7107 or H3B-8800 results in decreased production of mature mRNA transcripts (mat; blue) and concomitant accumulation of pre-mRNA transcripts (pre; red). In addition, alternative junctions (AJ; red) of certain genes are preferred upon splicing modulation resulting in decreased expression of constitutive spliced mRNAs (CJ; blue).Heatmap that presents dose-dependent modulation for both mature as well as pre-mature mRNA markers in *FLT3*/ITD^{neg} (left) and *FLT3*/ITD^{pos} (right) cells upon 4h treatment with H3B-8800. D, E - Correlation of response rate upon treatment with either H3B-8800 (D) or E7107 (E) to expression of mRNA transcripts normalized to control. Light blue dots (-) represent *FLT3*/ITD^{neg} patients, dark blue dots represent *FLT3*/ITD^{pos} patients (+).



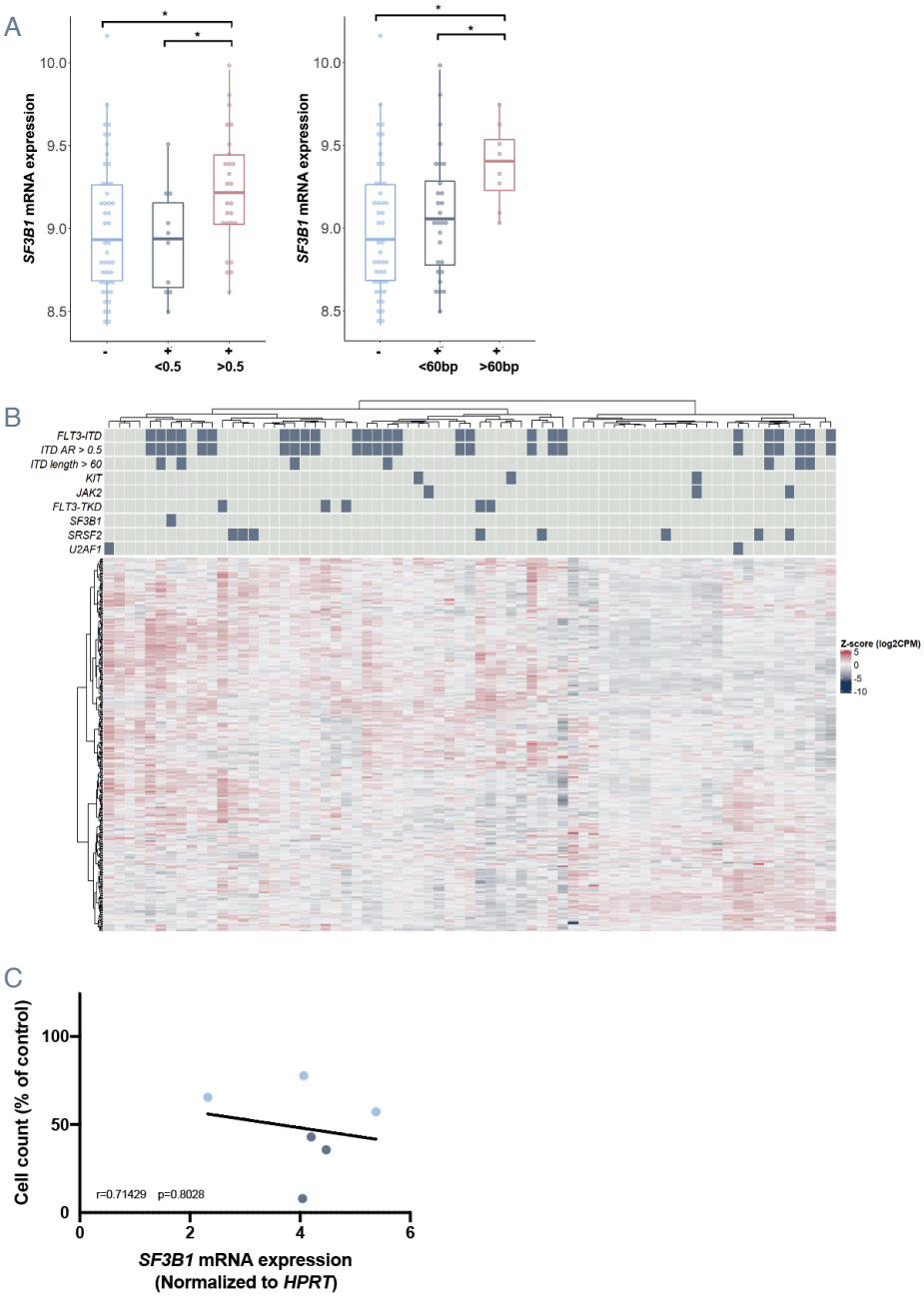
Supplemental Figure S3. Preferential sensitivity of *FLT3*/ITD^{pos} AML patients with high AR or long ITD length. A - Patient samples were grouped on their *FLT3*/ITD allelic ratio according to the ELN 2017. B - Patient samples were grouped on their ITD size. Cut off was based on mean size of ITD length of selected patients. C - Cell count of the total white blood cell population within healthy bone marrow was assessed upon treatment D - Lymphocytes within AML bone marrow samples were identified based on CD45, CD34 and CD7 expression. E - Cell count of CD34 positive cells within healthy bone marrow samples was determined upon treatment. Cells were treated with 250nM H3B-8800 for 48h. Cell counts are plotted as a percentage of untreated cells.



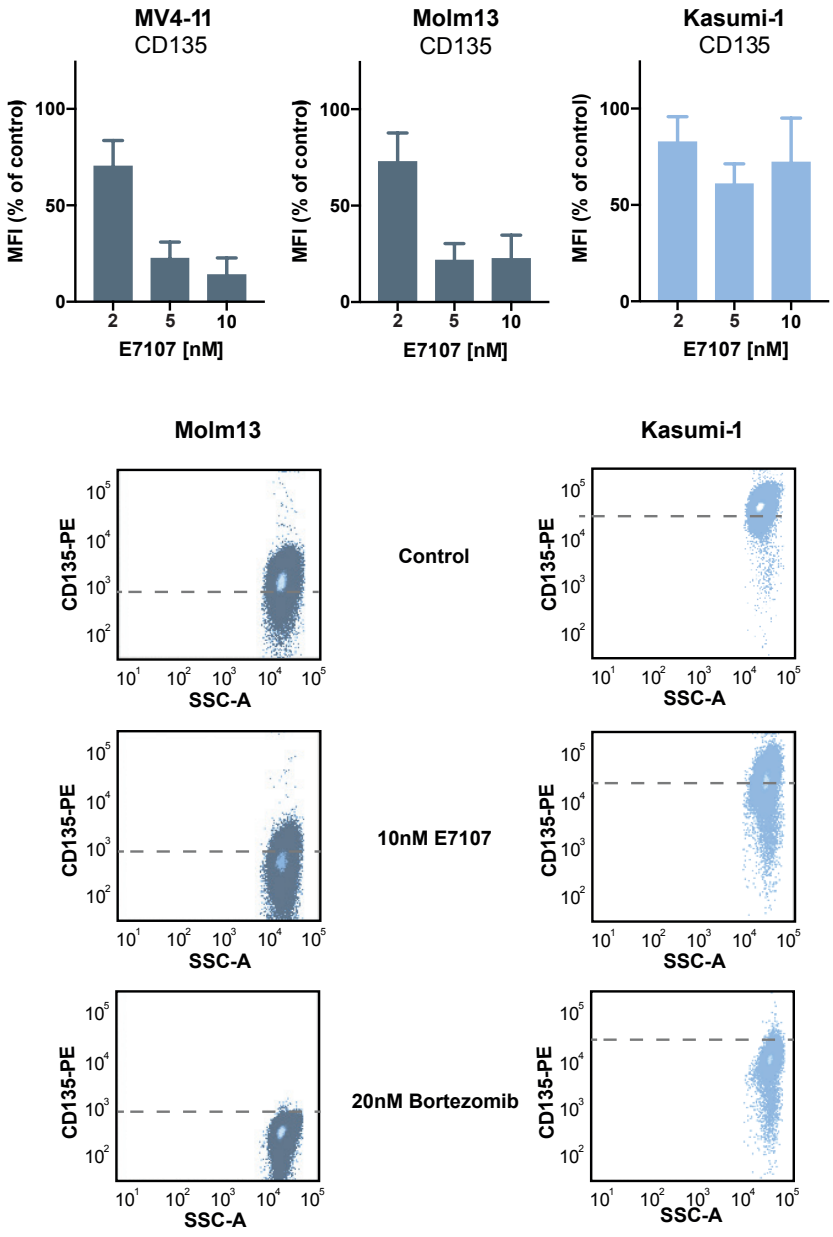
Supplemental Figure S4. MYC expression in *FLT3*/ITD^{pos} subgroups of AML patients. A - mRNA expression of *MYC* determined by RNAseq (left) or qPCR (right). Patient samples were grouped on their allelic ratio according to ELN2017 or ITD size (left). *FLT3*/ITD negative and positive cells with differences in sensitivity were subjected to qPCR (right). B - mRNA expression of *PRMT5* determined by RNAseq. C - mRNA expression of *N-MYC* determined by RNAseq (left) or qPCR (right). Patient samples were grouped on their allelic ratio according to ELN2017 or ITD size (left). *FLT3*/ITD negative and positive cells with differences in sensitivity were subjected to qPCR (right).



Supplemental Figure S5. Apoptosis induction of primary AML cells upon E7107 treatment. A - *MCL1* mRNA expression levels determined by RNA seq. B - *BCL2L1* mRNA expression levels determined by RNAseq. C - *Ex vivo* AML patients' cells, collected at diagnosis, were selected based on their *FLT3* mutation status. Cells were treated with various dosages of E7107 for 48h. Subsequently, flow cytometry-based apoptosis assays were performed. No differences in levels of apoptosis were measured in untreated cells or cells treated with 2nM or 5nM E7107, apoptosis induction is significantly higher in *FLT3*/ITD^{pos} cells upon incubation with 10nM E7107 (p-values based on Mann-Whitney U test). D - Cells were incubated with a range of E7107 concentrations for 6h, followed by RNA isolation, cDNA synthesis and qPCR performed on *MCL1*. The ratio of pro-apoptotic *MCL1-S* to anti-apoptotic *MCL1-L* is plotted. E - Association between response rates to E7107 (cell counts as percentage of control determined by Flow Cytometry) and the ratio of *MCL1-S* to *MCL1-L* (analyzed using qPCR). P-value determined by Spearman's rho test. Light blue dots represent *FLT3*/ITD^{neg} patients, dark blue dots represent *FLT3*/ITD^{pos} cells.



Supplemental Figure S6. Expression of SF3B1 is significantly increased in *FLT3*/ITD^{pos} AML patients with high AR or long ITD length. A - mRNA expression of *SF3B1* determined by RNAseq. B - Heatmap of genes involved in splicing in *FLT3*/ITD^{pos} patients with high allelic ratio (AR). C - Lack of correlation between response rates to E7107 (determined by Flow Cytometry) and *SF3B1* expression levels (measured by qPCR). P-value determined by Spearman's rho test. Light blue dots represent *FLT3*/ITD^{neg} specimen, dark blue dots represent *FLT3*/ITD^{pos} cells.



Supplemental Figure S7. Splicing modulation specifically decreases FLT3 expression levels decrease in FLT3/ITD^{pos} AML cell lines. Both MV4-11 and Molm13 represent *FLT3*/ITD^{pos} cell lines, Kasumi-1 cells are negative for this aberration. Cells were incubated with different dosages of E7107 for 24h. Subsequently, FLT3 (CD135) expression levels were measured using Flow Cytometry. Results are plotted as percentage of expression levels in control cells (N=3). In addition, FACS plots of cells stained for FLT3 (CD135) after 24h treatment with either E7107 or Bortezomib are presented.

CHAPTER 5

Maturation State-Specific Alternative Splicing in *FLT3*/ITD and *NPM1* Mutated AML

Anna Wojtuszkiewicz[1]
Inge van der Werf[1]
Stephan Hutter[2]
Wencke Walter[2]
Constance Baer[2]
Wolfgang Kern[2]
Jeroen J. W. M. Janssen[1]
Gert J. Ossenkoppele[1]
Claudia Haferlach[2]
Jacqueline Cloos[1#]
Torsten Haferlach[2#]

1. Dept. of Hematology, Amsterdam University Medical Center, VU University Medical Center, Cancer Center Amsterdam, Amsterdam, The Netherlands
2. MLL Munich Leukemia Laboratory, Munich, Germany

[#]Shared senior authorship

Cancers 2021, 13(16), 3929

INTRODUCTION

Whole genome profiling efforts in the last decade have defined the somatic mutational landscape of acute myeloid leukemia (AML).[1,2] While our knowledge regarding genetics of AML has largely increased, little progress has been made with respect to improvements in AML treatment outcome. Therefore, there is an urgent need to further deepen our understanding of how different genetic mutations affect phenotypic features of AML cells in order to better predict their responses to current treatments as well as to invent novel therapeutic approaches. The genetic lesions involved in AML pathogenesis include aberrations in transcription factors, epigenetic regulators and signalling molecules, which collaborate to promote a block in differentiation paralleled by enhanced survival, self-renewal and proliferation.[1,3,4] Interestingly, alternative pre-mRNA splicing (AS) is known to play a pivotal role in regulation of all of these processes.[5-7]

Pre-mRNA splicing is a crucial step in gene expression whereby the non-coding segments (introns) are excised and coding regions (exons) are joined together.[8] Tissue and organ development was documented to be driven by coordinated networks of AS events, which regulate various aspects of differentiation including cell cycle progression, DNA damage repair and apoptosis.[5-7,9-12] Accordingly, AS perturbations were shown to affect these key facets of development thereby facilitating oncogenesis.[13,14] In AML cells, it was estimated that almost 30% of expressed genes are aberrantly spliced as compared to non-malignant CD34+ progenitor cells.[15] Thus far, much attention has been dedicated to characterization of AS in AML samples carrying mutations in splicing factors (SF); however, AS in relation to mutations in other genes, indirectly linked to splicing regulation, remains poorly characterized.[16-19]

Internal tandem duplications (ITD) in the Fms-like tyrosine kinase 3 (*FLT3*) gene are among the most common genetic aberrations in AML affecting roughly 30% of patients.[1,2,20] FLT3 is a receptor tyrosine kinase, which, via downstream signalling pathways, controls growth and survival of myeloid progenitors and is rendered constitutively active upon ITD insertion.[21] *FLT3*/ITD rarely occurs alone and most frequently coincides with mutations in nucleophosmin (*NPM1*) with many of *FLT3*/ITD^{pos}/*NPM1*^{pos} AML patients eventually relapsing.[1] NPM1 is a multifunctional protein with diverse physiological roles that include regulation of cell cycle and DNA damage repair, maintenance of genomic stability and stress response.[22] The molecular synergy between *FLT3*/ITD and *NPM1* mutations was demonstrated to drive rapidly developing AML in mouse models.[23,24] In contrast, co-occurring *NPM1* and *NRAS* mutations induced less aggressive AML, underscoring the frequent occurrence and worse prognosis of *FLT3*/ITD^{pos}/*NPM1*^{pos} AML.[23,24] This was paralleled by the cooperative impact of these two aberrations on gene expression profiles.[23,25] Interestingly, both FLT3/ITD and NPM1 were also previously shown to shape the epigenome in AML.[22,26] As the process of splicing occurs co-transcriptionally and its regulation was shown to be influenced by chromatin status (including modifications to both histones and DNA), by shaping the epigenome *FLT3*/ITD and *NPM1* mutations have the potential to affect splicing.[27-29] Yet splicing profiles associated with these co-occurring aberrations have not been studied thus far.

Therefore, the aim of this study was to explore the differential splicing profiles associated with the presence of *FLT3*/ITD with concomitant *NPM1* mutations to characterize their potential oncogenic relevance. Furthermore, AS profiles as well as gene expression profiles orchestrate differentiation and maturation of cells and tissues and therefore can show much variability between various cell types and maturation stadia.[30,31] Since both *FLT3*/ITD and *NPM1* mutations occur in AML cells arrested in different maturation stadia, we evaluated whether differential splicing and differential expression signatures in relation to *FLT3*/ITD and *NPM1* mutations in AML showed FAB subtype specificity.

MATERIALS AND METHODS

PATIENT SAMPLES

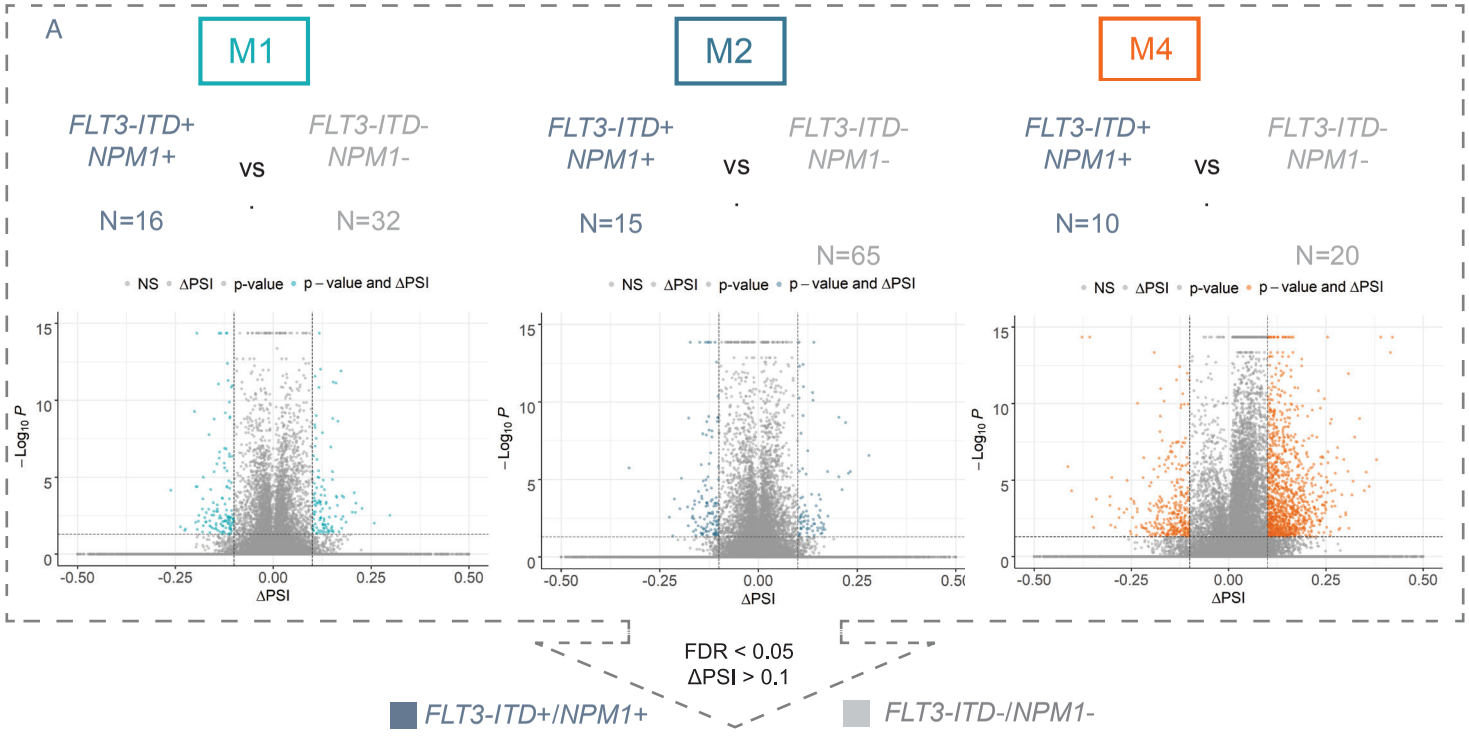
For splicing analysis 382 untreated bone marrow (BM) and peripheral blood (PB) samples collected at the time of diagnosis from AML patients were used. This included 327 samples in the discovery cohort (collected at MLL Munich Leukemia Laboratory) and 55 samples in an independent validation cohort (collected at Amsterdam University Medical Center, AUMC, location VUmc). All patients have signed a written informed consent. The cell type-specific analyses included the three most represented subtypes in the dataset according to the French-American-British (FAB) classification (72 M1, 92 M2 and 56 M4 samples). The validation cohort included 19 M1, 17 M2 and 19 M4 specimens (sample metadata are listed in Supplemental Tables S1-S2). This study was approved by the internal review board of the MLL and local ethics committee of Amsterdam UMC and was conducted in accordance with the Declaration of Helsinki.

GENETIC PROFILE

The presence of *FLT3*/ITD, *NPM1* mutations and SF mutations of patients from the MLL dataset were based on routine molecular diagnostics (including a combination of gene scan analysis, melting curve analysis, Sanger sequencing and next-generation amplicon sequencing as described previously) complemented by whole genome sequencing (Supplemental Methods).[32-35] The mutational status in the AUMC dataset was defined based on the molecular diagnostics as described previously, complemented with variant calling from RNA sequencing data (Supplemental Methods).[36,37] All samples carrying SF mutations (*SF1*, *SF3A1*, *SF3B1*, *SRSF2*, *U2AF1*, *U2AF2* and *ZRSR2*) or samples for which the average coverage in frequently mutated exons of SF genes was low were removed from the analysis (Supplemental Methods). All samples considered *FLT3*/ITD^{pos} based on the mutational analysis were included in the analysis (including the following fractions of samples with *FLT3*/ITD allelic ratio > 0.5 as determined on DNA by fragment analysis: 63.6% of M1, 47.1% of M2 and 58.3% of M4 samples).

RNA SEQUENCING

Total RNA was extracted from BM and PB samples using the MagNA Pure 96 Instrument and the MagNA Pure 96 Cellular RNA LV Kit (Roche LifeScience, Mannheim, Germany) for the discovery cohort, and using the RNeasy mini kit (Qiagen, Venlo, The Netherlands) for validation cohort. The TruSeq Total Stranded RNA kit was used to generate RNA libraries following the manufacturer's recommendations, starting with 250 ng of total RNA (Illumina, San Diego, CA, USA). 2x100 bp paired-end reads were sequenced on the NovaSeq 6000 with a median of 50


A
M1
FLT3-ITD+ NPM1+
vs
FLT3-ITD- NPM1-
N=16
N=32
M2
FLT3-ITD+ NPM1+
vs
FLT3-ITD- NPM1-
N=15
N=65
M4
FLT3-ITD+ NPM1+
vs
FLT3-ITD- NPM1-
N=10
N=20
NS ΔPSI p-value p – value and ΔPSI
−Log10 P
ΔPSI
FDR < 0.05
ΔPSI > 0.1
FLT3-ITD+/NPM1+
FLT3-ITD-/NPM1-

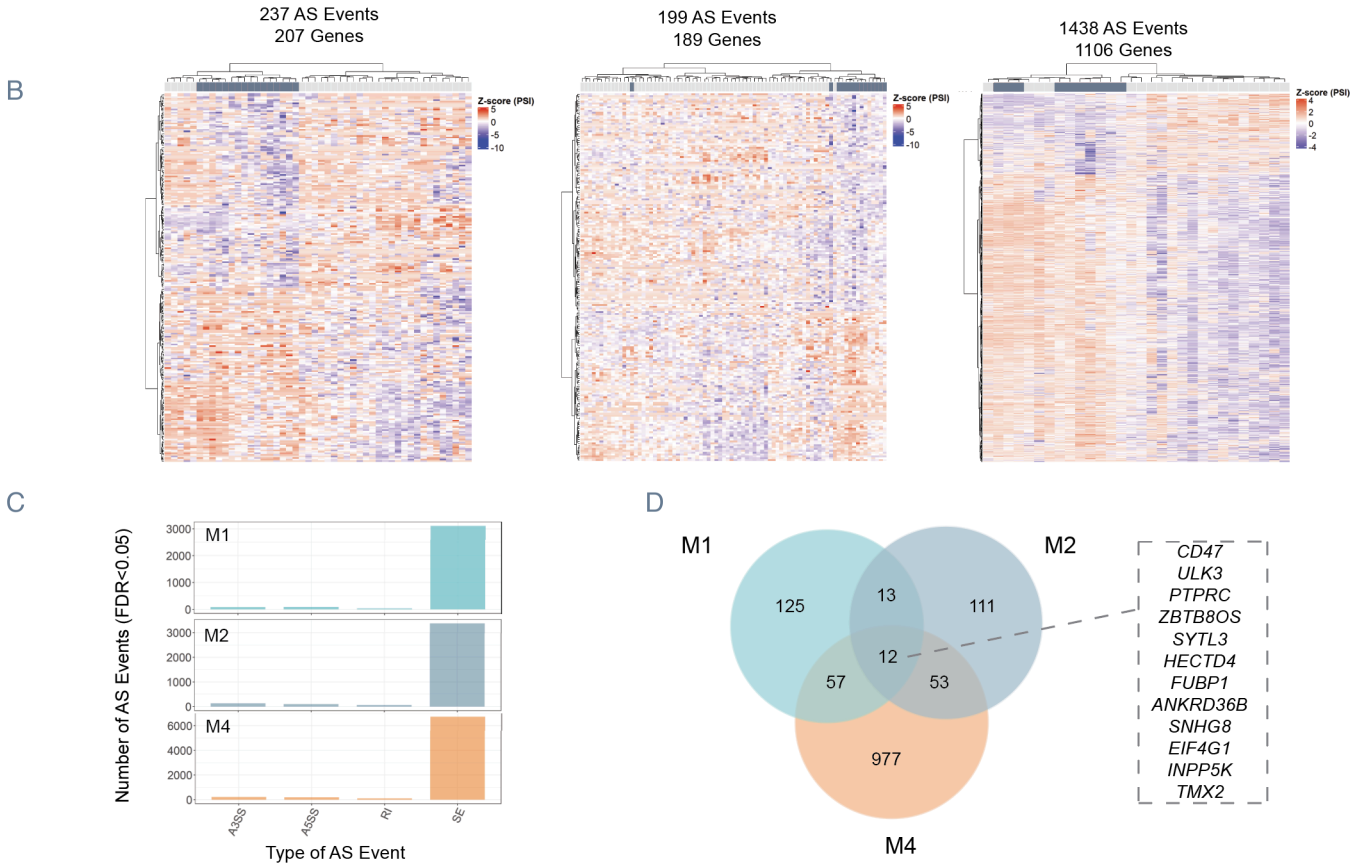


Figure 1. Concomitant *FLT3*/ITD and *NPM1* mutations are associated with a strong FAB type-specific splicing profile. An overview of the differential splicing analysis performed with rMATS in three major FAB subtypes of AML (M1, M2 and M4) including (A) volcano plots of PSI (proportion spliced-in) values and (B) hierarchical clustering performed using significant differential splicing events (FDR < 0.05) with a minimal splicing difference between the two groups of 0.1 (ΔPSI > 0.1). The numbers of *FLT3*/ITD^{pos}/*NPM1*^{pos} and FLT3/ITD^{neg}/NPM1^{neg} patients as well as the number of significant differential splicing (AS) events in each analysis are indicated. (C) The distribution of significant AS events in M1, M2 and M4 FAB types between the four main AS categories – skipped exons (SE), alternative 3' splice site selection (A3SS), alternative 5' splice site selection (A5SS) and retained introns (RI). (D) Overlap between significantly differentially spliced genes in relation to concomitant *FLT3*/ITD and *NPM1* mutations in M1, M2 and M4 FAB subtypes with a minimal splicing difference between the two groups of 0.1 (FDR < 0.05, |ΔPSI| > 0.1) . The 12 events over-lapping between all three FAB subtypes are indicated.

mln reads per sample (Illumina). Using BaseSpace's RNA-seq Alignment app (v2.0.1) with default parameters, reads were mapped with the STAR aligner (v2.5.0a) to the human reference genome hg19 (RefSeq annotation). For gene expression analysis, estimated gene counts were normalized applying Trimmed mean of M-values (TMM) normalization method of the edgeR package.[38] The resulting log2 counts per million (CPMs) were used as a proxy of gene expression. Genes with a CPM < 1 were filtered out.

DIFFERENTIAL GENE EXPRESSION AND SPLICING ANALYSIS

Gene expression differences were assessed using the limma package[39] with false discovery rate (FDR) correction for multiple testing. Genes with an FDR less than 0.05 and an absolute logFC greater than 1.5 were considered differentially expressed (DE).

rMATS version 4.0.2 was used to detect alternative splicing (AS) events.[40] rMATS is able to quantify four major types of alternative splicing events: skipped exons (SE), alternative 5' splice site selection (A5SS), alternative 3' splice site selection (A3SS) and retained introns (RI). The difference in splicing between the two groups is expressed as ΔPSI (Proportion Spliced-In). AS events supported by fewer than 10 counts per sample were filtered out. AS events were considered significantly differential when FDR < 0.05 and absolute ΔPSI > 0.1. Z-score calculation and hierarchical clustering for AS events were performed using PSI values. The data was visualized using ggplot2[41] and ComplexHeatmap packages (version 2.2.0)[42] in R (versions 3.5.3 and 3.6.2). Protein domains directly affected by splicing events were determined using the Maser package (version 1.0.0)[43] in R (version 3.5.3) upon conversion of the genomic coordinates from hg19 to hg38 assembly (using AnnotationHub v. 2.14.5, GenomicRanges v. 1.34.0 and rtracklayer v. 1.42.2 packages in R).[44-46] Motif enrichment analysis for the differentially spliced splicing factors was performed using rMAPS tool.[47,48] For gene ontology analysis, gene IDs for significant AS events were uploaded into the STRING tool (v11.0) to retrieve interactions.[49] STRING interaction networks were imported and annotated in Cytoscape (v3.8.1).[50] Gene ontology analysis was performed within Cytoscape using the ClueGO plugin.[51]

For validation of AS events in an independent sample cohort, AS events and their respective PSI values in the validation cohort were determined by rMATS. Subsequently, PSI values in the validation cohort, corresponding to significant AS events in the discovery cohort, were retrieved based on genomic coordinates and compared between *FLT3*/ITD+/*NPM1*+ and *FLT3*/ITD-/*NPM1*- sample groups using the Mann-Whitney U test. The raw results of differential splicing analyses are available in Supplemental Tables S3-S15 while the lists of differentially expressed genes are available in Supplemental Table S16.

RESULTS

ALTERNATIVE SPLICING PROFILES OF FLT3/ITD AND *NPM1* DOUBLE MUTATED CELLS SHOW HIGH FAB-TYPE SPECIFICITY

FLT3/ITD and *NPM1* mutations were previously reported to be associated with specific differential gene expression profiles in AML (Supplemental Figure S1),[25,52] however the influence of these aberrations on splicing was not studied thus far. Therefore, we applied rMATS al-

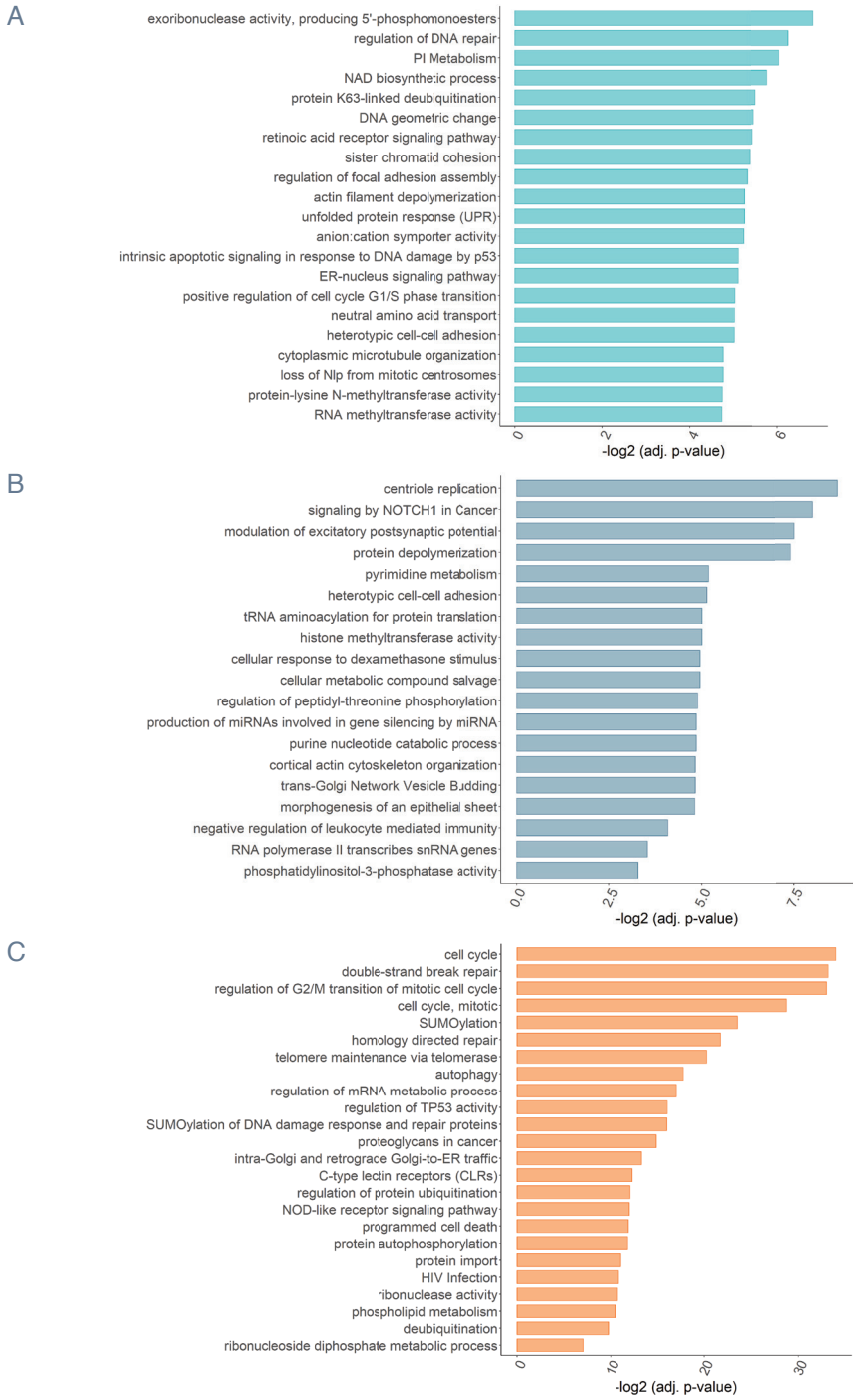


Figure 2. Functional analysis of differentially spliced genes in relation to concomitant *FLT3*/ITD and *NPM1* mutations. The figure depicts functional enrichment among significant differential splicing events (FDR < 0.05) in relation to concomitant *FLT3*/ITD and *NPM1* mutations in M1 (A), M2 (B) and M4 (C) with a minimal splicing difference between the two groups of 0.1 (ΔPSI > 0.1).

gorithm to analyse differential splicing in relation to the presence of these mutations in an RNA-sequencing dataset obtained from 327 diagnosis samples of *de novo* AML patients (with patients carrying splicing factor mutations excluded from the analysis to avoid bias). Since co-occurrence of *FLT3*/ITD and *NPM1* mutations was previously described to exert particularly strong synergistic effects on gene expression,[23,24,53] we primarily focused our splicing analysis on this double mutated subset of samples. This approach uncovered a total of 217 significant differential splicing events in *FLT3*/ITD^{pos}/*NPM1*^{pos} specimens as compared to *FLT3*/ITD^{neg}/*NPM1*^{neg} samples (Supplemental Figure S2). Hierarchical clustering did not reveal a specific cluster of *FLT3*/ITD^{pos}/*NPM1*^{pos} samples suggesting limited specificity of the identified splicing events for this subgroup (Supplemental Figure S2F).

Since cell type and maturation state of the cells are known to influence alternative splicing,[5,30,54-57] we stratified the cohort on the basis of individual (most common) FAB types, including 48 M1, 80 M2 and 30 M4 samples (Supplemental Figure S1D). Interestingly, within specific FAB types we could identify patterns of differential splicing more specific for *FLT3*/ITD^{pos}/*NPM1*^{pos} samples, suggesting highly maturation state-dependent splicing regulation in the context of these mutations (Figure 1A,B). The majority of significant events in all analyses constituted skipped exons (Figure 1C, Supplemental Figure S3B and S4B).

Furthermore, the splicing profiles of the double mutated *FLT3*/ITD^{pos}/*NPM1*^{pos} samples showed an improved clustering pattern as compared to either *FLT3*/ITD or *NPM1* overall (Figure 1 and Supplemental Figures S3-S4). This points to a possible synergy in splicing regulation between *FLT3*/ITD and *NPM1* mutations, similar to that observed for regulation of gene expression. Strikingly, the number of differentially spliced genes in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples was remarkably high in M4 subtype (1438 differential splicing events) as compared to M1 and M2 samples (approximately 200 events each, Figure 1A). In addition, the overlap between differentially spliced genes in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples in the three individual FAB types encompassed only 12 genes (Figure 1D), highlighting the impact of FAB type on differential splicing profiles associated with these genetic aberrations. The functional annotation of these genes does not give direct clues as to the relation with FLT3 or NPM1 but are more related to splicing and regulation of protein homeostasis. Interestingly, similar to splicing events, differentially expressed genes in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples were largely FAB type-specific (Supplemental Figure S5), including several regulators of hematopoietic differentiation previously reported to be associated with *FLT3*/ITD and *NPM1* mutations (i.e. differential expression of *FOXC1*, *MEIS1* and *FOXO1* in M1 and M2 but not in M4 samples, Supplemental Figure S6). Altogether, these findings demonstrate that differentially spliced (as well as aberrantly expressed) genes associated with *FLT3*/ITD and mutated *NPM1* might be relevant only in AML cells of specific differentiation stages.

FLT3/ITD AND *NPM1* DOUBLE MUTATED CELLS DISPLAY ALTERED SPLICING OF GENES INVOLVED IN CELL CYCLE CONTROL, DNA DAMAGE AND SIGNALLING PATHWAYS

To evaluate the biological functions of the uncovered differential splicing events in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples, we performed functional enrichment analysis. Remarkably, in the three FAB types M1, M2 and M4, the major affected processes included regulation of cell

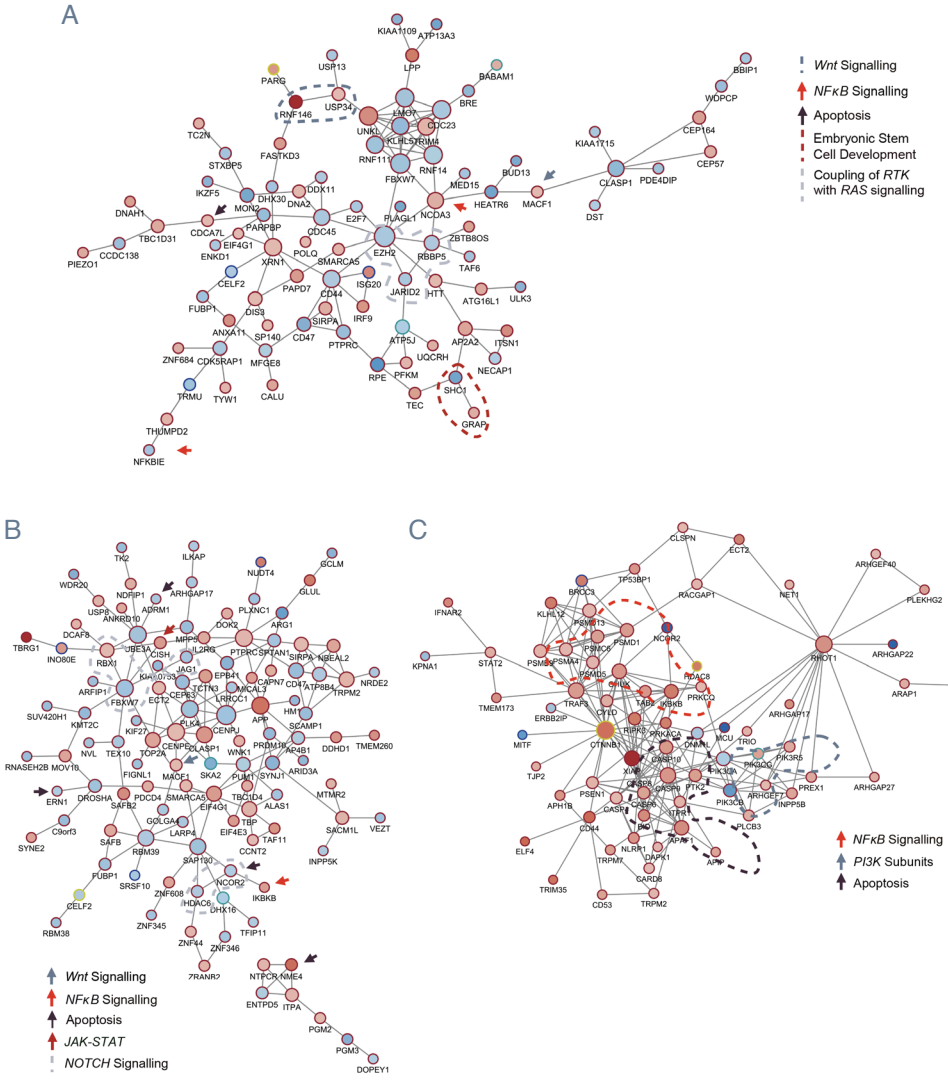


Figure 3. Networks of genes relevant to oncogenesis. A - Network of all significantly differential splicing events with the minimal splicing differ-ence between groups of 0.1 (FDR < 0.05, ΔPSI > 0.1) in relation to *FLT3*/ITD and mutated *NPM1* in M1 patients. B - Network of all significant differential splicing events with the minimal splicing difference between groups of 0.1 (FDR < 0.05, ΔPSI > 0.1) in relation to *FLT3*/ITD and mutated *NPM1* in M2 patients. C - Subnetwork of genes involved in signalling pathways significantly differentially spliced (FDR < 0.05, ΔPSI > 0.1) in relation to *FLT3*/ITD and mutated *NPM1* in M4 patients. In each panel genes relevant for particular signalling pathways are highlighted. Node size indicates connectivity of the genes. Node fill color signifies the ΔPSI value for each differential splicing event (ranging from -0.25 to 0.25; blue - red) while the color of the node edge codes for the type of differential splicing event: alternative 3'splice site selection (A3SS, blue), alternative 5' splice site selection (A5SS, turquoise), retained intron (RI, yellow) and skipped exon (SE, brown).

cycle and DNA damage repair (26 genes in M1, 29 genes in M2 and 171 genes in M4 subtype, Figure 2), although the particular repertoires of genes implicated in these functions were FAB type-specific. This included for instance two components of the BRCA1-A complex (*BABAM1* and *BABAM2/BRE*) and *CEP164* an ATR/ATM signalling regulator in M1 samples and two genes coding for centromeric proteins (*CENPE* and *CENPJ*) as well as *PLK4*, a kinase which plays a central role in centriole duplication in M2 specimens. Factors controlling DNA damage response and cell cycle constituted a large network among the numerous differentially spliced genes found in *FLT3*/ITD+/*NPM1*+ M4 samples. This network included for instance genes with an established role in oncogenesis, such as *ATR*, *BRCA2*, *TOP2A*, *TOP2B*, the Aurora kinases (*AURKA* and *AURKB*) as well as *MELK* kinase, an important regulator involved in both cell cycle control and self-renewal and apoptosis.

Next to cell cycle control and DNA damage repair, several genes differentially spliced in relation to *FLT3*/ITD and *NPM1* mutations were involved in signalling pathways that regulate survival and proliferation of AML cells (Figure 3). In M1 patients, which is the most undifferentiated of the three analysed FAB types, the network of significantly differentially spliced genes in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples included *EZH2*, an important regulator of hematopoietic stem cells, as well as two genes that regulate development of embryonic stem cells (*RBBP5* and *JARID2*). The most pronounced perturbation of signalling in (the more mature) *FLT3*/ITD^{pos}/*NPM1*^{pos} M2 specimens involved NOTCH signalling (*FBXW1*, *RBX1*, *JAG1*, *NCOR2* and *HDAC6*, Figure 3B) and apoptosis regulation (i.e. *NME4* and *APIP*). *FLT3*/ITD^{pos}/*NPM1*^{pos} samples of M4 FAB type displayed differential splicing of many factors involved in survival signalling (the entire network in Figure 3C). Prominent examples include genes coding for subunits of phosphoinositide 3-kinase (PI3K), including three catalytic (*PIK3CA*, *PIK3CB* and *PIK3CG*) and one regulatory subunits (*PIK3R5*). Remarkably, the PI3K/AKT signalling pathway was found to be perturbed by differential expression in M1 and M2 FAB types but not in M4 specimens.

Interestingly, overall we found very little overlap between differentially spliced and differentially expressed genes in relation to concomitant *FLT3*/ITD and *NPM1* mutations (Supplemental Figure S7). While both types of regulation affected genes involved in various survival signalling pathways, the major processes regulated by differential expression and splicing varied with differentially expressed genes primarily implicated in hematopoietic differentiation (i.e. *HOX* genes, *FOXC1*, *MEIS1* and *FOXO1*, Supplemental Figure S6). In summary, differential splicing in *FLT3*/ITD^{pos}/*NPM1*^{pos} cells perturbed reg-ulators of processes highly relevant for oncogenesis, including progression through cell cycle and DNA damage response as well as survival signalling. Furthermore, these two types of gene expression regulation (differential expression and splicing) appear to complement each other in the two important aspects of oncogenesis that is uncontrolled proliferation and impaired differentiation.

FACTORS REGULATING DIFFERENTIAL SPLICING IN THE CONTEXT OF *FLT3*/ITD AND *NPM1* MUTATIONS

The differential splicing profiles found in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples could be the result of differential expression of splicing factors in this context. We did not find any splicing regulators to be differentially expressed in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples in none of the FAB types.

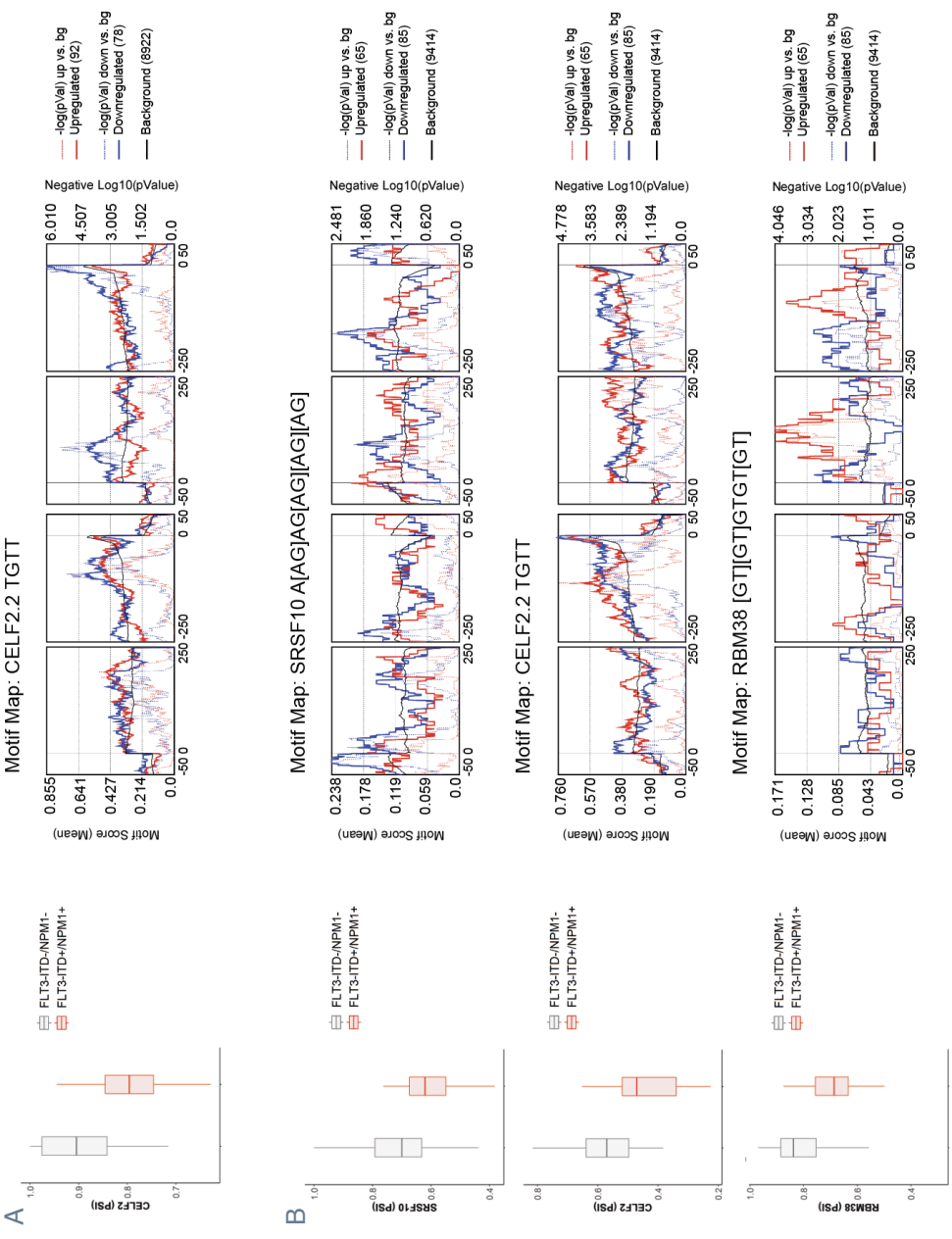
However, since splicing factors are known to autoregulate their own splicing, we also looked at differentially spliced splicing regulators (Figure 4 and Supplemental Figure S8). In the M1 subtype only *CELF2* splicing factor was found to be differentially spliced in *FLT3*/ITD^{pos}/*NPM1*^{pos} specimens. Notably, *FLT3*/ITD^{pos}/*NPM1*^{pos} M2 samples displayed altered splicing of 9 splicing regulators (*CELF2*, *RBM38*, *RBM39*, *DDX16*, *PUM1*, *SRSF10*, *PRMT7*, *ZRANB2* and *TFIP11*) and 18 splicing factors were differentially spliced in *FLT3*/ITD^{pos}/*NPM1*^{pos} M4 specimens (i.e. *CLK2*, *SRPK1*, *SRSF10*, *HNRNPC*, *HNRNPLL*, *PTBP1*, *RBM3* and *RBM5*).

To evaluate if these specific splicing factors possibly contributed to splicing regulation in the investigated sample set, we determined whether sequences (motifs) recognized and bound by these splicing regulators were enriched in the proximity to the significant differential splicing events, as compared to non-differentially spliced exons (using rMAPS tool, Figure 4 and Supplemental Figure S8). Remarkably, *CELF2* motifs were enriched in differential splicing events in *FLT3*/ITD^{pos}/*NPM1*^{pos} M1 and M2 samples. In M2, we also detected enrichment for *SRSF10*, *RBM38* and *PUM1* motifs. Similarly, differential splicing events identified in *FLT3*/ITD^{pos}/*NPM-1*^{pos} M4 specimens were enriched for *SRSF10*, *HNRNPC*, *PTBP1*, *RBM3*, *RBM5* and *HNRNPLL* motifs. Overall, these data indicate that differentially spliced splicing factors found in the current analysis are likely to at least partly contribute to the global changes in splicing profiles in the context of co-occurring *FLT3*/ITD and *NPM1* mutations in the three FAB types.

EVALUATION OF THE RELEVANCE OF DIFFERENTIAL SPLICING EVENTS

To gain more insight into the relevance of the identified splicing events, we next tested if they are likely to alter the function of the resultant proteins. To address this question, we evaluated whether sequences coding for functional protein domains were directly affected by selected splicing events using the Maser tool (Supplemental Figures S9-S10). In this analysis, we focused on genes involved in cell cycle, DNA repair and cell signalling specifically. Interestingly, the vast majority of AS events found in *FLT3*/ITD+/*NPM1*+ samples in all three FAB types were predicted to directly alter functional protein domains (M1: 83.6% , M2: 92.5%, M4: 86.2%) and are therefore likely to change or even abrogate the function of their corresponding proteins.

Finally, to further substantiate our findings, differentially spliced genes were validated in an independent sample set, which included 19 M1 samples, 17 M2 samples and 19 M4 specimens. We found that 33.3% of differentially spliced genes in *FLT3*/ITD^{pos}/*NPM1*^{pos} M1 samples showed a tendency (p-value < 0.2) towards differential splicing in the validation set (Supplemental Table S17). The same was true for 21.7% of genes in M2 subset (Supplemental Table S18) and 28.7% of M4 specimens (Supplemental Table S19 and selected examples that showed the same trend in discovery and validation cohorts in Figure 5 and Supplemental Figure S11). Taken together, while our analysis indicates large heterogeneity in splicing between samples, it does support the relevance of many of the identified differential splicing events in genes involved in regulation of cell cycle and DNA damage repair as well as signalling.


A
B
Motif Map: CELF2.2 TGTT
Motif Map: SRSF10 A[AG]AG[AG][AG][AG]
Motif Map: CELF2.2 TGTT
Motif Map: RBM38 [GT][GT]GTGT[GT]
Negative Log10(pValue)
Motif Score (Mean)
-log(pVal) up vs. bg
Upregulated (92)
-log(pVal) down vs. bg
Downregulated (78)
Background (8922)
Upregulated (65)
Downregulated (85)
Background (9414)
FLT3-ITD-/NPM1-
FLT3-ITD+/NPM1+
CELF2 (PSI)
SRSF10 (PSI)
RBM38 (PSI)

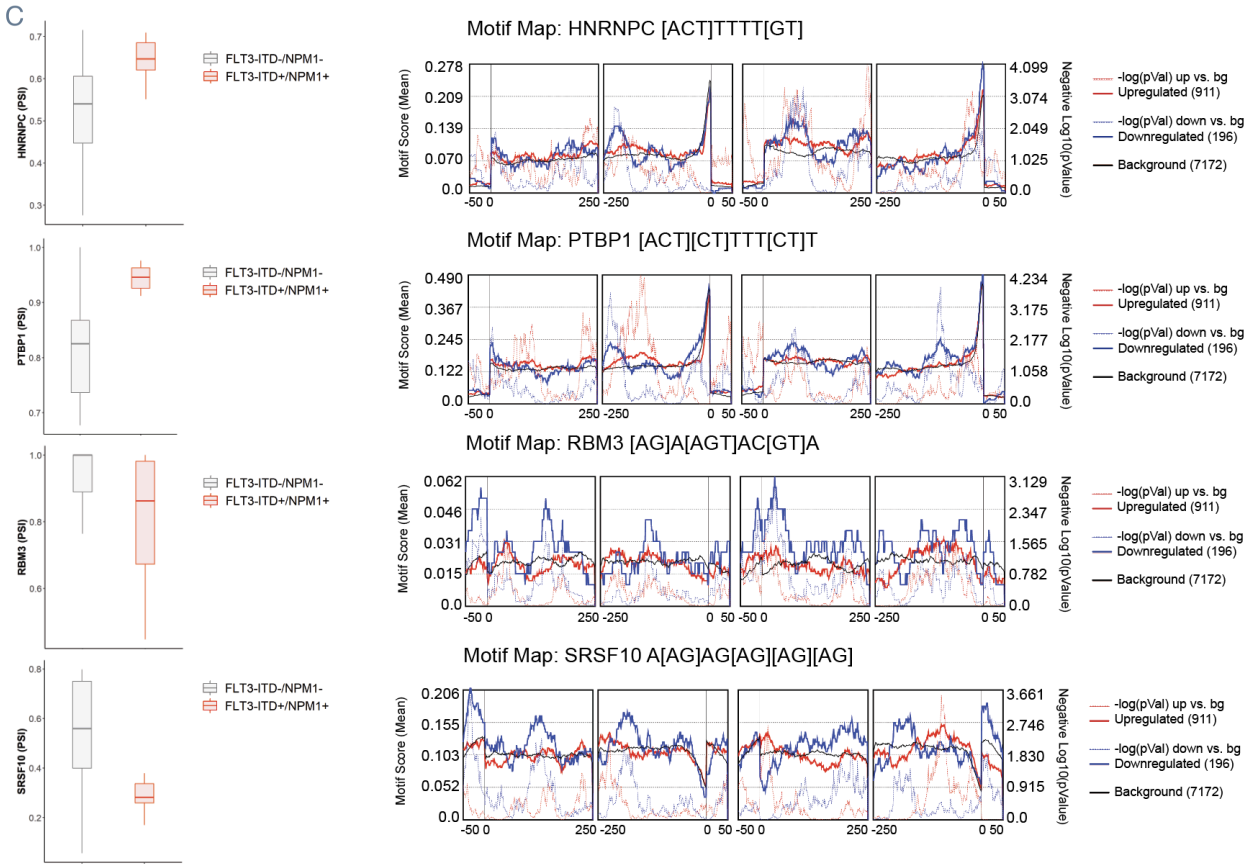


Figure 4. Differential splicing of splicing factors and motif enrichment analysis in relation to *FLT3*/ITD and *NPM1* mutations. The figure depicts PSI values for selected differentially spliced splicing regulators (FDR < 0.05 and ΔPSI > 0.1) in M1 (A), M2 (B) and M4 (C) FAB types as well as motif enrichment analysis for these splicing regulators performed by rMAPS2. This tool evaluates enrichment of motifs recognized by specific splicing factors in significantly differentially spliced events as compared to the background events (all splicing events detected by rMATS, including non-differential events). The motif enrichment is assessed in the differentially spliced exons as well as immediate upstream and downstream sequences. The motif enrichment score (the left y axis) is depicted by the solid blue (for events with ΔPSI < 0) and solid red (for events with ΔPSI > 0) lines. The solid black line indicates motif score for the background events. The negative logarithm of the p-value (right y axis) is depicted by the broken blue (for events with ΔPSI < 0) and red (for events with ΔPSI > 0) lines.

DISCUSSION

To the best of our knowledge, this is the first study to report differential splicing profiles associated with *FLT3*/ITD with a concomitant *NPM1* mutation in AML. It was previously demonstrated that *FLT3*/ITD collaborates with *NPM1* mutations in regulating chromatin state and gene expression profiles to drive AML.[23-25,52,53] Our data suggest that this cooperative regulation is further extended to alternative splicing. Importantly, we found that there appears to be no universal splicing profile associated with concomitant *FLT3*/ITD and *NPM1* mutations that would transcend all subtypes of AML cells. Instead, we find that co-occurrence of these two aberrations is associated with differential splicing of FAB subtype-specific sets of genes. This is in line with the crucial role of alternative splicing in differentiation of cells, including hematopoiesis.[22,23,52] While the FAB type specificity was very pronounced for differential splicing profiles, it also affected differential gene expression profiles. For instance, upregulation of *MEIS1* and *FOXC1*, previously reported to be associated with mutated *NPM1* and to regulate stem-like properties,[31] was only noted in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples of M1 and M2 FAB types, but not in M4 specimens. These observations indicate that the relevance of differential splicing and expression of important contributors to leukemogenesis is limited to certain differ-entiation stages of AML. Accordingly, a recent study reported that specific subsets of differentially expressed genes associated with relapse in AML only have prognostic value within specific molecular subsets (i.e. MLL rearranged) and FAB types.[22]

Interestingly, FAB type-specific differentially spliced genes were primarily involved in cell cycle control and DNA damage response, suggesting that perturbation of different genes involved in the same process could give similar outcome (i.e. deregulation of cell cycle). The normal physiological functions of NPM1 include maintenance of genomic stability by regulation of DNA repair and cell cycle progression.[4,22,23] Accordingly, mutated *NPM1* was previously linked to increased genomic instability and subsequent acquisition of additional mutations that activate signalling pathways (i.e. STAT or RAS).[58] Since *NPM1* mutations are thought to occur before *FLT3*/ITD, it is conceivable that differential splicing of genes involved in cell cycle regulation and DNA damage repair that we observed in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples arose due to *NPM1* mutations, or upon additional subsequent changes. The splicing perturbation of these processes could contribute to genomic instability, thereby facilitating acquisition of *FLT3*/ITD.

Strikingly, we noted functional divergence between differentially spliced and differentially expressed genes. While the first type of regulation primarily perturbed genes involved in cell cycle control and DNA damage response, the latter affected genes involved in hematopoietic differentiation. Since both of these processes constitute crucial and complementary aspects of oncogenesis, it appears that regulation at the level of gene expression and alternative splicing complement each other to drive development of AML. Finally, the extent of differential splicing in relation to *FLT3*/ITD and *NPM1* mutations in M4 FAB type was particularly large. This could be partly related to larger diversity of cells classified into this FAB type, which includes next to promyelocytes and more mature cells of the granulocytic lineage also more than 20% of cells with monocytic features.[5] FAB subtypes were used in the current study as an approximation of specific maturation stadia of AML cells. While this classification was useful for our pilot analy-

sis, subsequent studies should examine the maturation state specificity of differential expression and splicing in various purified immunophenotypic (and molecularly defined) subtypes of AML cells.

To a great extent, the detected splicing events were predicted to directly affect functional protein domains and therefore are likely to have an impact on the phenotype of AML cells. This should be further confirmed in functional studies. In addition, we found that 21.7 - 33.3% of differential splicing events from the discovery set showed a similar trend in an independent validation sample set. Although this analysis validated many events, it also suggests a relatively large heterogeneity in splicing between AML samples. Despite initially large number of AML patient specimens in the current study, the sample numbers in the subtype analyses were substantially lower. Therefore, splicing profiles of *FLT3*/ITD^{pos}/*NPM1*^{pos} cells should be further confirmed in larger datasets. This would also allow assessment of differential splicing in *FLT3*/ITD^{pos} samples without mutated *NPM1* (and vice versa), as well as focusing on *FLT3*/ITD^{pos} samples with high allelic ratio.

Finally, development and differentiation-related coordinated networks of alternative splicing events were previously reported to be orchestrated by RNA-binding proteins.[27–29] Although we did not find any differentially expressed splicing regulators in our dataset, we found differentially spliced splicing factors in each FAB subtype, for which binding motifs were enriched in the identified differential splicing events. Therefore, these regulators are likely to, at least partly, contribute to the differential splicing profiles of *FLT3*/ITD^{pos}/*NPM1*^{pos} AML samples. Future studies should confirm binding of these specific splicing factors in the vicinity of alternative splicing events. As the process of splicing occurs predominantly co-transcriptionally, its regulation is tightly coupled to transcription and chromatin status, including modifications to both histones and DNA.[59,60] Recent studies showed that mutated IDH2 as well as RUNX1 knockout alter splicing profiles.[61] Furthermore, dynamic changes in histone modifications were shown to predominantly occur in exons that were differentially spliced during differentiation of human embryonic stem cells, demonstrating that chromatin status can directly affect splicing, thereby driving cell differentiation.[61] Since both *FLT3*/ITD and *NPM1* mutations affect chromatin status, they are also likely to indirectly influence splicing profiles through changes in histone and DNA modifications. This could not be explored in the current dataset due to the lack of data on histone and DNA modification, but should be further evaluated to fully elucidate the mechanisms behind splicing regulation in the context of *FLT3*/ITD and mutated *NPM1*.

CONCLUSION

Altogether, these data shows that concomitant *FLT3*/ITD and *NPM1* mutations are associated with FAB type-specific altered splicing of genes with potential relevance for oncogenesis. Subgroup specific splicing analysis, stratified on FAB subtypes, pointed to important features (especially related to splicing) of cells carrying the same genetic aberrations (i.e. *FLT3*/ITD^{pos}/*NPM1*^{pos}) but arrested in a different stage of differentiation. Although FAB classification is not of prognostic use, maturation context-specific differential splicing analyses identified genes involved in critical cellular processes including regulation of DNA damage, and survival signalling. Based on functional relevance of such genes, alternative splicing could potentially affect the response of cells to (chemo- and targeted) therapy. Interesting examples of genes that could guide selection of cell type-specific therapeutic targets, include EZH2 in *FLT3*/ITD^{pos}/*NPM1*^{pos} M1 samples, Notch signalling in M2 patients and PI3K/AKT signalling or MELK kinase in M4 subtype. However, future studies should further explore the functional relevance of cell type-specific differential splicing in *FLT3*/ITD^{pos}/*NPM1*^{pos} AML cells in order to determine their impact on response to treatment and usefulness as novel therapeutic targets.

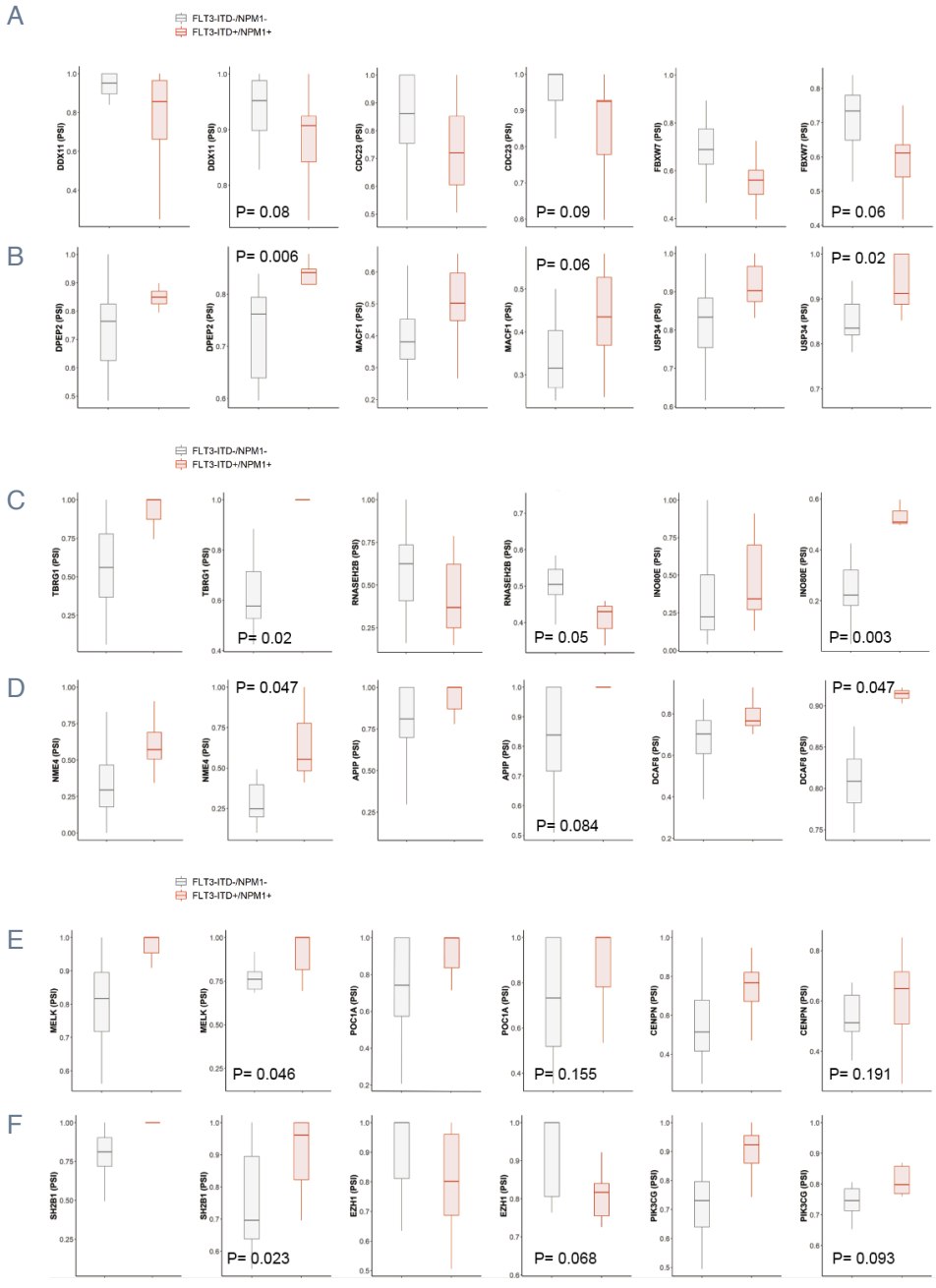


Figure 5. Validation of selected differential splicing events in an independent sample set. The figure depicts PSI values for selected differential splicing events in our initial discovery cohort (all events with FDR < 0.05 and ΔPSI > 0.1) and an independent validation cohort in M1 (A,B), M2 (C,D) and M4 (E,F) FAB types. A, C, E - Selected splicing events affecting genes involved in cell cycle regulation and DNA damage response in M1 (*DDX11*,*CDC23*, *FBXW7*), M2 (*TBRG1*, *RNASEH2B*, *INO80E*) and M4 (*MELK*, *POC1A*, *CENPN*) FAB types respectively. B, D, F - Selected splicing events affecting genes involved in signalling pathways in M1 (*DPEP2*, *MACF1*, *USP34*), M2 (*NME4*, *APIP*, *PCAF8*)and M4 (*SH2B1*, *EZH1*, *PIK3CG*) FAB types respectively.

REFERENCES

1. Papaemmanuil, E.; Gerstung, M.; Bullinger, L.; Gaidzik, V.I.; Paschka, P.; Roberts, N.D.; Potter, N.E.; Heuser, M.; Thol, F.; Bolli, N.; et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. N. Engl. J. Med. 2016, 374, 2209–2221, doi:10.1056/nejmoa1516192.
2. Genomic and Epigenomic Landscapes of Adult De Novo Acute Myeloid Leukemia. N. Engl. J. Med. 2013, 368, 2059–2074, doi:10.1056/nejmoa1301689.
3. Corces-Zimmerman, M.R.; Majeti, R. Pre-leukemic evolution of hematopoietic stem cells: The importance of early mutations in leukemogenesis. Leukemia 2014, 28, 2276–2282.
4. Corces-Zimmerman, M.R.; Hong, W.J.; Weissman, I.L.; Medeiros, B.C.; Majeti, R. Preleukemic mutations in human acute myeloid leukemia affect epigenetic regulators and persist in remission. Proc. Natl. Acad. Sci. U. S. A. 2014, 111, 2548–2553, doi:10.1073/pnas.1324297111.
5. Baralle, F.E.; Giudice, J. Alternative splicing as a regulator of development and tissue identity. Nat. Rev. Mol. Cell Biol. 2017, 18, 437–451.
6. Ajith, S.; Gazzara, M.R.; Cole, B.S.; Shankarling, G.; Martinez, N.M.; Mallory, M.J.; Lynch, K.W. Position-dependent activity of CELF2 in the regulation of splicing and implications for signal-responsive regulation in T cells. RNA Biol. 2016, 13, 569–581, doi:10.1080/15476286.2016.1176663.
7. Zhang, T.; Lin, Y.; Liu, J.; Zhang, Z.G.; Fu, W.; Guo, L.Y.; Pan, L.; Kong, X.; Zhang, M.K.; Lu, Y.H.; et al. Rbm24 Regulates Alternative Splicing Switch in Embryonic Stem Cell Cardiac Lineage Differentiation. Stem Cells 2016, 34, 1776–1789, doi:10.1002/stem.2366.
8. Wojtuszkiewicz, A.; Assaraf, Y.G.; Maas, M.J.P.; Kaspers, G.J.L.; Jansen, G.; Cloos, J. Pre-mRNA splicing in cancer: The relevance in oncogenesis, treatment and drug resistance. Expert Opin. Drug Metab. Toxicol. 2015, 11, doi:10.1517/17425255.2015.993316.
9. Pederiva, C.; Böhm, S.; Julner, A.; Farnebo, M. Splicing controls the ubiquitin response during DNA double-strand break repair. Cell Death Differ. 2016, 23, 1648–1657, doi:10.1038/cdd.2016.58.
10. Blencowe, B.J. Splicing regulation: The cell cycle connection. Curr. Biol. 2003, 13.
11. Moore, M.J.; Wang, Q.; Kennedy, C.J.; Silver, P.A. An alternative splicing network links cell-cycle control to apoptosis. Cell 2010, 142, 625–636, doi:10.1016/j.cell.2010.07.019.
12. Cambindo Botto, A.E.; Muñoz, J.C.; Giono, L.E.; Nieto-Moreno, N.; Cuenca, C.; Kornblihtt, A.R.; Muñoz, M.J. Reciprocal regulation between alternative splicing and the DNA damage response. Genet. Mol. Biol. 2020, 43, 1–6.
13. Kahles, A.; Lehmann, K. Van; Toussaint, N.C.; Hüser, M.; Stark, S.G.; Sachsenberg, T.; Stegle, O.; Kohlbacher, O.; Sander, C.; Caesar-Johnson, S.J.; et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. Cancer Cell 2018, 34, 211-224.e6, doi:10.1016/j.ccell.2018.07.001.
14. Climente-González, H.; Porta-Pardo, E.; Godzik, A.; Eyras, E. The Functional Impact of Alternative Splicing in Cancer. Cell Rep. 2017, 20, 2215–2226, doi:10.1016/j.celrep.2017.08.012.
15. Adamia, S.; Haibe-Kains, B.; Pilarski, P.M.; Bar-Natan, M.; Pevzner, S.; Avet-Loiseau, H.; Lode, L.; Verselis, S.; Fox, E.A.; Burke, J.; et al. AGenome-wide aberrantRNASplicing in patients with acute myeloid leukemia identifies novel potential disease markers and therapeutic targets. Clin. Cancer Res. 2014, 20, 1135–1145, doi:10.1158/1078-0432.CCR-13-0956.
16. 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. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. Cell Rep. 2015, 13, 1033–1045, doi:10.1016/j.celrep.2015.09.053.
17. Ilagan, J.O.; Ramakrishnan, A.; Hayes, B.; Murphy, M.E.; Zebari, A.S.; Bradley, P.; Bradley, R.K. U2AF1 mutations alter splice site recognition in hematological malignancies. Genome Res. 2015, 25, 14–26, doi:10.1101/gr.181016.114.
18. Kim, E.; Ilagan, J.O.; Liang, Y.; Daubner, G.M.; Lee, S.C.W.; Ramakrishnan, A.; Li, Y.; Chung, Y.R.; Micol, J.B.; Murphy, M.E.; et al. SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. Cancer Cell 2015, 27, 617–630, doi:10.1016/j.ccell.2015.04.006.

19. Zhang, J.; Lieu, Y.K.; Ali, A.M.; Penson, A.; Reggio, K.S.; Rabadan, R.; Raza, A.; Mukherjee, S.; Manley, J.L. Disease-associated mutation in SRSF2 misregulates splicing by altering RNA-binding affinities. Proc. Natl. Acad. Sci. U. S. A. 2015, 112, E4726–E4734, doi:10.1073/pnas.1514105112.
20. Metzeler, K.H.; Herold, T.; Rothenberg-Thurley, M.; Amler, S.; Sauerland, M.C.; Görlich, D.; Schneider, S.; Konstandin, N.P.; Dufour, A.; Bräundl, K.; et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. In Proceedings of the Blood; American Society of Hematology, 2016; Vol. 128, pp. 686–698.
21. Masson, K.; Rönnstrand, L. Oncogenic signaling from the hematopoietic growth factor receptors c-Kit and Flt3. Cell. Signal. 2009, 21, 1717–1726.
22. Zarka, J.; Short, N.J.; Kanagal-Shamanna, R.; Issa, G.C. Nucleophosmin 1 mutations in acute myeloid leukemia. Genes (Basel). 2020, 11, 1–16.
23. Dovey, O.M.; Cooper, J.L.; Mupo, A.; Grove, C.S.; Lynn, C.; Conte, N.; Andrews, R.M.; Pacharne, S.; Tzelepis, K.; Vijayabaskar, M.S.; et al. Molecular synergy underlies the co-occurrence patterns and phenotype of NPM1-mutant acute myeloid leukemia. Blood 2017, 130, 1911–1922, doi:10.1182/blood-2017-01-760595.
24. Rau, R.; Magoon, D.; Greenblatt, S.; Li, L.; Annesley, C.; Duffield, A.S.; Huso, D.; McIntyre, E.; Clohessy, J.G.; Reschke, M.; et al. NPMc+ cooperates with Flt3/ITD mutations to cause acute leukemia recapitulating human disease. Exp. Hematol. 2014, 42, 101-113. e5, doi:10.1016/j.exphem.2013.10.005.
25. Valk, P.J.M.; Verhaak, R.G.W.; Beijen, M.A.; Erpelinck, C.A.J.; van Doorn-Khosrovani, S.B. van W.; Boer, J.M.; Beverloo, H.B.; Moorhouse, M.J.; van der Spek, P.J.; Löwenberg, B.; et al. Prognostically Useful Gene-Expression Profiles in Acute Myeloid Leukemia. N. Engl. J. Med. 2004, 350, 1617–1628, doi:10.1056/nejmoa040465.
26. Cauchy, P.; James, S.R.; Zacarias-Cabeza, J.; Ptasinska, A.; Imperato, M.R.; Assi, S.A.; Piper, J.; Canestraro, M.; Hoogenkamp, M.; Raghavan, M.; et al. Chronic FLT3/ITD Signaling in Acute Myeloid Leukemia Is Connected to a Specific Chromatin Signature. Cell Rep. 2015, 12, 821–836, doi:10.1016/j.celrep.2015.06.069.
27. Luco, R.F.; Allo, M.; Schor, I.E.; Kornblihtt, A.R.; Misteli, T. Epigenetics in alternative pre-mRNA splicing. Cell 2011, 144, 16–26.
28. Luco, R.F.; Pan, Q.; Tominaga, K.; Blencowe, B.J.; Pereira-Smith, O.M.; Misteli, T. Regulation of alternative splicing by histone modifications. Science (80-.). 2010, 327, 996–1000, doi:10.1126/science.1184208.
29. Lev Maor, G.; Yearim, A.; Ast, G. The alternative role of DNA methylation in splicing regulation. Trends Genet. 2015, 31, 274–280.
30. Sciarrillo, R.; Wojtuszkiewicz, A.; Kooi, I.E.; Leon, L.G.; Sonneveld, E.; Kuiper, R.P.; Jansen, G.; Giovannetti, E.; Kaspers, G.J.L.; Cloos, J. Glucocorticoid resistant pediatric acute lymphoblastic leukemia samples display altered splicing profile and vulnerability to spliceosome modulation. Cancers (Basel). 2020, 12, doi:10.3390/cancers12030723.
31. Wiggers, C.R.M.; Baak, M.L.; Sonneveld, E.; Nieuwenhuis, E.E.S.; Bartels, M.; Creyghton, M.P. AML Subtype Is a Major Determinant of the Association between Prognostic Gene Expression Signatures and Their Clinical Significance. Cell Rep. 2019, 28, 2866-2877. e5, doi:10.1016/j.celrep.2019.08.012.
32. Höllein, A.; Jeromin, S.; Meggendorfer, M.; Fasan, A.; Nadarajah, N.; Kern, W.; Haferlach, C.; Haferlach, T. Minimal residual disease (MRD) monitoring and mutational landscape in AML with RUNX1-RUNX1T1: a study on 134 patients. Leukemia 2018, 32, 2270–2274, doi:10.1038/s41375-018-0086-0.
33. Cappelli, L.V.; Meggendorfer, M.; Dicker, F.; Jeromin, S.; Hutter, S.; Kern, W.; Haferlach, T.; Haferlach, C.; Höllein, A. DNMT3A mutations are over-represented in young adults with NPM1 mutated AML and prompt a distinct co-mutational pattern. Leukemia 2019, 33, 2741–2746.
34. Stengel, A.; Kern, W.; Haferlach, T.; Meggendorfer, M.; Fasan, A.; Haferlach, C. The impact of TP53 mutations and TP53 deletions on survival varies between AML, ALL, MDS and CLL: An analysis of 3307 cases. Leukemia 2017, 31, 705–711, doi:10.1038/leu.2016.263.
35. Baer, C.; Pohlkamp, C.; Haferlach, C.; Kern, W.; Haferlach, T. Molecular patterns in cytopenia patients with or without evidence of myeloid neoplasm—a comparison of 756 cases. Leukemia 2018, 32, 2295–2298, doi:10.1038/s41375-018-0119-8.

36. Pabst, T.; Vellenga, E.; Van Putten, W.; Schouten, H.C.; Graux, C.; Vekemans, M.C.; Biemond, B.; Sonneveld, P.; Passweg, J.; Verdonck, L.; et al. Favorable effect of priming with granulocyte colony-stimulating factor in remission induction of acute myeloid leukemia restricted to dose escalation of cytarabine. Blood 2012, 119, 5367–5373, doi:10.1182/blood-2011-11-389841.
37. Löwenberg, B.; Pabst, T.; Maertens, J.; Van Norden, Y.; Biemond, B.J.; Schouten, H.C.; Spertini, O.; Vellenga, E.; Graux, C.; Havelange, V.; et al. Therapeutic value of clofarabine in younger and middle-aged (18-65 years) adults with newly diagnosed AML. Blood 2017, 129, 1636–1645, doi:10.1182/blood-2016-10-740613.
38. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2009, 26, 139–140, doi:10.1093/bioinformatics/btp616.
39. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015, 43, e47, doi:10.1093/nar/gkv007.
40. Shen, S.; Park, J.W.; Lu, Z.X.; Lin, L.; Henry, M.D.; Wu, Y.N.; Zhou, Q.; Xing, Y. rMATS: Robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc. Natl. Acad. Sci. U. S. A. 2014, 111, E5593–E5601, doi:10.1073/pnas.1419161111.
41. Wickham, H. ggplot2 Elegant Graphics for Data Analysis; Use R!; Springer International Publishing: Cham, 2016; ISBN 978-3-319-24275-0.
42. Gu, Z.; Eils, R.; Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016, 32, 2847–2849, doi:10.1093/bioinformatics/btw313.
43. Bioconductor - maser Available online: https://bioconductor.org/packages/release/bioc/html/maser.html (accessed on Jun 24, 2021).
44. Bioconductor - AnnotationHub Available online: https://bioconductor.org/packages/release/bioc/html/AnnotationHub.html (accessed on Jun 24, 2021).
45. Lawrence, M.; Huber, W.; Pagès, H.; Aboyoun, P.; Carlson, M.; Gentleman, R.; Morgan, M.T.; Carey, V.J. Software for Computing and Annotating Genomic Ranges. PLoS Comput. Biol. 2013, 9, doi:10.1371/journal.pcbi.1003118.
46. Lawrence, M.; Gentleman, R.; Carey, V. rtracklayer: An R package for interfacing with genome browsers. Bioinformatics 2009, 25, 1841–1842, doi:10.1093/bioinformatics/btp328.
47. Park, J.W.; Jung, S.; Rouchka, E.C.; Tseng, Y.T.; Xing, Y. rMAPS: RNA map analysis and plotting server for alternative exon regulation. Nucleic Acids Res. 2016, 44, W333–W338, doi:10.1093/nar/gkw410.
48. Hwang, J.Y.; Jung, S.; Kook, T.L.; Rouchka, E.C.; Bok, J.; Park, J.W. rMAPS2: An update of the RNA map analysis and plotting server for alternative splicing regulation. Nucleic Acids Res. 2021, 48, W300–W306, doi:10.1093/NAR/GKAA237.
49. Szklarczyk, D.; Franceschini, A.; Kuhn, M.; Simonovic, M.; Roth, A.; Minguez, P.; Doerks, T.; Stark, M.; Muller, J.; Bork, P.; et al. The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011, 39, doi:10.1093/nar/gkq973.
50. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. Genome Res. 2003, 13, 2498–2504, doi:10.1101/gr.1239303.
51. Bindea, G.; Mlecnik, B.; Hackl, H.; Charoentong, P.; Tosolini, M.; Kirilovsky, A.; Fridman, W.H.; Pagès, F.; Trajanoski, Z.; Galon, J. ClueGO: A Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics 2009, 25, 1091–1093, doi:10.1093/bioinformatics/btp101.
52. Verhaak, R.G.W.; Goudswaard, C.S.; Van Putten, W.; Bijl, M.A.; Sanders, M.A.; Hugens, W.; Uitterlinden, A.G.; Erpelinck, C.A.J.; Delwel, R.; Löwenberg, B.; et al. Mutations in nucleophosmin (NPM1) in acute myeloid leukemia (AML): Association with other gene abnormalities and previously established gene expression signatures and their favorable prognostic significance. Blood 2005, 106, 3747–3754, doi:10.1182/blood-2005-05-2168.
53. Yun, H.; Vohra, S.; Mupo, A.; Giotopoulos, G.; Sasca, D.; Horton, S.J.; Agrawal-Singh, S.; Meduri, E.; Basheer, F.; Marando, L.; et al. Mutational Synergy Coordinately Remodels Chromatin Accessibility, Enhancer Landscape and 3-Dimensional DNA Topology to Alter Gene Expression during Leukemia Induction. Blood 2019, 134, 278–278, doi:10.1182/blood-2019-122413.
54. Gao, Y.; Vasic, R.; Halene, S. Role of alternative splicing in hematopoietic stem cells during development. Stem Cell Investig. 2018, 5.

55. Lemischka, I.R.; Pritsker, M. Alternative splicing increases complexity of stem cell transcriptome. Cell Cycle 2006, 5, 347–351.

56. Ullrich, S.; Guigó, R. Dynamic changes in intron retention are tightly associated with regulation of splicing factors and proliferative activity during B-cell development. Nucleic Acids Res. 2021, 48, 1327–1340, doi:10.1093/NAR/GKZ1180.

57. Reimer, K.A.; Neugebauer, K.M. Blood relatives: Splicing mechanisms underlying erythropoiesis in health and disease [version 1; referees: 3 approved]. F1000Research 2018, 7.

58. Qiu, J.; Zhou, B.; Thol, F.; Zhou, Y.; Chen, L.; Shao, C.; Deboever, C.; Hou, J.; Li, H.; Chaturvedi, A.; et al. Distinct splicing signatures affect converged pathways in myelodysplastic syndrome patients carrying mutations in different splicing regulators. RNA 2016, 22, 1535–1549, doi:10.1261/rna.056101.116.

59. Yoshimi, A.; Lin, K.T.; Wiseman, D.H.; Rahman, M.A.; Pastore, A.; Wang, B.; Lee, S.C.W.; Micol, J.B.; Zhang, X.J.; de Botton, S.; et al. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. Nature 2019, 574, 273–277, doi:10.1038/s41586-019-1618-0.

60. Huang, Y.-J.; Yan, M.; Chen, J.-Y.; Chen, L.; Kim, E.; Shima, T.; Davis, A.G.; Miyauchi, S.; Stoner, S.A.; Fu, X.-D.; et al. RUNX1 Deficiency Cooperates with SRSF2 Mutation to Further Disrupt RNA Splicing and Exacerbate Myelodysplastic Syndromes in Mouse Models. Blood 2018, 132, 1796–1796, doi:10.1182/blood-2018-99-119104.

61. Xu, Y.; Zhao, W.; Olson, S.D.; Prabhakara, K.S.; Zhou, X. Alternative splicing links histone modifications to stem cell fate decision. Genome Biol. 2018, 19, doi:10.1186/s13059-018-1512-3.

SUPPLEMENTAL INFORMATION

All Supplemental Information can be found online at https://www.mdpi.com/2072-6694/13/16/3929

CHAPTER 6

Selective Targeting of Pediatric Acute Myeloid Leukemia Stem and Progenitor Cell Splicing Deregulation

Inge van der Werf[1-3*]
Phoebe Mondala[1*]
Larisa Balaian[1]
Kathleen Steel[1]
Luisa Ladel[1]
Cayla Mason[1]
Raymond Diep[1]
Jacqueline Cloos[2]
Gertjan J.L. Kaspers[3,4]
Markus J. van Roosmalen[3]
Ruben van Boxtel[3]
Warren C. Chen[5]
Adam Mark[6]
Mary E. Donohoe[1]
James J. La Clair[5]
Peggy Wentworth[1]
Kathleen M. Fisch[6]
Leslie Crews[1]
Thomas Whisenant[6]
Michael D. Burkart[5]
Catriona Jamieson[1]

1. Division of Regenerative Medicine, Department of Medicine, Moores Cancer Center, University of California, San Diego, La Jolla, USA. 2. Dept. of Hematology, Amsterdam University Medical Center, VU University Medical Center, Cancer Center Amsterdam, Amsterdam, The Netherlands; 3. Princess Máxima Center for Pediatric Oncology, Utrecht,The Netherlands 4. Emma Children's Hospital Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Pediatric Oncology, Amsterdam, The Netherlands 5. Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, USA 6. Center for Computational Biology and Bioinformatics (CCBB), Department of Medicine, University of California, San Diego, La Jolla, USA

*These authors contributed equally to this work

Submitted

UNDER EMBARGO UNTILL PUBLISHED

CHAPTER 7

Decreased RBFOX2 Expression in Pediatric Acute Myeloid Leukemia

Inge van der Werf[1-3*]
Phoebe Mondala[1*]
Larisa Balaian[1]
Kathleen Steel[1]
Luisa Ladel[1]
Raymond Diep[1]
Jacqueline Cloos[2]
Gertjan J.L. Kaspers[3,4]
James J. La Clair[5]
Peggy Wentworth[1]
Kathleen M. Fisch[6]
Leslie Crews[1]
Thomas Whisenant[6]
Michael D. Burkart[5]
Catriona Jamieson[1]

1. Division of Regenerative Medicine, Department of Medicine, Moores Cancer Center, University of California, San Diego, La Jolla, USA. 2. Dept. of Hematology, Amsterdam University Medical Center, VU University Medical Center, Cancer Center Amsterdam, Amsterdam, The Netherlands; 3. Princess Máxima Center for Pediatric Oncology, Utrecht, The Netherlands 4. Emma Children's Hospital Amsterdam, Amsterdam UMC, Vrije Universiteit Amsterdam, Pediatric Oncology, Amsterdam, The Netherlands 5. Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, USA 6. Center for Computational Biology and Bioinformatics (CCBB), Department of Medicine, University of California, San Diego, La Jolla, USA

*These authors contributed equally to this work

Manuscript in Preparation

UNDER EMBARGO UNTILL PUBLISHED

CHAPTER 8

Summary and Discussion

Because of the expanded use of RNA sequencing of sorted cell populations and single cell RNA sequencing, cell identity is being increasingly defined by transcriptional profiles.[1] Transcriptional networks play a central role in governing stem cell function and fate. However, we now appreciate that transcriptional events do not entirely determine cellular identity. Recent studies have revealed that diverse post-transcriptional mechanisms, including alternative splicing, heavily influence the functional output of genetic programs required by (stem) cells (Chapter 2).[2-12] Concurrently, in acute myeloid leukemia (AML), a disease that originates from a primitive hematopoietic stem cell, genome-wide sequencing studies found the process of splicing to be disrupted on a surprisingly large scale.[13-15] Here, we set out to explore the relevance of disrupted pre-mRNA splicing in AML from different angles.

First, we aimed to characterize the clinical features associated with and prognostic impact of splicing factor mutations which have been recently found to be recurrently mutated in AML. Next, we studied the potential of splicing modulation in different AML subtypes and characterized differential splicing profiles of leukemic cells independent of these splicing factor mutations, as alternative splicing unveils to be a global phenomenon in hematological malignancies. Finally, we employed different strategies to determine the underlying biology and vulnerability of splicing dysregulation hematopoietic stem cells derived from pediatric AML specifically.

SPLICING FACTOR MUTATIONS IN AML

The impact of *SF3B1* mutations on prognosis in myelodysplastic syndrome (MDS) is already well established. In 2016, an updated version of the World Health Organization (WHO) classification for myeloid neoplasms and acute leukemia incorporated *SF3B1* mutations as a diagnostic criterion for MDS.[16-19] In this disease entity, mutations in *SF3B1* were linked with the presence of ring sideroblasts and are associated with a better prognosis. In contrast, in chronic lymphocytic leukemia (CLL), *SF3B1* mutations co-occur with *ATM* mutations and are indicative of poor treatment outcome and therapy resistance.[14,20,21] While recent studies suggested that SRSF2 mutations should be considered for incorporation into prognostic guidelines in AML, they have not been endorsed as adverse risk factors yet.[15,22,23] Chapter 3 entails the largest study to date focusing specifically on biological features, genetic background, and clinical relevance of splicing factor mutations in AML. Our findings underscore the overall unfavourable characteristics and inferior treatment outcome of AML patients with splicing factor mutations and suggest that incorporation of these aberrations into the current prognostic schemes could be beneficial.

In our study, we identified two factors that could drive further improvements in the prognostication of AML. Firstly, *SF3B1* mutations among patients classified as having favorable risk according to the ELN 2017 classification and, in particular, intermediate risk, marked individuals with worse overall survival (OS) and event-free survival (EFS), suggesting that these patients would benefit from more intensive treatment or innovative therapies, such as splicing modulation. Secondly, we found that patients carrying concomitant mutations in any of the studied splicing factors (either *SF3B1*, *SRSF2*, *U2AF1* or *ZRSR2*) and *RUNX1* mutations had a particularly poor prognosis in univariable and multivariable analyses. This effect was largely due to the particularly strong interaction of *SRSF2* and *RUNX1* mutations. Strikingly, patients

with *RUNX1* mutations without any of the four common splicing factor mutations had longer OS and EFS compared with the rest of patients within the adverse risk group according to the ELN 2017 classification. Hence, *RUNX1* mutations may not be relevant for risk assessment without the co-occurrence of any of the four most common splicing factor mutations.

Exclusion of patients carrying any of the four most common splicing factor mutations from the *RUNX1*-mutated subgroup had a stronger effect than did exclusion of only *SRSF2*-mutated patients only, highlighting the importance of considering mutations in any splicing factor as a subgroup. In addition, we found that this subgroup of patients with splicing factor mutations, associated with shorter OS and EFS in univariable and multivariable models. Just as reported by Papaemanuil et al, this suggests that splicing deregulation in general, via mutations in any splicing factor, confers a poor prognosis in AML.[15] Interestingly, the prognostic impact of splicing factor mutations was shown to be dependent on the disease entity. Not only within hematological disorders as illustrated above, but also between different cancer types.[24-26] Furthermore, while splicing factor mutations do occur in other types of cancer, they are extremely common in hematological disorders.[27-29] Accordingly, future studies should elucidate why hematopoietic cells have increased vulnerability for acquisition of these mutations.

In addition, studies should unravel how splicing factor mutations contribute to disease pathogenesis. Mutations in splicing factor genes were demonstrated to result in widespread changes in the transcriptome accompanied by altered hematopoiesis and are generally considered to be early leukemogenic events.[28,29] However, less is known about how these splicing alterations affect downstream cellular processes and oncogenic properties of mutant cells.[17,30,31] Despite this still nascent understanding of these mutations, a clinical trial to evaluate the safety, pharmacokinetics, and pharmacodynamics of splicing modulator H3B-8800 in MDS, CMML, or AML patients that carry a missense *SF3B1* mutation at a variant allele frequency of five percent or higher, has recently been initiated (NCT02841540).

THERAPEUTIC VULNERABILITY TO SPLICING MODULATION

Thus far, treatment with splicing modulators has mainly focused on MDS or AML carrying spliceosome mutations that have been studied in Chapter 3. As cancer cells bearing these mutations are dependent on wild-type spliceosome function, these cells have been shown to be preferentially killed by splicing modulators E7107 and H3B-8800.[32-36] In addition, oncogenic activation of MYC or high expression of MCL1 or BCL2A1 have been associated with hypersensitivity to splicing modulation.[37,38]

In Chapter 4 we demonstrated that AML patients carrying *FLT3*/ITD show increased sensitivity to SF3B1 modulators. These findings hold true for both E7107 and its newly developed analogue H3B-8800. Several lines of evidence from cell lines and primary patient material did indicate that splicing modulation holds potential as a novel therapeutic option for *FLT3*/ITD^{pos} AML with high allelic ratio (AR) and/or long ITD length, both subsets of AML that respond poorly to current treatment options. The increased response to SF3B1 modulators in cells with high AR and/or long ITD length specifically, did suggest a direct link between the presence of *FLT3*/ITD and the preferential sensitivity of these cells to splicing modulation. This hypothesis was

further supported by the drastic decrease of FLT3 protein expression levels that we observed upon treatment with splicing modulation.

However, the direct downstream events that drive vulnerability of *FLT3*/ITD^{pos} cells remain unclear. We established that neither splicing factor mutations nor elevated expression of *MYC* or *MCL1-L*, which have been associated with hypersensitivity to splicing modulation, did drive increased sensitivity of *FLT3*/ITD^{pos} cells in our study. Furthermore, we analysed the expression of an extensive set of splicing regulators using RNA-seq, but no significant differential expression was found between *FLT3*/ITD^{pos} and *FLT3*/ITD^{neg} diagnostic samples either. To gain more insight into differences in splicing between these two subgroups, we explored differential splicing profiles associated with *FLT3*/ITD^{pos} AML in Chapter 5.

While we did not identify clear differential splicing patterns that could distinguish *FLT3*/ ITD^{pos} from and *FLT3*/ITD^{neg} AML, this study showed that concomitant *FLT3*/ITD and *NPM1* mutations are associated with FAB type-specific altered splicing of genes with potential relevance for oncogenesis.Interestingly, we observed functional divergence between alternatively spliced and differentially expressed genes in *FLT3*/ITD^{pos}/*NPM1*^{pos} samples in three analyzed FAB types (M1, M2 and M4). While the first type of regulation primarily perturbed genes involved in cell cycle control and DNA damage response, the latter affected genes involved in hematopoietic differentiation. Since perturbations in both gene expression and alternative splicing affected crucial and complementary aspects of oncogenesis, it appears that both drive development of AML.

In agreement, alternative splicing was recently shown to be tightly coupled to transcription and chromatin status, including modifications to both histones and DNA.[39] In addition, both mutated *IDH2* as well as RUNX1 knockout were found to alter splicing profiles.[40,41] Furthermore, dynamic changes in histone modifications were shown to predominantly occur in exons that were differentially spliced during differentiation of human embryonic stem cells, demonstrating that chromatin status can directly affect splicing, thereby driving cell differentiation.[42-45] Since both *FLT3*/ITD and *NPM1* mutations affect chromatin status, identified splicing patterns could be explained by changes in histone and DNA modifications, which should be studied in future studies.

Thus, while diagnostics and prognostics in leukemia are mainly based on DNA mutations, cellular behavior is determined by a complex combination of the genome, transcription, post-transcriptional regulation, and translation. Moreover, the process of alternative splicing massively expands the coding capacity of our genome. Accordingly, we might miss a lot of information if we solely define a disease based on its DNA. Therefore, a better understanding on how genetics and transcriptome interact to drive leukemia is required.

In addition, future studies should further explore the functional relevance of cell type-specific differential splicing in *FLT3*/ITD^{pos}/*NPM1*^{pos} AML. However, the functional consequences of alternative (normal) or aberrant (mis-spliced) splicing have shown to be difficult to interpret as both our study and several other studies identified many genes to be affected by alternative

splicing.[46-50] As a consequence, a complex scenario with multi-transcriptomic changes might be responsible for the leukemia phenotype.

While the functional cell type and context specific consequences of splicing disruption will require further investigation, identified splice variants might already be of interest. For example, several recent studies have shown that tumor-specific alternative splicing may generate peptides that contribute to epitope repertoires.[25,51,52] These splicing-derived epitopes might serve as neoepitopes and elicit an endogenous immune response. Similarly, splicing modulation has been shown to generate neoantigens that can provoke an effective anti-tumor immune response.[53,54] Thus, despite the functional consequences of splicing disruption need to be elucidated, we have learned that innovative therapeutic strategies could be based on a synthetic lethal approach (e.g. splicing modulation in combination with immunotherapy or potentially in in *FLT3*/ITD^{pos} AML).[31,55] Importantly, we should to take into account that although SF3B1 modulation is a targeted therapy, it elicits a pleiotropic effect since alternative splicing affects 95% of the human genome as well as many crucial aspects of human physiology, including homeostasis, cell differentiation as well as cell survival.[2,3,50,56] As a result, it will be important to assess the right context for therapeutic vulnerability and the right dose for differential toxicity.

SPLICING DEREGULATION IN PEDIATRIC AML

In Chapter 6 we showed that pediatric AML samples are highly sensitive to splicing modulation regardless of genetic background. In this study we used splicing modulator 17S-FD-895. Strikingly, pediatric AML samples were shown to be significantly more sensitive to splicing modulation than adult *de novo* AML in both short and long-term clonogenicity assays which was further supported by differential gene expression and splicing analysis. Firstly, while pediatric AML-derived hematopoietic stem cells did not show differential expression of genes relative to adult AML-derived hematopoietic stem cells, they did show an abundance of differentially spliced genes. Next, we aimed to identify genes that were differentially spliced in pediatric samples only. Accordingly, we investigated the overlap between the skipped exon splicing events in the pediatric AML compared with non-leukemic samples and pediatric AML compared with adult AML samples. Subsequently, events unique to pediatric AML were retained and used as input into an enrichment analysis to identify associated pathways and functions. The subsequent analysis of pediatric unique splicing events uncovered widespread differential splicing of genes involved in mRNA splicing in hematopoietic stem and progenitor cells. Notably, splicing factors are known to autoregulate their own splicing and could possibly contribute to global splicing deregulation, which indicates that a concert of splicing regulators might be involved in global differential splicing patterns detected in pediatric AML.[57-59] Thus, while previous studies already pointed at substantial differences between the genomic landscapes of adult and pediatric AML, we showed that regulation of alternative splicing is distinct in both subtypes as well.[60]

Hence, beyond classes of genes expected to be associated with cancer like cell cycle and metabolic pathways, the process of mRNA splicing is enriched in performed network analysis of pediatric unique splicing events. In addition, we identified significantly increased expression of the interferon-inducible splice isoform *ADAR1 p150* compared to the non-interferon-responsive isoform *ADAR1 p110* which is vital for A-I editing. A-I editing can affect and modify splicing

decisions which could result in widespread splicing changes.[61-65] In particular, HPCs derived from pediatric AML present relative high RNA editing levels near exon junctions, which have previously been shown to influence splicing efficiency in a position independent manner.[63,66]

Concurrently, both pediatric leukemia derived hematopoietic stem and progenitor cells showed downregulation of RNA-Binding Fox 2, RBFOX2 compared to non-leukemic counterparts. Since this was the only significantly differentially expressed genes related to splicing in both cell subtypes, we explored the contribution of RBFOX2 to splicing deregulation in Chapter 7. Together with Muscleblind-Like RNA Binding (MBNL) proteins, as well as Pyrimidine Tract Binding (PTBP) proteins, decreased expression of RBFOX2 has been associated with an alternative splicing pattern that is differentially regulated between embryonic stem cells (ESCs) and differentiated cell types.[7,67] Accordingly, in this study we performed comprehensive RNA-seq analysis of RBFOX2 target genes which were identified in the oRNAment database, which uses an algorithm to predict genes with target binding sites, as well as an eCLIP study that reported RBFOX2 target genes.[68,69] In this analysis, we detected splicing events reminiscent of embryonic reversion. Furthermore, network analysis of RBFOX2 target genes revealed that the top signaling pathways affected by this crucial regulator corresponded with processes enriched among pediatric unique splicing events, further emphasizing the central role of RBFOX2 in pediatric AML biology. In agreement, we identified increased clonal capacity and sensitivity to splicing modulation upon RBFOX2 knockdown corresponding to features of pediatric AML derived cells. In addition, we found increased expression of *PTK2B-202* and *MCL1-L* upon knockdown, consonant to our findings in Chapter 6. Future studies should study the underlying mechanism of RBFOX2 downregulation and the contribution of key validated targets to AML biology.

Altogether, we are the first to present global splicing deregulation in pediatric AML independent of cytogenetic and mutational landscape. In addition, we indicate that alternative splicing of splicing regulators, RNA editing and decreased RBFOX2 expression could drive this phenomenon. However, a functional role for splicing deregulation in pediatric leukemogenesis has been described before. Recently, the splicing regulator MBNL1 was found to be deleted in a small proportion of pediatric AML patients. Interestingly, MBNL1 has also been shown to regulate essential alternative RNA splicing patterns in MLL-rearranged infant leukemia.[70] In addition, the fusion oncogene *RUNX1/RUNX1T1* was found to mediate alternative splicing and reorganize the transcriptional landscape in leukemia.[71] Furthermore, a recent study did associate specific splicing profiles with glucocorticoid-resistance in primary pediatric acute lymphoblastic leukemia.[72] Together, these studies establish a foundation for differential splicing as an important facet of pediatric leukemia.

GENOME-WIDE DETECTION OF ALTERNATIVE SPLICING

In this thesis, RNA-seq has enabled detailed analysis of splicing patterns. Accordingly, it is important to understand and interpret its limitations. In Chapter 6,7 and 8, we used the bioinformatic tool named rMATS to detect splicing events. rMATS is an event-based method that quantifies unique reads that support specific splicing events.[73,74,75] The common splice event types rMATS can identify include skipped exons, intron retentions, alternative 3' splice sites

and alternative 5' splice sites (A5SS). In addition, we used the software RSEM to acquire gene level counts based on transcript level abundance.[76] This way, we were able to obtain alignment of transcripts of interest including *MCL1-S* and *MCL1-L*, *CD44* splice variants which are important markers of cancer stem cells, or splice variant *PTK2B-202* which has been implicated in pathogenesis of secondary AML and other hematological malignancies.

While we can estimate gene and isoform expression levels from RNA-Seq data using RSEM based on transcript level abundance, we must consider that RNA-seq does not provide sufficient information as sequencing read in our studies only span 150 base pairs.[73,77] Consequently, a single read can align to different alternatively spliced transcripts derived from the same gene. Accordingly, both rMATS and RSEM only use unique reads to estimate abundance of a specific splicing event or transcript. In addition, it has to be emphasized that the output of our rMATS analysis in this thesis involves thousands of genes that are differentially spliced between specified groups. Subsequentely, we can explore the affected processes which are enriched in differentially spliced gene sets, however functional consequences of most splice variants are still unknown.

In future studies, the relatively recently developed method Ribosome Profiling (also known as ribosome sequencing; Ribo-seq, or active mRNA translation sequencing; ART-seq) could be used to increase our knowledge on the functional concequences of leukemia specific alternative splicing profiles. This method provides a 'snapshot' of all the ribosomes active in a cell at a specific time point.[78] Accordingly, this method can be used to estimate translation efficiency of alternatively spliced transcripts and help determine which transcripts are actively translated. In addition, we should study the role of non-coding RNAs such as long non-coding RNAs (lncRNAs) and microRNAs (miRNAs). However, it will still be difficult to determine the contribution of key validated targets to AML biology, since a complex scenario involving alternative splicing of thousands of genes might be responsible for the leukemia phenotypes.

CONCLUSION

AML is associated with a relatively low mutational load compared to other malignancies. In addition, children generally have less DNA damaging environmental exposures and do not exhibit clonal hematopoiesis related to advanced age. These observations suggest that changes on epigenetic level are heavily involved in pathogenesis of both adult and, in particular, pediatric AML. Accordingly, in this thesis, we identified a vital role for splicing deregulation in several aspects of AML biology. Furthermore, we determined that the use of splicing modulation has great potential in several AML subtypes. Firstly, we found that AML patients carrying spliceosome mutations have a relatively poor prognosis. This underscores the pivotal role of deregulated splicing in AML and suggests that those individuals could benefit from more intensive treatment regimens or innovative therapies such as splicing modulation. Secondly, application of splicing modulators should not be limited to this subgroup, as our data suggests that both adult AML patients carrying *FLT3*/ITD and pediatric AML patients could also benefit from this innovative treatment option. Furthermore, we established a foundation for differential splicing as an important facet of pediatric leukemia and adult AML patients carrying *FLT3*/ITD. The mechanisms underlying the identified splicing events and therapeutic vulnerability in both contexts require further investigation. Meanwhile, the results described in this thesis highlight the importance of alternative splicing in AML and provide valuable evidence which can stimulate further research in the future. These insights will increase our understanding of AML pathogenesis and most importantly, provide an opportunity to fulfil the unmet need for new therapeutics to improve outcome for these patients.

REFERENCES

1. Boyer LA, Lee TI, Cole MF, et al. Core transcriptional regulatory circuitry in human embryonic stem cells. Cell. 2005;122(6):947-956.
2. Chua BA, Van Der Werf I, Jamieson C, Signer RAJ. Post-Transcriptional Regulation of Homeostatic, Stressed, and Malignant Stem Cells. Cell Stem Cell. 2020;26(2):138-159.
3. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010;463(7280):457-463.
4. van den Berg PR, Budnik B, Slavov N, Semrau S. Dynamic post-transcriptional regulation during embryonic stem cell differentiation. bioRxiv. 2017:123497.
5. Williamson AJ, Smith DL, Blinco D, et al. Quantitative proteomics analysis demonstrates post-transcriptional regulation of embryonic stem cell differentiation to hematopoiesis. Mol Cell Proteomics. 2008;7(3):459-472.
6. Atlasi Y, Mowla SJ, Ziaee SA, Gokhale PJ, Andrews PW. OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells. Stem Cells. 2008;26(12):3068-3074.
7. Han H, Irimia M, Ross PJ, et al. MBNL proteins repress ES-cell-specific alternative splicing and reprogramming. Nature. 2013;498(7453):241-245.
8. Holm F, Hellqvist E, Mason CN, et al. Reversion to an embryonic alternative splicing program enhances leukemia stem cell self-renewal. Proceedings of the National Academy of Sciences. 2015;112(50):15444-15449.
9. Lu Y, Loh YH, Li H, et al. Alternative splicing of MBD2 supports self-renewal in human pluripotent stem cells. Cell Stem Cell. 2014;15(1):92-101.
10. Ng AHM, Khoshakhlagh P, Rojo Arias JE, et al. A comprehensive library of human transcription factors for cell fate engineering. Nat Biotechnol. 2021;39(4):510-519.
11. Bergen V, Lange M, Peidli S, Wolf FA, Theis FJ. Generalizing RNA velocity to transient cell states through dynamical modeling. Nat Biotechnol. 2020;38(12):1408-1414.
12. Kainov YA, Makeyev EV. A transcriptome-wide antitermination mechanism sustaining identity of embryonic stem cells. Nat Commun. 2020;11(1):361.
13. Darman RB, Seiler M, Agrawal AA, et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. Cell Rep. 2015;13(5):1033-1045.
14. Rossi D, Bruscaggin A, Spina V, et al. Mutations of the SF3B1 splicing factor in chronic lymphocytic leukemia: association with progression and fludarabine-refractoriness. Blood. 2011;118(26):6904-6908.
15. Papaemmanuil E, Gerstung M, Bullinger L, et al. Genomic Classification and Prognosis in Acute Myeloid Leukemia. New England Journal of Medicine. 2016;374(23):2209-2221.
16. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. Blood. 2016;127(20):2391-2405.
17. Nikpour M, Scharenberg C, Liu A, et al. The transporter ABCB7 is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts. Leukemia. 2013;27(4):889-896.
18. Damm F, Thol F, Kosmider O, et al. SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. Leukemia. 2012;26(5):1137-1140.
19. Malcovati L, Papaemmanuil E, Bowen DT, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood. 2011;118(24):6239-6246.
20. Zhang Z, Chen S, Chen S, et al. SF3B1 mutation is a prognostic factor in chronic lymphocytic leukemia: a meta-analysis. Oncotarget. 2017;8(41):69916-69923.
21. Damm F, Nguyen-Khac F, Fontenay M, Bernard OA. Spliceosome and other novel mutations in chronic lymphocytic leukemia and myeloid malignancies. Leukemia. 2012;26(9):2027-2031.

22. Metzeler KH, Herold T, Rothenberg-Thurley M, et al. Spectrum and prognostic relevance of driver gene mutations in acute myeloid leukemia. Blood. 2016;128(5):686-698.
23. Hou HA, Liu CY, Kuo YY, et al. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. Oncotarget. 2016;7(8):9084-9101.
24. Zhou Z, Gong Q, Wang Y, et al. The biological function and clinical significance of SF3B1 mutations in cancer. Biomarker Research. 2020;8(1):38.
25. Jayasinghe RG, Cao S, Gao Q, et al. Systematic Analysis of Splice-Site-Creating Mutations in Cancer. Cell Rep. 2018;23(1):270-281.e273.
26. Fu X-D, Ares M. Context-dependent control of alternative splicing by RNA-binding proteins. Nature Reviews Genetics. 2014;15(10):689-701.
27. Sebestyén E, Singh B, Miñana B, et al. Large-scale analysis of genome and transcriptome alterations in multiple tumors unveils novel cancer-relevant splicing networks. Genome Res. 2016;26(6):732-744.
28. Yoshida K, Ogawa S. Splicing factor mutations and cancer. Wiley Interdiscip Rev RNA. 2014;5(4):445-459.
29. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nature Medicine. 2016;22(9):976-986.
30. Seiler M, Peng S, Agrawal AA, et al. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. Cell Rep. 2018;23(1):282-296.e284.
31. Dalton WB, Helmenstine E, Walsh N, et al. Hotspot SF3B1 mutations induce metabolic reprogramming and vulnerability to serine deprivation. J Clin Invest. 2019;129(11):4708-4723.
32. Seiler M, Yoshimi A, Darman R, et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nature Medicine. 2018;24(4):497-504.
33. Buonamici S, Yoshimi A, Thomas M, et al. H3B-8800, an Orally Bioavailable Modulator of the SF3b Complex, Shows Efficacy in Spliceosome-Mutant Myeloid Malignancies. Blood. 2016;128(22):966-966.
34. Lee SC, North K, Kim E, et al. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018;34(2):225-241.e228.
35. Bonnal SC, López-Oreja I, Valcárcel J. Roles and mechanisms of alternative splicing in cancer — implications for care. Nature Reviews Clinical Oncology. 2020;17(8):457-474.
36. Desterro J, Bak-Gordon P, Carmo-Fonseca M. Targeting mRNA processing as an anticancer strategy. Nature Reviews Drug Discovery. 2020;19(2):112-129.
37. Hsu TY, Simon LM, Neill NJ, et al. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature. 2015;525(7569):384-388.
38. Aird D, Teng T, Huang C-L, et al. Sensitivity to splicing modulation of BCL2 family genes defines cancer therapeutic strategies for splicing modulators. Nature Communications. 2019;10(1).
39. Yoshimi A, Lin K-T, Wiseman DH, et al. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. Nature. 2019;574(7777):273-277.
40. Huang Y-J, Yan M, Chen J-Y, et al. RUNX1 Deficiency Cooperates with SRSF2 Mutation to Further Disrupt RNA Splicing and Exacerbate Myelodysplastic Syndromes in Mouse Models. Blood. 2018;132(Supplement 1):1796-1796.
41. Xu Y, Zhao W, Olson SD, Prabhakara KS, Zhou X. Alternative splicing links histone modifications to stem cell fate decision. Genome Biology. 2018;19(1):133.
42. Cauchy P, James SR, Zacarias-Cabeza J, et al. Chronic FLT3-ITD Signaling in Acute Myeloid Leukemia Is Connected to a Specific Chromatin Signature. Cell Rep. 2015;12(5):821-836.
43. Luco RF, Allo M, Schor IE, Kornblihtt AR, Misteli T. Epigenetics in alternative pre-mRNA splicing. Cell. 2011;144(1):16-26.
44. Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of alternative splicing by histone modifications. Science. 2010;327(5968):996-1000.

45. Lev Maor G, Yearim A, Ast G. The alternative role of DNA methylation in splicing regulation. Trends Genet. 2015;31(5):274-280.
46. Chen SL, Dai YJ, Hu F, Wang Y, Li H, Liang Y. Effects of Alternative Splicing Events on Acute Myeloid Leukemia. DNA Cell Biol. 2020;39(11):2040-2051.
47. Bowman TV. Improving AML Classification Using Splicing Signatures. Clin Cancer Res. 2020;26(14):3503-3504.
48. Jin P, Tan Y, Zhang W, Li J, Wang K. Prognostic alternative mRNA splicing signatures and associated splicing factors in acute myeloid leukemia. Neoplasia. 2020;22(9):447-457.
49. Anande G, Deshpande NP, Mareschal S, et al. RNA Splicing Alterations Induce a Cellular Stress Response Associated with Poor Prognosis in Acute Myeloid Leukemia. Clin Cancer Res. 2020;26(14):3597-3607.
50. Sciarrillo R, Wojtuszkiewicz A, Assaraf YG, et al. The role of alternative splicing in cancer: From oncogenesis to drug resistance. Drug Resist Updat. 2020;53:100728.
51. Kahles A, Lehmann KV, Toussaint NC, et al. Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. Cancer Cell. 2018;34(2):211-224.e216.
52. Smart AC, Margolis CA, Pimentel H, et al. Intron retention is a source of neoepitopes in cancer. Nature Biotechnology. 2018;36(11):1056-1058.
53. Bowling EA, Wang JH, Gong F, et al. Spliceosome-targeted therapies trigger an antiviral immune response in triple-negative breast cancer. Cell. 2021;184(2):384-403.e321.
54. Lu SX, De Neef E, Thomas JD, et al. Pharmacologic modulation of RNA splicing enhances anti-tumor immunity. Cell. 2021;184(15):4032-4047.e4031.
55. Zhou Q, Derti A, Ruddy D, et al. A chemical genetics approach for the functional assessment of novel cancer genes. Cancer Res. 2015;75(10):1949-1958.
56. Wojtuszkiewicz A, Assaraf YG, Maas MJ, Kaspers GJ, Jansen G, Cloos J. Pre-mRNA splicing in cancer: the relevance in oncogenesis, treatment and drug resistance. Expert Opin Drug Metab Toxicol. 2015;11(5):673-689.
57. Damianov A, Black DL. Autoregulation of Fox protein expression to produce dominant negative splicing factors. Rna. 2010;16(2):405-416.
58. Gonçalves V, Jordan P. Posttranscriptional Regulation of Splicing Factor SRSF1 and Its Role in Cancer Cell Biology. Biomed Res Int. 2015;2015:287048.
59. Sun S, Zhang Z, Sinha R, Karni R, Krainer AR. SF2/ASF autoregulation involves multiple layers of post-transcriptional and translational control. Nat Struct Mol Biol. 2010;17(3):306-312.
60. Brunner AM, Graubert TA. Genomics in childhood acute myeloid leukemia comes of age. Nature Medicine. 2018;24(1):7-9.
61. Jiang Q, Isquith J, Zipeto MA, et al. Hyper-Editing of Cell-Cycle Regulatory and Tumor Suppressor RNA Promotes Malignant Progenitor Propagation. Cancer Cell. 2019;35(1):81-94.e87.
62. Jiang Q, Isquith J, Ladel L, et al. Inflammation-driven deaminase deregulation fuels human pre-leukemia stem cell evolution. Cell Reports. 2021;34(4).
63. Kapoor U, Licht K, Amman F, et al. ADAR-deficiency perturbs the global splicing landscape in mouse tissues. Genome Research. 2020;30(8):1107-1118.
64. Zipeto MA, Court AC, Sadarangani A, et al. ADAR1 Activation Drives Leukemia Stem Cell Self-Renewal by Impairing Let-7 Biogenesis. Cell Stem Cell. 2016;19(2):177-191.
65. Hartner JC, Walkley CR, Lu J, Orkin SH. ADAR1 is essential for the maintenance of hematopoiesis and suppression of interferon signaling. Nat Immunol. 2009;10(1):109-115.
66. Beghini A, Ripamonti CB, Peterlongo P, et al. RNA hyperediting and alternative splicing of hematopoietic cell phosphatase (PTPN6) gene in acute myeloid leukemia. Hum Mol Genet. 2000;9(15):2297-2304.
67. Venables JP, Lapasset L, Gadea G, et al. MBNL1 and RBFOX2 cooperate to establish a splicing programme involved in pluripotent stem cell differentiation. Nature Communications. 2013;4(1).

68. Benoit Bouvrette LP, Bovaird S, Blanchette M, Lécuyer E. oRNAment: a database of putative RNA binding protein target sites in the transcriptomes of model species. Nucleic Acids Res. 2020;48(D1):D166-d173.
69. Van Nostrand EL, Freese P, Pratt GA, et al. A large-scale binding and functional map of human RNA-binding proteins. Nature. 2020;583(7818):711-719.
70. Itskovich SS, Gurunathan A, Clark J, et al. MBNL1 regulates essential alternative RNA splicing patterns in MLL-rearranged leukemia. Nat Commun. 2020;11(1):2369.
71. Grinev VV, Barneh F, Ilyushonak IM, et al. RUNX1/RUNX1T1 mediates alternative splicing and reorganises the transcriptional landscape in leukemia. Nature Communications. 2021;12(1):520.
72. Sciarrillo R, Wojtuszkiewicz A, Kooi IE, et al. Glucocorticoid Resistant Pediatric Acute Lymphoblastic Leukemia Samples Display Altered Splicing Profile and Vulnerability to Spliceosome Modulation. Cancers. 2020;12(3):723.
73. Mehmood A, Laiho A, Venäläinen MS, McGlinchey AJ, Wang N, Elo LL. Systematic evaluation of differential splicing tools for RNA-seq studies. Briefings in Bioinformatics. 2019;21(6):2052-2065.
74. Shen S, Park JW, Huang J, et al. MATS: a Bayesian framework for flexible detection of differential alternative splicing from RNA-Seq data. Nucleic Acids Res. 2012;40(8):e61.
75. Muller IB, Meijers S, Kampstra P, et al. Computational comparison of common event-based differential splicing tools: practical considerations for laboratory researchers. BMC Bioinformatics. 2021;22(1):347.
76. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323.
77. Hu Y, Huang Y, Du Y, et al. DiffSplice: the genome-wide detection of differential splicing events with RNA-seq. Nucleic Acids Res. 2013;41(2):e39.
78. Andreev DE, O'Connor PB, Loughran G, Dmitriev SE, Baranov PV, Shatsky IN. Insights into the mechanisms of eukaryotic translation gained with ribosome profiling. Nucleic Acids Res. 2017;45(2):513-526.

APPENDIX &

Dutch Introduction and Summary
Curriculum Vitae
Publications
Awards
Acknowledgements

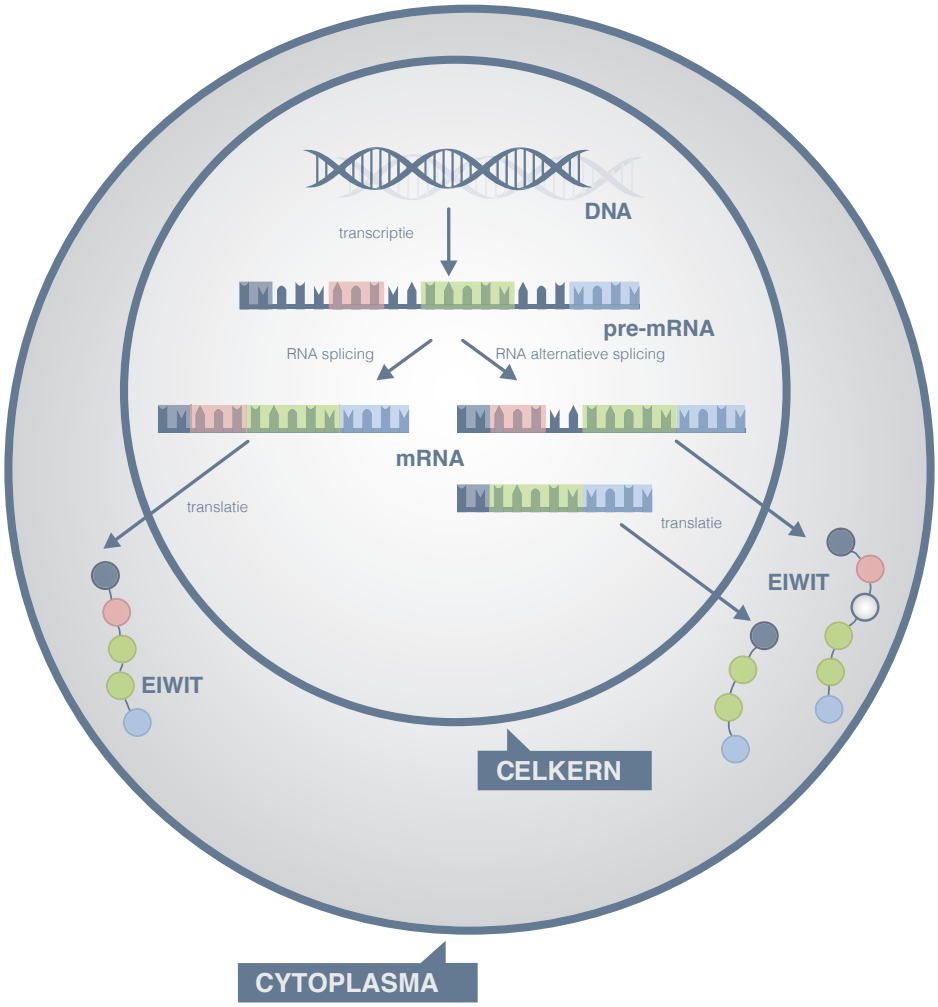
ACUTE LEUKEMIE

Leukemie is een vorm van bloedkanker waarbij de aanmaak van cellen in het beenmerg in het gedrang komt door een toename van abnormale, onrijpe, witte bloedcellen met ongecontroleerde groei (leukemicellen, ook wel blasten genoemd).[1] Als de ziekte geconstateerd wordt, is het zaak dat de patiënt snel wordt behandeld met chemotherapie, waarbij de leukemiecellen met ongecontroleerde groei worden geëlimineerd. Ondanks dat de behandeling in de loop der jaren aanzienlijk is verbeterd, is de terugkeer van de ziekte, een recidief, na een in eerste instantie ogenschijnlijk succesvolle therapie, een veelvoorkomend probleem.[2-4] De terugkeer van de ziekte lijkt te worden veroorzaakt door een populatie van cellen met stamcelcapaciteiten. Deze zeer laagfrequente celpopulatie is na de chemotherapie meestal niet meer detecteerbaar, maar is in staat zichzelf in stand te houden, en op een later moment, opnieuw uit te groeien tot een recidief.[5] Deze celpopulatie kan de chemotherapie overleven omdat ze ongevoelig is voor huidige therapieën. Het recidief is dus moeilijk behandelbaar met de chemotherapie die momenteel ter beschikking is. Daarom is van het van cruciaal belang om nieuwe therapieën te ontwikkelen met andere werkingsmechanismen. We richten ons daarbij op processen die verstoord zijn in leukemische cellen in vergelijk met gezonde cellen, zodat er bij systemische behandeling, minder bijwerkingen te verwachten zijn. Standaardtherapie richt zich vooral op het veroorzaken van DNA-schade in de cellen. Het onderzoek in dit proefschrift zich op het RNA, wat wordt afgeleid van het DNA maar een meer functionele rol heeft in de biologische functionaliteit van de cel.

WAT IS 'SPLICING'?

De blueprint van alle biologische processen ligt vast in het DNA. DNA bevindt zich, opgerold in de chromosomen, in de kern van een cel. Het DNA bevat onze genen die coderen voor RNA, wat codeert voor de functionele eiwitten. Eind jaren 70 ontdekten Richard John Roberts en Phillip Allen Sharp dat humane genen bestaan uit eiwit-coderende regio's (exonen) onderbroken door niet-coderende regio's (intronen), een ontdekking waarvoor zij in 1993 een Nobelprijs ontvingen.[6,7] Voordat het RNA kan worden omgezet naar eiwit heeft het een extra bewerking nodig, waarbij prematuur RNA (pre-mRNA) wordt omgezet naar matuur RNA (mRNA). Ofwel, tijdens de vertaling van DNA naar RNA, transcriptie, worden eerst zowel de exonen als intronen omgezet naar pre-mRNA. Vervolgens worden de intronen verwijderd en de exonen samengevoegd in een proces genaamd 'splicing' (zie Figuur 1).[8] Via deze weg wordt een functioneel mRNA gevormd dat kan worden omgezet naar eiwit, dit noemen wij translatie.

Inmiddels is bekend dat dit proces uitermate belangrijk is voor het functioneren van de cel en zeer stringent wordt gereguleerd. Strikte regulatie is van groot belang aangezien door 'splicing' uiteindelijk veel verschillende eiwitvarianten worden gemaakt. Intronen kunnen achterblijven of exonen kunnen geheel of gedeeltelijk worden verwijderd, waardoor vanaf één enkel gen vele alternatieve transcripten (varianten) kunnen worden gevormd. Meer dan 90% van onze genen wordt door deze alternatieve 'splicing' bewerkt, waardoor ons genoom een veel uitgebreider palet van functionaliteit tot beschikking heeft dan op basis van het aantal genen zou worden verwacht.[9]



Figuur 1. Alternatieve Splicing. DNA bevindt zich in de chromosomen die zitten opgeslagen in de kern van een cel. Het DNA codeert voor mRNA, wat codeert voor de functionele eiwitten. Voordat het mRNA kan worden omgezet naar eiwit heeft het een extra bewerking nodig, pre-mRNA wordt omgezet naar mRNA. Dit proces noemen we 'splicing'. Hierbij worden delen die niet voor eiwit coderen (intronen) uit het pre-mRNA geknipt, zodat alleen de coderende delen (exonen) overblijven en kunnen worden samen- gevoegd. Hiernaast kunnen de exonen en intronen op verschillende manieren worden gecombineerd. Dit proces kennen wij als alternatieve 'splicing'. Via deze weg kunnen verschillende transcripten, ofwel mRNA-varianten, worden gevormd. Als resultaat kan ons lichaam meer dan 100.000 eiwitten produceren, terwijl wij maar 20.000 genen tot onze beschikking hebben.

Het proces van 'splicing' wordt gefaciliteerd door het spliceosoom, een groot en dynamisch complex bestaande uit verschillende RNA-moleculen, wat de specifieke binding aan het pre-mRNA verzorgt, samen met meer dan 150 functionele eiwitten.[8,10,11] Het spliceosoom herkent specifieke sequenties die het begin en het einde van intronen aanduiden. Door middel van een wisselwerking van de eiwitten met deze specifieke RNA-sequenties worden intronen geknipt en exonen samengevoegd. In Chapter 2 wordt de rol van dit proces in stamcel functie beschreven.

MUTATIES IN 'SPLICING' REGULATOREN

Aangezien de transcripten van meer dan 90% van onze genen door middel van alternatieve 'splicing' wordt bewerkt, heeft een afwijking in het proces grote gevolgen. Mutaties in genen die coderen voor eiwitten die 'splicing' initiëren en faciliteren, resulteren in enorme veranderingen over het gehele mRNA-spectrum (transcriptoom) van een cel, wat grote gevolgen heeft voor het functioneren van de cel. Deze mutaties worden waargenomen in verschillende typen kanker, maar komen in het bijzonder vaak voor in hematologische maligniteiten.[10,12-15] In Chapter 3 beschrijven we de implicaties van deze mutaties in volwassen patiënten met AML. Een van de mutaties die tot deze categorie behoort, is een mutatie in 'splicing factor 3b subunit 1' (*SF3B1*) gen. Ditzelfde geldt voor mutaties in het 'U2 small nuclear RNA auxiliary factor 1' (*U2AF1*) gen, dat net als *SF3B1* codeert voor een belangrijke component van het spliceosoom, en het 'serine and arginine rich splicing factor 2' (*SRSF2*) gen, dat een belangrijke rol vertolkt in initiatie van 'splicing' in een scala aan genen.

ALTERNATIEVE SPLICING IN AML

Ondanks dat mutaties in *SF3B1*, *U2AF1* en *SRSF2* lang niet bij alle patiënten met acute leukemie worden waargenomen, speelt alternatieve 'splicing' praktisch altijd een rol in de ontwikkeling van leukemie. Ons genoom bevat ongeveer 20.000 genen, maar wij kunnen meer dan 100.000 eiwitten produceren. Alternatieve 'splicing' is hiervoor essentieel en de leukemische cellen gebruiken specifieke eiwitvarianten in hun voordeel.

Een van de eerste genen waarvan alternatieve varianten zijn geïdentificeerd is het apoptotische 'B-cell lymphoma-extra' (*BCL-X*) gen. 'Splicing' van het *BCL-X* pre-mRNA kan resulteren in de vorming van twee varianten: een korte variant met pro-apoptotische eigenschappen (*BCL-XS*) en een lange variant gekenmerkt door anti-apoptotische effecten (*BCL-XL*). Bij leukemie is de expressie van BCL-XL gerelateerd aan resistentie tegen conventionele therapieen, mede doordat deze variant apoptose onderdrukt en de leukemische cel op deze manier beschermt. Naast *BCL-X* zijn verschillende genen met tegenstrijdige functies geïdentificeerd, waaronder *MCL-1*, *Caspase-2*, *FAS* en *Survivin-2B*, waarvan de anti-apoptotische variant wordt geassocieerd met therapieresistentie bij verschillende tumoren.[16,17] Deze veranderingen in alternatieve 'splicing' van apoptotische genen laat zien dat 'splicing' een belangrijke rol speelt tijdens leukemogenese.

Hiernaast zijn in de loop der jaren aberrante varianten geïdentificeerd die uitsluitend in leukemische cellen worden waargenomen. Ongeveer 30% van de transcripten in het beenmerg van patiënten met AML is verschillend ten opzichte van gezonde donoren. Zodoende worden ook

zonder mutaties in genen zoals *SF3B1*, *U2AF1* en *SRSF2*, vele afwijkingen in het transcriptoom waargenomen.

BEHANDELINGEN GERICHT TEGEN 'SPLICING' REGULATOREN

In de afgelopen jaren zijn meerdere middelen ontwikkeld die de uitkomst van 'splicing' kunnen beïnvloeden.[18] Veel van deze middelen, zoals Spliceostatin A, Meayamycin B, Pladienolide B, en meer recent ontwikkelde E7107, 17S-FD-895 en H3B-8800, moduleren het proces van 'splicing' via binding aan zowel wildtype als gemuteerd SF3B1.[19-21] Als gevolg van blootstelling van cellen aan deze middelen wordt de binding van het SF3B1-eiwit met het spliceosoom ontwricht, wat resulteert in een ophoping van het pre-mRNA, een blok in de celcyclus en DNA-schade met celdood als gevolg.

Door het globale effect van de verschillende SF3B1-modulatoren zal moeten worden gezocht naar patiënten die baat hebben bij modulatie van dit complexe proces. Tot dusver lijken cellen met een mutatie in het *MYC*-gen relatief sensitief voor 'splicing'-modulatie. Een mutatie in het *MYC*-gen resulteert in verhoogde transcriptie, wat leidt tot druk op het proces van 'splicing', wat dit fenomeen kan verklaren.[22] Hiernaast lijken leukemiecellen van patiënten met hoge expressie van MCL1 of met een mutatie in *SF3B1*, *U2AF1* en *SRSF2* zeer gevoelig voor behandeling met E7107.[19,20,23,24] Met deze kennis is H3B-8800 ontwikkeld, waarmee recentelijk een nieuwe fase 1-studie is gestart bij MDS- en AML-patiënten met een mutatie in een van deze genen (NCT02841540).[25]

Aangezien alternatieve 'splicing' ook een rol speelt in patiënten zonder een mutatie in een van deze splicing regulator genen, zijn we in Chapter 4 op zoek gegaan naar een subtype AML dat ook van deze innovatieve modulatoren zou kunnen profiteren. Hierbij hebben we cellen met een *FLT3*/ITD geïdentificeerd. Cellen met deze mutatie zijn zeer gevoelig voor de SF3B1 modulatoren E7107 en H3B-8800. Daarbij zijn cellen met een hoge 'allelic ratio', een maat voor de hoeveelheid cellen met de mutatie, en lange ITD-lengte, extra gevoelig. Deze bevinding wijst op een directe link tussen de aanwezigheid van een *FLT3*-mutatie en het effect van de splicing modulator. In Chapter 5 bekijken we hoe de relatief veelvoorkomende *FLT3*/ITD en *NPM1* mutaties het proces van alternatieve splicing beïnvloeden. In dit hoofdstuk laten we zien dat *FLT3*/ITD en *NPM1* gemuteerde cellen geassocieerd zijn met de rijpingstoestand-specifieke differentiële splicing van genen met potentieel oncogene relevantie.

Hiernaast hebben we in Chapter 6 cellen van kinderen met AML behandeld met splicing modulator 17S-FD-895. Ook deze cellen zijn zeer sensitief, zeker in vergelijking tot leukemiecellen geïsoleerd van volwassenen met AML. Analyse van alternatieve splicing patronen in kinderen met AML laat zien dat deze afwijken van de patronen in het normale hematopoietische systeem en dat van volwassenen met AML. In kinderen met AML vertonen waargenomen splicing patronen overeenkomsten met die van embryonaly stamcellen. Hiernaast zien wij dat alternatieve splicing van veel genen, die belangrijk zijn voor de overleving van leukemiecellen, anders is in kinderen met AML. In Chapter 7 laten we zien dat relatief lage expressie van RBFOX2 de verschillen in gevoeligheid na behandeling met 17S-FD-895, tussen non-leukemische en leukemische cellen, zou kunnen verklaren.

In Chapter 8 worden de algemene conclusies van dit onderzoek besproken en belicht ik hoe we de verkregen inzichten in praktijk kunnen brengen. Want, juist bij AML, waarbij patiënten nog altijd met aspecifieke en intensieve chemotherapie worden behandeld, en de ziekte in veel gevallen terugkeert na een in eerste instantie ogenschijnlijk succesvolle behandeling, kunnen innovatie therapiën als splicing modulatie mogelijk in de toekomst een verschil maken. Hiernaast zijn onze studies ook op mechanistisch gebied van groot belang. Vanuit een mechanistisch oogpunt kunnen onze resultaten helpen met het begrijpen van het ontstaan van AML bij volwassen en bij kinderen. Daarbij inspireren en stimuleren onze resultaten verder onderzoek om de huidige behandelingsopties te verbeteren.

REFERENCES

1. Döhner H, Weisdorf DJ, Bloomfield CD. Acute Myeloid Leukemia. N Engl J Med. 2015;373(12):1136-1152.
2. Brunner AM, Graubert TA. Genomics in childhood acute myeloid leukemia comes of age. Nature Medicine. 2018;24(1):7-9.
3. Walter RB, Appelbaum FR, Tallman MS, Weiss NS, Larson RA, Estey EH. Shortcomings in the clinical evaluation of new drugs: acute myeloid leukemia as paradigm. Blood. 2010;116(14):2420-2428.
4. Reedijk AMJ, Klein K, Coebergh JWW, et al. Improved survival for children and young adolescents with acute myeloid leukemia: a Dutch study on incidence, survival and mortality. Leukemia. 2019;33(6):1349-1359.
5. Bachas C, Schuurhuis GJ, Assaraf YG, et al. The role of minor subpopulations within the leukemic blast compartment of AML patients at initial diagnosis in the development of relapse. Leukemia. 2012;26(6):1313-1320.
6. Chow LT, Gelinas RE, Broker TR, Roberts RJ. An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. Cell. 1977;12(1):1-8.
7. Berget SM, Moore C, Sharp PA. Spliced segments at the 5′ terminus of adenovirus 2 late mRNA. Proceedings of the National Academy of Sciences. 1977;74(8):3171.
8. Chen M, Manley JL. Mechanisms of alternative splicing regulation: insights from molecular and genomics approaches. Nat Rev Mol Cell Biol. 2009;10(11):741-754.
9. Nilsen TW, Graveley BR. Expansion of the eukaryotic proteome by alternative splicing. Nature. 2010;463(7280):457-463.
10. Bonnal SC, López-Oreja I, Valcárcel J. Roles and mechanisms of alternative splicing in cancer — implications for care. Nature Reviews Clinical Oncology. 2020;17(8):457-474.
11. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. Nat Rev Cancer. 2016;16(7):413-430.
12. Seiler M, Peng S, Agrawal AA, et al. Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. Cell Rep. 2018;23(1):282-296.e284.
13. Malcovati L, Papaemmanuil E, Bowen DT, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood. 2011;118(24):6239-6246.
14. Damm F, Thol F, Kosmider O, et al. SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. Leukemia. 2012;26(5):1137-1140.
15. Hou HA, Liu CY, Kuo YY, et al. Splicing factor mutations predict poor prognosis in patients with de novo acute myeloid leukemia. Oncotarget. 2016;7(8):9084-9101.
16. Wojtuszkiewicz A, Assaraf YG, Maas MJ, Kaspers GJ, Jansen G, Cloos J. Pre-mRNA splicing in cancer: the relevance in oncogenesis, treatment and drug resistance. Expert Opin Drug Metab Toxicol. 2015;11(5):673-689.
17. Sciarrillo R, Wojtuszkiewicz A, Assaraf YG, et al. The role of alternative splicing in cancer: From oncogenesis to drug resistance. Drug Resist Updat. 2020;53:100728.
18. Lee SC-W, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nature Medicine. 2016;22(9):976-986.
19. Seiler M, Yoshimi A, Darman R, et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. Nature Medicine. 2018;24(4):497-504.
20. Lee SC, North K, Kim E, et al. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. Cancer Cell. 2018;34(2):225-241.e228.
21. Crews LA, Balaian L, Delos Santos NP, et al. RNA Splicing Modulation Selectively Impairs Leukemia Stem Cell Maintenance in Secondary Human AML. Cell Stem Cell. 2016;19(5):599-612.
22. Hsu TY, Simon LM, Neill NJ, et al. The spliceosome is a therapeutic vulnerability in MYC-driven cancer. Nature. 2015;525(7569):384-388.
23. Buonamici S, Yoshimi A, Thomas M, et al. H3B-8800, an Orally Bioavailable Modulator of the SF3b Complex, Shows Efficacy in Spliceosome-Mutant Myeloid Malignancies. Blood. 2016;128(22):966-966.

24. Aird D, Teng T, Huang C-L, et al. Sensitivity to splicing modulation of BCL2 family genes defines cancer therapeutic strategies for splicing modulators. Nature Communications. 2019;10(1).

25. Steensma DP, Wermke M, Klimek VM, et al. Results of a Clinical Trial of H3B-8800, a Splicing Modulator, in Patients with Myelodysplastic Syndromes (MDS), Acute Myeloid Leukemia (AML) or Chronic Myelomonocytic Leukemia (CMML). Blood. 2019;134(Supplement_1):673-673.

CURRICULUM VITAE

Inge van der Werf was born on October 13th 1992 in Voorschoten, the Netherlands. In 2014 she received her Bachelor's degree in Biomedical Sciences at the VU University in Amsterdam. The research for her Bachelor's thesis focused on alternative splicing and drug resistance in acute lymphoblastic leukemia and was awarded the Grant Award at the Student Research Conference 2015. In 2016 she graduated her Master's program Cum Laude with specialization in oncology research at the VU University in Amsterdam. The research for her Master's thesis was conducted at Harvard Medical School and focused on stem cell targeted therapies in Infantile Hemangioma under supervision of prof. Joyce Bischoff for which she received support of the Bekker-La Bastide Foundation. During her Master training, she was awarded the Diamond Grant with a PhD research project based on her own ideas.

In September 2016 she started her PhD research project that resulted in the current thesis under supervision of prof. dr. Gertjan Kaspers and prof. dr. Jacqueline Cloos. During her PhD trajectory, she visited the lab of prof. dr. Catriona Jamieson at the University of California (San Diego) for seven months, for which she received support from the Cancer Center Amsterdam and the Nijbakker Morra Foundation. In 2019, her research performed in San Diego was awarded the Tom Voûte Award for Young Investigators. In the final stage of her PhD research, she visited the lab of dr. Ruben van Boxtel at the Princess Máxima Center for six months. Currently, Inge is working as a post-doc on stem cell failure in children at the Princess Máxima Center for pediatric oncology under the supervision of dr. Mirjam Belderbos.

PUBLICATIONS

INGE VAN DER WERF*, Phoebe Mondala*, Larisa Balaian, Luisa Ladel, Kathleen Fish, Cayla Mason, Raymond Diep, Jacqueline Cloos, Gertjan Kaspers, Jim La Clair, Peggy Wentworth, Leslie Crews, Kathleen Fisch, Thomas Whisenant, Michael Burkart, Catriona Jamieson. Selective Targeting of Alternative Splicing Deregulation in Pediatric Acute Myeloid Stem and Progenitor Cells. Manuscript in preparation

INGE VAN DER WERF*, Phoebe Mondala*, Larisa Balaian, Luisa Ladel, Kathleen Fish, Cayla Mason, Raymond Diep, Jacqueline Cloos, Gertjan Kaspers, Jim La Clair, Peggy Wentworth, Leslie Crews, Kathleen Fisch, Thomas Whisenant, Michael Burkart, Catriona Jamieson. Global Splicing Disruption in Pediatric AML by RBFOX2 Downregulation. Manuscript in preparation

Anna Wojtuszkiewicz, INGE VAN DER WERF, Stephan Hutter, Wencke Walters, Constance Baer, Wolfgang Kern, Jeroen J. W. M. Janssen, Gert J. Ossenkoppele, Claudia Haferlach, Jacqueline Cloos* and Torsten Haferlach*. Maturation State-Specific Alternative Splicing in FLT3-ITD and NPM1 Mutated AML. Cancers 2021, 13(16), 3929

INGE VAN DER WERF*, Anna Wojtuszkiewicz*, Manja Meggendorfer, Stephan Hutter, Constance Baer, Martijn Heijmans, Peter Valk, Wolfgang Kern, Claudia Haferlach, Jeroen Janssen, Gert J. Ossenkoppele, Torsten Haferlach[#] and Jacqueline Cloos[#]. Splicing Factor Gene Mutations in Acute Myeloid Leukemia Offer Additive Value If Incorporated in Current Risk Classification. Blood Advances. 2021. 14;5(17):3254-3265

INGE VAN DER WERF. The Process of Splicing is, as always, Surprising. 'Behind the Paper' Nature Portfolio Cancer Community. 2021.

INGE VAN DER WERF, Anna Wojtuszkiewicz, Huilan Yao, Rocco Sciarillo, Manja Meggendorfer, Stephan Hutter, Wencke Walters, Jeroen Janssen, Peter J. M. Valk, Wolfgang Kern, Claudia Haferlach, Torsten Haferlach, Gerrit Jansen, Gertjan L. Kaspers, Richard Groen, Jacqueline Cloos. SF3B1 as Therapeutic Target in FLT3/ITD Positive Acute Myeloid Leukemia. Leukemia. 2021. 35(9):2735

INGE VAN DER WERF*, Bernadette A. Chua*, Catriona Jamieson, Robert A.J. Signer. Post-Transcriptional Regulation of Normal, Stressed and Diseased Stem Cells. Cell Stem Cell. 2020. 6;26(2):138-159

INGE VAN DER WERF and Catriona Jamieson. The Yin and Yang of RNA Methylation: An Imbalance of Erasers Enhances Sensitivity to FTO Demethylase Small-Molecule Targeting in Leukemia Stem Cells. Cancer Cell. 2019. Volume: 35, Issue: 4, Page: 540-541

Elisabeth Huijbers, INGE VAN DER WERF, Lisette Faber, Lena Sialino, Pia van der Laan, Hanna Holland, Anca Cimpean, Victor Thijssen, Judy van Beijnum and Arjan Griffioen. Targeting Tumor Vascular CD99 Inhibits Tumor Growth. Frontiers in Immunology. 2019. Volume: 10. Article: 651

INGE VAN DER WERF, J. Cloos. (2018). Verassende 'Splicing' in Hematologische Maligniteiten. Nederlands Tijdschrift voor Hematologie. Volume: 15, Page: 388-95

AWARDS

van der Werf I, awarded the ASH Abstract Achievement Award for the abstract titled Selective Targeting of Alternative Splicing Deregulation in Pediatric Acute Myeloid Stem and Progenitor submitted to the 62nd ASH Annual Meeting and Exposition, December 5-8, 2020, Virtual

van der Werf I, awarded the Tom Voûte Young Investigator Award in a competition with ten preselected candidates by presenting my research 'Splicing Deregulation in Pediatric Acute Myeloid Leukemia' to a room of researchers and a three-person jury. November 2019

van der Werf I, awarded the CCA Travel Grant to visit the lab of prof. dr. Catriona Jamieson at the Sanford Consortium at the University of San Diego, San Diego, 2017

van der Werf I, awarded the Nijbakker Morra Foundation Stipend to visit the lab of prof. dr. Catriona Jamieson at the Sanford Consortium, University of San Diego, San Diego, 2017

van der Werf I, awarded the Travel Award from the Bermuda Principles Foundation Fund for attendance of the Bermuda Principles Impact on Splicing Meeting 2018, January 21-25, 2018, Southampton, Bermuda

van der Werf I, awarded the OOA Diamond Program Grant from the graduate school for oncology Amsterdam and Dutch Research Council for the design of a PhD research project based on her own ideas. April 2016

van der Werf I, awarded the 'Grant Award' for Best Bachelor Research of all Dutch and Flemish Universities at the Student Research Conference 2015 in the category Medical and Health. November 2015.

ACKNOWLEDGEMENTS

Once upon a time, I started my PhD.

More than four years later it is time to write the acknowledgements. My journey was unique, eventful and turbulent, but in the end, it made me love science even more. I want to thank all the people involved in my journey. The Department of Hematology at the Cancer Center Amsterdam, the members of the lab of prof.dr. Catriona Jamieson at the University of California San Diego and the van Boxtel Group at the Princess Máxima Center in Utrecht. In addition, I want to use this section to highlight and thank most important people.

To start, prof.dr. Gertjan Kaspers, March 12, 2013, the second year of my Bachelor Biomedical Sciences. Do you remember? I visited you in your office asking you for an internship and advice how to pursuit a career in Pediatric Oncology. Today, I can say I succeeded. I want to thank you for your continuous support, encouragement, and clinical perspective.

Prof.dr. Jacqueline Cloos, not long after I met Gertjan, I was in your office. We always stayed in touch. You were my supervisor during my bachelor internship, you delivered my graduation speech at the end of my master studies and together we wrote the research proposal for the OOA Diamond Grant. During my PhD, you gave me the opportunity to grow. The freedom you provided made it possible to travel to San Diego and spend my final months in the Princess Maxima Center. I appreciate your generosity a lot, thank you for the opportunities, guidance and encouragement.

Dr. Anna Wojtuszkiewicz, dear Anna, we know each other already quite some years, we have worked on many different projects, and I believe we are a very successful team. I believe, you deserve a lot of recognition for your contribution to this thesis. Dr. Richard Groen, I want to thank you for your critical and professional perspective on my research projects.

Prof. Dr. Catriona Jamieson, dear Cat, you are an amazing mentor. You made my dream come true by supporting my plans in your lab. You have gone beyond the limits to support me after I went back to the Netherlands. You are knowledgeable, inspiring and a great motivator. I can't thank you enough and I look forward to continue our collaboration in the future.

Phoebe, you are not only a colleague, but you are also a friend. You made my time in San Diego unforgettable. Afterwards, you always had my back. I want to thank you for all your hard work to make our project successful. I think your dedication to science, your eagerness to learn and your friendly personality will bring you a beautiful future. Also, a special thanks to Tom, I have told you already plenty of times how much I loved our meetings. Late at night, early in the morning, both tired but always ready to discuss science. I hope we will be able to spend more time working together, I wish you all the best.

Ruben, I really appreciate your encouraging enthusiasm. You have an amazing group of people working on the most relevant question in Pediatric Oncology, why do children get cancer? It

has been an honor to spend the final months of my project in your lab. I am curious what the future will bring.

I would also like to thank the members of my doctoral committee, prof.dr. Sonja Zweegman, prof.dr. Michel Zwaan, prof.dr. Arjan Griffioen, prof.dr. Olaf Heidenreich and dr. Marcel Spaargaren. I look forward to defending this thesis in front of all of you.

Last but not least, I want to add a personal note.

My parents, I know how proud the two of you are. Same holds true for my brother, Martijn. Reineke and Daan, in hard times you have always been there, and we always have a lot of fun. Neil and Yoriko, my amazing American family members. Welcoming, supportive and one of the main reasons I dared to move to San Diego. Kat and Steve, thanks to you I have a second home on the other side of the world. I wish you all the best. Lieve, we have a unique friendship. We know that we have each other, and I definitely know that I will treasure that forever. Rob, Ashleigh, Chris and Elzemieke, you are my dearest friends. Thank you for standing by my side when times were hard. Actually, my words can't express how thankful I am to have you in my life. Though, I will add some special words for Chris. Whenever I come to you for advice or help, you are always there. I am lucky I had you around during this journey.

I saved my final words for Willem, I know there is no one as proud as you are, with this little booklet in its hands. You were already there at the start of my scientific career, and I know for sure that my career path would have looked different if you wouldn't have been there. You kept me grounded, pushed me in the right direction, you know what's best for me and even created my own room to work on my research in our home. You resemble the foundation of my accomplishments and together with Gandalf we form a great team. I love you.

I finalized my PhD, and for sure, will live happily ever after.