UNIVERSITÀ
DI PAVIA

PhD IN BIOMEDICAL SCIENCES
DEPARTMENT OF BRAIN AND BEHAVIORAL SCIENCES
DEPARTMENT OF MOLECULAR MEDICINE

SF3B1-mutant myelodysplastic syndromes: clonal trajectories and therapeutic targeting

PhD Tutor: Prof. Luca Malcovati

PhD dissertation of
Martina Sarchi

a.a. 2021-2022

Table of Contents

Table of figures

Summary

Myeloid neoplasms are driven by acquisition of clonal genetic lesions, yet it remains unclear how these lesions cooperate to transform hematopoietic stem and progenitor cells (HSPCs) programs and modify drug responses. The integration of genomic profiling with clinical and cytogenetic data in the recent IPSS-M has significantly improved risk stratification of patients with myelodysplastic syndromes (MDS). MDS with mutations in the splicing factor *SF3B1* (SF3B1m) have been proposed as a distinct disorder characterized by ring sideroblasts, ineffective erythropoiesis and good prognosis. However, selected co-occurring genetic abnormalities can modulate patients outcome, transforming this relatively indolent condition to high-risk malignancy. Despite the prevalence of *SF3B1* lesions, mechanisms of oncogenesis and progression remain poorly understood, and targeted therapies are not currently available.

Mutational analysis of *SF3B1* co-mutated genes in patients with SF3B1m MDS or secondary AML (sAML) showed that *RUNX1* and *STAG2* mutations were strikingly associated with aggressive disease and reduced overall survival. To understand how these events cooperate with *SF3B1* mutations, we modeled these clonal trajectories in primary human HSPCs. *SF3B1* K700E mutation was introduced in $CD34^+$ cells isolated from cord blood or peripheral blood of healthy donors, using CRISPR/Cas9 editing and AAV6-mediated homology directed repair. To model disease progression, we introduced secondary *RUNX1* (S-R), *STAG2* (S-S), or control *AAVS1* (S-A) mutations into *SF3B1* K700E knock-in $CD34^+$ HSPCs, leading to a reduction of RUNX1 or STAG2 expression in *SF3B1*m cells (50% and 90%, respectively).

The S-R and S-S trajectories induced divergent alterations in lineage specification of SF3B1m HSPCs. S-R promoted myeloid skewing at the expense of the erythroid lineage, while S-S induced a block in differentiation, impairing both myeloid and erythroid differentiation. Consistently, introduction of secondary *RUNX1* or *STAG2* mutations in an induced pluripotent stem cell model of *SF3B1*m MDS recapitulated S-R myeloid expansion and S-S maturation arrest identified in primary HSPCs. S-R and S-S genotypes resulted in divergent transcriptional programs including inflammation, immune response, and myeloid differentiation. However, both high-risk clonal trajectories expanded the immature $CD34^+CD38^-$ HSCs/multipotent progenitors. These data suggest that, despite divergent

molecular pathways, high-risk mutations converge on expansion of stem cell potential of SF3B1m HSPCs.

To understand how genetic heterogeneity can influence drug responses, we profiled a panel of spliceosome inhibitors and 166 FDA approved and investigational compounds in gene edited K562 cell line. Both S-R and S-S genotypes conferred decreased response to single and combinatorial agents conventionally used for the treatment of myeloid neoplasms. High-risk genotypes maintained elevated sensitivity to SF3b inhibition, but conferred differential response to novel classes of spliceosome modulators. By contrast, CHK1 inhibitor Prexasertib was highly selective for SF3B1m cells irrespective of co-mutations. CHK1 is a key regulator of DNA damage and cell cycle. Splicing analysis and functional testing in SF3B1m genotypes revealed that coordinated mis-splicing of BUB1B and CDC27 by mutant-SF3B1 delayed mitotic progression leading to CHK1 activation and sensitizing SF3B1m cells to CHK1 inhibition. To determine how HSPCs state further modulates drug response, SF3b and CHK1 inhibitor were tested in primary $CD34^+$ HSPCs. Although both inhibitors preferentially targeted SF3B1m $CD34^+$ HSPCs, SF3b inhibitor showed elevated toxicity on wild-type $CD34^+CD133^+$ HSCs. By contrast, CHK1 inhibitor spared wild-type HSCs while effectively killing SF3B1m HSCs.

In conclusion, progression from low-risk SF3B1m MDS to high-risk disease is mediated by molecularly distinct trajectories driven by *RUNX*1 and *STAG2* mutations that converge on expansion of SF3B1m HSC compartment. Moreover, clonal progression is associated with genotype-specific drug responses that are further modulated by HSPCs epigenetic states. Our study uncovers CHK1 inhibition as a potential therapeutic strategy to target HSCs across the different SF3B1m genotypes.

Introduction

Myelodysplastic syndromes: definition, classification and prognosis.

Myelodysplastic syndromes (MDS) are myeloid neoplasms (MNs) characterized by clonal expansion of hematopoietic stem cells (HSCs), ineffective hematopoiesis and increased risk of leukemic evolution. They represent heterogeneous MNs that range from indolent conditions with a long natural history to subtypes analogous to acute myeloid leukemia (AML) [1,2].

MDS are most common in the elderly population with a median age at diagnosis of 70 years[1,3]. The incidence rate of MDS is about 5 cases per 100.000 persons per year in the general population, but increases to 20-50 cases per 100.000 persons per year after age 60. Men have higher incidence rate than women, and white individuals higher than other racial/ethnic groups.[2,3] Both congenital and acquired risk factors, such as previous treatment with chemotherapy or radiotherapy, can predispose to MDS onset but the etiology for most patients is often unknown.[1,4,5]

Cytopenia, i.e. the reduction in one or more peripheral blood (PB) lineages, and dysplasia, i.e. morphological abnormalities of the cells, are the key features of MDS and their characterization is required for MDS diagnosis and classification.[4] Cytopenia in at least one lineage is an essential diagnostic criterion for MDS, and it is defined by the following thresholds: hemoglobin (Hb) < 10 g/dL, platelets < 100 x 10^9/L and absolute neutrophil count < 1.8 x 10^9/L. The threshold to define dysplasia is 10% dysplastic cells in any hematopoietic lineage. [4]

The first classification of MDS was introduced in 1982 by the French-American-British (FAB) group and it became the reference standard for subsequent MDS classification schemes. It considered 5 different categories of MDS: refractory anemia (RA), refractory anemia with ring sideroblasts (RARS), refractory anemia with excess of blasts (RAEB), refractory anemia with excess of blasts in transformation (RAEB-t), chronic myelomonocytic leukemia (CMML). The distinguishing features between these groups were largely based on the proportion of blast cells in the PB and/or the bone marrow (BM) together with specific morphologic abnormalities, such as ring sideroblasts, erythroid precursors with abnormal accumulation of mitochondrial iron.[6,7,8]

In 2001, the World Health Organization (WHO) proposed a new classification for MDS, that maintained the structure of FAB classification while integrating it with more recently acquired information about BM cytogenetics, immunologic markers and molecular genetics. For the first time, the presence of a specific cytogenetic abnormality, the deletion of the long arm of chromosome 5 (5q-), was incorporated as a diagnostic parameter that exclusively defined an MDS category.[9]

WHO classification was subsequently revised in 2008 and 2016 to incorporate clinical and scientific findings that better defined previously described categories and introduced newly recognized entities.[10,11,12] Specifically, in the 2016 revision, the focus shifted from lineages manifesting cytopenia to lineages manifesting dysplasia. Therefore, terms such as "refractory anemia" or "refractory cytopenia" were replaced by "MDS" followed by the appropriate modifiers, i.e. single lineage dysplasia (MDS-SLD) or multilineage dysplasia (MDS-MLD). BM blasts remained critical in defining WHO MDS categories: 2-4% PB blasts or 5-9% BM blasts defined "MDS with excess of blasts-1" (MDS-EB-1), while 5-19% PB blasts or 10-19% BM blasts defined "MDS with excess of blasts-2" (MDS-EB-2). Blast proportion over 20% was instead diagnostic for AML. MDS with ring sideroblasts (MDS-RS) was defined by a ring sideroblasts proportion of at least 15% of BM erythroid elements, but the strong association between mutations in *SF3B1* gene and ring sideroblasts formation lowered this threshold to 5% in presence of *SF3B1* mutation.[4] Even though not included in 2016 WHO classification, clonal cytopenia of unknown significance (CCUS) emerged as a new category on the borderland of MDS, defined by cytopenia and clonal abnormalities, in absence of dysplasia or other clonal BM neoplasms.[13] The 2016 WHO classification has been adopted in this study.

In 2022, two classification proposals of myeloid malignancies appeared: the International Consensus Classification (ICC)[14] and 5th edition of the WHO classification[15]. Despite differences in terminology and addition/upgrade of new entities, both classifications recognized premalignant clonal cytopenias as diagnostic entities and included categories defined on the basis of gene mutations, e.g. *SF3B1* and *TP53*. Future efforts will be required to harmonize these classifications in order to facilitate patients diagnosis and treatment.[16]

MDS display variable clinical course, and morbidity and mortality are related primarily to complications arising from cytopenias and transformation to AML, whose risk can be assessed

through several prognostic systems.[17] The International Prognostic Scoring System (IPSS) and its revised version (IPSS-R) are the most common risk stratification tools for MDS.[18, 19] They are based on cytogenetic abnormalities, BM blast percentage and peripheral blood parameters, such as hemoglobin, platelets and neutrophils count. According to IPSS-R score, patients are stratified into 5 categories: Very Low, Low, Intermediate, High and Very High risk. Conventionally, patients with IPSS-R Very Low, Low and Intermiediate <=3.5 points are considered lower-risk, while IPSS-R Intermediate > 3.5 points, High, Very High indicates higher-risk cases.[19] In addition, in 2022 a molecular IPSS score (IPPS-M) has been developed. This score integrates cytogenetic abnormalities and hematologic parameters with the mutational status in 31 MDS driver genes, improving risk stratification of patients with MDS and supporting clinical decision-making.[20]

Therapeutic strategies for myelodysplastic syndromes.

Therapeutic objectives and strategies are strongly influenced by the variability of MDS presentation. Thus, risk stratification is the first step for establishing the treatment goals and plan.[1]

Response criteria to MDS treatment are based on therapy goals. If treatments are aimed at modifying the disease course, the response is distinguished in complete remission, partial remission, stable disease and progression. If patients are treated exclusively with supportive care to improve cytopenias, hematologic improvement and quality of life become the most relevant response criteria.[1,21]

The only potentially curative treatment option for MDS is allogeneic hematopoietic stem cell transplantation (alloHSCT). Eligibility should be always assessed in patients with higher-risk MDS, and it can be considered in a portion of patients with lower-risk MDS.[22,23] However, the majority of patients with MDS are not eligible for alloHSCT, because of age and/or comorbidities.[24]

In lower-risk patients, the goal of therapy has traditionally been ameliorating PB cytopenias and quality of life, even though an increasing interest for early intervention has been emerging to delay disease progression and improve survival.[24] Standard of care includes erythropoiesis-stimulating agents to increase red-cell production as well as red-cell transfusion and iron chelation to prevent

iron overload. In transfusion-dependent patients with hypoplastic MDS, immunosuppressive therapy can lead to transfusion independence.[22]

In higher-risk patients, median life expectancy is less than 2 years. As so, the goal of treatment is not only improving cytopenias but also preventing progression to AML and prolonging overall survival. For higher-risk patients that are ineligible for transplant, hypomethylating agents, such as azacytidine and decitabine, currently represent the most common initial treatment.[22,21] However, there are patients that fail to respond to frontline treatment with azacytidine, and also in azacytidine responders hematopoiesis remains clonal, without eradication of the mutant clone.[25] With the advent of hypomethylating agents, intensive chemotherapy is used less frequently in patients with MDS, while it still represents the treatment of choice for AML.[22,21] Suggested regimens include combinations of cytarabine with idarubicin, or fludarabine.[26] CPX-351, a liposomal encapsulation of cytarabine and daunorubicin, is currently in clinical trials for patients with higher-risk MDS, after showing prolonged survival in patients with sAML, as compared to conventional chemotherapy.[22,27]

The characterization of genetic drivers of MDS and the identification of recurrently dysregulated pathways has provided the rationale for investigating novel therapeutic approaches.[22,21] Novel strategy for *SF3B1*-mutant MDS will be discussed in the dedicated section of this chapter.

Hematopoietic hierarchy and genetic heterogeneity in myelodysplastic syndromes.

Genetic diversity and functional heterogeneity within each genetic subclone are key determinants of tumor evolution.[28] Consistently, MDS pathogenesis derives from the interplay between hematopoietic stem and progenitor cells (HSPCs) state and genetic networks supporting clonal expansion.[29]

Hematopoiesis is organized in a hierarchical structure where HSPCs are characterized by distinct self-renewal and differentiation capacity. At the top of this hierarchy, long-term HSCs (LT-HSCs) are capable of durable engraftment in primary and secondary transplantation assays, while intermediate and short-term HSCs as well as multipotent progenitors (MPP) engraft transiently in primary transplants. These populations can differentiate into myeloid, lymphoid and

megakaryocyte-erythroid progenitors, producing all differentiated blood types.[29,30] While initial models in early 2000s represented these stages as distinct and subsequent populations, more recent studies proposed acquisition of lineage specific fates as a continuous process, and supported the heterogeneous nature of HSPC populations.[31] This process is tightly regulated at steady-state and can adapt to drastic changes of demand. However, the balance between self-renewal and differentiation is disrupted in myeloid malignancies.[29]

MDS arise from a small pool of disease-initiating HSCs that are able to sustain generation of myeloid progenitors in vitro and in vivo, resist to conventional therapy and drive disease progression. [17,32,33] This process is driven by acquisition of genetic abnormalities that support clonal expansion and give rise to highly diverse subclonal architecture in MDS stem cells.[34] The onset of MDS is often preceded by a premalignant state known as clonal hematopoiesis (CH), characterized by expansion of mutant stem cells in absence of an overt neoplasm.[34,22] Stage specific alterations in HSPCs subpopulations have been identified in MDS evolution: while lower-risk MDS is characterized by expansion of common myeloid progenitors (CMP), higher-risk MDS show expansion of stem cell compartment and granulocyte monocyte progenitors (GMP), and suppression of megakaryocyte erythrocyte progenitors (MEP).[35] In addition, different genetic abnormalities can sustain distinct HSPCs architectures in MDS and modulate response to therapy.[33]

The identification of somatic mutations of *TET2* in myeloid neoplasms[36] has been the breakthrough in the elucidation of the genetic basis of MDS, and it gave rise to an important number of studies aimed at defining genes and mutations involved in MDS pathogenesis. Advancement of sequencing technologies enabled identification of at least one mutation in a recurrently mutated gene in 75-90% of patients with MDS.[17,37] These mutations affect several cellular pathways: RNA splicing, DNA methylation, chromatin modification, transcription regulation, DNA repair and replication, signal transduction, and cohesin complex.[38,39] While somatic mutations in spliceosome components and epigenetic regulators are generally early genetic events, the remaining mutated genes mainly contribute to clonal progression.[40,41,42] Interestingly, selected mutated genes can affect prognosis of MDS patients, and this can be further modulated by co-occurring genetic lesions.[20]

***SF3B1*-mutant myelodysplastic syndromes.**

Genes encoding splicing factors are recurrently mutated in MDS, with up to 60% of MDS cases carrying spliceosome mutations.[17] The spliceosome is a ribonucleoprotein (RNP) complex that removes noncoding introns from precursor messenger RNAs (pre-mRNAs) and ligates exons, to form mRNA. Alternative splicing, the process by which different splice sites are recognized and different exons included in mRNA, is a major source of transcript diversity in mammalian cells.[43,44,45] Nonetheless, in presence of mutations affecting spliceosome genes, this process is disrupted, and aberrant splicing events are induced, resulting in dysregulation of cellular processes.[45]

Among spliceosome components, *SF3B1* is the most frequently mutated in MDS, with ~30% of patients harboring *SF3B1* mutations.[46] SF3B1 is a core component of the U2 spliceosome that recognizes 3' splice sites, and *SF3B1* mutations alter splicing by promoting recognition of alternative 3' splice sites (a3ss).[47,48] *SF3B1* mutations in MDS are nearly universally associated with formation of ring sideroblasts (RS), erythroid precursors with iron-laden mitochondria, and >80% of patients diagnosed with MDS-RS harbor *SF3B1* mutation.[46] Key genes metabolic mitochondrial iron pathways (e.g. PPOX, TMEM14C, ABCB7) are mis-spliced in *SF3B1*-mutant patients[48,49,50,51], and the causative role of TMEM14C and ABCB7 reduction in RS formation was probed in a recent study.[52]

In addition to RS formation, *SF3B1*-mutant MDS is commonly associated with ineffective erythropoiesis and indolent clinical course. As so, it was proposed as a distinct nosologic entity, and both ICC and 2022 WHO classification introduced MDS with *SF3B1* mutation as a separate diagnostic category.[14,15,46] Despite the majority of patients with *SF3B1* mutation have favorable prognosis, selected co-occuring genetic lesions affect patients outcome.[20,46] The IPSS-M further distinguished *SF3B1*-mutant patients into 3 groups: SF3B1^{5q} for concomitant presence of del(5q), SF3B1^{b} in presence of co-mutation in *BCOR*, *BCORL1*, *NRAS*, *RUNX1*, *SRSF2* or *STAG2* genes, and SF3B1^{a} in the remaining cases. The favorable outcome was confined to SF3B1^{a}, but not in SF3B1^{5q} and SF3B1^{b}.[20]

Low-risk *SF3B1*-mutant MDS can be managed with supportive treatments, but high-risk cases require disease-modifying therapy.[46] Luspatercept, a transforming growth factor beta (TGF-b) ligand trap, has been approved by FDA for transfusion-dependent patients with MDS-RS.[53] In

addition, the identification of NF-kb hyperactivation as a consequence of IRAK4 alternative splicing pointed at IRAK4 inhibition as potential strategy in *SF3B1*-mutant MDS, and IRAK4 inhibitor CA-4948 is currently in clinical trial for Very Low, Low and Intermediate risk MDS. [24,54] Preclinical studies also showed increased sensitivity to spliceosome modulation in splicing factor mutant MDS/AML, including *SF3B1*-mutant cases. [55] However, first line inhibitors of core spliceosome component SF3b had elevated toxicity (E7107[56]) or limited response (H3B-8800[57]) in clinical trials. Despite 88% patients in the study carried splicing factor mutations, no complete or partial responses were observed after treatment with H3B-8800.[57] More recent strategies (e.g. RBM39 degraders, type I PRMT inhibitors, PRMT5 inhibitors) perturb splicing by acting on spliceosome accessory proteins and post-translational modification, and they have recently entered clinical trials.[55,58,59,60] Overall, despite the progress in understanding the pathophysiology and managing lower risk patients, therapeutic advances for higher-risk patients have been limited so far, and no targeted treatments are currently available.[22]

Aims of the study.

Ineffective erythropoiesis, RS formation and indolent clinical course are the main features of *SF3B1*-mutant MDS. However, some patients with *SF3B1* mutation progress to AML and have poor prognosis.[20,46] Despite the prevalence of *SF3B1* lesions, mechanisms of oncogenesis and progression remain poorly understood, and no targeted therapies have been approved for high-risk *SF3B1*-mutant patients.

Our study has the following goals:

1) Identification of genetic drivers of disease progression and modeling of clonal evolution.

2) Characterization of the impact of high-risk mutations on stem cell compartment and lineage specification of *SF3B1*-mutant HSPCs.

3) Identification of therapeutic vulnerabilities of high-risk genotypes with *SF3B1* mutation.

Methods

Patients, samples collection and Next Generation Sequencing

This study included 176 patients investigated for suspected myeloid neoplasms (MNs) at the Department of Hematology, IRCCS Policlinico San Matteo Foundation and University of Pavia, between 2001 and 2020.

In all patients, complete clinical evaluation was performed, and PB samples and BM aspirate, BM biopsy or both were obtained. The procedures followed were in accordance with the Declaration of Helsinki of 1975, as revised in 2000, and samples were obtained after patients provided written informed consent. Diagnostic procedures aimed at the identification of the MNs were in accord to recent recommendations[2,4,61], and diagnosis of MN was based on the criteria proposed in the WHO classification of MNs and its revisions. [4,9,10]

Peripheral blood granulocytes were isolated by standard density gradient centrifugation, followed by red blood cell lyses with hypotonic solution and immunomagnetic selection on MiniMACS separation columns using anti-CD15 antibody (Miltenyi Biotec, Bergisch Gladbach, Germany), as previously reported.[62] Genomic DNA was extracted by following standard protocols for human tissue.

Targeted capture DNA sequencing of recurrently mutated genes in MNs was performed locally using four different validated panels.[63,64,65] Functionally annotated variants were filtered based on the information retrieved from public databases (dbSNP, 1000genome, and ESP6500), the expected germ line allele frequency. The remaining variants were considered as candidate somatic mutations, and were finally tagged as oncogenic, based on the information derived from the literature, the Catalog of Somatic Mutations in Cancer (COSMIC; http://cancer.sanger.ac.uk/cancergenome/projects/cosmic), and on in silico prediction effect. Variant allele frequency was calculated as the number of variant reads divided by the total reads. The analysis was limited to variants with allele frequency of equal or greater than 2%, as previously described.[66] Analysis was restricted to a core of 24 genes evaluated in all panels. These include: *ASXL1*, *BCOR*, *CBL*, *CEBPA*, *DNMT3A*, *ETV6*, *EZH2*, *FLT3*, *IDH1*, *IDH2*, *JAK2*, *KDM6A*, *KIT*, *KRAS*, *NRAS*, *RUNX1*, *SETBP1*, *SF3B1*, *SRSF2*, *STAG2*, *TET2*, *TP53*, *U2AF1*, *ZRSR2*.

AAV generation: vector constructs and AAV production

Gene editing plasmids were generated using pAAV-MCS2 (was a gift from Steve Jackson (Addgene plasmid #46954).

pAAV.SF3B1.BFP plasmid was generated with SF3B1 homology arms containing a total of 1.6kb of homology comprising coordinates chr2:198,266,105-198,267,704 (hg37). Each homology arm was 800bp in length and were synthesized as gBlocks (IDT). The right homology arm also contains the K700E mutation at genomic coordinate 198,266,834 (A>G, hg37) and a synonymous PAM mutation at genomic coordinate 198,266,826 (G>C, hg37). The PGK.BFP.WPRE.SV40 cassette was amplified out of a pLKO.1 derived BFP vector using primers FP: 5'-GGGTTGCGCCTTTTCCAAG and RP: 5' - TAAGATACATTGATGAGTTTGGACAA. pAAV-MCS2 was digested with NotI-HF and MluI-HF (NEB) and Gibson Assembly (NEB) was used to insert the two 0.8kb homology arms and the PGK.BFP.WPRE.SV40 cassette.

pAAV.AAVS1.BFP plasmid was generated with AAVS1 homology arms containing a total of 1.6kb of homology comprising coordinates chr19:55,626,320-55,627,919 (hg37). Homology arms synthesis, PGK.BFP.WPRE.SV40 cassette generation and cloning were done as per pAAV.SF3B1.BFP plasmid.

AAV stocks were produced by double transfection of pAAV.BFP vector and pDGM6 (a gift from David Russell, addgene plasmid #110660) in HEK 293T cells. Transfected cells were collected 48 hours later, lysed by freeze-thaw, benzonase-treated, and purified over iodixanol density gradient as previously described.[67] Titers of the viral stocks were determined by qPCR of AAV genomes[68] and ranged from 1×10^{11} to 1×10^{12} per microliter.

Isolation, culture and gene editing of human cord blood HSPCs

Human umbilical cord blood mononuclear cells were collected by centrifugation with Ficoll-Paque Plus (GE), followed by red cell lysis with ammonium chloride (StemCell Technologies) and viably cryopreserved in 10% DMSO, 40% FBS. $CD34^+$ cells were extracted using the Miltenyi CD34 Microbead kit according to manufacturer's instructions.

Cryopreserved $CD34^+$ HSPCs were thawed following the Lonza Poietics protocol (www.lonza.com) and cultured for 3-4 days in HSPC expansion media consisting of StemSpan SFEM II (StemCell

Technologies) plus SCF (100 ng/ml), FLT3 (100ng/ml), TPO (100ng/ml), IL-6 (100ng/ml), UM171 35nM and SR1 0.75uM. Confluency was maintained between 4 x10^5 and 1 x10^6 per ml. After 72-96 hours, HSPCs underwent CRISPR editing. RNP complexes were generated by combining 1.875ul 20uM Cas9 protein (Synthego) and 0.94ul 100uM sgRNA (Synthego) in 4.685 P3 Primary Cell Nucleofector Solution with Supplement 1. For combined gene editing of multiple loci, 0.94ul of each 100uM sgRNA were added to 1.875ul 20uM Cas9 protein and 3.715ul (for 2 sgRNAs) or 2.805ul (for 3 sgRNAs) P3 buffer. The sgRNA sequences are as follows:

SF3B1	5' – UGGAUGAGCAGCAGAAAGUU – 3'
AAVS1	5' – GGGCCACUAGGGACAGGAU – 3'
RUNX1	5' – GAGCCCAGGCAAGAUGAGCG – 3'
STAG2	5' – AUACCUUGUGGAUAGCAUGU – 3'
TET2	5' – UUGUAGCCAGAGGUUCUGUC – 3'

Complex was incubated for 15 minutes at room temperature. Between 4 x10^5 and 1 x10^6 HSPCs per condition were resuspended in 17.5 P3 Buffer and added to the RNP complex. $CD34^+$ HSPCs were electroporated using Lonza 4D (DZ-100 program) and returned to HSPCs media containing AAV particles (up to 20% culture volume). HPSCs were washed 16-24 hours later and replated in HSPCs media for 48 hours. Next, editing efficiency was assessed by flow cytometry and PCR amplification of edited loci, followed by Sanger sequencing and ICE Synthego deconvolution[69].

Lentivirus generation: lentiviral constructs and preparation

For lentivirus mediated CRISPR/Cas9 editing, oligos were cloned in pLentiCRISPR-v2-mCherry vector (#99154) following the cloning vector protocol (Addgene).
sgRNA oligonucleotide sequences are as follows:

AAVS1	5' – GGGCCACTAGGGACAGGAT – 3'
RUNX1#A	5' – CACTTCGACCGACAAACCTG – 3'
RUNX1#C	5' – GAGCCCAGGCAAGATGAGCG – 3'
STAG2#B	5' – ATACCTTGTGGATAGCATGT – 3'
STAG2#C	5' – AATACTAACCTTGAACCGAC – 3'

For lentivirus mediated gene knockdown, shRNA oligonucleotides were cloned in the TRC cloning vector pLKO.1 (Addgene, plasmid #10878) following the pLKO.1 TRC cloning vector protocol (Addgene).
shRNA sense oligonucleotide sequences used are as follows:

Control Luciferase (M1)	5' – CTTACGCTGAGTACTTCGAC – 3'
BUB1B#2 (B2)	5' – GAGACAACTAAACTGCAAATT – 3'
BUB1B#3 (B3)	5' – GTGGAACACTGAAACTGTATG – 3'
CDC27#2 (C2)	5' – CAAGTACCTAATCATAGTTTA – 3'
CDC27#3 (C3)	5' – GCCTATAACAGTGACTTGATT – 3'

Virus was prepared using third generation packaging plasmids pMDLg/RRE and pRSV/Rev. pMD.G was used for VSV-G pseudotyping. Plasmids were transfected into HEK 293T cells with calcium phosphate (Takara). Lentivirus was concentrated via ultracentrifugation at 23,000 rpm for 2:15 hrs, resuspended in SFEM and stored at -80°C. All viruses were titered by serial dilution on 293T cells.

K562 culture, infection and gene editing

K562 cells with the SF3B1 K700E mutation or K700K control were a gift from Robert K. Bradley (Fred Hutchinson Cancer Center). K562 cells were grown in Iscove's modified Dulbecco's media (IMDM) with 10% FBS and penicillin/streptomycin and cultures were maintained at $<1.5 \times 10^6$ cells/ml. For lentivirus infection, K562 cells were transduced in a 48-well plate in standard media with polybrene. $5x10^5$ cells were transduced at an MOI of 1 and washed 16-24 hours later. Transduced cells were purified on a BD FACS Aria III sorter. SF3B1 WT control cells with AAVS1 editing and SF3B1m subclones (S-A, S-R, S-S) were established from single cell sorting of edited K562 cells.

iPSC-HPCs culture, infection and gene editing

iPSC-HPCs were generated as previously described.[52,70] iPSC-HPCs cells were cultured in SFEM (StemCell Technologies) with 50 ng/ml SCF, 50 ng/ml FLT3, 50 ng/ml TPO, 50 ng/ml IL-6, 10 ng/ml IL-3 (all Peprotech) and penicillin/streptomycin. Doxycycline was added at 2 μg/ml (Sigma). Cultures

were maintained at <1.5 x 10^6 cells/ml, and media was exchanged every 3-4 days. iPSC-HPC lines were maintained in culture between 60-100 days prior to CRISPR/Cas9 editing and differentiation. iPSC-HPCs cells were transduced in a 48-well plate in standard media with polybrene. 6 x 10^5 cells were infected at an MOI of 5 in each well and spun at 2300 rpm, 30 minutes prior to incubation at 37°C. Cells were washed 16-24 hours later and resuspended in fresh iPSC-HPCs media.

Erythroid/myeloid differentiation

Erythroid/myeloid differentiation culture was adapted from Dutt S. et al., 2011. iPSC-HPCs and primary $CD34^+$ HSPCs were cultured for 7 days in IMDM + 10% FBS + 1% BSA with 100 U/mL penicillin/streptomycin, 2mM glutamine, 100 ng/mL SCF, 10 ng/mL IL-3, 15 ng/mL G-CSF, 40 ng/mL FLT3 ligand and 0.5 U/mL erythropoietin. Cells were plated at a density of 2-3 x 10^5/ml and maintained <10^6/ml.

***In vitro* drug treatment**

For high-throughput drug screening, K562 cells were plated onto 6 x 384well plates plus one clear bottom plate, and placed at 37°C, 5% CO2. Serial dilutions of each compound were added 24 hours after plating, and cell were incubated for 72 hours. Viability was measured with Promega's CellTiter Glo 2.0 Lot# 000472252. Data from wells of each plate was normalized for 32 plate blank values and 16 DMSO solvent control wells with the formula: Well %Viability = ((Well Relative Light Unit – Plate Blank)/(DMSO Control – Plate Blank)) * 100. IC50s were determined for each genotype from fitted curves of viabilities at serial dilution of each compound.

For targeted drug testing, K562 cells were plated on 96-well tissue culture-treated plates at a density of 5,000 cells per 150ul per well respectively. Cells were treated with serial dilutions of SF3b inhibitor Pladienolide B (0-100nM, Tocris), type I PRMTs inhibitor MS-023 (0-100uM, Sigma), PRMT5 inhibitor GSK3326595 (0-100uM, MedChemExpress), CHK1/CHK2 inhibitor AZD7762 (0-1uM, SelleckChem), CHK1 inhibitor LY2606368 Prexasertib HCl (0-100nM, SelleckChem), CHK2 inhibitor BML-277 (0-1uM, SelleckChem). After 5-7 days, cell viability read-out was performed using the CellTiter-Glo 2.0 Cell Viability Assay (Promega, G9242) as per manufacturer instructions. The proportion of viable cells with drug treatment was calculated relative to DMSO control. A three-parameter nonlinear fit

of log(inhibitor) versus response was performed in GraphPad Prism v7.0 (GraphPad Software, San Diego, CA) to determine IC50 values.

For testing SF3b and CHK1 inhibitor in $CD34^+$ HSPCs, cells were cultured in HSPCs expansion media with 2.5nM Pladienolide B, 2.5nM Prexasertib or DMSO vehicle control for 7 days. Fresh media and drugs were added at day 4. Total viable cell counts and proportion of HSPCs populations were determined by flow cytometry, as described below.

Flow cytometry and cell sorting

Erythroid/myeloid differentiation was analyzed using antibodies CD71 APC-H7 (M-A712; BD), CD235a/Glycophorin A PE-Cy7 (11E4B-7-6; Coulter), CD11b APC (ICRF44; BD), CD14 BV711 (MϕP9; BD) and CD15 BV785 (HI98; BD). HSPCs populations were analyzed using stem cell markers CD34 AF700 (581; BD), CD38 PE-Cy7 (HB7; BD) and CD133 APC (AC133; Mylteni). For quantifying drug response in gene edited HSPCs, CountBright™ Absolute Counting Beads (Invitrogen™) were added to cells resuspension as per manufacturer instructions. For both staining panels, wells were incubated for 20 minutes at room temperature in PBS with 2% FBS with 1:100 antibody dilution (1:50 for CD11b and CD133).

For intracellular flow cytometry, cells were stained with CD34 PE-Cy7 surface marker as described above. Then, cells were fixed (BD Cytofix) for 20 min at 4C and washed twice with BD Perm/Wash buffer. For PU.1 staining, cells were incubated with 0.5ug/ul antibody (PE anti-PU.1; 7C2C34; Biolegend) in BD Perm/Wash buffer for 30 min at 4C. Cell were washed twice and resuspended in PBS + 2% FBS for flow cytometry analysis.

For cell cycle analysis, cultured cells were incubated in 20uM EdU for 2.5 hours. Cells were then harvested and fixed (BD Cytofix) 15 min RT, permeabilized with BD Perm/Wash buffer for 15 mins, followed by EdU detection as per Click-iT EdU cell proliferation kit protocol (ThermoFisher cat. # C10340), and DNA detection with DAPI acquired in linear mode.

For all flow cytometry applications, acquisition was performed on the BD LSRII and Symphony A3 cytometers. All data were analyzed using FlowJo software (10.7.0).

For molecular characterization, gene edited primary HSPCs and iPSC-HPCs were isolated through Fluorescence-Activated Cell Sorting (FACS) on BD Aria. Samples were resuspended in PBS with 50% FBS, and $CD34^{+}BFP^{+}$ or $mCherry^{+}DAPI^{-}$ populations were isolated, respectively.

gDNA extraction, PCR amplification and Sanger sequencing

gDNA was extracted from gene edited cells using QIAamp DNA Micro Kit (QIAGEN) as per manufacturer instructions. Edited loci were PCR amplified using Taq 5X Master Mix (NEB) or Primestar GXL Polymerase (Takara). PCR products were purified using NucleoSpin gel and PCR Clean-Up kit (Macherey-Nagel) and submitted to Genewiz for Sanger sequencing. Percentage of frameshift mutations was estimated from Sanger sequencing using ICE Synthego tool.[69] Primer details are as follow (* indicates sequencing primer):

AAVS1: *F 5' – ATCCTCTCTGGCTCCATCGT – 3'
R 5' – CCGGTTAATGTGGCTCTGGT – 3'
RUNX1: *F 5' – AAGGCCCCTGAACGTGTATG – 3'
R 5' – GGCCAGTACCTTGAAAGCGA – 3'
STAG2: *F 5' – GGCAGTTTCTTCTCTGTCCT – 3'
R 5' – AACAAAACTATGCACGAAGTAAGA – 3'
TET2: *F 5' – AACTAGAGGGCAGCCTTGTG – 3'
R 5' – TGGCTTACCCCGAAGTTACG – 3'
SF3B1: F 5' – GCTGCTGGTCTGGCTACTAT – 3'
*R 5' – TCTATCCAAAGCCATCCTGTGC – 3'

Quantitative RT-PCR

RNA was extracted from 50,000-300,000 cells using TRIzol (ThermoFisher) and resuspended in water for cDNA synthesis using iScript cDNA synthesis kit (BioRad) according to manufacturer instructions. Primer details are as follow:

GAPDH: F 5' – TCCTGCACCACCAACTGCTTA – 3'
R 5' – TCTTCTGGGTGGCAGTGATGG – 3'

RUNX1: F 5’ – AGATGAGCGAGGCGTTG – 3’

R 5’ – CGGAGCAGAGGAAGTTGG – 3’

STAG2: F 5’ – GGATAGCATGTGGGACTGTG – 3’

R 5’ – GCACTCTCTTGCCTATCTGTT – 3’

TET2: F 5’ – TCACCTCCCATTTGCCAGAC – 3’

R 5’ – AGGAGCCCAGAGAGAGAAGG – 3’

BUB1B: F 5’ – AAATGACCCTCTGGATGTTTGG – 3’

R 5’ – GCATAAACGCCCTAATTTAAGCC – 3’

CDC27: F 5’ – CCCGTCCAGGCTGCTATATG – 3’

R 5’ – AAAGGCGTTCTGCGAGGAAAA – 3’

Western Blot

Protein lysates were prepared by resuspending 100,000 washed cells/10µl of RIPA buffer (Thermo Fisher) with protease and phosphatase inhibitor (Sigma). Lysates were resolved by 4-20% SDS-PAGE (Biorad) and immunoblotted with antibodies for GAPDH (1: 10,000; Abcam, ab9485), HSP90 (1: 10,000; BD Biosciences), RUNX1 (1:1,000, Santa Cruz sc-365644), STAG2 (1:1,000, Santa Cruz sc-81852), BUB1B (1:1,000, BD Biosciences), pCHK1 S345 (1:1,000, Cell Signalling Technology #2348), CHK1 (1:1,000, Cell Signalling Technology #2360). Membranes were washed in TBST and then incubated with antirabbit or anti-mouse HRP-conjugated secondary antibodies for visualization on a BioRad ChemiDoc. Images were exported to ImageJ for analysis.

RNA sequencing, gene expression and splicing analysis

Gene edited $CD34^+$ HSPCs and iPSC-HPCs were FACS sorted as described above. For K562 cells, 2 independent single cell derived clone per genotype were analyzed. CD34+ HSPCs, iPSC-HPCs and K562 cells were resuspended in Trizol (ThermoFisher, cat. #15596026) and RNA was extracted as

per manufacturer protocol. Library preparation and sequencing was performed at Omega Bioservices using the standard pipeline with polyA selection. Samples were processed on the Illumina HiSeq platform in the 2x150bp configuration with 25-50 million reads per sample.

For gene expression analysis, reads were pseudoaligned to hg38 reference genome using kallisto.[71] Differential gene expression was determined using the DESeq2 package.[72] Gene set enrichment analysis (GSEA) was performed on the Gene Ontology datasets in the Broad MSigDB[73] and curated datasets from the literature.[74]

For $CD34^+$ HSPCs and K562 cells splicing analysis, isoform expression levels for SF3B1m genotypes and control WT lines were estimated as previously described.[75] In brief, a transcriptome annotation for the hg19/GRCh37 human genome assembly was created by merging gene annotations from Ensembl v71.1[76], the UCSC knownGene track[77], and MISO v2.0 isoform annotations[78]. RSEM v1.2.4[79] was used to map all RNA-seq reads to this transcriptome annotation. Remaining unaligned reads were then mapped a database of all possible junctions between annotated 5' and 3' splice sites of the transcriptome annotation, as well as to the genome sequence, with TopHat v2.0.8b[80], and the resulting aligned reads were merged with the RSEM output. The expression levels of isoforms annotated in MISO v2.0's annotation were estimated with MISO v2.0. Events that were differentially spliced in SF3B1m versus control samples were identified as previously described. [52,75] Briefly, we defined the metric deltaPSI (ΔPSI) as the isoform ratio (absolute percentage of mRNA) in SF3B1m samples - isoform ratio in control cells, and computed a significance statistic using Wagenmakers's Bayesian framework[81] for single-sample comparisons or the Mann-Whitney U test for group comparisons. An event was classified as differentially spliced if it exhibited an absolute change of $\Delta PSI \geq 10\%$ and Bayes factor ≥ 5 or $p \leq 0.05$.

Statistical analysis

Statistical analysis was performed with GraphPad Prism software. Data are shown as the mean with standard deviation unless noted. For all analyses, $p < 0.05$ was considered statistically significant. Investigators were not blinded to the different groups.

Results

SF3B1-mutant MDS is a genetically heterogeneous disorder.

SF3B1-mutant (SF3B1m) MDS is a genetically heterogeneous disorder, with a subset of patients in the IPSS-M classification having poor prognosis and high-risk of AML transformation[20]. To identify molecular drivers of SF3B1m MDS evolution, we compared co-mutated genes in low-risk (LR) patients with MDS with ring sideroblasts (MDS-RS) without evidence of progression at follow-up (n=66) to high-risk (HR) patients with MDS with excess blasts (MDS-EB) or secondary AML (n=31) (**Fig. 1A**). Consistent with previous studies,[20,46] *TET2* was the most frequently co-mutated gene (HR 35%, LR 29%), followed by *DNMT3A* (15%) and *ASXL1* (12%) in the LR group, and *ASXL1* (23%), *STAG2* (23%) and *RUNX1* (23%) in the HR group (**Fig. 1B**). However, *TET2* mutations did not affect prognosis of SF3B1m patients, in line with the favorable outcome of SF3B1^{a} group in the IPSS-M classification.[20] By contrast, co-mutations in *RUNX1* (OR=18.36 (2.18-862.91), p=0.0013) or *STAG2* (OR= Inf (3.57-Inf), p=0.0002) were significantly enriched in HR SF3B1m patients (**Fig. 1C**), and significantly associated with increased BM blast count (**Fig. 1E**) and reduced overall survival (p=0.0002 and p<0.0001, respectively) (**Fig. 1F**). Median variant allele frequency (VAF) was 42% for *SF3B1* and 15.7% and 11% for *RUNX1* and *STAG2*, respectively (**Fig. 1D**), suggesting that these are mainly secondary events. Our data parallel the recent IPSS-M study[20] and implicate clonal heterogeneity as a major driver of outcomes in SF3B1m MDS.

Generation of CD34^{+} HSPC model of SF3B1m MDS.

To model SF3B1m MDS in CD34^{+} umbilical cord blood (CB) or adult mobilized blood (PB) HSPCs, we combined CRISPR/Cas9 gene editing and AAV6-mediated template delivery for homologous recombination to introduce the *SF3B1* K700E mutation at the endogenous locus (**Fig. 2A**). This system also allows introduction of BFP intronic marker for isolation and tracking of edited cells (**Fig. 2B**). We confirmed *SF3B1* K700E mutation (chr2: 198,266,834 T>C; c.2098A>G) and PAM mutation in BFP+ cells (**Fig. 2C**), and heterozygous expression of the mutant K700E allele (**Fig. 2D**). *SF3B1* mutations are associated with highly stereotyped aberrant splicing marked by recognition of alternative 3' splice sites (a3'ss).[47,48] Splicing analysis of CD34^{+} HSPCs showed that *SF3B1* K700E knock-in HSPCs recapitulated a3'ss mis-splicing of canonical genes identified in SF3B1m MDS patients, including MAP3K7, TMEM14C, and DYNLL1 (**Fig. 2E**).[48,82] We found significant correlation

between quantitative levels of a3'ss in $CD34^+$ HSPCs with K700E mutation knock-in compared to previously established SF3B1m cellular models (**Fig.2F**). Next, to model genetic heterogeneity in patients with SF3B1m MDS, we combined *SF3B1* K700E knock-in with sgRNA-mediated targeting of *RUNX1* or *STAG2* (termed S-R and S-S, respectively), or *AAVS1* safe harbor locus as a control (termed S-A). The efficiency of frameshift mutations was 34% for *AAVS1*, 60% for *RUNX1*, and 69% for *STAG2* (**Fig. 2G**), leading to ~50% reduction in RUNX1 expression and ~90% reduction in STAG2 expression in SF3B1m HSPCs (**Fig. 2H-I**). This model provides a defined genetic background to study the impact of genetic heterogeneity on the biology and therapeutic responses of primary human SF3B1m HSPCs.

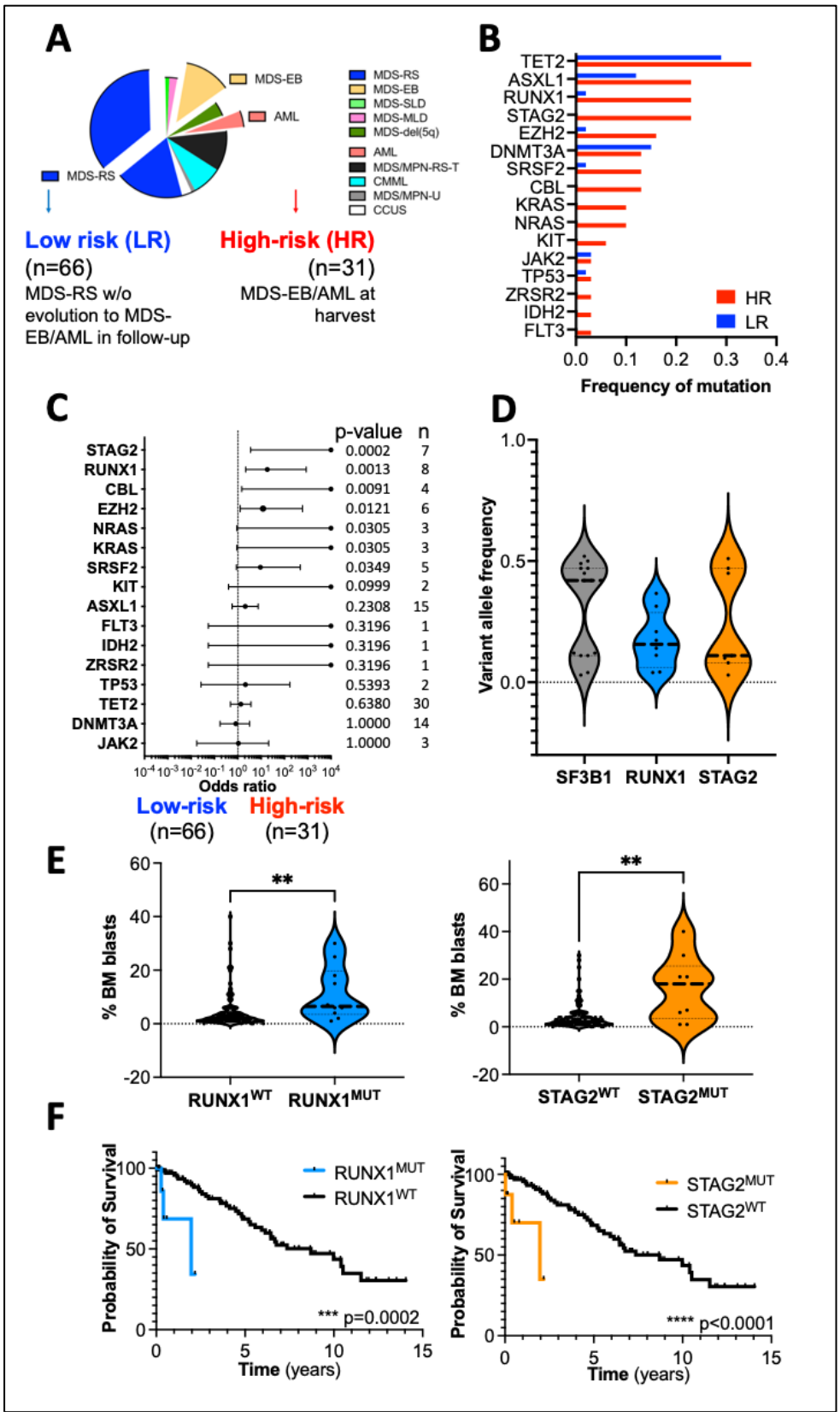

Figure 1. RUNX1 and STAG2 are high-risk genes in SF3B1m MDS.

(A) SF3B1m cohort characteristics. (B) Frequency of SF3B1 co-mutated genes in LR and HR groups. (C) Odds ratio of SF3B1 co-mutated genes in LR and HR groups. Fisher Exact Test. (D) Variant allele frequency of SF3B1, RUNX1 and STAG2 mutations. (E) Proportion of BM blasts in SF3B1m patients with or without high-risk co-

mutations. Unpaired t test. (F) Overall survival of SF3B1m patients with or without high-risk co-mutations. Mantel-Cox test.

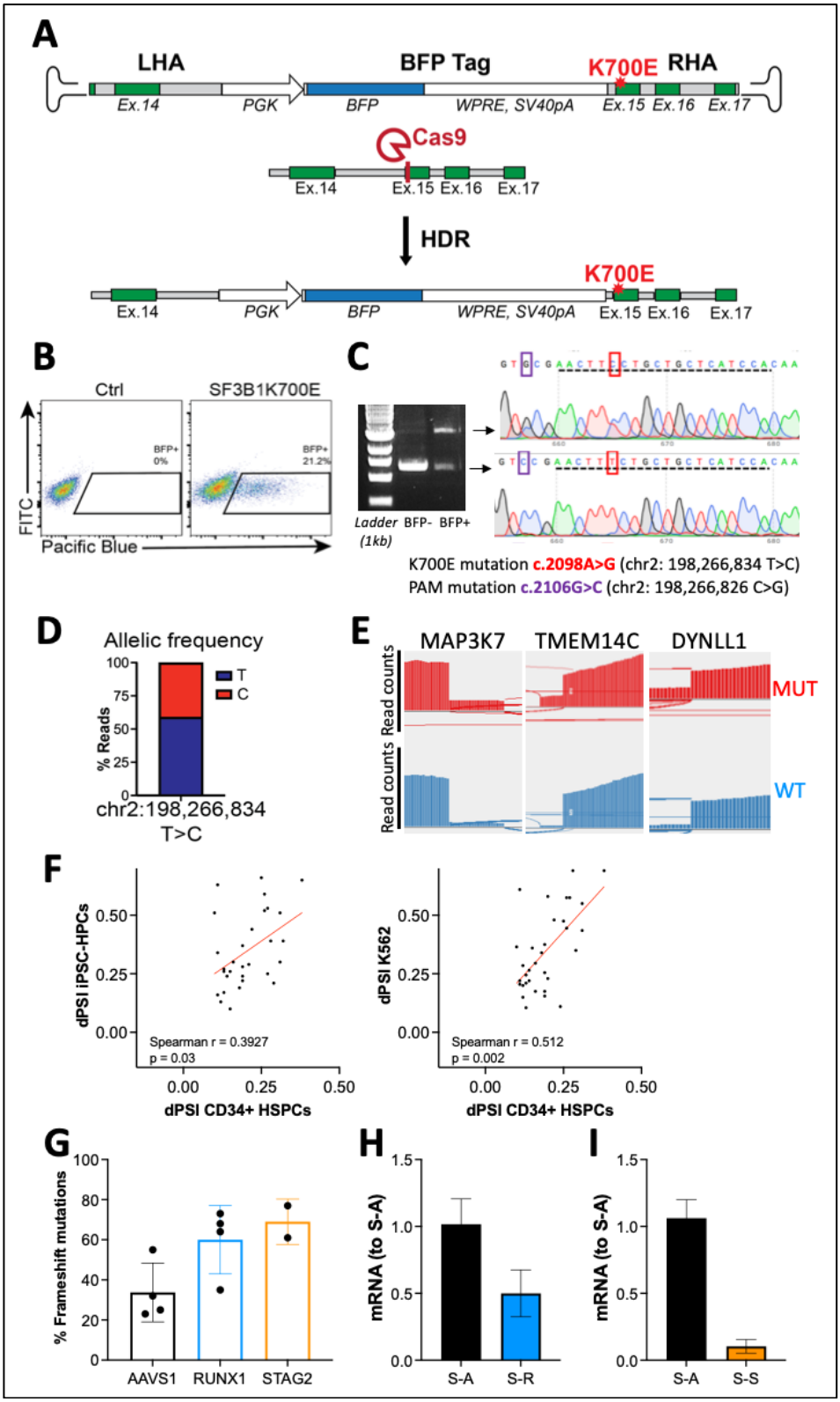

Figure 2. Generation of a primary CD34⁺ HSPCs model of SF3B1m MDS.

(A) Schematics of AAV6 template and CRISPR/Cas9 editing. (B) Flow cytometry plot of edited BFP$^+$ HSPCs. (C) PCR and Sanger sequencing of SF3B1 wild-type and mutant alleles. (D) Proportion of reads in support of SF3B1 wild-type and mutant alleles from RNA-seq. (E) Aberrant splicing (a3ss) of previously identified targets[48,82] of mutant-SF3B1. (F) Spearman correlation between the quantitative levels of aberrant splicing (a3ss) in CD34$^+$ HSPCs and iPSC-HPCs (left) and K562 cell lines (right) with SF3B1m. (G) Proportion of AAVS1, RUNX1 and STAG2 frameshift mutations in CD34$^+$ HSPCs as quantified by Sanger sequencing and ICE Synthego deconvolution.[69] (H), (I) Expression of RUNX1 (H) or STAG2 (I) as measured by quantitative PCR. Expression was normalized to S-A, mean ± s.d., n = 2 independent experiments.

High-risk mutations induce divergent lineage outcomes.

MDS pathogenesis is marked by aberrant HSPCs differentiation into mature hematopoietic lineages, especially the erythroid and neutrophil lineages, and accumulation of undifferentiated precursor cells.[35,83] In order to dissect the effect of high-risk mutations on SF3B1m HSPCs lineage specification, we differentiated edited HSPCs into erythroid and myeloid lineages and quantified the differentiation efficiency by flow cytometry using erythroid (CD71$^+$CD235a$^-$, CD71$^+$CD235a$^+$) and myeloid (CD11b$^+$CD14$^+$, CD15$^+$CD14$^-$) lineage markers (**Fig. 3A**). RUNX1 loss in SF3B1m cells (termed S-R) tended to reduce erythroid differentiation compared to SF3B1-only S-A control (S-A 30.6%, S-R 19.4%, p=0.1) while promoting a moderate increase in myeloid lineage output (S-A 31.4%, S-R 39%, p=0.1) (**Fig.3B**). By contrast, STAG2 loss in SF3B1m cells (termed S-S) significantly reduced both erythroid (S-A 30.6%, S-S 15.1%, p=0.02) and myeloid (S-A 31.4%, S-S 20.2%, p=0.003) differentiation and induced accumulation of precursor cells, suggesting a block in differentiation (S-A 38%, S-S 64.6%, p=0.007) (**Fig.3C**). Thus, *SF3B1* co-mutations further impair erythroid differentiation but have divergent effects on the myeloid lineage.

To validate these effects in an orthogonal model, we targeted *RUNX1*, *STAG2*, or *AAVS1* control in SF3B1m or isogenic SF3B1 wild-type (WT) iPSC-HPCs by lentiviral-mediated CRISPR/Cas9 editing (**Fig. 4A-I; Fig. 5A-B**). *RUNX1* mutation in SF3B1m iPSC-HPCs significantly reduced erythroid differentiation (S-A 52.7%, S-R 37.2%, p=0.0007), while expanding the myeloid compartment (S-A 16.5%, S-R 39.7%, p=0.0001) (**Fig. 5C**). By contrast, STAG2 loss caused a reduction in erythroid (S-A 52.7%, S-R 35.2%, p=0.016) and myeloid lineage output (S-A 16.5%, S-S 9.6%, p=0.014), leading to the accumulation of precursor cells (S-A 32.8%, S-S 54.1%, p=0.002) (**Fig. 5D**). Transcriptome profiling revealed divergent regulation of immune response and signal transduction genes in S-R and S-S genotypes (**Fig. 6A**). S-R displayed positive enrichment of granulocyte-monocyte progenitor (GMP)[74] genes (**Fig. 6B**) and PU.1 target genes that were instead negatively enriched in S-S (**Fig. 6C-**

E). Interestingly, when RUNX1 was disrupted in the isogenic SF3B1 WT iPSC-HPCs, no significant changes were observed (**Fig. 5E**). STAG2 loss in the WT iPSC-HPCs had significantly lower erythroid differentiation, but did not impair myeloid differentiation (**Fig. 5F**). The limited changes observed in the WT lines support a cooperative effect between R/S and *SF3B1* mutations. These data are concordant with the findings in the primary SF3B1m HSPCs, and indicate that high-risk mutations have different effects on the myeloid lineage specification of SF3B1m HSPCs.

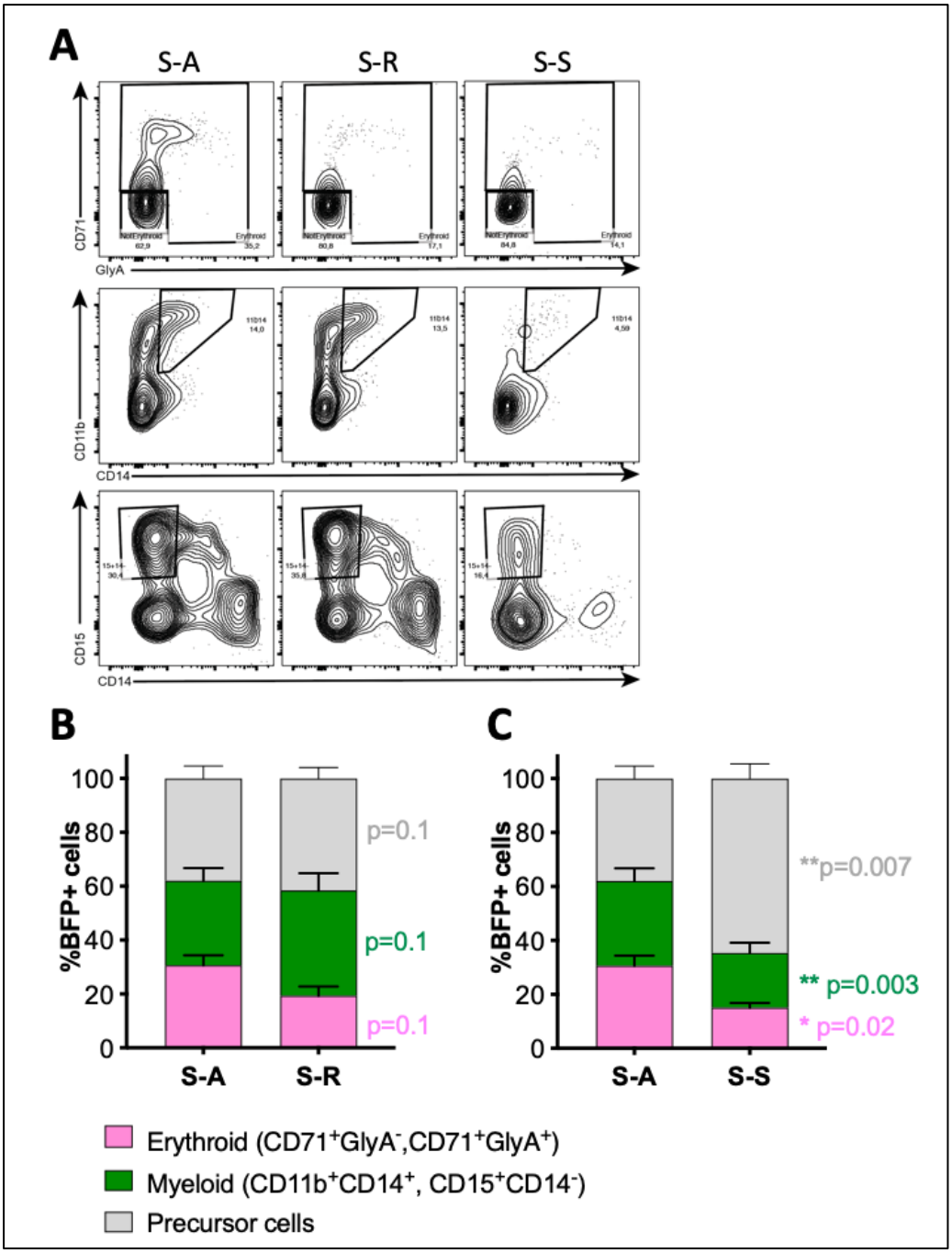

Figure 3. S-R and S-S induce divergent lineage outcomes in primary HSPCs.

(A) Representative flow cytometry plots of erythroid and myeloid lineage markers. (B), (C) Quantification of erythroid/myeloid output in SF3B1m HSPCs Proportion of erythroid, myeloid and precursor cells in S-A and S-R genotypes (B) or S-A and S-S genotypes (C).

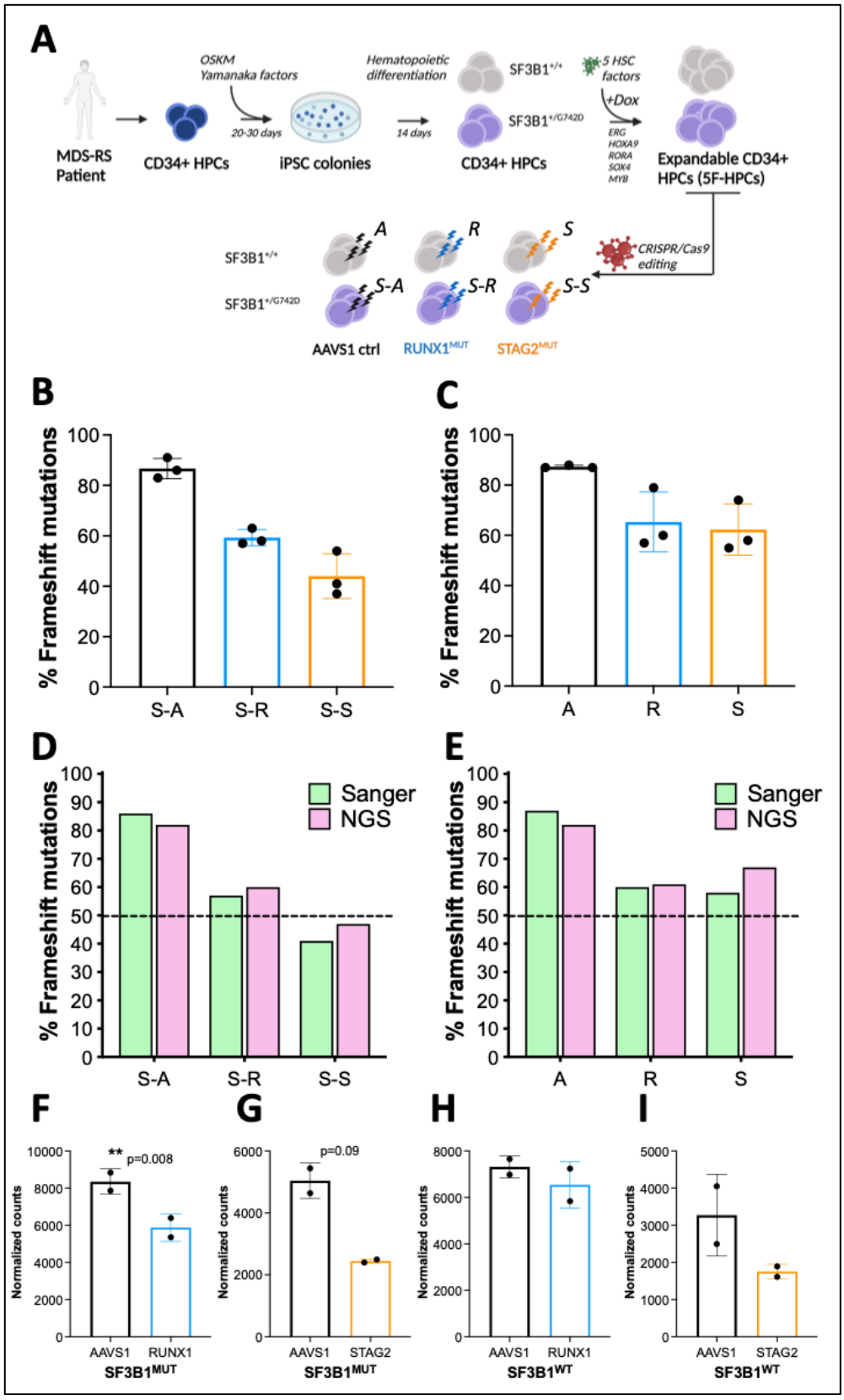

Figure 4. Generation of an iPSC-derived model of high-risk SF3B1m MDS.

(A) Schematics for iPSC-HPCs generation. (B), (C) Proportion of frameshift mutations in SF3B1m (B) and SF3B1 wt (C) iPSC-HPCs as quantified by Sanger sequencing and ICE Synthego deconvolution. (D), (E) NGS validation of editing efficiency in SF3B1m (D) and SF3B1 WT (E) iPSC-HPCs. (F) RUNX1 normalized counts in SF3B1m iPSC-HPCs. Paired t-test. (G) STAG2 normalized counts in SF3B1m iPSC-HPCs. Paired t-test. (H) RUNX1 normalized counts in SF3B1 WT iPSC-HPCs. (I) STAG2 normalized counts in SF3B1 WT iPSC-HPCs.

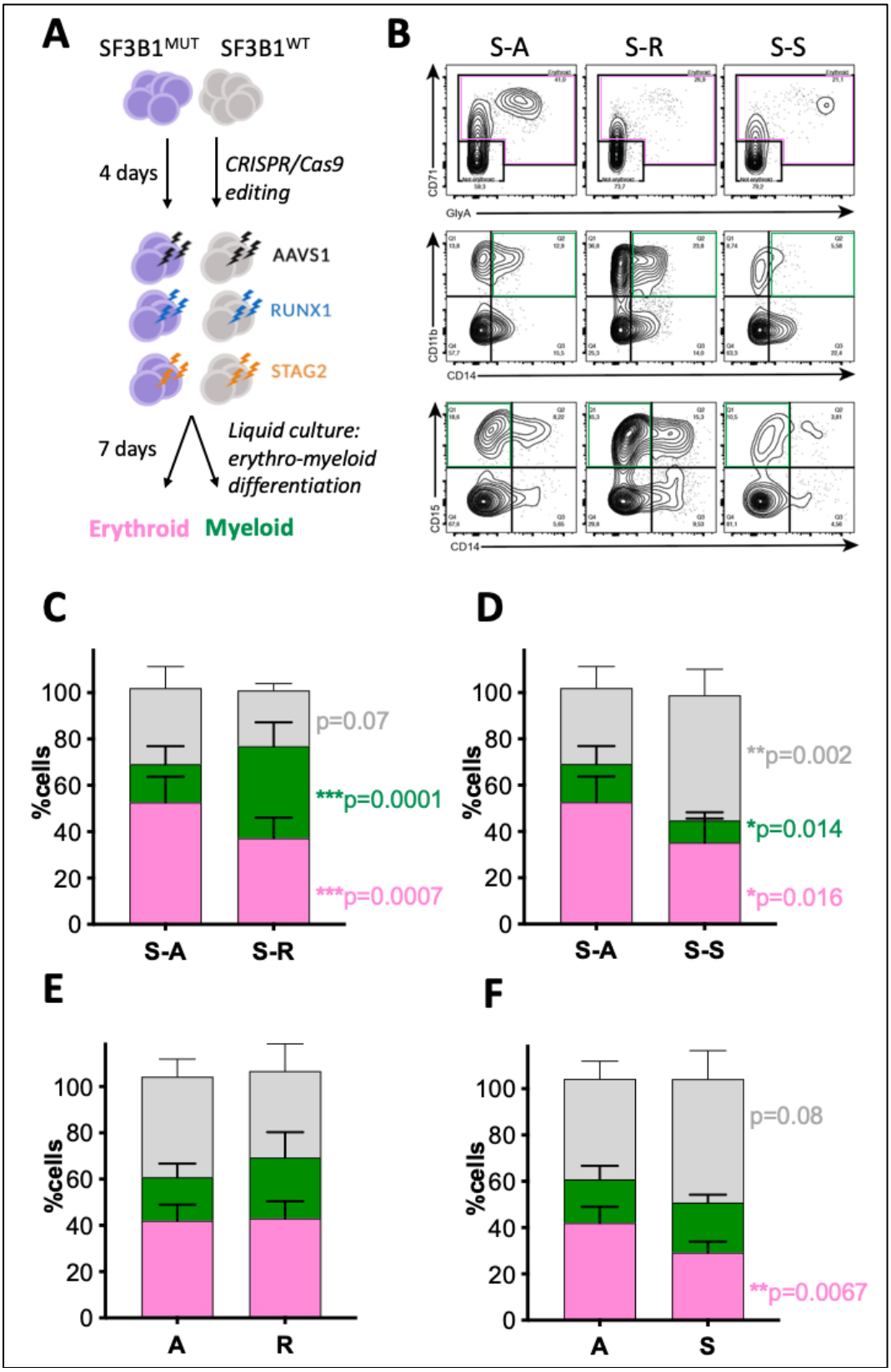

Figure 5. S-R and S-S induce divergent lineage outcomes in iPSC-HPCs.

(A) Schematics for iPSC-HPCs CRISPR/Cas9 editing and erythroid/myeloid differentiation. (B) Representative flow cytometry plots of erythroid and myeloid lineage markers. (C), (D) Quantification of erythroid/myeloid output in SF3B1m iPSC-HPCs: proportion of erythroid, myeloid and precursor cells in S-A and S-R genotypes (C) or S-A and S-S genotypes (D). (E), (E) Quantification of erythroid/myeloid output in SF3B1 WT iPSC-HPCs: proportion of erythroid, myeloid and precursor cells in A and R genotypes (E) or A and S genotypes (F).

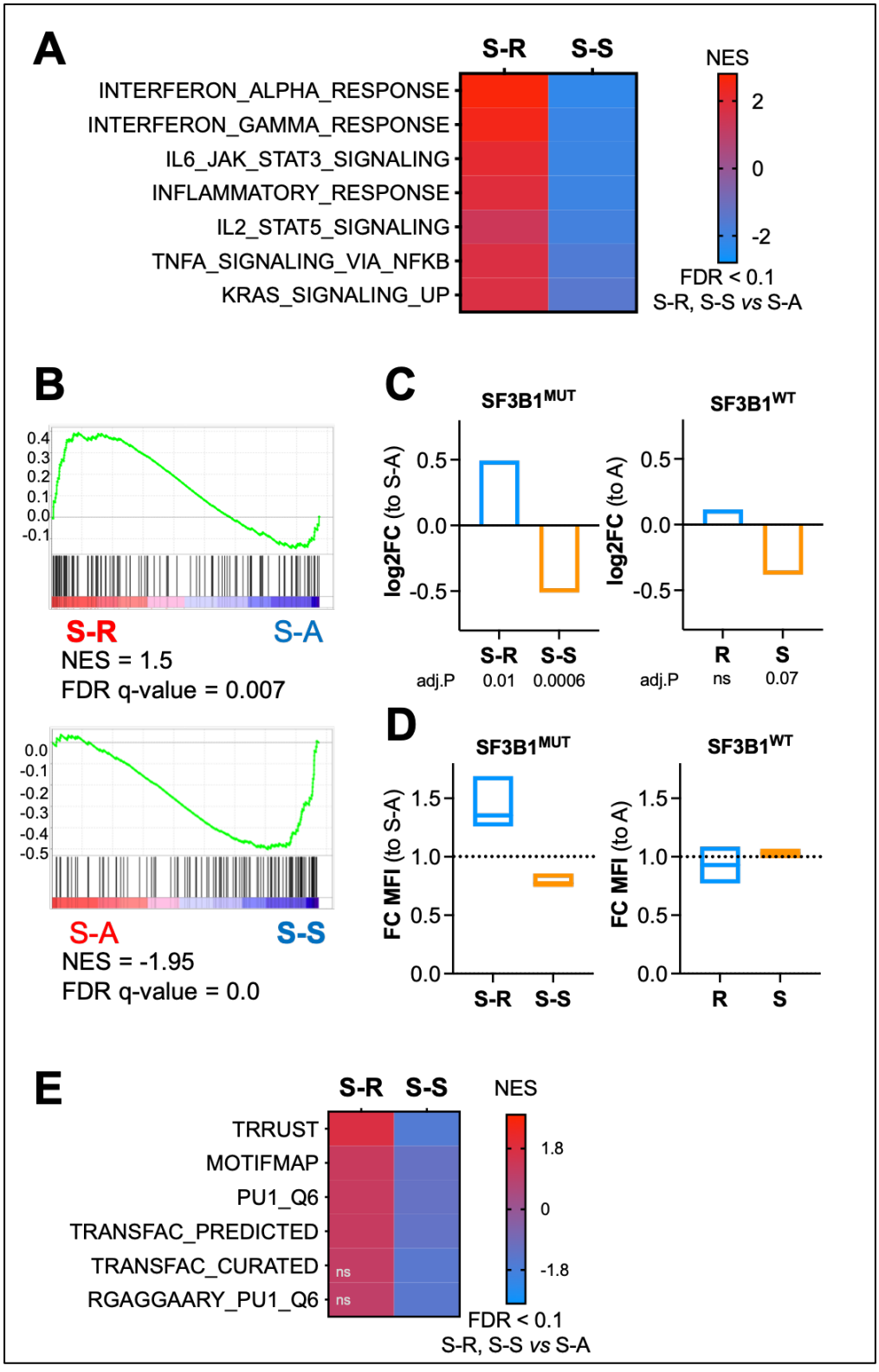

Figure 6. S-R and S-S induce divergent dysregulation of myeloid genes and inflammatory pathways.

(A) Enrichment score of hallmark gene sets in GSEA of S-R versus S-A or S-S versus S-A. FDR < 0.1. (B) GSEA of GMP gene signature from Laurenti et al.[74] of S-R versus S-A (top) or S-S versus S-A (bottom). (C) Expression of PU.1 transcript in SF3B1m (left) and SF3B1 WT (right) iPSC-HPCs, relative to S-A or A, respectively. (D) Expression of PU.1 protein as determined by intracellular flow cytometry in SF3B1m (left) and SF3B1 WT (right) iPSC-HPCs, relative to S-A or A, respectively. (E) Enrichment score of PU.1 target gene sets in GSEA of S-R versus S-A or S-S versus S-A. FDR < 0.1.

High-risk mutations converge on expansion of SF3B1m HSCs.

Expansion of the HSC compartment is a hallmark of MDS progression.[35] Previous studies identified distinct HSPCs architectures in MDS driven by different genetic abnormalities.[33] To study the effect of high-risk *RUNX1* or *STAG2* genotypes on SF3B1m CD34$^+$ HSPCs pool, we cultured edited cells for 14 days in media supporting HSPC maintenance and evaluated HSC/MPP population using cell surface markers CD34, CD38, and CD133 (**Fig. 7A**). Normal differentiation is accompanied by depletion of CD34$^+$CD133$^+$ and CD34+CD38$^-$ HSC/MPPs and differentiation into CD34$^+$CD133$^-$ precursors and finally CD34$^-$ mature cells.[30] As expected, we observed significant variability in the kinetics of HSPC differentiation in individual CB/PB donors. Despite the divergent dysregulation of myeloid lineage specification, both RUNX1 and STAG2 co-mutations converged on maintenance or expansion of immature HSPCs, with a higher proportion of CD34$^+$38$^-$ HSC/MPPs than SF3B1-only S-A control (**Fig. 7B-C**). Interestingly, RUNX1 but not STAG2 loss expanded the most immature CD34$^+$CD38$^-$CD133$^+$ population (**Fig. 7D-E**).

TET2 is the most frequent co-mutation with *SF3B1* but does not affect prognosis in our patient cohort or IPSS-M[20] (**Fig. 8A-B**). To test if expansion of SF3B1m HSC/MPPs in S-R and S-S was specific to high-risk genotypes or was an additive effect of co-occurring mutations, we combined *SF3B1* K700E mutation knock-in with targeting of the *TET2* locus in primary CD34$^+$ HSPCs. The mean percentage of frameshift mutations in *TET2* was 43% (**Fig. 8C**), with ~50% reduction in TET2 expression in SF3B1m HSPCs (**Fig. 8D**). We found no significant changes in CD34$^+$CD38$^-$ (**Fig. 8E**) or CD34$^+$ CD133$^+$ HSC/MPP populations when TET2 mutations were introduced in SF3B1m HSPCs (termed S-T) (**Fig. 8F**). Interestingly, the combination of TET2 and RUNX1 mutations in SF3B1m HSPCs (termed S-R-T) promoted further expansion of the CD34$^+$CD133$^+$ HSC/MPP pool compared to S-R with a control AAVS1 edit (termed S-R-A) (**Fig. 8G**). Taken together, our data show that high-risk but not low-risk genotypes promote maintenance or expansion of SF3B1m HSC/MPPs. Moreover,

distinct co-mutation patterns preferentially expand phenotypically distinct HSC/MPP subpopulations, indicating that SF3B1m MDS/AML is a heterogeneous disease at the stem cell level.

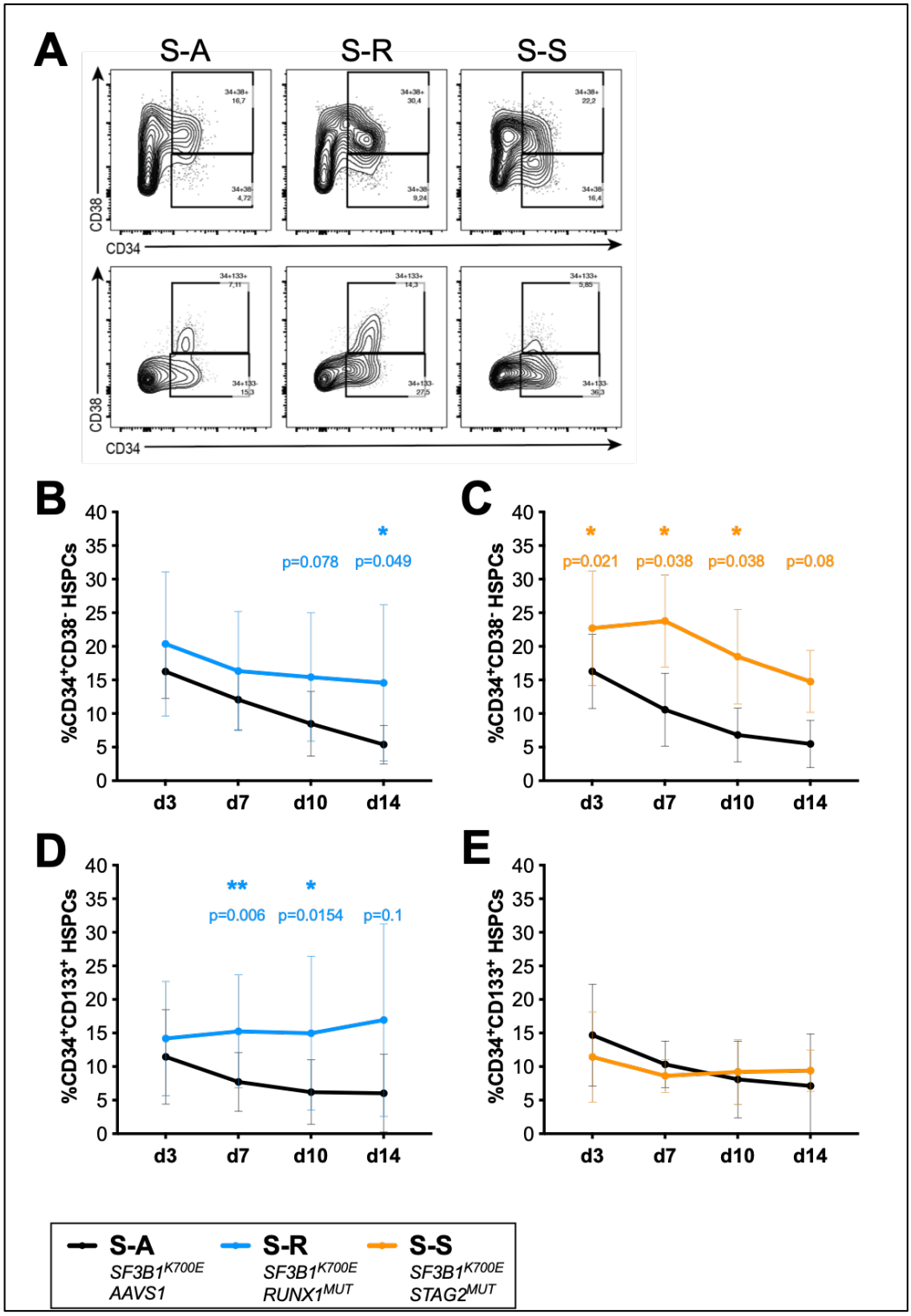

Figure 7. S-R and S-S promote expansion of SF3B1m HSC/MPPs.

(A) Representative flow cytometry plots for stem cells markers CD34/CD38 (top) and CD34/CD133 (bottom). (B) Proportion of CD34⁺CD38⁻ HSC/MPPs in S-A and S-R. Ratio paired t test. (C) Proportion of CD34⁺CD38⁻ HSC/MPPs in S-A and S-S. Ratio paired t test. (D) Proportion of CD34⁺CD133⁺ HSC in S-A and S-R. Ratio paired t test. (E) Proportion of CD34⁺CD133⁺ HSC in S-A and S-R. Ratio paired t test.

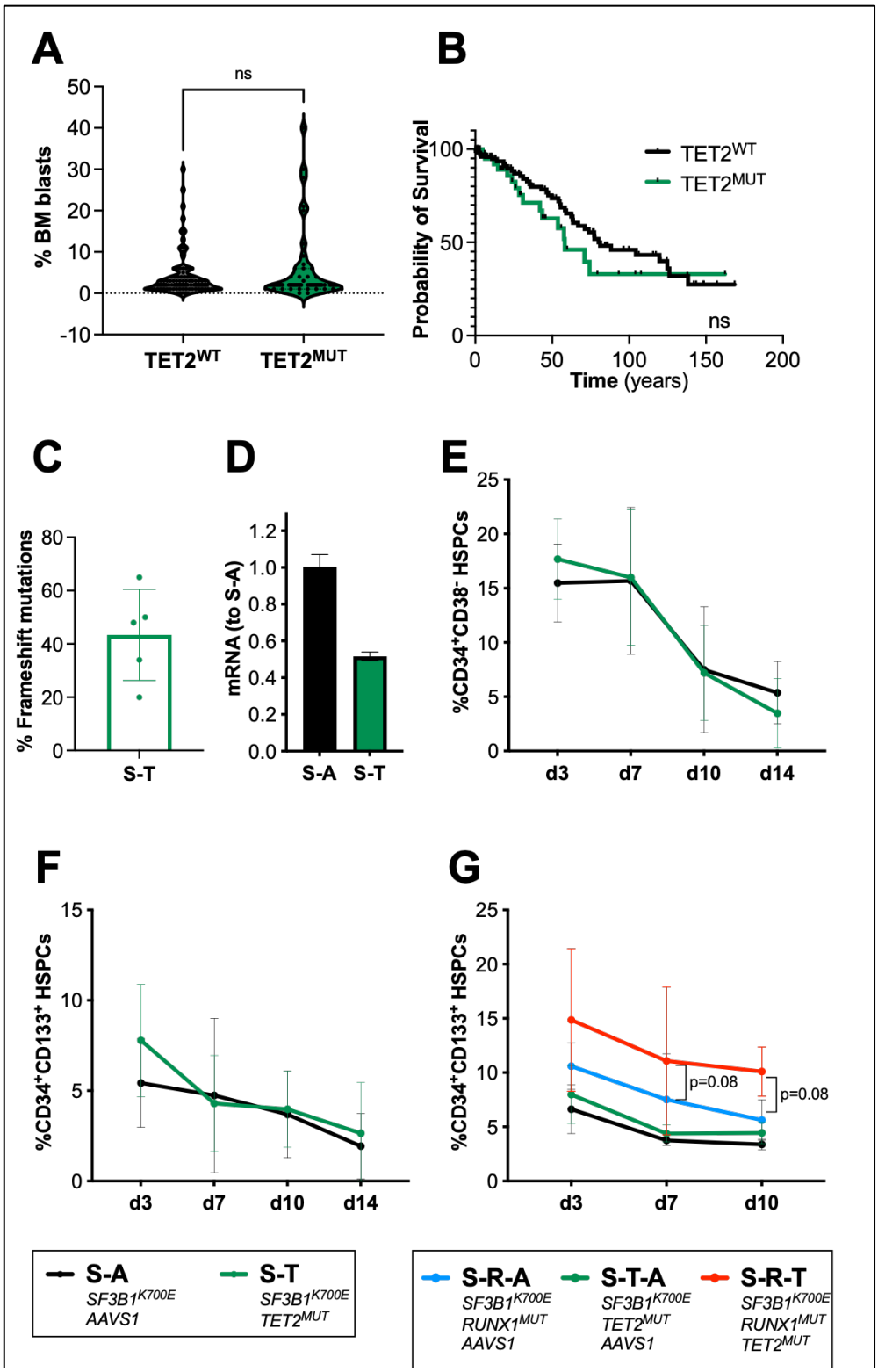

Figure 8. TET2 mutations do not affect outcome of SF3B1m patients and do not alter SF3B1m HSC/MPPs in absence of high-risk genes.

(A) Proportion of BM blasts in SF3B1m patients with or without TET2 co-mutations. Unpaired t test. (B) Overall survival of SF3B1m patients with or without TET2 co-mutation. Mantel-Cox test. (C) Proportion of TET2 frameshift mutations in CD34+ HSPCs as quantified by Sanger sequencing and ICE Synthego deconvolution. (D) Expression of TET2 as measured by quantitative PCR. Expression was normalized to S-A, mean ± s.d. (E) Proportion of $CD34^{+}CD38^{-}$ HSC/MPPs in S-A and S-T. Ratio paired t test. (F) Proportion of $CD34^{+}CD133^{+}$ HSC in

S-A and S-T. Ratio paired t test. (G) Proportion of $CD34^{+}CD133^{+}$ HSC in S-A, S-R-A, S-T-A and S-R-T. Ratio paired t test.

Genetic heterogeneity impacts response to spliceosome inhibition.

Genetic heterogeneity is a major determinant of therapeutic responses.[84,85] While patients with low-risk SF3B1m MDS usually have an indolent course that can be managed with supportive therapies, high-risk genotypes require disease-modifying treatment.[21] In order to determine how genetic heterogeneity modulates drug sensitivity, we edited K562 cells and isolated single cell-derived SF3B1m subclones (S-A, S-R, S-S) and SF3B1 wild-type (WT) lines with AAVS1 disruption as a control (**Fig. 9A-B**). To study how these genotypes respond to conventional and novel therapeutics, we screened K562 lines with the library of 166 FDA-approved and investigational compounds. Hypomethylating agents and chemotherapy are front-line treatments for myeloid neoplasms, including conditioning to hematopoietic stem cell transplantation that is currently the only curative option.[21,22] Comparison of the half maximal cell growth inhibitory concentration (IC50) between SF3B1 WT control and SF3B1m subclones revealed that high-risk genotypes generally have inferior response to chemotherapeutics (e.g., doxorubicin, idarubicin, cytarabine) as single agents or combinations (**Fig. 9C**), pointing to the need to identify alternative strategies.

Spliceosome-mutant cells require normal spliceosome function for viability, providing the rationale for therapeutic development of spliceosome-inhibiting compounds.[55,58,59,60] However, the initial clinical trial of H3B-8800, a derivative of SF3b inhibitor pladienolide B, showed minimal clinical responses.[57] To investigate if genetic heterogeneity affects response to spliceosome inhibitors, we first profiled mis-splicing patterns in isogenic SF3B1 WT control and SF3B1m K562 lines (S-A, S-R, S-S). No major differences were identified in the relative proportions of mis-splicing events between SF3B1m subclones (**Fig.10A**). The proportion of a3'ss events was similar irrespective of co-mutation status (**Fig.10B**), and the degree of gene mis-splicing was correlated between SF3B1m genotypes (**Fig.10C**). Over 60% of mis-spliced genes in high-risk genotypes were shared with SF3B1-only S-A and only ~10% were unique to S-R or S-S genotypes (**Fig.10D-E**). These data show that splicing alterations induced by mutant-SF3B1 are stereotyped and largely independent of co-mutations, suggesting that spliceosome modulators may target heterogeneous SF3B1m subclones.

Indeed, *RUNX1* or *STAG2* mutated SF3B1m cells maintained elevated response to SF3b inhibitor pladienolide B, with a ~3-fold reduction in IC50 compared to SF3B1 WT control (IC50 control WT 5.22 nM, S-A 1.65nM, S-R 1.65 nM, S-S 1.26nM) (**Fig.11A-B**). However, co-mutations altered the response to PRMT inhibitors which modulate splicing through distinct mechanisms.[60,86] STAG2, but not RUNX1 loss, selectively sensitized SF3B1m cells to type I PRMTs inhibitor MS023 (IC50 S-A 20.0 uM, S-R 21.6 uM, S-S 1.97 uM) (**Fig.11C-D**). By contrast, S-A and S-R displayed higher sensitivity to PRMT5 inhibition than SF3B1 wild-type cells (IC50 control WT 11.5 uM, S-A 2.48 uM, S-R 0.50 uM), while S-S had essentially no response to this compound (**Fig.11E-F**). These data suggest that genetic heterogeneity affects response of SF3B1m cells to spliceosome modulators, however they retain elevated sensitivity to SF3b inhibition.

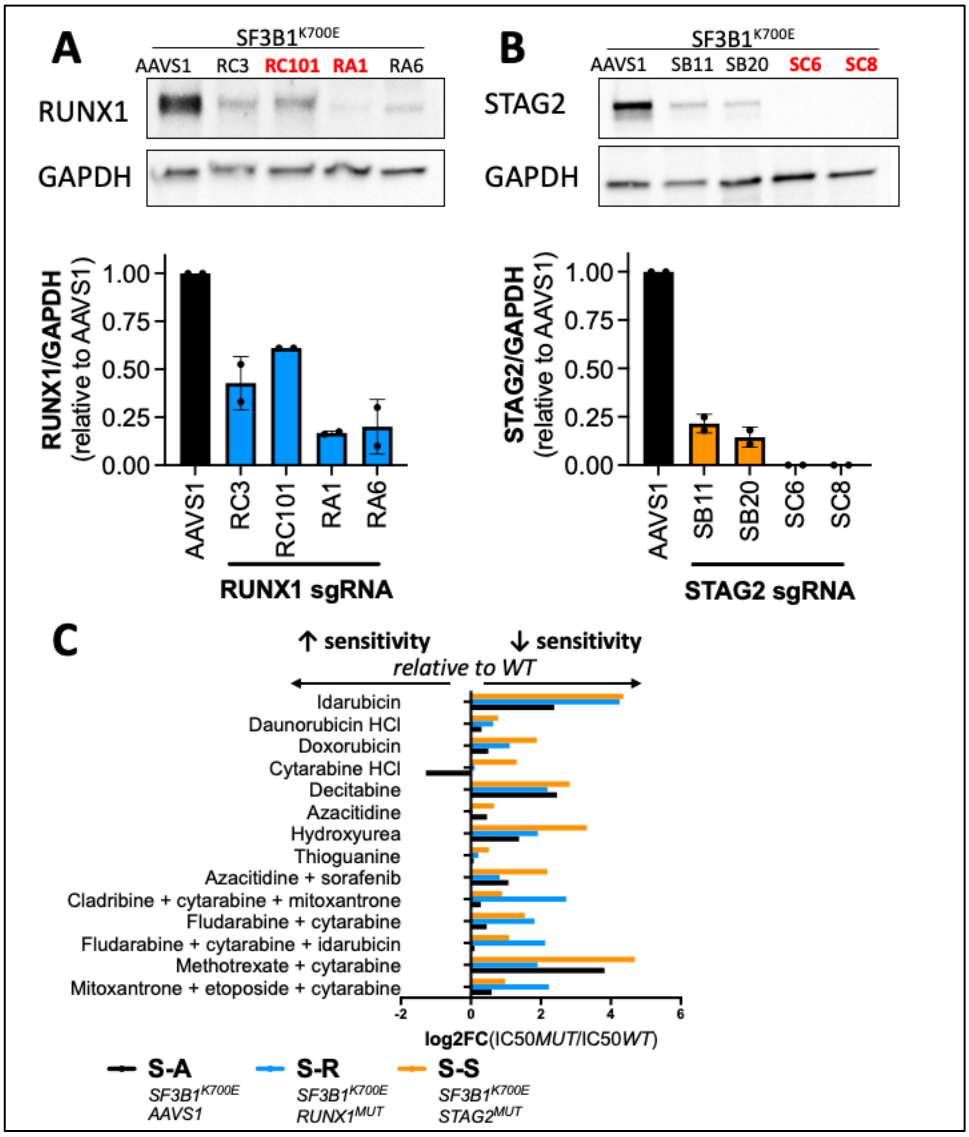

Figure 9. High-risk genotypes have poor response to chemotherapeutics.

(A), (B) Single cell derived clones generation: Western Blot analysis of RUNX1 (A) or STAG2 (B) protein, with quantification relative to AAVS1 control. n = 2 independent experiments. (C) Quantification of drug response to chemotherapeutics. Log2 Fold Change of IC50 value in S-A, S-R and S-S relative to control WT cells is represented.

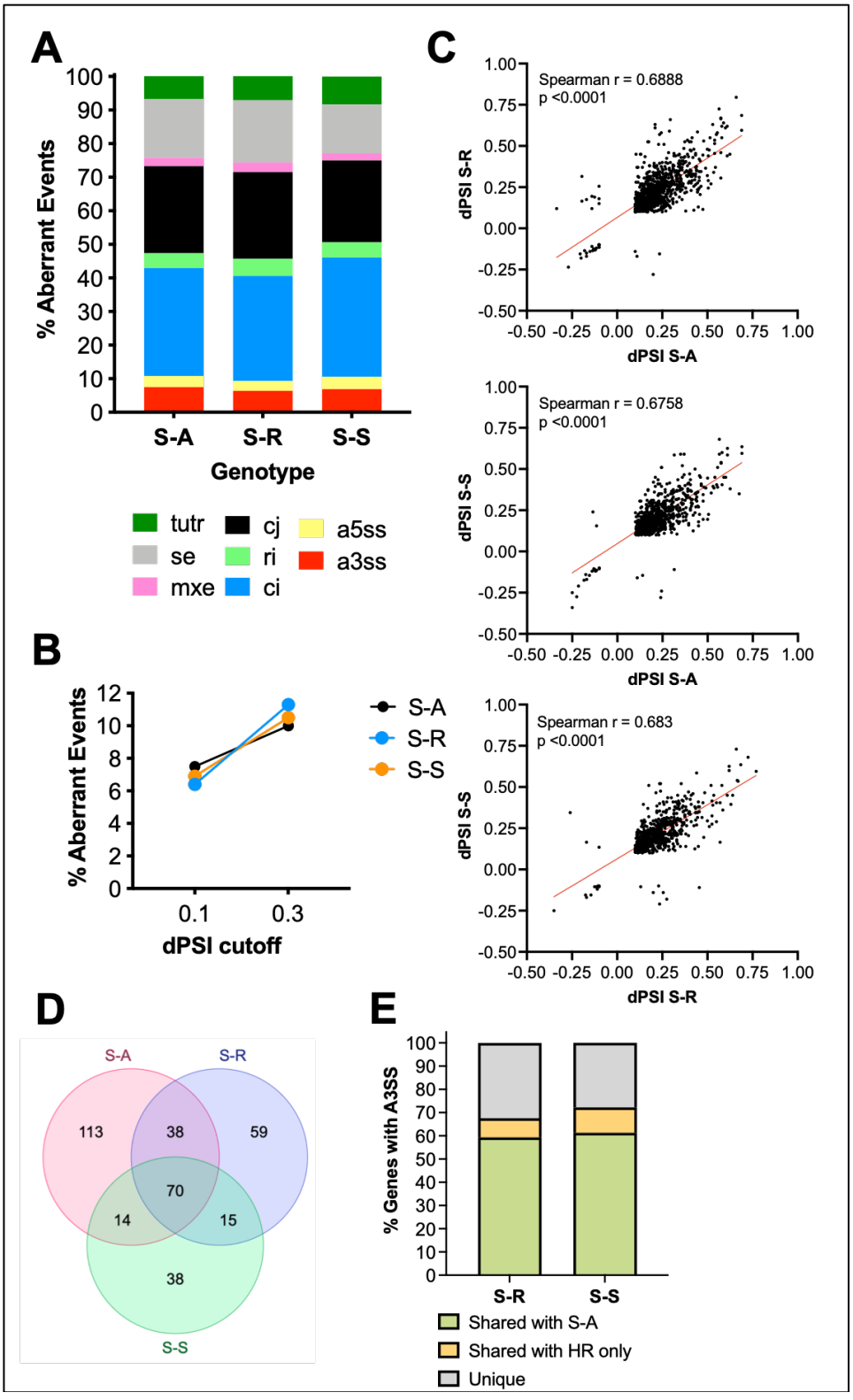

Figure 10. Mutant-SF3B1 induced mis-splicing is conserved in presence of high-risk co-mutations.

(A) Proportion of mis-spliced isoforms in S-A, S-R, S-S. Splicing events classified as arising from tandem 3' UTRs (tutr), cassette or skipped exons (se), retained introns (ri), mutually exclusive exons (mxe), alternative usage of normally constitutively spliced junctions (cj), alternative retention of normally constitutively spliced introns (ci), alternative 5'ss (a5ss), and alternative 3' ss (a3ss). (B) Proportion of a3ss in S-A, S-R and S-S at a dPSI cut-off of 0.1 or 0.3. (C) Spearman correlation between the quantitative levels of aberrant splicing in S-A versus S-R (top), S-A versus S-S (middle) and S-R versus S-S (bottom). (D) Analysis of the overlap between genes with a3ss in S-A, S-R and S-S. (E) Proportion of shared versus uniquely mis-spliced genes in high-risk genotypes.

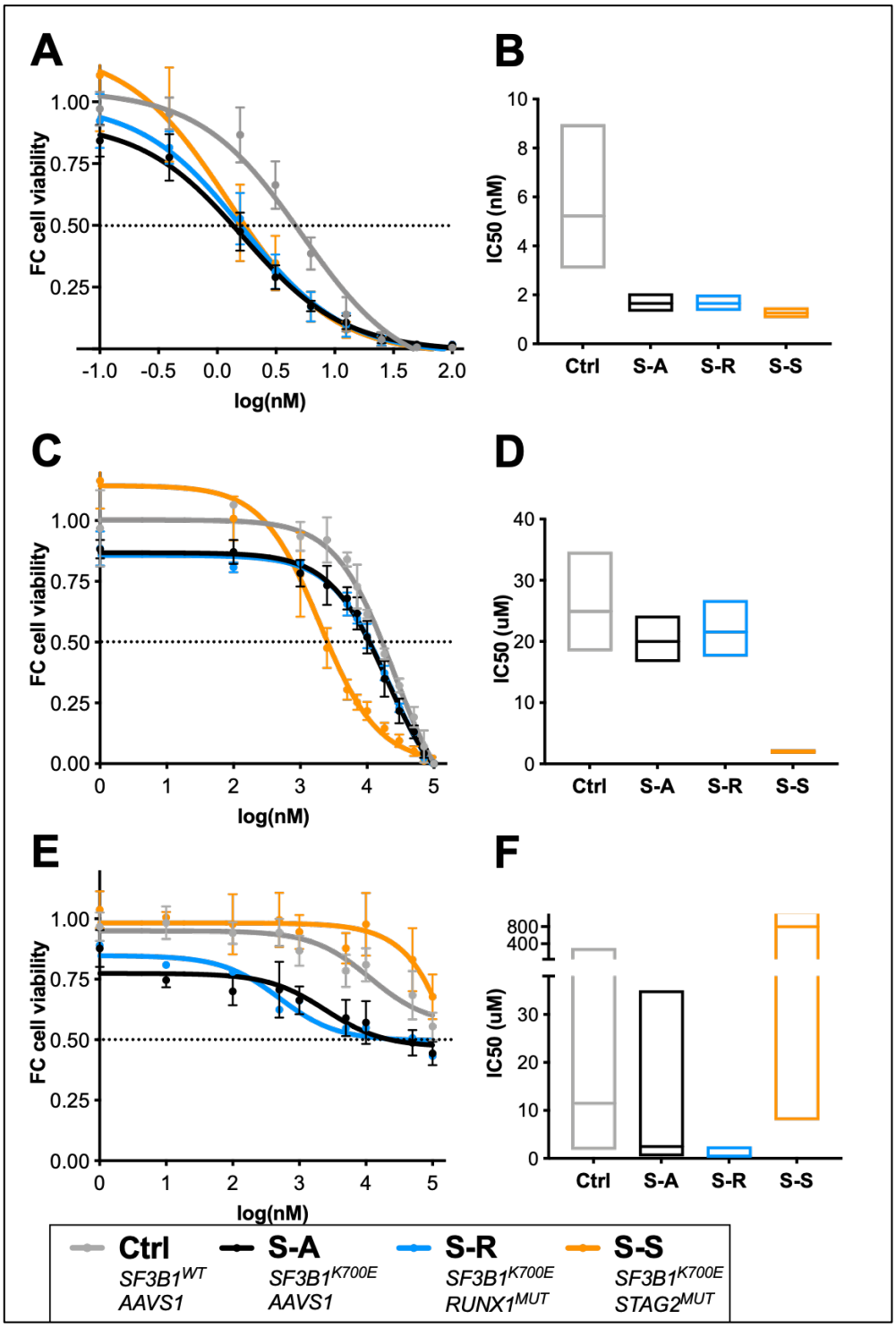

Figure 11. Genetic heterogeneity impacts response to spliceosome inhibition.

(A) Dose-response curve for SF3b inhibitor Pladienolide B. n = 3-4 independent experiments, 2 clones per genotype. (B) Pladienolide B IC50 values with 95% CI, relative to fig.11A. (C) Dose-response curve for type I PRMTs inhibitor MS-023. n = 2-3 independent experiments, 2 clones per genotype. (D) MS-023 IC50 values with 95% CI, relative to fig.11C. (E) Dose-response curve for PRMT5 inhibitor GSK3326595. n = 3-4 independent experiments, 2 clones per genotype.

CHK1 kinase inhibition targets SF3B1m genotypes.

We next sought to uncover novel compounds with broad activity against genetically heterogeneous SF3B1m clones. Our unbiased drug screen revealed AZD-7762, a checkpoint kinase (CHK) inhibitor,[87,88] as a promising candidate (**Fig. 12A**). Independent dose-response validation showed that SF3B1m genotypes were ~8-fold more sensitive to AZD-7762 compared to control cells (IC50 control WT 472 nM, S-A 55.4 nM, S-R 62.0 nM, S-S 57.9 nM) (**Fig. 12B-C**). AZD-7762 has equal potency against both CHK1 and CHK2.[89] To determine which component was driving response in SF3B1m cells, we next tested a specific CHK1 inhibitor Prexasertib and CHK2 inhibitor BML-277.[89,90,91] SF3B1m cells were not sensitized to BML-277 (**Fig. 12D-E**). By contrast, SF3B1m cells had ~7-fold higher sensitivity to Prexasertib irrespective of co-mutation status (IC50 control 18.8 nM, S-A 2.60 nM, S-R 3.29 nM, S-S 2.89 nM) (**Fig. 12F-G**). These data indicate that CHK1 inhibition (CHK1i) can be leveraged to target high-risk genotypes.

High selectivity of SF3B1m cells to Prexasertib prompted us to investigate the biological alterations underlying vulnerability to CHK1i. CHK1 is a central regulator of DNA damage and cell cycle. In response to DNA damage, ATR phosphorylates CHK1 on conserved S317/S345 residues. CHK1 in turn phosphorylates downstream effectors, such as CDC25A, CDC25C, and WEE1, to engage the intra-S and G2/M checkpoints.[92,93,94] CHK1 also has a critical albeit less well characterized function in the mitotic spindle assembly checkpoint,[95] and CHK1-knockout cells undergo mitotic arrest.[96] In line with previous reports,[97] all SF3B1m subclones displayed a modest increase in CHK1 S345 phosphorylation, which was enhanced by Prexasertib (**Fig. 13A-B**). To determine if cell cycle progression is altered in SF3B1m cells, we performed EdU cell cycle analysis (**Fig. 13C**). All SF3B1m subclones showed significantly reduced proportion of cells in S phase (**Fig. 13D, left**), and increased proportion of cells in G2/M phase with the exception of S-S (**Fig. 13E, left**). These differences were strongly enhanced after 24-hour Prexasertib treatment, with all SF3B1m subclones displaying reduced proportion of cells in S-phase (**Fig. 13D, right**) and increased proportion in G2/M (**Fig. 13E,**

right). These data reveal a previously unidentified G2/M block engaged by CHK1 inhibition in SF3B1m cells.

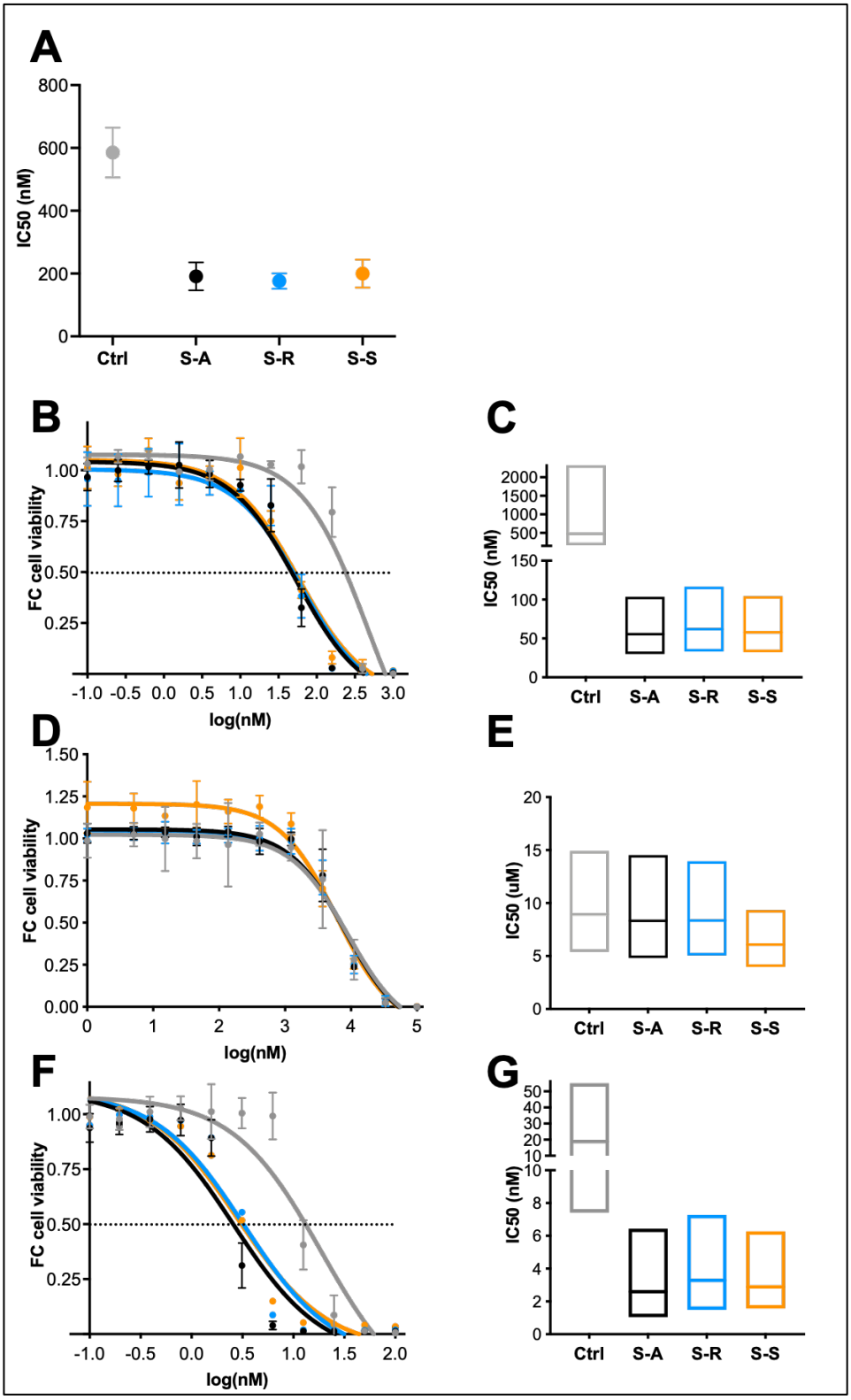

Figure 12. CHK1 inhibition is a vulnerability of SF3B1m cells.

(A) AZD-7762 IC50 value with standard error from high-throughput drug screen. (B) Dose-response curve for CHK1/CHK2 inhibitor AZD-7762. n = 3 independent experiments, 1 clone per genotype. (C) AZD-7762 IC50

values with 95% CI, relative to fig.12B. (D) Dose-response curve for CHK2 inhibitor BML-277. n = 3 independent experiments, 1 clone per genotype. (E) BML-277 IC50 values with 95% CI, relative to fig.12D. (F) Dose-response curve for CHK1 inhibitor Prexasertib. n = 3-4 independent experiments, 2 clones per genotype. (G) Prexasertib IC50 values with 95% CI, relative to fig.12D.

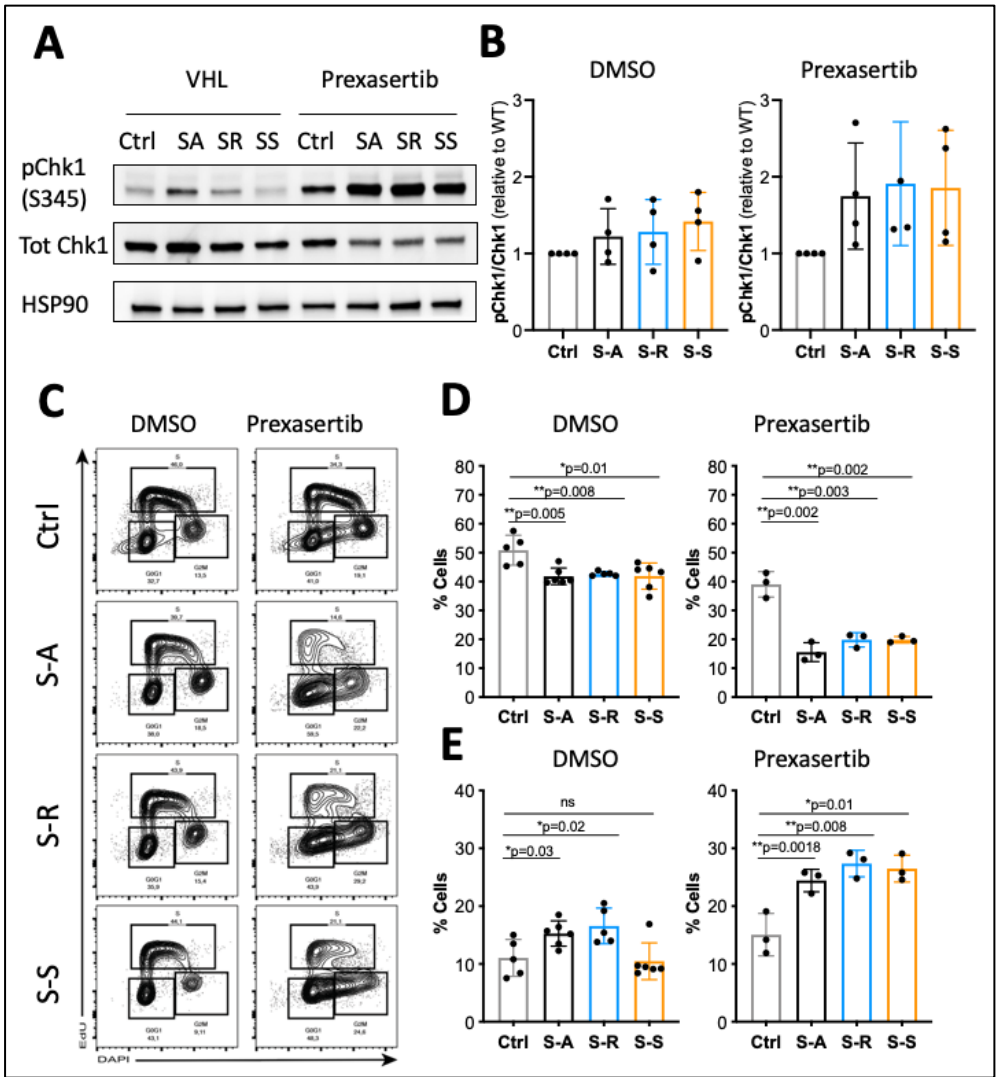

Figure 13. SF3B1m cells have increased CHK1 phosphorylation and alterations in cell cycle.

(A) Representative Western Blot of control WT, S-A, S-R and S-S after 5 hours of treatment with DMSO or 10nM Prexasertib. (B) Quantification of pCHK1/CHK1 ratio with DMSO (left) or Prexasertib treatment (right). n = 2 independent experiments, with 2 technical replicates each. (C) Representative flow cytometry plot of cell cycle analysis of control WT, S-A, S-R and S-S after 24 hours of treatment with DMSO or 10nM Prexasertib. (D) Proportion of cells in S phase with DMSO (left) or Prexasertib (right) treatment. n = 3-5, t test. (E) Proportion of cells in G2M phase with DMSO (left) or Prexasertib (right) treatment. n = 3-5, t test.

Coordinated mis-splicing of BUB1B and CDC27 sensitizes SF3B1m cells to CHK1 inhibition.

Since all SF3B1m genotypes were highly sensitive to Prexasertib irrespective of co-mutated genes, we hypothesized that genes mis-spliced by mutant-SF3B1 induce vulnerability to CHK1i. To address

this question, we first carried out RNA sequencing and splicing analysis of SF3B1m CD34^{+} HSPCs. Strikingly, pathway annotation of a3'ss mis-spliced genes in *SF3B1* K700E mutant cells showed that gene ontology (GO) categories (**Fig. 14A**) related to cell cycle regulation, mitotic checkpoint signaling, and chromosome segregation were most highly enriched in SF3B1m HSPCs. Similar gene categories were enriched in SF3B1m subclones derived from K562 cells (**Fig. 14B-D**). To identify candidate drivers, we annotated recurrently mis-spliced genes in the broad cell cycle GO category (**Fig. 14E**). BUB1B and CDC27 were consistently mis-spliced (a3'ss events) in all SF3B1m genotypes. Notably, BUB1B is regulated by CHK1 signaling during mitosis.[95] By contrast, none of the other annotated CHK1 targets were mis-spliced in SF3B1m cells, with the exception of mutually exclusive exon usage in CDC25C, which does not result in expression change (data not shown). BUB1B and CDC27 mRNA levels were significantly downregulated in all SF3B1m subclones (**Fig. 14F**), and BUB1B protein expression was concordantly decreased in SF3B1m cells irrespective of co-mutations (**Fig. 14G**).

BUB1B also known as BUBR1 is a CHK1-regulated mitotic kinase with dual function in chromosome alignment and mitotic checkpoint complex that monitors the spindle assembly.[95,98,99,100] CDC27 is a component of the anaphase-promoting complex that directly interacts with BUB1B.[100,101] We hypothesized that mis-splicing of both BUB1B and CDC27 delays G2/M and sensitizes SF3B1m cells to CHK1 inhibition. To test if reduced expression of BUB1B and CDC27 alters the sensitivity of normal cells to CHK1i, we transduced K562 SF3B1 WT cells with individual shRNAs targeting BUB1B, CDC27, or validated luciferase control. We achieved 75-90% knockdown for BUB1B (**Fig. 15A**) and 50% knockdown for CDC27 (**Fig. 15B**). As before, SF3B1m cells transduced with control shRNA were highly sensitive to Prexasertib (IC50 WT-M1 11.2 nM, MUT-M1 3.08 nM) (**Fig. 15C-F**). Knockdown of BUB1B sensitized SF3B1 WT cells to Prexasertib treatment (IC50 shRNA-B3 3.82 nM, shRNA-B2 6.13 nM) (**Fig. 15C-D**). Similarly, knockdown of CDC27 partially sensitized SF3B1 WT cells to Prexasertib (IC50 shRNA-C3 5.15 nM, shRNA-C2 7.60 nM) (**Fig. 15E-F**). WT cells with BUB1B or CDC27 knockdown showed a moderate increase of CHK1 S345 phosphorylation that was enhanced with Prexasertib treatment (**Fig. 15G-J**), suggesting that loss of mitotic regulators activates CHK1. Furthermore, knockdown of BUB1B or CDC27 in WT cells increased the proportion of cells in G2/M phase, partially recapitulating cell cycle alterations in SF3B1m cells (**Fig. 15K-L**). These data indicate that mis-splicing of BUB1B and CDC27 by mutant-SF3B1 delays mitotic progression leading to CHK1 activation and sensitizing SF3B1m cells to CHK1 inhibition.

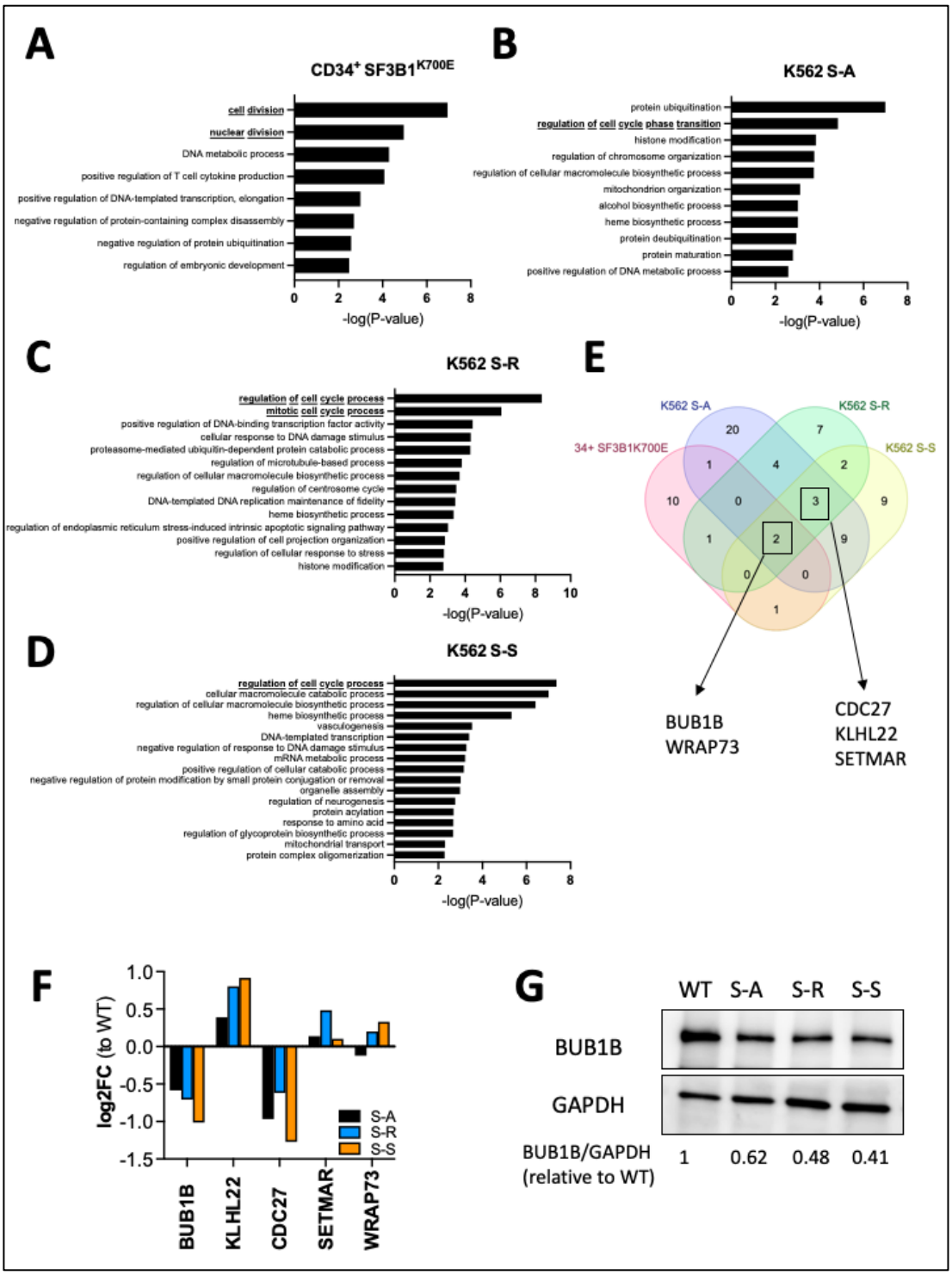

Figure 14. Regulators of cell cycle progression are mis-spliced by mutant-SF3B1.

(A) Gene ontology (GO) analysis of genes with a3ss in CD34⁺ HSPCs with SF3B1m relative to control WT CD34⁺ HSPCs. (B) GO analysis of genes with a3ss in K562 S-A cells relative to control WT cells. (C) GO analysis of genes with a3ss in K562 S-R cells relative to control WT cells. (D) GO analysis of genes with a3ss in K562 S-S cells relative to control WT cells. (E) Analysis of the overlap between mitotic regulators with a3ss in CD34⁺ HSPCs and K562 S-A, S-R and S-S. (F) Relative gene expression of recurrently mis-spliced mitotic regulators in K562 S-A, S-R and S-S compared to control WT cells. (G) Western Blot quantification of BUB1B protein in K562 control WT, S-A, S-R and S-S.

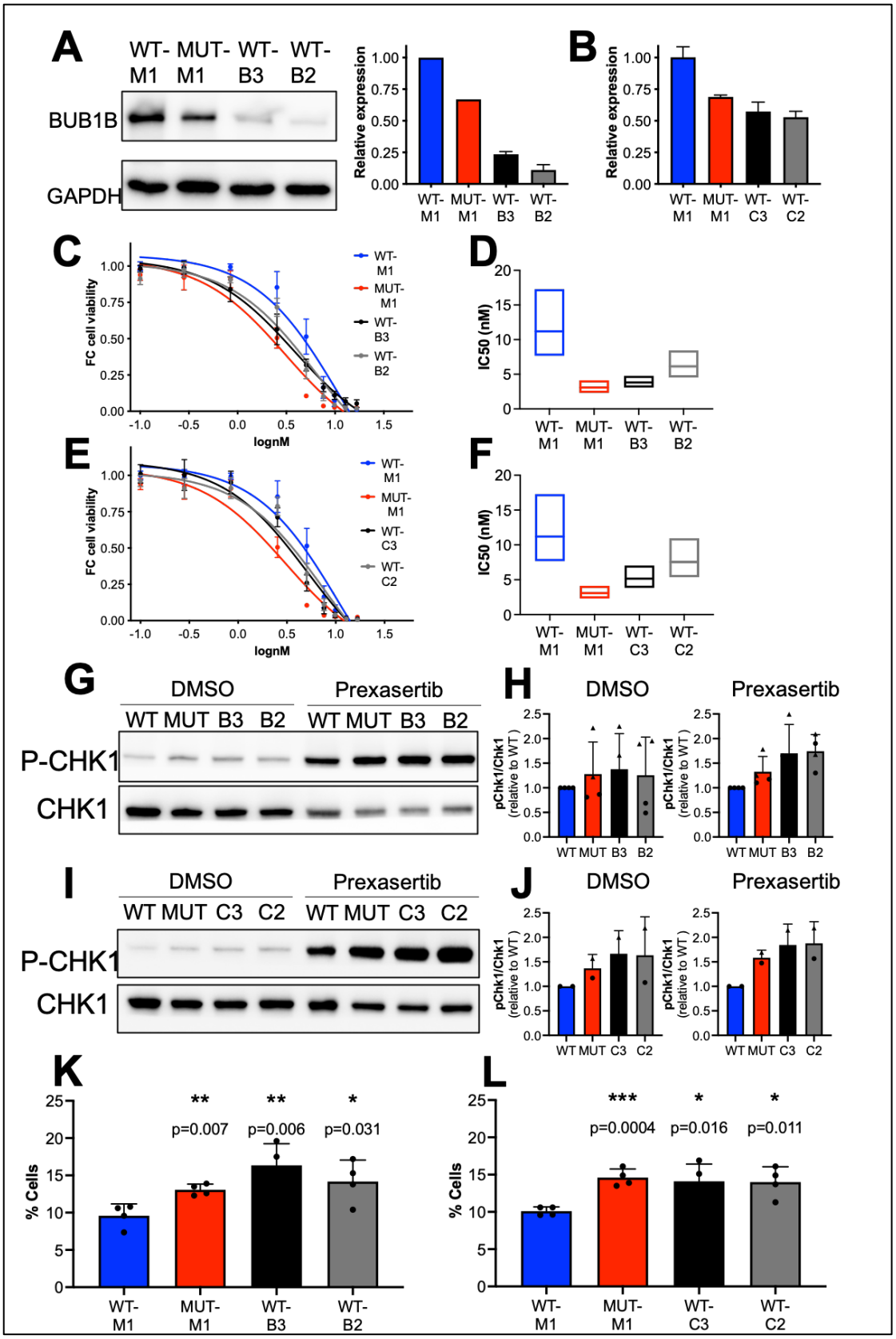

Figure 15. Coordinated mis-splicing of BUB1B and CDC27 sensitizes SF3B1m cells to CHK1 inhibition.

(A) BUB1B protein expression in WT and SF3B1m K562 cells and WT K562 cells transduced with shRNAs targeting BUB1B (B2 or B3). Representative Western Blot and quantification of normalized expression (to GAPDH), relative to WT K562 cells. n = 2 independent experiments. (B) CDC27 transcript expression in WT and

SF3B1m K562 cells and WT K562 cells transduced with shRNAs targeting CDC27 (C2 or C3). n = 2 independent experiments. (C) Dose-response curve for Prexasertib treatment in K562 SF3B1 WT cells with BUB1B KD. n = 2 sets of cell lines, with 3 independent experiments each. (D) Prexasertib IC50 values with 95% CI, relative to fig.15C. (E) Dose-response curve for Prexasertib treatment in K562 SF3B1 WT cells with CDC27 KD. n = 2 sets of cell lines, with 3 independent experiments each. (F) Prexasertib IC50 values with 95% CI, relative to fig.15E. (G) Representative WB of K562 cells with SF3B1 WT, SF3B1m and SF3B1 WT with BUB1B KD (B2, B3) after 5 hours of treatment with DMSO or 10nM Prexasertib. (H) Quantification of pCHK1/CHK1 ratio with DMSO (left) or Prexasertib treatment (right), relative to Fig 15 G. n = 2 independent experiments, with 2 technical replicates each. (I) Representative WB of K562 cells with SF3B1 WT, SF3B1m and SF3B1 WT with CDC27 KD (C2, C3) after 5 hours of treatment with DMSO or 10nM Prexasertib. (J) Quantification of pCHK1/CHK1 ratio with DMSO (left) or Prexasertib treatment (right), relative to Fig 15 I. n = 2 independent experiments. (K) Proportion of K562 cells with SF3B1 WT, SF3B1m and SF3B1 WT with BUB1B KD (B2, B3) in G2M phase. (L) Proportion of K562 cells with SF3B1 WT, SF3B1m and SF3B1 WT with CDC27 KD (C2, C3) in G2M phase.

CHK1 but not SF3b inhibition selectively targets SF3B1m HSCs.

Genetic heterogeneity and 'stemness' are major drivers of drug response in myeloid neoplasms, and it is essential that prospective therapeutics effectively target genetically diverse MDS HSCs.[33,85,102] To determine whether SF3B1m HSCs are sensitive to SF3bi or CHK1i, we tested Pladienolide B and Prexasertib in primary HSPCs with knock-in *SF3B1* K700E mutation. Edited cells were cultured for 7 days in presence of 2.5 nM Pladienolide B or 2.5 nM Prexasertib, and the relative proportion and absolute cell number of HSC/MPPs ($CD34^+CD133^+$), precursors ($CD34^+CD133^-$), and mature cells ($CD34^-$) were quantified by flow cytometry. SF3B1 WT HSPCs with BFP cassette knock-in in AAVS1 locus were used as control. SF3bi led to ~10-30% relative enrichment of $CD34^+CD133^-$ precursors in both SF3B1 WT and mutant cells (**Fig. 16A**). By contrast, CHK1 inhibitor led to a 2.5-fold relative enrichment of normal but not SF3B1m $CD34^+CD133^+$ HSCs, suggesting that normal HSCs are resistant to killing by Prexasertib (**Fig. 16B**). Next, we quantified the absolute number of viable cells after drug treatment relative to vehicle control. SF3B1m $CD34^+$ HSPCs (containing both HSCs and precursors) showed comparable sensitivity to SF3Bi or CHK1i (**Fig. 16C-D**). Pladienolide B treatment reduced the number of SF3B1m $CD34^+$ cells by ~50% (48% S-A; p=0.08, 47% S-R; p=0.1 and 40% S-S; p=0.1) compared to 33% for normal control CD34+ (**Fig. 16C**). Prexasertib similarly reduced SF3B1m $CD34^+$ cell number by ~50% (48% S-A, p=0.05; 54% S-R, p=0.1; 45% S-S, p=0.05) compared to 25% for normal control $CD34^+$ (**Fig. 16D**). Lastly, we evaluated the sensitivity of SF3B1m and normal $CD34^+CD133^+$ HSC/MPPs to SF3Bi or CHK1i. Pladienolide B treatment reduced the number of SF3B1m HSC/MPPs by ~75%, however normal control HSCs were decreased by ~50% (**Fig. 16E**), suggesting that both mutant and normal HSCs are highly sensitive to SF3bi. Reduction of Pladienolide B dose to 1 nM to preserve wild-type HSCs also reduced SF3B1m response, and no

significant differences were found between the genotypes (**Fig. 16G-H**). By contrast, Prexasertib entirely spared wild-type HSC/MPPs, while reducing the number of SF3B1m HSCs by about ~50% (HSCs vs vehicle: control WT 126%, S-A 65.3%, S-R 65.8%, S-S 49.6%) (**Fig. 16F**). Prexasertib treatment of primary SF3B1m BM samples reduced total $CD34^+$ HSPCs count by ~50% in 3 out of 4 patients (**Fig. 16I-K**). Taken together, these data indicate that SF3b inhibition is toxic to normal human HSCs and does not preferentially target SF3B1m HSCs. By contrast, CHK1 inhibition targets HSCs across the different SF3B1m genotypes while sparing normal HSCs.

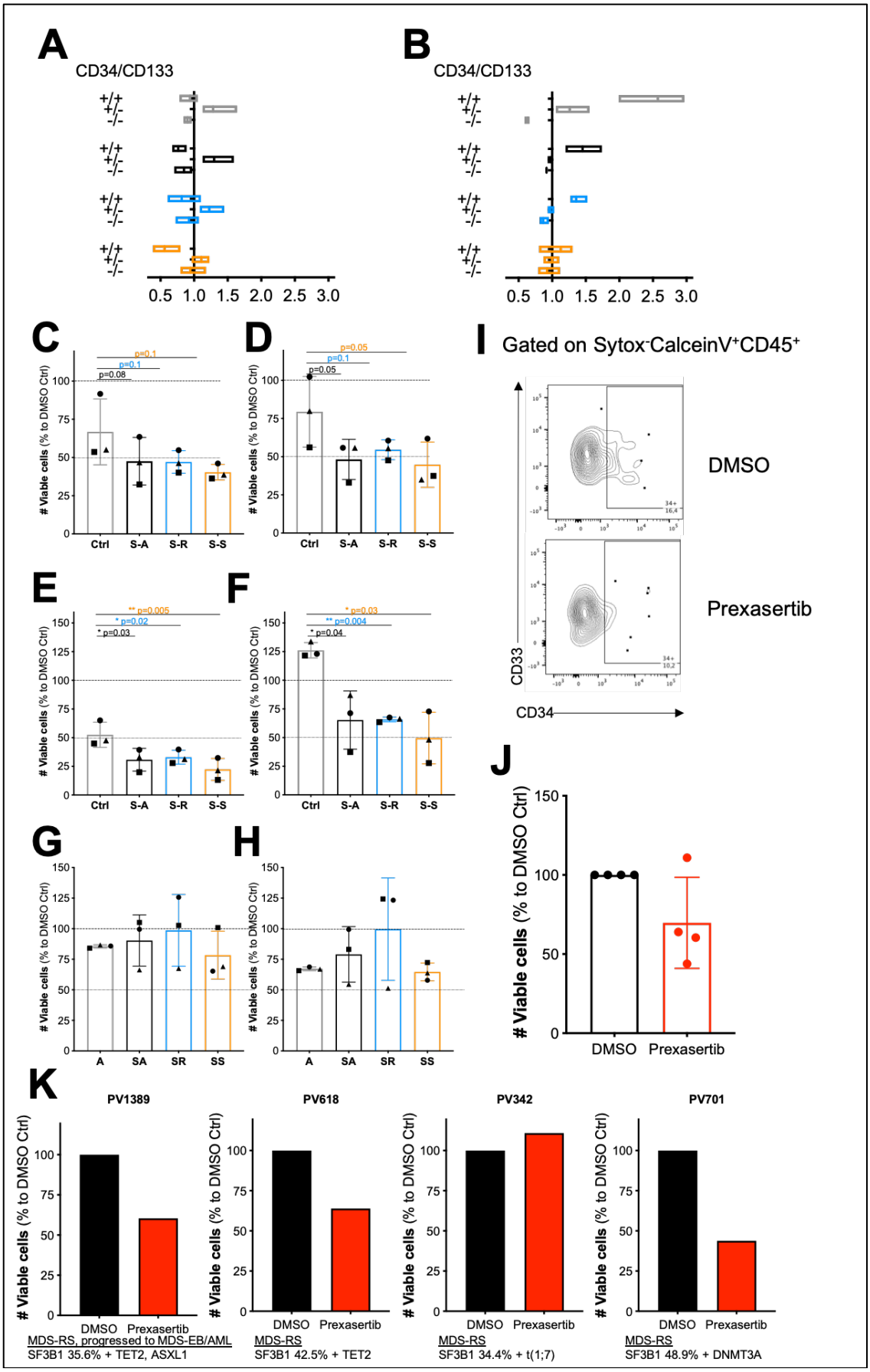

Figure 16. CHK1 but not SF3b inhibition selectively targets SF3B1m HSCs.

(A), (B) Relative proportion of $CD34^+CD133^+$ HSCs/MPPs, $CD34^+$ precursors, $CD34^-$ mature cells after 7 days with 2.5nM Pladienolide B (A) or 2.5nM Prexasertib (B) treatment. (C), (D) Total $CD34^+$ HSPCs count after 7 days with 2.5nM Pladienolide B (C) or 2.5nM Prexasertib (D) treatment. Percentage of viable cells relative to DMSO control is represented (mean ± s.d.). n = 3 independent experiments, t test. (E), (F) Total $CD34^+CD133^+$ HSC/MPPs count after 7 days with 2.5nM Pladienolide B (E) or 2.5nM Prexasertib (F) treatment. Percentage of viable cells relative to DMSO control is represented (mean ± s.d.). n = 3 independent experiments, t test. (G), (H) Total $CD34^+$ HSPCs (G) and $CD34^+CD133^+$ HSC/MPPs (H) count after 7 days with 1nM Pladienolide B. Percentage of viable cells relative to DMSO control is represented (mean ± s.d.) n = 3 independent experiments, t test. (I) Representative flow cytometry plot from Prexasertib treatment (2.5nM, 4 days) of primary SF3B1m patient samples. (J) Total $CD34^+$ HSPCs count in primary SF3B1m patient samples after 4 days of treatment with 2.5nM Prexasertib. Percentage of viable cells relative to DMSO control is represented (mean ± s.d.). (K) Total $CD34^+$ HSPCs count in individual patients samples, relative to Fig 16J. For each patient, WHO category at sampling, follow-up and co-mutation status were annotated.

Discussion

MDS are heterogenous disorders driven by acquisition of clonal genetic lesions.[1,2] Integration of these genetic abnormalities with clinical and cytogenetic data in IPSS-M has significantly improved patients stratification.[20] *SF3B1* mutations are commonly associated with alterations in the erythroid lineage, including ring sideroblasts formation and ineffective erythropoiesis, lower incidence of disease progression and better overall survival.[20,46,65] As so, SF3B1m MDS was proposed as a distinct nosologic entity,[46] and included in the most recent classifications of myeloid neoplasms as an independent diagnostic category.[14,15] Selected co-occurring genetic alterations can modulate the prognostic effect of *SF3B1* mutations, and affect patients outcome.[20,46] Consistently with IPSS-M indications,[20] we find that acquisition of *RUNX1* or *STAG2* mutations in SF3B1m patients transforms this relatively indolent condition into a high-risk malignancy, with increased BM blasts count and reduced overall survival. By leveraging gene editing technologies, we model evolution of SF3B1m MDS in primary $CD34^+$ HSPCs. We show that progression of SF3B1m MDS occurs through molecularly distinct clonal trajectories converging on expansion of SF3B1m HSC/MPPs that can be targeted through CHK1 inhibition.

Architecture of HSPC compartment and gene mutations interacting during differentiation contribute to phenotypic heterogeneity in MDS.[103] MDS initiating and propagating cells were identified in $CD34^+CD38^-$ HSCs pool.[103,104,32] Consistently, *SF3B1* mutations are acquired in $CD34^+CD38^-$ HSCs and are propagated to mature progenitors.[105,106] Altered hematopoietic differentiation is a hallmark of MDS, and stage specific changes in HSPC architecture occur during disease evolution.[35,83] RUNX1 and STAG2 have been implicated in the regulation of HSPCs self-renewal and differentiation by modulating gene expression and chromatin accessibility.[107,108] We find that *RUNX1* and *STAG2* high-risk mutations induce divergent lineage outcomes in SF3B1m HSPCs, with *RUNX1* promoting myeloid skewing and *STAG2* inducing a differentiation block. These functional changes are supported by divergent transcriptomic dysregulation, including key myeloid genes. However, both high-risk genotypes display increased proportion of $CD34^+CD38^-$ HSC/MPPs compared to isolated SF3B1m cells, consistent with the expansion of stem cells compartment in higher-risk MDS patients.[35] CD133 is a marker of HSCs and cancer stem cells, and decline in expression correlates with HSCs differentiation in vitro.[30] Primitive $CD34^+CD133^+$ SF3B1m HSCs are expanded by *RUNX1* but not *STAG2* mutations, implying a direct effect of genetic diversity on

SF3B1m HSPC architecture. As so, our study provides biological evidences of the genetic heterogeneity of SF3B1m MDS that emerged from IPSS-M.[20]

Lower-risk SF3B1m MDS is usually managed with transfusions, erythropoiesis-stimulating agents and Luspatercept, while higher-risk cases require disease-modifying therapy.[21,53] Hypomethylating agents are the frontline treatment for these patients, but aberrant HSPCs can persist through clinical remission, driving disease relapse.[83,25] Since *SF3B1* mutations are funding events,[105,106,109] the mutant clone needs to be targeted to eradicate the disease. Recognition that splicing factor mutant cells, including SF3B1m, retain dependency on the wild-type allele has driven the development of compounds that target spliceosome activity.[55] Preclinical studies showed that spliceosome modulation preferentially targeted splicing factor mutant cells *in vitro* and *in vivo*.[55,58,59,60] However, the first clinical trial with the SF3b inhibitor H3B-8800 showed minimal responses.[57] We find that genetic heterogeneity affects response of SF3B1m cells to spliceosome modulators, but high-risk genotypes retain elevated sensitivity to SF3b inhibition. Genetic alterations alone do not fully account for therapeutic outcomes, and cellular architecture of MDS plays a major role in modulating drug response.[33] Progenitor and stem cell fractions differentially respond to drugs.[102] We find that SF3bi positively selects $CD34^{+}CD133^{-}$ precursor cells in both SF3B1m and wild-type genotypes, and has elevated toxicity on wild-type HSCs, indicating poor selectivity for SF3B1m cells in the HSC pool. The restricted therapeutic index in HSCs emerging from our study could provide a biological basis for lack of partial/complete remission in H3B-8800 clinical trial[57], supporting the need to find alternative approaches to eradicate SF3B1m HSCs.

In this study, we identify CHK1i as a promising strategy to target SF3B1m genotypes. CHK1 is a major regulator of DNA damage response and cell cycle.[92] Shalini et al. previously showed increased activation of ATR/CHK1 signaling in SF3B1m cells as a consequence of R loops accumulation, and increased response to ATRi and CHK1i.[97] Here, we describe a previously unidentified G2/M block in SF3B1m cell engaged by CHK1i. We find that aberrant splicing induced by mutant-SF3B1 affects key regulators of chromosome alignment and mitotic progression, and we propose that coordinated mis-splicing of BUB1B and CDC27 sensitizes SF3B1m cells to CHK1i. Treatment of primary HSPCs with CHK1i selects for wild-type $CD34^{+}CD133^{+}$ HSCs but not SF3B1m, indicating that wild-type but not mutant HSCs are resistant to CHK1i. Consistently, CHK1i significantly reduces SF3B1m HSCs count without impairing SF3B1 wild-type HSCs. Incomplete eradication of MDS clones with current

treatments has been attributed to inability to eliminate clonally abnormal HSC pool.[25,35,83,109] Our study identifies CHK1 inhibition as an effective strategy to target SF3B1m HSCs irrespective of the co-mutation status, while sparing normal HSCs.

In conclusion, genetic and functional complexity in MDS modulates patients outcome and therapeutic response,[20,110] but dissecting the impact of precise co-mutation patterns has been hampered by heterogeneity of primary patients samples. Through gene editing of $CD34^+$ HSPCs, we define HSPC architectural changes during SF3B1m MDS evolution, and we identify CHK1i as a promising strategy to eradicate SF3B1m HSCs.

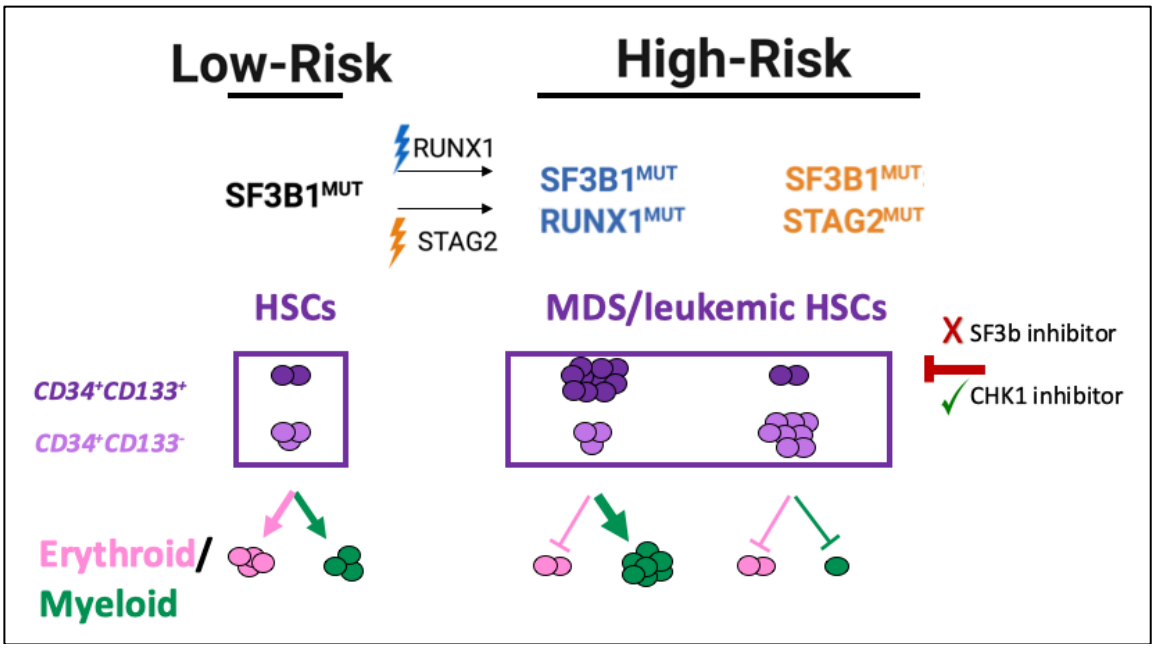

Figure 17. Model for SF3B1m MDS evolution and therapeutic targeting.

References

1. Fenaux, P., Haase, D., Sanz, G. F., *et al.* Myelodysplastic syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **25**, 57–69 (2014).
2. Malcovati, L., Hellström-Lindberg, E., Bowen, D., *et al.* Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood.* **122**, 2943–2964 (2013).
3. Xiaomei, M. Epidemology of myelodysplastic syndromes. *Am J Med.* **125**, S2-5 (2012).
4. Arber, D. A., Orazi, A., Hasserjian, R., *et al.* The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood.* **127**, 2391–2406 (2016).
5. Leone, G., Fianchi, L., Voso M.T. Therapy-related myeloid neoplasms. *Curr Opin Oncol.* **23**, 32834 (2011).
6. Bennett, J. M., Catovsky, D., Daniel, M.T., *et al.* Proposals for the classification of the myelodysplastic syndromes. *Br J Haematol.* **51**, 189–199 (1982).
7. Cazzola, M., Invernizzi, R. Ring sideroblasts and sideroblastic anemias. *Haematologica* **96**, 789–791 (2011).
8. Bain, B. J. Auer rods or McCrae rods? *Am J Hematol.* **86**, 689 (2011).
9. Vardiman, J. W., Harris, N. L., Brunning, R. D. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood* **100**, 2292–2302 (2002).
10. Vardiman, J. W., Thiele, J., Arber, D. A., *et al.* The 2008 revision of the WHO classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood* **114**, 937–952 (2008).
11. Bennett, J. M. A comparative review of classification systems in myelodysplastic syndromes (MDS). *Semin Oncol* **32**, S3-10 (2005).
12. Cazzola, M. Myelodysplastic syndrome with isolated 5q deletion (5q- syndrome). A clonal stem cell disorder characterized by defective ribosome biogenesis. *Haematologica* **93**, 967–972 (2008).
13. Valent, P., Orazi, A., Steensma, D.P., *et al.* Proposed minimal diagnostic criteria for myelodysplastic syndromes (MDS) and potential pre-MDS conditions. *Oncotarget* **8**, 73483–73500 (2017).
14. Arber, D. A., Orazi, A., Hasserjian, R.P., *et al.* International Consensus Classification of Myeloid

Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. *Blood* **140**, 1200–1228 (2022).

15. Khoury, J. D., Solary, E., Abla, O., *et al.* The 5th edition of the World Health Organization Classi fi cation of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia* **36**, 1703-1719 (2022).
16. Falini, B., Martelli, M. P. Comparison of the International Consensus and 5th WHO edition classifications of adult myelodysplastic syndromes and acute myeloid leukemia. *Am J Hematol* (2023).
17. Sperling, A. S., Gibson, C. J., Ebert, B. L. The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat Rev Cancer* **17**, 5-19 (2017).
18. Greenberg, P. L., Cox, C., LeBeau, M.M., *et al.* International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* **89**, 2079–88 (1997).
19. Greenberg, P. L., Tuechler, H., Schanz, J., *et al.* Revised international prognostic scoring system for myelodysplastic syndromes. *Blood* **120**, 2454–2465 (2012).
20. Bernard, E., Tuechler, H., Greenberg, P. L., *et al.* Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *N Engl J Med Evid.* **1**, (2022).
21. Fenaux, P., Haase, D., Santini, V., *et al.* Myelodysplastic syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* **32**, 142–156 (2021).
22. Cazzola, M. Myelodysplastic Syndromes. *N Engl J Med.* **383**, 1358–1374 (2020).
23. de Witte, T., Bowen, D., Robin, M., *et al.* Allogeneic hematopoietic stem cell transplantation for MDS and CMML : recommendations from an international expert panel. *Blood.* **129**, 1753–1762 (2017).
24. Bazinet, A., Bravo, G. M. New Approaches to Myelodysplastic Syndrome Treatment. *Curr Treat Options Oncol*. **23**, 668–687 (2022).
25. Unnikrishnan, A., Papaemmanuil, E., Beck, D., *et al.* Integrative Genomics Identifies the Molecular Basis of Resistance to Azacitidine Therapy in Myelodysplastic Syndromes Article Integrative Genomics Identifies the Molecular Basis of Resistance to Azacitidine Therapy in Myelodysplastic Syndromes. *Cell Rep.,* **20**, 572–585 (2017).
26. Estey, E. H., Thall, P.F., Cortes, J.E., *et al.* Comparison of idarubicin + ara-C-, fludarabine + ara-C-, and topotecan + ara-C- based regimens in treatment of newly diagnosed acute myeloid leukemia , refractory anemia with excess blasts in transformation , or refractory anemia with excess blasts. *Blood* **98**, 3575–3583 (2001).

27. Lancet, J. E., Uy, G. L., Cortes, J.E., *et al.* CPX-351 (cytarabine and daunorubicin) Liposome for Injection Versus Conventional Cytarabine Plus Daunorubicin in Older Patients With Newly Diagnosed Secondary Acute Myeloid Leukemia. *J Clin Oncol.* **36**, 2684-2692 (2018).
28. Kreso, A., Dick, J. E. Evolution of the Cancer Stem Cell Model. *Cell Stem Cell* **14**, 275–291 (2014).
29. Laurenti, E., Göttgens, B. From haematopoietic stem cells to complex differentiation landscapes. Nature. **553**, 418-426 (2018).
30. Doulatov, S., Notta, F., Laurenti, E., *et al.* Hematopoiesis: A Human Perspective. *Cell Stem Cell*. **10**, 120-136 (2012).
31. Velten, L., Haas, S., Raffael, S., *et al.* Human haematopoietic stem cell lineage commitment is a continuous process. *Nat Cell Biol.* **19**, 271-281 (2017).
32. Woll, P.S., Kjällquist, U., Chowdhury, O. *et al.* Myelodysplastic Syndromes Are Propagated by Rare and Distinct Human Cancer Stem Cells In Vivo. *Cancer Cell.* **25**, 794–808 (2014).
33. Ganan-Gomez, I., Yang, H., Ma, F. *et al.* Stem cell architecture drives myelodysplastic syndrome progression and predicts response to venetoclax-based therapy. *Nat Med.* **28**, 557-567 (2022).
34. Chen, J., Kao, Y., Sun, D. *et al.* Myelodysplastic Syndrome Progression to Acute Myeloid Leukemia at the Stem Cell Level. *Nat Med.* **25**, 103–110 (2019).
35. Will, B., Zhou, L., Vogler, T. O., *et al.* Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations. *Blood*. **120**, 2076–2086 (2012).
36. Delhommeau, F., Dupont, S., Della Valle, V. *et al.* Mutation in *TET2* in Myeloid Cancers. *N Engl J Med.* **360**, 2289–2301 (2009).
37. Haferlach, T., Nagata, Y., Grossman, V., *et al.* Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia* **28**, 241–247 (2014).
38. Cazzola, M., Della Porta, M. G., Malcovati, L. The genetic basis of myelodysplasia and its clinical relevance. *Blood.* **122**, 4021–4035 (2013).
39. Papaemmanuil, E., Gerstung, M., Malcovati, L., *et al.* Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* **122**, 3616–3627 (2013).
40. Ogawa, S. Genetics of MDS. *Blood* **133**, 1049–1059 (2019).
41. Makishima, H., Yoshizato, T., Yoshida, K, *et al.* Dynamics of clonal evolution in myelodysplastic syndromes. *Nat. Genet.* **49**, 204–212 (2016).

42. Lindsley, R. C., Mar, B.G., Mazzola, E., *et al.* Acute myeloid leukemia ontogeny is defined by distinct somatic mutations. *Blood* **125**, 1367–1377 (2015).

43. Wahl, M. C., Will, C. L., Lu, R. The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell.* **136**, 701–718 (2009).

44. Keren, H., Lev-Maor, G., Ast, G. Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet.* **11**, 345-355 (2010).

45. Seiler, M., Peng, S., Agrawal, A.A., *et al.* Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across Resource Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Rep.* **23**, 282–296 (2018).

46. Malcovati, L., Stevenson, K., Papaemmanuil, E., *et al.* SF3B1-mutant myelodysplastic syndrome as a distinct disease subtype - A Proposal of the International Working Group for the Prognosis of Myelodysplastic Syndromes (IWG-PM). *Blood.* **136**, 157-170 (2020).

47. Darman, R. B., Seiler M., Agrawal, A. A., *et al.* Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point Article Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3 0 Splice Site Selection through Use of a Different Branch Point. *Cell Rep.* **13**, 1033–1045 (2015).

48. Shiozawa, Y., Malcovati L., Gallì, A., *et al.* Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat. Commun.* **9**, 3649 (2018).

49. Madan, V., Li, J., Zhou, S., *et al.* Distinct and convergent consequences of splice factor mutations in myelodysplastic syndromes. *Am J Hematol* **95**, 133-143 (2019).

50. Pellagatti, A., Armstrong, R. N., Steeples, V., *et al.* Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood.* **132**, 1225–1240 (2018).

51. Dolatshad, H., Pellagatti, A., Fernandez-Mercado, M., *et al.* Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. *Leukemia.* **29**, 1092–1103 (2015).

52. Clough, C. A., Pangallo, J., Sarchi, M., *et al.* Coordinated missplicing of TMEM14C and ABCB7 causes ring sideroblast formation in SF3B1-mutant myelodysplastic syndrome. *Blood.* **139**, 2038-2049 (2022).

53. Fenaux, P., Platzbecker, U., Mufti, G. J., *et al.* Luspatercept in Patients with Lower-Risk Myelodysplastic Syndromes. *N Eng J Med.* **382**, 140-151 (2020).

54. Choudhary, G. S., Pellagatti, A., Agianian, B. *et al.* Activation of targetable inflammatory immune signaling is seen in myelodysplastic syndromes with SF3B1 mutations. *Elife*. (2022).
55. Lee, S. C., Dvinge, H., Kim, E., *et al.* Modulation of splicing catalysis for therapeutic targeting of leukemias with spliceosomal mutations. *Nat Med.* **22**, 672–678 (2016).
56. Hong, D.S., Kurzrock, R., Naing, A., *et al.* A phase I, open-label, single-arm, dose-escalation study of E7107, a precursor messenger ribonucleic acid (pre-mRNA) splicesome inhibitor administered intravenously on days 1 and 8 every 21 days to patients with solid tumors. *Invest New Drugs* **32**, 436–444 (2014).
57. Steensma, D. P., Wermke, M., Klimek, V. M., *et al.* Phase I First-in-Human Dose Escalation Study of the oral SF3B1 modulator H3B-8800 in myeloid neoplasms. *Leukemia.* **35**, 3542-3550 (2021).
58. Stanley, R. F., Abdel-Wahab, O. Dysregulation and therapeutic targeting of RNA splicing in cancer. *Nat Cancer.* **3**, 536-546 (2022).
59. Wang, E., Lu, S.X., Pastore, A., *et al.* Targeting an RNA-Binding Protein Network in Acute Article Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. *Cancer Cell* **35**, 369-384 (2019).
60. Fong, J. Y., Pignata, L., Goy, P. A., *et al.* Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation Article Therapeutic Targeting of RNA Splicing Catalysis through Inhibition of Protein Arginine Methylation. *Cancer Cell* **36**, 194-209 (2019).
61. Döhner, H., Estey, E., Grimwade, D., *et al.* Diagnosis and Management of AML in Adults: 2017 ELN Recommendations from an International expert panel. *Blood* **129**, 424–447 (2017).
62. Malcovati, L., Gallì, A., Travaglino, E., *et al.* Clinical significance of somatic mutation in unexplained blood cytopenia. *Blood.* **129**, 3371–3378 (2017).
63. Karimi, M., Nilsson, C., Dimitriou, M., *et al.* High-throughput mutational screening adds clinically important information in myelodysplastic syndromes and secondary or therapy-related acute myeloid leukemia. *Haematologica*. **100**, 223–225 (2015).
64. Shiozawa, Y., Malcovati, L., Gallì, A., *et al.* Gene expression and risk of leukemic transformation in myelodysplasia. *Blood*. **130**, 2642–2653 (2017).
65. Malcovati, L., Karimi, M., Papaemmanuil, E., *et al.* SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood.* **126**, 233–241 (2015).
66. Todisco, G., Creignou, M., Gallì, A., *et al.* Co-mutation pattern, clonal hierarchy , and clone size concur to determine disease phenotype of SRSF2 P95-mutated neoplasms. *Leukemia*. **35**,

2371–2381 (2021).

67. Khan, I. F., Hirata, R. K., Russell, D. W. AAV-mediated gene targeting methods for human cells. *Nat Protoc.* **6**, 482–501 (2013).
68. Aurnhammer, C., Haase, M., Muether, N., *et al.* Universal real-time PCR for the detection and quantification of adeno-associated virus serotype 2-derived inverted terminal repeat sequences. *Hum Gene Ther Methods.* **23**, 18-28 (2012)
69. Synthego Performance Analysis, ICE Analysis. 2019. v3.0. Synthego.
70. Hsu, J., Reilly, A., Hayes, B. J., *et al.* Reprogramming identi fi es functionally distinct stages of clonal evolution in myelodysplastic syndromes. *Blood.* **134**, 186-198 (2019).
71. Bray, N. L., Pimentel, H., Melsted, P., *et al.* Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol.* **34**, 4–8 (2016).
72. Love, M. I., Huber, W., Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
73. Broad MSigDB. https://www.broadinstitute.org/msigdb
74. Laurenti, E., Doulatov, S., Zandi, S., *et al.* The transcriptional architecture of early human hematopoiesis identifies multilevel control of lymphoid commitment. *Nat Immunol.* **14**, 756-763 (2013).
75. Inoue, D., Chew, G-L., Liu, B., *et al.* Spliceosomal disruption of the non-canonical BAF complex in cancer. *Nature***. 574**, 432-436 (2019)
76. Flicek, P., Ahmed, I., Amode, M.R., *et al.* Ensembl 2013. *Nucleic Acids Res.* **41**, 48–55 (2013).
77. Goldman, M., Craft, B., Swatloski, T., *et al.* The UCSC Cancer Genomics Browser: update 2013. *Nucleic Acids Res.* **41**, 949–954 (2013).
78. Katz, Y., Wang, E. T., Airoldi, E. M., *et al.* Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nat Methods.* **7**, 1009-1015 (2010).
79. Li, B., Dewey, C. N. RSEM : accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics.* (2011).
80. Trapnell, C., Pachter, L., Salzberg, S. L. TopHat : discovering splice junctions with RNA-Seq. *Bioinformatics.* **25**, 1105–1111 (2009).
81. Wagenmakers, E., Lodewyckx, T., Kuriyal, *et al.* Bayesian hypothesis testing for psychologists : A tutorial on the Savage–Dickey method. *Cogn. Psychol.* **60**, 158–189 (2010).
82. Lee, S.C., North, K., Kim, E., *et al.* Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. *Cancer Cell* **34**, 225-241 (2018).

83. Craddock, C., Quek, L., Goardon, N., *et al.* Azacitidine fails to eradicate leukemic stem / progenitor cell populations in patients with acute myeloid leukemia and myelodysplasia. *Leukemia.* **27**, 1028–1036 (2013).

84. Bejar, R., Lord, A., Stevenson, K., *et al.* TET2 mutations predict response to hypomethylating agents in myelodysplastic syndrome patients. *Blood.* **124**, 2705–2712 (2014).

85. Drusbosky, L. M., Singh, KN. K., Hawkins, K. E., *et al.* A genomics-informed computational biology platform prospectively predicts treatment responses in AML and MDS patients. *Blood Adv.* **3**, 1837-1847 (2019).

86. Ito, K., Thodima, V., Carter, J., *et al.* PRMT5 inhibition regulates alternative splicing and DNA damage repair pathways in SF3B1 R625G expressing uveal melanoma cells. Cancer Res **81**, 1137 (2021).

87. Oza, V., Ashwell, S., Almeida, L., *et al.* Discovery of Checkpoint Kinase Inhibitor (S)-5-(3-Fluorophenyl)-N-(piperidin-3-yl)-3-ureidothiophene-2-carboxamide (AZD7762) by Structure-Based Design and Optimization of Thiophenecarboxamide Ureas. *J. Med. Chem.* **55**, 5130–5142 (2012).

88. Zabludoff, S. D., Deng, C., Grondine, M. R., *et al.* AZD7762, a novel checkpoint kinase inhibitor, drives checkpoint abrogation and potentiates DNA-targeted therapies. **7**, 2955–2966 (2008).

89. Rundle, S., Bradbury, A., Drew, Y., *et al.* Targeting the ATR-CHK1 Axis in Cancer Therapy. Cancers. **9**, 1–25 (2017).

90. Smith, H. L., Southgate, H., Tweddle, D. A., *et al.* DNA damage checkpoint kinases in cancer. *Expert Rev Mol Med.* (2020).

91. Arienti, K. L., Burnmark, A., Axe, F. U., *et al.* Checkpoint kinase inhibitors: SAR and radioprotective properties of a series of 2-arylbenzimidazoles. J Med Chem. **48**, 1873-1875 (2005)

92. Patil, M., Pabla, N., Dong, Z. Checkpoint kinase 1 in DNA damage response and cell cycle regulation. *Cell Mol Life Sci.* **70**, 4009–4021 (2014).

93. Niida, H., Tsuge, S., Katsuno, Y., *et al.* Depletion of CHK1 Leads to Premature Activation of Cdc2-cyclin B and Mitotic Catastrophe. *J. Biol. Chem.* **280**, 39246–39252 (2005).

94. Rodriguez, R., Gagou, M. E., Eyers, P. A., *et al.* DNA replication stress in CHK1-depleted tumour cells triggers premature (S-phase) mitosis through inappropriate activation of Aurora kinase B. *Cell Death Dis.* **5**, (2014).

95. Zachos, G., Black, E. J., Walker, M., *et al.* CHK1 Is Required for Spindle Checkpoint Function.

Dev Cell. **12**, 247–260 (2007).

96. Tang, J., Erikson, R. L., Liu, X. Checkpoint kinase 1 (CHK1) is required for mitotic polo-like kinase 1 (Plk1). Proc Natl Acad Sci. **103**, 11964-11969 (2006).
97. Singh, S., Ahmed, D., Dolatshad, H., *et al.* SF3B1 mutations induce R-loop accumulation and DNA damage in MDS and leukemia cells with therapeutic implications. *Leukemia.* **34**, 2525-2530 (2020)
98. Zhang, G., Mendez, B. L., Sedgwick, G. G., *et al.* Two functionally distinct kineotochore pools of BubR1 ensure accurate chromosome segregation. *Nat Commun.* **7**, 12256 (2016).
99. Ditchfield, C., Johnson, V. L., Tighe, A. *et al.* Aurora B couples chromosome alignment with anaphase by targeting BubR1, Mad2, and Cenp-E to kinetochores. *J Cell Biol*. **161**, 267–280 (2002).
100. Elowe, S., Dulla, K., Uldschmid, A., *et al.* Uncoupling of the spindle-checkpoint and chromosome-congression functions of BubR1. *J Cell Sci.* **123**, 84–94 (2010).
101. Peters, J. The anaphase promoting complex/cyclosome: a machine designed to destroy. *Nat Rev Mol Cell Biol*. **7**, 644-656 (2006).
102. Tehranchi, R., Woll, P. S., Anderson, K., *et al.* Persistent Malignant Stem Cells in del(5q) Myelodysplasia in Remission. *N Engl J Med.* **363**, 1025–1037 (2010).
103. Chesnais, V., Arcangeli, M-L., Delette, C., *et al.* Architectural and functional heterogeneity of hematopoietic stem/progenitor cells in non-del (5q) myelodysplastic syndromes. *Blood.* **129**, 484-496 (2017).
104. Nilsson, L., Astrand-Grundström, I., Arvidsson, I., *et al.* Isolation and characterization of hematopoietic progenitor/stem cells in 5q-deleted myelodysplastic syndromes: evidence for involvement at the hematopoietic stem cell level. Blood. **96**, 2012–2021 (2000).
105. Mian, S. A., Roualt-Pierre, K., Smith, A. E., *et al.* SF3B1 mutant MDS-initiating cells may arise from the haematopoietic stem cell compartment. *Nat Commun.* (2015).
106. Mortera-Blanco, T., Dimitriou, M., Woll, P. S., *et al.* SF3B1 -initiating mutations in MDS-RSs target lymphomyeloid hematopoietic stem cells. *Blood.* **130**, 881–890 (2017).
107. Bellissimo, D. C., Speck, N. A. RUNX1 Mutations in Inherited and Sporadic Leukemia. *Front Cell Dev Biol.* (2017).
108. Jann, J., Tothova, Z. Cohesin mutations in myeloid malignancies. *Blood.* **138**, 649-661 (2021)
109. Mossner, M., Jann J-C., Wittig, J., *et al.* Mutational hierarchies in myelodysplastic syndromes dynamically adapt and evolve upon therapy response and failure. *Blood*. **128**, 1246–1259

(2016).

110. Shastri, A., Will, B., Steidl, U., *et al.* Stem and progenitor cell alterations in myelodysplastic syndromes. *Blood.* **129**, 1586–1594 (2017).

Acknowledgements

My PhD journey has come to an end, and this would not have been possible without the amazing people around me.

First, I'd like to thank my mentors for the constant guidance on the project, for supporting my scientific growth and helping me set my future goals. To Prof. Luca Malcovati, thank you for transmitting your passion for hematology. You have always been there when I had to take difficult decisions, and I'm so grateful for your time and wise feedbacks (and patience), looking forward to the next chapters! To Prof. Sergei Doulatov, thank you for everything I learnt during these past 3 years, whether in the lab or in the way to approach big scientific questions. I'll make treasure of all our exciting discussions and your suggestions, I hope this was just the beginning of many future collaborations.

A huge thanks to Dr. Anna Gallì, this project would have never started without you, and our fun summer of variants & candies. To Dr. Elisabetta Molteni, I've learnt so much from you since the first day I joined the lab, and to Dr. Sara Pozzi, for always offering your help.

Thanks to all the members of the Doulatov Lab, for creating such a collaborative environment, for the precious scientific inputs and for all the lab outings – we are not the bowling champions, but we are an amazing team! Thanks Dr. Andreea Reilly, Dr. Massiel Stolla, Raisa Stolitenko, Rachel Wellington, Dr. Courtnee Clough, Jason Kim, Dr. J Philip Creamer, Laura Baquero Galvis, Rochelle Bergantinos, Sintra Stewart, and Nelli Aydinyan! Raisa, thank you for always taking care of me and making sure I had everything I needed, inside and outside the lab – orange chocolate is the best! Jason, thank you for all your help in these last crazy months, it's been a pleasure working together and watching you grow so fast. Courtnee, thank you for being the best SF3B1 buddy and an amazing friend. To all the experiments we have done together, happy hours, trips – can't wait for our next adventures together!

Thanks to all my friends for our never-ending phone calls, moving to Seattle at the beginning of the pandemic was not easy, but I knew I could always count on you. To Valentina, Elena, Anna, Giulia

and Gabriele, thank you for making me smile also in the toughest moments. To Giovanna and Federico, thank you for sharing joys and sorrows of our PhD journeys.

Finally, the biggest thanks go to my family, for giving me all I needed to follow my dreams. Thanks to my grandma, for always being my first supporter. Thanks to my parents for believing in me also when I thought I couldn't make it. I know how hard it was being so far away, but I could feel all your love – also shipped in the form of boxes of food. Thanks to Riccardo, because it doesn't matter how many (thousands) kilometers you are away, you always know how to make me feel loved and supported. To all the hikes, ski trips, fancy meals, flights and hours on the phone (including my hundredth practice of the same talk), I can't imagine arriving here without you. Looking forward to the next adventures of our life together.

Thanks to all of you for shaping the person and scientist I've become. There are no words to say how grateful I am for having you in my life.