Dissecting the role of aberrant splicing in mutant-*SF3B1* myelodysplastic syndromes

Courtnee A. Clough

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2022

Reading Committee:
Sergei Doulatov, Chair
Robert K. Bradley
Suzanne Hoppins

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Dissecting the role of aberrant splicing in mutant-*SF3B1* myelodysplastic syndromes

Courtnee A. Clough

Chair of the Supervisory Committee:

Sergei Doulatov, Associate Professor

Hematology, University of Washington

Mutations in RNA splicing factors are the most recurrent genetic aberrations found in myeloid malignancies, such as myelodysplastic syndromes (MDS). These mutations are found in "hotspot" regions and most occur within the core U2 spliceosome factors *SF3B1, U2AF1,* and *SRSF2.* One common MDS subtype, MDS with ring sideroblasts (MDS-RS) is highly correlated with heterozygous gain-of-function mutations in *SF3B1*, which occur in ~80% of these patients. This subtype is characterized by ineffective erythroid differentiation and ring sideroblast (RS) formation, erythroid precursors with iron-laden mitochondria. Here, we describe the development of an *in vitro* induced pluripotent stem cell (iPSC) model of MDS-RS that has robust erythroid differentiation, mutant-*SF3B1* mis-splicing, and reproducible RS formation. We leverage this model to functionally test the long-standing hypothesis that mutant *SF3B1* mis-splicing mimics inherited sideroblastic anemias through the dysregulation of heme synthesis and iron sulfur cluster (ISC) biogenesis. We find that mutant-*SF3B1* mis-splices *TMEM14C, PPOX,* and *ABCB7,* key genes in these metabolic mitochondrial iron pathways. We functionally determine that

reduction of TMEM14C and ABCB7 results in RS formation. To investigate factors contributing to ineffective erythropoiesis, we profile heme production and gene expression in maturing erythroid cells and identify dysregulation of heme synthesis which results in metabolic reprogramming. Future work will focus on the development of primary human models of mutant-*SF3B1* to functionally interrogate mitochondrial iron trafficking and ineffective erythropoiesis in MDS-RS. Taken together, this work provides the first direct evidence of the connection between mutant-*SF3B1* and RS formation and ineffective erythropoiesis in MDS.

Table of Contents

List of Figures

Acknowledgements

I want to start by thanking all of my teachers and professors over the years who always saw in me what I couldn't see within myself. Their constant encouragement to always seek the next challenge got me to where I am today.

I have a lot of people to thank who helped me go from an undergraduate interested in research to being accepted into graduate school. First, I'd like to thank Dr. Stacy Alvares who gave me my first research opportunity as an undergrad. Next, I'd like to thank Dr. David Rawlings and his lab at Seattle Children's Research Institute for prepping me for graduate school success during my three years as a research scientist in the lab. I'd like to especially thank Brenda Seymour, Malika Hale, Matthew MacQuivey, Danny Thomson, and King Hung for their constant friendship and support as we navigated the journey together. And a special thanks to Drs. Iram Khan and Karen Sommer for their mentorship.

My outstanding support network have helped carve the path to my PhD.

First, I'd like to thank my mentor, Dr. Sergei Doulatov. Being the first graduate student in the lab has had its difficult moments but we navigated through them together and you always had my best interest at heart. I will take my experiences in helping to shape the lab over the past years and apply them to my own lab in the future. I look forward to our new relationship as colleagues and will always look back on my years in the lab in fondness. To the other members of the Doulatov lab past and present, Jasper Hsu, Stephanie Busch, Dr. Andreea Reilly, Dr. Massiel Stolla, Raisa Stolitenko, Rachel Wellington, Dr. Martina Sarchi, Dr. J Philip Creamer, Rochelle Bergantinos, and Sintra Stewart, thank you for creating such a supportive and fun environment. Massiel, thank you for all of the mentorship and friendship over the year. I'll always treasure our long talks and gym sessions. Raisa, thank you for always looking after me and for sharing all of your delicious canned goods with me over the years. Martina, you've only been in the lab for two years but it feels like you've been here the whole time. I've loved working

together on all of our *SF3B1* projects in the lab and for all of our happy hours after a long week in the lab. I look forward to my future trips to come visit you in Italy.

To my committee members Drs. Robert Bradley, Suzanne Hoppins, Bonny Brewer, and Jan Abkowitz, thank you for all of the advice and mentorship to navigate my academic journey. Each of you have played fundamental role in shaping both the project and myself.

I'd also like to thank my MCB family for helping me navigating through both graduate school and life. First, I'd like to thank the Dub Club, Marcella Waugh, Jilliane Bruffey, and Jeet Patel. Our countless conversations and coffee chats kept me going all these years and I look forward to seeing all that we accomplish. It's fitting that we are all graduating in the same quarter, graduate school wouldn't look the same without you all. A big thanks to my trivia group to make sure that I got out of the lab and had plenty of fun. To my great british bake off group, our dinner parties were always the highlight of the week. To the rest of my MCB cohort, our instant friendships during our first year made the transition back into school seamless. To the rest of the friends that I've made in MCB throughout the years, thanks for helping create such a family. To Maia Low, thanks for being the mama bear who always has all of our backs. And a huge thanks to the rest of the MCB staff who have always made each step as easy as possible and did everything they could to support us.

To John Wang and Taylor Wang, thank you for the constant friendship throughout graduate school and as our COVID bubble buddies. All of the time we spent together whether it was watching football (Go Hawks!), food adventures, or just hanging out, was wonderful and I look forward to our continued friendship as we all finish school and move into our next chapters.

And lastly, a huge thanks to all of my family for constantly cheering me on. To my grandparents, I will always appreciate having grown up with you and thank you for all of the support throughout my life. To my parents, thank you for always making sure I had what I needed to achieve my dreams. To my brother Mitch, thanks for all of the laughter and hijinks over the years, I love that we always turn right back into kids when we see each. To my

extended family, growing up next to all of my aunts, uncles, and cousins was such a blessing, thank you for all of the fun times. To Bill, Sheri, Breanna, Patrick, and Taylor, thank you for all of the love, support, and crazy nights of Pictionary.

And lastly, to my husband, Sean, you've been with me through this entire journey. Thank you for always making sure I'm taken care of whether it's cooking dinner, cleaning the house, or making an ice cream run. I would not have survived the chaos of graduate school without you and I cannot thank you enough.

To everyone who has every inspired and supported me, I love you all.

Dedication

To my grandmother, Yurn Fish, who didn't have the chance to see me finish, but gave me everything I needed to make it here

Chapter 1. Introduction

1.1 Characterization of Myelodysplastic Syndromes and development of diagnostic criteria

Myelodysplastic syndromes (MDS) are a collection of heterogenous hematopoietic stem and progenitor cell (HSPC) clonal disorders. The majority of diagnoses occur in the elderly populations[1] with the median age of diagnosis of 70 years of age[2] with a higher prevalence in men and white individuals. Each year, ~5 people per 100,000 are diagnosed in the general population, but this is likely an underestimate due to underreporting[1, 3]. Both congenital and acquired risk factors can predispose to MDS onset but the etiology for most patients is often unknown[2]. MDS is characterized by ineffective hematopoiesis driven by apoptosis of cells during maturation and risk for leukemic progression. Ineffective erythropoiesis results in anemia, the most common cytopenia within these patients[3].

The first description of myelodysplastic syndrome (MDS) is attributed to German physician, von Leube in 1900 who described a patient with “leukanamie,” a macrocytic anemia progressing to acute leukemia[4]. While similar pre-leukemia patients were reported in the subsequent decades, little was done to formally classify the disease of these patients. They were often described as “pre-leukemias” but as only ~30% of MDS patients go on to develop acute myeloid leukemia (AML)[5], by the 1970s, these disorders were reclassified as myelodysplastic syndromes, named for the misshapen shape of the myeloid cells derived from these patients.

In 1982, to better classify and treat this heterogenous group of patients, the French-American-British (FAB) cooperative group established the first set of MDS classification guidelines that striated these patients into five classifications that are still the framework for the classifications used today: refractory anemia (RA), refractory anemia with ringed sideroblasts (RARS), refractory anemia with excess blasts (RAEB), refractory anemia with excess blasts in

transformation RAEB-T), and chronic myelomonocytic leukemia (CMML). The striation of these patients was largely based on the proportion and morphology of blast cells in the peripheral blood (PB) and/or the bone marrow (BM) from these patients[6]. Importantly, these guidelines established the separate categories of refractory anemia with ringed sideroblasts identified by their abnormal accumulation of mitochondrial iron in erythroid precursors and the inclusion of CMML as a separate MDS subcategory. While these guidelines helped classify these patients into different treatment paradigms, the gross simplification of these initial criteria could not accurately predict patient leukemic progression or survival.

With the improvement in cytogenetic analysis tools during the early 1990s, there was a clear need to update MDS prognostic criteria to include additional clinical and cytogenetic data. After a surge in prognostic MDS studies were published in the late 1980s and early 1990s, in 1997 an International MDS Risk Analysis Workshop was convened to create the database of the International Prognostic Scoring System (IPSS) to create a standardized classification system. This group integrated data from seven previously published MDS risk analysis studies and reanalyzed this large patient dataset to generate consensus risk criterion[7]. Analysis of this integrated data identified three major drivers of progression to AML: cytogenetic abnormalities, % bone marrow myeloblasts, and the number of cytopenias. These analyses also sought to delineate MDS classification to create subgroups of patient outcomes, good, intermediate and poor. In addition to the three major driving factors, patient age and gender were also predictive variables for survival but did not have a significant role in predicting evolution to AML[7].

In 2001, the World Health Organization (WHO) published their own set of guidelines which built upon the FAB structure but formally incorporated cytogenetic analyses as outlined in the IPSS 1997 guidelines. The biggest difference between the classification system was the stratification between MDS and AML through the lowering of required blast counts from 30% to 20%. Low-risk categories including refractory anemia and refractory anemia with ring sideroblasts were also refined[8]. The most recent update to these guidelines came in 2016,

which normalized nomenclature to diagnose all patients with MDS and then a descriptor of the subcategory of MDS such as number of lineages affected (single or multi lineage), cytogenetics, excessive blasts, or ring sideroblasts. MDS with ring sideroblasts classification was refined to incorporate *SF3B1* mutations which lowered the threshold of %RS to 5% in patients harboring *SF3B1* mutations acknowledging the predominance of *SF3B1* mutations in MDS-RS patients[9]. In 2020, the International Working Group for the Prognosis of MDS proposed to classify *SF3B1*-mutant as a separate subtype of MDS as analyses show that these mutations are often the initiating event for MDS-RS development and generally portend good prognosis[10]. With the rapid onset of sequencing technologies, further integration of genetic and clinical analyses have refined these analyses and in 2022, the IPSS guidelines were updated to include an interactive risk calculator[11, 12].

1.2 Driver mutations in MDS

A defining feature of MDS is the selective advantage of mutant HSPC clones. This process is thought to begin as a clinically indolent condition known as clonal hematopoiesis (CH) in which certain HSPC clones begin to expand and populate the bone marrow before any overt disease is present. CH is associated with aging but refinement of the genetic and environmental factors that result in clonal expansion and evolution to MDS/AML is needed[13]. Advancements in sequencing technologies and access to CH and MDS patient samples have identified ~30 driver mutations of HSPC clonal expression that are associated with clonal progression to MDS or AML. Most of these genetic abnormalities are found in genes involved in RNA splicing, epigenetic regulation, DNA repair and regulation of gene expression[14, 15].

Interestingly, only a small subset of driver mutations, *SF3B1, SRSF2, TET2, ASXL1, DNMT3A,* and *RUNX1* are mutated in >10% of MDS patients[16] suggesting that these genes may play a critical role in progression from HSPC clonality to overt myeloid neoplasms.

Strikingly, mutations in splicing factor *SF3B1* are found in ~30% of MDS patients and occur in >80% of patients with a subtype of the MDS, MDS with ring sideroblasts (MDS-RS)[10]. This near universality implies a strong causative link between *SF3B1* mutations and mitochondrial iron dysregulation in MDS. To date, how *SF3B1* mutations result in clonality is not known as no animal model fully recapitulates the disease. The role of mutant-*SF3B1* in ring sideroblast formation will be discussed in the acquired sideroblastic anemia section of this chapter.

1.3 Sideroblastic anemias result from the dysregulation of mitochondria

Sideroblastic anemias (SA) are a heterogenous group of inherited and acquired disorders characterized by ineffective erythropoiesis and the marked accumulation of ring sideroblasts (RS) in erythroid precursors. Ineffective erythropoiesis, which results in aberrant apoptosis of erythroid cells during differentiation, causes anemia within these patients and can result in transfusion dependence[17]. Dysregulation of mitochondrial iron metabolism drives ring sideroblast formation, perinuclear iron-laden mitochondria, during the later stages of erythroid maturation. RS can be detected by Prussian blue iron staining and staining of SA patient bone marrow shows prevalent Prussian blue staining compared to controls[18]. Genetically, development of SA is associated with mutations in several mitochondrial pathways including heme synthesis, iron sulfur cluster (ISC,Fe-S) synthesis, mitochondria protein synthesis, and oxidative phosphorylation[19]. Interestingly, within the hematopoietic system, this mitochondrial iron accumulation appears to be unique to the erythroid lineage and only occurs in dysregulated erythropoiesis such as MDS-RS[20]. The specificity of RS formation in developing erythroid cells is likely a result of the high metabolic iron demands for heme synthesis and iron sulfur cluster generation required for differentiation. Roughly two thirds of the total iron in the body is metabolized by the erythron for hemoglobin synthesis making these cells uniquely vulnerable to any alterations to mitochondrial iron metabolism[21].

1.4 The majority of inherited sideroblastic anemias result from loss of heme synthesis

The most common form of inherited sideroblastic anemias (ISA) are X-linked sideroblastic anemias (XLSA) found in > 40% of patients[19]. These cases are caused by mutations in *ALAS2*[22], which encodes the enzyme that mediates the first rate-limiting step of heme synthesis and metabolizes glycine and succinyl-CoA to form aminolevulinate[23]. These mutations result in the reduction of ALAS2 either through loss of function (LOF) mutations or by gene modulation via alterations of transcription factor binding motifs and intronic enhancer elements[24, 25]. Heme insufficiency in these patients results in microcytic anemias, an adaptive response by erythroid cells to maintain the mean corpuscular cell size to hemoglobin ratio[26]. Mutations in *ABCB7*, essential for cytoplasmic ISC biosynthesis, result in a syndromic form of XLSA in which patients also exhibit spinocerebellar ataxia[27]. *ABCB7* loss impacts both ISC biogenesis and heme synthesis as Fe-S clusters are critical regulators of translation or enzymatic activity for several heme synthesis proteins including *ALAS2*[28], *ALAD*[29], and *FECH*[30]. While much rarer than mutations in *ALAS2*, a few dozen cases of mutations in the highly expressed, erythroid specific mitochondrial glycine importer *SLC25A38* have been identified in ISAs[31]. ALAS2 which has relatively low affinity for glycine is dependent on SLC25A38 import of glycine to sustain high levels of heme synthesis[32]. This subset of patients often has more overt disease and are often transfusion dependent likely due to substantial loss of heme synthesis[33]. Mutations in the rate limiting, terminal heme synthesis gene *FECH* often results in porphyria, a build-up of heme precursors, but ring sideroblast accumulation is occasionally observed in these patients as well[34]. Patients harboring mutations in ISC trafficking proteins are very rare, with only a couple dozen families reported[19]. While these proteins do not alter the biogenesis of *ABCB7,* mutations within *HSPA9*[35], *HSCB*[36], and *GLRX5*[37], alter the trafficking of nascent ISCs resulting in SA. Mutations in the zebrafish *GLXR5* homolog, *shiraz,* results in embryonic lethality[38] further supporting the need for ISCs for heme synthesis and erythroid development.

1.5 Acquired sideroblastic anemias are primarily driven by mutations in *SF3B1*

MDS-RS represents the majority of acquired SAs but non-genetic causes such as alcoholism, vitamin B6 deficiency, or heavy metal poisoning can also result in SA[10, 39]. Acquired SAs are far more common than ISAs but the genetic etiology of their pathogenesis was largely unknown until the sequencing advancements of the early 2010s. These seminal studies identified change of function mutations in core U2 splicing factor *SF3B1* in ~30% of total MDS patients and strikingly, were identified in 70-90% of MDS-RS patients suggesting a causative nature in RS formation[10, 40]. SF3B1 modulates pre-mRNA adenosine branch point binding and the majority of mutations within *SF3B1* have been shown to disrupt this process through dysregulation of SUGP1 binding[41]. This results in aberrant binding to upstream adenosine branch points, which results in alternative 3' splice site (a3ss) utilization[42, 43]. Many of these aberrant transcripts can be degraded via non-sense mediated RNA decay[44] resulting in a reduction of protein of mis-spliced genes but instances of 5'UTR mis-splicing resulting in over expression of proteins have also been reported[45] (Tam 2021).

1.6 Mutant *SF3B1* mis-splices several genes often mutated in inherited sideroblastic anemia

One of the challenges in dissecting the functional role of mis-splicing in mutant-*SF3B1* cells is generating RNA sequencing datasets with the required read depth to confidently identify changes in splicing junctions. To overcome some of these issues, several mutant-*SF3B1* hematopoietic cell lines have been engineered and sequenced to identify aberrant splicing signatures[42-44, 46-49]. Of note, the mis-splicing of *ABCB7*, the causative gene in XLSA with ataxia, results in rapid nonsense mediated RNA decay likely resulting in a reduction in iron sulfur cluster biogenesis[43, 44]. Mis-splicing of key heme synthesis components *TMEM14C* and *PPOX* have also been identified but their connection to SAs is less clear as neither have been directly

linked to RS formation. While RNA sequencing of mutant-*SF3B1* cell lines and patient samples has identified genes linked to iron metabolism, to date, no direct functional studies have been performed due to a lack of *in vitro* models which recapitulate both splicing and ring sideroblast biology.

In this dissertation I will address two outstanding questions:

1. How does mutant-*SF3B1* mediated aberrant splicing result in RS formation?

In chapter 2, I describe an *in vitro* mutant-*SF3B1* induced pluripotent stem cell approach to generate patient derived cell lines that recapitulate mis-splicing and erythroblast ring sideroblast formation. We leverage this novel model to test the hypothesis that mutant-*SF3B1* mimics inherited sideroblastic anemia by mis-splicing genes related to iron sulfur cluster biogenesis and heme synthesis. We perform RNA-sequencing based splicing analysis on these cell lines to identify mis-spliced genes and then functionally validate these targets using an *in vitro* erythroid differentiation assay to interrogate their roles in RS formation.

2. What is the role of mutant-*SF3B1* in driving ineffective erythropoiesis?

In chapter 3, I build upon the work in chapter 2 to focus on the dysregulation of heme synthesis and the potential contributions to ineffective erythropoiesis observed in MDS-RS patients. Leveraging our MDS-RS cell lines, we perform total porphyrin and heme analysis to measure the efficiency of heme synthesis in mutant-*SF3B1* cells. Utilizing differential gene expression from a mid- and late- stage timepoint of erythroid differentiation, we identify altered gene patterns that may drive ineffective erythropoiesis in MDS-RS.

Chapter 2. Coordinated missplicing of TMEM14C and ABCB7 causes ring sideroblast formation in SF3B1-mutant myelodysplastic syndrome

This work was originally published in *Blood*. Clough CA, Pangallo J, Sarchi M, et al. Coordinated missplicing of TMEM14C and ABCB7 causes ring sideroblast formation in SF3B1-mutant myelodysplastic syndrome. *Blood*. 2022-03-31 2022;139(13):2038-2049. doi:10.1182/blood.2021012652 © the American Society of Hematology**.**

2.1 Abstract

SF3B1 splicing factor mutations are near-universally found in myelodysplastic syndromes (MDS) with ring sideroblasts (RS), a clonal hematopoietic disorder characterized by abnormal erythroid cells with iron-loaded mitochondria. Despite this remarkably strong genotype-to-phenotype correlation, the mechanism by which mutant SF3B1 dysregulates iron metabolism to cause RS remains unclear due to an absence of physiological models of RS formation. Here, we report an induced pluripotent stem cell model of *SF3B1*-mutant MDS that for the first time recapitulates robust RS formation during in vitro erythroid differentiation. Mutant SF3B1 induces missplicing of ~100 genes throughout erythroid differentiation, including proposed RS driver genes *TMEM14C, PPOX,* and *ABCB7*. All 3 missplicing events reduce protein expression, notably occurring via 5′ UTR alteration, and reduced translation efficiency for *TMEM14C*. Functional rescue of TMEM14C and ABCB7, but not the non–rate-limiting enzyme PPOX, markedly decreased RS, and their combined rescue nearly abolished RS formation. Our study demonstrates that coordinated missplicing of mitochondrial transporters *TMEM14C* and *ABCB7* by mutant SF3B1 sequesters iron in mitochondria, causing RS formation.

2.2 Introduction

Sideroblastic anemias comprise a heterogeneous group of inherited and acquired disorders characterized by ring sideroblasts (RS), erythroid precursors harboring perinuclear iron-laden mitochondria[50]. The presence of RS in the marrow is a common pathological finding in myelodysplastic syndromes (MDS). Papaemmanuil et al first identified a strong association with somatic mutations in the core spliceosome factor SF3B1 in MDS with ring sideroblasts (MDS-RS)[40]. SF3B1 is part of the U2 small nuclear ribonucleoproteins complex, which mediates 3′ splice site selection by binding to intronic branchpoint sequences. SF3B1 change-of-function mutations in MDS and other cancers promote usage of alternative 3′ splice sites (a3′ss), generating a repertoire of aberrant transcripts[42, 43]. *SF3B1*-mutant MDS has been proposed to be a distinct nosologic entity marked by ineffective erythropoiesis and RS formation, lower incidence of multilineage dysplasia and leukemic progression, and higher overall survival[10, 51]. Despite this remarkably strong association, the molecular basis of ring sideroblast formation in *SF3B1*-mutated MDS remains incompletely understood.

Inherited sideroblastic anemias (ISAs) are often characterized by defects in mitochondrial iron and heme metabolism[52]. Mutations in *ALAS2*, a rate-limiting step of heme synthesis, are the most common cause of ISAs[22]. Mutations in *ABCB7*, the only known mitochondrial iron-sulfur cluster (Fe-S) exporter, cause X-linked sideroblastic anemia (XLSA) with ataxia[43, 53]. A compelling hypothesis is that *SF3B1* mutations in MDS mimic ISAs by altering splicing of iron-related genes[54], inducing iron accumulation in mitochondria in the form of mitochondrial ferritin[55]. Consistent with this hypothesis, molecular studies of *SF3B1*-mutated MDS bone marrow cells have identified pervasive missplicing of *ABCB7*, as well as *TMEM14C* and *PPOX,* which mediate sequential reactions in heme biosynthesis[44, 56]. Overexpression of ABCB7 in MDS-RS patient cells rescues erythroid differentiation defects[57, 58], suggesting that loss of *ABCB7* is a major driver in MDS-RS. To directly test the role of

aberrant gene transcripts in RS formation requires cell models of *SF3B1*-mutated MDS. SF3B1^{K700E} conditional knock-in mice develop ineffective erythropoiesis and anemia but not RS[59, 60]. MDS-RS patient cells give rise to RS, but only after transplantation into immunodeficient mice or on 3D scaffolds[61-63]. Although an induced pluripotent stem cell (iPSC)-derived model of XLSA has been developed, to date, there is no in vitro model of MDS-RS[48, 64]. Because current mouse and human cell models with *SF3B1* mutations do not give rise to RS in vitro, the role of missplicing of heme and iron metabolism genes in RS formation in MDS remains unresolved.

Here, we describe a patient-derived iPSC model of *SF3B1*-mutant MDS, which recapitulates missplicing patterns in MDS-RS patients and, for the first time, robust RS formation during in vitro erythroid differentiation. Using this model, we demonstrate that missplicing of *TMEM14C* and *ABCB7* causes RS formation in *SF3B1*-mutant MDS.

2.3 Results

2.3.1 iPSC model of SF3B1-mutant MDS with RS

To model MDS-RS, we derived induced pluripotent stem cells (iPSCs) from CD34$^+$ hematopoietic stem and progenitor cells from an MDS-RS patient with pathogenic *SF3B1*$^{G742D/+}$ and *EZH2*$^{R685H/+}$ mutations and a high percentage of RS (40-50% of marrow erythroblasts). Although comparatively rare in MDS-RS relative to the most common *SF3B1* mutation (*SF3B1*$^{K700E/+}$), the *SF3B1*$^{G742D/+}$ mutation has been previously observed in MDS-RS patients with *SF3B1* mutations[65, 66]. Following reprogramming, which can capture normal cells and clonal intermediates from individual patients ensuring identical genetic background[48], we isolated isogenic wild-type (WT), *SF3B1*-mutant only, and *SF3B1;EZH2*-mutant iPSC lines. iPSCs were differentiated into CD34$^+$ hematopoietic progenitor cells (HPCs) and conditionally immortalized by lentiviral delivery of 5 transcription factors (herein 5F-HPCs)[67]. 5F-HPCs were maintained as

undifferentiated $CD34^+$ cells and induced to undergo erythroid differentiation by doxycycline removal (**Fig. 2.5.1A**). We differentiated isogenic WT, *SF3B1*-mutant, and *SF3B1;EZH2* 5F-HPCs into erythroid cells using a two-stage protocol that includes transferrin-bound but not exogenous iron, and tracked differentiation using CD71 and CD235a/Glycophorin A (GlyA), as well as cell morphology by May-Grunwald Giemsa staining. *SF3B1*-mutant and WT 5F-HPCs followed a similar kinetics of erythroid differentiation, with *SF3B1*-only cells showing slightly accelerated differentiation (**Fig. 2.5.S1A**) consistent with previous reports[68]. All HPCs efficiently differentiated into late-stage $CD71^{lo}GlyA^+$ (**Fig. 2.5.1B,D**) polychromatic and orthochromatic erythroblasts (**Fig. 2.5.1C,E**). At the end of differentiation, *SF3B1*-only cells showed a slightly increased proportion of morphologically mature orthochromatic erythroblasts (59.7%) compared to WT (36.8%), and *SF3B1;EZH2* lines (38.8%) (**Fig. 2.5.1E**). WT and *SF3B1*-only HPCs had comparable total erythroid cell output, while the output of *SF3B1;EZH2* HPCs was reduced by ~2-fold (**Fig. 2.5.S1B**). To determine if iPSC-derived erythroid cells form RS, we stained cells from day 14 to 18 of culture with the Prussian Blue iron stain, and evaluated RS following the clinical criteria of iron-stained mitochondria encircling >1/3 of the nuclear perimeter. Strikingly, *SF3B1*-mutant but not isogenic WT erythroid cells formed high numbers of RS during terminal differentiation that were morphologically identical to primary patient RS (**Fig. 2.5.1F)**. Total RS count increased starting at day 15, coinciding with basophilic to polychromatic transition, and peaked at ~35% on the final day, day 18, of erythroid culture (**Fig. 2.5.1G)**. Most RS were observed on days 16-18 when orthochromatic erythroblasts become dominant (**Fig. 2.5.1E**). These data show that iPSC-derived MDS-RS progenitors efficiently form ring sideroblasts in erythroid culture, the first demonstration of this phenomenon *in vitro* for an *SF3B1*-mutant cell line.

2.3.2 Mutant SF3B1 missplicing during iPSC erythroid differentiation

Patient bone marrow cells with *SF3B1* mutations show a highly recurrent pattern of aberrant splicing marked by noncanonical usage of cryptic 3′ splice sites[42-44, 69, 70]. To determine if mis-splicing in the iPSC model was concordant with splicing changes found in patient cells, we profiled global splicing patterns with deep RNA-seq on isogenic WT, *SF3B1*-mutant, and $CD34^+$ 5F-HPC progenitor cells and compared with $CD34^+$ HPCs from MDS patients with or without a $SF3B1^{K700E/+}$ mutation[43, 65], as well as K562 cells with or without a knock-in $SF3B1^{K700E/+}$ mutation. This analysis identified thousands of misspliced transcripts in 5F-HPCs (**Figure 2.5.S2.A-B**). Global patterns of missplicing were largely concordant between our iPSC-based models and patient $CD34^+$ HPCs, as well as between cells with $SF3B1^{K700E/+}$ and $SF3B1^{G742D/+}$ mutations, suggesting that *SF3B1*-mutant 5F-HPCs recapitulate missplicing observed in primary MDS-RS (**Figure 2.5.2A**). The quantitative extent of missplicing was also modestly ($r = 0.45$) but significantly ($P < .0001$) correlated between 5F erythroblasts and K562 cells (**Figure 2.5.S2C**). These data demonstrate the stereotyped nature of *SF3B1* mutation-induced missplicing and validate iPSC-derived progenitors as a model of mutant *SF3B1* biology.

To assess missplicing changes during erythropoiesis, we profiled the transcriptomes of isogenic WT and SF3B1-mutant CD34+ progenitors, CD71+GlyA- proerythroblasts, and CD71+GlyA+ erythroblasts (**Figure 2.5.2B**). Although approximately 1000 transcripts were significantly misspliced at each stage of erythroid differentiation, only a small subset (<100) of genes exhibited very strong missplicing (change in isoform ratio >40%; **Figure 2.5.2C**). Mutant SF3B1-dependent missplicing was significantly correlated across all 3 stages of differentiation (**Figure 2.5.2D-E**), although some transcripts exhibited strong stage-specific differences (**Figure 2.5.S2D**). Further restricting to a3ss isoforms, the canonical consequence of *SF3B1* mutations[42], similarly showed high concordance between top misspliced genes and erythroid

stages and cell types, suggesting missplicing is conserved during erythroid differentiation (**Figure 2.5.2F**).

2.3.3 Missplicing downregulates expression of SF3B1 target genes

SF3B1-mutated MDS bone marrow cells display pervasive mis-splicing of *TMEM14C, PPOX, ABCB7,* and *MAP3K7*[42-44, 46, 47], key mediators of iron and heme metabolism and SF3B1-driven pathogenesis. *TMEM14C* and *PPOX*, encoding proteins that mediate sequential steps in heme biosynthesis, were both highly misspliced in *SF3B1*-mutant cells and upregulated during normal erythroid differentiation (**Figure 2.5.3A**). Cryptic 3′ splice site usage in *TMEM14C* and *PPOX* was strongly and significantly associated with mutant *SF3B1* in each system, with the added complexity of additional *PPOX* intron retention independent of genotype in primary MDS and K562 samples (**Figure 2.5.3B; Figure 2.5.S3A**). *ABCB7* has been proposed as a driver of RS formation in MDS[43, 56, 57]. *ABCB7* was modestly misspliced in our analysis (Figure 3A), likely due to efficient degradation of the misspliced messenger RNA (mRNA) by nonsense-mediated mRNA decay[44]. *MAP3K7* is implicated in *SF3B1*-mutant MDS pathogenesis[47] and was highly misspliced in our model. Missplicing of each of these 4 candidate genes was significantly associated with *SF3B1* mutations in each model system, including 2 distinct cohorts of primary MDS samples[43, 65] (**Figure 2.5.3B**). Based on this analysis, we selected *TMEM14C, PPOX*, *MAP3K7*, and *ABCB7* for functional studies of their potential roles in RS formation.

We next assessed the molecular consequence of each mis-splicing event. For *PPOX, ABCB7*, and *MAP3K7*, *SF3B1* mutations cause cryptic 3′ splice site usage that disrupts the open reading frame; in contrast, cryptic 3′ splice site usage in *TMEM14C* adds 14 nucleotides to the 5′UTR without altering the protein-coding sequence (**Figure 2.5.S3A**). Consistent with this, total RNA levels of *PPOX, ABCB7*, and *MAP3K7* were decreased in *SF3B1*-mutant 5F-HPCs, whereas RNA levels of *TMEM14C* were unchanged (**Figure 2.5.3C**). As 5′UTR sequence

influences translation efficiency[71], we next determined how *TMEM14C* 5′UTR missplicing altered protein output. We cloned the WT or mutant *SF3B1*–promoted (+14nt) 5′UTR into a tandem luciferase construct, which allows for accurate measurement of relative translational efficiency. The +14nt *TMEM14C* 5′UTR drove a moderate (20%), but highly significant (P = .015), reduction in LUC translational efficiency (**Figure 2.5.3D**). We therefore measured how altered *TMEM14C* 5′UTR usage influenced endogenous TMEM14C protein levels. We performed unbiased shotgun proteomics, which revealed decreased levels of TMEM14C, PPOX, MAP3K7, and ABCB7 proteins in 5F-HPCs (**Figure 2.5.S3B**). We then validated these results by probing individual proteins. Because expression of loading controls can change during differentiation, we measured protein levels in staged day-14 erythroid precursors. We raised an antibody to an N-terminal epitope of TMEM14C and found that TMEM14C protein levels were consistently reduced by an average of 40% in SF3B1-mutant cells (**Figure 2.5.3E**). We similarly confirmed reduced PPOX, MAP3K7, and ABCB7 protein levels in SF3B1-mutant erythroid cells (**Figure 2.5.3E**). Taken together, these data demonstrate that mutant SF3B1-dependent missplicing decreases expression of TMEM14C, PPOX, MAP3K7, and ABCB7 by reducing transcript stability or translation efficiency.

2.3.4 Rescue of TMEM14C and ABCB7 reduces RS formation

We next tested the hypothesis that missplicing of mitochondrial iron genes causes RS formation in *SF3B1*-mutated MDS with functional rescue experiments. *SF3B1*-mutant 5F-HPCs were transduced (at 60% to 90% efficiency) with individual lentiviral vectors encoding WT ORFs for TMEM14C, PPOX, ABCB7, MAP3K7, or LUC control (**Figure 2.5.4A**). Transduced *SF3B1*-mutant 5F-HPCs were differentiated for 18 days into erythroid cells. Overexpression level of each ORF was approximately twofold (**Figure 2.5.4B**). Overexpression of any of the ORFs did not impact the kinetics of erythroid differentiation assessed by flow cytometry and cell

morphology as all conditions had ~25% $CD71^{lo}GlyA^{+}$ cells and ~50% morphologically mature orthochromatic erythroblasts by day 18 of culture (**Figure 2.5.4C-F**). Cytospins were collected daily starting on day 15, when RS were first observed in untransduced cells, until day 18 (**Figure 2.5.1G**), and RS were quantified by staining with Perl's Prussian blue (**Figure 2.5.4G**). As expected, the proportion of RS in all ORF conditions increased between days 15 and 18 (**Figure 2.5.4H**), with most RS observed in orthochromatic erythroblasts (**Figure 2.5.S4A-B**). Erythroid cells transduced with control LUC ORF reached 37.3 ± 10.4% of RS (**Figure 2.5.4I**), similar to untransduced cells (**Figure 2.5.1G**). Overexpression of MAP3K7, a canonical mutant SF3B1 target with no known roles in iron metabolism, did not alter the frequency of RS (38.2 ± 9.7%) (**Figure 2.5.4I**). Overexpression of PPOX, not a rate limiting enzymatic step in heme synthesis, led to a modest decrease in RS formation (31.5 ± 6.9%) (**Figure 2.5.4I**). Strikingly, *SF3B1*-mutant erythroid cells overexpressing TMEM14C or ABCB7 displayed consistently lower RS formation throughout late-stage differentiation (23.84 ± 7.1% and 16.9 ± 9.8%, respectively) (**Figure 2.5.4G,I**). Because TMEM14C and PPOX are sequential components in heme synthesis and may display epistatic effects, we also analyzed RS in *SF3B1*-mutant cells transduced with a combination of TMEM14C and PPOX lentiviruses. Combined overexpression of TMEM14C and PPOX had similar RS counts compared with TMEM14C alone, suggesting that missplicing of *TMEM14C* but not *PPOX* drives RS formation (**Figure 2.5.4J**).

TMEM14C is an inner mitochondrial protein that transports protoporphyrinogen IX into the mitochondrial matrix, where PPOX converts it to protoporphyrin IX (PPIX) for heme synthesis[72]. Consistently, *SF3B1*-mutant erythroid cells transduced with TMEM14C or PPOX exhibited a small increase in PPIX fluorescence (**Figure 2.5.S4C-D**). We tested if increasing PPIX availability by blocking the nonspecific efflux pump ABCG2[73], not predicted to be misspliced by SF3B1, could rescue RS formation. Cells transduced with an shRNA targeting ABCG2 with ~65% knockdown showed a small increase in PPIX fluorescence and ~60% reduction in RS counts compared with an shRNA control, with no change in erythroid

differentiation (**Figure 2.5.S4E-I**). These data suggest that overexpression of TMEM14C may partially rescue RS formation by increasing the porphyrin pool available for heme synthesis.

Because rescue of TMEM14C and ABCB7 each partially rescued RS formation, we next assessed whether combinatorial rescue would further reduce RS. To test this, we cotransduced 5F-HPCs with dual fluorescent ORF lentiviruses and purified double-transduced $GlyA^+$ erythroid precursors by flow sorting on day 10 of erythroid differentiation. Control double-transduced sorted erythroid cells had higher baseline RS formation (60.1 ± 4.6%) compared with unsorted cells (37.3 ± 10.4%) (**Figure 2.5.4I,K**). Overexpression of TMEM14C or ABCB7 alone reduced percent RS compared with LUC control, with ABCB7 showing significantly stronger rescue compared with TMEM14C (13.4 ± 2.3% and 37.4 ± 7.9%, respectively) (**Figure 2.5.4K**), as seen with unsorted cells. Combined rescue of both TMEM14C and ABCB7 showed the highest reduction in RS counts (10.8 ± 2%) but was not significantly different compared with ABCB7 alone (**Figure 2.5.4K**), demonstrating that ABCB7 is the major driver of RS formation. Taken together, these data show that mutant *SF3B1*-dependent missplicing mimics ISAs by converging upon key mitochondrial iron pathways to drive RS formation in MDS.

2.4 Discussion

Mutations in splicing factor *SF3B1* lead to pervasive missplicing of mitochondrial iron metabolic genes *TMEM14C, PPOX*, and *ABCB7*[42-44, 46, 74]. Missplicing of *ABCB7*, the causative gene for XLSA with ataxia, and heme synthesis genes has been proposed to drive RS formation in SF3B1-mutant MDS[43, 46, 57]. However, these hypotheses have not been tested due to lack of physiological model systems of RS formation. Here, we develop a patient-derived iPSC model of *SF3B1*-mutant MDS which recapitulates global missplicing observed in primary MDS-RS cells and for the first time efficiently gives rise to RS *in vitro*. In our model, RS formation starts at the polychromatic stage and intensifies in orthochromatic erythroblasts, possibly reflecting a

perinuclear distribution of mitochondria in late-stage erythroblasts[75]. We demonstrate that reduced expression due to missplicing of mitochondrial transporters *TMEM14C* and *ABCB7* causes RS formation. Our findings explain the remarkably strong clinical association between *SF3B1* mutations and RS in MDS.

Inherited sideroblastic anemias are caused by mutations in key mitochondrial pathways, such as heme synthesis, mitochondrial DNA replication, and Fe-S cluster biogenesis[19, 50]. *ABCB7*, the only known mitochondrial Fe-S exporter[27], is mutated in XLSA with ataxia[53], and its missplicing has been considered a candidate driver of RS formation in MDS. Boultwood et al first reported decreased ABCB7 expression in MDS-RS[74]. Dolatshad et al showed that missplicing of *ABCB7*, due to usage of an alternative 3′ splice site, occurs in the HSPCs and erythroblasts of MDS patients with *SF3B1* mutations[43]. ABCB7 downregulation induced mitochondrial ferritin and reduced erythroid colony capacity in normal bone marrow cells, whereas its overexpression partially rescued erythroid defects of MDS-RS precursors[57]. Our functional rescue studies confirm the long-standing hypothesis that missplicing and reduced expression of ABCB7 is the major driver of RS formation in *SF3B1*-mutant MDS.

TMEM14C and *PPOX* are the only heme pathway components misspliced in MDS-RS patients, although their role has been less clear[43, 44, 55]. Unlike *ABCB7* or *PPOX* in which missplicing disrupts the coding sequence, missplicing occurs in the 5′UTR of *TMEM14C*. Missplicing within the 5′UTR is pervasive in *SF3B1*-mutant cells, although the functional consequences of these events are less clear. Although previous studies have not detected a change in TMEM14C expression in *SF3B1*-mutant cells[44], we show that *TMEM14C* missplicing reduces translation efficiency and total protein. Our functional rescue studies show that TMEM14C, but not PPOX, reduces RS formation. TMEM14C imports protoporphyrinogen IX, which is oxidized by PPOX into porphyrin[72]. However, PPOX is not rate limiting, and inherited *PPOX* mutations cause hepatic but not erythroid porphyria[76]. Independently increasing porphyrin levels via knockdown of *ABCG2*, a nonspecific PPIX efflux pump, also reduced RS

counts, suggesting that TMEM14C regulates the porphyrin pool. More refined measurements of PPIX and other heme pathway intermediates are needed to confirm this hypothesis.

Iron imported into mitochondria is rapidly incorporated into heme or Fe-S[77]. In our model, missplicing of *TMEM14C* and *ABCB7* perturbs both branches of this pathway, resulting in excess iron. ABCB7 exports an unidentified cytosolic-specific Fe-S precursor[27]. Fe-S availability modulates heme synthesis through the regulation of ALAS2[28], ALAD[78], and FECH[30]. Because Fe-S clusters are required for activation of FECH[30], the terminal rate-limiting step of heme synthesis, loss of ABCB7 may impact both heme and Fe-S pathways. Accordingly, combined rescue of TMEM14C and ABCB7 reduces RS by more than sixfold, but is not significantly lower than ABCB7 alone, suggesting an epistatic relationship. Rescue of TMEM14C and ABCB7 does not completely abolish RS formation, suggesting that other misspliced genes may play a role. Mutations or deletions of mitochondrial DNA have been implicated in ISAs[19], and we find that *MTERFD3*, required for mitochondrial DNA replication[79], is highly misspliced in mutant *SF3B1* cells. Taken together, our findings demonstrate that TMEM14C and ABCB7 are the major drivers of RS formation in MDS.

SF3B1 mutations are prevalent in cancers, and downstream targets remain poorly characterized. Because missplicing often has modest effects on protein levels or function, oncogenic phenotypes are likely driven by cooperative effects of multiple misspliced target genes. We show that RS formation is the result of combinatorial missplicing of *TMEM14C* and *ABCB7* that perturbs both heme and Fe-S pathways. Our study thus highlights the paradigm that combined missplicing of multiple cooperating genes leads to cancer-associated phenotypes.

2.5 Figures and legends

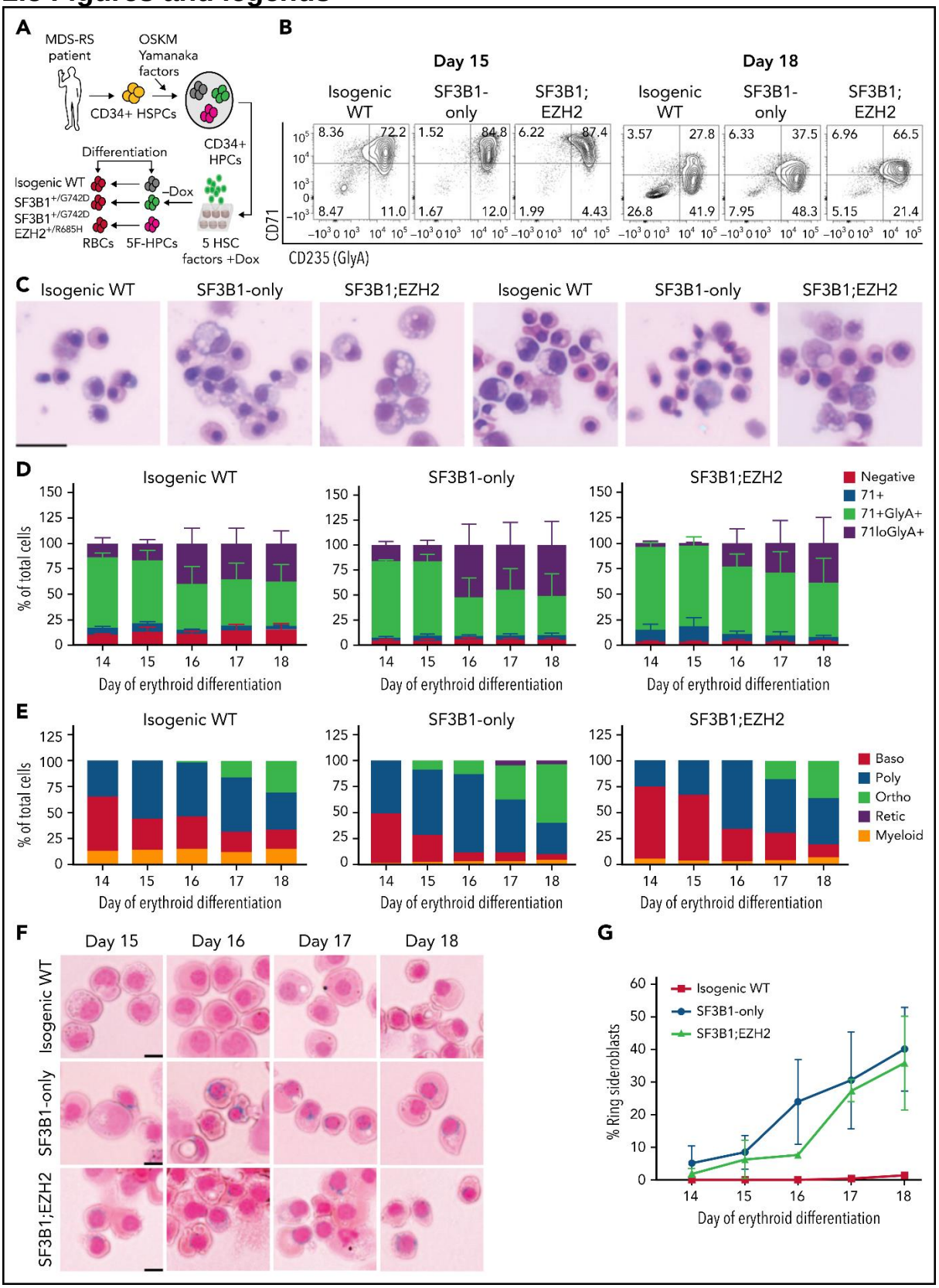

Figure 2.5.1 iPSC model of *SF3B1*-mutant MDS with RS

A. Schematic of iPSC reprogramming to generate an *SF3B1*-mutant MDS-RS model. MDS-RS patient $CD34^+$ cells were reprogrammed with episomal factors and 3 iPSC lines were selected: normal isogenic WT, $SF3B1^{G742D/+}$, and $SF3B1^{G742D/+}$;$EZH2^{/R685H/+}$. iPSCs were differentiated into $CD34^+$ HPCs and transduced with 5 transcription factors (ERG, HOXA9, RORA, SOX4, MYB) to establish doxycycline-expandable 5F-HPCs, which serve as progenitor lines for erythroid differentiation studies. **B.** Representative flow plots of CD71 and CD235 (Glycophorin A) expression in isogenic WT, $SF3B1^{G742D/+}$, and $SF3B1^{G742D/+}$;$EZH2^{R685H/+}$ cells on days 15 and 18 of erythroid differentiation. **C.** Representative May-Grunwald Giemsa staining of erythroid cell morphology on days 15 (left 3 images) and 18 (right 3 images) of differentiation. Scale bar, 25 μm. **D.** Erythroid differentiation staging using CD71 and GlyA levels measured by flow cytometry using the gating strategy shown in 1B from days 14 to 18 of erythroid differentiation; mean ± standard error of the mean of n = 3. **E.** Quantification of erythroid cell morphology using May-Grunwald Giemsa staining from days 14 to 18. A representative experiment was counted; >30 cells counted for each day of differentiation. **F.** Representative images of RS iron staining in *SF3B1*-WT and *SF3B1*-mutant cells. Erythroid cells were collected on days 15 through 18 of erythroid differentiation and stained with Perls Prussian blue. Scale bar, 10 μm. **G.** Quantification of RS counts in the isogenic WT and $SF3B1^{G742D/+}$ erythroid cells; >200 cells counted; mean ± standard deviation (SD) of n = 3 independent experiments and $SF3B1^{G742D/+}$;$EZH2^{/R685H/+}$ cells; mean ± SD of n = 2; >150 cells counted. Baso, basophilic erythroblast; poly, polychromatic erythroblast; ortho, orthochromatic erythroblast; retic, reticulocyte.

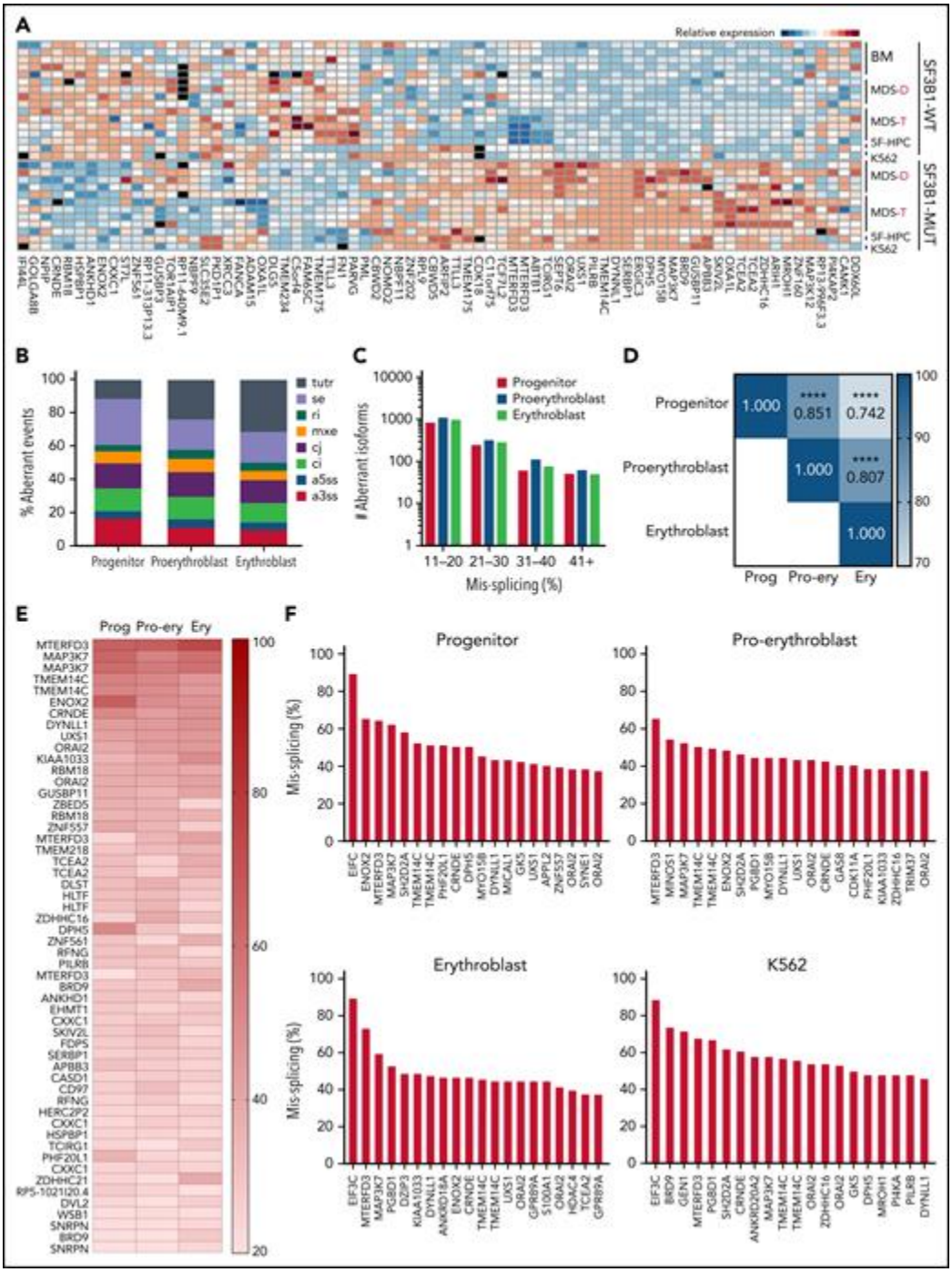

Figure 2.5.1 Mutant *SF3B1* missplicing during iPSC erythroid differentiation.

A. Relative expression of the indicated differentially spliced isoforms in *SF3B1*-WT and *SF3B1*-mutant cells. Plot restricted to competing 3′ splice sites, cassette exons, and retained introns. Normal samples include: bone marrow HPCs (BM), MDS without *SF3B1* mutations (WT), K562 cells, and normal isogenic iPSC-derived 5F-HPCs. *SF3B1*-mutant samples include: *SF3B1*-mutant MDS, K562 *SF3B1*$^{K700E/+}$ cells, and SF3B1$^{G742D/+}$ iPSC-derived 5F-HPCs. Events were restricted to ≥20% missplicing, Bayes factor ≥5, and a minimum counts cutoff ≥20. MDS-D refers to *SF3B1*$^{K700E/+}$ patient data from Dolatshad et al[43] and MDS-T refers

to *SF3B1*$^{K700E/+}$ patient data from Taylor et al.[65] **B.** The proportion of misspliced isoforms in 5F-HPCs: progenitor CD34^{+}, proerythroblast CD71^{+}GlyA^{-}, and erythroblast CD71^{+}GlyA^{+}. 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). **C.** The number of misspliced isoforms grouped by absolute value of missplicing at the 3 stages of erythroid differentiation of 5F-HPCs: progenitor CD34^{+}, proerythroblast CD71^{+}GlyA^{-}, and erythroblast CD71^{+}GlyA^{+}. **D.** Pearson correlation matrix of the level of missplicing between the 3 stages of erythroid differentiation of *SF3B1*$^{G742D/+}$ 5F-HPCs. ****$P < .0001$. **E.** The level of missplicing of individual gene isoforms during erythroid differentiation of *SF3B1*$^{G742D/+}$ 5F-HPCs; included isoforms were detected at all the stages of differentiation. **F.** Top 20 misspliced a3ss isoforms in *SF3B1*-mutant 5F-HPCs and 3 stages of erythroid cells: CD71^{+} proerythroblasts, CD71^{+}GlyA^{+} erythroblasts, and K562s. Unless otherwise stated, for all panels, differential missplicing between *SF3B1*-mutant and WT cells is defined based on the absolute value of missplicing ≥20%, Bayes factor >5, and supported by >5 transcript counts. Prog, progenitor; Pro-ery, proerythroblast; Ery, erythroblast.

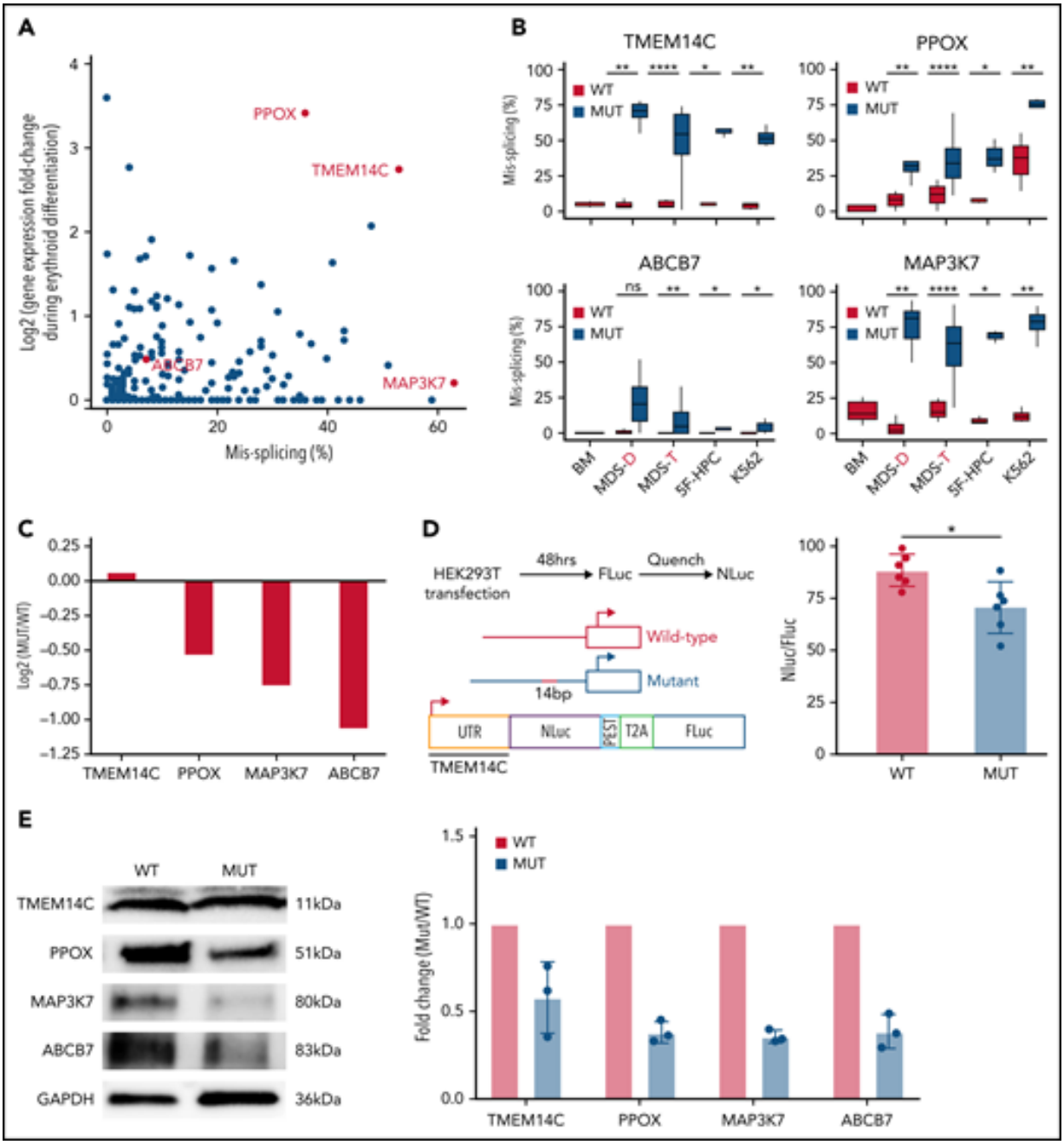

Figure 2.5.2 Missplicing downregulates expression of *SF3B1* target genes.
A. Identification of erythroid-specific mutant *SF3B1* targets based on the level of missplicing vs fold change of transcript expression during normal erythroid differentiation. Selected genes were misspliced in both *SF3B1*-mutant MDS patient cells and iPSC 5F-HPC derived erythroid cells. **B.** Total level of missplicing of *TMEM14C*, *PPOX*, *ABCB7,* and *MAP3K7* in normal bone marrow (BM), *SF3B1*-WT and mutant MDS patients, 5F-HPCs, and K562s. **P* < .5, ***P* < .01, ****P* < .001, *****P* < .0001, 1-sided Mann-Whitney *U* test. MDS-D refers to *SF3B1*$^{K700E/+}$ patient data from Dolatshad et al[43] and MDS-T refers to *SF3B1*$^{K700E/+}$ patient data from Taylor et al.[65] **C.** The change in RNA expression (log fold) by RNA-seq in *SF3B1*-mutant compared with isogenic WT

5F-HPC-derived erythroid cells. **D.** The efficiency of LUC translation in reporter assay with WT or mutant *TMEM14C* 5′UTR. Left: The schematic of WT and mutant *TMEM14C* 5′UTR and dual LUC 5′UTR reporter design. Right: The ratio of nano-LUC to firefly LUC fluorescence with WT or mutant *TMEM14C* 5′UTR. N = 6 independent experiments, **P* = .015, Student *t* test. **E.** Western blot analysis of TMEM14C, PPOX, MAP3K7, and ABCB7 protein levels at day 14 of erythroid differentiation of *SF3B1*-mutant and isogenic WT 5F-HPC derived erythroid cells. The expression was normalized to glyceraldehyde-3-phosphate dehydrogenase and shown as fold-change of *SF3B1*-mutant vs WT; 3 independent experiments, mean ± SD.

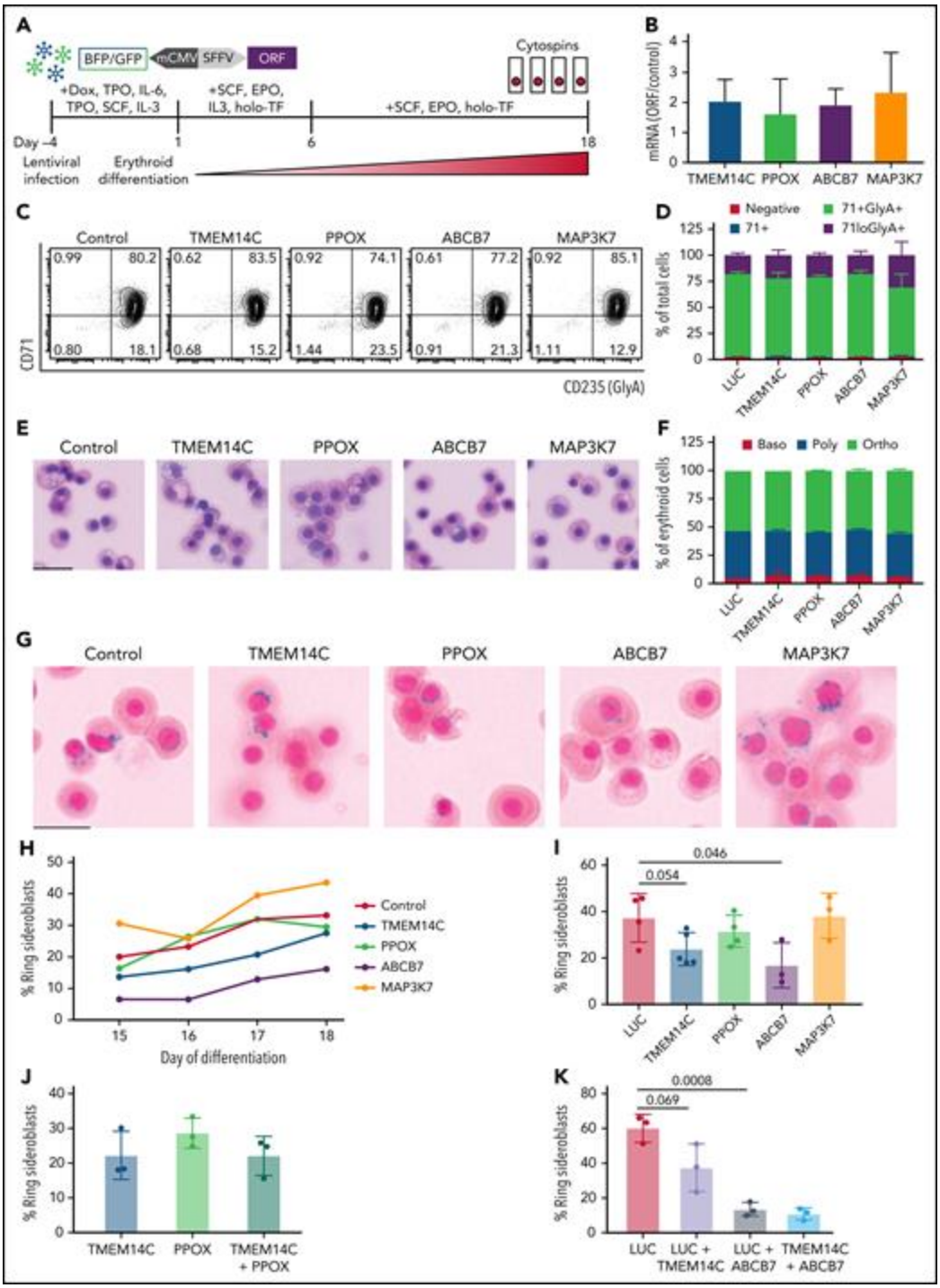

Figure 2.5.3 Rescue of TMEM14C and ABCB7 reduces RS formation.

A. Schematic of the functional RS rescue experiments. SF3B1-mutant 5F-HPCs were transduced with lentiviruses overexpressing individual ORFs or LUC control to 50-70% transduction efficiency. To induce differentiation, doxycycline was removed and cells were moved into erythroid differentiation media containing SCF, EPO, IL-3, and holo-TF (without exogenous iron) for 6 days prior to the removal of IL-3 from culture. Cytospins were collected between days 15 and 18 as indicated and stained for RS with Prussian blue. **B.** ORF mRNA

overexpression level measured by quantitative PCR and normalized to LUC control at day 15 of erythroid differentiation; mean ± SD of n = 3 independent experiments. **C.** Representative CD71 and GlyA flow plots of SF3B1-mutant cells expressing each ORF or LUC control on day 18 of erythroid differentiation. CD71 and GlyA staining is indicative of the stage of erythroid differentiation. **D.** Erythroid differentiation staging using CD71 and GlyA levels as measured by flow cytometry using gating strategy shown in 4C at day 18 of erythroid differentiation; mean ± SEM of n = 3. **E.** Representative May-Grunwald Giemsa staining of erythroid cell morphology on day 18 of erythroid differentiation. Scale bar = 25 mm. **F.** Quantification of erythroid cell morphology using May-Grunwald Giemsa staining from day 18 of erythroid differentiation; 2 independent experiments. **G.** Representative cytospin images of erythroid cells expressing each ORF or LUC control stained with Prussian blue iron staining for RS quantification. Scale bar = 25 mm. **H.** Quantification of RS counts in SF3B1-mutant cells expressing each ORF or LUC control over the course of days 15-18 of erythroid differentiation. Mean values are shown for each day, TMEM14C, ABCB7, LUC control n = 3 independent experiments; PPOX and MAP3K7 n = 2 independent experiments. **I.** Maximum RS counts between days 15-18 of differentiation for each ORF or LUC control. Each point represents an independent experiment (n = 3 - 5 independent experiments), >500 erythroid cells counted per experiment; mean± SD. P values calculated using a t test. **J.** RS counts on day 18 of erythroid differentiation for the combination of TMEM14C and PPOX, compared with TMEM14C or PPOX alone. Each point represents an independent experiment, >500 erythroid cells counted per experiment; mean± SD. **K.** RS counts on day 18 of erythroid differentiation for TMEM14C and ABCB7 combination compared to LUC combinations or LUC alone in SF3B1-mutant erythroid cells. Each point represents an independent experiment (n = 3), >500 erythroid cells counted per experiment; mean ± SD. P values calculated using a t test.

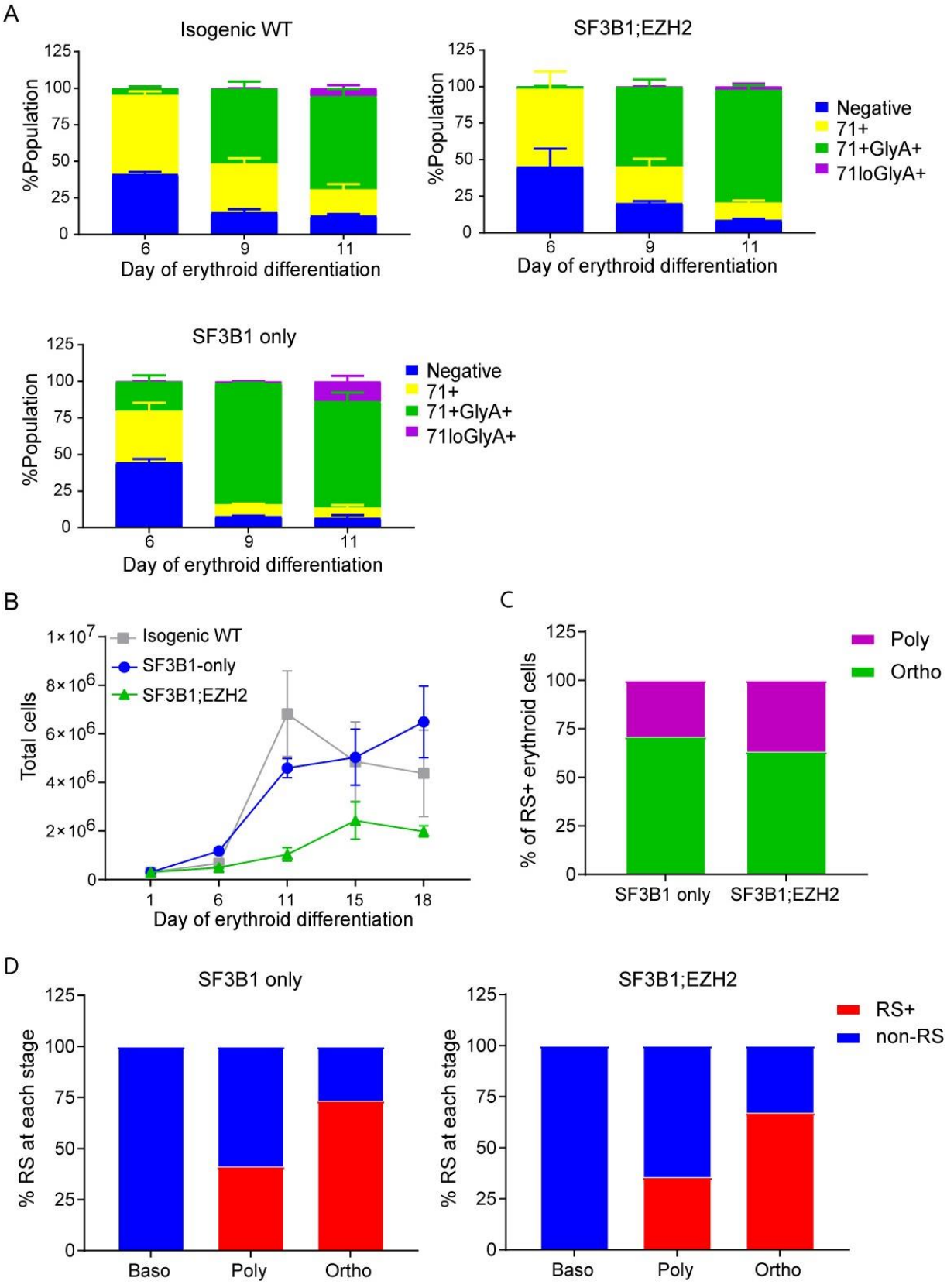
A
Isogenic WT
SF3B1;EZH2
SF3B1 only
%Population
Day of erythroid differentiation
Negative
71+
71+GlyA+
71loGlyA+
B
Total cells
Isogenic WT
SF3B1-only
SF3B1;EZH2
Day of erythroid differentiation
C
% of RS+ erythroid cells
Poly
Ortho
SF3B1 only
SF3B1;EZH2
D
% RS at each stage
RS+
non-RS
Baso
Poly
Ortho

Figure 2.5.S1 Related to Figure 2.5.1.

A. Relative CD71 and GlyA levels as measured by flow cytometry using gating strategy shown in 1B for days 6, 9, and 11 of erythroid differentiation for normal isogenic WT, *SF3B1*$^{G742D/+}$, and *SF3B1*$^{G742D/+}$;*EZH2*$^{R685H/+}$ lines; mean ± SEM of n = 3. **B.** Total erythroid cell number over 18 days of erythroid differentiation of 5F-HPC derived erythroid cells. **C.** Quantification of ring sideroblast erythroid cell morphology. Erythroid morphology was evaluated in RS cells stained with Perls Prussian blue on day 18 of culture. Poly = polychromatic, ortho = orthochromatic erythroblasts; a representative experiment was counted. **D.** Representative quantification of ring sideroblasts (RS+) at each stage of differentiation defined by erythroid cell morphology. Erythroid morphology and RS were evaluated in cells stained with Perls Prussian blue on day 18 of culture. Baso = basophilic, poly = polychromatic, ortho = orthochromatic erythroblasts; a representative experiment was counted.

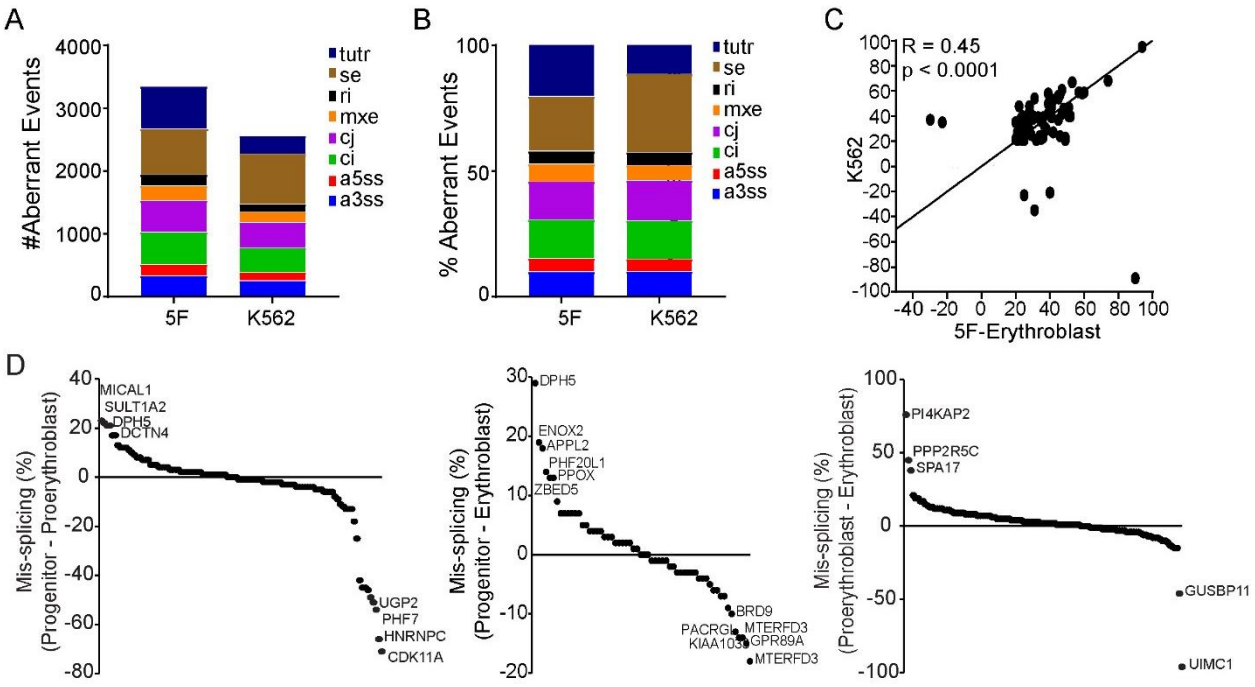

Figure 2.5.S2 Related to Figure 2.5.2.

A. Total number and **B.** proportion of differentially spliced events in all 5F SF3B1$^{G742D/+}$ cell populations compared with K562 SF3B1$^{K700E/+}$ cells (change in isoform ratio ≥ 10%). 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). **C**. Spearman correlation between the quantitative level of differential splicing in SF3B1$^{G742D/+}$ 5F erythroid cells compared with K562 SF3B1$^{K700E/+}$ cells. Line of best fit is plotted. **D.** Differential mis-splicing between 5F-HPCs and erythroid cells, and between CD71^{+} proerythroblasts and CD71^{+}GlyA^{+} erythroblasts. Isoforms are ranked by % mis-splicing in HPCs compared with erythroblasts or by %mis-splicing in proerythroblast compared with erythroblasts respectfully. Genes with mis-splicing ≥ 0.2 and total read count >5 were included in the analysis.

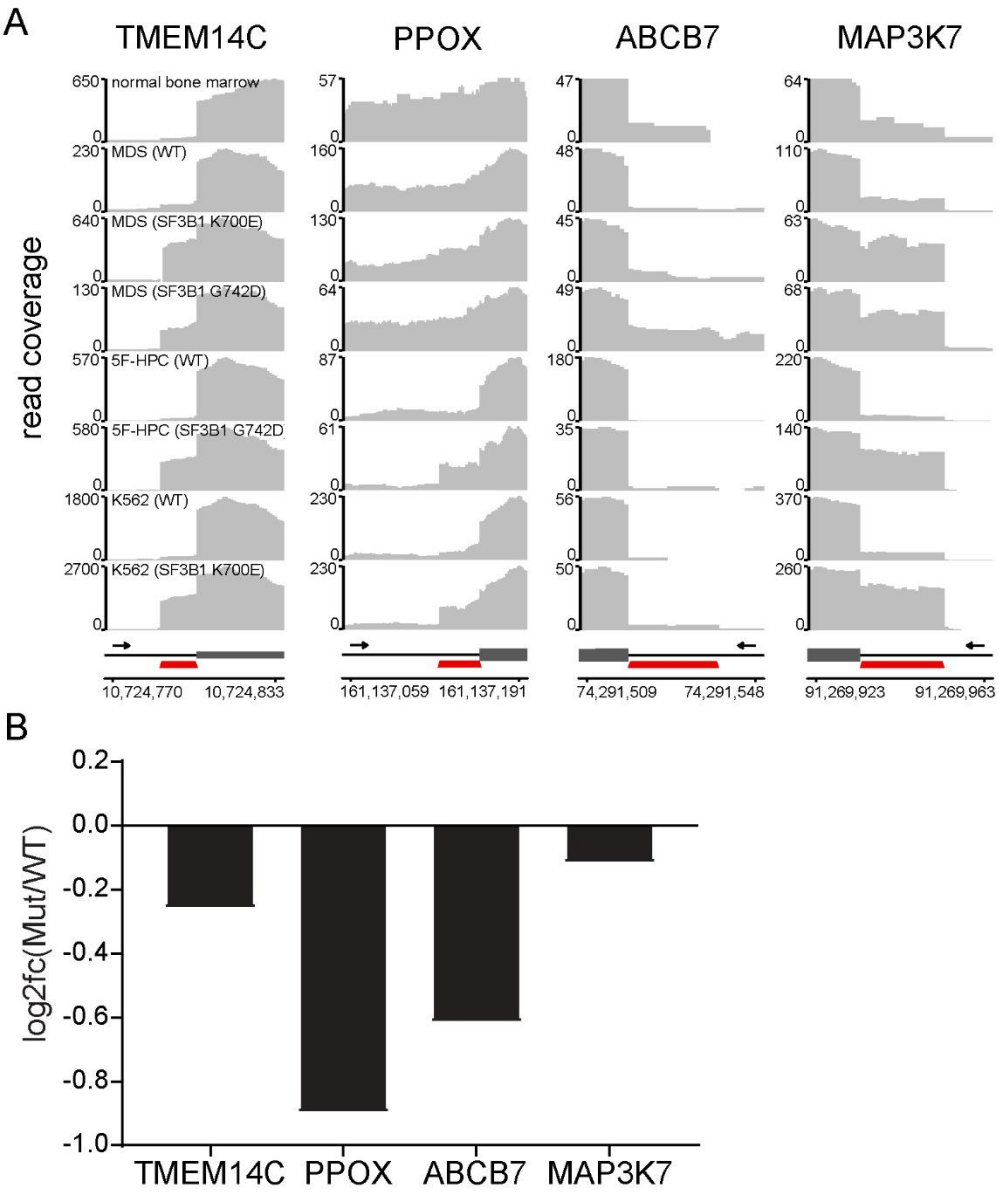

Figure 2.5.S3. Related to Figure 2.5.3.

A. Identification of erythroid-specific mutant *SF3B1* targets based on the level of mis-splicing vs. fold change of transcript expression during normal erythroid differentiation. Selected genes were mis-spliced in both *SF3B1*-mutant MDS patient cells and iPSC 5F-HPC derived erythroid cells. Arrows indicate direction of transcription. Red bar indicates location of aberrant splicing event. Genomic coordinates relative to respective chromosomes are indicated for each mis-splicing event. **B.** Log fold change of protein expression in *SF3B1*-mutant compared to *SF3B1*-wild type 5F-HPCs for the selected candidate mis-spliced genes.

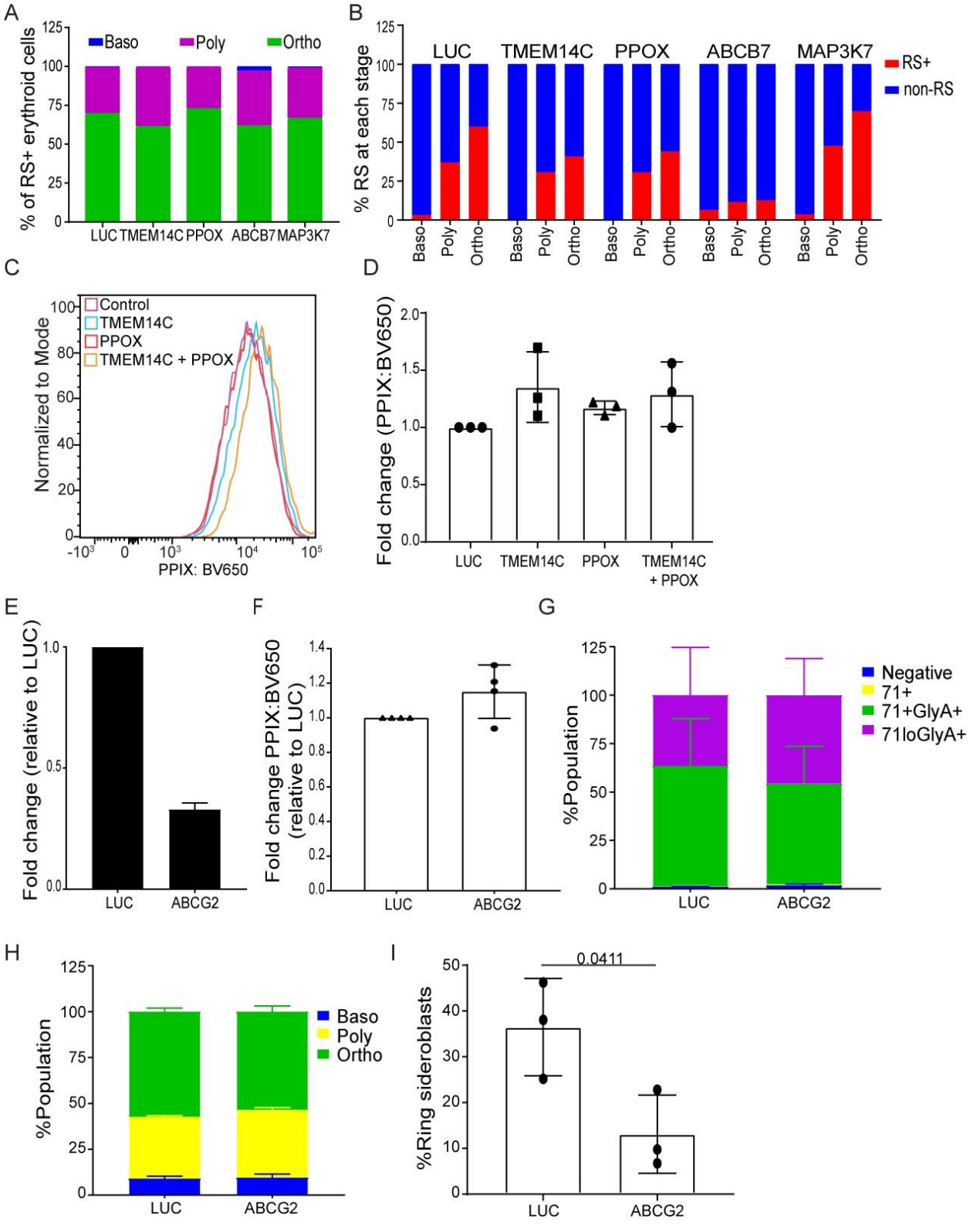
A
% of RS+ erythroid cells
Baso
Poly
Ortho
LUC TMEM14C PPOX ABCB7 MAP3K7
B
% RS at each stage
LUC
TMEM14C
PPOX
ABCB7
MAP3K7
RS+
non-RS
Baso
Poly
Ortho
C
Normalized to Mode
Control
TMEM14C
PPOX
TMEM14C + PPOX
PPIX: BV650
D
Fold change (PPIX:BV650)
LUC
TMEM14C
PPOX
TMEM14C + PPOX
E
Fold change (relative to LUC)
LUC
ABCG2
F
Fold change PPIX:BV650 (relative to LUC)
LUC
ABCG2
G
%Population
Negative
71+
71+GlyA+
71loGlyA+
LUC
ABCG2
H
%Population
Baso
Poly
Ortho
LUC
ABCG2
I
0.0411
%Ring sideroblasts
LUC
ABCG2

Figure 2.5.S4. Related to Figure 2.5.4

A. Quantification of ring sideroblast erythroid cell morphology. Erythroid morphology was evaluated in RS cells stained with Perls Prussian blue on day 18 of culture. Baso = basophilic, poly = polychromatic, ortho = orthochromatic erythroblasts; a representative experiment was counted. **B.** Representative quantification of ring sideroblasts (RS+) at each stage of differentiation defined by erythroid cell morphology. Erythroid morphology and RS were evaluated in cells stained with Perls Prussian blue on day 18 of culture. Baso = basophilic, Poly = polychromatic, Ortho = orthochromatic. **C**. Representative flow cytometry plot of PPIX fluorescence measured in the BV650 channel (405 nm excitation, 670/30 nm emission). **D**. Fold-change of PPIX levels in *SF3B1*-mutant cells transduced with TMEM14C and/or PPOX ORFs normalized to *SF3B1*-mutant cells transduced with LUC control; mean ± s.d. n = 3 independent experiments. **E**. Expression of *ABCG2* as measured by quantitative PCR relative in cells transduced with shRNA targeting luciferase control or *ABCG2.* Expression was normalized to LUC shRNA, mean ± s.d., n = 3 independent experiments. **F.** Fold-change of PPIX levels in *SF3B1*-mutant cells transduced with shRNA targeting LUC control or *ABCG2.* Each point represents an independent experiment, mean ± s.d. **G.** Relative CD71 and GlyA levels as measured by flow cytometry at day 18 of erythroid differentiation for *SF3B1*-mutant cells transduced with shRNA targeting LUC or *ABCG2*, mean ± SEM of n = 3. **H.** Quantification of May-Grunwald Giemsa staining of erythroid cell morphology from day 18 of erythroid differentiation; n = 2. **I.** RS counts on day 18 of erythroid differentiation in *SF3B1*-mutant erythroid cells transduced with shRNA constructs targeting luciferase (LUC) control or *ABCG2,* mean ± s.d. of n = 3 independent experiments. P-value indicated, t-test.

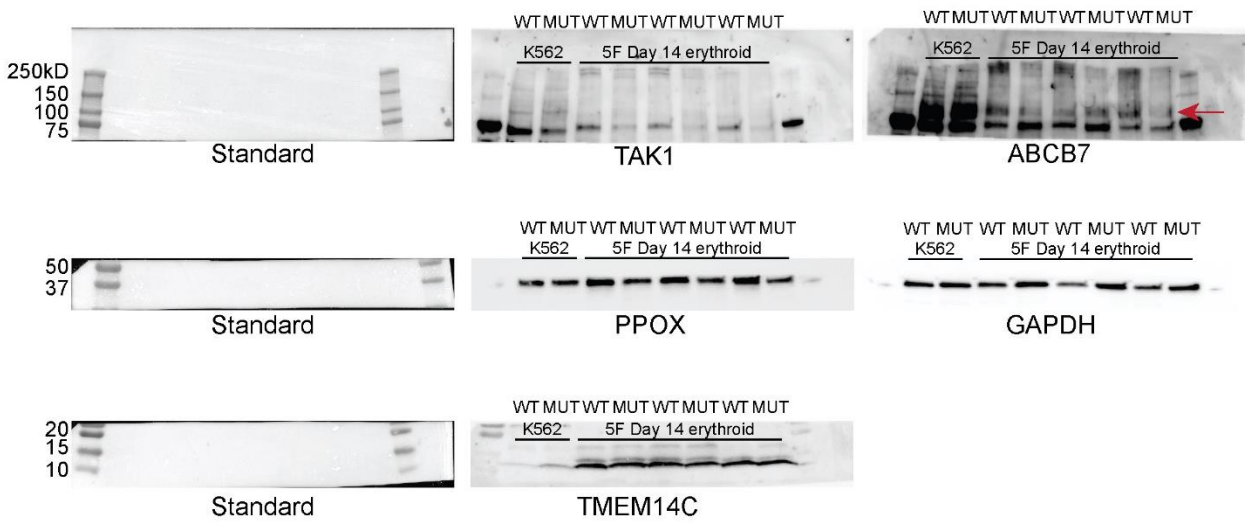

Figure 2.5.S5 Related to Figure 2.5.3E

A. Uncropped western blot images from **Figure 2.5.3E**

2.6 Methods

iPSC reprogramming: Bone marrow mononuclear cells were obtained from an MDS-RS patient according to the institutional guidelines approved by the University of Washington and Fred Hutchinson Cancer Research Center Institutional Review Board. Written informed consent was received per the Declaration of Helsinki. Clinical sequencing was performed using a TruSeq Custom Amplicon panel, clinically validated for detection of mutations in genes commonly mutated in myeloid neoplasms. iPSC reprogramming, hematopoietic differentiation, and 5F-HPC generation was performed as previously described[48]. iPSC colonies were genotyped for patient mutations by polymerase chain reaction (PCR) and Sanger sequencing**.**

5F-HPC culture: 5F-HPCs cells were cultured in StemSpan Serum-Free Expansion Medium (StemCell Technologies) with 50 ng/mL stem cell factor (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 × 106 cells/mL, and media was exchanged every 3 to 4 days. 5F-HPC lines were maintained in culture between 60 to 100 days prior to induction of erythroid differentiation.

Erythroid differentiation: Differentiation protocol was adapted based on Lee et al.[80] 5F-HPC cells expanded in progenitor media with doxycycline were placed directly into erythroid stage 1 media without doxycycline or exogenous iron. Stage 1 (6 days): Iscove modified Dulbecco medium + 1% bovine serum albumin (Gibco), 20% fetal bovine serum (FBS) (Sigma), 1 mM l-glutamine, penicillin/streptomycin, 500 μg/mL holo-transferrin (Sigma), 10 μg/mL human insulin (CellSciences), 6U Epo (Procrit), 100 ng/mL SCF, and 5 ng/mL IL-3. Cells were seeded at a density of 2 to 3 × 105 cells/mL in 24-well plates. Stage 2 (up to 12 days): Media were changed to Iscove modified Dulbecco medium + 1% bovine serum albumin, 20% FBS, 1 mM l-glutamine, penicillin/streptomycin, 500 μg/mL holo-transferrin, 10 μg/mL human insulin, 6U Epo, and 50 ng/mL SCF. Cells were seeded at a density of 3 × 105 cells/mL in 24-well plates as cell number

permitted. During stage 2, cells were maintained at a density of <2 × 106/mL in a maximum volume of 2.5 mL of media in a 24-well plate. For larger numbers of cells, plates and dish sizes were scaled up accordingly to maintain the 3 × 105 cells/mL density. Erythroid differentiation was carried out until day 18 (6 days stage 1 + 12 days stage 2), and differentiation efficiency evaluated by flow cytometry and cell morphology as described below prior to proceeding with RS staining.

Flow cytometry: Erythroid differentiation was analyzed at the end of stage 1 or during stage 2 using antibodies CD71 APC-H7 or APC (M-A712; BD) and CD235a/Glycophorin A PE-Cy7 (11E4B-7-6; Coulter). Stain was performed with < 1 × 105 cells per 50 µl staining buffer (phosphate-buffered saline + 2% FBS) with 1:100 dilution of each antibody for 30 minutes at 37°C in the dark. Acquisition was performed on BD LSRII flow cytometer with the optical configuration installed here: (http://www.pathology.washington.edu/research/flow/images/LSR%20Detector%20Array.pdf).

May-Grunwald Giemsa staining: 5×10^4 to 1×10^5 were collected onto microscope slides with a cytocentrifuge (Shandon/Thermo Scientific) and air dried. Slides were stained in May-Grunwald solution (Sigma) and then stained in 1:33 dilution of Giemsa solution (Sigma) for 12 minutes. Slides were washed twice in water and air dried prior to Vectamount (Vector Laboratories) and coverslip. Erythroid morphology was quantified as basophilic, polychromatic, orthochromatic erythroblasts, or reticulocytes.

RS iron staining: RS staining was performed starting on day 14, with the highest efficiency of RS formation observed on days 17 and 18 of erythroid culture. For RS staining, 5×10^4 to 1×10^5 cells were collected onto microscope slides with a cytocentrifuge (Thermo Scientific) and air dried. Cells were fixed in −20°C methanol for 12 minutes and air dried prior to iron staining. Slides were stained with Prussian Blue kit (Abcam, ab150674). Solutions were prepared according to the manufacturer's protocol. Slides were incubated in water for 30 seconds prior to a 12-minute incubation in iron stain followed by a 30-second water wash. Cells were counter-

stained in nuclear fast red for 5 minutes and rinsed for 4 × 30-second washes in water. Slides were preserved with Vectamount (Vector laboratories) and coverslipped prior to image acquisition.

Microscopy analysis: All images were collected with an EVOS M5000 microscope (ThermoFisher Scientific). Cells were imaged at x60 objective with a brightfield phase ring with red, green, and blue combined excitation. Brightness, saturation, and contrast were adjusted to ensure iron staining was visible for RS analysis. Images were collected by tiling across slides, and RS were manually counted. A RS was defined as an erythroid cell with iron-laden mitochondria covering greater than one-third of the nuclear perimeter. Greater than 300 cells were counted for RS analysis.

RNA-sequencing: $CD34^+$ progenitors were isolated from 5F-HPCs using the CD34 enrichment kit (Miltenyi). $CD34^-CD71^+GlyA^-$ proerythroblasts were isolated by flow sorting on day 4 after doxycycline removal. $CD71^+GlyA^+$ erythroblasts were isolated by flow sorting on day 9 after doxycycline removal. RNA was extracted from 3×10^5 to 5×10^5 cells of each population using the TRIzol reagent. Five hundred nanograms of total RNA was used as input to make poly(A)-selected RNA-sequencing (RNA-seq) libraries with the TruSeq RNA library prep kit v2 (Illumina). Libraries with DNA fragments of ~300 bp were selected using AMPure XP bead purification. Purified libraries were sequenced on an Illumina Hi-Seq 2000 using paired-end, 50 bp reads. RNA-seq splicing analysis methods can be found in the supplement.

RNA-seq splicing analysis: Isoform expression levels for all *SF3B1*-mutant and *SF3B1*-WT lines were estimated as previously described[81]. In brief, a transcriptome annotation for the hg19/GRCh37 human genome assembly was created by merging gene annotations from Ensembl v71.1[82], the UCSC knownGene track[83], and MISO v2.0 isoform annotations[84]. RSEM v1.2.4[85] 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[86], 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 *SF3B1*-mutant versus *SF3B1*-WT samples were identified as previously described. In brief, we defined the metric deltaPSI (ΔPSI) as the isoform ratio (absolute percentage of mRNA) in *SF3B1*-mutant samples - isoform ratio in *SF3B1*-WT cells, and computed a significance statistic using Wagenmakers's Bayesian framework[87] 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 ΔPSI ≥ 10% and Bayes factor ≥ 5 or $p \leq 0.05$.

Luciferase assay: $5x10^4$ HEK293T cells were plated into a 24-well plate 16-24 hours before DNA transfection. 1µg of DNA was transfected using PEI-MAX™ (Polysciences) at a 3:1 PEI:DNA ratio. After 48-72 hours, each transfected well was split into 3-wells of a 96-well plate and luciferase was measured using the Nano-Glo® Dual-Luciferase® Reporter Assay (Promega) per manufacturer's instructions. Luminescence was quantified using an Infinite M200 plate reader (TECAN), triplicate values were averaged prior to analysis.

Western Blot analysis: Protein lysates were prepared by resuspending 100,000 washed cells/10µl of RIPA buffer (Thermo Fisher). Lysates were resolved by 4-20% SDS-PAGE (Biorad) and immunoblotted with antibodies for TMEM14C (Custom, YenZym Antibodies LLC), PPOX (Santa Cruz Biotechnology, sc-271768), MAP3K7 (CST 5206), ABCB7 (LS-B13035), and GAPDH (Abcam, ab9485). 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.

Proteomics: Total protein was extracted from $CD34^+$ 5F-HPCs using urea lysis buffer that contained 6M Urea, 25 mM Tris (pH 8.0), 1 mM EDTA, 1 mM EGTA, protease and phosphatase inhibitors (Sigma-Aldrich). Total protein content of protein lysates from each sample was determined using BCA assay (ThermoFisher Scientific). Protein lysates (30µg each) were

processed for digestion and TMT labeled as per manufacturer's instructions (ThermoFisher Scientific). Tryptic peptides were pooled and desalted by solid-phase extraction using Oasis HLB sorbent (Waters, MA, USA). The desalted tryptic digests were fractionated by high-pH reverse phase (RP) liquid chromatography and concatenated into 24 samples. Fractionated samples were analyzed by LC-MS/MS on an Easy-nLC 1200 (Thermo Scientific) coupled to an Orbitrap Eclipse mass spectrometer (Thermo Scientific) operated in positive ion mode. Raw MS/MS spectra were searched using proteome discoverer 2.1 against the reviewed Human Universal Protein Resource (UniProt) sequence database. The search was performed with tryptic enzyme constraint set for up to two missed cleavages, oxidized methionine set as a variable modification, and carbamidomethylated cysteine set as a static modification. The precursor ion tolerance was set to 10 ppm and the fragment ion tolerance was set to 0.6 Da. Peptide identifications were filtered to a peptide FDR of 1%. Data normalization was performed by scaling each TMT channel to the channel median. Differential expression analysis was performed with R package. For these analyses we choose to keep features that had values in all TMT channels leaving us with 6782 quantified proteins. t-statistics were used on log2 scale to assess differences in expression/abundance.

Lentiviral vectors: Open reading frames (ORFs) for TMEM14C, PPOX, ABCB7, and MAP3K7 were purchased from the Mission® TRC3 Human LentiORF Collection and subcloned into the pSMAL overexpression vector (Addgene, Plasmid #161785). A luciferase ORF was synthesized and subcloned into pSMAL as a control. The TRC cloning vector pLKO.1 was used to generate the ABCG2 shRNA vectors. shRNA sequences were selected from the GPP web portal (Broad Institute). A shRNA targeting luciferase was used as a control. Sequences are as follows:

ABCG2-1:5'

CCGGCCTTCTTCGTTATGATGTTTACTCGAGTAAACATCATAACGAAGAAGGTTTTTG

ABCG2-2: 5’

CCGGCCTGCCAATTTCAAATGTAATCTCGAGATTACATTTGAAATTGGCAGGTTTTTG

LUC: 5’ CCGGCTTACGCTGAGTACTTCGACTCGAGTCGAAGTACTCAGCGTAAGTTTTTG

Lentiviral preparation/5F-HPC infection: 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 HEK293T cells with calcium phosphate (Takara). Lentivirus was concentrated via ultracentrifugation at 23,000 rpm for 3 hrs and resuspended in SFEM and stored at -80°C. All viruses were titered by serial dilution on 293T cells. CD34-5F cells were transduced in a 96-well plate in standard CD34-5F media (outlined above). 2-3 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 cells were resuspended in fresh CD34-5F media.

Quantitative RT-PCR: RNA was extracted from 100-300,000 cells using TRIzol (ThermoFisher) and resuspended in water for cDNA synthesis using iScript cDNA synthesis kit (BioRad) according to manufacturer instructions. *ABCB7* primers were as previously described[57]. *ABCG2* primers were as previously described[88]. Primers for *PPOX*, *TMEM14C*, and *MAP3K7* were selected from https://pga.mgh.harvard.edu/primerbank/. Sequences are as follows:

PPOX: F – 5’ TCTGCCGTGGAGTGTTTGC R – 5’ ATGGAACGATGGGTTTGCTCA

TMEM14C: F—5’ CACTGGCTCAGTAGTGCCTTT R – 5’ AGTGGTAGAACCTCATTCCCAT

MAP3K7: F—5’ AAACCACCAAACTTACTGCTGG R – 5’ CGCGTTATCACTTCCCAAAGAA

GAPDH: F – 5’ TCCTGCACCACCAACTGCTTA R – 5’ TCTTCTGGGTGGCAGTGATGG

Statistical Analysis: Statistical analysis was performed with GraphPad Prism software. Data are shown as the mean with standard deviation unless noted. For all analyses, $P < .05$ was considered statistically significant. Investigators were not blinded to the different groups.

2.7 Acknowledgements

The authors thank Sioban B. Keel (UW Hematology) and Lorinda Soma (UW Laboratory Medicine) for assistance with RS stains and analysis. Chenwei Lin (Fred Hutch Proteomics Core), Xiaoping Wu (Pathology Flow Core), and Nathaniel Peters (Keck Imaging Center) provided technical assistance.

S.D. and R.K.B. were supported in part by the National Institutes of Health (NIH) National Heart, Blood, and Lung Institute (NHLBI) (grant R01 HL151651) and Edward P. Evans Foundation. S.D. was supported in part by the NIH New Innovator Award (DP2 HL147126), NIH NHLBI (grant R21 HL139864), Wayne D. Kuni and Joan E. Kuni Foundation, and American Society of Hematology Scholar Award. R.K.B. was supported in part by the NIH National Cancer Institute (grant R01 CA251138), NIH NHLBI (grant R01 HL128239), NIH National Institute of Diabetes and Digestive and Kidney Diseases (grant R01 DK103854), and Blood Cancer Discoveries Grant program through the Leukemia & Lymphoma Society, Mark Foundation for Cancer Research, and Paul G. Allen Frontiers Group (8023-20). R.K.B. is a Scholar of the Leukemia & Lymphoma Society (1344-18) and holds the McIlwain Family Endowed Chair in Data Science. Computational studies were supported in part by Fred Hutchinson Cancer Research Center's Scientific Computing Infrastructure (Office of Research Infrastructure Programs grant S10 OD028685). This research was supported in part by the NIH NCI (Cancer Center Support Grant P30 CA015704) and Seattle Translational Tumor Research grant. E.K. is supported by The National Research Foundation of Korea (NRF) grant funded by the Korean government (NRF-2018R1A6A1A03025810 and NRF-2020R1A2C2012878). C.A.C. is supported by the Institute for Stem Cell and Regenerative Medicine fellowship.

2.8 Authorship

Contribution: C.A.C., O.A.-W., J.L.A., R.K.B., and S.D. designed the study; C.A.C., J.P., M.S., J.O.I., K.N., R.B., J.N., P.N., and E.K. performed experiments; C.A.C., J.P., M.S., J.O.I., R.K.B., M.C.S., A.R.S., and S.D. analyzed and interpreted data; D.L.S. and J.L.A. provided clinical samples and clinical history; C.A.C., S.D., and R.K.B. wrote the manuscript.

Chapter 3. Leveraging iPSC derived erythroid cells to dissect alterations to erythroid metabolism in mutant-*SF3B1* cells

3.1 Introduction

Ineffective erythropoiesis, or the inability to produce mature red blood cells, is common in many inherited blood disorders such as thalassemias and sideroblastic anemias. These disorders result from a dysregulation of hemoglobin production and result in iron overload through increased intestinal absorption and/or excess iron from chronic blood transfusions. In these patients, organ and tissue damage likely results from an increase in reactive oxygen species (ROS) from non-transferrin bound iron[17, 89]. Interestingly, ineffective erythropoiesis is a hallmark of myelodysplastic syndrome with ring sideroblasts (MDS-RS), an acquired sideroblastic anemia associated with mutations in *SF3B1*[10], but the etiology of ineffective erythropoiesis is unclear. We[49] and others[42-44, 46, 47] have identified mutant-*SF3B1* mediated mis-splicing of genes related to inherited sideroblastic anemias but the contribution of mis-splicing to erythroid defects has yet to be studied. Thus, whether the mechanisms of ineffective erythropoiesis are conserved between inherited and acquired blood disorders is unknown.

While inherited sideroblastic anemias result in microcytic anemias due to insufficient heme synthesis and a compensatory decrease in cellular volume[19], many MDS-RS patients present with normocytic or macrocytic anemias[5] suggesting that heme is synthesized at sufficient levels. Understanding how heme synthesis is regulated in MDS-RS may help elucidate the cause of ineffective erythropoiesis in MDS-RS as abnormal heme levels have been implicated in other types of MDS[90]. How sufficient or even excess levels of heme is synthesized in mutant-*SF3B1* MDS-RS erythroid cells is particularly perplexing as mutant-*SF3B1* mis-splices several heme synthesis components including *UROD*[91], *TMEM14C*, and *PPOX*[42-44, 46, 47, 49]. Additionally, mis-splicing of *ABCB7* results in the reduction of Fe-S biogenesis[27], a necessary

co-factor for the enzymatic function of *FECH*[30], the final step of heme generation. Thus, dissecting heme synthesis in MDS-RS will be crucial to further understanding the mechanism of ineffective erythropoiesis in these patients.

A major limitation of MDS-RS erythropoiesis studies to date is the inability to capture sufficient numbers of mutant-*SF3B1* erythroblasts for functional studies. The clonal heterogeneity of patients at the time of diagnosis likely precludes the capture of clones harboring only an *SF3B1* mutation, and loss of *TET2,* the most commonly co-mutated gene[92], can result in dysfunctional erythroid progenitors independently of *SF3B1* mutations[93]. Recent advances in RNA-sequencing technologies, which combine long-read sequencing and CITE-seq have been used to simultaneously profile RNA splicing and expression, genotype, and cell surface marker expression at a single cell level[91]. Gaiti et al. applied this approach to sequence several mutant-*SF3B1* samples and identify erythroid stage specific splicing alterations and changes to gene expression profiles compared to healthy controls. While this approach may identify novel mis-splicing candidates or altered genes, functional studies are still needed to determine the contribution of these events to ineffective erythropoiesis.

Here, we begin to dissect the mechanism of altered heme metabolism in MDS-RS erythroid cells by profiling porphyrin and heme synthesis in wild-type and mutant-*SF3B1* late-stage erythroid cells differentiated from our MDS-RS patient-derived induced pluripotent stem cell progenitor cell lines[48, 49]. We identify a strong upregulation of intermediate porphyrins, PPIX, heme (FePPIX), and ZnPPIX in mutant-*SF3B1* cells. We then profile gene expression profiles of wild type and mutant-*SF3B1* erythroid cells from two timepoints of erythroid differentiation, to capture the beginning and end of heme synthesis, and identify the dysregulation of several metabolic pathways in mutant-*SF3B1* cells that may contribute to ineffective erythropoiesis.

3.2 Results

3.2.1 Heme synthesis is upregulated in mutant-*SF3B1* cells

To determine whether heme synthesis is dysregulated in mutant-*SF3B1* cells, we differentiated erythroid cells from our iPSC derived MDS-RS cell lines. ~100 million erythroid cells were collected at day 15 of erythroid differentiation to measure both cytoplasmic intermediate porphyrins and mitochondrial, PPIX and Heme (FePPIX) in wild-type and mutant-*SF3B1* cells. We observe a greater than 2-fold increase in intermediate porphyrins in mutant cells (**Fig 3.4.1A**) with a proportional increase in Uroporphyrin (C8OOH) and a decrease in Coproporphyrin (C4OOH) levels in mutant-*SF3B1* samples (**Fig 3.4.S1A)**. In concordance with an increase in upstream porphyrin levels, we also observe a significant increase in both PPIX (**Fig 3.4.1B)** and heme **(Fig 3.4.1C)** levels in mutant-*SF3B1* cells. We also observe very high levels of ZnPPIX which is generated by the addition of zinc instead of iron into PPIX in mutant cells which is indicative of excessive cellular PPIX content[94] (**Fig 3.4.1D)**. Taken together, these results show that as expected, *SF3B1* mutations result in the dysregulation of heme synthesis. This dysregulation results in excessive heme production which explains the macrocytic anemia observed in MDS-RS patients. How heme synthesis is upregulated in mutant-*SF3B1* erythroid cells even though several heme related proteins are decreased[49] remains to be determined.

3.2.2 Generation of an erythroid differentiation gene expression resource to study the role of mutant-SF3B1 in human erythropoiesis

While *SF3B1*$^{K700E/+}$ murine models recapitulate macrocytic anemia as seen in MDS-RS patients, they do not recapitulate ring sideroblast formation[58, 59]. Thus, it is likely that murine ineffective erythropoiesis may not reflect what is seen in humans due to differences in mitochondrial iron dysregulation. In addition, comparative gene expression analysis of differentiating human and murine erythroid cells, sorted on cell surface markers to distinguish stages of differentiation, reveal drastically different expression patterns between species. While

humans have strong patterns of both upregulation and downregulation of genes during late-stage differentiation, there is a global downregulation pattern of gene expression in murine cells[95]. Because of this, we sought to generate a gene expression dataset comprised of mid-stage (basophilic and polychromatic erythroid cells and late-stage (late polychromatic, orthochromatic erythroid cells). To do this, we collected erythroid cells derived from our isogenic wild-type and mutant-*SF3B1* MDS-RS iPSC lines at days 9 and 16 of differentiation (**Fig 3.4.2A).** To control for the accelerated erythroid differentiation of mutant-*SF3B1* cells[68], we used the expression level of erythroid cell surface markers CD71, the transferrin receptor, and GlyA, glycophorin A, in wild-type cells to exclude mutant cells that were more differentiated (**Fig 3.4.2B**). Bulk RNA-sequencing samples clustered by day of differentiation confirm that we are capturing similar stages of differentiation of wild-type and mutant cells (**Fig 3.4.2C**).

3.2.3 iPSC-derived erythroid cells recapitulate erythroid differentiation

To validate that our iPSC-derived erythroid cells properly recapitulate major gene expression changes during the maturation from basophilic/polychromatic (day 9) to orthochromatic (day 16) erythroid cells, we performed gene set enrichment analysis (GSEA) to look for differentially expressed genes over time. Comparison of wild-type and mutant-*SF3B1* cells reveal a high level of similarity in overall pathway enrichment. As expected, heme synthesis is the top differentially expressed pathway as the cells mature to orthochromatic cells (**Fig 3.4.3A,B)**. We observe an upregulation of heme synthesis related genes in mutant-*SF3B1* compared to wild-type cells at day 9 further suggesting that heme synthesis is upregulated in these cells (**Fig 3.4.S2A**). Gene ontology analysis was applied to upregulated genes at day 16 compared to day 9 of differentiation. We observe an expected increase in genes involved in macromolecule turnover such as autophagy[96] and protein catabolism in both wild-type and mutant *SF3B1* cells indicating that our iPSC derived erythroid cells are recapitulating erythropoiesis (**Fig 3.4.3C,D).**

3.2.4 Mutant-SF3B1 cells have altered gene expression patterns related to erythroid maturation

Interestingly, IL6/JAK-STAT3 signaling is enriched in orthochromatic mutant cells (**Fig 3.4.3B, Fig 3.4.S2B)**. Future work to explore the impact of this pathway in mutant cells is warranted as dysregulation of both IL-6 signaling[97] and JAK-STAT3 signaling[98] have been shown to prevent erythroid maturation. While chromatin reorganization within erythroid cells is thought to be fairly limited[99], since it was the second most upregulated pathway, we explored the top upregulated genes related to chromatin organization in wild-type and mutant cells. A couple of genes were strongly upregulated in both genotypes such as *HDAC5,* which is required for terminal differentiation, likely through chromatin compaction,[100] and *CBX7,* which is likely related to iPSC reprogramming[101]. Analysis of top upregulated genes during erythroid maturation revealed a skew toward erythroid promoting genes in wild-type *SF3B1* cells and a skew toward erythroid inhibition in mutant *SF3B1* cells (**Fig 3.4.3E,F)**. Wild-type cells show an increase in *IGF2* expression which has been shown to promote erythropoiesis[102] and an increase in EZH1 expression, a sign that the orthochromatic cells have undergone the expected *EZH2* to *EZH1* switch[103]. Contrastingly, mutant-*SF3B1* erythroid cells show an upregulation of *WBP5* and *PHF1*. Elevated *WBP5* expression has been linked to acute myeloid leukemia through upregulation of HOX gene expression[104]. Consistent with this, we observe a strong upregulation of *HOXA9* expression in mutant-*SF3B1* cells (**Fig 3.4.S2C)**. *HOXA9* overexpression can limit erythroid potential through an upregulation of NF-kB which is known to occur in mutant-*SF3B1* cells[105]. We also observe elevated expression of *PHF1* which has been shown to induce apoptosis in a p53-dependent mechanism in several cell lines[106]. Taken together, functional analyses are needed to determine if any of these dysregulated genes contribute to alterations to erythropoiesis. Future studies will focus on the role of genes which inhibit terminal erythroid differentiation through modulation of apoptosis or other pathways that can contribute to ineffective erythropoiesis in mutant-*SF3B1* MDS-RS.

3.2.5 Metabolism is altered in mutant-SF3B1 cells during erythroid differentiation

To further elucidate differences between wild-type and mutant-*SF3B1* erythroid cells during differentiation, we performed GSEA analysis comparing mutant and wild-type gene expression at both day 9 and day 16 to look for stage specific differences. Pathway analysis revealed alterations to metabolism with an increase in glycolysis in day 9 mutant-*SF3B1* cells (**Fig 3.4.4A)** and a decrease in fatty acid metabolism in day 16 cells (**Fig 3.4.4B, 3.4.S3A)**. Previous studies have suggested differentiating cells undergo a metabolic switch from glycolysis to OXPHOS during erythroid maturation[107] indicating that upregulation of glycolysis in mutant cells may contribute to dysregulation of erythroid maturation. Additionally, downregulation of fatty acid metabolism has been shown to limit enucleation and terminal differentiation[108]. We hypothesize that excessive heme synthesis within mutant-*SF3B1* cells drives metabolic reprogramming of phosphatidyl choline, glycolytic, and TCA pathways to generate sufficient amounts of glycine and succinyl-CoA for heme production (**Fig 3.4.4C**). It is possible that mutant-*SF3B1* results in the metabolic reprogramming of hematopoietic stem and progenitor cells independent of heme synthesis demands. These changes may predispose cells toward excessive heme synthesis as a result of a buildup of heme related metabolites.

3.2.6 Mutant-SF3B1 erythroid cells are metabolically skewed toward 3P pyruvate production to increase glycine synthesis

We hypothesize that excessive heme synthesis requires an increase in the mitochondrial glycine pool in mutant-*SF3B1* cells. To test this, we utilized our RNA-sequencing dataset to assess changes in gene expression in glycine synthesis pathways. We measured the expression of *SLC25A38*, the mitochondrial glycine importer, and observed a significant upregulation in mutant-*SF3B1* cells at day 9 of differentiation and *SLC25A38* levels remained elevated throughout differentiation (**Fig 3.4.5A**). Glycine is synthesized through the breakdown of choline or through shunting glycolysis into serine synthesis (**Fig 3.4.S4A)**. We first looked at

the contribution of choline production by measuring the levels of *PHOSPHO1,* which is necessary for choline production. Expression of *PHOSPHO1* is very low at day 9, suggesting that glycine synthesis is primarily driven through glycolytic serine synthesis for heme production. Expression of *PHOSPHO1* significantly increases by day 16 but remains lower in mutant-*SF3B1* cells (**Fig 3.4.5B**). Surprisingly, we did not observe any significant increase in any genes involved in the glycolytic conversion of glucose to pyruvate suggesting that changes in glycolysis are specifically linked to the shunt to serine and glycine production through 3P pyruvate production (**Fig 3.4.S4B,C**). As expected, we see a strong upregulation of *PHGDH* expression (**Fig 3.4.5C**) and *PSAT1* and *PSPH* which are required for serine synthesis (**Fig 3.4.5D**). We also observe upregulated expression *SHMT2* which converts serine into glycine (**Fig 3.4.5E**). Taken together, excessive heme production results in an increased demand for mitochondrial glycine causing a glycolytic shift toward serine and glycine synthesis in mutant-*SF3B1* erythroid cells.

3.2.7 Methylglyoxal generation drives pyruvate production in mutant-SF3B1 cells through D-lactate dehydrogenase for succinyl-CoA production

Accumulation of genotoxic metabolites, such as methylglyoxal, a byproduct of glycolysis, contributes to DNA damage, alterations to hematopoiesis, and progression to leukemic states[109]. Several enzymes including AKR1A1, ALDH1B1, GLO1, and GLO2 have been shown to protect against this toxicity by converting methylglyoxal to inert lactate[110]. As glycolytic demands are increased in mutant-*SF3B1* cells, we assessed the level of these protective enzymes at day 9, when GSEA showed a strong upregulation of glycolysis (**Fig 3.4.4A**). Of the four enzymes, we observe a significant upregulation of *GLO2*, the enzyme that converts S-D-lactoylglutathione to lactate indicative that methylglyoxal production is elevated in these cells (**Fig 3.4.6A**). *GLO2* expression is erythroid specific[111] and is upregulated during erythroid

maturation with peak expression observed in $CD71^{+}GlyA^{+}$ cells likely reflecting the cell stage with the highest glycolytic demands (**Fig 3.4.6B**).

Pyruvate is required for succinyl-CoA production through the TCA cycle. As succinyl-CoA, together with glycine, is required for the initial step of heme synthesis we hypothesize that mutant-*SF3B1* cells need to produce high levels of pyruvate for succinyl-CoA generation. Since the glycolytic shift to serine synthesis likely lowers standard glycolytic pyruvate production, we assessed whether increased lactate from the increase in methylglyoxal production is metabolized through lactate dehydrogenase. We measured the levels of *LDHA, LDHB,* and *LDHD* which convert lactate into pyruvate. We observe no changes in the relative gene expression levels of *LDHA* or *LDHB* between wild-type and mutant-*SF3B1* cells at either day 9 or at day 16 of differentiation and observe a strong reduction in total RNA levels as the cells mature (**Fig 3.4.S5A,B**). Strikingly, we see a drastic increase in the levels of *LDHD* in mutant-*SF3B1* cells over time with more than a 200-fold increase in total *LDHD* expression in mutant-*SF3B1* expression in day 16 erythroid cells (**Fig 3.4.6C,D**). LDHD is thought to specifically metabolize D-lactate[112], the lactate byproduct of methylglyoxal degradation, and its unique upregulation in mutant cells suggests an increased pyruvate demand that exceeds glycolytic and canonical *LDHA/LDHB* production. To maintain excessive heme production, we hypothesize that the increase in *LDHD* pyruvate production is metabolized to generate acetyl-CoA for succinyl-CoA production through the tricarboxylic acid cycle (TCA) (**Fig 3.4.6E**). Consistent with this hypothesis, we observe an increase in pyruvate dehydrogenase, *PDHA1,* with no significant alterations to the rest of the TCA enzymes involved in citrate to succinyl-CoA production (**Fig 3.4.6F, Figure 3.4.S5C**).

3.3 Discussion

Sideroblastic anemias are characterized by ineffective erythropoiesis and the presence of ring sideroblasts (RS), erythroid cells with iron-laded mitochondria[19]. The extent to which

pathophysiology is conserved between acquired sideroblastic anemias, such as MDS-RS, and inherited sideroblastic anemias is unclear. While we have shown that mutant-*SF3B1* mis-splicing results in a reduction in protein levels of heme synthesis protein, TMEM14C, and key Fe-S cluster protein, ABCB7, which collectively drive RS formation in MDS-RS[49], it is unclear whether dysregulation of these mitochondrial pathways also result in other aberrant changes that may drive ineffective erythropoiesis. The development of normocytic or macrocytic anemia in most MDS-RS patients[5] suggests that unlike inherited sideroblastic disorders which are almost exclusively microcytic[26], MDS-RS erythroid cells produce sufficient or excessive levels of heme despite a reduction in heme and Fe-S cluster biogenesis genes. Here, we leverage our iPSC derived MDS-RS progenitor cell lines to measure heme synthesis in late stage (polychromatic/orthochromatic) erythroid cells collected at day 15 of *in vitro* differentiation. Heme synthesis is aberrantly upregulated in mutant-*SF3B1* cells resulting in elevated levels of cytosolic porphyrins, mitochondrial PPIX, heme (FePPIX), and ZnPPIX. This finding explains the macrocytic anemia that is common to these patients and the high level of porphyrin and ZnPPIX accumulation suggests that heme synthesis is likely uncoupled from mitochondrial iron availability[94]. To dissect how this upregulation of heme synthesis alters mutant-*SF3B1* erythroid cells, we leveraged our MDS-RS cell lines to profile gene expression during erythroid maturation. We find a general trend toward expression of erythroid promoting genes in wild-type cells with an upregulation of erythroid inhibitory genes in mutant-*SF3B1* cells. Future functional studies will focus on the contributions of these genes to hematopoietic progenitor lineage commitment and erythroid maturation. We find that excessive heme production, which requires sufficient levels of glycine and succinyl-CoA substrates, drives a metabolic shift toward glycolytic serine synthesis. This shift generates both glycine and succinyl-CoA through the conversion of methylglyoxal into pyruvate (**Fig 3.4.7**). Future work to determine if this metabolic shift drives ineffective erythropoiesis by preventing proper metabolic reprogramming such as the upregulation of fatty acid synthesis which is required for terminal erythroid maturation[108], is

needed. Whether mutant-*SF3B1* mis-splicing contributes also alters these metabolic pathways is currently unknown as splicing analysis of erythroid cells has been limited.

While we find clear evidence of aberrant upregulation of heme synthesis in mutant-*SF3B1* cells, but what drives excessive heme levels is unknown. Heme has been shown to regulate the erythroid transcription factor, *GATA1*[113], suggesting that altered heme levels might dysregulate erythroid differentiation and contribute to ineffective erythropoiesis. Recent work in a K562 SF3B1$^{K700E/+}$ cell line identifies mis-splicing of *MAP3K7,* which results in the dysregulation of the MAP3K7-p38 MAPK axis and prematurely downregulates *GATA1* resulting in erythroid apoptosis[68]. Further exploration to dissect how mutant-*SF3B1* results in excessive heme production and how this cooperates with loss of mutant-*SF3B1* mediated mis-splicing of *MAP3K7* to drive ineffective erythropoiesis is needed.

One major limitation of this work is the lack of metabolomic data to validate that gene expression changes result in alterations of metabolite levels. Follow-up studies will focus on profiling additional stages of erythroid differentiation to better characterize the metabolic shift that we observe and determine whether it occurs independently of heme synthesis. While our findings align with previous reports that mutant-*SF3B1* cells are sensitive to serine and glycine deprivation, prior studies focused on mutant-*SF3B1* metabolism in breast epithelial cells, which do not have the same heme synthesis demands[114]. Whether metabolic sensitivity is driven by heme or if other pathways converge on serine and glycine synthesis in other mutant-*SF3B1* cancers is unknown. Taken together, we describe excessive upregulation of heme synthesis in and metabolic reprogramming to favor glycine production in mutant-*SF3B1* erythroid cells. Compounds that modulate the glycolytic shunt to serine synthesis should be explored as potential ineffective erythropoiesis therapies in MDS-RS.

3.4 Figure and Legends

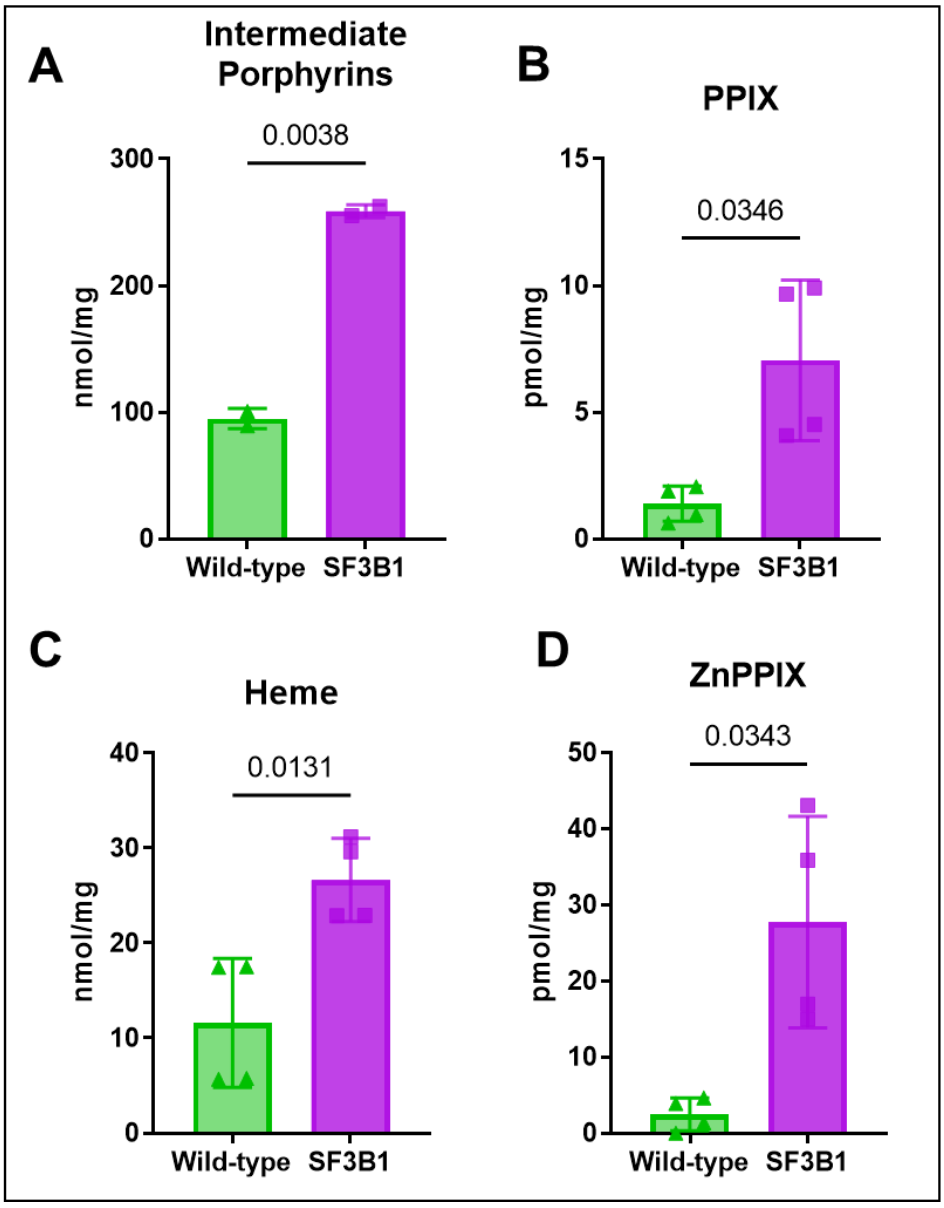

Figure 3.4.1. Heme synthesis is upregulated in mutant-*SF3B1* erythroid cells
A. Sum of cytosolic porphyrins uroporphyrin, heptacarboxylporphyrin, hexacarboxylporphyrin, pentacarboxylporphyrin, and corproporphyrin in wild-type and mutant-*SF3B1* erythroid cells collected at day 15 of erythroid differentiation. N = 2, Welch's corrected t-test **B.** Porphyrin IX (PPIX) concentration in wild-type and mutant-*SF3B1* erythroid cells collected at day 15 of erythroid differentiation. N = 4, Welch's corrected t-test **C.** Heme (FePPIX) concentration in wild-type and mutant-*SF3B1* erythroid cells collected at day 15 of erythroid differentiation. N = 4, Welch's corrected t-test D. Total ZnPPIX concentration in wild-type and mutant-*SF3B1* erythroid cells collected at day 15 of erythroid differentiation. N = 4, Welch's corrected t-test

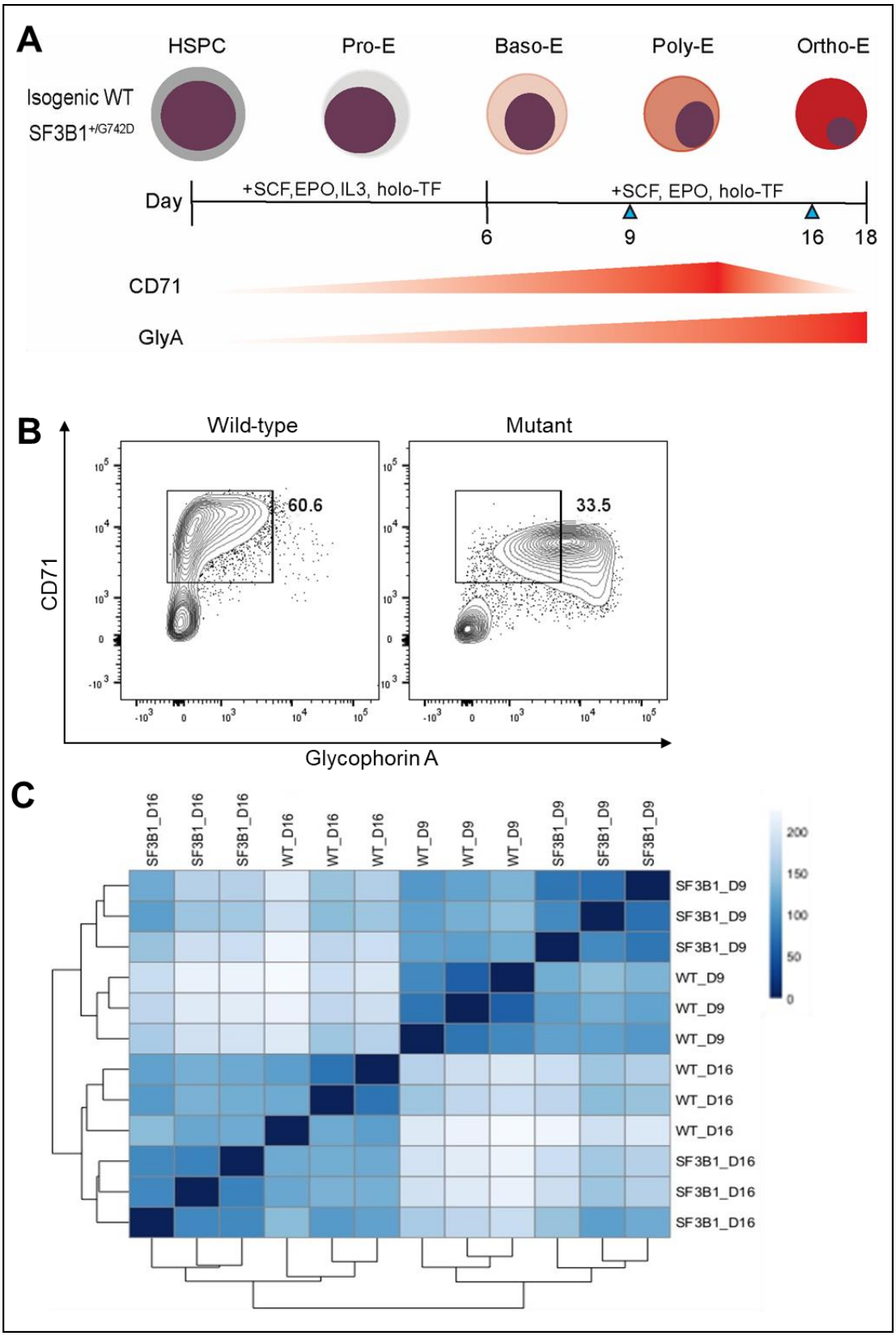
A
HSPC
Pro-E
Baso-E
Poly-E
Ortho-E
Isogenic WT
SF3B1+/G742D
Day
+SCF,EPO,IL3, holo-TF
+SCF, EPO, holo-TF
6
9
16
18
CD71
GlyA
B
Wild-type
Mutant
60.6
33.5
CD71
Glycophorin A
C
SF3B1_D16
WT_D16
WT_D9
SF3B1_D9
200
150
100
50
0

Figure 3.4.2. Generation of an erythroid RNA sequencing dataset to study maturation of wild-type and mutant-*SF3B1* cells

A. Erythroid differentiation schematic. Isogenic wild-type and SF3B1$^{G742D/+}$ progenitor cells were differentiated in a two-stage culture medium. Triangles indicate day of differentiation in which cells were collected for RNA-sequencing. HSPC: Hematopoietic stem and progenitor cell. Pro-E: Proerythroblast. Baso-E: Basophilic erythroblast. Poly-E: Polychromatic erythroblast. Ortho-E: Orthochromatic erythroblast. Expression patterns of surface markers CD71, transferrin receptor, and GlyA, glycophorin A, are indicated. **B.** Fluorescence-activated Cell Sorting strategy to collect wild-type and mutant-*SF3B1* cells with similar levels of CD71+ and GlyA expression. Representative flow plots from day 9 and 16 are plotted. Levels of CD71 and GlyA on wild-type cells were used to draw sorting gates **C.** Distance matrix heatmap of normalized RNA counts from isogenic wild-type and mutant-SF3B1 lines from days 9 and 16 of erythroid differentiation. Each day and line had 3 independent replicates. WT = wild type, SF3B1 = SF3B1$^{G742D/+}$

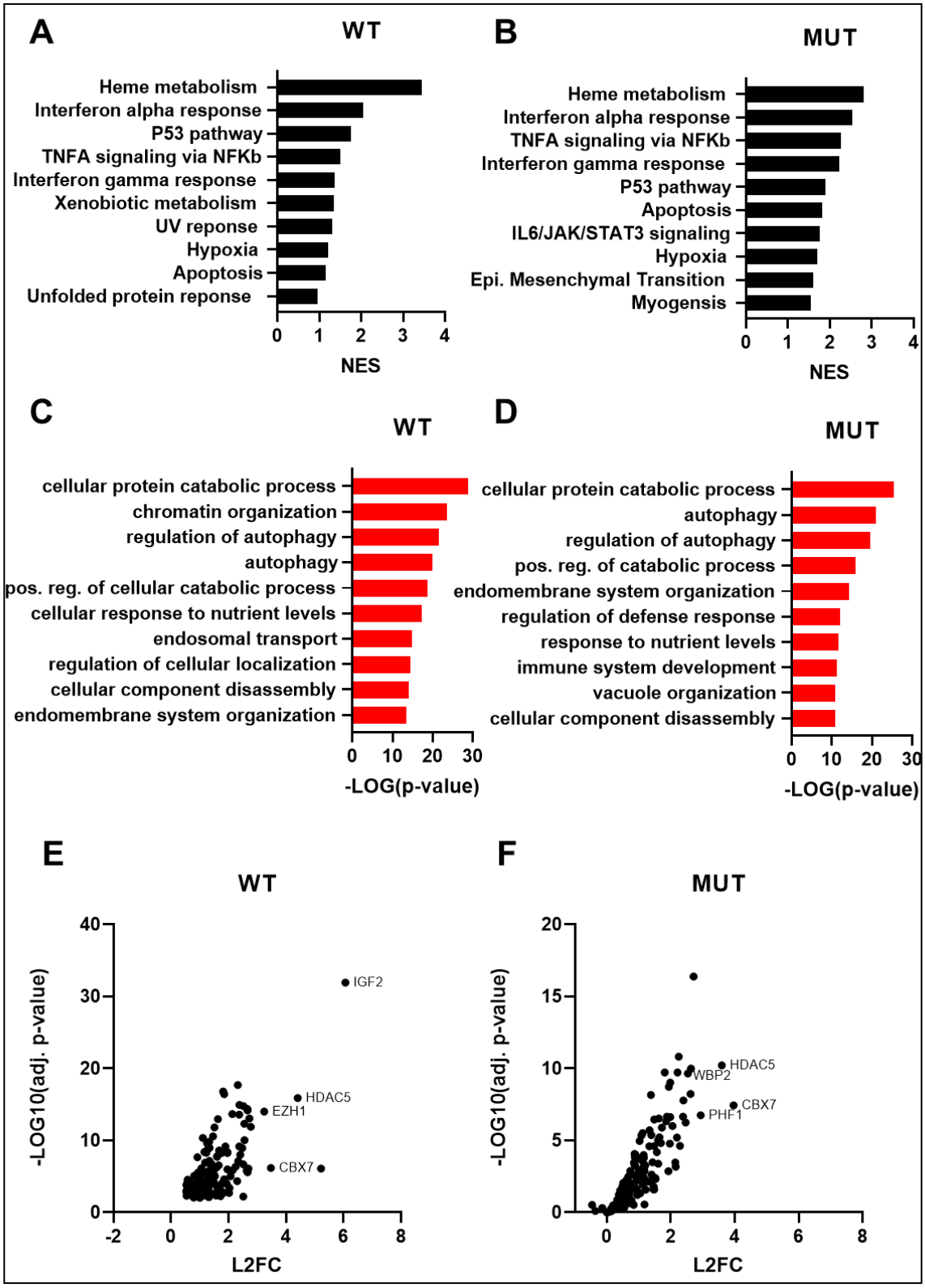
A
WT
Heme metabolism
Interferon alpha response
P53 pathway
TNFA signaling via NFKb
Interferon gamma response
Xenobiotic metabolism
UV reponse
Hypoxia
Apoptosis
Unfolded protein reponse
0 1 2 3 4
NES
B
MUT
Heme metabolism
Interferon alpha response
TNFA signaling via NFKb
Interferon gamma response
P53 pathway
Apoptosis
IL6/JAK/STAT3 signaling
Hypoxia
Epi. Mesenchymal Transition
Myogensis
0 1 2 3 4
NES
C
WT
cellular protein catabolic process
chromatin organization
regulation of autophagy
autophagy
pos. reg. of cellular catabolic process
cellular response to nutrient levels
endosomal transport
regulation of cellular localization
cellular component disassembly
endomembrane system organization
0 10 20 30
-LOG(p-value)
D
MUT
cellular protein catabolic process
autophagy
regulation of autophagy
pos. reg. of catabolic process
endomembrane system organization
regulation of defense response
response to nutrient levels
immune system development
vacuole organization
cellular component disassembly
0 10 20 30
-LOG(p-value)
E
WT
-LOG10(adj. p-value)
0 10 20 30 40
IGF2
HDAC5
EZH1
CBX7
-2 0 2 4 6 8
L2FC
F
MUT
-LOG10(adj. p-value)
0 5 10 15 20
HDAC5
WBP2
CBX7
PHF1
0 2 4 6 8
L2FC

Figure 3.4.3. Erythroid differentiation of iPSC derived progenitors reveals altered gene expression patterns related to maturation

A,B. Top upregulated pathways from gene set enrichment analysis (GSEA) (**Liberzon**) in day 16 erythroid cells compared to day 9 erythroid cells as ranked by normalized enrichment score value in **A.** wild type and **B.** mutant-*SF3B1* cells. **C,D.** Gene ontology (GO) analysis (**Zhou 2019**) of upregulated genes in day 16 versus day 9 comparisons. Upregulation was characterized as a log2 fold change ≥ 0.5 and an adjusted p-value < 0.01 as calculated using DESEQ2 (**Love 2014**). Upregulated genes in **C.** wild-type erythroid cells. N = 2731 genes. **D.** mutant-*SF3B1* erythroid cells. N = 1905 genes. **E,F.** Log2 fold-change and adjusted p-value of genes taken from the GO term, chromatin organization. Genes were analyzed with DESEQ2 to determine log2 fold-change and adjusted p-value. Expression of day 16 was compared to day 9 **E.** wild-type erythroid cells. **F.** mutant- *SF3B1* erythroid cells.

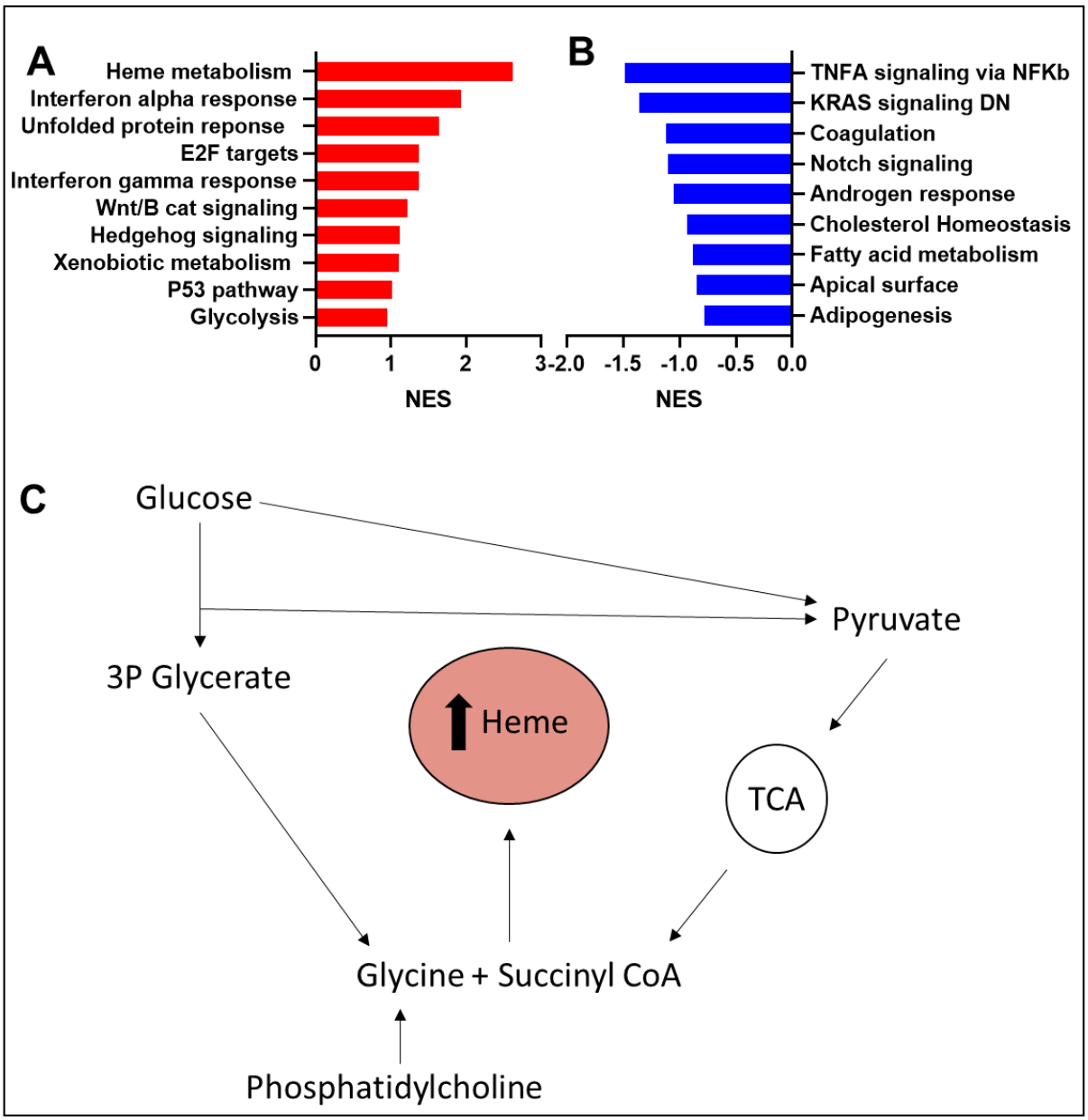

Figure 3.4.4. Gene expression analysis reveals metabolic alterations in mutant-*SF3B1* erythroid cells

A. Top upregulated pathways at day 9 of erythroid differentiation in mutant-*SF3B1* cells compared to wild-type controls ranked by GSEA calculated normalized enrichment score. **B.** top downregulated pathways at day 16 of erythroid differentiation in mutant-*SF3B1* cells compared to wild-type controls ranked by GSEA calculated normalized enrichment score. **C.** Basic schematic of the convergence of metabolic pathways glycolysis, the tricarboxylic acid (TCA) cycle, and phosphatidylcholine metabolism to generate glycine and succinyl-CoA for heme synthesis

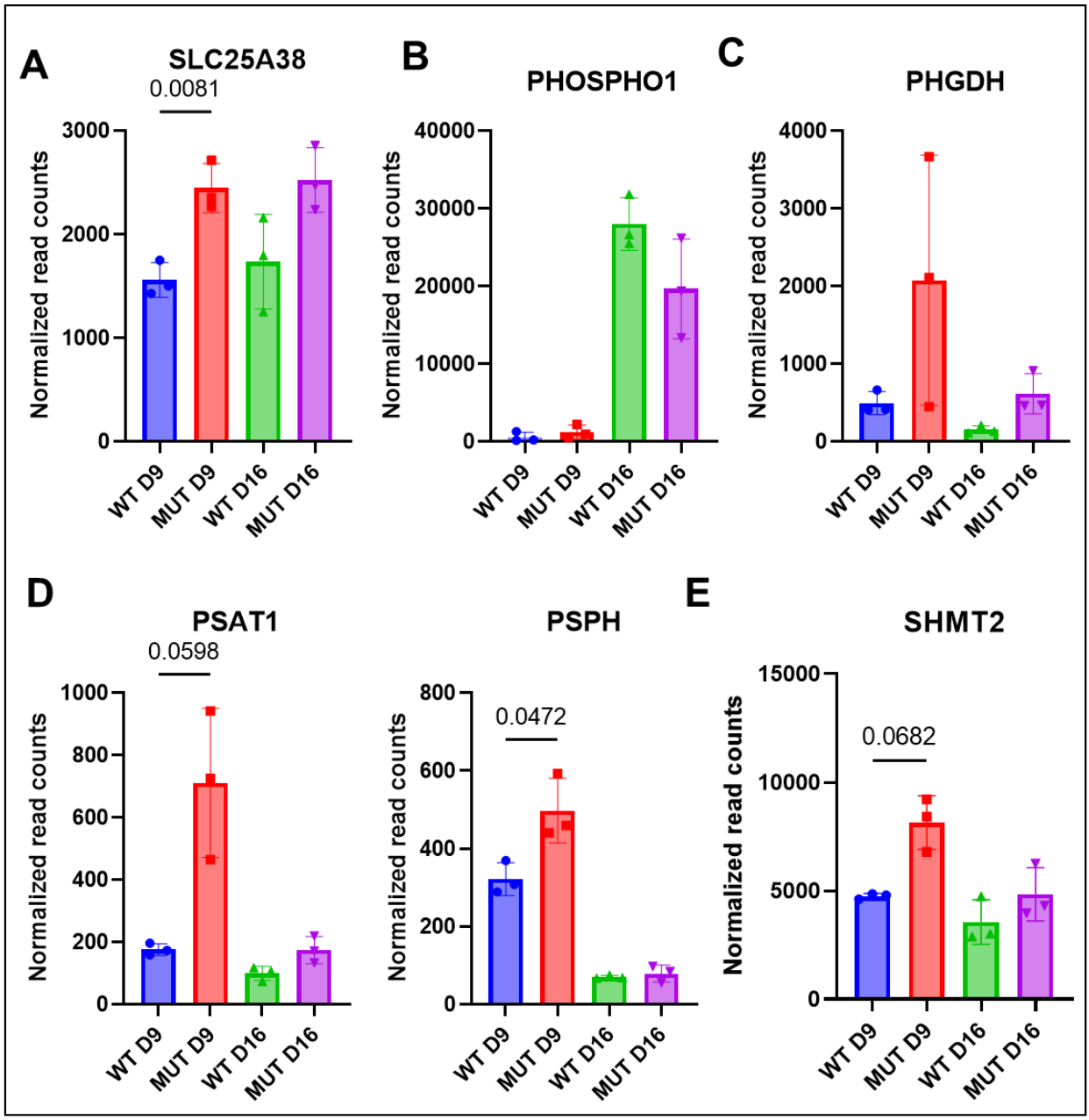

Figure 3.4.5. Mutant-*SF3B1* erythroid cells are metabolically skewed toward 3P pyruvate production to increase glycine synthesis

Normalized RNA-sequencing read counts of wild-type and mutant-*SF3B1* erythroid cells at days 9 and 16 of differentiation. N=3, Welch's corrected t-test. **A.** Mitochondrial glycine importer, *SLC25A38*. **B.** Phosphoethanolamine/Phosphocholine Phosphatase 1, *PHOSPHO1* **C.** Phosphoglycerate Dehydrogenase, *PHDGH* **D.** Phosphoserine Aminotransferase 1, *PSAT1,* Phosphoserine Phosphatase, *PSPH* **E.** Serine Hydroxymethyltransferase 2, *SHMT2*

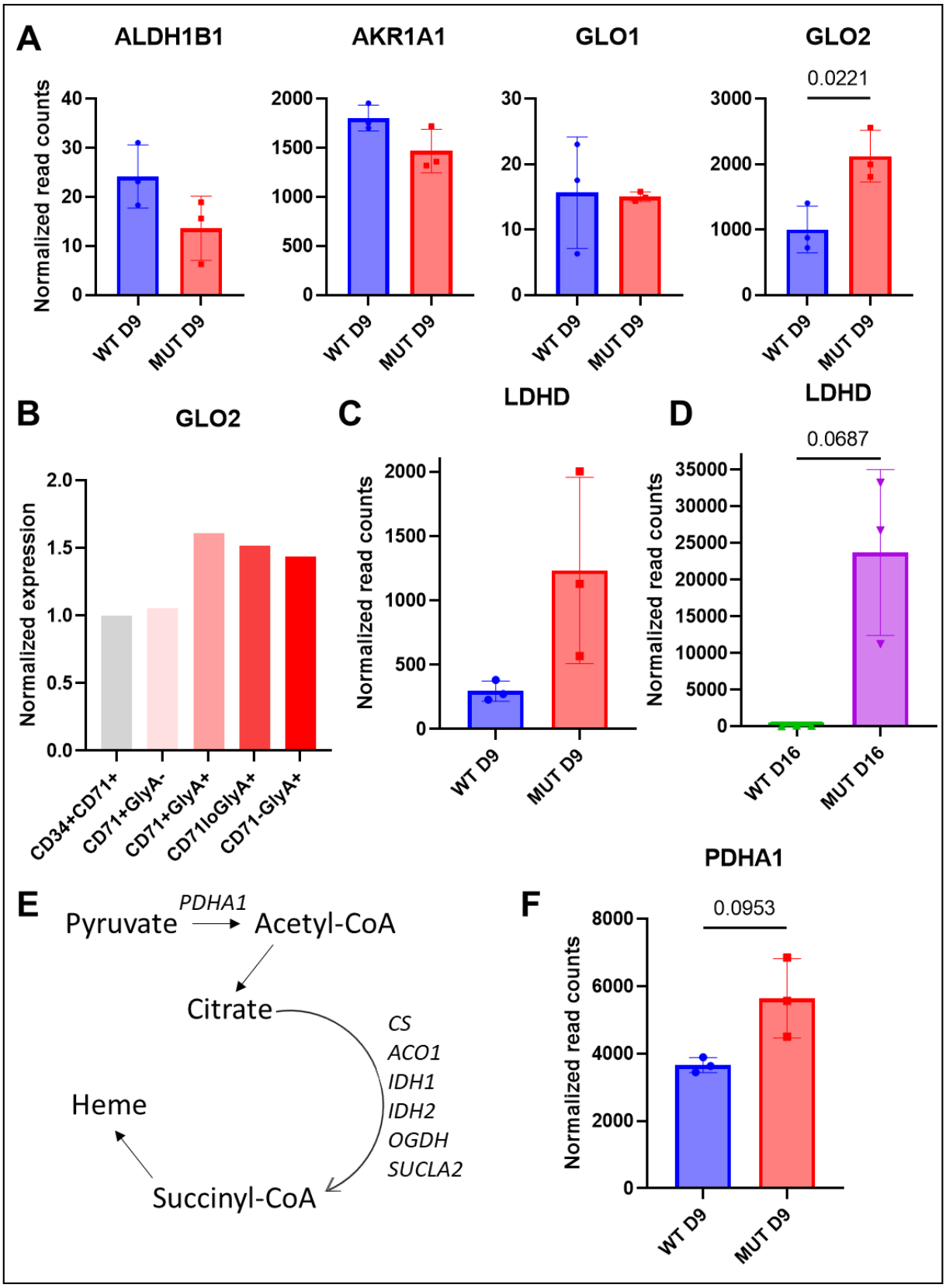
A
ALDH1B1
AKR1A1
GLO1
GLO2
Normalized read counts
0.0221
WT D9
MUT D9
B
GLO2
Normalized expression
CD34+CD71+
CD71+GlyA-
CD71+GlyA+
CD71loGlyA+
CD71-GlyA+
C
LDHD
Normalized read counts
D
LDHD
0.0687
WT D16
MUT D16
E
Pyruvate
PDHA1
Acetyl-CoA
Citrate
CS
ACO1
IDH1
IDH2
OGDH
SUCLA2
Succinyl-CoA
Heme
F
PDHA1
0.0953

Figure 3.4.6. Methylglyoxal generation drives pyruvate production in mutant-*SF3B1* cells through D-lactate dehydrogenase for succinyl-CoA production

A. Normalized RNA-sequencing read counts of methylglyoxal detoxifying enzymes, *ALDH1*, Aldehyde Dehydrogenase 1 Family Member A1, *AKR1A1*, Aldo-Keto Reductase Family 1 Member A1, *GLO1*, Glyoxalase I, and *GLO2*, Glyoxalase II in wild-type and mutant-*SF3B1* cells at day 9 of erythroid differentiation. N=3, Welch's corrected t-test. **B.** Normalized expression of *GLO2* during erythroid differentiation in erythroid progenitor, $CD34^{+}CD71^{-}$, pro-erythroblast, $CD71^{+}GlyA^{-}$, basophilic/polychromatic, $CD71^{+}GlyA^{+}$, polychromatic/orthochromatic, $CD71^{lo}GlyA^{+}$, and orthochromatic/enucleated erythrocyte, $CD71^{-}GlyA^{+}$ cells. Expression was analyzed using Bloodspot[111], normal hematopoiesis[115] is plotted. **C,D.** Normalized RNA-sequencing read counts of lactase dehydrogenase, *LDHD,* in wild-type and mutant-*SF3B1* cells **C.** at day 9 of erythroid differentiation **D.** at day 16 of erythroid differentiation. N = 3, Welch's correct t-test. **E.** Schematic outlining the connection between pyruvate and the tricarboxylic acid cycle which results in the production of succinyl-CoA for heme synthesis. **F.** Normalized RNA-sequencing read counts of *PDHA1,* Pyruvate Dehydrogenase E1 Subunit Alpha 1 at day 9 of erythroid differentiation in wild-type and mutant-*SF3B1* cells. N = 3, Welch's corrected t-test

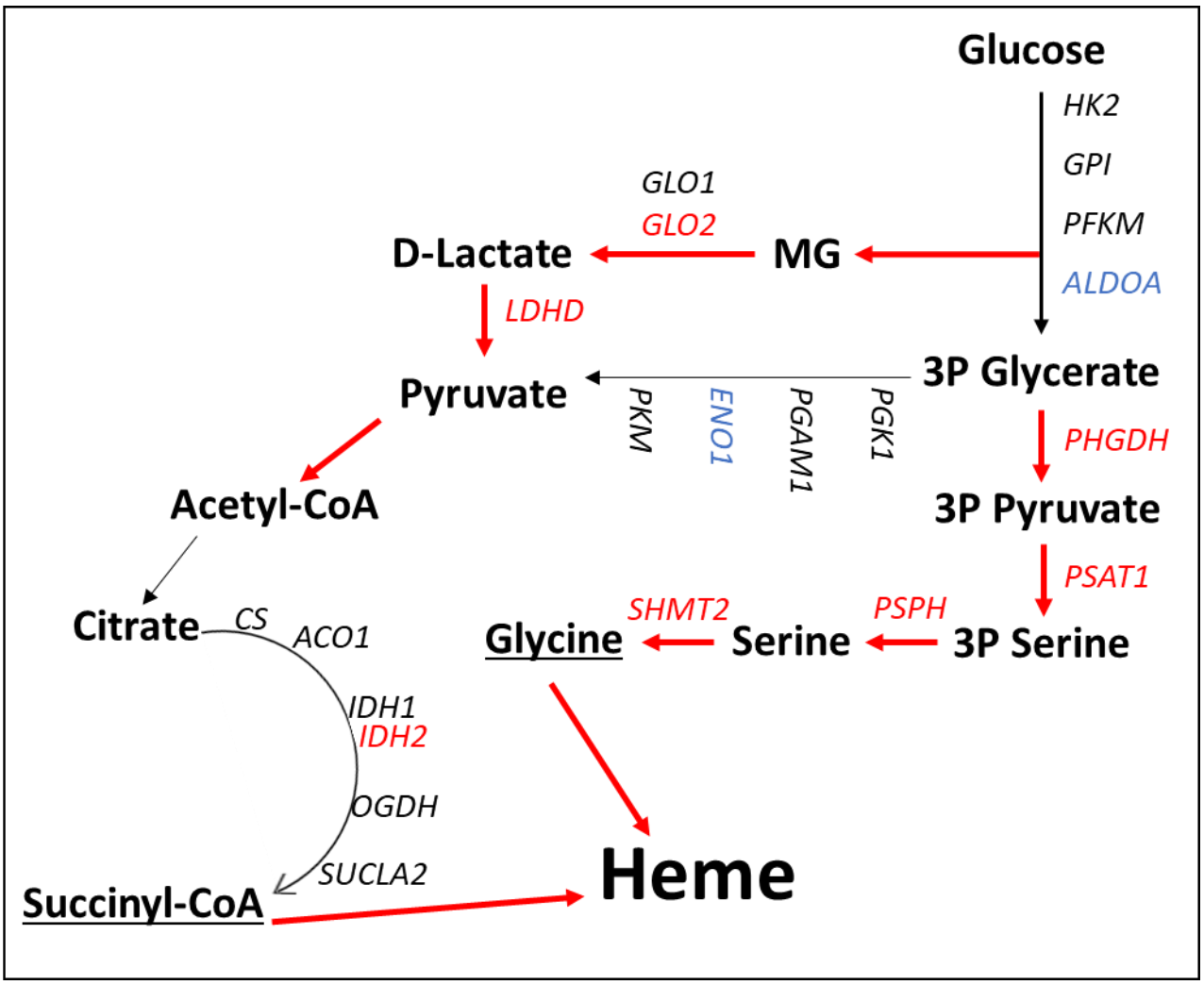

Figure 3.4.7. Summary of metabolic changes in mutant-*SF3B1* cells which converge on heme synthesis substrates glycine and succinyl-CoA

Standard glycolysis converts a glucose molecule to 3P Glycerate which can generate pyruvate which feeds into the tricarboxylic acid cycle (TCA) or is shunted into serine synthesis which results in the production of glycine which can feed into heme synthesis. Methylglyoxal (MG) is a toxic byproduct of glycolysis which is detoxified into D-lactate. D-lactate is further metabolized into pyruvate which can be fed through the TCA. The TCA cycle metabolizes citrate into succinyl-CoA which feeds into heme synthesis. Collectively, this glycolytic shift results in the tandem generation of glycine and succinyl-CoA to allow for excessive heme synthesis in mutant-*SF3B1* cells. Genes are italicized, upregulated expression is indicated in red font, downregulated expression is indicated in blue font. Some metabolic intermediates and genes are omitted for clarity.

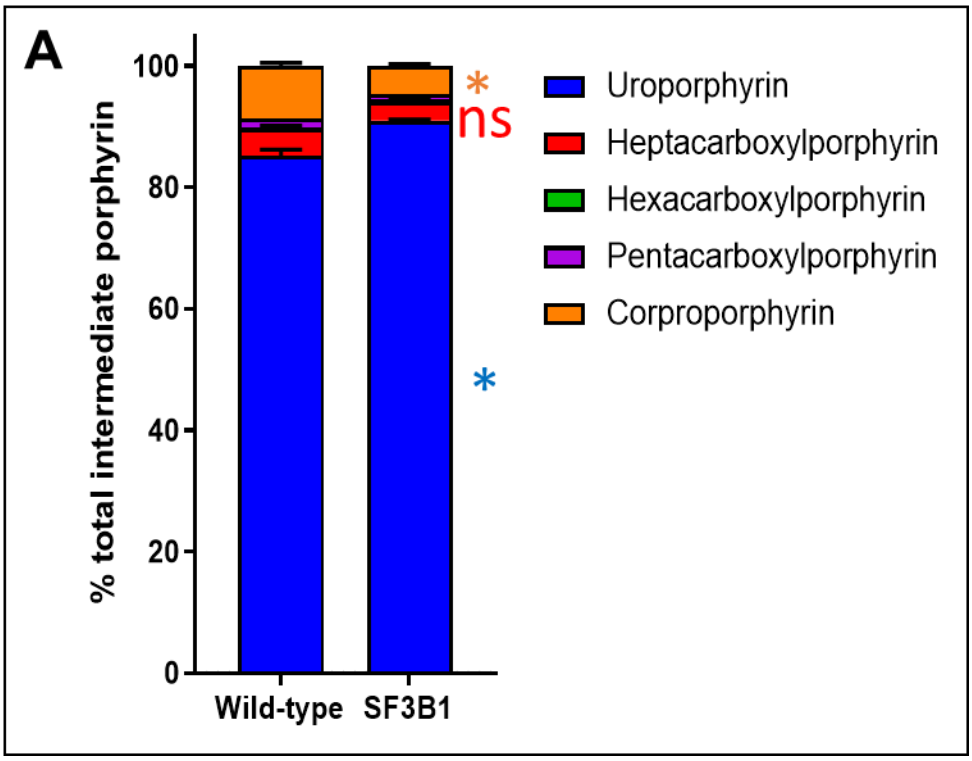

Figure 3.4.S1 – related to Figure 3.4.1

A. Proportion of cytosolic porphyrins in wild-type and mutant-*SF3B1* erythroid cells collected at day 15 of erythroid differentiation. N = 2, Welch's corrected t-test

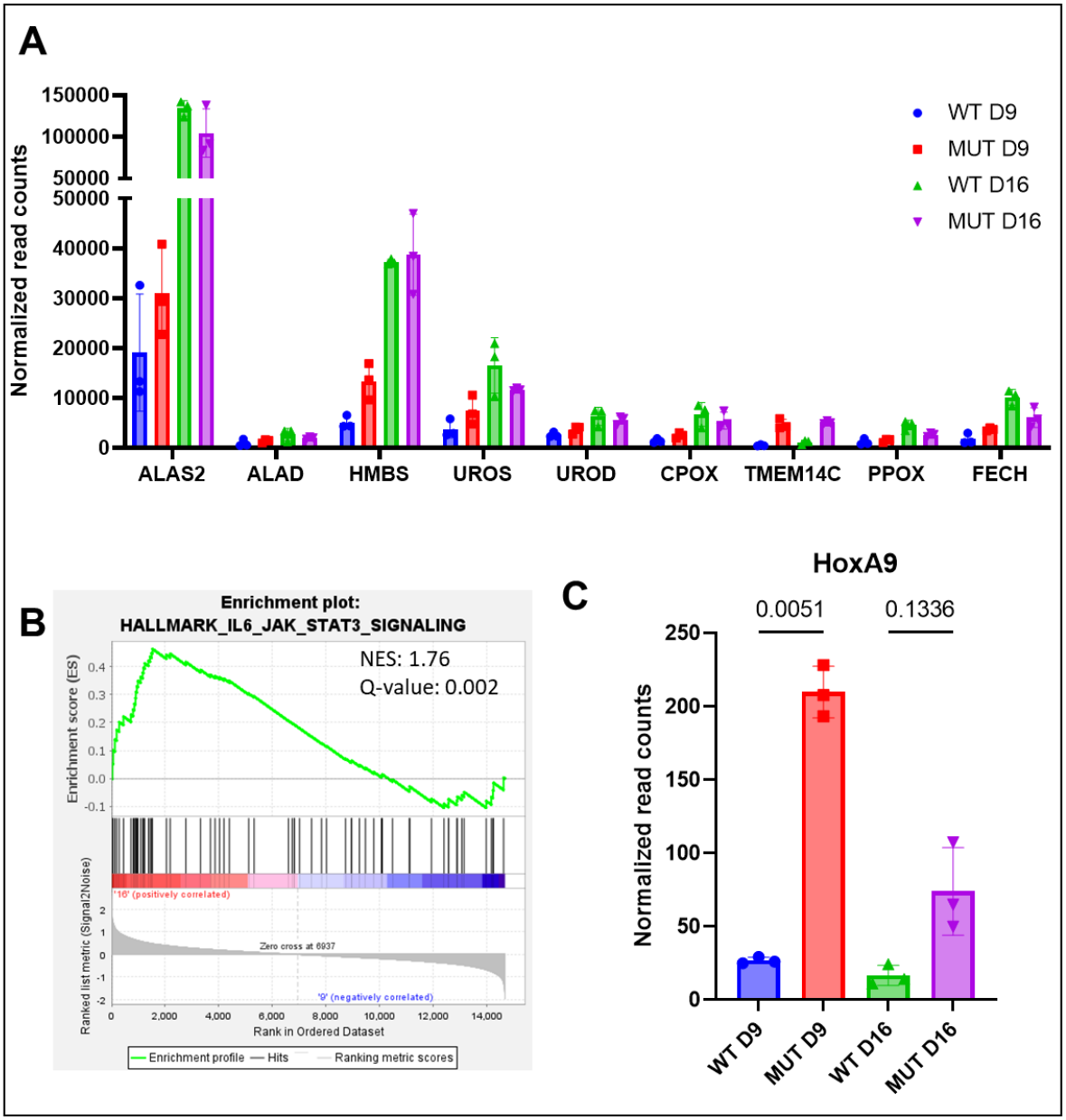

Figure 3.4.S2 – related to Figure 3.4.3

A. Normalized RNA-sequencing read counts of heme synthesis genes at both day 9 and day 16 of erythroid differentiation in isogenic wild-type and mutant-*SF3B1* cells. *ALAS2*, 5'-Aminolevulinate Synthase 2, *ALAD*, Aminolevulinate Dehydratase, *HMBS*, Hydroxymethylbilane Synthase, UROS, Uroporphyrinogen III Synthase, *UROD*, Uroporphyrinogen Decarboxylase, *CPOX*, Coproporphyrinogen Oxidase, *TMEM14C*, Transmembrane Protein 14C, *PPOX,* Protoporphyrinogen Oxidase, *FECH*, Ferrochelatase. **B.** GSEA plot of GO hallmark gene list IL6_JAK_STAT3_SIGNALING comparing day 16 to day 9 mutant-*SF3B1* erythroid cells. NES = normalized enrichment score. Q-value indicates false discovery rate. **C.** Normalized RNA-sequencing read counts of HOXA9 in day 9 and day 16 wild-type and mutant-*SF3B1* erythroid cells. N = 3, p-value calculated using Welch's correct t-test

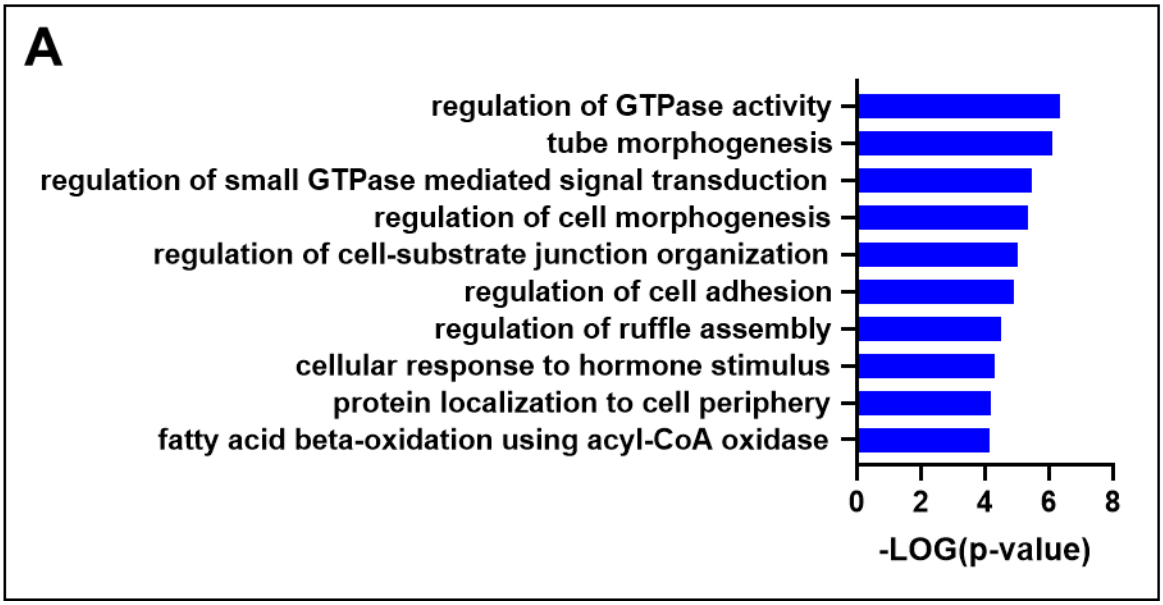

Figure 3.4.S3 related to Figure 3.4.4

A. GO analysis of mutant-*SF3B1* downregulated genes at day 16 of erythroid differentiation compared to wild-type cells. Genes with a log2fold change ≤ -0.5 and p value < 0.05 as calculated in DESEQ2 analysis were included. N= 616 genes

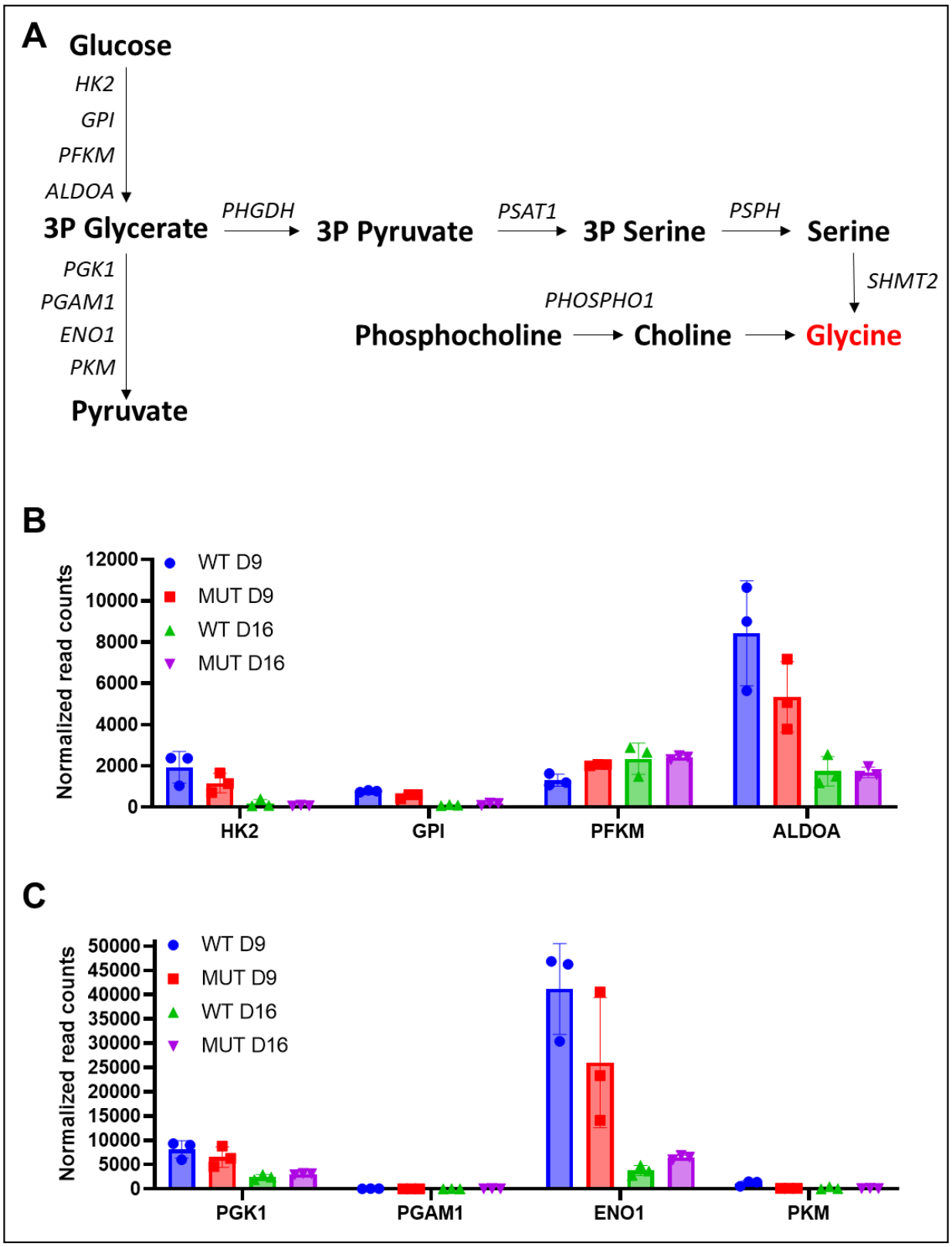
A
Glucose
HK2
GPI
PFKM
ALDOA
3P Glycerate
PHGDH
3P Pyruvate
PSAT1
3P Serine
PSPH
Serine
SHMT2
PHOSPHO1
Phosphocholine
Choline
Glycine
PGK1
PGAM1
ENO1
PKM
Pyruvate
B
Normalized read counts
12000
10000
8000
6000
4000
2000
0
WT D9
MUT D9
WT D16
MUT D16
HK2
GPI
PFKM
ALDOA
C
Normalized read counts
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
WT D9
MUT D9
WT D16
MUT D16
PGK1
PGAM1
ENO1
PKM

Figure 3.4.S4 related to Figure 3.4.5

A. Detailed schematic of genes related to glycine production. Cells can metabolize glucose using standard glycolysis to produce pyruvate or 3P Glycerate can be shunted into serine synthesis which can be metabolized to make glycine. Alternatively, cells can breakdown phosphocholine to generate glycine. Genes involved in each pathway are italicized. Some genes and metabolic intermediates of each pathway are omitted for clarity. **B,C.** Normalized RNA-sequencing read counts of glycolytic genes in wild-type and mutant-*SF3B1* erythroid cells at days 9 and 16 of erythroid differentiation. **B.** genes upstream of 3P glycerate. **C.** genes downstream of 3P glycerate.

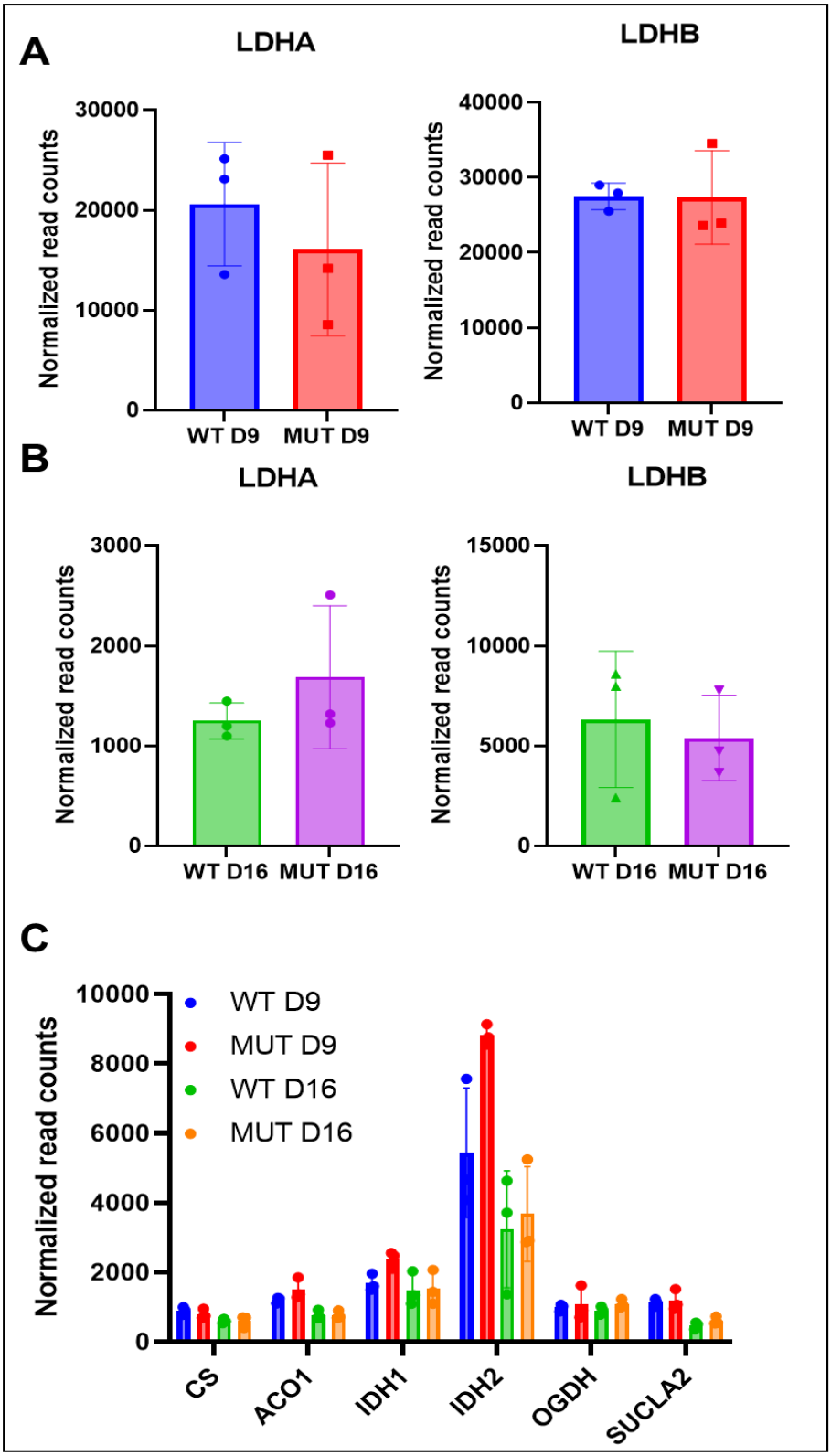
A
LDHA
LDHB
Normalized read counts
30000
20000
10000
0
WT D9
MUT D9
40000
30000
20000
10000
0
WT D9
MUT D9
B
LDHA
LDHB
Normalized read counts
3000
2000
1000
0
WT D16
MUT D16
15000
10000
5000
0
WT D16
MUT D16
C
Normalized read counts
10000
8000
6000
4000
2000
0
WT D9
MUT D9
WT D16
MUT D16
CS
ACO1
IDH1
IDH2
OGDH
SUCLA2

Figure 3.4.S5 related to Figure 3.4.6

A,B Normalized RNA-sequencing read counts of lactase dehydrogenases *LDHA* and *LDHB* in wild type and mutant-*SF3B1* cells **A.** at day 9 and **B.** at day 16 of erythroid differentiation. **C.** Normalized RNA-sequencing read counts of tricarboxylic acid cycle enzymes between citrate and succinyl CoA in wild type and mutant-*SF3B1* cells at days 9 and 16 of erythroid differentiation. *CS,* citrate synthase, *ACO1,* aconitase 1, *IDH1/2,* isocitrate dehydrogenase (NADP(+)) 1/2, *OGDH,* oxoglutarate dehydrogenase, and *SUCLA2,* Succinate-CoA Ligase ADP-Forming Subunit Beta

3.5 Methods

Porphyrin, FePPIX, and ZnPPIX analysis: Porphyrins (excluding PPIX) were analyzed using a Waters ultra-performance liquid chromatography (UPLC) system. Porphyrins are separated on a reverse-phase C18 column, 2.1x 100 mm BEH, and visualized with the tandem high sensitivity photodiode array and fluorescence detectors of the UPLC. Molecules analyzed include: the biosynthetic intermediate porphyrins: uroporphyrin, heptacarboxylporphyrin, hexacarboxylporphyrin, pentacarboxylporphyrin, and coproporphyrin. (Standards are available from Frontier Scientific, Logan, UT.) A UPLC system equipped with a reverse phase C18 column, a photodiode array detector and a fluorescence detector is used to determine total heme (FePPIX), ZnPPIX, and protoporphyrin IX (PPIX). Samples are subjected to solvent extraction and all noncovalently associated heme and PPIX are transferred into the organic phase and quantified. Hemin chloride (the Fe+3 equivalent of heme formed during sample processing) and PPIX are used as standards for quantitation. Heme/porphyrin/metal analysis was performed at the Iron and Heme Core facility at the University of Utah, supported in part by a grant from the NIH National Institute of Diabetes and Digestive and Kidney Diseases, Grant number U54DK110858.

Erythroid differentiation: Differentiation protocol was adapted based on Lee et al.(**Lee et al)**. 5F-HPC cells expanded in progenitor media with doxycycline were placed directly into erythroid stage 1 media without doxycycline or exogenous iron. Stage 1 (6 days): IMDM + 1% BSA (Gibco), 20% FBS (Sigma), 1 mM L-glutamine, penicillin/streptomycin, 500 µg/ml holo-transferrin (Sigma), 10 µg/ml human insulin (CellSciences), 6U Epo (Procrit), 100 ng/ml SCF, and 5 ng/ml IL-3. Cells were seeded at a density of 2-3 x 10^5 cells/ml in 6-well plates. Stage 2 (10 days): Media were changed to: IMDM + 1% BSA, 20% FBS, 1 mM L-glutamine, penicillin/streptomycin, 500 µg/ml holo-transferrin, 10 µg/ml human insulin, 6U Epo, and 50 ng/ml SCF. Cells were seeded at a density of 3 x 10^5 cells/ml in 6-well plates as cell number

permitted. During stage 2, cells were maintained at a density of <2 x 10^6/ml in a maximum volume of 6ml of media in a 6-well plate. For larger numbers of cells, plates and dish sizes were scaled up accordingly to maintain the 3 x 10^5 cells/ml density. Erythroid differentiation was carried out until day 16 (6 days stage 1 + 10 days stage 2).

Fluorescence activated cell sorting: Erythroid cells were stained at day 9 and day 16 using antibodies CD71 APC (M-A712; BD) and CD235a/Glycophorin A PE-Cy7 (11E4B-7-6; Coulter). Stain was performed with <1x10^6 cells per 50-100 µl sorting buffer (PBS + 50% FBS) with 1:100 dilution of each antibody, 30 min at 37C in the dark. Acquisition was performed on BD ARIA III.

RNA sequencing: $CD71^+$ $GlyA^+$ progenitors were isolated by flow sorting on day 9 and day 16 of erythroid differentiation. RNA was extracted from 3x10^5-1.5x10^6 cells of each population using the TRIzol reagent (Thermofisher). Up to 500ng of total RNA was used as input to make poly(A)-selected RNA-seq libraries with the TruSeq RNA library prep kit v2 (Illumina). Purified libraries were sequenced on an Illumina Hi-Seq 2000 using paired-end, 150 bp reads. Reads were mapped to hg38 genome using the kallisto package (**Bray 2016**). Differential gene expression was performed using the DESEQ2 R package (**Love 2014**). Gene set enrichment analysis (GSEA) was performed on the hallmark gene set in the Broad MSigDB (https://www.broadinstitute.org/msigdb) (**Liberzon 2015**). Metascape (**Zhou 2019**) was also used for gene ontology analyses.

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.

Chapter 4. Discussion

4.1 Low-risk MDS is characterized by ring sideroblast formation and altered erythropoiesis

Myelodysplastic syndromes (MDS) are a heterozygous group of disorders characterized by hematopoietic stem and progenitor cell (HSPC) clonal expansion. This clonal expansion is often accompanied by ineffective differentiation of megakaryocytic erythroid progenitors (MEP) which results in anemia and/or thrombocytopenia[54]. The median age of onset of MDS is ~ 70 years of age and patient prognosis varies depending on the spectrum of acquired mutations[2]. To date, approximately 30 driver mutations have been identified primarily in pathways related to splicing and control of gene expression including epigenetic modelers[14]. Recently, revised guidelines to the International Prognostic Scoring System (IPSS-R) have been published to integrate patient mutational status to further striate patients into high-and low- risk groups[11] to aid with treatment plans and better inform predicted patient outcomes.

One subtype of low-risk MDS, MDS with ring sideroblasts (MDS-RS), is characterized by ineffective erythropoiesis and ring sideroblast (RS) formation, iron-laden mitochondria that form perinuclear rings in polychromatic and orthochromatic erythroblasts. These patients near-universally harbor mutations within the splicing factor, SF3B1, a U2 spliceosome component that mediates adenosine branch point binding[40]. Mutations in SF3B1 result in mis-splicing through aberrant branch point and 3' cryptic splice site utilization[42, 43]. While this strong genotype to phenotype connection has been known for the past decade, how mutant-SF3B1 mis-splicing results in clonal advantage in the HSPC compartment, mitochondrial iron accumulation and RS formation, and the loss of erythroid maturation is unclear[10].

Therapy options for MDS are very limited and are typically reserved for high-risk patients with the worst predicted outcomes. Hematopoietic stem cell transplant is the only curative option available to these patients but general patient age and health often limits the ability to

transplant. The use of other traditional therapies such as chemotherapy or hypomethylating agents aimed at eliminating the mutant HSPC cells are often restricted to the individuals within the high-risk group[116]. In contrast, low risk patient treatment strategies are limited to supportive therapies which focus on restoration of erythropoiesis to limit transfusions to lower iron overload[117]. These erythroid modulating therapies include erythropoietin (EPO) and the recently approved, luspatercept, a fusion protein which sequesters TGF ligands to limit SMAD signaling[118]. While these treatments do improve erythroid output and function, many patients only partially respond resulting in chronic transfusion dependence. An SF3B1 inhibitor, H3B-8800, showed great promise in eliminating mutant-*SF3B1* HSPC clones in pre-clinical studies[119] but failed to eliminate mutant cells in a phase 1 trial[120]. However, a subset of patients became transfusion independent suggesting that this may provide some therapeutic benefit to these patients[120]. Thus, there is a clear need to further dissect the underlying biology of low-risk MDS to develop new therapeutics to improve patient outcomes.

In this chapter, I will discuss our contributions to dissecting the role of mutant-*SF3B1* in RS formation and erythroid maturation defects and will outline our future directions to dissect the role of mutant-*SF3B1* in HSPCs.

4.2 Mutant-*SF3B1* dysregulates mitochondrial iron metabolism to drive RS formation

Given the strong connection between mutant-*SF3B1* and ring sideroblast formation we aimed to dissect the connection between mutant-*SF3B1* mis-splicing and dysregulation of mitochondrial iron. As outlined in chapter 2, we developed MDS-RS patient derived induced pluripotent stem cell hematopoietic progenitor cell lines. From this patient, we isolated wild-type *SF3B1*$^{+/+}$ and mutant *SF3B1*$^{G742D/+}$ and *SF3B1*$^{G742D/+}$;*EZH2*$^{R685H/+}$ lines. These lines have robust progenitor expansion and erythroid potential. These isogenic cell lines allow for the first time, the direct comparison between wild-type and mutant-*SF3B1* erythroid cells during differentiation

and allow for a systematic daily evaluation of RS formation to better understand the kinetics of iron accumulation. Using these lines, we show the first *in vitro* recapitulation of mutant-*SF3B1* mediated ring sideroblast formation. To identify causative drivers of RS formation, we isolated $CD34^+$ progenitor cells and two early erythroid cell populations sorted on $CD71^+$ or $CD235^+$ for bulk RNA-sequencing and splicing analysis. To narrow our list of target genes, we focused on genes related to inherited sideroblastic anemias which often harbor mutations in heme and iron-sulfur cluster related genes[19]. Similar to previous studies[42-44, 46, 47], we identify high levels of mis-splicing in heme synthesis genes *TMEM14C* and *PPOX*, and *ABCB7*, a mitochondrial transporter essential for cytosolic iron sulfur (Fe-S) cluster biogenesis[27]. We show that mis-splicing results in a reduction of protein, notably, via a 5' UTR mechanism for *TMEM14C*. Traditionally, studies have focused on mis-splicing resulting in nonsense mediated RNA decay but other 5' UTR mis-splicing events, notably, mis-splicing of the 5' UTR of *DYNLL1*[45] have been identified. Mis-splicing of *DYNLL1* results in the overexpression of protein and thus, it is clear that mis-splicing impacts gene expression beyond nonsense mediated RNA decay. Through functional rescue experiments, we show that the mis-splicing of *TMEM14C* and *ABCB7* contribute to RS formation in mutant-SF3B1 cells.

While our study confirms the longstanding hypothesis that mutant-*SF3B1* mimics inherited sideroblastic anemias by mis-splicing genes involved in heme and Fe-S synthesis to drive RS formation, the exact mechanism remains to be determined. We and others have observed an accelerated erythroid differentiation pattern in mutant-*SF3B1* cells[68] but if this accelerated differentiation results in altered mitochondrial iron patterns resulting in mitochondrial ferritin (FTMT) formation is unclear. FTMT is thought to provide a protective mechanism against iron mediated oxidative stress but most of this work has focused on the role in the brain[121]. We and others[57] have proposed that mutant-*SF3B1* results in mitochondrial ferritin accumulation is a response to decreased mitochondrial iron processing through heme synthesis and Fe-S

biogenesis since FTMT is not observed in normal erythroid cells. Whether alterations to these pathways drives oxidative stress in erythroid cells that also results in FTMT accumulation is unclear but increased levels of mitochondrial ferritin are associated with aberrant upregulation of the transferrin receptor and is associated with apoptosis irrespective of mutation[122]. Future work focusing on the role of iron trafficking into erythroid cells and into the mitochondria is needed to determine the timing and mechanism of mitochondrial ferritin formation and its contributions to ineffective erythropoiesis.

4.3 A metabolic shift in mutant-*SF3B1* may drive ineffective erythropoiesis

Ineffective erythropoiesis is the loss of erythroid cells due to apoptosis of late-stage erythroblasts which is common to both inherited and acquired sideroblastic anemias[123]. These anemias result in iron-loading through the downregulation of hepcidin production which results in chronically elevated serum iron levels[124]. Whether mitochondrial ferritin production is partially a response to chronically elevated iron levels remains to be determined. In chapter 3, we explored the contribution of mutant-*SF3B1* to heme production in late-stage erythroid cells and discover that although heme synthesis is dysregulated, mutant cells produce more heme than wild-type cells. These findings complicate interpretation of FTMT accumulation as it is less clear whether dysregulated heme synthesis contributes to its accumulation. Taken together, further explorations into the extent that systemic iron overload, heme synthesis, and Fe-S production result in FTMT accumulation is needed to further dissect what drives RS formation in these patients.

While the contributions of FTMT to erythroid apoptosis has been shown[122], we hypothesized that other factors may be contributing to ineffective erythropoiesis in mutant-*SF3B1* cells. We sought to answer this question by generating an RNA-sequencing data set at mid- and late stages of erythroid maturation which roughly capture the beginning and end of heme synthesis. Utilizing our MDS-RS iPSC lines, erythroid cells were sorted based on the

expression of surface markers CD71, transferrin receptor, and Glycophorin A at day 9 and day 16 of differentiation for bulk RNA sequencing analysis. Given the excessive amount of heme synthesis in mutant-*SF3B1* cells, we hypothesized that dysregulation of this pathway may drive ineffective erythropoiesis. Through differential gene expression analysis, we identify a metabolic shift toward the glycolytic pathway which results in the coupled production of heme substrates glycine and succinyl-CoA. Combined with the identification of a skew toward erythroid inhibitory gene signatures, our findings suggest mutant-*SF3B1* cells are predisposed to a loss of erythroid differentiation which is exacerbated through metabolic reprogramming. In order to dissect the consequences of this metabolic reprogramming, metabolomic and lipidomic analyses will be conducted to better assess the level of metabolic alterations in these cells.

It is possible that ineffective erythropoiesis is driven by cell extrinsic factors within the bone marrow niche. Dysregulation of the stroma can result in oxidative stress and inflammatory dysregulation to drive clonal expansion and leukemic progression[125]. Several inflammatory signatures have been reported in mutant-*SF3B1* cells[126] and the contributions of altered cytokine secretion in inhibiting erythroid differentiation has been largely understudied. Since iron overload can exacerbate oxidate stress, future work will focus on the intersection of iron accumulation and cellular stress which may drive ineffective erythropoiesis in MDS-RS.

4.4 Development of a primary human cell model to study mutant-*SF3B1* hematopoietic stem and progenitor cells

In MDS patients, HSPC clonal evolution is often predicted using deep whole genome sequencing (WGS) to determine variant allele frequency of known driver mutations to infer the order of mutational acquisition[13]. Recently other methods, such as iPSC reprogramming of patient cells have been leveraged to capture and sequence individual clones to infer this order[48]. Using these methods, *SF3B1* mutations have been found to predominantly be the

initiating clonal event[10] but how mutant-*SF3B1* promotes clonal expansion and drives progression to MDS is unknown.

In the last decade, several mutant-*SF3B1* models have been developed in an attempt to address this question (**Figure 4.1**). The first murine *SF3B1* studies were in a heterozygous knock out (KO) mouse[127-129] which exhibited no clear evidence of splicing defects, impairment of the erythroid compartment or ring sideroblast formation but had impaired hematopoietic stem cell proliferation and repopulating potential. As MDS-RS patients harbor gain of function mutations in MDS, the next generation of these mice were conditional heterozygous knock-in (KI) models that expressed the *SF3B1*^{K700E} allele, the most prevalent mutation found in MDS[58, 59]. In contrast to the KO models, these mice exhibited aberrant splicing and impaired erythroid cells but had no induction of RS likely driven by a lack of overlap in mis-spliced genes due to a lack of intronic homology[130]. Interestingly, these mice have an expanded long-term HSC (LT-HSC) compartment but exhibit a repopulation disadvantage when competed against wild-type cells complicating interpretation of whether the models are recapitulating clonal expansion.

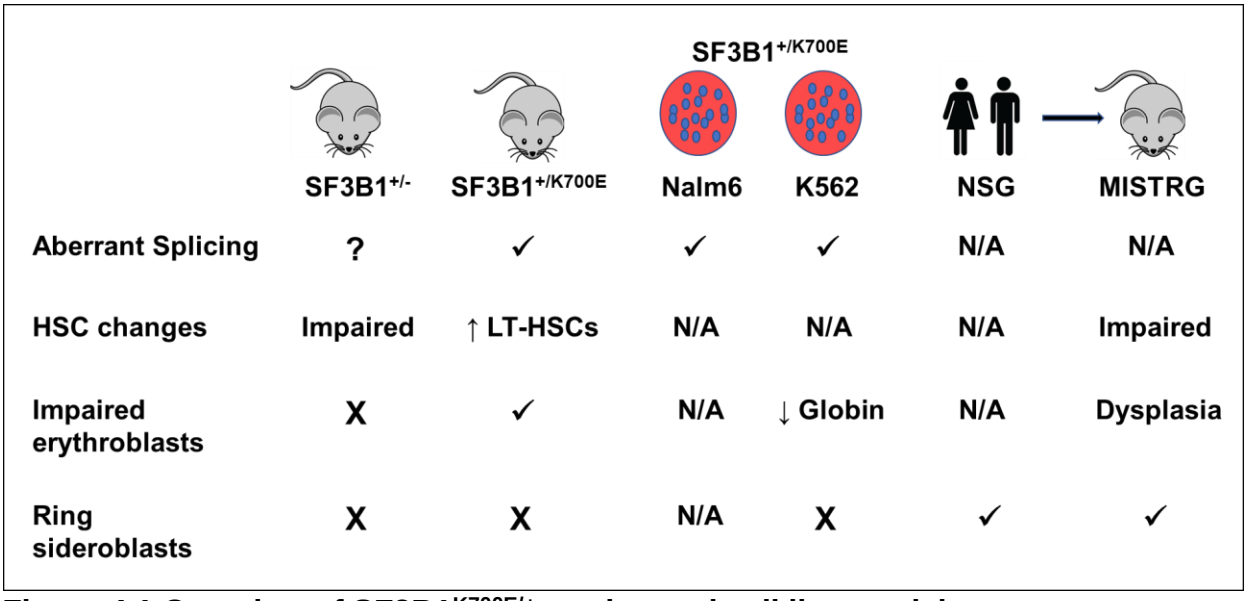

Figure 4.1 Overview of SF3B1$^{K700E/+}$ murine and cell line models

Since murine models do not recapitulate mutant-*SF3B1* mis-splicing biology, several human mutant-*SF3B1* have been established[42, 43, 81, 131] to better define the mis-spliced transcriptome of MDS-RS. However, these models have been established from erythroid or Pro-B cell populations and thus do not capture HSPC gene expression. To try to overcome these limitations, patient derived xenograft (PDX) models have been developed to directly assess MDS-RS cells *in vivo.* Initial studies in NSG mice were the first system to capture ring sideroblast formation and found that only LT-HSCs were capable of propagating the *SF3B1* mutation supporting the LT-HSC expansion observed in the KI murine models[60]. PDX models were further improved with the use of MISTRG mice, immunodeficient mice that express humanized M-CSF, IL3/GM-CSF, SIRP alpha, and thrombopoietin to better support human stem cell engraftment[132]. This model robustly captures myeloid dysplasia, RS formation, multi-lineage reconstitution, and the ability to be serially transplanted confirming that mutant-*SF3B1* LT-HSCs are capable of propagating the mutation[61]. While this updated model provides a robust system for testing potential therapeutics, the availability of MISTRG mice and available patient samples limit mechanistic studies.

To date, there have been no functional studies to assess whether mutant-*SF3B1* is sufficient to drive clonal advantage as murine models do not fully recapitulate MDS and MDS-RS xenograft models rely on patient samples which harbor additional mutations. Recently, updated RNA sequencing and long-read analysis of MDS-RS patient samples allow for the quantification of gene expression, splicing analysis and genotyping on a single cell level[91]. While this work identifies a fitness advantage in mutant-*SF3B1* erythroid progenitors, limited hematopoietic stem cell numbers prevent the further dissection of whether mutant-*SF3B1* confers a fitness advantage in the stem cell compartment.

Because of the collective limitations of these approaches, we developed a cellular engineering approach to create a primary human HSPC model to study how *SF3B1* mutations

drive clonal advantage. Successful strategies to efficiently and precisely edit primary human HSPCs have been developed over the last decade[133, 134]. Most of these approaches utilize CRISPR/Cas9 and adeno-associated virus (AAV) to deliver repair templates to guide homology directed repair (HDR). This AAV approach improves upon other methods which used single stranded oligonucleotides (ssODN)[135], as the packaging limit of AAV allowed for the simultaneous integration of selectable markers such as fluorescent proteins or antibiotic resistance genes for selection. We utilized these previous approaches to develop an HDR based approach to knock-in SF3B1^{K700E} into primary human CD34^{+} cells (**Fig 4.2**). These SF3B1$^{K700E/+}$ CD34^{+} HSPCs present a unique opportunity to study the role of mutant-*SF3B1* in HSPC clonal advantage and erythroid skewing observed in MDS-RS patients. To date, high throughput splicing analysis of mutant-*SF3B1* cells is limited by the clonal heterogeneity and progenitor bias within these patients which prevents the ability to easily capture wild-type and mutant-*SF3B1* stem cells.

While the scRNA-seq approach discussed previously[91] allows for the ability to striate wild-type and mutant cells, the approach cannot easily capture all genetic aberrations within each cell. Our editing approach will allow for high-throughput RNA splicing analysis to determine mutant-*SF3B1* mediated mis-spliced genes within CD34^{+} stem and progenitor cells. We hypothesize that this analysis will identify novel mis-spliced genes within these HSPCs. Follow-up functional analysis will focus on the contribution of these genes to clonal expansion or erythroid progenitor fitness advantage that is seen in MDS-RS patients. These *SF3B1*$^{K700E/+}$ CD34^{+} HSPCs will also be engrafted into mice to track clonal dynamics and lineage output of these cells. We expect that mutant cells will exhibit a fitness advantage in long-term stem cells and erythroid progenitors with erythroid differentiation defects. Should these studies show no advantage of mutant-*SF3B1* cells, we will focus on other approaches to better mimic MDS-RS such as the introduction of other common driver mutations such as *TET2*[92] or chronic

inflammatory stimuli to better mimic the bone marrow niche within these patients which may be needed to drive significant clonal advantage.

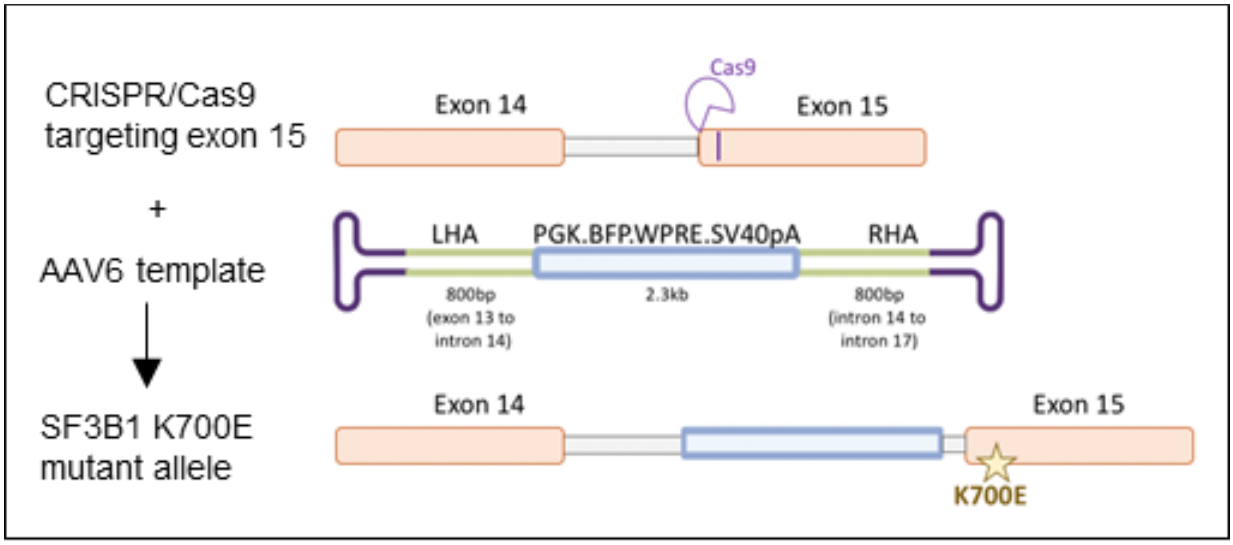

Figure 4.2 CRISPR/Cas9 and AAV editing schematic.
A. Schematic of CRISPR/Cas9 guide RNA which cuts near SF3B1^{K700E} which is combined with an AAV6 template which contains 1.6kb of total homology which inserts a gene cassette for BFP expression into the adjacent intron. The right homology arm contains the K700E A>G mutation and a mutation to abolish the gRNA PAM sequence.

In conclusion, the findings outlined in this thesis provide a solid foundation for future MDS-RS studies. Our MDS-RS cell line model should be leveraged to dissect the mechanism of mitochondrial ferritin accumulation in mutant-*SF3B1* cells to better understand the role of mitochondrial ferritin in ring sideroblast formation and ineffective erythropoiesis. Our identification of a glycolytic shift in mutant-*SF3B1* cells presents an exciting therapeutic avenue in MDS-RS as glycolysis is a druggable pathway[136] and its modulation may rescue erythroid maturation in these cells. Lastly, the development of a gene editing approach to study the role of mutant-*SF3B1* within the stem cell compartment may discover novel approaches to develop curative treatment therapies for these low-risk MDS patients.

References

1. Ma X. Epidemiology of Myelodysplastic Syndromes. *The American Journal of Medicine*. 2012-07-01 2012;125(7):S2-S5. doi:10.1016/j.amjmed.2012.04.014

2. Fenaux P, Haase D, Sanz GF, Santini V, Buske C. Myelodysplastic syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology*. 2014-09-01 2014;25(suppl 3):iii57-iii69. doi:10.1093/annonc/mdu180

3. Malcovati L, Hellström-Lindberg E, Bowen D, et al. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood*. 2013-10-24 2013;122(17):2943-2964. doi:10.1182/blood-2013-03-492884

4. *Berliner Klinische Wochenschrift*. 1900;37:851.

5. Swerdlow SC, E; Harris, NL; Jaffe, ES; Pileri, SA; Stein, H; Thiele, J; Vardiman, JW *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues* 4th ed. WHO Press 2008.

6. Bennett JM CD, Daniel MT, Flandrin G, Galton DA, Gralnick HR, Sultan C. Proposals for the classification of the myelodysplastic syndromes. *British Journal Haematology*. Jun 1982;51(2):189-99.

7. Greenberg P, Cox C, Lebeau MM, et al. International Scoring System for Evaluating Prognosis in Myelodysplastic Syndromes. *Blood*. 1997-03-15 1997;89(6):2079-2088. doi:10.1182/blood.v89.6.2079

8. Vardiman JW, Harris NL, Brunning RD. The World Health Organization (WHO) classification of the myeloid neoplasms. *Blood*. 2002-10-01 2002;100(7):2292-2302. doi:10.1182/blood-2002-04-1199

9. G HMaH. The 2016 Revision to the World Health Organization Classification of Myelodysplastic Syndromes. *J Transl Int Med*. 2017;5(3):139-143. doi:doi: 10.1515/jtim-2017-0002

10. Malcovati L, Stevenson K, Papaemmanuil E, et al. SF3B1-mutant MDS as a distinct disease subtype: a proposal from the International Working Group for the Prognosis of MDS. *Blood*. 2020-07-09 2020;136(2):157-170. doi:10.1182/blood.2020004850

11. Bernard E, Tuechler H, Greenberg PL, et al. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes. *NEJM Evidence*. 2022-06-28 2022;1(7)doi:10.1056/evidoa2200008

12. foundation M. https://www.mds-foundation.org/ipss-r-calculator/

13. Bowman RL, Busque L, Levine RL. Clonal Hematopoiesis and Evolution to Hematopoietic Malignancies. *Cell Stem Cell*. 2018-02-01 2018;22(2):157-170. doi:10.1016/j.stem.2018.01.011

14. Papaemmanuil E, Gerstung M, Malcovati L, et al. Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*. 2013-11-21 2013;122(22):3616-3627. doi:10.1182/blood-2013-08-518886

15. Ogawa S. Genetics of MDS. *Blood*. 2019-03-07 2019;133(10):1049-1059. doi:10.1182/blood-2018-10-844621

16. Cazzola M, Rossi M, Malcovati L. Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms. *Blood*. 2013-01-10 2013;121(2):260-269. doi:10.1182/blood-2012-09-399725

17. Abu-Zeinah G, Desancho MT. Understanding Sideroblastic Anemia: An Overview of Genetics, Epidemiology, Pathophysiology and Current Therapeutic Options. *Journal of Blood Medicine*. 2020-09-01 2020;Volume 11:305-318. doi:10.2147/jbm.s232644

18. Cazzola M, Invernizzi R. Ring sideroblasts and sideroblastic anemias. *Haematologica*. 2011-06-01 2011;96(6):789-792. doi:10.3324/haematol.2011.044628

19. Ducamp S, Fleming MD. The molecular genetics of sideroblastic anemia. *Blood*. 2019-01-03 2019;133(1):59-69. doi:10.1182/blood-2018-08-815951

20. Levi S, Ripamonti M, Dardi M, Cozzi A, Santambrogio P. Mitochondrial Ferritin: Its Role in Physiological and Pathological Conditions. *Cells*. 2021-08-03 2021;10(8):1969. doi:10.3390/cells10081969

21. Braymer JJ, Lill R. Iron–sulfur cluster biogenesis and trafficking in mitochondria. *Journal of Biological Chemistry*. 2017-08-01 2017;292(31):12754-12763. doi:10.1074/jbc.r117.787101

22. Cotter PD, Baumann M, Bishop DF. Enzymatic defect in "X-linked" sideroblastic anemia: molecular evidence for erythroid delta-aminolevulinate synthase deficiency. *Proceedings of the National Academy of Sciences*. 1992-05-01 1992;89(9):4028-4032. doi:10.1073/pnas.89.9.4028

23. Shoolingin-Jordan PM, Al-Daihan S, Alexeev D, et al. 5-Aminolevulinic acid synthase: mechanism, mutations and medicine. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*. 2003/04/11/ 2003;1647(1):361-366. doi:https://doi.org/10.1016/S1570-9639(03)00095-5

24. Ducamp S, Kannengiesser C, Touati M, et al. Sideroblastic anemia: molecular analysis of the ALAS2 gene in a series of 29 probands and functional studies of 10 missense mutations. *Human Mutation*. 2011-06-01 2011;32(6):590-597. doi:10.1002/humu.21455

25. Campagna DR, De Bie CI, Schmitz-Abe K, et al. X-linked sideroblastic anemia due to ALAS2 intron 1 enhancer element GATA-binding site mutations. *American Journal of Hematology*. 2014-03-01 2014;89(3):315-319. doi:10.1002/ajh.23616

26. Iolascon A, De Falco L, Beaumont C. Molecular basis of inherited microcytic anemia due to defects in iron acquisition or heme synthesis. *Haematologica*. 2009-03-01 2009;94(3):395-408. doi:10.3324/haematol.13619

27. Pondarre C, Campagna DR, Antiochos B, Sikorski L, Mulhern H, Fleming MD. Abcb7, the gene responsible for X-linked sideroblastic anemia with ataxia, is essential for hematopoiesis. *Blood*. 4 2007;109:3567-3569. doi:10.1182/blood-2006-04-015768

28. Ye H, Jeong SY, Ghosh MC, et al. Glutaredoxin 5 deficiency causes sideroblastic anemia by specifically impairing heme biosynthesis and depleting cytosolic iron in human erythroblasts. *Journal of Clinical Investigation*. 2010-05-03 2010;120(5):1749-1761. doi:10.1172/jci40372

29. Liu Z, Yoshimi A, Wang J, et al. Mutations in the RNA Splicing Factor SF3B1 Promote Tumorigenesis through MYC Stabilization. *Cancer Discovery*. 2020-06-01 2020;10(6):806-821. doi:10.1158/2159-8290.cd-19-1330

30. Crooks DR, Ghosh MC, Haller RG, Tong W-H, Rouault TA. Posttranslational stability of the heme biosynthetic enzyme ferrochelatase is dependent on iron availability and intact iron-sulfur cluster assembly machinery. *Blood*. 2010-01-28 2010;115(4):860-869. doi:10.1182/blood-2009-09-243105

31. Guernsey DL, Jiang H, Campagna DR, et al. Mutations in mitochondrial carrier family gene SLC25A38 cause nonsyndromic autosomal recessive congenital sideroblastic anemia. *Nature Genetics*. 2009-06-01 2009;41(6):651-653. doi:10.1038/ng.359

32. Lunetti P, Damiano F, De Benedetto G, et al. Characterization of Human and Yeast Mitochondrial Glycine Carriers with Implications for Heme Biosynthesis and Anemia. *Journal of Biological Chemistry*. 2016-09-01 2016;291(38):19746-19759. doi:10.1074/jbc.m116.736876

33. Leblanc MA, Bettle A, Berman JN, et al. Study of Glycine and Folic Acid Supplementation to Ameliorate Transfusion Dependence in Congenital SLC25A38 Mutated Sideroblastic Anemia. *Pediatric Blood & Cancer*. 2016-07-01 2016;63(7):1307-1309. doi:10.1002/pbc.25981

34. Rademakers LHPM, Koningsberger JC, Sorber CWJ, Faille HBDL, Hattum JV, Marx JJM. Accumulation of iron in erythroblasts of patients with erythropoietic protoporphyria. https://doi.org/10.1111/j.1365-2362.1993.tb00752.x. *European Journal of Clinical Investigation*. 1993/02/01 1993;23(2):130-138. doi:https://doi.org/10.1111/j.1365-2362.1993.tb00752.x

35. Schmitz-Abe K, Ciesielski SJ, Schmidt PJ, et al. Congenital sideroblastic anemia due to mutations in the mitochondrial HSP70 homologue HSPA9. *Blood*. 2015-12-17 2015;126(25):2734-2738. doi:10.1182/blood-2015-09-659854

36. Crispin A, Guo C, Chen C, et al. Mutations in the iron-sulfur cluster biogenesis protein HSCB cause congenital sideroblastic anemia. *Journal of Clinical Investigation*. 2020-08-31 2020;130(10):5245-5256. doi:10.1172/jci135479

37. Camaschella C, Campanella A, De Falco L, et al. The human counterpart of zebrafish shiraz shows sideroblastic-like microcytic anemia and iron overload. *Blood*. 2007-08-15 2007;110(4):1353-1358. doi:10.1182/blood-2007-02-072520

38. Wingert RA, Galloway JL, Barut B, et al. Deficiency of glutaredoxin 5 reveals Fe–S clusters are required for vertebrate haem synthesis. *Nature*. 2005-08-18 2005;436(7053):1035-1039. doi:10.1038/nature03887

39. Ashorobi D CA. Sideroblastic Anemia. StatPearls Publishing; 2021.

40. Papaemmanuil E, Cazzola M, Boultwood J, et al. Somatic SF3B1 Mutation in Myelodysplasia with Ring Sideroblasts. *New England Journal of Medicine*. 10 2011;365:1384-1395. doi:10.1056/nejmoa1103283

41. Zhang J, Ali AM, Lieu YK, et al. Disease-Causing Mutations in SF3B1 Alter Splicing by Disrupting Interaction with SUGP1. *Molecular Cell*. 2019-10-01 2019;76(1):82-95.e7. doi:10.1016/j.molcel.2019.07.017

42. Darman RB, Seiler M, Agrawal AA, et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3 Splice Site Selection through Use of a Different Branch Point Splice Site Selection through Use of a Different Branch Point. *Cell Reports* 2015;13:1033-1045. doi:10.1016/j.celrep.2015.09.053

43. Dolatshad H, Pellagatti A, Liberante F, et al. Cryptic splicing events in the iron transporter ABCB7 and other key target genes in SF3B1-mutant myelodysplastic syndromes. *Leukemia*. 2016;30doi:10.1038/leu.2016.149

44. Shiozawa Y, Malcovati L, Gallì A, et al. Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nature Communications*. 12 2018;9:1-16. doi:10.1038/s41467-018-06063-x

45. Tam AS, Tsai S, Chang EY-C, et al. DYNLL1 mis-splicing is associated with replicative genome instability in SF3B1 mutant cells. *bioRxiv*. 2021:2021.05.26.445839. doi:10.1101/2021.05.26.445839

46. Conte S, Katayama S, Vesterlund L, et al. Aberrant splicing of genes involved in haemoglobin synthesis and impaired terminal erythroid maturation in SF3B1 mutated refractory anaemia with ring sideroblasts. *British Journal of Haematology*. 11 2015;171:478-490. doi:10.1111/bjh.13610

47. Lee SCW, North K, Kim E, et al. Synthetic Lethal and Convergent Biological Effects of Cancer-Associated Spliceosomal Gene Mutations. *Cancer Cell*. 8 2018;34:225-241.e8. doi:10.1016/j.ccell.2018.07.003

48. Hsu J, Reilly A, Hayes BJ, et al. Reprogramming identifies functionally distinct stages of clonal evolution in myelodysplastic syndromes. *Blood*. 7 2019;134:186-198. doi:10.1182/blood.2018884338

49. Clough CA, Pangallo J, Sarchi M, et al. Coordinated missplicing of TMEM14C and ABCB7 causes ring sideroblast formation in SF3B1-mutant myelodysplastic syndrome. *Blood*. 2022-03-31 2022;139(13):2038-2049. doi:10.1182/blood.2021012652

50. Cazzola M, Malcovati L. Diagnosis and treatment of sideroblastic anemias: from defective heme synthesis to abnormal RNA splicing. *Hematology*. 2015-12-05 2015;2015(1):19-25. doi:10.1182/asheducation-2015.1.19

51. Malcovati L, Karimi M, Papaemmanuil E, et al. SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood*. 7 2015;126:233-241. doi:10.1182/blood-2015-03-633537

52. Cazzola M, Invernizzi R, Bergamaschi G, et al. Mitochondrial ferritin expression in erythroid cells from patients with sideroblastic anemia. Blood; 2003. p. 1996-2000.

53. Allikmets R, Raskind WH, Hutchinson A, Schueck ND, Dean M, Koeller DM. Mutation of a putative mitochondrial iron transporter gene (ABC7) in X-linked sideroblastic anemia and ataxia (XLSA/A). *Pediatric Research*. 5 1999;8:743-749. doi:10.1093/hmg/8.5.743

54. Cazzola M. Myelodysplastic Syndromes. *New England Journal of Medicine*. 2020-10-01 2020;383(14):1358-1374. doi:10.1056/nejmra1904794

55. Pellagatti A, Armstrong RN, Steeples V, et al. Impact of spliceosome mutations on RNA splicing in myelodysplasia: dysregulated genes/pathways and clinical associations. *Blood*. 2018-09-20 2018;132(12):1225-1240. doi:10.1182/blood-2018-04-843771

56. Nikpour M, Pellagatti A, Liu A, et al. Gene expression profiling of erythroblasts from refractory anaemia with ring sideroblasts (RARS) and effects of G-CSF. *British Journal of Haematology*. 2009-10-08 2009;149(6):844-854. doi:10.1111/j.1365-2141.2010.08174.x

57. Nikpour M, Scharenberg C, Liu A, et al. The transporter ABCB7 is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts. *Leukemia*. 4 2013;27:889-896. doi:10.1038/leu.2012.298

58. Obeng EA, Chappell RJ, Seiler M, Wu CJ, Fleming MD, Ebert BL. Physiologic Expression of Sf3b1 K700E Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell*. 2016;doi:10.1016/j.ccell.2016.08.006

59. Mupo A, Seiler M, Sathiaseelan V, et al. Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts. *Leukemia*. 2017;31:720-727. doi:10.1038/leu.2016.251

60. Mortera-Blanco T, Dimitriou M, Woll PS, et al. SF3B1-initiating mutations in MDS-RSs target lymphomyeloid hematopoietic stem cells. *Blood*. 2017-08-17 2017;130(7):881-890. doi:10.1182/blood-2017-03-776070

61. Song Y, Rongvaux A, Taylor A, et al. A highly efficient and faithful MDS patient-derived xenotransplantation model for pre-clinical studies. *Nature Communications*. 2019-12-01 2019;10(1)doi:10.1038/s41467-018-08166-x

62. Elvarsdóttir EM, Mortera-Blanco T, Dimitriou M, et al. A three-dimensional in vitro model of erythropoiesis recapitulates erythroid failure in myelodysplastic syndromes. *Leukemia*. 2020-01-01 2020;34(1):271-282. doi:10.1038/s41375-019-0532-7

63. Hatta S, Fujiwara T, Yamamoto T, et al. A defined culture method enabling the establishment of ring sideroblasts from induced pluripotent cells of X-linked sideroblastic anemia. *Haematologica*. 2018-05-01 2018;103(5):e188-e191. doi:10.3324/haematol.2017.179770

64. Saito K, Fujiwara T, Hatta S, et al. Generation and Molecular Characterization of Human Ring Sideroblasts: a Key Role of Ferrous Iron in Terminal Erythroid Differentiation and Ring Sideroblast Formation. *Molecular and Cellular Biology*. 2019-04-01 2019;39(7)doi:10.1128/mcb.00387-18

65. Taylor J, Mi X, North K, et al. Single-cell genomics reveals the genetic and molecular bases for escape from mutational epistasis in myeloid neoplasms. *Blood*. 2020-09-24 2020;136(13):1477-1486. doi:10.1182/blood.2020006868

66. Seo JY, Lee K-O, Kim S-H, et al. Clinical significance of SF3B1 mutations in Korean patients with myelodysplastic syndromes and myelodysplasia/myeloproliferative neoplasms with ring sideroblasts. *Annals of Hematology*. 2014;93(4):603-608. doi:10.1007/s00277-013-1915-x

67. Doulatov S, Vo LT, Macari ER, et al. Drug discovery for Diamond-Blackfan anemia using reprogrammed hematopoietic progenitors. *Sci Transl Med*. Feb 08 2017;9(376)doi:10.1126/scitranslmed.aah5645

68. Lieu YK, Liu Z, Ali AM, et al. SF3B1 mutant-induced missplicing of MAP3K7 causes anemia in myelodysplastic syndromes. *Proc Natl Acad Sci U S A*. Jan 4 2022;119(1)doi:10.1073/pnas.2111703119

69. DeBoever C, Ghia EM, Shepard PJ, et al. Transcriptome Sequencing Reveals Potential Mechanism of Cryptic 3' Splice Site Selection in SF3B1-mutated Cancers. *PLOS Computational Biology*. 3 2015;11:e1004105. doi:10.1371/journal.pcbi.1004105

70. Alsafadi S, Houy A, Battistella A, et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nature Communications*. 2 2016;7:10615. doi:10.1038/ncomms10615

71. Araujo PR, Yoon K, Ko D, et al. Before It Gets Started: Regulating Translation at the 5' UTR. *Comparative and Functional Genomics*. 2012/05/28 2012;2012:475731. doi:10.1155/2012/475731

72. Yien YY, Robledo RF, Schultz IJ, et al. TMEM14C is required for erythroid mitochondrial heme metabolism. *Journal of Clinical Investigation*. 2014-10-01 2014;124(10):4294-4304. doi:10.1172/jci76979

73. Zhou S, Zong Y, Ney PA, Nair G, Stewart CF, Sorrentino BP. Increased expression of the Abcg2 transporter during erythroid maturation plays a role in decreasing cellular protoporphyrin IX levels. *Blood*. 2005-03-15 2005;105(6):2571-2576. doi:10.1182/blood-2004-04-1566

74. Boultwood J, Pellagatti A, Nikpour M, et al. The Role of the Iron Transporter ABCB7 in Refractory Anemia with Ring Sideroblasts. *PLoS ONE*. 2008-04-09 2008;3(4):e1970. doi:10.1371/journal.pone.0001970

75. Liang R, Menon V, Qiu J, et al. Mitochondrial localization and moderated activity are key to murine erythroid enucleation. *Blood Advances*. 2021-05-25 2021;5(10):2490-2504. doi:10.1182/bloodadvances.2021004259

76. Deybach J. Mutations in the protoporphyrinogen oxidase gene in patients with variegate porphyria. *Human Molecular Genetics*. 1996-03-01 1996;5(3):407-410. doi:10.1093/hmg/5.3.407

77. Richardson DR, Lane DJR, Becker EM, et al. Mitochondrial iron trafficking and the integration of iron metabolism between the mitochondrion and cytosol. *Proceedings of the National Academy of Sciences*. 2010-06-15 2010;107(24):10775-10782. doi:10.1073/pnas.0912925107

78. Liu G, Sil D, Maio N, et al. Heme biosynthesis depends on previously unrecognized acquisition of iron-sulfur cofactors in human amino-levulinic acid dehydratase. *Nature Communications*. 2020-12-01 2020;11(1)doi:10.1038/s41467-020-20145-9

79. Hyvärinen AK, Pohjoismäki JLO, Holt IJ, Jacobs HT. Overexpression of MTERFD1 or MTERFD3 impairs the completion of mitochondrial DNA replication. *Molecular Biology Reports*. 2011-02-01 2011;38(2):1321-1328. doi:10.1007/s11033-010-0233-9

80. Lee HY, Gao X, Barrasa MI, et al. PPAR-α and glucocorticoid receptor synergize to promote erythroid progenitor self-renewal. Nature Publishing Group; 2015. p. 474-477.

81. Inoue D, Chew G-L, Liu B, et al. Spliceosomal disruption of the non-canonical BAF complex in cancer. *Nature*. 2019-10-17 2019;574(7778):432-436. doi:10.1038/s41586-019-1646-9

82. Flicek P, Ahmed I, Amode MR, et al. Ensembl 2013. *Nucleic Acids Research*. 2012-11-30 2012;41(D1):D48-D55. doi:10.1093/nar/gks1236

83. Goldman M, Craft B, Swatloski T, et al. The UCSC Cancer Genomics Browser: update 2013. *Nucleic Acids Research*. 2013-01-01 2013;41(D1):D949-D954. doi:10.1093/nar/gks1008

84. Katz Y, Wang ET, Airoldi EM, Burge CB. Analysis and design of RNA sequencing experiments for identifying isoform regulation. *Nature Methods*. 2010-12-01 2010;7(12):1009-1015. doi:10.1038/nmeth.1528

85. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics*. 2011-12-01 2011;12(1):323. doi:10.1186/1471-2105-12-323

86. Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics*. 2009-05-01 2009;25(9):1105-1111. doi:10.1093/bioinformatics/btp120

87. Wagenmakers E-J, Lodewyckx T, Kuriyal H, Grasman R. Bayesian hypothesis testing for psychologists: A tutorial on the Savage–Dickey method. *Cognitive Psychology*. 2010-05-01 2010;60(3):158-189. doi:10.1016/j.cogpsych.2009.12.001

88. Tang Y, Hou J, Li G, et al. ABCG2 regulates the pattern of self-renewing divisions in cisplatin-resistant non-small cell lung cancer cell lines. *Oncology Reports*. 2014-11-01 2014;32(5):2168-2174. doi:10.3892/or.2014.3470

89. Rivella S. Ineffective erythropoiesis and thalassemias. *Current Opinion in Hematology*. 2009-05-01 2009;16(3):187-194. doi:10.1097/moh.0b013e32832990a4

90. Yang Z, Keel SB, Shimamura A, et al. Delayed globin synthesis leads to excess heme and the macrocytic anemia of Diamond Blackfan anemia and del(5q) myelodysplastic syndrome. *Science Translational Medicine*. 2016-05-11 2016;8(338):338ra67-338ra67. doi:10.1126/scitranslmed.aaf3006

91. Gaiti F, Chamely P, Hawkins AG, et al. Single-cell multi-omics defines the cell-type specific impact of splicing aberrations in human hematopoietic clonal outgrowths. *bioRxiv*. 2022:2022.06.08.495292. doi:10.1101/2022.06.08.495292

92. Janusz K, Izquierdo MM, Cadenas FL, et al. Clinical, biological, and prognostic implications of SF3B1 co-occurrence mutations in very low/low- and intermediate-risk MDS patients. *Annals of Hematology*. 2021-08-01 2021;100(8):1995-2004. doi:10.1007/s00277-020-04360-4

93. Qu X, Zhang S, Wang S, et al. TET2 deficiency leads to stem cell factor–dependent clonal expansion of dysfunctional erythroid progenitors. *Blood*. 2018-11-29 2018;132(22):2406-2417. doi:10.1182/blood-2018-05-853291

94. Sachar M, Anderson KE, Ma X. Protoporphyrin IX: the Good, the Bad, and the Ugly. *Journal of Pharmacology and Experimental Therapeutics*. 2016-01-07 2016;356(2):267-275. doi:10.1124/jpet.115.228130

95. An X, Schulz VP, Li J, et al. Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*. 2014-05-29 2014;123(22):3466-3477. doi:10.1182/blood-2014-01-548305

96. Betin VMS, Singleton BK, Parsons SF, Anstee DJ, Lane JD. Autophagy facilitates organelle clearance during differentiation of human erythroblasts. *Autophagy*. 2013-06-06 2013;9(6):881-893. doi:10.4161/auto.24172

97. Mccranor BJ, Kim MJ, Cruz NM, et al. Interleukin-6 directly impairs the erythroid development of human TF-1 erythroleukemic cells. *Blood Cells, Molecules, and Diseases*. 2014-02-01 2014;52(2-3):126-133. doi:10.1016/j.bcmd.2013.09.004

98. Hegde S, Ni S, He S, et al. Stat3 promotes the development of erythroleukemia by inducing Pu.1 expression and inhibiting erythroid differentiation. *Oncogene*. 2009-09-24 2009;28(38):3349-3359. doi:10.1038/onc.2009.202

99. Ryzhkova A, Battulin N. Genome Reorganization during Erythroid Differentiation. *Genes*. 2021-06-30 2021;12(7):1012. doi:10.3390/genes12071012

100. Wang Y, Li W, Schulz VP, et al. Impairment of human terminal erythroid differentiation by histone deacetylase 5 deficiency. *Blood*. 2021-10-28 2021;138(17):1615-1627. doi:10.1182/blood.2020007401

101. Li D, Liu J, Yang X, et al. Chromatin Accessibility Dynamics during iPSC Reprogramming. *Cell Stem Cell*. 2017;21(6):819-833.e6. doi:10.1016/j.stem.2017.10.012

102. Sanders M SS, Dainiak N. Insulin-like growth factors stimulate erythropoiesis in serum-substituted umbilical cord blood cultures. *Exp Hematol*. 1993;21(1):25-30.

103. Xu J, Shao Z, Li D, et al. Developmental Control of Polycomb Subunit Composition by GATA Factors Mediates a Switch to Non-Canonical Functions. *Molecular Cell*. 2015-01-01 2015;57(2):304-316. doi:10.1016/j.molcel.2014.12.009

104. Ward C, Cauchy P, Garcia P, Frampton J, Esteban MA, Volpe G. High WBP5 expression correlates with elevation of HOX genes levels and is associated with inferior survival in patients with acute myeloid leukaemia. *Scientific Reports*. 2020-12-01 2020;10(1)doi:10.1038/s41598-020-60480-x

105. Zeng J, Yi D, Sun W, et al. Overexpression of HOXA9 upregulates NF-κB signaling to promote human hematopoiesis and alter the hematopoietic differentiation potentials. *Cell Regeneration*. 2021-12-01 2021;10(1)doi:10.1186/s13619-020-00066-0

106. Yang Y, Wang C, Zhang P, et al. Polycomb Group Protein PHF1 Regulates p53-dependent Cell Growth Arrest and Apoptosis. *Journal of Biological Chemistry*. 2013-01-01 2013;288(1):529-539. doi:10.1074/jbc.m111.338996

107. Richard A, Vallin E, Romestaing C, Roussel D, Gandrillon O, Gonin-Giraud S. Erythroid differentiation displays a peak of energy consumption concomitant with glycolytic metabolism rearrangements. *PLOS ONE*. 2019-09-04 2019;14(9):e0221472. doi:10.1371/journal.pone.0221472

108. Huang N-J, Lin Y-C, Lin C-Y, et al. Enhanced phosphocholine metabolism is essential for terminal erythropoiesis. *Blood*. 2018-06-28 2018;131(26):2955-2966. doi:10.1182/blood-2018-03-838516

109. Garaycoechea JI, Crossan GP, Langevin F, Daly M, Arends MJ, Patel KJ. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature*. 2012-09-01 2012;489(7417):571-575. doi:10.1038/nature11368

110. Wang M, Dingler FA, Patel KJ. Genotoxic aldehydes in the hematopoietic system. *Blood*. 2022-04-07 2022;139(14):2119-2129. doi:10.1182/blood.2019004316

111. Bagger FO, Kinalis S, Rapin N. BloodSpot: a database of healthy and malignant haematopoiesis updated with purified and single cell mRNA sequencing profiles. *Nucleic Acids Research*. 2019-01-08 2019;47(D1):D881-D885. doi:10.1093/nar/gky1076

112. Monroe GR, Van Eerde AM, Tessadori F, et al. Identification of human D lactate dehydrogenase deficiency. *Nature Communications*. 2019-12-01 2019;10(1)doi:10.1038/s41467-019-09458-6

113. Doty RT, Yan X, Lausted C, et al. Single-cell analyses demonstrate that a heme–GATA1 feedback loop regulates red cell differentiation. *Blood*. 2019-01-31 2019;133(5):457-469. doi:10.1182/blood-2018-05-850412

114. Dalton WB, Helmenstine E, Walsh N, et al. Hotspot SF3B1 mutations induce metabolic reprogramming and vulnerability to serine deprivation. *Journal of Clinical Investigation*. 2019-09-30 2019;129(11):4708-4723. doi:10.1172/jci125022

115. Novershtern N, Subramanian A, Lawton LN, et al. Densely Interconnected Transcriptional Circuits Control Cell States in Human Hematopoiesis. *Cell*. 2011;144(2):296-309. doi:10.1016/j.cell.2011.01.004

116. Sekeres MA, Cutler C. How we treat higher-risk myelodysplastic syndromes. *Blood*. 2014-02-06 2014;123(6):829-836. doi:10.1182/blood-2013-08-496935

117. Fenaux P, Adès L. How we treat lower-risk myelodysplastic syndromes. *Blood*. 2013-05-23 2013;121(21):4280-4286. doi:10.1182/blood-2013-02-453068

118. Fenaux P, Platzbecker U, Mufti GJ, et al. Luspatercept in Patients with Lower-Risk Myelodysplastic Syndromes. *New England Journal of Medicine*. 2020-01-09 2020;382(2):140-151. doi:10.1056/nejmoa1908892

119. Seiler M, Yoshimi A, Darman R, et al. H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nature Medicine*. 2018-04-01 2018;24(4):497-504. doi:10.1038/nm.4493

120. Steensma DP, Wermke M, Klimek VM, et al. Phase I First-in-Human Dose Escalation Study of the oral SF3B1 modulator H3B-8800 in myeloid neoplasms. *Leukemia*. 2021-12-01 2021;35(12):3542-3550. doi:10.1038/s41375-021-01328-9

121. Wang Y-Q, Chang S-Y, Wu Q, et al. The Protective Role of Mitochondrial Ferritin on Erastin-Induced Ferroptosis. Original Research. *Frontiers in Aging Neuroscience*. 2016-December-20 2016;8doi:10.3389/fnagi.2016.00308

122. Invernizzi R, Travaglino E, Della Porta MG, et al. Effects of mitochondrial ferritin overexpression in normal and sideroblastic erythroid progenitors. *British Journal of Haematology*. 2013-06-01 2013;161(5):726-737. doi:10.1111/bjh.12316

123. Cazzola M. Ineffective erythropoiesis and its treatment. *Blood*. 2022-04-21 2022;139(16):2460-2470. doi:10.1182/blood.2021011045

124. Ambaglio I, Malcovati L, Papaemmanuil E, et al. Inappropriately low hepcidin levels in patients with myelodysplastic syndrome carrying a somatic mutation of SF3B1. *Haematologica*. 2013-03-01 2013;98(3):420-423. doi:10.3324/haematol.2012.077446

125. Tosato G, Feng J-X, Ohnuki H, Sim M. Bone marrow niches in myelodysplastic syndromes. *Journal of Cancer Metastasis and Treatment*. 2021-01-01 2021;doi:10.20517/2394-4722.2021.120

126. Kornblau SM, Mccue D, Singh N, Chen W, Estrov Z, Coombes KR. Recurrent expression signatures of cytokines and chemokines are present and are independently prognostic in acute myelogenous leukemia and myelodysplasia. *Blood*. 2010-11-18 2010;116(20):4251-4261. doi:10.1182/blood-2010-01-262071

127. Visconte V, Tabarroki A, Zhang L, et al. Splicing factor 3b subunit 1 (Sf3b1) haploinsufficient mice display features of low risk Myelodysplastic syndromes with ring sideroblasts. BioMed Central Ltd.; 2014. p. 89.

128. Wang C, Sashida G, Saraya A, et al. Depletion of Sf3b1 impairs proliferative capacity of hematopoietic stem cells but is not sufficient to induce myelodysplasia. *Blood*. 2014-05-22 2014;123(21):3336-3343. doi:10.1182/blood-2013-12-544544

129. Matsunawa M, Yamamoto R, Sanada M, et al. Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia. *Leukemia*. 2014-09-01 2014;28(9):1844-1850. doi:10.1038/leu.2014.73

130. William Roy S, Gilbert W. The evolution of spliceosomal introns: patterns, puzzles and progress. *Nature Reviews Genetics*. 2006/03/01 2006;7(3):211-221. doi:10.1038/nrg1807

131. Jin S, Su H, Tran N-T, et al. Splicing factor SF3B1K700E mutant dysregulates erythroid differentiation via aberrant alternative splicing of transcription factor TAL1. *PLOS ONE*. 2017-05-18 2017;12(5):e0175523. doi:10.1371/journal.pone.0175523

132. Rongvaux A, Willinger T, Martinek J, et al. Development and function of human innate immune cells in a humanized mouse model. *Nat Biotechnol*. Apr 2014;32(4):364-72. doi:10.1038/nbt.2858

133. Bak RO, Dever DP, Reinisch A, Cruz Hernandez D, Majeti R, Porteus MH. Multiplexed genetic engineering of human hematopoietic stem and progenitor cells using CRISPR/Cas9 and AAV6. *eLife*. 2017-09-28 2017;6doi:10.7554/elife.27873

134. Ferrari S, Jacob A, Beretta S, et al. Efficient gene editing of human long-term hematopoietic stem cells validated by clonal tracking. *Nature Biotechnology*. 2020/11/01 2020;38(11):1298-1308. doi:10.1038/s41587-020-0551-y

135. Antony JS, Latifi N, Haque AKMA, et al. Gene correction of HBB mutations in CD34+ hematopoietic stem cells using Cas9 mRNA and ssODN donors. *Molecular and Cellular Pediatrics*. 2018-12-01 2018;5(1)doi:10.1186/s40348-018-0086-1

136. Clem B, Telang S, Clem A, et al. Small-molecule inhibition of 6-phosphofructo-2-kinase activity suppresses glycolytic flux and tumor growth. *Molecular Cancer Therapeutics*. 2008-01-01 2008;7(1):110-120. doi:10.1158/1535-7163.mct-07-0482