From the Department of Medicine

Karolinska Institutet, Stockholm, Sweden

GENOTYPE-PHENOTYPE RELATIONS IN *SF3B1* MUTATED MYELODYSPLASTIC SYNDROMES WITH RING SIDEROBLASTS

Simona Conte

Karolinska Institutet

Stockholm 2016

Published by Karolinska Institutet.
Printed by Eprint AB 2016

ISBN 978-91-7676-439-8

Front cover: Watercolor by Simona Conte, 2016

GENOTYPE-PHENOTYPE RELATIONS IN *SF3B1* MUTATED MYELODYSPLASTIC SYNDROMES WITH RING SIDEROBLASTS

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Simona Conte

Principal Supervisor:
Professor Eva Hellström-Lindberg, M.D., Ph.D.
Karolinska Institutet
Department of Medicine
Center for Hematology and Regenerative Medicine

Co-supervisor(s):
Docent Julian Walfridsson, Ph.D
Karolinska Institutet
Department of Medicine
Center for Hematology and Regenerative Medicine

Liselotte Vesterlund, M.D., Ph.D.
Karolinska Institutet
Department of Biosciences and Nutrition

Opponent:
Professor David P. Steensma, M.D.
Harvard Medical School, Boston, MA, USA
Department of Medicine
Division of Hematology/Oncology

Examination Board:
Docent Johan Flygare, M.D., Ph.D.
Lund University
Department of Molecular Medicine and Gene Therapy

Docent Martin Höglund, M.D.
Uppsala University
Department of Medical Sciences

Professor Martin Bergö, M.D., Ph.D.
Karolinska Institutet
Department of Biosciences and Nutrition

Διά τὸ θαυμάζειν ἡ σοφία

“Wisdom begins in wonder”

Plato, Theaetetus, ca. 369 BC

To mom and dad

ABSTRACT

Myelodysplastic syndrome with ring sideroblasts (MDS-RS) is a clonal hematopoietic stem cell disorder characterized by hyperplastic and ineffective erythropoiesis, ring sideroblasts in the bone marrow, and anemia. Heterozygous mutations in the spliceosome gene *SF3B1* are found in a majority of MDS-RS cases and are associated with the ring sideroblast phenotype and a favorable prognosis. MDS-RS is a usually slowly progressing disease of the elderly distinguished by an initial phase of erythroid hyperplasia and macrocytic anemia that can remain stable for several years, but which usually evolve into progressive erythroid failure and a need for regular red blood cell transfusions.

The focus of this thesis was to study the genotype-phenotype relations in MDS-RS patients in order to understand the abnormal iron accumulation and erythroid failure, and in particular the role of *SF3B1* in MDS pathogenesis.

In study I, we tested whether *ABCB7* was a key mediator of the aberrant iron accumulation, by modulating its expression in normal bone marrow. *ABCB7* down-regulation reduced erythroid differentiation, growth and colony formation, and resulted in a gene expression pattern similar to that observed in MDS-RS erythroblasts, and in the accumulation of mitochondrial ferritin. Conversely, forced *ABCB7* expression restored erythroid colony growth and decreased mitochondrial ferritin in MDS-RS $CD34^+$ progenitor cells. Also, we showed altered exon usage of *ABCB7* as possible explanation of the reduced expression in MDS-RS and a potential link of this gene with *SF3B1*.

In study II we sought to identify potential downstream targets of *SF3B1* mutations and understand how these affect RNA splicing and gene expression profile of MDS-RS patients. In particular, we detected a significant up-regulation of genes involved in hemoglobin synthesis and in the oxidative phosphorylation process, and down-regulation of mitochondrial ABC transporters compared to normal bone marrow. These findings together with mis-splicing of hemoglobin genes indicated a compromised hemoglobinization during MDS-RS erythropoiesis. Importantly, we demonstrated that anemia in MDS-RS patients develops during terminal differentiation into reticulocytes.

Ultimately in study III, we investigated the mechanistic effects of *SF3B1* mutation in MDS-RS pathogenesis and explored if the most frequent *SF3B1* mutation K700E confers a loss-of-function with regard to gene expression and splicing of key genes in MDS-RS. Loss or reduction of SF3B1 normal protein in human myeloid cells resulted unexpectedly in an MDS-RS-like phenotype with reduced *ABCB7* expression and altered exon usage, increased levels of *ALAS2* and mis-splicing of *TMEM14C*. Identical effects were observed when we expressed *SF3B1* K700E mutation at physiological levels. Additionally, loss of *SF3B1* compromised cell growth but did not increase apoptosis. Overall, our findings support an essential role of *ABCB7* in the MDS-RS phenotype and offer insights into the mechanistic role of *SF3B1* mutations. Via altered gene expression or mis-splicing of key genes in the heme and hemoglobin synthesis, these mutations disturb mitochondrial iron handling in a way that lead to mitochondrial iron accumulation in MDS-RS.

LIST OF PUBLICATIONS

I. Nikpour M.*, Scharenberg C.*, Liu A., **Conte S.**, Karimi M., Mortera-Blanco T., Giai V., Fernandez-Mercado M., Papaemmanuil E., Hogstrand K., Jansson M., Vedin I., Wainscoat J.S., Campbell P., Cazzola M., Boultwood J., Grandien A., Hellström-Lindberg E. (2013). The transporter ABCB7 is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts. *Leukemia,* 27, 889-896.

II. **Conte S.**, Katayama S., Vesterlund L., Karimi M., Dimitriou M., Jansson M., Mortera-Blanco T., Unneberg P., Papaemmanuil E., Sander B., Skoog T., Campbell P., Walfridsson J., Kere J., Hellström-Lindberg E. (2015). Aberrant splicing of genes involved in haemoglobin synthesis and impaired terminal erythroid maturation in SF3B1 mutated refractory anaemia with ring sideroblasts. *Br J Haematol.* 2015 Nov;171(4):478-90

III. **Conte S.**, Heshmati Y., Harisankar A., Dimitriou M., Jansson M., Vesterlund L., Elvarsdottir E., Mortera-Blanco T., Larsson J., Shiozawa Y., Ogawa S., Karimi M., Walfridsson J., Hellström-Lindberg E. Mechanistic role of *SF3B1* in MDS-RS. *Manuscript*

* Authors contributed equally to this study

Related publications not included in the thesis

I. Malcovati L.*, Papaemmanuil E.*, Bowen D.T., Boultwood J., Della Porta M.G., Pascutto C., Travaglino E., Groves M.J., Godfrey A.L., Ambaglio I., Galli A., Da Via M.C., **Conte S.**, Tauro S., Keenan N., Hyslop A., Hinton J., Mudie L.J., Wainscoat J.S., Futreal P.A., Stratton M.R., Campbell P.J., Hellström-Lindberg E., Cazzola M. (2011). Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myelo-proliferative neoplasms. *Blood,* 118, 6239-6246

II. Malcovati L., Karimi M., Papaemmanuil E., Ambaglio I., Jädersten M., Jansson M., Elena C., Gallì A., Walldin G., Della Porta M.G., Raaschou-Jensen K., Travaglino E., Kallenbach K., Pietra D., Ljungström V., **Conte, S.**, Boveri E., Invernizzi R., Rosenquist R., Campbell P.J., Cazzola M., Hellström Lindberg E. (2015). *SF3B1* mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood,* 9;126(2):233-41

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ABCB7	ATP Binding Cassette B7
ALA	Aminolevulinic Acid
ALAS2	Aminolevulinic Acid Synthase 2
AML	Acute Myeloid Leukemia
ATG	Anti-Thymocyte Globulin
BFU-E	Burst Forming Unit-Erythroid
BM	Bone Marrow
CFU-E	Colony Forming Unit-Erythroid
CLL	Chronic Lymphocytic Leukemia
CMML	Chronic Myelomonocytic Leukemia
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
DMT1	Divalent Metal Transporter 1
EPO	Erythropoietin
ESA	Erythropoiesis Stimulating Agents
ET	Essential Thrombocythemia
FAB	French American British
Fe-S	Iron-Sulphur
FPKM	Fragments Per Kilobase of Exon per Million of Fragments Mapped
FTMT	Mitochondrial Ferritin
G-CSF	Granulocyte Colony-Stimulating Factor
GFP	Green Fluorescent Protein
GLRX5	Glutaredoxin 5
GMP	Granulocytes Macrophage Progenitors
GPA	Glycophorin A
GSEA	Gene Set Enrichment Analysis
HEAT	Huntingtin Elongation Factor EF2 Protein Phosphatase PP2A yeast kinase TOR1
HMA	Hypomethylating Agents
HSC	Hematopoietic Stem Cells
IL-3	Interleukin-3
IPSS	International Prognostic Scoring System
IRP	Iron Regulatory Proteins
JAK	Janus Kinases
MAPK	Mitogen-Activated Protein Kinases
MDS	Myelodysplastic Syndromes
MDS-RS	Myelodysplastic Syndromes with Ring Sideroblasts

MDS/MPN	Myelodysplastic Syndromes/Myeloproliferative Neoplasms
MNC	Mononuclear Cells
mTOR	Mechanistic Target Of Rapamycin
NBM	Normal Bone Marrow
NMD	Nonsense Mediated Decay
NSG	Nod Scid Gamma
PAM	Protospacer Adjacent Motif
PB	Peripheral Blood
PMF	Primary Myelofibrosis
qPCR	Quantitative Real Time PCR
RAEB	Refractory Anemia with Excess of Blasts
RARS	Refractory Anemia with Ring Sideroblasts
RARS-T	Refractory Anemia with Ring Sideroblast with Thrombocytosis
RBC	Red Blood Cell
RCMD	Refractory Cytopenia with Multilineage Dysplasia
ROS	Reactive Oxygen Species
RS	Ring Sideroblasts
SA	Sideroblastic Anemia
SCF	Stem Cell Factor
SF1	Splicing Factor 1
SF3B1	Splicing Factor 3b subunit 1
shRNA	short hairpin RNA
SLC25A37	Mitoferrin-1
snRNP	small nuclear Ribonucleic Protein
SOLiD	Sequencing by Oligo Ligation Detection
SS	Splice Site
STAT	Signal Transducer and Activator of Transcription
STEAP3	Ferroreductase 3
TF	Transferrin
TFCP2	Transcription Factor CP2
TGF-β	Transforming Growth Factor β
TLDA	Taqman Low Density Array
TMEM14C	Transmembrane protein 14C
TPO	Thrombopoietin
WHO	World Health Organization
WT	Wild-Type
XLSA	X-Linked Sideroblastic Anemia

1 INTRODUCTION

1.1 ERYTHROPOIESIS

The process of erythropoiesis entails the generation of mature red blood cells from multipotent stem cells. The erythrocyte is the most common cell type in blood. Mature erythrocytes have a life span of approximately 120 days and need to be continuously produced in order to renew the red cell mass. In mammals, definitive erythropoiesis first occurs in the fetal liver with progenitor cells from the yolk sac (Palis *et al.* 1999). Within the fetal liver and the adult bone marrow, hematopoietic cells are formed continuously from a small population of pluripotent stem cells that generate progenitors committed to one or a few hematopoietic lineages (Hattangadi *et al.* 2011).

The erythroid lineage consists of erythroid progenitors and precursors. Erythroid progenitors can be divided in early or late progenitors based on their colony forming capacity *in vitro*. The earliest committed progenitors, the burst forming unit-erythroid (BFU-E), are the first solely erythroid restricted cells and give rise to multi-clusters colonies (or bursts) comprising 200 to many thousands of erythroblasts (Testa 2004). The late progenitors are the colony forming unit-erythroid (CFU-E), more differentiated than the BFU-E and able to generate smaller colonies in one or few separate clusters. Importantly, BFU-E and CFU-E can be distinguished based on their differential growth factor requirements. The proliferation and survival of BFU-E is dependent on the stem cell factor (SCF) and interleukin-3 (IL-3) signaling, while CFU-E progenitors are exclusively dependent on erythropoietin (EPO). CFU-Es differentiate into morphologically distinguishable erythroid precursor, which in contrast to the erythroid progenitors, are less EPO responsive, but strongly iron dependent (Camaschella *et al.* 2016). The first recognizable precursors, proerythroblasts, undergo 3-5 cell divisions and give rise to basophilic, polychromatic and orthochromatic erythroblasts. These cell divisions are characterized by a rapid G1 cell phase resulting in a progressive decrease in cell size (von Lindern 2006). Simultaneously, maturating precursors undergo to substantial changes in morphology due to the accumulation of erythroid-specific proteins and chromatin condensation (Hattangadi *et al.* 2011). Then, orthochromatic erythroblasts form reticulocytes by extruding their nuclei and reticulocytes lose their mitochondria and ribosomes within a couple of days before finally maturing into erythrocytes (Testa 2004) (Figure 1).

Each stage of the erythroid maturation process can be distinguished by the expression of

surface markers/receptors, transcription factors and effector molecules. The main hormone involved in erythropoiesis is EPO, which interacts with cells carrying its receptor EpoR, i.e. all cells of the erythroid lineage except for reticulocytes. SCF is the ligand for the c-kit receptor, which it is present on a large spectrum of hematopoietic stem and progenitors cells. Both BFU-E and CFU-E exhibit high levels of c-kit (or CD117) and the levels decrease at the proerythroblast stage. The expression of glycoprotein CD36, transferrin receptor (TFR or CD71) and glycophorin A (GPA or CD235a) divide erythroid precursors into three erythroblasts populations: basophilic/polychromatic (CD36+ CD71+ GPA+), orthochromatic (CD36- CD71+ GPA+) and reticulocytes/red blood cells (CD36- CD71- GPA+) (Merryweather-Clarke *et al.* 2011). GATA1 is the main transcription factor driving terminal erythroid differentiation, while GATA2 is highly expressed in hematopoietic stem and progenitor cells. Finally, the hemoglobin production that starts gradually at the erythroblast stage reaches its maximum expression at the erythrocyte stage (Figure 1).

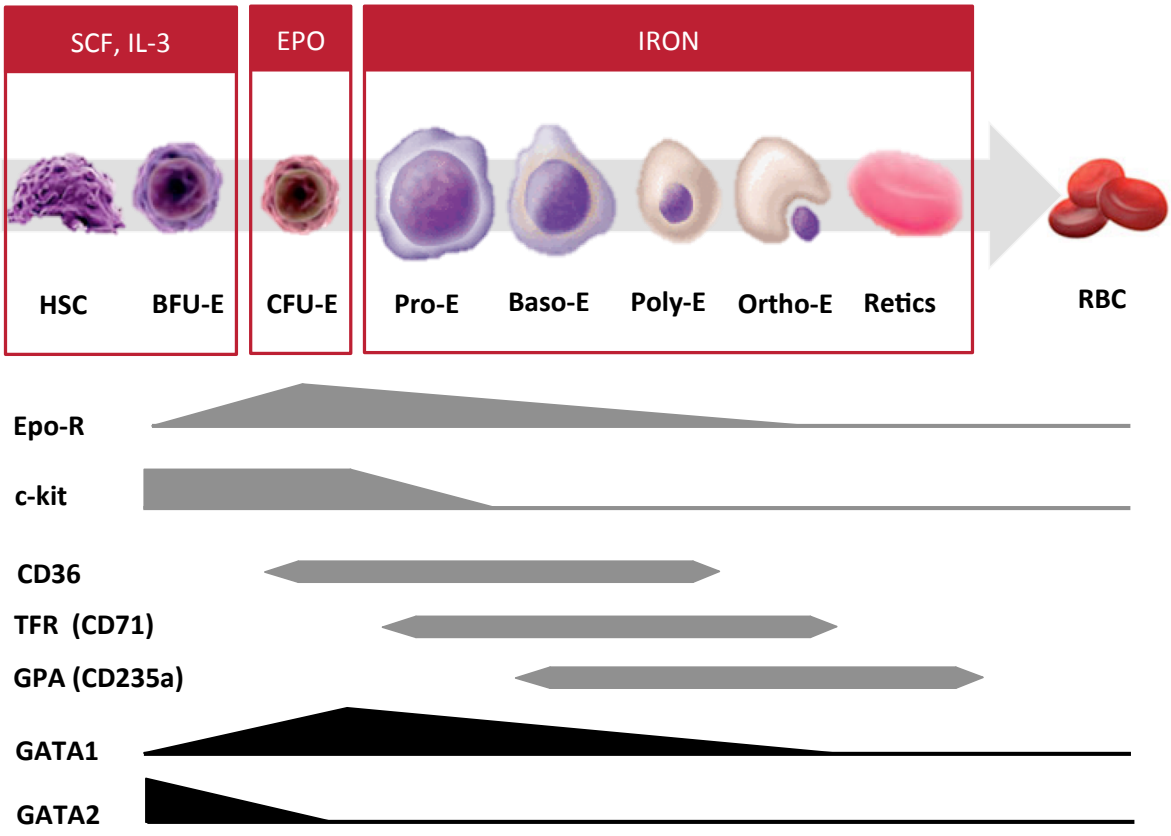

Figure 1 Outline of the erythroid differentiation. SCF, IL-3, EPO are the growth factors required during the erythroid maturation process. Each stage is distinguished by the expression of surface marker/receptors (grey) and transcription factors (black).

1.2 IRON METABOLISM AND HEME SYNTHESIS

The second phase of the erythropoiesis from proerythroblasts to erythrocytes is strongly iron dependent. Iron is essential for heme and iron-sulfur cluster synthesis in every cell of the body but is required in larger amounts for hemoglobin synthesis in maturing erythroblasts (Camaschella *et al.* 2016). To produce about 200 billion erythrocytes every day, approximately 25 mg of iron must be supplied to the bone marrow. The vast majority is provided by macrophages, who recycle hemoglobin-derived iron from the breakdown of senescent erythrocytes, while in the absence of anemia, <1–2 mg iron daily derives from intestinal absorption (Camaschella *et al.* 2011). As we recycle much more iron than what we absorb, iron has a very conservative metabolism.

Iron enters the body through the diet and most iron absorption takes place in the duodenum (Figure 2). In the enterocytes, iron is solubilized and converted from trivalent into bivalent by the duodenal cytochrome B (DcytB). Iron is transported across the enterocyte's cell membrane by the divalent metal transporter 1 (DMT1) into the cell where heme is degraded by heme oxygenase and iron is either stored in cellular ferritin or released into the portal circulation by the ferroportin (Figure 2A). The hepatocyte serves as the main storage for the iron surplus, but most body iron is present in erythrocytes and macrophages. In the hepatocyte, hepcidin controls and regulates the entry of iron in the circulation (Figure 2B). The macrophage recycles hemoglobin-derived iron from senescent erythrocytes. This iron can either be stored in the macrophage as hemosiderin or ferritin or be delivered to the erythroid progenitor as an ingredient for new erythrocytes (Figure 2C). The iron exporter ferroportin is responsible for the efflux of Fe into the circulation. In the erythroid progenitor, transferrin saturated with 2 iron molecules is endocytosed via the transferrin receptor 1. After endocytosis, the iron is released from transferrin, converted from trivalent to divalent Fe by the ferroreductase STEAP3 and transported to the cytosol by DMT1, where it is available mainly for the heme synthesis (Figure 2D) (Donker *et al.* 2014). Iron is transported within the serum bound to transferrin (TF). Cellular Fe homeostasis is maintained by iron regulatory proteins (IRP1 and IRP2) that recognize the iron responsive elements (IREs) located in untranslated regions of a subset of mRNA involved in Fe homeostasis (Hentze *et al.* 2010). Within the mitochondrion iron can be used to synthesize heme and Fe-S clusters, or incorporated into mitochondrial ferritin (Levi *et al.* 2001).

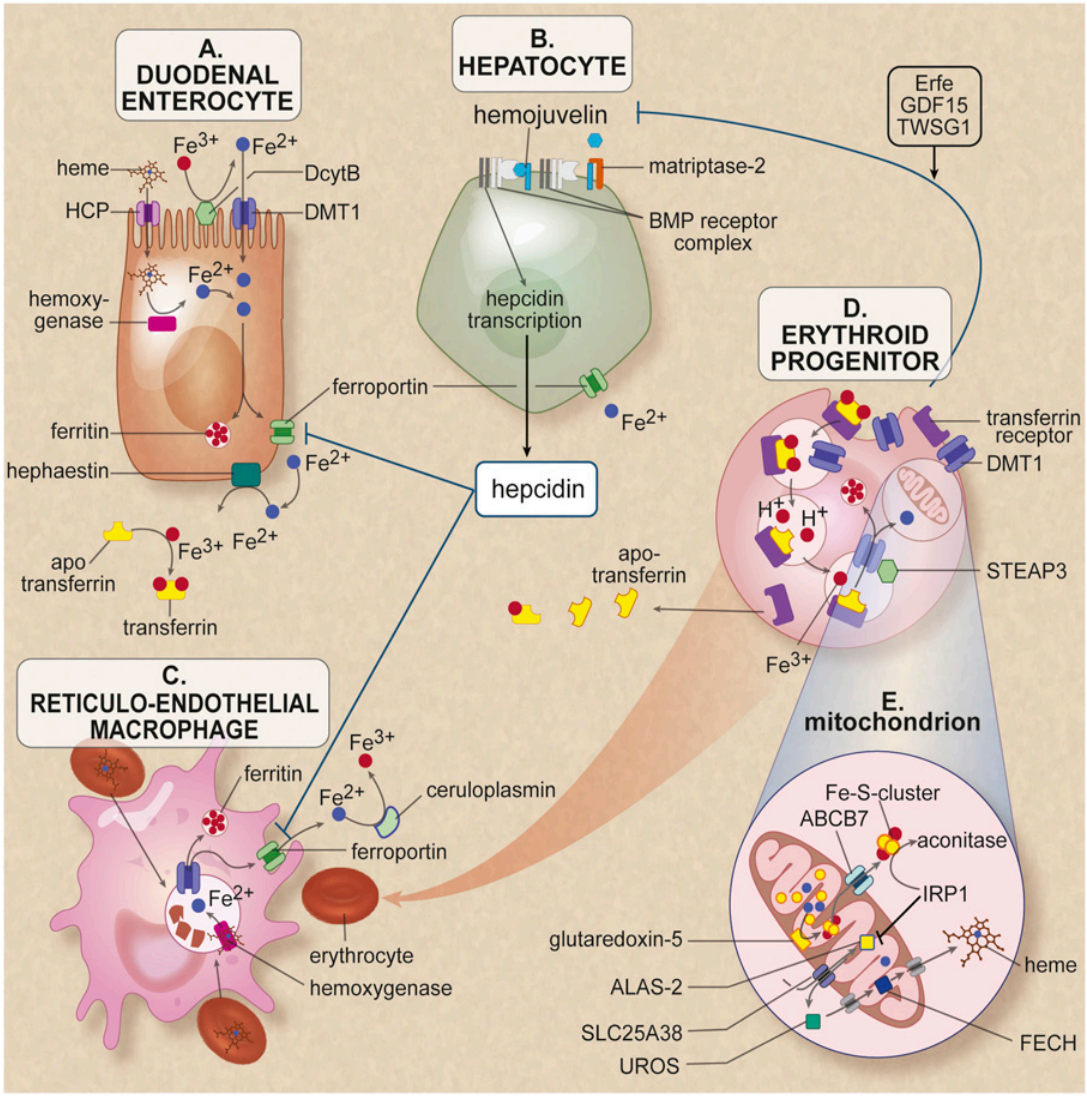

Figure 2 Iron homeostasis and heme synthesis. Figure reprinted with permission from the publisher (Donker *et al.* 2014)

Heme synthesis in erythroid cells is finely coordinated with that of alpha (α) and beta (β)-globin, resulting in the production of hemoglobin, and heme as the prosthetic group. Heme is not only the structural component of hemoglobin, but it plays multiple regulatory roles during the differentiation of erythroid precursors since it controls its own synthesis and regulates the expression of several erythroid-specific genes (Chiabrando *et al.* 2014). Heme is synthesized in developing erythroid progenitors by the stage of proerythroblast, through a series of eight enzymatic reactions divided between mitochondria and cytosol. The heme synthesis begins with the formation of the aminolevulinic acid (ALA) from glycin and succinyl-CoA, catalyzed by ALAS2 in the mitochondrial matrix (Figure 2E). ALA is exported to the cytosol to be converted into porphobilinogen, hydroxymethylbilane, uroporphyrinogen III and finally coporphyrinogen III through a series of enzymatic reactions (Figure 3).

Coporphyrinogen III is transferred into mitochondrial intermembrane space to where it is modified to form protoporphyrinogen IX. Fe is inserted into protoporphyrinogen IX by ferrochetalase (FECH) to produce heme (Chiabrando *et al.* 2014). Fe-S clusters are required in the heme synthesis when a large amount of heme is needed for hemoglobinazation. The enzyme glutaredoxin-5 (GLRX5) plays a role in the synthesis of the Fe-S clusters, which are transported via the transmembrane transporter ABCB7 (Donker *et al.* 2014), however its substrate is unknown (Bekri *et al.* 2000). Fe-S clusters serve as cofactors for molecules involved in electron transfer, catalysis, redox reactions, ribosome assembly, DNA damage repair, telomere maintenance, DNA replication and other processes (Lill *et al.* 1999).

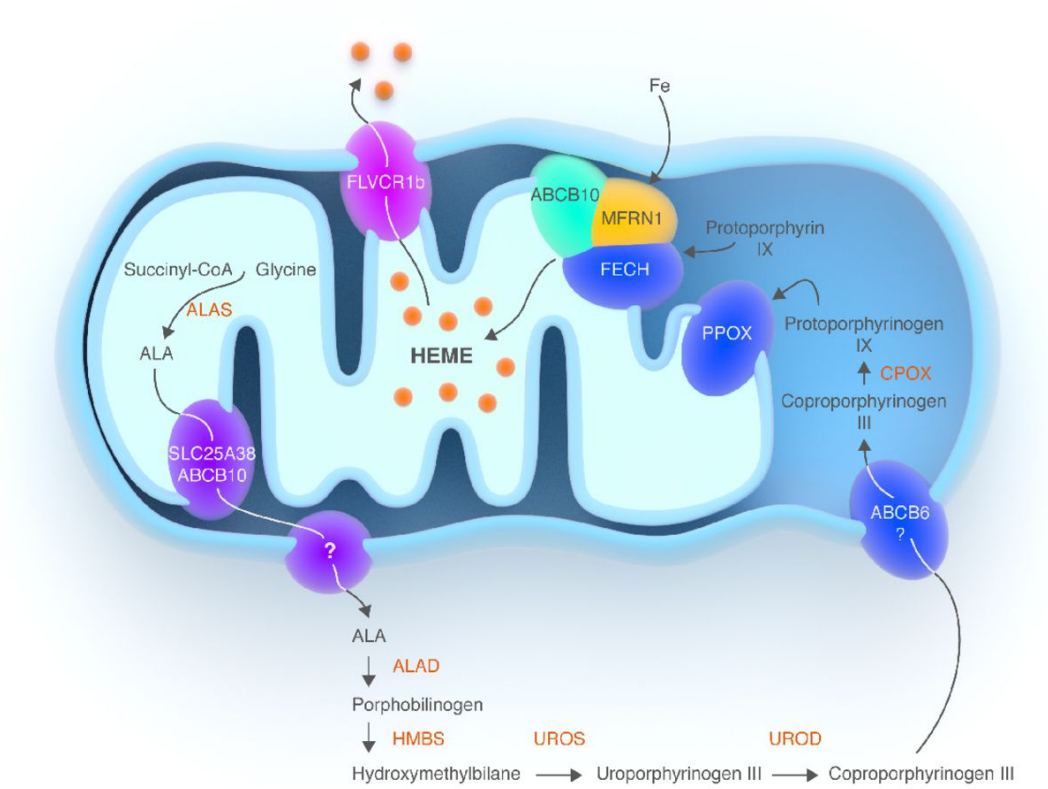

Figure 3 Schematic representation of the heme biosynthetic pathway. Heme synthesis starts in the mitochondrion with the condensation of succynil-CoA and glycine to form ALA, catalyzed by ALAS. There are two isoforms of ALAS*, ALAS1* and *ALAS2*, *ALAS1* is ubiquitously expressed and provides heme in non-erythroid tissues while *ALAS2* gene is located is expressed exclusively in erythroid cells. ALA is transported into the cytosol where it is converted to coporphyrinogen III through a series of enzymatic reactions. Then, coporphyrinogen III is transferred into mitochondrial intermembrane space to where it is modified to form protoporphyrinogen IX. Finally, ferrous iron is incorporated into PPIX to form heme in the mitochondrial matrix, a reaction catalyzed by FECH. Figure adapted from (Chiabrando *et al.* 2014) and reprinted with permission from Haematologica Journal.

1.3 MYELODYSPLASTIC SYNDROMES

The term myelodysplastic syndromes (MDS) encompass a heterogeneous group of myeloid neoplasms characterized by ineffective clonal hematopoiesis and dysplasia in one or more cell lineages, resulting in peripheral blood cytopenia(s). MDS is a disease of the elderly, with a median age of 75 years and with less than 10% of the patients being younger than 50 years of age. The incidence in Europe is 4 cases/100 000 inhabitants/year with a slight dominance of men (Greenberg *et al.* 2012) and no known ethnic differences; however, MDS in Asia tend to occur at an earlier age (Fenaux *et al.* 2014).

MDS pathophysiology is a multistep process involving cytogenetic changes, somatic mutations, and epigenetic alterations (Jiang *et al.* 2009). Ninety percent of MDS are idiopathic and approximately 10% of cases occur secondary to previous treatment with cytostatic drugs or radiotherapy (Neukirchen *et al.* 2011). Epidemiological studies identified occupational exposure to solvents or agricultural chemicals and tobacco use as risk factors; however these factors have a poor association with the disease (Nisse *et al.* 1995, Strom *et al.* 2005). Interestingly, the risk of MDS increases in different bone marrow failure syndromes such as Fanconi Anemia, Diamond-Blackfan anemia and dyskeratosis congenital (Holme *et al.* 2012, West *et al.* 2014). Importantly, the MDS classification from 2008 has recently been revised and the 2016 update contains several changes (i.e. refinements in morphologic interpretation, cytopenia assessment and genetic information), not the least with importance for MDS with ring sideroblasts, Table 1, (Arber *et al.* 2016).

The clinical presentation and the natural course of MDS vary between subgroups and individuals. Symptoms are often related to the cytopenia(s), in the vast majority represented by anemia, often macrocytic, but also leukopenia and/or thrombocytopenia. Half of the patients have severe anemia (<100 g/L) at diagnosis (Greenberg *et al.* 1997) and anemia related symptoms such as fatigue, dyspnea and headache. Other symptoms include infections and bleedings as a result of leukopenia and thrombocytopenia. A minority of patients suffers for autoimmune manifestations such as arthritis, pulmonary effusions and vasculitis (Billstrom *et al.* 1995, Marisavljevic *et al.* 2006). Approximately 30% of the patients will progress to acute myeloid leukemia, AML (Greenberg *et al.* 1997).

Table 1 The 2016 revision to the WHO classification of MDS

Name	Dysplastic lineages	Cytopenias*	RS as % of marrow erythroid elements	BM and PB blasts	Cytogenetics by conventional karyotype analysis
MDS with single lineage dysplasia (MDS-SLD)	1	1 or 2	<15%/<5%**	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS with multilineage dysplasia (MDS-MLD)	2 or 3	1-3	<15%/<5%**	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS with ring sideroblasts (MDS-RS)					
MDS-RS with single lineage dysplasia (MDS-RS-SLD)	1	1 or 2	≥15%/≥5%**	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS-RS with multilineage dysplasia (MDS-RS-MLD)	2 or 3	1-3	≥15%/≥5%**	BM <5%, PB <1%, no Auer rods	Any, unless fulfills all criteria for MDS with isolated del(5q)
MDS with isolated del(5q)	1-3	1-2	None or any	BM <5%, PB <1%, no Auer rods	del(5q) alone or with 1 additional abnormality except −7 or del(7q)
MDS with excess blasts (MDS-EB)					
MDS-EB-1	0-3	1-3	None or any	BM 5%-9% or PB 2%-4%, no Auer rods	Any
MDS-EB-2	0-3	1-3	None or any	BM 10%-19% or PB 5%-19% or Auer rods	Any
MDS, unclassifiable (MDS-U)					
with 1% blood blasts	1-3	1-3	None or any	BM <5%, PB = 1%, ‡no Auer rods	Any
with single lineage dysplasia and pancytopenia	1	3	None or any	BM <5%, PB <1%, no Auer rods	Any
based on defining cytogenetic abnormality	0	1-3	<15% §	BM <5%, PB <1%, no Auer rods	MDS-defining abnormality
Refractory cytopenia of childhood	1-3	1-3	None	BM <5%, PB <2%	Any

(*****) Cytopenias defined as: hemoglobin, <10 g/dL; platelet count, <100 × 10^9/L; and absolute neutrophil count, <1.8 × 10^9/L. Rarely, MDS may present with mild anemia or thrombocytopenia above these levels. PB monocytes must be <1 × 10^9/L; (**) If *SF3B1* mutation is present; (**‡**) One percent PB blasts must be recorded on at least 2 separate occasions; (**§**) Cases with ≥15% ring sideroblasts by definition have significant erythroid dysplasia, and are classified as MDS-RS-SLD.

MDS diagnosis is based on peripheral blood counts including a full differential count and cytomorphology, bone marrow aspiration and biopsy, and cytogenetics of bone marrow cells. The bone marrow in MDS patients is typically hypercellular with significant dysplasia (≥10% of all nucleated cells) in at least one of the hemopoietic lineages (i.e. erythroid, myeloid or megakaryocytic). The percentage of bone marrow blasts may be increased but should not amount to 20% or more. Morphological features of dysplasia(s) are summarized in Table 2. Cytogenetic abnormalities have an important impact on therapeutic decision-making in individual patients (Pozdnyakova *et al.* 2008). Karyotypic abnormalities are seen in approximately 50% of patients (Malcovati *et al.* 2013). Unbalanced chromosomal abnormalities reflecting gain or loss of chromosomal material are more prevalent in comparison to AML, while balanced translocations are rare. The most frequent abnormality is the deletion of the long arm of chromosome 5, with or without additional karyotypic abnormalities (Schanz *et al.* 2012). Finally, laboratory values such as lactate dehydrogenase (LDH), ferritin, transferrin, reticulocytes counts, haptoglobin, vitamin B12, folate concentrations, endogenous erythropoietin (EPO) and creatinine levels, are important to support or exclude the diagnosis of MDS from iron deficiency anemia, hemolytic anemia or vitamin B12 or folate deficiency (Fenaux *et al.* 2014).

Due to the variable natural history of MDS, risk stratification is critical for both prognostic assessment and treatment goals. Prognosis is based on the revised International Prognostic Scoring System (IPSS-R) (Greenberg *et al.* 2012), although the conventional IPSS (Greenberg *et al.* 1997) still remains widely used. Based on the percentage of bone marrow blasts, number of cytopenias and karyotype, the IPSS stratifies patients in four risk categories: low, intermediate-1 (Int-1), intermediate 2 (Int-2), and high.

Table 2 Morphological features of dysplasia in MDS

Dyserythropoiesis	**Dysgranulopoiesis**	**Dysmegakaryocytopoiesis**
Nuclear budding	Small or unusual large size	Micromegakaryocytes
Hyperlobation	Nuclear hypolobation (pseudo Pelger Huet anomaly)	Nuclear hyperlobation
Internuclear bridging	Irregular hypersegmentation	Multinucleation
Karyorrexis	Decreased granules	
Multinuclearity	Agranularity	
Megaloblastic changes	Pseudo Chediak-Higashi granules	
Ring Sideroblasts	Auer rods	
Vacuolization		
Periodic acid-Schiff positivity		

1.4 MYELODYSPLASTIC SYNDROMES WITH RING SIDEROBLASTS

The MDS umbrella includes a subgroup called MDS with ring sideroblasts (MDS-RS), initially described by Bjorkman (Bjorkman 1956) in 1956 and subsequently categorized as idiopathic acquired sideroblastic anemia in the French American British (FAB) classification (Bennett *et al.* 1982). The same subgroup of disease was later defined as refractory anemia with ring sideroblasts (RARS) or refractory cytopenia with multi-lineage dysplasia and ring sideroblasts (RCMD-RS) by the International Agency for Research and Cancer in 2011 (www.iarc.fr 9982/3) (Jaffe *et al.* 2001). According to the WHO 2016 Classification, if an SF3B1 mutation is identified, a diagnosis of MDS-RS may be made if ring sideroblasts comprise as few as 5% of nucleated erythroid cells, whereas at least 15% ring sideroblasts are still required in cases lacking *SF3B1* mutation (Arber *et al.* 2016). As shown in Table 1, MDS-RS subgroup is subdivided into cases with single lineage dysplasia, MDS-SLD (previously classified as RARS) and cases with multi-lineage dysplasia, MDS-MLD (previously classified as RCMD).

1.4.1 MORPHOLOGICAL AND CLINICAL FEATURES OF MDS-RS

MDS-RS is a low-risk subtype of myelodysplastic syndromes and it accounts for around 10% of all MDS cases. Ninety percent of MDS-RS patients have no cytogenetic abnormalities, while 10% may have various but not high-risk aberrations (Komrokji *et al.* 2003).

Ring sideroblasts (RS) are erythroid precursors with iron-loaded mitochondria visualized as a perinuclear ring of blue granules by Prussian blue staining (Cazzola *et al.* 2003). A cell is defined as ring sideroblast, when there is a minimum of 5 siderotic granules covering at least a third of the nuclear circumference (Mufti *et al.* 2008) (Figure 4A). Ring sideroblasts are found exclusively in pathological conditions and should not be confused with ferritin sideroblasts (Figure 4B), which are present in normal bone marrow. These latter are normal erythroblasts that, after Prussian blue staining, show a few blue granules scattered in the cytoplasm, representing endosomes or siderosomes, filled with excess iron not utilized for heme synthesis (Harrison *et al.* 1996). Although the presence of ring sideroblasts is one of the hallmarks of MDS-RS, none of the liquid culture systems reported so far have been able to reproduce them *in vitro*.

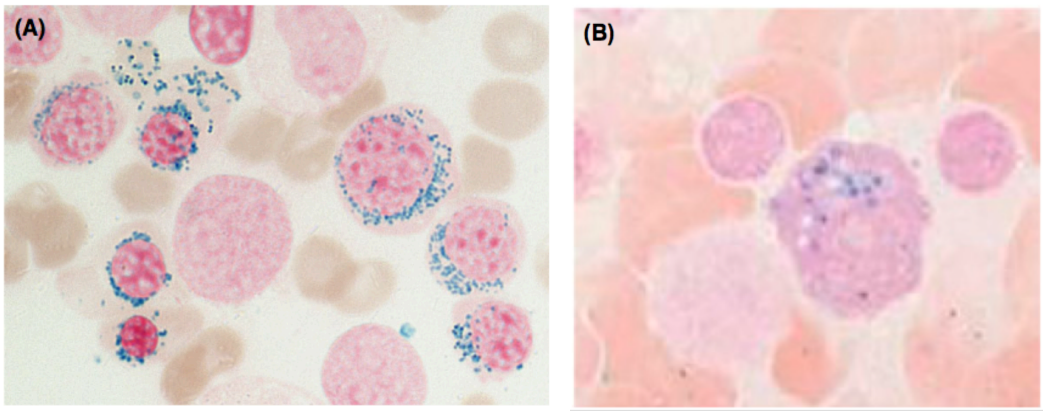

Figure 4 (A) Bone marrow ring sideroblasts in MDS-RS. Prussian blue staining shows that most erythroblasts have five or more positive granules disposed in a ring surrounding the nucleus and covering at least one-third of nuclear circumference. (B) Ferritin sideroblasts. Images A and B are adapted with permission from the publisher (Malcovati *et al.* 2016) (Cazzola *et al.* 2013a), respectively.

In MDS-RS, aberrant iron accumulation in erythroblasts is present in the form of mitochondrial ferritin (FTMT), while normal erythroblasts lack this protein (Cazzola *et al.* 2003). Data suggested that the expression of FTMT was a response to mitochondrial iron accumulation, but the finding that bone marrow cells of MDS-RS patients start to express FTMT in the early stages of erythroid differentiation, before evident iron accumulation, indicated a more complex mechanism of gene regulation (Invernizzi *et al.* 2013), which is not completely understood. The fact that in K562 cell line, FTMT affects JAK/STAT pathway, a significant regulator of erythropoiesis, supports the hypothesis that FTMT may have an important role in the ineffective erythropoiesis (Santambrogio *et al.* 2011). In the same study, FTMT induced apoptosis in human $CD34^+$ cells.

Regardless the etiology, the presence of ring sideroblasts usually signifies ineffective erythropoiesis and mitochondrial iron overload, in the form of mitochondrial ferritin. However, for a correct diagnosis of MDS-RS, it is important to exclude nonclonal conditions associated with ring sideroblasts formation; excess of alcohol, lead toxicity, copper or pyridoxine deficiency, isoniazid therapy and hereditary sideroblastic anemias (Camaschella 2009, Willekens *et al.* 2013). The most frequent hereditary sideroblastic anemia is the X-linked sideroblastic anemia (XLSA) occurring secondary to missense mutations in conserved amino acids in the aminolevulinate synthase 2, *ALAS2*, a gene encoding the first enzyme of the heme biosynthetic pathway. Loss of function mutation in *SLC25A38* has been shown to cause a sideroblastic anemia similar to XLSA, but with an autosomal recessive inheritance (Guernsey *et al.* 2009). XLSA with ataxia (XLSA-A) is another inherited disorder associated

with sideroblastic anemia and non-progressive spinocerebellar ataxia. This form is caused by mutations involving the ATP binding cassette gene *ABCB7* (Steensma *et al.* 2007) not found in acquired cases. Similarly mutations in glutaredoxin-5 *GLRX5*, a gene that encodes a mitochondrial protein involved in iron-sulfur cluster biogenesis can give rise to sideroblastic anemia (SA) (Rouault *et al.* 2008). Additional rare causes of SA include mitochondrial deletion syndromes, usually diagnosed in pediatric patients (Patnaik *et al.* 2015).

MDS-RS is a slowly progressing disease characterized by an initial phase of erythroid hyperplasia and ineffective erythropoiesis that can remain stable for several years, and in some cases for decades. Usually, however, the stable phase is followed by a phase of increasing erythroid failure eventually resulting in a need for regular erythrocyte transfusions (Cazzola *et al.* 2013b). While pancytopenia and leukemic transformation is a rare event, anemia, chronic transfusion need and a subsequent risk of iron overload are the main clinical complications in MDS-RS and quality of life is usually severely reduced due to anemia (Bennett *et al.* 1982, Greenberg *et al.* 1997, Schmid 2009). Anemia is the most common cytopenia in all MDS, usually accompanied by increased apoptosis of bone marrow precursors, low reticulocytes counts and mild hemolysis (Santini 2015). Notably, defective erythropoiesis is a common denominator for most subtypes of MDS. Whereas hypoplastic erythropoiesis, characterized by a decreased relative number of erythroid progenitors in the bone marrow, is typically seen in advanced MDS (i.e. in 5q-syndrome), hyperplastic erythropoiesis, such as in MDS-RS, is characterized by an increased percentage of marrow erythroblasts, of which many undergo apoptosis before they mature into erythrocytes. This latter is typically observed in lower risk MDS, but is also common in a subset of RA and RAEB with a moderate increase of marrow blasts (Hellstrom-Lindberg *et al.* 2013). In the bone marrow cells of patients with low-risk MDS, the ineffective erythropoiesis has been attributed to the an excess of apoptosis and to a significantly higher DNA instability especially in the erythroid fraction, compared with age-matched controls (Novotna *et al.* 2009). Previous studies have shown that erythroid apoptosis in low-risk MDS is mediated via mitochondrial release of cytochrome *c* (Figure 5) and subsequent caspase activation (Tehranchi *et al.* 2003, Tehranchi *et al.* 2005a). The same studies demonstrated that growth factors such as erythropoietin and granulocyte colony-stimulating factor (G-CSF) can act both via inhibition of apoptosis of myelodysplastic erythroid precursors and via selection of cytogenetically normal progenitors. However, one of the aims of this thesis was to clarify at which stage of erythroid differentiation apoptosis is executed.

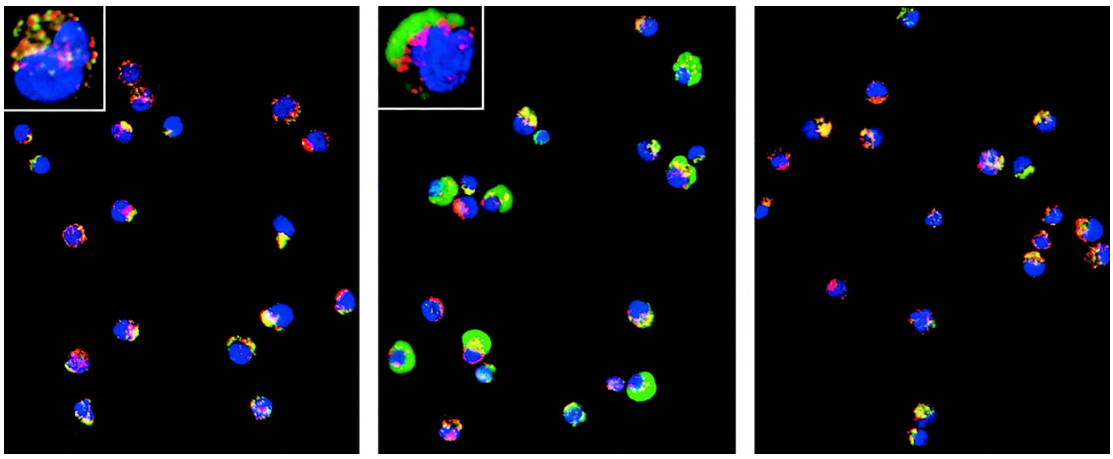

Figure 5 Mitochondrial release of cytochrome *c* in MDS-RS progenitors cells: inhibition by G-CSF. Mitochondria were stained with MitoTracker (red fluorescence), and cytochrome *c* localization was revealed by indirect immunofluorescence (green fluorescence). The yellow pattern denotes mitochondrial localization of cytochrome *c,* and the diffuse green pattern indicates cytochrome *c* that has been released into the cytosol. Data are derived from normal bone marrow (NBM; left panel), RARS progenitors (middle panel), and RARS progenitors treated with G-CSF (100 ng/mL; right panel), at day 7 of culture. Original magnification × 40; insets, × 100. Image adapted with permission from the publisher (Tehranchi *et al.* 2003).

1.4.2 TREATMENT OF MDS-RS

The predominant problem in MDS-RS is anemia, which alters quality of life in the elderly patients. Erythropoiesis stimulating agents, ESAs (i.e., recombinant erythropoietin, EPO or darbepoetin, DAR) constitute the first-line treatment of anemia (Fenaux *et al.* 2013, Malcovati *et al.* 2013). The efficacy of ESA can be improved by adding granulocyte colony-stimulating factor. Indeed, G-CSF in combination with EPO reduces the number of apoptotic bone marrow precursors by blocking the cytochrome c release during all stages of erythroid differentiation (Tehranchi *et al.* 2003). The treatment with EPO and G-CSF improves anemia and therefore, survival of patients, with a median duration of response of 23 months (Tehranchi *et al.* 2003, Jadersten *et al.* 2005, Jadersten *et al.* 2008, Hellstrom-Lindberg *et al.* 2013). When treatment after ESA fails due to primary resistance or relapse after a response, patients are usually confined to long-term red blood cell (RBC) transfusions and second-line treatments need to be considered.

Hypomethylating agents (HMAs), lenalidomide (LEN), and anti-thymocyte globulin (ATG) have been investigated as second-line treatment for ESA-refractory MDS-RS, but with very limited success. Hence, the main second-line therapy is chronic transfusion therapy, which will lead to iron overload, deleterious to organs such as heart and liver (Steensma *et al.*

2013). Therefore, chronic transfusion therapy is usually combined with iron chelation agents such as, deferoxamine and deferasirox (Malcovati *et al.* 2013, Steensma *et al.* 2013). Allogeneic stem cell transplantation is usually not indicated in MDS-RS due to its low risk for progression, but may be an option for patients who at relatively young age develop a very high transfusion need.

Recently, studies have indicated Sotatercept (ACE-011) and Luspatercept (ACE-536), as a new alternative to alleviate anemia (Dussiot *et al.* 2014, Suragani *et al.* 2014). These molecules are ligand traps for the transforming growth factor (TGF-β) superfamily, that by targeting a pathway fundamentally distinct from EPO (Raje *et al.* 2010), bind and block the interaction of ligands to their receptors and therefore inhibit their signaling (Attie *et al.* 2014). Pre-clinical studies in mice were associated with an increase of hemoglobin level and red blood cell count making ACE-011 and ACE-536, interesting novel compounds currently evaluated in phase II clinical trials on low and intermediate-1 risk MDS patients (Komrokji *et al.* 2014, Giagounidis *et al.* 2015).

1.4.3 MUTATIONAL LANDSCAPE IN MDS-RS

The introduction of next generation sequencing technology has allowed us to define the genetic landscape in MDS. In this scenario, spliceosome mutations were identified as the most common molecular abnormalities, affecting mainly the initial steps of the splicing process, such as the recognition of the 3' SS of the pre-mRNA target intron (SFSR2 and U2AF1) or the recruitment of the U2 snRNP to the branch point proximal to the 3'SS that contains *SF3B1* (Hahn *et al.* 2015).

Splicing mutations occur entirely as heterozygous missense mutations at highly restricted residues and in a mutually exclusive manner with one another (Inoue *et al.* 2016) making them an hallmark of MDS. Among these, *SF3B1* mutations are the most common in patients with MDS-RS (Papaemmanuil *et al.* 2011, Yoshida *et al.* 2011, Patnaik *et al.* 2012, Papaemmanuil *et al.* 2013, Haferlach *et al.* 2014). Interestingly, splicing mutations, while mutually exclusive with one another, show a tendency to co-occur with mutations in DNA methylators, chromatin modifiers and transcription factors.

A recent study by *Malcovati et al.* analyzed a cohort of 243 MDS patients with ring sideroblasts and showed that the most frequent mutations were splicing factors (181 of 243, 74.5%), DNA methylators (*TET2* & *DNMT3A*, 81 of 243, 33%), chromatin modifiers

(*ASXL1* & *EZH2*, 35 of 243, 14.4%), and the transcription factor *RUNX1* (28 of 243, 11.5%). In particular, a significantly higher prevalence of mutations in genes involved in DNA methylation was observed in RARS/RCMD-RS compared with RAEB. Conversely, a significantly lower rate of mutations in chromatin modifiers and transcription factors was found in sideroblastic categories compared with RAEB (17 of 159 vs 11 of 43, and 9 of 159 vs 13 of 43, respectively). In addition, a significantly higher prevalence of mutations in *TP53* was found in RA/RCMD (4 of 34, 11.8%) and RAEB (4 of 43, 9.3%) compared with RARS/RCMD-RS (2 of 159, 1.2%). Finally, a significantly higher prevalence of mutations in genes involved in signaling was observed in MDS/MPN with ring sideroblasts compared with RARS/RCMD-RS (Malcovati *et al.* 2015) (Figure 6).

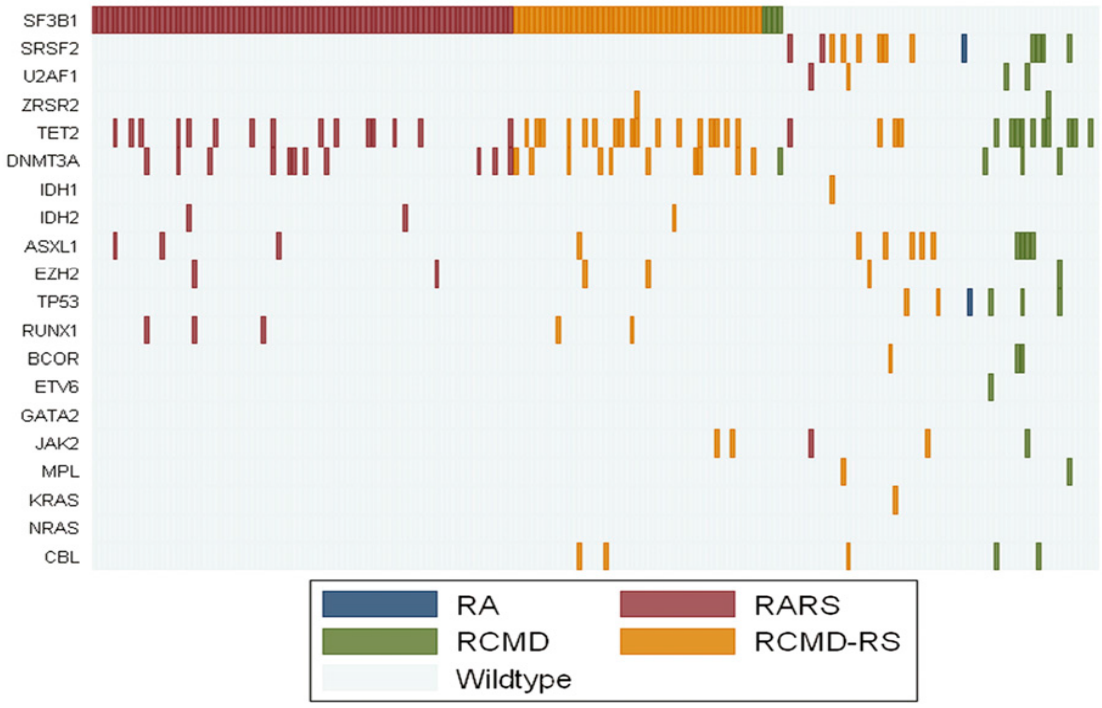

Figure 6 Mutations pattern in MDS patients. Figure adapted with permission from the publisher (Malcovati *et al.* 2015).

1.4.4 SPLICEOSOME AS THE "INDICTED CONSPIRATOR" IN MDS-RS

The discovery of mutations in components of the spliceosome, defined as the "indicted conspirator" (Abdel-Wahab *et al.* 2011) of myeloid malignancies, has led to the necessity of understanding the role of splicing and splicing factors in cancer. Splicing is an essential process that allows the generation of mature mRNA from pre-mRNA and where exon-intron junctions are recognized and intronic sequences removed. RNA splicing is highly coordinated throughout different reactions catalyzed by the spliceosome, a complex of five small ribonucleoproteins (snRNPs), U1, U2, U4/U6 and U5 together with several associated proteins (Hahn *et al.* 2015). In contrast, there are only two types of introns, U2 and U12, which are spliced by the U2-type and U12 type-dependent spliceosome, referred as major and minor spliceosomes, respectively.

The splicing process starts with the binding of the snRNP U1 to the 5' splice site of introns and at the same time at the 3' splice site, splicing factor 1 (SF1) binds to the branch point (BP), and the serine/arginine (SR) rich proteins SRSF2 and ZRSR2 bind to the exon splicing enhancer (ESE) site of the next exon to aid the binding and the stability of U2AF1 and U2AF2. Then, U2snRNP replaces SF1 and binds to the branch point forming complex A. SF3B1 protein, part of the U2 snRNP, is involved in the recognition of the branch point and prevents the premature attack at the site, by the splicing machinery (Schellenberg *et al.* 2011). U4/U6.U5 tri-snRNPs are recruited to form the B complex. After release of U1 and U4 snRNP, activated B complex performs the catalytic process of splicing (C complex) in two steps: firstly, the 5' splice site is cleaved and the 5' residue of the intron is linked to near the 3' splice junction forming the lariat loop; secondly, the 3' splice site is cleaved releasing the lariat and joining the two exons together (Yoshida *et al.* 2014), (Figure 7). After splicing is complete, the edited mRNA and introns are released, the spliceosome disassembles for recycling and the lariat loop is degraded in the nucleus.

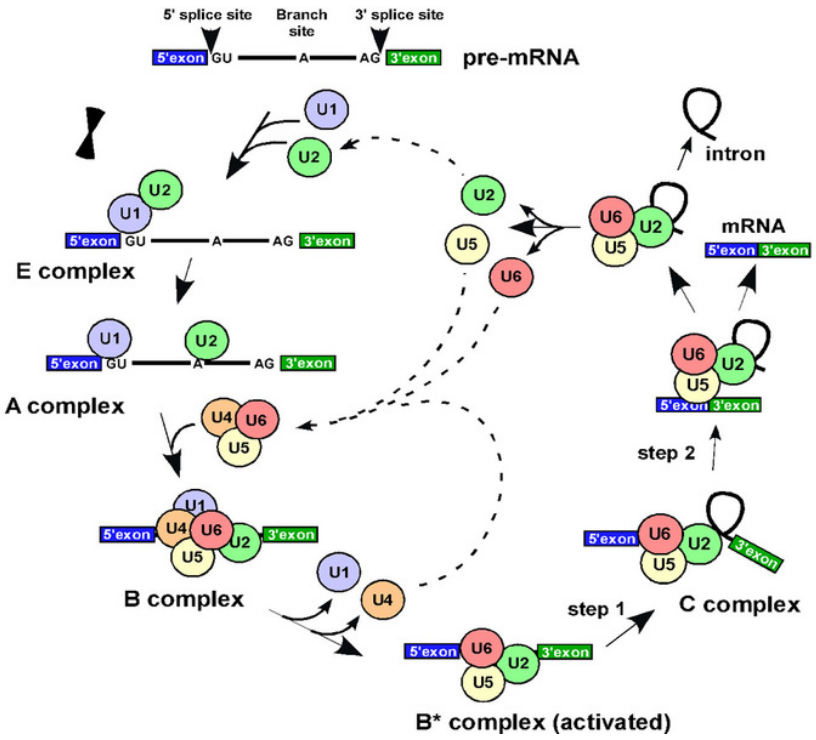

Figure 7 The splicing process. Image from the Max Planck Institute for Biophysical Chemistry http://www.mpibpc.mpg.de/

1.4.5 MUTATIONS IN THE SPLICING FACTOR *SF3B1*

SF3B1 is located at the chromosomal band 2q33 and encodes a core component of the U2 snRNP that binds the branch point sequences to stabilize the interaction between the U2 snRNP and the pre-mRNA (Hahn *et al.* 2015). Most mutations in *SF3B1* are heterozygous substitutions and tend to cluster on the C-terminal region consisting of 22 HEAT domains between exon 12 and exon 16 of *SF3B1* (Papaemmanuil *et al.* 2011). The most common recurrent *SF3B1* mutation affects amino acid K700 (45-68% of reported cases) followed by H662 (10%), K666 (10%), E622 (7%) and R625 (6%) (Malcovati *et al.* 2011, Papaemmanuil *et al.* 2011, Damm *et al.* 2012b, Papaemmanuil *et al.* 2013) (Figure 8). No significant differences were reported in regards to WHO subtype, clinical features and other hematologic parameters (Malcovati *et al.* 2011). However, significant relationships in MDS patients were found between *SF3B1* mutation and marrow erythroblasts, soluble transferrin receptor and serum growth differentiation factor 15. A multivariable analysis showed that the hepcidin to ferritin ratio, a measure of adequacy of hepcidin levels relative to body iron stores, was inversely related to the *SF3B1* mutation (Ambaglio *et al.* 2013).

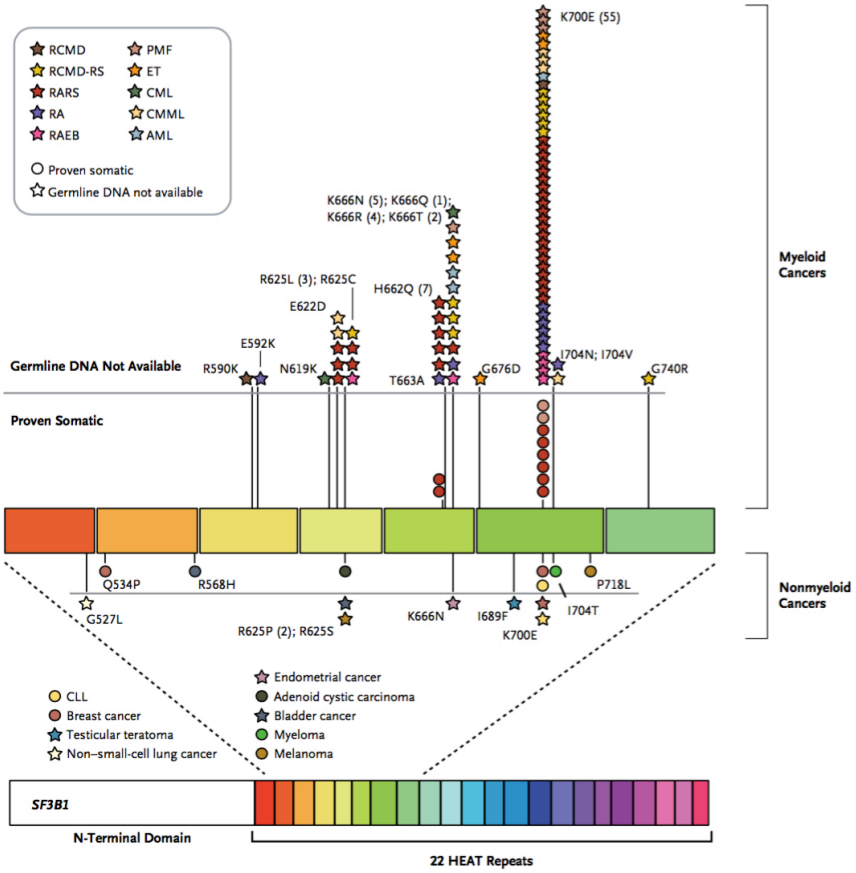

Figure 8 Distribution of mutations in *SF3B1*. Image reproduced with permission from (Papaemmanuil *et al*, 2011), Copyright Massachusetts Medical Society.

Average *SF3B1* allele burden in MDS-RS is approximately 40% with a few cases with low mutant allele burden. The median allele burden is similar in purified $CD34^+$, granulocytes and bone marrow GPA^+ cells, but the mutant clone confers a growth disadvantage at the terminal maturation into reticulocytes, as will be discussed later (Conte *et al.* 2015). Somatic mutations in *SF3B1* were reported in 14-28% of MDS and 19% of MDS/MPN cases (Papaemmanuil *et al.* 2011, Damm *et al.* 2012a, Damm *et al.* 2012b, Thol *et al.* 2012, Mian *et al.* 2013, Papaemmanuil *et al.* 2013), but also in other hematological malignancies such as primary myelofibrosis (PMF) (4%), therapy related AML, *de novo* AML (5%), essential thrombocythemia (ET) (3%), CLL (5%), multiple myeloma (3%) and chronic myelomonocytic leukemia (CMML) (5%). Interestingly, *SF3B1* mutations in CLL confer a poor prognosis (Quesada *et al.* 2012). Mutations were also found in solid tumors such as breast cancer (1%), renal cancer (3%) and adenoid cystic carcinoma (4%) (Papaemmanuil *et al.* 2011). Although *SF3B1* mutations were reported in a variety of myeloid malignancies, a remarkable high mutation rate has been reported among MDS patients with RS, including RARS (64-

90%) and RCMD-RS (37-76%) (Malcovati *et al.* 2011, Yoshida *et al.* 2011, Patnaik *et al.* 2012, Malcovati *et al.* 2015) and in RARS associated with thrombocytosis, RARS-T (67-85%) (Broseus *et al.* 2013), classified as a unique form of MDS/MPN.

SF3B1 mutated lower-risk MDS patients present a distinct clinical and morphological phenotype associated with the presence of RS (Malcovati *et al.* 2011). *SF3B1* mutations correlate with a favorable prognosis and confer a positive predictive value of 97.7% for the RS phenotype, whereas the absence of RS has a negative predictive value of 98% (Malcovati *et al.* 2011). Importantly, MDS-RS *SF3B1* wild type represents a subset with high prevalence of TP53 mutation and worse outcome (Malcovati *et al.* 2015). Clinical correlation with *SF3B1* mutation showed that MDS patients with the mutations have significantly lower incidence of multilineage dysplasia, lower proportion of dysplastic myeloid cells and megakaryocytes, higher white blood cell and platelets counts, marked erythroid hyperplasia and lower bone marrow blasts and incidence of chromosomal abnormalities, compared to unmutated patients (Malcovati *et al.* 2015). In MDS-RS patients, *SF3B1* mutations co-occur with mutations in genes involved in DNA methylation, chromatin modification, and transcription factor (i.e. *RUNX1*). Interestingly, among *SF3B1* mutated patients, coexisting mutations in DNA methylation genes were also associated with a higher degree of multilineage dysplasia. On the contrary, exclusivity was observed between *SF3B1* mutations and other splicing factors (*SRSF2* and *U2AF1*) and *TP53* (Malcovati *et al.* 2015).

The prognostic impact of *SF3B1* mutations in MDS-RS is slightly controversial, with some studies demonstrating a favorable independent prognostic impact (Malcovati *et al.* 2011, Broseus *et al.* 2013, Malcovati *et al.* 2015) and with others not confirming this pattern.

In particular, an early study on 161 low-risk MDS with RS (Malcovati *et al.* 2011) and more importantly, a recent multicenter study pursued by *Hellström-Lindberg* and *Cazzola* (Malcovati *et al.* 2015), analyzed a large cohort of patients with 1% or more RS and found that *SF3B1* had a positive prognostic value on survival and risk of disease progression and that this independent value was retained when the analysis was restricted to WHO subtypes without excess of blasts or to sideroblastic categories. Conversely, in one of the studies not confirming the prognostic relevance of *SF3B1* mutations, the presence of the mutation was associated with better overall surival and leukemia-free survival in univariable analysis, but this prognostic value was lost in multivariable analysis (Patnaik *et al.* 2012). In another study, the disassociation between *SF3B1* mutation and good prognosis was partially related to the heterogeneity of the cohort investigated and to the inclusion of a lower number of

MDS-RS cases (Damm *et al.* 2012a). In the study conducted by *Ebert et al.*, *SF3B1* mutations were under-represented in category 3 with a median survival of 1.11 years and of note, mutations were overlapping more than expected by chance with mutation in *DNMT3A*, indicating an unappreciated molecular synergy between these two genetic lesions. In fact, since *SF3B1* had a trend toward longer survival, co-occurrence may have mitigated any negative effect of *DNMT3A*, which showed longer median overall survival than *DNMT3A* alone (Bejar *et al.* 2012).

1.4.6 FUNCTIONAL IMPACT OF *SF3B1* MUTATIONS IN MDS-RS

Although substantial progress has been made in understanding the impact of several of these splice factor mutations, functional connections linking the mechanistic changes in splicing induced by these mutations to the phenotypic consequences of inefficient erythropoiesis have not been elucidated yet.

1.4.6.1 Mouse models of *SF3B1* mutations

Several studies in mice have sought to understand the functional role of *Sf3b1* (Visconte *et al.* 2012, Matsunawa *et al.* 2014, Visconte *et al.* 2014, Wang *et al.* 2014). *Sf3b1* heterozygous (*Sf3b1*$^{+/-}$) mice had normal hematopoiesis and did not develop hematological malignancies during a long observation period (Matsunawa *et al.* 2014, Wang *et al.* 2014). *Sf3b1*$^{+/-}$ cells had a significantly impaired capacity to reconstitute HSC in a competitive setting, but showed normal homing capability, a mild increase in apoptosis and no defects in differentiation. Additional depletion of *Sf3b1* using shRNAs compromised *Sf3b1*$^{+/-}$ HSC proliferative capacity compared to WT HSC *in vitro* and *in vivo*, implying that *Sf3b1* is required for the self-renewal capacity of HSC but not critical for the differentiation capabilities of HSC (Wang *et al.* 2014). Gene set enrichment analysis (GSEA) between *SF3B1*$^{+/-}$ and WT reported a negative enrichment of the non-sense mediated mRNA decay pathway, a system that surveys and degrades abnormal transcripts containing premature stop codons (Maquat 2004). This pathway was also reported as the mechanism through which *SF3B1* mutant aberrant transcripts is degraded, causing a down-regulation of key genes (i.e. *ABCB7*) in MDS-RS (Darman *et al.* 2015, Yusuke 2015).

Intriguingly, the correlation between *Sf3b1* mutations and the RS phenotype is controversial. No changes were observed in the frequencies of sideroblasts in *Sf3b1*$^{+/-}$

erythroblasts or cultured *Sf3b1*$^{+/-}$ erythroblasts expressing shRNAs targeting *Sf3b1* (Wang *et al.* 2014), showing that *Sf3b1* haploinsufficiency in mice is not enough to induce a RARS-like phenotype. Conversely, in two studies from *Visconte et al.* (Visconte *et al.* 2012, Visconte *et al.* 2014) BM aspirates from *SF3B1*$^{+/-}$ compared to C57BL/6 mice in one study and *Sf3b1*$^{+/-}$ mice compared to *Sf3b1*$^{+/+}$ showed the RS phenotype upon staining with Prussian blue. Also in these mice, hematological parameters between heterozygous and WT mice were comparable. In addition, the presence of RS has been preliminarily reported in NOD scid gamma (NSG) mice engrafted with MDS-RS purified HSC (Mortera-Blanco 2016), suggesting that the bone marrow microenvironment may be essential for the RS formation. Importantly, recent data from conditional knock-in mouse models of Sf3b1 K700E (Mupo *et al.* 2016, Obeng *et al.* 2016) sought to explain how *Sf3b1* alterations affect normal and malignant hematopoiesis.

Obeng et al. observed that *Sf3b1*^{K700E} conditional knock-in mice, had progressive macrocytic anemia associated with a terminal block in erythroid maturation outlined by the expression of CD71 and Ter119, which subdivided cells into stages R1-R4 of erythroblast maturation. As in previous mice models, no changes were observed in the total count of WBC, platelet and reticulocytes and no animals developed acute leukemia. *Sf3b1*$^{+/K700E}$ expression resulted in expansion of the long-term stem cell (LT-HSC) compartment and a decrease of GMPs, but an impaired capacity to reconstitute hematopoiesis in a competitive transplantation setting compared to *Sf3b1*$^{+/+}$ cells. Of note, *Obeng et al.* did not find ring sideroblasts or circulating siderocytes in the bone marrow consistent with other mouse models of congenital sideroblastic anemia (including a conditional knock-out model of *Abcb7*) (Pondarre *et al.* 2007), potentially suggesting an unknown difference between human and murin erythroid precursors with respect to mitochondrial iron metabolism (Inoue *et al.* 2016).

Moreover, RNA sequencing on myeloid progenitor cells sorted from 3 *Sf3b1*$^{+/K700E}$ and 3 *Sf3b1*$^{+/+}$ in the study by *Obeng et al*. showed only a minimal overlap in mis-spliced genes and events between human MDS and murine samples. For instance, *Abcb7* cryptic splice site was not seen in *Sf3b1*$^{+/K700E}$ mice because of the distinct intronic sequence at this region of *Abcb7*. In addition, authors identified that loss of *Tet2*, which plays a key role in DNA demethylation, in combination with *Sf3b1*$^{+/K700E}$ aggravated the macrocytic anemia and impaired terminal maturation and accelerated the expansion of LT-HSC, erythroid and megakaryocytes dysplasia. The combination of *Tet2* loss and *Sf3b1*^{K700E} was sufficient to

rescue the competitive repopulation disadvantage induced by *Sf3b1*^{K700E} alone, offering an explanation to why the mutations in both of these genes are well-tolerated in MDS patients. Also, bone marrow progenitor cells expressing *Sf3b1*^{K700E} showed an increased sensitivity to the spliceosome modulator E7107 compared to *Sf3b1* wild-type.

Sf3b1^{K700E} knock-in mouse model from *Mupo et al.* presented progressive normocytic anemia compared to wild-type, no significant dysplasia and no ring sideroblasts, despite of a moderate overall increase in iron deposits. Interestingly, *Sf3b1*^{K700E} animals showed no differences in the early erythroid cells, but a relative decrease of the more mature erythroid cells and a decrease in HSC in mutant mice with no change in the size of LMPP, GMP, CMP and MEP progenitors compartments compared to wild-type mice.

Similarly to *Obeng et al*, *Mupo et al.* looked at the pattern of aberrant splicing promoted by *SF3B1*^{K700E}, and found that mutant animals have an increased use of cryptic 3' splice sites (SS) throughout their genome, similarly as in human cancers carrying *SF3B1* mutations. Although the splicing defects observed in *SF3B1*^{K700E} closely mimicked those of human MDS, the murine orthologues of genes associated with ring sideroblasts in human MDS, including *Abcb7* and *Tmem14c*, were not mis-spliced as a result of the relatively poor interspecies conservation of intronic sequences able to function as aberrant splice sites. Overall, the mild phenotype observed by the authors in comparison to *Sf3b1* mutant human MDS has been explained as the requirement for additional mutations to develop MDS and that different phenotypic effects (i.e. anemia, ring sideroblasts, clonal expansion) of splicing mutations may have distinct molecular causes (Mupo *et al.* 2016).

1.4.6.2 Effects of *SF3B1* mutations on splicing

In addition to experimental evidences in mice, transcriptome sequencing of MDS-RS BM MNC (Visconte *et al.* 2015) and CD34^{+} cells (Dolatshad *et al.* 2015, Dolatshad *et al.* 2016) allowed the study of normal bone marrow, *SF3B1* mutated and WT patients, in some case from different WHO subgroups, and to identify mis-spliced events functionally linked to MDS pathogenesis and downstream targets of *SF3B1* mutations. *SF3B1* mutated RARS/-T showed more abundant iron deposits in the mitochondria that unexpectedly, did not correlate with increased reactive oxygen species (ROS) or DNA damage; instead, the distinct iron distribution was attributed to intron retention event in the splicing pattern of *SLC25A37*, a crucial importer of Fe^{2+} into the mitochondria (Visconte *et al.* 2015). Also, disruption of *SF3B1* resulted in differential exon usage of genes known to be involved in

MDS (*TP53* and *EZH1*), erythroid genes (*ALAD, UROD* and *EPB42*) and genes associated with cell cycle (*AURKB* and *CRNDE*) and RNA processing (*RBM5, RBM25, PRPF40A* and *HNRNPD*) when comparing uncultured BM $CD34^+$ cells of *SF3B1* mutated and those of healthy controls (Dolatshad *et al.* 2015). In the same study, altered splicing of genes involved in MDS pathophysiology (*CBL, ASXL1* and *DNMT3A*), mitochondrial function (*ALAS2, NDUFAF6*), erythroid differentiation (*NFE2L2, PPOX, HMBS*) was reported when comparing *SF3B1* mutated and WT cases (these last all RCMD patients). Differential exon usage was also detected in genes (i.e. *NUMA1, SMAD4, BRIC2* etc.) regulated by the DNA damage-induced mRNA splicing complex composed by *SF3B1-BRCA1* and *BCLAF1*, implying a role of *SF3B1* in the maintenance of genomic stability (Savage *et al.* 2014). Notably, the lack of MDS-RS *SF3B1* wild type patients without other MDS co-recurrent mutations, made genome editing for functional studies and transcriptome on genome edited primary cells, highly relevant to highlight the global mechanisms through which *SF3B1* mutations contribute to changes in the splicing process. Recently, results from RNA sequencing in human *SF3B1* mutated cells (Darman *et al.* 2015, Shiozawa *et al.* 2015) have demonstrated that most of the aberrant splicing events selectively observed in *SF3B1* mutated samples are caused by misrecognition of the 3'splice sites (SS), consistent with the function of the wild type protein and mostly by utilizing a cryptic AG sequence and resulting in frameshift mutations and subsequent transcripts degradation by the NMD pathway (Figure 9 adapted from (Malcovati *et al.* 2016). Furthermore, treatment of cultured *SF3B1* mutant MDS erythroblasts and CRISPR/Cas9-generated *SF3B1* mutant cell line with the NMD inhibitor cycloheximide, showed that the aberrantly spliced *ABCB7* transcripts is targeted by NMD pathway (Dolatshad *et al.* 2016), possibly explaining the increased mitochondrial iron accumulation found in MDS patients with RS.

Additionally, computational analysis suggested that *SF3B1* mutations cause a change in the size of the sterically protected region downstream of the branch point without any altered BP recognition allowing the utilization of the cryptic AG sequence (DeBoever *et al.* 2015). It was also hypothesized that *SF3B1* mutations could alter interaction with U2AF2, affecting the ability to recognize the canonical 3'SS and leading to cryptic 3'SS selection. In contrast, *Darman et al*, demonstrated that the short and weak Py tract upstream of the cryptic AG is insufficient for the recruitment of the U2AF2 complex and that the integrity of the canonical 3'SS, and therefore interaction with U2AF2 is required to observe aberrant splicing as well. Also, they showed that *SF3B1* mutant U2 complex utilizes a BP

that is different from the one used by the WT U2-complex. However, despite the recent progress, the understanding of how mutations in *SF3B1* result in altered 3'SS usage as well as, the understanding of the normal function of SF3B1 HEAT domains, where the majority of the mutations are located, is still incomplete.

In parallel, other studies identified novel roles for SF3B1 in binding to chromatin or chromatin-modifying proteins, for instance SF3B1 binds to nucleosomes on exonic DNA, independently of RNA, and the association of SF3B1 with nucleosomes facilitates the splicing recognition of exons (Kfir *et al.* 2015). Therefore, disruption of SF3B1 binding to such nucleosomes affects splicing of these exons similarly to *SF3B1* knock-down. In a similar way, SF3B1 and Polycomb repressive complex 1 (PRC1) component interaction (Isono *et al.* 2005) was shown to be required to repress Hox genes, transcription factors closely linked to normal hematopoiesis and leukemogenesis (Argiropoulos *et al.* 2007).

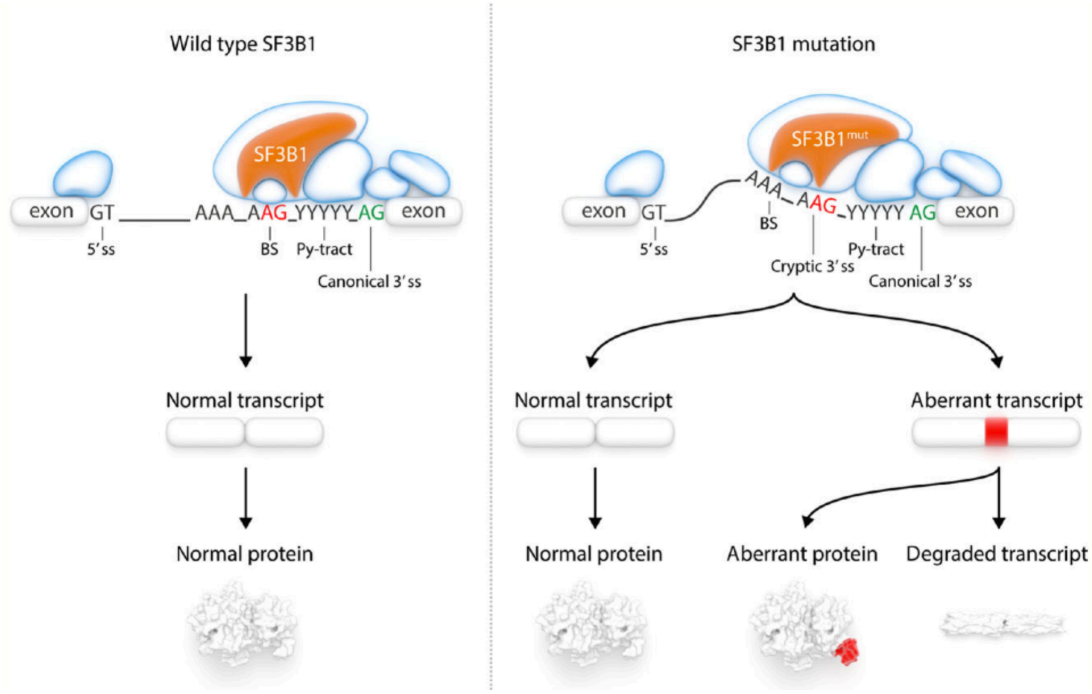

Figure 9 RNA splicing in wild-type and mutant *SF3B1* cells. On the left, *SF3B1*, core component of the U2 snRNP, is responsible of the 3'SS recognition. U2 snRNP is directed to the 3'SS by short pre-mRNA sequences, including the branch point (BP) a sequence motif upstream the 3'SS, the Py tract and the AG dinucleotide at the intron-exon junction. On the right panel, *SF3B1* mutations result in misrecognition of 3'SS, by utilizing a cryptic AG sequence. Based on the position of premature termination codons, aberrant mRNA transcripts that contain premature stop codons are then subjected to degradation by NMD pathway. Figure reprinted with permission from the publisher (Malcovati *et al.* 2016).

1.4.6.3 Gene expression profile in *SF3B1* mutated cells

Gene expression profile offers an image of the transcriptional activity of genes at a specific time point. For a decade, it has provided a powerful tool to identify common downstream targets, predictors of response and prognostic markers in spite of the clinical heterogeneity of MDS patients. An early GE study on RARS and NBM $CD34^+$ cells at day 0 and at day 7 of erythroid differentiation reported a marked down-regulation of several pathways including apoptosis, DNA damage repair, mitochondrial function and the JAK/Stat pathway, a mediator of cellular responses to growth factors (Nikpour *et al.* 2010). In the same study, effects of G-CSF on gene expression to further understand its anti-apoptotic role in RARS erythropoiesis showed that the mitochondrial pathway including mitofusin 2 (*MFN2*), involved in mitochondrial membrane integrity (Sugioka *et al.* 2004), was significantly modified by G-CSF, and several heat shock protein genes were up-regulated, as evidence of anti-apoptotic protection of erythropoiesis. By contrast, G-CSF had no effect on iron-transport (i.e. *ABCB7*), erythropoiesis-associated genes or the expression level of *MAP3K7*, a negative regulator of apoptosis and down-regulated in RARS erythroblasts (Nikpour *et al.* 2010). A key finding of the report from *Pellagatti et al.* was the up-regulation of several interferon-induced genes (*IRF2, IRF6, IRF2BP2, TOR1AIP2, IFNA17, AEN* and *ISG20L2*) that may be responsible for enhancement of apoptosis in erythroid progenitors, hence in ineffective erythropoiesis of MDS-RS (Pellagatti *et al.* 2006).

Gene Set Enrichment Analysis (GSEA) of *SF3B1* mutated samples and in *SF3B1* mutant (RARS and RCMD-RS) versus wild type and control (RCMD) showed under-expressed mRNA profiles with emphasis on mitochondrial pathways (Papaemmanuil *et al.* 2011) and up-regulated gene sets related to cell cycle checkpoints and mRNA splicing (Dolatshad *et al.* 2015). While the comparison of *SF3B1* mutant with healthy control highlighted down-regulated gene sets related to cell differentiation and apoptosis (Dolatshad *et al.* 2015). Although most cells, including normal erythroid cells, express extremely low levels of mitochondrial ferritin (*FTMT*), this intron-less gene is highly expressed in erythroblast. Overexpression of *FTMT* in RARS erythroblasts occurs at a very early stage of eryhtroid differentiation (Tehranchi *et al.* 2005b), suggesting that it might be a cause rather than a consequence of mitochondrial iron overload. Differential gene expression analysis of *SF3B1* mutant versus wild type and controls highlighted genes linked to MDS-RS pathogenesis such as, *ALAS2* (up-regulated) and *ABCB7* (down-regulated) and also an up-regulation of

SLC25A37, encoding a mitochondrial iron importer and *GLRX5*, essential for the iron-sulfur cluster biosynthesis (Ye *et al.* 2010). Finally, MDS-RS progenitors up-regulate several genes involved in the hemoglobin pathway (i.e. *HBQ1, HBA2, HBA1, HBB*) compared to healthy controls (Dolatshad *et al.* 2015). The cause of this up-regulation is not entirely clear. However, corresponding studies in NBM $CD34^+$ cells reported that the process of differentiation-associated with hemoglobin accumulation occurred first at day 6 followed by a rapid raise in the ratio between fetal and adult hemoglobin production between day 7 and 9 during the transition from pro to basophilic erythroblasts (Wojda *et al.* 2002). Therefore the early up-regulation of these genes may indicate a compromised hemoglobinization during the subsequent phases of erythropoiesis.

2 AIM OF THE THESIS

The overall aim of this thesis was to investigate the genotype-phenotype relations in patients with myelodysplastic syndromes and ring sideroblasts.

Specific aims were:

- STUDY I

Test the hypothesis that *ABCB7* is a key mediator of ineffective erythropoiesis in MDS-RS patients.

- STUDY II

Identify potential downstream targets of *SF3B1* mutations and understand how these affect RNA splicing and gene expression profile of MDS-RS patients.

- STUDY III

Understand the mechanistic effects of *SF3B1* mutation in MDS-RS pathogenesis and explore if the mutation confers a loss-of-function with regard to gene expression and splicing of key genes in MDS-RS.

3 METHODOLOGICAL APPROACHES

The laboratory techniques used in this thesis are described in detail in study I-III. Hereby, only principal methods and their potential limitations are listed and commented.

3.1 $CD34^+$ CELLS ISOLATION AND ERYTHROBLAST CULTURE

$CD34^+$ cells were isolated from the BM samples of MDS-RS patients and healthy controls using magnetic-activated cell sorting (MACS) columns according to the manufacturer's instructions. Following positive selection for CD34, cells (0.1×10^6/mL) were cultured for 14 days in Iscoves medium supplemented with 15% BIT9500 serum substitute (containing bovine serum albumin, bovine pancreatic insulin and iron/saturated human transferrin), recombinant human interleukin-3 (rh-IL-3, 10 ng/mL), rh-IL-6 (10 ng/mL) and rh–stem cell factor (rh-SCF, 25 ng/mL). The medium was replenished every second day with the aforementioned cytokines to maintain the cultures at the same cell concentration. EPO (2 U/mL) was added to the medium in the second week, beginning at day 7, and was replenished at days 9 and 11 only to the amount of media added to keep the same cell concentration.

This cell culture system is a useful model for studies of early erythroid maturation as well as maturation defects in MDS-RS. Indeed, it allows for ample proliferation of intermediate erythroblasts in which it is possible to detect aberrant mitochondrial ferritin accumulation in MDS-RS cells, but it does not produce mature red blood cells or morphologically typical ring sideroblasts.

3.2 COLONY FORMING UNIT ASSAY

The first CFU assay was published in 1966 (Bradley *et al.* 1966, Pluznik *et al.* 1966). This assay estimates the frequencies of multi-potential and lineage-committed hematopoietic progenitors cells (HPCs) in the BM, PB and other hematopoietic tissues. A colony is defined to consist of at least 50 cells. The primary use of this assay is to detect the differentiation potential of hematopoietic progenitors cells, therefore it can be defined as a qualitative proliferation assay based on number and morphology of colonies.

In study I, lentiviral-transduced $CD34^+$ cells from healthy controls were plated in a concentration of 500 cells/plate in methylcellulose medium (H4434, StemCell Technologies) on polystyrene 35 mm petri dishes and cultured according to manufacturer's guidelines. The cells were incubated at 37^oC and 5% CO^2 for 14 days, after which the number and type of

colonies was determined by manual counting as per established protocols (Goh *et al.* 2005). Burst forming unit-erythroid (BFU-Es), colony forming unit-erythroid (CFU-Es), colony forming unit-granulocyte-macrophage (CFU-GMs) and colony forming unit-granulocyte-erythrocyte-monocyte-macrophage (CFU-GEMMs) were scored according to color, morphology and number of cells with an inverted or a fluorescence microscope, if cells were carrying a fluorescent reporter. Three replicates were counted for each sample. CFU assay offers the possibility of re-isolating cells from the methylcellulose-based medium by 'picking' individual colonies and use them for further experiments (i.e. mutational analysis).

3.3 Q-PCR

Quantitative real time PCR was used to: i) measure mRNA levels of selected genes (Study I-III), ii) validate gene expression data from RNA sequencing (study II), iii) assess exon usage (study I and III) and finally, iv) monitor the efficiency of gene knock-out (study I and III). Both Sybr green and Taqman probes were used to address the aforementioned purposes. To normalize the differences in cDNA input, data were normalized versus the expression of the following housekeeping genes: *GAPDH, ACTIN* and *HPRT*. Fold change was calculated using the comparative Ct method described by *Livak* and *Schmittgen* (Livak *et al.* 2001, Schmittgen *et al.* 2008). Overall, qPCR is highly reproducible but is influenced by the method used for nucleic acids extraction. Hence, in order to compare gene/transcript numbers from different samples, it must be ensured that the same extraction procedure is used for each sample. Also, the target specificity of the qPCR assay is determined by the choice of probes and the design of the primers, which has to satisfy certain criteria such as the length (90-150 bp) or the GC content (20-80%). Finally, due to the high sensitivity of the fluorescence-based qPCR method, it is important that RNA templates are free from contaminating DNA that could interfere with the final amplification signal. This feature is particularly important when quantifying intron-less genes, such as mitochondrial ferritin. Importantly, sensitivity is limited by the fact that qPCR measures only a part of the transcript (i.e. the part corresponding to the primer or the probe location), but not an entire isoform. Therefore, in case of RNA sequencing splicing validation data, it is important to identify the specific mis-spliced region of a transcript isoform before designing the assay.

3.4 TAQMAN LOW DENSITY ARRAY

Low Density Array is a q-PCR based method introduced by Applied Biosystems. Taqman®

LDA is a modified method of real-time TaqMan® PCR that uses micro plates based on micro-fluidic systems. The advantages of this technology include the possibility of simultaneous quantification of a large numbers of target genes in a single sample retaining the sensitivity of a Taqman qPCR.

In study II, the gene expression profile of BM $CD34^+$ cells from 7 RARS and 4 NBM samples at day 0 and 4 of erythroid differentiation was investigated using this technology. Samples were run in triplicate. Each TLDA card allowed simultaneous quantification of 61 targets and 3 housekeeping genes per sample. Loading and setting up the LDA cards is straightforward, with no needs for robotic liquid handling. Previous studies demonstrate that LDA can accurately detect 2-fold changes in target copy number, at both low Cts (high expression) and high Cts (low expression), within the approximate range of $1 \times 10^2 – 1 \times 10^6$ copies (Goulter *et al.* 2006). However, other studies sustain that reproducibility decreased in low copy genes. Hence, caution is recommended when analyzing LDA results of those low copy genes (Lu *et al.* 2008). In our hands, TLDA was a sensitive and high-throughput approach to analyze changes in gene expression during erythroid differentiation. However, RNA extraction methods and high quality RNA are absolutely necessary and RNA input requirements were significantly high considering the low amount of RNA in $CD34^+$ cells from RARS and NBM samples (Hofmann *et al.* 2002).

3.5 PYROSEQUENCING

Pyrosequencing was applied to follow the fate of the *SF3B1* clone during erythroid differentiation (study I). DNA from cultured BM $CD34^+$, freshly isolated BM MNC and GPA^+ of RARS patients was used with DNA from NBM, as wild type control, while cDNA was used to assess allele burden in reticulocytes. Mutation-specific assays were designed using PyroMark assay design software and performed according to the PyroMark Q24 user manual (Qiagen). PCR reactions and mutation analysis by pyrosequencing were performed according to procedures and reagents provided by manufacturer. Pyrosequencing involves several steps: i) a sequencing primer is added to a single strand PCR amplicon. ii) the hybridized primer and single-stranded template are incubated with the enzymes DNA polymerase, ATP sulfurylase, luciferase, and apyrase, as well as the substrates adenosine 5' phosphosulfate (APS) and luciferin. iii) dNTPs are added to the reaction and every time a dNTP is incorporated in the strand, a pyrophosphate (Ppi) is released. iv) the ATPsulfurylase converts PPi to ATP; ATP mediates conversion of luciferin into oxyluciferin that generates a

light signal. The height of each peak (light signal) is proportional to the number of nucleotides incorporated while apyrase degrades all the unincorporated nucleotides.

Overall, pyrosequencing is a useful method to quantify mutated allele burden of target genes (i.e. *SF3B1*) and it can be used as validation method after targeted gene sequencing. In the mutational analysis, we set as thresholds, mutation read above 25% as positive and read below 5% as negative. Of note, the sensitivity of this method varies depending on cell type, for instance pyrosequencing on colonies picked from methylcellulose may encounter risk of contamination from other colonies or negative measurements due to scarce amount and low quality of the DNA isolated from the colonies. Today more sensitive methods (i.e. droplet digital PCR) are available to address similar research questions, but at the time of our studies, we employed well-established and appropriate methods.

3.6 RNA SEQUENCING

Spliceosome mutations are expected to exert their effects through alterations in RNA expression and splicing, making the study of the MDS transcriptome of critical importance. The type of information that can be gained from RNA sequencing include "annotated" information such as expressed transcripts, exon/intron boundaries, transcriptional start sites (TSS), and poly-A sites; as well as quantitative data such as, differences in expression, alternative splicing, alternative TSS, and alternative polyadenylation between two group of samples (Wang *et al.* 2009). RNA sequencing involves RNA extraction, cDNA synthesis, fragmentation and library synthesis. The library is then sequenced using short-read sequencing which produces millions of short sequence reads that correspond to individual cDNA fragments. The precision of the RNA sequencing gene quantification is directly dependent on the number of reads that are mapped to transcripts. The coverage of the transcriptome affects the accuracy of transcripts abundance estimation and the sensitivity of transcripts detection, which are critical features of all gene expression studies.

In study II, RNA sequencing was used for transcriptome profiling as well as for the identification of alternative splicing events. Total RNA from BM $CD34^+$ cells of 3 MDS-RS and 3 NBM samples during early erythroid differentiation was depleted of ribosomal RNA and strand-specific libraries were constructed using Sequencing by Oligo Ligation Detection (SOLiD™) system. Ribosomal depletion, in contrast to PolyA enrichment, preserves non-polyadenylated RNAs allowing the investigation of immature mRNAs and non-polyadenylated non-coding RNAs (ncRNAs). Due to the very labile nature of RNA, we

extracted RNA immediately after the $CD34^+$ enrichment to assure high quality and a good yield. Small amounts of the RNA samples were used to assess RNA integrity before starting library synthesis. Libraries were sequenced on the SOLiD 5500 Wildfire system. The sequencing by ligation approach used in SOLiD involves hybridization and ligation of primed template and fluorescently labeled anchor-probe with known base(s). In each cycle, the emitted fluorescence of ligated probe is captured for imaging the template. Unlike other sequencing platforms, SOLiD utilizes two-base-encoding probes, in which each ligation signal represents two bases of nucleotides, defined as color-space (Mardis 2008, Metzker 2010).

The sequencing performed generated a read length of 75bp bases that were mapped to the human genome hg19 and assembled into transcripts and differential expression estimation, exon usage and relative abundance (fragments per kilobase of exon per million of fragments mapped, FPKM) was calculated. Genes were considered as expressed if FPKM>1.

In conclusion, RNA sequencing usually requires complex analysis including filtering of sequencing reads, assembling reads into transcripts or aligning them to reference sequences, annotating the putative transcripts, quantifying the number of reads per transcript, and performing a statistical comparison of transcript abundance across samples. RNA sequencing studies based on clear and explicit hypothesis should be preferred and pursued over hypothesis-free studies that focus on finding significant correlations in very large data sets. This is especially true in the prospective of subsequent functional data validation required after sequencing.

3.7 LENTIVIRAL TRANSDUCTION

In study I and III lentiviral vectors were used to transduce both human K562 cell line and human bone marrow progenitors cells. Our plasmids were 2^{nd} generation lentiviral systems both constitutive (i.e. pLKO.1 and Lenti V2-GFP) and inducible plasmids and they were used in knock –out and over-expressing experiments, respectively.

For decades, these vectors have been largely employed in gene therapy and functional studies for their ability to integrate into dividing but also, quiescent cells (Naldini *et al.* 1996b) such as HSC. The advantageous feature of lentiviral vectors is the ability to mediate transduction and stable expression both *in vivo* and *in vitro* (Naldini *et al.* 1996a). This feature makes these vectors a good model for instance for long-term differentiation assays,

often crucial in disorders characterized by ineffective hematopoiesis, such as MDS. The efficiency of transgene expression is also depending on the type of promoter driving the expression. It has been reported that vectors containing the EF1α and to lesser extent, the PGK promoter, govern high-level gene expression in hematopoietic progenitors (Salmon *et al.* 2000). Also, when the viral vector enters the cell, it integrates into the genome and is passed on to the progeny of the cells. However, the site of integration is random and it may disturb the function of other genes and/or lead to the activation of oncogenes (Cattoglio *et al.* 2007). Interestingly, while retroviral vectors integrate preferably near transcriptional start sites, lentiviral vectors show a potentially safer integration profile in terms of probability of oncogene activation (Modlich *et al.* 2009).

Notably unlike cell lines, $CD34^+$ cells tend to be less susceptible to manipulation and require a high concentration of virus inoculum thereby demanding time-consuming virus stock preparation where batch variation can occur.

3.8 SHORT HAIRPIN RNA

In study I we used shRNAs to knock-down *ABCB7* in human bone marrow progenitors. shRNAs require the use of an expression vector, therefore they were cloned into a pLKO.1 vector carrying GFP as reporter. Flow cytometry was used to separate cells expressing fluorescence thereby assessing transduction efficiency. When the vector has integrated into the host genome, the shRNA is transcribed in the nucleus by polymerase III. The product mimics the pri-microRNA (pri-miRNA) and is processed by Drosha. Exportin 5 exports the resulting pre-shRNA from the nucleus. This product is then processed by Dicer and loaded into the RNA-induced silencing complex (RISC). The sense strand is degraded. The antisense (guide) strand directs RISC to mRNA that has a complementary sequence. In the case of perfect complementarity, RISC cleaves the mRNA. In the case of imperfect complementarity, RISC represses translation of the mRNA. In both of these cases, the shRNA leads to target gene silencing.

Since our shRNAs vectors carried a fluorescent marker, we could measure the knock-down efficiency over time by quantifying the target mRNA in sorted GFP^+ cells by qPCR without the need of a clonal isolation. Although shRNAs have been the most common tool for loss-of-function studies, the transduction efficiency can be very different between various cell-types, and for some gene target there is no effective shRNA. Also, in some circumstances shRNAs can trigger non-specific effects and/or sequence-specific off-target effects

(Barrangou *et al.* 2015). These off-target effects are very difficult to monitor, thereby forcing the researcher to adopt several shRNA constructs per gene and to pursue "rescue experiments". In such experiment, knock-down of the endogenous gene of interest is phenotypically rescued by induced expression of the modified gene, that is not targeted by the knock-down construct.

3.9 CRISPR/CAS9 GENOME EDITING

One of the aims in study III was to understand if the phenotype observed in MDS-RS patients was the result of a loss-of-function mutation. Therefore, we designed knock-out experiments using CRISPR/Cas9 technology for its high efficiency, low off-target effects and for its possibility to introduce K700E allowing for studies of the effects of this mutation. Study III is in manuscript form and some experiments are ongoing.

The Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) is a bacterial immune system that has been modified for genome engineering. CRISPR consists of two components: a "guide" RNA (gRNA) and a non-specific CRISPR-associated endonuclease (Cas9). The gRNA is a short synthetic RNA composed of a targeting and a scaffold sequence (derived from endogenous bacterial crRNA and tracRNA) necessary for Cas9-binding and a user-defined 20 nucleotide sequence which defines the genomic target to be modified (target sequence). The sequence has to be unique compared to the rest of the genome and the target sequence must precede a 5'-NGG sequence called Protospacer Adjacent Motif (PAM). The PAM sequence is necessary for target binding and the exact sequence is dependent upon the species of Cas9 (Figure 10A) (Ran *et al.* 2013). Once expressed, the Cas9 protein and the gRNA form a riboprotein complex, Cas9 undergoes a conformational change upon gRNA binding that shifts the molecule from an inactive, non-DNA binding conformation, into an active DNA-binding conformation. The Cas9-gRNA complex will bind any genomic sequence with a PAM, but the extent to which the gRNA spacer matches the target DNA determines whether Cas9 will cut. Once the Cas9-gRNA complex binds a putative DNA target, a "seed" sequence at the 3' end of the gRNA targeting sequence begins to anneal to the target DNA (Larson *et al.* 2013). Cas9 will only cleave the target if sufficient homology exists between the gRNA spacer and target sequences. Cas9 undergoes a second conformational change upon target binding that positions the nuclease domains to cleave opposite strands of the target DNA. The Cas9-mediated DNA cleavage results in a double

strand break (DSB) within the target DNA, 3-4 nucleotides upstream of the PAM sequence (Figure 10B).

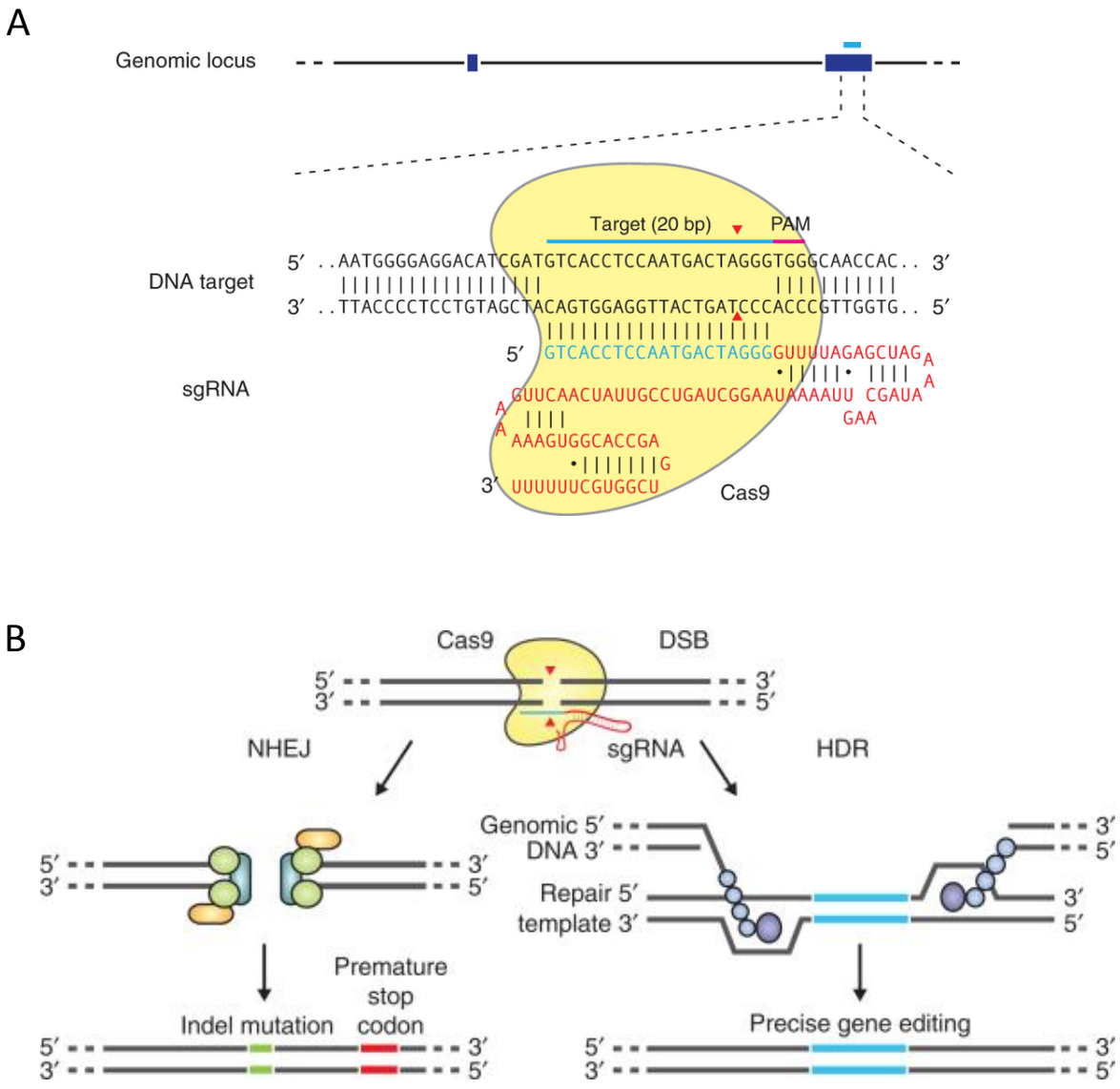

Figure 10 A The Cas9 nuclease (in yellow) is targeted to genomic DNA by a gRNA sequence (blue) and a scaffold (red). The guide sequence pairs with the DNA target (blue bar on top strand), directly upstream of a required 5′-NGG motif (PAM; pink). Cas9 mediates a DSB 3 bp upstream of the PAM (red triangle). **B** Schematic representation of the HDR and NHEJ repair pathways. Figures adapted with the permission from the publisher (Ran *et al.* 2013).

The DSB is then repaired by one of two repair pathways: the Non-Homologous End Joining (NHEJ) pathway and the Homology Directed Repair (HDR) pathway. The NHEJ repair pathway is very efficient but error-prone. It generates small indels in the target DNA or frameshift mutations leading to premature stop codons within the open reading frame

(ORF) of the targeted gene. Ideally, the end result is a loss-of-function mutation within the targeted gene; but the efficacy of the knock-out phenotype is ultimately determined by the amount of residual gene function, that is why it is important to evaluate for example the KO efficiency at protein level.

By contrast, the HDR pathway can be used to generate specific nucleotide changes, for example to introduce a point mutation such as K700E in *SF3B1*. This pathway is less efficient but more precise. However, to induce HDR, a DNA repair template containing the desired sequence must be delivered into the cells along with the gRNA and Cas9. The repair template must contain the desired edit as well as additional homologous sequence upstream and downstream of the target. It can be a single or a double-stranded oligonucleotide, or a double-stranded DNA plasmid depending on the specific application. Importantly, the template must lack the PAM sequence that is present in the genomic DNA to avoid a possible cleavage mediated by Cas9. Therefore, it is wise to introduce a silent mutation in the repair template, so that the PAM is no longer present. However, the efficiency of HDR in $CD34^+$ cells has been reported to be low (<10%) (Genovese *et al.* 2014) due to the fact that human progenitor cells poorly tolerate Cas9 and generally, the introduction of vectors or oligonucleotides by nucleofection. Also, whereas the efficiency of Cas9 cleavage is relatively high and that of HDR is relatively low, a portion of the Cas9-induced DSBs will be repaired via NHEJ, meaning that the resulting cells population will contain a combination of wild-type alleles, NHEJ-repaired alleles, and/or the HDR-edited allele. Thus, considering the fact that we study differentiating cells and that CRISPR/Cas9 modifies gene (via knock-out or knock-in) at DNA genomic level and with sometime very low efficiency rate, it is crucial to confirm the presence of the desired edits experimentally (i.e. by sequencing technologies or ddPCR or western blot), to distinguish NHEJ and HRD events and to isolate clones (both homozygous and heterozygous) containing the desired changes. However, generating clones may be difficult especially in primary cells (i.e. $CD34^+$ cells) if the gene of interest is crucial for cell proliferation or survival, thereby making the knock-out efficiency validation very challenging. In conclusion, while CRISPR/Cas9 allows the generation of a full knock-out, (i.e. the complete loss of function), shRNAs reduce the gene function following a normal distribution that may lead to clonal selection.

4 RESULTS AND DISCUSSIONS

4.1 STUDY I

In study I, in order to investigate whether *ABCB7* is an important mediator of the aberrant iron accumulation and ineffective erythropoiesis in MDS-RS, we modulated its expression in human bone marrow progenitors cells. First, down-regulation of *ABCB7* in normal progenitors by shRNAs fused with GFP significantly reduced erythroid colony growth and to a lesser extent myeloid colony formation (N=3) (Figure 1A, study I). *ABCB7* silenced cells were cultured for 14 days in our well-established erythroblast culture system (Tehranchi *et al.* 2003, Pellagatti *et al.* 2007). While no significant effects on cell survival were observed during the first 7 days, the percentage of silenced *ABCB7* GFP^+ cells dropped drastically at day 14 compared to approximately 8-10% in the scramble controls (Figure 1B, study I). During the 14 days of culture, cells expressing the early erythroid CD36 marker did not decrease, but showed a decrease of GPA^+ suggesting that *ABCB7* may affect more the later stages of erythroid maturation (Figure 1C-D, study I).

Next, we explored potential changes in gene expression in $CD34^+$ GFP^+ sorted *ABCB7* down-regulated cells during differentiation. After 3 days of culture, we could observe a 2-fold increase of mitochondrial ferritin (*FTMT*). At day 7 the *ALAS2* expression was significantly decreased, indicating a reduced heme synthesis at this time point. Furthermore, we observed a reduction of *FOXO3A*, involved in the protection from oxidative stress, and *MAP3K7*, a negative regulator of apoptosis were also reduced upon *ABCB7* down-regulation (Figure 1E, study I).

Vice versa, over-expression of *ABCB7* using a constitutive vector expressing YFP rescued erythroid maturation and restored colony growth of these cells. $CD34^+$ cells from 4 MDS-RS patients over-expressing *ABCB7* showed enhanced erythroid colony growth, in particular of the BFU-E colonies, compared to cells transduced with the empty vector (Figure 2A-B, study I). Importantly, *ABCB7* forced expression decreased the mRNA level of *FTMT*, suggesting that ABCB7 may have a role in the aberrant mitochondrial iron accumulation (Figure 2C-D, study I).

RNA samples from 5 *SF3B1* mutated MDS-RS patients and 3 healthy controls were obtained at day 7 and day 14 of *in vitro* erythroid differentiation and used to investigate a potential differential exon usage previously detected in a pilot RNA sequencing experiment. The latter analysis highlighted an altered usage between exons 8 and 9. Primers were designed to

specifically quantify different regions of *ABCB7* mRNA transcript. Exon usage was more heterogeneous within the MDS-RS population compared to the healthy controls. In spite of a profound down-regulation in all investigated regions at both time points (Figure 3A-B, study I), it was difficult to detect a strong effect on this specific exonic region 7J8-9 (Figure 3E, study I). Furthermore, transient siRNA-mediated down-regulation of SF3B1 in K562 cells resulted in reduced expression of *ABCB7*, a reduction that was even more pronounced when cells were treated with hemin to induce erythroid differentiation (Figure 4, study I).

To summarize, in study I we examined the role of *ABCB7*, an ATP-binding cassette (ABC) transporter located at the inner membrane of the mitochondria. Missense mutations in this gene lead to X-linked sideroblastic anemia and ataxia (XLSA/A) (Allikmets *et al.* 1999). Notably, the existence of an inherited disorder XLSA/A caused by mutations in ABCB7 is very revealing of the probable consequences of reduced expression of the same gene in acquired disorders. Direct sequencing of *ABCB7*, including the promoter region, was performed on DNA from 13 RARS patients and did not show any mutations, nor significant differences in the methylation status of *ABCB7* promoter region of RARS patients compared to NBM (Nikpour *et al.* 2010). Previous studies performing *ABCB7* silencing and conditional gene targeting in mice resulted in siderocytosis without ring sideroblasts formation and highlighted the importance of *ABCB7* in hematopoiesis and cytosolic iron-sulfur cluster biogenesis (Pondarre *et al.* 2006, Pondarre *et al.* 2007). Additionally, silencing of *ABCB7* in HeLa cells resulted in an iron-deficient phenotype with mitochondrial iron accumulation (Cavadini *et al.* 2007).

At the time when study I was designed, *ABCB7* had been reported to be significantly down-regulated in MDS-RS patients in comparison with healthy controls and other MDS subtypes (i.e. RA and RAEB) (Boultwood *et al.* 2008). Furthermore, we observed that *ABCB7* expression levels decreased during differentiation of MDS-RS cultured erythroblasts, in contrast to a continuous increase during normal erythroid differentiation (Nikpour *et al.* 2010), raising the hypothesis that low *ABCB7* levels would reflect abnormal mitochondrial iron homeostasis. Intriguingly, *ABCB7* expression increased in normal erythroblast cultures in parallel with hemoglobinization, in contrast to the gradual decrease of *ABCB7* expression from day 0 to day 14 in MDS-RS, thereby supporting a critical role of *ABCB7* in aberrant iron accumulation. Interestingly, although a functional relation between *ABCB7* and RS formation has not been established in MDS-RS, *ABCB7* expression level is inversely correlated with the percentage of marrow ring sideroblasts (Boultwood *et al.* 2008). As the

erythroblast culture model does not produce ring sideroblasts, we used the expression of *FTMT* as surrogate marker for aberrant iron accumulation, which accumulates during early erythroid differentiation of RARS progenitors (Tehranchi *et al.* 2003, Tehranchi *et al.* 2005b). Specifically, *FTMT* gene and protein expression increased after *ABCB7* silencing in normal progenitors during erythroid differentiation. *ABCB7* down-regulation and *FTMT* over-expression both lead to increased RNA-binding activity of iron regulatory proteins (IRPs) and consequently an increase in cellular iron uptake from transferrin (Nie *et al.* 2005, Cavadini *et al.* 2007), which mainly incorporates into *FTMT*. The avidity of *FTMT* for iron is stronger than that of its cytosolic counterpart (Nie *et al.* 2005). Hence, *FTMT* over-expression may lead to the functional iron deficiency in RARS erythrocytes, evident by the finding of hypochromic erythrocytes in RARS (Ljung *et al.* 2004). Moreover, *FTMT* expression decreased in RARS samples with forced expression of *ABCB7*, in spite of relatively moderate transduction frequencies.

Silencing of *ABCB7* in NBM cells inhibited erythropoiesis more than myelopoiesis, indicating that the ABCB7 protein may be more important for differentiation of the erythroid lineage than the myeloid lineage. Interestingly, this finding corroborates with a recent study showing that mitochondrial damage impairs elimination of mitochondria during erythroid differentiation leading to enhanced erythrocyte destruction (Ahlqvist *et al.* 2015).

In particular, *ABCB7* down-regulation impaired late erythroid differentiation and survival, while earlier and intermediate stages were unaffected. This pattern corresponds well to the MDS-RS morphology, which is dominated by intermediate and late dysplastic erythroblasts. Thus, it is possible that reduced *ABCB7* levels results in marked cellular phenotypic changes first when iron turnover or heme synthesis increases above a critical threshold. Our finding of a strong association between MDS-RS and low gene expression levels of *ABCB7* and the fact that over-expression of *ABCB7* restored erythropoiesis and survival of RARS progenitors while decreasing the expression of aberrant mitochondrial ferritin, raises the possibility of therapeutic approaches for MDS-RS for instance by using drugs or genetics approaches (i.e. *in vitro* transcribed mRNAs) to boost the expression of *ABCB7*.

Importantly, studies from *Papaemmanuil* and *Yoshida et al.* reported high frequencies of mutations in *SF3B1* in MDS-RS patients and encouraged us to explore a potential association between *SF3B1* gene mutation and *ABCB7* down-regulation in RARS. Transient siRNA-mediated down-regulation of *SF3B1* during erythroid differentiation resulted in reduced expression of *ABCB7*, suggesting a possible link between the two genes. Finally, we reported

altered usage of *ABCB7* exons as a possible explanation of the reduced *ABCB7* expression in MDS-RS, which might have been secondary to the altered *SF3B1* function due to mutation. Whether this relationship plays a role in RARS pathogenesis was not clear at the time of this publication but was later confirmed by other studies (Darman *et al.* 2015, Shiozawa *et al.* 2015).

4.2 STUDY II

In study II RNA sequencing comparing MDS-RS *SF3B1* mutated and NBM $CD34^+$ cells during early erythroid differentiation (day 0 and 4) showed marked dissimilarities in gene expression, pathways and splicing. In particular, several of the top 20 up-regulated genes (such as *HBM, HBA2, HBA1, ALAS2*) in MDS-RS were involved in the heme and hemoglobin biosynthesis (Table I, study II). The μ-globin (*HBM*) gene was the most up-regulated among the genes involved in the hemoglobin pathway. *HBM* was initially identified as a pseudogene and later detected particularly abundant in cord blood reticulocytes (Goh *et al.* 2005). *HBA1* and *HBA2* encode the α-globin protein and mutations in these genes are generally the most common cause of α-talassemia (Hardison *et al.* 2002). Of note, *HBA2* was reported down-regulated in differentiating erythroblasts of MDS patients with del(5q) after treatment with lenalidomide (Pellagatti *et al.* 2007). Among the most up-regulated genes in MDS-RS, we reported also *ALAS2*, a gene known for its fundamental role in the heme synthesis and expressed exclusively in erythroid cells. The gene expression levels of HBM, *HBA2, HBA1, ALAS2* were similar in *SF3B1* mutated (n=6) and wild-type (n=2) patients. However, transcriptome data from *Dolatshad et al.* (Dolatshad *et al.* 2015) reported these genes to be differentially expressed between SF3B1 mutated and wild-type patients.

The expression of several de-regulated genes (*TIMD4, ALAS2, CD5L, LYVE1, CXCL18, MSR1* and *CXCL12*) detected by RNA sequencing was confirmed by TLDA analysis (Figure 4C-D, study II). Out of 61 genes selected for TLDA, 20 genes at day 0 and 17 genes at day 4 were differentially expressed between MDS-RS and NBM. Interestingly, several of these genes (*ABCB7, ALAS2, APOE, HBB, MFN2, SEPP1*) have been shown to be involved in defense against oxidative stress suggesting that MDS-RS progenitors are challenged by ROS during early differentiation (Miyata *et al.* 1996, Saito *et al.* 1999, Shen *et al.* 2007, Liesa *et al.* 2012, Bayeva *et al.* 2013) (Supplementary Table 9, study II). This finding is in line with a positive enrichment of the oxidative phosphorylation pathway, and together with a down-regulation of the ABC transporters strongly implies a mitochondrial dysfunction in MDS-RS (Figure 1E-F,

study II). Previous studies have shown that low-risk MDS patients exhibit a significant increase of oxidative damage to DNA, which could contribute to genomic instability and disease progression (Ghoti *et al.* 2007, Novotna *et al.* 2009).

In addition, the up-regulation of genes involved in the heme biosynthesis pathway may indicate a compromised hemoglobinization and an impaired ability of mitochondrial pathways to use iron, leading to the increased mitochondrial iron accumulation observed in MDS patients with ring sideroblasts. Although the altered expression of these genes is an intriguing finding in MDS-RS patients, mutations found in congenital sideroblastic anemia (i.e. *ALAS2* mutations) have not been described in MDS.

The fact that splicing and gene expression patterns undergo major changes during normal differentiation (An *et al.* 2014, Cheng *et al.* 2014, Pimentel *et al.* 2016), made this longitudinal study during erythroid maturation highly relevant. We employed RNA sequencing to identify mis-spliced events functionally linked to MDS-RS disease pathogenesis in its earliest stages. In particular, 18 genes were mis-spliced at day 0 and 2 genes (*ANKHD1* and *TMEM14C*) showed altered splicing at day 4. Only *TMEM14C* was differentially spliced at both time points (Supplementary Table 7, study II). Of note, the coverage of *ABCB7* for which we previously reported altered exon usage during erythroid maturation, was suboptimal for this analysis. Among those, the function of *TMEM14C* and *TFCP2* was closely related to mitochondria and in particular to heme biosynthesis (Yien *et al.* 2014, Yien *et al.* 2015) (Figure 3A, study II). *Tmem14c* encodes a mitochondrial inner membrane protein required for the transport of mitochondrial porphyrin (Yien *et al.* 2015). Recent evidence suggests that protoporphyrin IX synthesis in *Tmem14c*-deficient eryhtroid cells was blocked causing an accumulation of porphyrin precursors. In these cells, the heme synthesis defect was ameliorated with a protoporphyrin IX analogue, indicating that *Tmem14c* functions mainly in the terminal stages of the heme synthesis facilitating the import of protoporphyrinogen IX into the mitochondrial matrix, and is required for heme synthesis and late erythropoiesis (Yien *et al.* 2014, Yien *et al.* 2015). Also, *Tmem14c* knock-down in zebrafish led to profound anemia without affecting erythroid lineage specification (Nilsson *et al.* 2009). Our transcriptome analysis reported two isoforms of TMEM14C: TCONS15 strongly up-regulated and TCONS16, significantly down-regulated in MDS-RS patients versus NBM. Of note, *TMEM14C* TCONS15 was significantly down-regulated in *SF3B1* wild-type compared to mutated patients (Figure 3A, study II). On the other hand, the shorter isoform TCONS16 lacked 14 bases in the 5'UTR within the predicted internal

ribosome entry site (IRES) important for translation of specific mRNAs during mitosis, apoptosis and hypoxia (Hellen *et al.* 2001) (Figure 3B, study II). Therefore, a sequence change in this region could affect the translation of *TMEM14C*, with consequences for mitochondrial function and haemoglobin synthesis (Yien *et al.* 2015). Moreover, we detected 3 different isoforms of *TFCP2* differentially expressed between MDS-RS and healthy controls (Figure 3A, study II). *TFCP2* is a transcription factor that activates the transcription of the α-globin gene, thus regulating the erythroid gene expression (Kang *et al.* 2005). Among the 3 isoforms found, TCONS_0012739 lacked 153 bases in the sixth exon of the mRNA corresponding to the CP2 DNA binding domain that binds the α-globin promoter (Figure 3C, study II). CP2 is a major factor in the regulation of globin expression in human and mouse erythroid cells, and CP2 binding to the promoter is essential for the enhanced transcription of globin genes in erythroid differentiation (Chae *et al.* 2003). We used qPCR to measure the abundance of transcripts containing exon 6 and we observed a 1.3 fold-decrease in MDS-RS versus NBM, whereas the expression of TCONS00112740 (referred as TCONS40 in study II) showed the same trend as in the RNA Sequencing, however inter-patient variation was high and the comparison was not statistically significant. The additional 16 mis-spliced genes were related to various functions, such as DNA replication (*DUT*), MAPK signalling pathway (*MBIP*), transcription regulation (*ZNF518B*, *EIF2B4, TGFI2*), blood coagulation (*PROCR*), intracellular trafficking and energy metabolism (*ACBD3, SNX3, CYP2S1, NUDT7*) (Supplementary Table 7, study II). In particular, EIF2B4 acts within a tri-subunits complex that includes EIF2α kinase. In normal erythropoiesis, heme induces globin transcription and translation by binding the inhibitors BACH1 and EIF2α kinase, respectively, so as soon as heme is present, globin synthesis begins (Yang *et al.* 2016). This inhibition occurs when the intracellular concentration of heme declines, thereby preventing the synthesis of globin peptides in excess of heme (Han *et al.* 2001).

Next, we investigated whether *SF3B1* mutated progenitors cells had a growth disadvantage compared to *SF3B1* wild-type cells. MDS-RS is characterized by ineffective erythropoiesis and anaemia, but the exact stage of erythroid maturation at which this defect occurs it is poorly understood (Hattangadi *et al.* 2011). Previous investigations have described apoptotic features of RARS erythroblasts in bone marrow biopsies as well as in erythroid culture systems (Tehranchi *et al.* 2003, Tehranchi *et al.* 2005a, Tehranchi *et al.* 2005b). Interestingly, this does not translate into a decreased proportion of marrow erythroid cells, but rather to erythroid expansion, indicating concomitant stimulation of erythroid

proliferation and survival (Nikpour *et al.* 2013). First, we measured *SF3B1* allele burden in $CD34^+$ cells isolated from 5 MDS-RS patients at day 0, 4, 7, 11 and 14 of the erythroid and myeloid differentiation (Figure 4E, study II). Allele burden was stable over time and a similar pattern was observed in patients with or without transfusion need, i.e. the production or not of erythrocytes did not influence the results. Then, to exclude the possibility that the culture system would have conferred a growth advantage of the mutated clone, we selected GPA^+ cells from freshly obtained BM $CD34^-$ cells of 7 MDS-RS patients carrying different *SF3B1* mutations. Allele burden in the GPA^+ fraction was similar to the $BMCD34^+$ and erythroblast D14 fractions previously evaluated. Finally, we assessed *SF3B1* allele burden in reticulocytes obtained from peripheral blood of non-transfused MDS-RS patients with stable anemia and a median of 47% (33-70) RS of total BM erythroblasts. The *SF3B1* allele burden was 1.5 fold lower in reticulocytes compared to corresponding fresh $CD34^+$ cells and cells cultured for 14 days, and 1.4 fold lower than in freshly isolated GPA^+ cells (Figure 4F, study II). These results imply reduced differentiation of *SF3B1* mutated erythroblasts at the stage when they mature into erythrocytes, and indicate that *SF3B1* mutations do not inhibit growth and maturation of marrow erythroblasts, in spite of the mitochondrial iron accumulation. Conversely, the decreased allele burden in reticulocytes indicates that the main threat to red cell formation develops during terminal differentiation to erythrocytes. The fate of *SF3B1* clone has been also assessed in an ongoing study using an *ex-vivo* long-term culture model able to reproduce self-renewal potential of HSC and to generate ring sideroblasts, erythroid islands and red blood cells. In this culture, the *SF3B1* mutant clone was stable throughout 4 weeks of culture (Elvarsdóttir *et al.* 2016).

To summarize, study II offers a number of potential targets affected by the *SF3B1* mutations, and demonstrates how these targets may contribute to the MDS-RS pathogenesis. Mutations of *SF3B1* may, via altered gene expression or mis-splicing of key genes in the heme and hemoglobin synthesis, disturb mitochondrial iron handling in a way that causes mitochondrial iron accumulation in MDS-RS. In addition, this study indicates that potential therapeutic targets such as ACE-536 or ACE-011 (Dussiot *et al.* 2014, Suragani *et al.* 2014), targeting the late stages of erythropoiesis and currently evaluated in clinical trials may be suitable to treat MDS-RS patients. A major challenge would be to establish if *SF3B1* mutations are causative of the development of MDS-RS, or if they are involved in the disease progression, merely altering the phenotype and in some cases co-operating with other mutations, with no effect on the natural history of the disease.

4.3 STUDY III

Our aim was to investigate the impact of the *SF3B1* K700E mutation in myeloid cells and explore if this mutation confers a loss-of-function in regards of gene expression and splicing of key genes in MDS-RS. From a therapeutic point of view it is important to understand if the clinical phenotype derives from a loss or a gain of function of SF3B1. To date, SF3B1 heterozygous mouse models have failed to develop the human MDS-RS phenotype and knock-in mice generated only a mild phenotype in terms of RS formation and specific mis-splicing events observed in MDS-RS progenitors. Also, it is still unclear if *SF3B1* alone is sufficient to initiate the disease in human primary cells. In this study, we used CRISPR/Cas9 technology to knock-out *SF3B1* and a tetracycline inducible system to express the wild-type and the recombinant *SF3B1* K700E in human K562 cells. First, we transduced K562 *SF3B1* wild-type cells with an empty vector (Lentiv2-GFP) and with Lentiv2-GFP targeting either exon 1 (Ex1-intr) or exon 2 (Ex2) of *SF3B1*. The targeting vector efficiently knocked-out *SF3B1* in K562 cells causing 83% and 49% protein reduction in Ex1-intr and Ex2 respectively (Figure 1A, study III), and 98.6% deletion in exon 1 and 68% in exon 2 at genomic level (Figure 1B-C, study III). Moreover, the cutting efficiency of CRISPR/Cas9 was evaluated at mRNA level using primers specific for the predicted deleted regions in GFP^+ cells and a significant reduction was detected in both exons (Figure 1D-E, study III). Then, by targeting exon1-intron, we generated two single cell-derived heterozygous clones, EX1-int N5 and EX1-int N7, both carrying a frameshift deletion, while gRNAs against exon 2 produced two clones, Ex2 N15 homozygous and Ex2 N2 heterozygous (Figure 2A-C, study III). Protein alignment of Lentiv2-GFP N1 (wild-type for Exon1-int and Ex2) against each clone showed that frameshift deletions induced by CRISPR/Cas9 led to various amino acids substitutions in the heterozygous clones, but not in Ex2 N15 where the deletion did not change the amino-acid sequence (Figure 2D-E, study III). These results were supported by western blot analysis showing a 50% reduction of SF3B1 protein in all heterozygous clones (EX1-int N5, EX1-int N7 and Ex2 N2) in addition to an almost complete loss of protein in the homozygous clone Ex2 N15 (Figure 2F, study III). Conversely, to test if *SF3B1* K700E mutation may induce an MDS-like molecular phenotype in K562 cells, we cloned *SF3B1* WT or mutant K700E into a Tet-on inducible vector containing GFP. Upon transduction and treatment with doxycycline, cells were sorted for high and low GFP levels, and part of the cells were cultured to monitor possible toxic effects of the vector, which did not affect cell viability during 14 days of

culture. Of note, the mRNA expression level of *SF3B1* recombinant and wild-type were similar to that of physiological levels (Figure 3 A-B, study III).

To evaluate how K700E expression, and loss or reduction of the normal SF3B1 protein affect gene expression and splicing, we focused on three well-known genes in MDS-RS for their potential role in the disease pathogenesis: *ABCB7, ALAS2* and *TMEM14C*.

Loss or reduction of SF3B1 normal protein resulted unexpectedly in an MDS-RS-like phenotype with reduced *ABCB7* expression and altered exon usage (Figure 4B and Figure 5E, study III), increased levels of *ALAS2* expression (Figure 4C, study III) and mis-splicing of *TMEM14C* (Figure 5C, study III). Furthermore, loss of SF3B1 compromised cell growth in K562, but did not increase apoptosis suggesting that the reduced cell growth instead might be due to an impaired regulation of the cell cycle (Figure 6A-B, study III).

Intriguingly, expression of K700E mutation at physiological level, induced mis-splicing of *TMEM14C* (Figure 5B, study III), decrease *ABCB7* and increase *ALAS2* expression levels (Figure 4D-E, study III), as observed in MDS-RS patients and in knock-out cells; however *ALAS2* expression in the mutant samples was less pronounced compared to the that of patients, or that detected in the KO clones. Importantly, *SF3B1* WT expression resulted in a moderate increase of *ABCB7* as well as an unexpected increase in *ALAS2*. These results may be explained by the fact that *ABCB7* has been previously reported to be mis-spliced in MDS-RS and to be one of the targets of mutant *SF3B1*-induced nonsense-mediated decay (NMD pathway) (Darman *et al.* 2015). Conversely, *ALAS2* was characterized by differential exon usage (Dolatshad *et al.* 2015), however according to *Darman et al*. study (Darman *et al.* 2015) *ALAS2* was not identified as a target of the NMD surveillance system. According to this study, *SF3B1* heterozygous mutation in MDS-RS patients causes splicing events, half of which are operated by normal SF3B1 and the other half by abnormal SF3B1. Splicing mediated by normal SF3B1 leads to canonical transcripts and proteins, whereas the mutant splicing factor mis-recognizes the 3' splice site and generates frameshift mutations resulting in an unproductive isoform that would shunt the gene from the normally translated pathway into the NMD pathway (Darman *et al.* 2015). Therefore, we could speculate that the haploinsufficiency in MDS-RS *SF3B1* mutated patients may lead to gain-of-function at stem cell level (i.e. proliferation of HSC), but loss-of-function at precursors level, perhaps caused by *ABCB7* mis-splicing resulting in ineffective erythropoiesis (Boultwood *et al.* 2008, Nikpour *et al.* 2013). These observations may be relevant for the treatment of MDS-RS as recent studies have shown that SF3B1 haploinsufficiency is responsible for the proliferative

capacity of HCS. It also implies that the use of splicing factor modulators may selectively target MDS-RS cells with somatic mutations in *SF3B1*, as recently shown by the E7107 treatment of progenitor cells from *Sf3b1*^{K700E} mice (Obeng *et al.* 2016).

In summary, our data indicates that *SF3B1* mutation exerts its effects through a loss-of-function. The degree to which the function is lost varied in the different experimental systems, however both knock-out and expression of the mutant result in gene expression patterns similar to those observed in MDS-RS patients. Detailed RNA sequencing and genome editing experiments in primary cells will be essential to decipher whether *SF3B1* mutations act though a gain or a loss of function. In contrast to our data, the fact that *SF3B1* mutations cluster at hotspots, in a mutation pattern that is characteristic for oncogenes, suggest a possible gain-of-function (Zhou *et al.* 2015). Yet it remains to be unraveled what functional role *SF3B1* mutations may play during oncogenesis and whether dys-regulated *SF3B1* activity is required for cancer maintenance. SF3B1 small-molecule inhibitors are currently under development and have entered clinical trials (Webb *et al.* 2013). Moreover, the absence of frameshift deletions among hundreds of MDS-RS patients, and the fact that the observed mutations are less deleterious than expected on the basis of chance (Papaemmanuil *et al.* 2011), may imply that the mutant SF3B1 protein retains structural integrity, but with altered function (Malcovati *et al.* 2016), therefore arguing for *SF3B1* as a neomorphic mutation (Darman *et al.* 2015). However, loss-of-function effects have been described as a consequence of common point mutations, such as R882 *DNMT3A* in acute myeloid leukemia. This mutation acts as dominant negative against the WT enzyme causing hypomethylation throughout the genome of these patients (Russler-Germain *et al.* 2014).

Future studies should aim to optimize methods to introduce *SF3B1* point mutations and genetically modify normal primary bone marrow CD34^{+} cells in order study the role of SF3B1 in disease initiation. Importantly, considering that *SF3B1* is shared by two distinct spliceosome complexes (major and minor) owing to its importance in the removal of any type of intron, a disturbance in the splicing mechanism by mutations may have dramatic consequences on transcription and expression of thousand of gene. Therefore, it may be relevant to pursue a combined analysis of transcriptiome and proteomics and to investigate the biochemical properties of transcripts, including protein-RNA interactions and domain activity, which may have a great impact on translational efficiency and offer insights into therapeutic opportunities able to modulate SF3B1.

5 CONCLUDING REMARKS AND FUTURE DIRECTIONS

MDS-RS is characterized by ineffective erythropoiesis, the presence of bone marrow ring sideroblasts and anemia. Undeniably, MDS-RS being a disease of the elderly, other factors may have an impact on its development. At present, the available treatments aim primarily to improve erythroid output.

Before the discovery that *SF3B1* splicing mutations were associated with MDS-RS, the general hypothesis was that X-linked sideroblastic anemia and acquired RARS were sharing the same molecular basis. Indeed, the existence of an inherited disorder XLSA/A caused by mutations in *ABCB7* led us to investigate the possible consequences of reduced expression of the same gene in acquired disorders, with the major aim of understanding the reason for the aberrant iron accumulation in MDS-RS.

In study I, we unraveled the role of *ABCB7* and reported evidences that this gene is a key mediator of the MDS-RS phenotype. In normal bone marrow *ABCB7* down-regulation reduced erythroid differentiation, growth and colony formation and resulted in a gene expression pattern similar to that observed in RARS erythroblasts, and in the accumulation of mitochondrial ferritin. Conversely, forced *ABCB7* expression restored erythroid colony growth and decreased mitochondrial ferritin in MDS-RS $CD34^+$ progenitor cells. In addition, we suggested a relation between *SF3B1* mutation and *ABCB7* down-regulation and showed altered usage of *ABCB7* exons as a possible explanation of the reduced *ABCB7* expression in MDS-RS and a secondary effect to the altered SF3B1 function; an hypothesis later supported by other authors in a detailed RNA sequencing study (Darman *et al.* 2015).

The discovery of splicing mutations in MDS has provoked the field with more challenging tasks. These mutations occur in a mutually exclusive manner with one another and at highly restricted residues. Interestingly, each of these mutations is associated with distinct clinical subtypes and phenotypes of MDS (i.e. *SF3B1* mutations with the ring sideroblasts). Undeniably, evidence that aberrant pre-mRNA splicing contributes to the disease phenotype is strong and it has raised an array of mechanistic, functional and biological questions. However, understanding the functional role of mutations in spliceosomal genes is just beginning. Most reports used RNA sequencing to study splicing throughout normal hematopoiesis. One common concept is that each stage of hematopoiesis is defined by cell type-specific differential splicing, which may reflect into changes in gene expression and result in altered function and/or stability of the encoded proteins. This idea inspired us to perform study II, where we addressed aberrancies during erythroid maturation of normal

and MDS-RS *SF3B1* mutated patients with regard to splicing and gene expression profile. In fact, a very recent longitudinal transcriptome study of late human erythropoiesis, the splicing in a differentiation stage-specific context showed striking increases in levels of intron-containing transcripts in the final two steps of erythroid maturation (Pimentel *et al.* 2016).

In study II, the comparison of MDS-RS *SF3B1* mutated and NBM $CD34^+$ cells during early erythroid differentiation (day 0 and 4) revealed a marked up-regulation of genes involved in the hemoglobin synthesis and the oxidative phosphorylation process and a down-regulation of the mitochondrial ABC transporters most likely as a compensatory reaction to the defective terminal erythroid maturation with accumulation of iron in the mitochondria instead of incorporation into heme. More importantly, we discovered that via mis-splicing of genes involved in the heme and hemoglobin synthesis, *SF3B1* mutation might disturb the mitochondrial iron handling and the hemoglobinization process in MDS-RS. Finally, we found that *SF3B1* did not confer a growth disadvantage until terminal maturation into reticulocytes, implying a reduced differentiation of *SF3B1* mutated erythroblasts into mature erythrocytes and suggesting that in MDS-RS, anemia develops during the late stages of erythropoiesis. This finding provides a logical link to new compounds such as ACE 536, which targets the late stage of erythropoiesis and increases the levels of red blood cells and hemoglobin (Attie *et al.* 2014, Suragani *et al.* 2014).

Ultimately, in study III we evaluated the mechanistic effects of *SF3B1* K700E mutation in myeloid cells and explored if the mutation confers a loss-of-function in regards of gene expression and splicing of key genes in MDS-RS. Unexpectedly, we observed that both reduced and abrogated SF3B1 protein, using CRISPR/Cas9 technology, as well as expression of *SF3B1* K700E mutation at physiological levels, resulted in a typical MDS-RS gene expression pattern with reduced levels of *ABCB7* and altered exon usage, increased *ALAS2* expression and mis-spliced *TMEM14C*. To this end, our preliminary results suggest that the *SF3B1* mutation acts through a loss-of-function, however additional experiments are needed to clarify the role of these mutations. Although initial mice models of several of these mutations, as well as many insights into the global effects of *SF3B1* mutations on splicing have been described, it remains to be clarified if *SF3B1* mutations are causative of MDS-RS or if they are passenger mutations only contributing to a small extent to the disease progression. The answer may be found in targeted RNA sequencing using primary genetically modified normal BM cells; however progress is limited by the difficulty of

delivering the mutant and the wild-type SF3B1 in human progenitor cells.

Our and others studies have tried to identify pathways and key target genes linked to the disease phenotype in order to understand the mechanisms of the ineffective erythropoiesis. Along this line, functional studies should aim to validate findings and investigate further the key genes pinpointed by several recent genome wide studies. Interestingly, the fact that the *Abcb7* mouse model as well as *Sf3b1* knock-in mice (Mupo *et al.* 2016, Obeng *et al.* 2016) were unable to effectively generate ring sideroblasts may suggest an unknown difference between human and murine erythroid precursors with respect to mitochondrial iron metabolism. Also, common splicing events in MDS-RS were not observed in these models maybe because *SF3B1* mutations exert their effect on well-conserved functions of *SF3B1* unrelated to splicing. Alternatively, the iron accumulation may also be the effect of other mutations or biological mechanisms. For instance, *Obeng et al.* reported that loss of *Tet2* cooperates with $Sf3b1^{K700E}$ to worsen anemia, and causes an earlier expansion of long-term hematopoietic stem cells, but rescues the impaired competitive repopulating activity conferred by $Sf3b1^{K700E}$. The additive effect of *Tet2* may explain why these mutations are well tolerated in MDS patients, thereby making this combinatorial model extremely useful to recapitulate the mutational pattern of patients. Certainly, the high frequency of *SF3B1* mutations in MDS-RS patients have shifted the efforts towards the understanding of the splicing mechanisms and how they contribute to disease pathogenesis. Indeed, considering that SF3B1 is acting within a complex, it would be crucial to understand in further detail how mutations in *SF3B1* result in altered 3' SS usage, how *SF3B1* is directed to their pre-mRNA targets, how other proteins are interacting with SF3B1, and the structural and biochemical function of the HEAT domains of SF3B1, where the majority of *SF3B1* mutations are located. Thus, a proteomic approach would provide a mechanistic link between the mutant spliceosomal protein, alteration in the splicing and in impaired erythropoiesis, hence offering new therapeutic opportunities such as splicing modulators and other compounds.

6 ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to all who supported me during my Ph.D. First and foremost to my supervisors, who have taken the responsibility to guide me throughout these years; each has contributed with their own perspective, experience and skills to direct my development as a scientist.

To my main supervisor, **Eva Hellström-Lindberg**, for giving me the opportunity to pursue a Ph.D. in your group, for your guidance and support throughout my years as a student. In addition to your vast expertise on MDS, I have admired your enthusiasm, perseverance and your optimism, all of which have been instrumental during this time. Thank you for establishing such a warm, productive and creative research environment at HERM.

My co-supervisor **Julian Walfridsson**, for all the encouragement you have given me and for literally shaping my Ph.D. Your trust, constructive criticism, optimism and your sense of humor were greatly appreciated and always reassuring and made me feel like the confident student that I am today. Thank you for showing me that science is fun and "not about feelings" and for simply being an amazing co-supervisor.

My co-supervisor **Liselotte Vesterlund**, for always showing great interest in my projects and giving me valuable inputs and ideas. Thank you for being always helpful and promptly replying to my emails and for all the nice chats about life and future plans. You are truly great!

Petter Woll and **Sten Eirik Jacobsen**, thank you for the interesting scientific discussions and the constructive inputs during our joint meetings and MDS retreats.

Elli Papaemmanuil, it is no understatement to say that your work on *SF3B1* has revolutionized the entire field of MDS and my Ph.D. Thank you for sharing your expertise and knowledge throughout these years. All the encounters we had at conferences have encouraged me to keep on going and were a true source of inspiration.

Seishi Ogawa for being a great collaborator and for sharing your vast expertise in genetics, MDS and sequencing. Your help for my last manuscript is invaluable.

My co-authors, **Shintaro Katayama**, for your patience and the enormous help with the RNA sequencing data, **Per Unneberg** and **Juha Kere** for the assistance and the collaboration in paper II.

A sincere gratitude to the MDS group, for sharing every part of this journey from the very beginning, with its ups and downs. A special thanks to **Christian Scharenberg, Edda Elvarsdottir, Marios Dimitriou, Mohsen Karimi** and **Teresa Mortera-Blanco** for the scientific inputs, the support and your cheerful attitude, **Magnus Tobiasson**, for all the NBMs I asked you to take, **Lalla Forsblom**, **Maryam Nikpour** for welcoming me to the MDS group, **Asmaa Ben Azenkoud**, for your kindness, **Annette Engström** for your continuous help and for making sure that everything was ready for my defense way in advance. To **Gunilla Walldin,** for impeccably coordinating and arranging samples and for not giving up on my Swedish! Last but not least, to **Monika Jansson**, for your thoughtfulness, your generous help in the lab and for having a solution to almost all my problems.

Thanks to **Johanna Ungerstedt**, for your enthusiasm, your positive energy and your commitment in providing samples; **Martin Jädersten**, for your kindness, your interest in my projects and your valuable advises for my future career; **Evren Alici** for your motivational and friendly guidance since the very beginning of my Ph.D.; **Hong Qian** and **Robert Månsson**, for answering to all my questions anytime I knocked on your door; **Sören Lehman**, for the scientific inputs you have given me at my half time; **Petter Höglund** for accepting to be the chairman at my defense; **Iyadh Douagi** for helping me with the FACS and the sorting; **Katarina Le Blanc** for supporting a great research environment; **Cecilia Götherström** for kindly providing cord blood samples and **Nina Heldring**, for taking care of the SAP seminars and for always giving me constructive feedback.

From Lund University, thanks to **Niels Bjarne Woods**, for welcoming me in your lab and introducing me to the iPS cells; **Jonas Larsson** for providing us with indispensable vectors to paper III; **Sofie Singbrant**, for your kindness and your genuine interest in my projects and **Johan Flygare**, for accepting to be part of my half time committee and to critically revise my thesis. Your expertise on erythropoiesis will add a great value to my defense.

To **Astrid Olsnes Kittang** for warmly welcoming me in Bergen six years ago and for all the nice chats at the NMDS meetings.

To **Anethe Mansen** and **Tina Persson** at the KI Career Service, for the hard work and the strong commitment in helping students with their first career steps after PhD. Thank you for organizing the Career Skills for Scientist course, thanks to which I had the pleasure to interview **Maria Kakoulidou** and **Andreas Lennartsson** and to learn from their interesting and inspiring career stories.

To the Department of Medicine, and especially to **Jan Bolinder**, **Anastasia Urban, Therese Lind, Elenor Nyman**, **Klas Karlsson** and **Edgardo Faune** for easing the long bureaucracy and for helping me with any kind of technical issue.

I want to express my heartfelt appreciation to the members of **HERM.** You all have been part of this thesis in one way or another, thank you for bettering my day-to-day life even with a simple smile in the corridor and for sharing lunches, fika, julbord, afterworks, dinners, parties, tears, frustration and so many laughs! You all made this journey memorable! Thanks in particular to:
Adil Doğanay Duru, **Ayla De Paepe**, **Carin Dahlberg**, **Caroline Gavin**, **Makoto Kondo**, **Thibault Bouderlique** the best chef of HERM and **Sridharan Ganesan**, the true soul of HERM, for having been part of fun afterworks, parties and beautiful Swedish summer evenings; my "sister" **Monika Dolinska** for all the enjoyable conversation we had in the corridor and over a glass of wine; **Erle Refsum**, **Aleksandra Krstić** and **Eva Hellqvist** for your nice company at lunch and on the train; **Stephan Meinke** my spiritual supporter during the FACS staining; **Thuy Luu Thanh**, for your concern for others and for being always so positive; **Anton Törnqvist** and **Srinivasa Rao Nagubothu** for our nice encounters in the corridor and your genuine smiles; **Charlotte Gustafsson, Xiaoze Li, Lakshmi Sandhow**, **Pignan Xiao** and **Anne-Sofie Johansson**, for all the fun in the molecular lab and the good music; **Nadir Kadri**, for cheering me up with all your jokes and your loud and deep laughs; **Esmat Kamali Dolatabadi**, for the nice time we spent in the culture room and for letting me taste the amazing Persian *poolaki*; **Shabnam Kharazi**, for thinking of me even on vacation and for bringing me delicious jam; **Allegra Matheson Lord** and **Christer Nilsson** for keeping a nice atmosphere in the office; **Mari Gilljam**, **Birgitta Stellan** and **Tolga Sütlü** for being wonderful neighbors at KFC and for your help whenever I had questions; **Inger Vedin**, for your kindness

and wisdom; **Alf Grandien** and **Kari Högstrand**, for helping me with cloning issues; **Minna Suomela**, *kiitos* for being always so smiley and baking amazing cookies; **Karin Mellström** for giving me the chance to follow the GMP course; **Ann Wallblom**, **Sofia Bengtzen**, **Dimitra Vasileiadi, Lili Andersson** and **Linnea Acaralp** for your continuous assistance and help at HERM. To the new Hermi's **Giovanna Perinetti Casoni**, **Lamberto Torralba Raga**, **Martha-Lena Müller**, **Matthias Voss**, **Hongya Han**, **Lucia Pena Perez**, **Stina Virding** and **Desmond Chin** and the former graduate students **Lisamarie Roch**, **Athina Totomi**, **Flavia Peci**, **Kathleen Heck** and **David Chang**, that even only for short time, have largely contributed to the nice atmosphere at HERM.

A special thanks to the V.I.P. members of the "Oompa Loompa Virus Factory"…

Yaser Heshmati, my favorite young scientist! I truly admire your scientific independency as well as your positive energy and perseverance. Thank you for your constant support, your friendship and for teaching me so much in the lab. I am so "*felice*" to have worked with you. In a few years you will be a great PI with an amazing group! Best of luck Prof!

Aditya Harisankar, for your "contagious" positive energy, your help and kindness, for the loud metal music in the lab, for a culture room in the darkness, for thousands of after-works and coffees, for teaching me the rudiments of cloning and for wiping my sweat during the trans Blot.

Gözde Türköz, for your kindness and your genuine concern for others, for brain-storming with me during our CRISPR/Cas9 sessions, for our "tap-tap" trip to Lund… and for sharing the long days in the virus lab.

To my dear colleagues…

Hani Abdulkadir and **Deepika Nair** for your continuous support, for cheering me up anytime I needed, for the long and nice chats on our commute home, for being simply amazing and for reminding me this: "Everything will be fine in the end, if it's not fine, it's not the end". **Huthayfa Mujahed** it was great share the office with you, thank you for making our corner paper-free and for keeping the plants alive. **Hero Nikdin,** my beautiful "sare piaz", I am so grateful for the time we have spent at HERM. Thank you for your friendship and for accepting to be my toastmaster! It's finally time to celebrate my dear!!! **Michael Chrobok,** my favorite German and SFI buddy! This PhD was just the beginning of a beautiful friendship. **Ying Qu,** we have come a long way since you used to silently come to our office on the 6th floor. Thank you for having been there, "for better or for worse" and for all the times you have helped me with statistics and bioinformatics. **Stephanie Darmanin,** my true mentor! *Grazie di cuore* for the millions of times you have listened to me, encouraged me and given me precious advises.

…and my wonderful friends

Ivana Bulatović, when I first met you in the quiet corridor at KFC, I did not imagine you would have become one of the most important people of my life. I cannot imagine my time in Stockholm and my PhD without you on my side. They say, "hardship make or break people", but I can say that I made it thanks to you, my *draga prijateljice*!

Valentina Giai, I cannot express how much I value your friendship and how grateful I am for all you have done for me when I first arrived in Stockholm. August 20th 2010, through a big wall-glass in Arlanda was not only the beginning of my adventure in Stockholm, but the day I met an amazing friend.

Erica Miraglia, **Lorenza Capantini**, **Elisa Dioguardi**, **Valeria Messina** and **Valentina Pampanini** a.k.a Miranda, Lorenz, Fitu, Messina and Pampas, for your constant presence throughout these years. You have truly created a sense of belonging for which I feel very fortunate. *Grazie bimbe per aver condiviso davvero di tutto!*
Paola Violi and **Giuseppe Cappellano**, for helping me settle down at the beginning of my Swedish adventure. I will never forget our fun times in Flempan.
Paola Longati and **Rainer Heuchel,** it has always been a pleasure bumping into you two at NOVUM. Thank you for being there for me.
Thanks to my Finnish friends, **Fanni Salmela** for the laughs and the beers, and **Niklas von Lüders** for your bone marrow donations ☺ and for entertaining me with your incredible stories.
Gitta and **Jarmo,** *kiitos* for being part of my family in Stockholm, for your concern and kindness.

Thanks to my friends in Italy, for making me feel like if I had never left Turin…
Elisa Migliore, for being probably one of the reasons why I ended up in Sweden, **Costanza Frari, Corrado Laronga**, **Fiorella Ballerini**, for your authentic and precious friendship and **Leone Gioria,** for the unforgettable time in Göteborg with Bruce**.**

A **Michele Aliberti**, perché oggi potresti finalmente dire che "quella di Torino" è davvero una dottoressa.

To **Victor**, for being my biggest fan despite my emotional storms. Thank you for your love, patience and understanding.

A **Ilaria**, **mamma e papà**, dedico il lavoro e il successo di questa tesi. Non avrei mai raggiunto questo traguardo senza di voi. In questi ultimi anni abbiamo affrontato tante sfide, ma voi non avete mai perso il coraggio, anzi, siete stati la mia forza. Mi avete insegnato ad affrontare gli ostacoli e a rialzarmi più forte di prima. Sebbene la distanza che ci tiene separati l'uno dall'altra sia enorme, il vostro amore e supporto sono stati e saranno sempre il motore di tutto.

This thesis includes research financially supported by Karolinska Institutet, Swedish Cancer Society, Cancer Society in Stockholm and by the Scientific Research Council.

7 REFERENCES

Abdel-Wahab, O. and R. Levine (2011). "The spliceosome as an indicted conspirator in myeloid malignancies." *Cancer Cell* 20(4): 420-423.

Ahlqvist, K. J., S. Leoncini, A. Pecorelli, et al. (2015). "MtDNA mutagenesis impairs elimination of mitochondria during erythroid maturation leading to enhanced erythrocyte destruction." *Nat Commun* 6: 6494.

Allikmets, R., W. H. Raskind, A. Hutchinson, et al. (1999). "Mutation of a putative mitochondrial iron transporter gene (ABC7) in X-linked sideroblastic anemia and ataxia (XLSA/A)." *Hum Mol Genet* 8(5): 743-749.

Ambaglio, I., L. Malcovati, E. Papaemmanuil, et al. (2013). "Inappropriately low hepcidin levels in patients with myelodysplastic syndrome carrying a somatic mutation of SF3B1." *Haematologica* 98(3): 420-423.

An, X., V. P. Schulz, J. Li, et al. (2014). "Global transcriptome analyses of human and murine terminal erythroid differentiation." *Blood* 123(22): 3466-3477.

Arber, D. A., A. Orazi, R. Hasserjian, et al. (2016). "The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia." *Blood* 127(20): 2391-2405.

Argiropoulos, B. and R. K. Humphries (2007). "Hox genes in hematopoiesis and leukemogenesis." *Oncogene* 26(47): 6766-6776.

Attie, K. M., M. J. Allison, T. McClure, et al. (2014). "A phase 1 study of ACE-536, a regulator of erythroid differentiation, in healthy volunteers." *Am J Hematol* 89(7): 766-770.

Barrangou, R., A. Birmingham, S. Wiemann, et al. (2015). "Advances in CRISPR-Cas9 genome engineering: lessons learned from RNA interference." *Nucleic Acids Res* 43(7): 3407-3419.

Bayeva, M., A. Khechaduri, R. Wu, et al. (2013). "ATP-binding cassette B10 regulates early steps of heme synthesis." *Circ Res* 113(3): 279-287.

Bejar, R., K. E. Stevenson, B. A. Caughey, et al. (2012). "Validation of a prognostic model and the impact of mutations in patients with lower-risk myelodysplastic syndromes." *J Clin Oncol* 30(27): 3376-3382.

Bekri, S., G. Kispal, H. Lange, et al. (2000). "Human ABC7 transporter: gene structure and mutation causing X-linked sideroblastic anemia with ataxia with disruption of cytosolic iron-sulfur protein maturation." *Blood* 96(9): 3256-3264.

Bennett, J. M., D. Catovsky, M. T. Daniel, et al. (1982). "Proposals for the classification of the myelodysplastic syndromes." *Br J Haematol* 51(2): 189-199.

Billstrom, R., H. Johansson, B. Johansson, et al. (1995). "Immune-mediated complications in patients with myelodysplastic syndromes--clinical and cytogenetic features." *Eur J Haematol* 55(1): 42-48.

Bjorkman, S. E. (1956). "Chronic refractory anemia with sideroblastic bone marrow; a study of four cases." *Blood* 11(3): 250-259.

Boultwood, J., A. Pellagatti, M. Nikpour, et al. (2008). "The role of the iron transporter ABCB7 in refractory anemia with ring sideroblasts." *PLoS One* 3(4): e1970.

Bradley, T. R. and D. Metcalf (1966). "The growth of mouse bone marrow cells in vitro." *Aust J Exp Biol Med Sci* 44(3): 287-299.

Broseus, J., T. Alpermann, M. Wulfert, et al. (2013). "Age, JAK2(V617F) and SF3B1 mutations are the main predicting factors for survival in refractory anaemia with ring sideroblasts and marked thrombocytosis." *Leukemia* 27(9): 1826-1831.

Camaschella, C. (2009). "Hereditary sideroblastic anemias: pathophysiology, diagnosis, and treatment." *Semin Hematol* 46(4): 371-377.

Camaschella, C. and A. Pagani (2011). "Iron and erythropoiesis: a dual relationship." *Int J Hematol* 93(1): 21-26.

Camaschella, C., A. Pagani, A. Nai, et al. (2016). "The mutual control of iron and erythropoiesis." *Int J Lab Hematol* 38 Suppl 1: 20-26.

Cattoglio, C., G. Facchini, D. Sartori, et al. (2007). "Hot spots of retroviral integration in human CD34+ hematopoietic cells." *Blood* 110(6): 1770-1778.

Cavadini, P., G. Biasiotto, M. Poli, et al. (2007). "RNA silencing of the mitochondrial ABCB7 transporter in HeLa cells causes an iron-deficient phenotype with mitochondrial iron overload." *Blood* 109(8): 3552-3559.

Cazzola, M., R. Invernizzi, G. Bergamaschi, et al. (2003). "Mitochondrial ferritin expression in erythroid cells from patients with sideroblastic anemia." *Blood* 101(5): 1996-2000.

Cazzola, M., M. G. Della Porta and L. Malcovati (2013a). "The genetic basis of myelodysplasia and its clinical relevance." *Blood* 122(25): 4021-4034.

Cazzola, M., M. Rossi, L. Malcovati, et al. (2013b). "Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms." *Blood* 121(2): 260-269.

Chae, J. H. and C. G. Kim (2003). "CP2 binding to the promoter is essential for the enhanced transcription of globin genes in erythroid cells." *Mol Cells* 15(1): 40-47.

Cheng, A. W., J. Shi, P. Wong, et al. (2014). "Muscleblind-like 1 (Mbnl1) regulates pre-mRNA alternative splicing during terminal erythropoiesis." *Blood* 124(4): 598-610.

Chiabrando, D., S. Mercurio and E. Tolosano (2014). "Heme and erythropoieis: more than a structural role." *Haematologica* 99(6): 973-983.

Conte, S., S. Katayama, L. Vesterlund, et al. (2015). "Aberrant splicing of genes involved in haemoglobin synthesis and impaired terminal erythroid maturation in SF3B1 mutated refractory anaemia with ring sideroblasts." *Br J Haematol* 171(4): 478-490.

Damm, F., O. Kosmider, V. Gelsi-Boyer, et al. (2012a). "Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes." *Blood* 119(14): 3211-3218.

Damm, F., F. Thol, O. Kosmider, et al. (2012b). "SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications." *Leukemia* 26(5): 1137-1140.

Darman, R. B., M. Seiler, A. A. Agrawal, et al. (2015). "Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point." *Cell Rep* 13(5): 1033-1045.

DeBoever, C., E. M. Ghia, P. J. Shepard, et al. (2015). "Transcriptome sequencing reveals potential mechanism of cryptic 3' splice site selection in SF3B1-mutated cancers." *PLoS Comput Biol* 11(3): e1004105.

Dolatshad, H., A. Pellagatti, M. Fernandez-Mercado, et al. (2015). "Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells." *Leukemia* 29(8): 1798.

Dolatshad, H., A. Pellagatti, F. G. Liberante, et al. (2016). "Cryptic splicing events in the iron transporter ABCB7 and other key target genes in SF3B1 mutant myelodysplastic syndromes." *Leukemia*.

Donker, A. E., R. A. Raymakers, L. T. Vlasveld, et al. (2014). "Practice guidelines for the diagnosis and management of microcytic anemias due to genetic disorders of iron metabolism or heme synthesis." *Blood* 123(25): 3873-3886; quiz 4005.

Dussiot, M., T. T. Maciel, A. Fricot, et al. (2014). "An activin receptor IIA ligand trap corrects ineffective erythropoiesis in beta-thalassemia." *Nat Med* 20(4): 398-407.

Elvarsdóttir, E., T. Mortera-Blanco, M. Dimitriou, et al. (2016). A culture model mimicking the BM-niche allows for studying the Dysregulated erythropoiesis of SF3B1 mutated

myelodyspastic syndrome with ringed sideroblasts. European Hematology Association. Copenhagen, EHA Learning Center.

Fenaux, P. and L. Ades (2013). "How we treat lower-risk myelodysplastic syndromes." *Blood* 121(21): 4280-4286.

Fenaux, P., D. Haase, G. F. Sanz, et al. (2014). "Myelodysplastic syndromes: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up." *Ann Oncol* 25 Suppl 3: iii57-69.

Genovese, P., G. Schiroli, G. Escobar, et al. (2014). "Targeted genome editing in human repopulating haematopoietic stem cells." *Nature* 510(7504): 235-240.

Ghoti, H., J. Amer, A. Winder, et al. (2007). "Oxidative stress in red blood cells, platelets and polymorphonuclear leukocytes from patients with myelodysplastic syndrome." *Eur J Haematol* 79(6): 463-467.

Giagounidis, A., U. Platzbecker, U. Germing, et al. (2015). "Luspatercept Treatment Leads to Long Term Increases in Hemoglobin and Reductions in Transfusion Burden in Patients with Low or Intermediate-1 Risk Myelodysplastic Syndromes (MDS): Preliminary Results from the Phase 2 PACE-MDS Extension Study." *Blood* 126(23): 92-92.

Goh, S. H., Y. T. Lee, N. V. Bhanu, et al. (2005). "A newly discovered human alpha-globin gene." *Blood* 106(4): 1466-1472.

Goulter, A. B., D. W. Harmer and K. L. Clark (2006). "Evaluation of low density array technology for quantitative parallel measurement of multiple genes in human tissue." *BMC Genomics* 7: 34.

Greenberg, P., C. Cox, M. M. LeBeau, et al. (1997). "International scoring system for evaluating prognosis in myelodysplastic syndromes." *Blood* 89(6): 2079-2088.

Greenberg, P. L., H. Tuechler, J. Schanz, et al. (2012). "Revised international prognostic scoring system for myelodysplastic syndromes." *Blood* 120(12): 2454-2465.

Guernsey, D. L., H. Jiang, D. R. Campagna, et al. (2009). "Mutations in mitochondrial carrier family gene SLC25A38 cause nonsyndromic autosomal recessive congenital sideroblastic anemia." *Nat Genet* 41(6): 651-653.

Haferlach, T., Y. Nagata, V. Grossmann, et al. (2014). "Landscape of genetic lesions in 944 patients with myelodysplastic syndromes." *Leukemia* 28(2): 241-247.

Hahn, C. N., P. Venugopal, H. S. Scott, et al. (2015). "Splice factor mutations and alternative splicing as drivers of hematopoietic malignancy." *Immunol Rev* 263(1): 257-278.

Han, A. P., C. Yu, L. Lu, et al. (2001). "Heme-regulated eIF2alpha kinase (HRI) is required for translational regulation and survival of erythroid precursors in iron deficiency." *EMBO J* 20(23): 6909-6918.

Hardison, R. C., D. H. Chui, B. Giardine, et al. (2002). "HbVar: A relational database of human hemoglobin variants and thalassemia mutations at the globin gene server." *Hum Mutat* 19(3): 225-233.

Harrison, P. M. and P. Arosio (1996). "The ferritins: molecular properties, iron storage function and cellular regulation." *Biochim Biophys Acta* 1275(3): 161-203.

Hattangadi, S. M., P. Wong, L. Zhang, et al. (2011). "From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications." *Blood* 118(24): 6258-6268.

Hellen, C. U. and P. Sarnow (2001). "Internal ribosome entry sites in eukaryotic mRNA molecules." *Genes Dev* 15(13): 1593-1612.

Hellstrom-Lindberg, E. and A. van de Loosdrecht (2013). "Erythropoiesis stimulating agents and other growth factors in low-risk MDS." *Best Pract Res Clin Haematol* 26(4): 401-410.

Hentze, M. W., M. U. Muckenthaler, B. Galy, et al. (2010). "Two to tango: regulation of Mammalian iron metabolism." *Cell* 142(1): 24-38.

Hofmann, W. K., S. de Vos, M. Komor, et al. (2002). "Characterization of gene expression of CD34+ cells from normal and myelodysplastic bone marrow." *Blood* 100(10): 3553-3560.

Holme, H., U. Hossain, M. Kirwan, et al. (2012). "Marked genetic heterogeneity in familial myelodysplasia/acute myeloid leukaemia." *Br J Haematol* 158(2): 242-248.

Inoue, D., R. K. Bradley and O. Abdel-Wahab (2016). "Spliceosomal gene mutations in myelodysplasia: molecular links to clonal abnormalities of hematopoiesis." *Genes Dev* 30(9): 989-1001.

Invernizzi, R., E. Travaglino, M. G. Della Porta, et al. (2013). "Effects of mitochondrial ferritin overexpression in normal and sideroblastic erythroid progenitors." *Br J Haematol* 161(5): 726-737.

Isono, K., Y. Mizutani-Koseki, T. Komori, et al. (2005). "Mammalian polycomb-mediated repression of Hox genes requires the essential spliceosomal protein Sf3b1." *Genes Dev* 19(5): 536-541.

Jadersten, M., S. M. Montgomery, I. Dybedal, et al. (2005). "Long-term outcome of treatment of anemia in MDS with erythropoietin and G-CSF." *Blood* 106(3): 803-811.

Jadersten, M., L. Malcovati, I. Dybedal, et al. (2008). "Erythropoietin and granulocyte-colony stimulating factor treatment associated with improved survival in myelodysplastic syndrome." *J Clin Oncol* 26(21): 3607-3613.

Jaffe, E. S., H. S. Harris, J.W.; and V. (eds) (2001). *Pathology and genetics of tumours of haematopoietic and lymphoid tissues*, Iarc.

Jiang, Y., A. Dunbar, L. P. Gondek, et al. (2009). "Aberrant DNA methylation is a dominant mechanism in MDS progression to AML." *Blood* 113(6): 1315-1325.

Kang, H. C., J. H. Chae, Y. H. Lee, et al. (2005). "Erythroid cell-specific alpha-globin gene regulation by the CP2 transcription factor family." *Mol Cell Biol* 25(14): 6005-6020.

Kfir, N., G. Lev-Maor, O. Glaich, et al. (2015). "SF3B1 association with chromatin determines splicing outcomes." *Cell Rep* 11(4): 618-629.

Komrokji, R. and J. M. Bennett (2003). "The myelodysplastic syndromes: classification and prognosis." *Curr Hematol Rep* 2(3): 179-185.

Komrokji, R. S., G. Garcia-Manero, L. Ades, et al. (2014). "An Open-Label, Phase 2, Dose-Finding Study of Sotatercept (ACE-011) in Patients with Low or Intermediate-1 (Int-1)-Risk Myelodysplastic Syndromes (MDS) or Non-Proliferative Chronic Myelomonocytic Leukemia (CMML) and Anemia Requiring Transfusion." *Blood* 124(21): 3251-3251.

Larson, M. H., L. A. Gilbert, X. Wang, et al. (2013). "CRISPR interference (CRISPRi) for sequence-specific control of gene expression." *Nat Protoc* 8(11): 2180-2196.

Levi, S., B. Corsi, M. Bosisio, et al. (2001). "A human mitochondrial ferritin encoded by an intronless gene." *J Biol Chem* 276(27): 24437-24440.

Liesa, M., W. Qiu and O. S. Shirihai (2012). "Mitochondrial ABC transporters function: the role of ABCB10 (ABC-me) as a novel player in cellular handling of reactive oxygen species." *Biochim Biophys Acta* 1823(10): 1945-1957.

Lill, R., K. Diekert, A. Kaut, et al. (1999). "The essential role of mitochondria in the biogenesis of cellular iron-sulfur proteins." *Biol Chem* 380(10): 1157-1166.

Livak, K. J. and T. D. Schmittgen (2001). "Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method." *Methods* 25(4): 402-408.

Ljung, T., R. Back and E. Hellstrom-Lindberg (2004). "Hypochromic red blood cells in low-risk myelodysplastic syndromes: effects of treatment with hemopoietic growth factors." *Haematologica* 89(12): 1446-1453.

Lu, B., J. Xu, J. Chen, et al. (2008). "TaqMan low density array is roughly right for gene expression quantification in colorectal cancer." *Clin Chim Acta* 389(1-2): 146-151.

Malcovati, L., E. Papaemmanuil, D. T. Bowen, et al. (2011). "Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms." *Blood* 118(24): 6239-6246.

Malcovati, L., E. Hellstrom-Lindberg, D. Bowen, et al. (2013). "Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet." *Blood* 122(17): 2943-2964.

Malcovati, L., M. Karimi, E. Papaemmanuil, et al. (2015). "SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts." *Blood* 126(2): 233-241.

Malcovati, L. and M. Cazzola (2016). "Recent advances in the understanding of myelodysplastic syndromes with ring sideroblasts." *Br J Haematol*.

Maquat, L. E. (2004). "Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics." *Nat Rev Mol Cell Biol* 5(2): 89-99.

Mardis, E. R. (2008). "The impact of next-generation sequencing technology on genetics." *Trends Genet* 24(3): 133-141.

Marisavljevic, D., N. Kraguljac and Z. Rolovic (2006). "Immunologic abnormalities in myelodysplastic syndromes: clinical features and characteristics of the lymphoid population." *Med Oncol* 23(3): 385-391.

Matsunawa, M., R. Yamamoto, M. Sanada, et al. (2014). "Haploinsufficiency of Sf3b1 leads to compromised stem cell function but not to myelodysplasia." *Leukemia* 28(9): 1844-1850.

Merryweather-Clarke, A. T., A. Atzberger, S. Soneji, et al. (2011). "Global gene expression analysis of human erythroid progenitors." *Blood* 117(13): e96-108.

Metzker, M. L. (2010). "Sequencing technologies - the next generation." *Nat Rev Genet* 11(1): 31-46.

Mian, S. A., A. E. Smith, A. G. Kulasekararaj, et al. (2013). "Spliceosome mutations exhibit specific associations with epigenetic modifiers and proto-oncogenes mutated in myelodysplastic syndrome." *Haematologica* 98(7): 1058-1066.

Mitchell, M., S. D. Gore and A. M. Zeidan (2013). "Iron chelation therapy in myelodysplastic syndromes: where do we stand?" *Expert Rev Hematol* 6(4): 397-410.

Miyata, M. and J. D. Smith (1996). "Apolipoprotein E allele-specific antioxidant activity and effects on cytotoxicity by oxidative insults and beta-amyloid peptides." *Nat Genet* 14(1): 55-61.

Modlich, U., S. Navarro, D. Zychlinski, et al. (2009). "Insertional transformation of hematopoietic cells by self-inactivating lentiviral and gammaretroviral vectors." *Mol Ther* 17(11): 1919-1928.

Mortera-Blanco, T. D., M.; Woll, P.; Elvarsdottir, E.; Karimi, M.; Conte, S.; Jansson, M.; Douagi,I.; Papaemmanuil, E.; Jacobsen, SE., Hellström-Lindberg,E. (2016). SF3B1 mutations in MDS-RS arise in multipotent hemopoietic stem cells European Hematology Association Copenhagen, EHA Learning Center.

Mufti, G. J., J. M. Bennett, J. Goasguen, et al. (2008). "Diagnosis and classification of myelodysplastic syndrome: International Working Group on Morphology of myelodysplastic syndrome (IWGM-MDS) consensus proposals for the definition and enumeration of myeloblasts and ring sideroblasts." *Haematologica* 93(11): 1712-1717.

Mupo, A., M. Seiler, V. Sathiaseelan, et al. (2016). "Hemopoietic-specific Sf3b1-K700E knock-in mice display the splicing defect seen in human MDS but develop anemia without ring sideroblasts." *Leukemia*.

Naldini, L., U. Blomer, F. H. Gage, et al. (1996a). "Efficient transfer, integration, and sustained long-term expression of the transgene in adult rat brains injected with a lentiviral vector." *Proc Natl Acad Sci U S A* 93(21): 11382-11388.

Naldini, L., U. Blomer, P. Gallay, et al. (1996b). "In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector." *Science* 272(5259): 263-267.

Neukirchen, J., W. M. Schoonen, C. Strupp, et al. (2011). "Incidence and prevalence of myelodysplastic syndromes: data from the Dusseldorf MDS-registry." *Leuk Res* 35(12): 1591-1596.

Nie, G., A. D. Sheftel, S. F. Kim, et al. (2005). "Overexpression of mitochondrial ferritin causes cytosolic iron depletion and changes cellular iron homeostasis." *Blood* 105(5): 2161-2167.

Nikpour, M., A. Pellagatti, A. Liu, et al. (2010). "Gene expression profiling of erythroblasts from refractory anaemia with ring sideroblasts (RARS) and effects of G-CSF." *Br J Haematol* 149(6): 844-854.

Nikpour, M., C. Scharenberg, A. Liu, et al. (2013). "The transporter ABCB7 is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts." *Leukemia* 27(4): 889-896.

Nilsson, R., I. J. Schultz, E. L. Pierce, et al. (2009). "Discovery of genes essential for heme biosynthesis through large-scale gene expression analysis." *Cell Metab* 10(2): 119-130.

Nisse, C., C. Lorthois, V. Dorp, et al. (1995). "Exposure to occupational and environmental factors in myelodysplastic syndromes. Preliminary results of a case-control study." *Leukemia* 9(4): 693-699.

Novotna, B., Y. Bagryantseva, M. Siskova, et al. (2009). "Oxidative DNA damage in bone marrow cells of patients with low-risk myelodysplastic syndrome." *Leuk Res* 33(2): 340-343.

Obeng, E. A., R. J. Chappell, M. Seiler, et al. (2016). "Physiologic Expression of Sf3b1(K700E) Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation." *Cancer Cell* 30(3): 404-417.

Palis, J., S. Robertson, M. Kennedy, et al. (1999). "Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse." *Development* 126(22): 5073-5084.

Papaemmanuil, E., M. Cazzola, J. Boultwood, et al. (2011). "Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts." *N Engl J Med* 365(15): 1384-1395.

Papaemmanuil, E., M. Gerstung, L. Malcovati, et al. (2013). "Clinical and biological implications of driver mutations in myelodysplastic syndromes." *Blood* 122(22): 3616-3627; quiz 3699.

Patnaik, M. M., T. L. Lasho, J. M. Hodnefield, et al. (2012). "SF3B1 mutations are prevalent in myelodysplastic syndromes with ring sideroblasts but do not hold independent prognostic value." *Blood* 119(2): 569-572.

Patnaik, M. M. and A. Tefferi (2015). "Refractory anemia with ring sideroblasts and RARS with thrombocytosis." *Am J Hematol* 90(6): 549-559.

Pellagatti, A., M. Cazzola, A. A. Giagounidis, et al. (2006). "Gene expression profiles of CD34+ cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype." *Blood* 108(1): 337-345.

Pellagatti, A., M. Jadersten, A. M. Forsblom, et al. (2007). "Lenalidomide inhibits the malignant clone and up-regulates the SPARC gene mapping to the commonly deleted region in 5q- syndrome patients." *Proc Natl Acad Sci U S A* 104(27): 11406-11411.

Pimentel, H., M. Parra, S. L. Gee, et al. (2016). "A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis." *Nucleic Acids Res* 44(2): 838-851.

Pluznik, D. H. and L. Sachs (1966). "The induction of clones of normal mast cells by a substance from conditioned medium." *Exp Cell Res* 43(3): 553-563.

Pondarre, C., B. B. Antiochos, D. R. Campagna, et al. (2006). "The mitochondrial ATP-binding cassette transporter Abcb7 is essential in mice and participates in cytosolic iron-sulfur cluster biogenesis." *Hum Mol Genet* 15(6): 953-964.

Pondarre, C., D. R. Campagna, B. Antiochos, et al. (2007). "Abcb7, the gene responsible for X-linked sideroblastic anemia with ataxia, is essential for hematopoiesis." *Blood* 109(8): 3567-3569.

Pozdnyakova, O., P. M. Miron, G. Tang, et al. (2008). "Cytogenetic abnormalities in a series of 1,029 patients with primary myelodysplastic syndromes: a report from the US with a focus on some undefined single chromosomal abnormalities." *Cancer* 113(12): 3331-3340.

Quesada, V., L. Conde, N. Villamor, et al. (2012). "Exome sequencing identifies recurrent mutations of the splicing factor SF3B1 gene in chronic lymphocytic leukemia." *Nat Genet* 44(1): 47-52.

Raje, N. and S. Vallet (2010). "Sotatercept, a soluble activin receptor type 2A IgG-Fc fusion protein for the treatment of anemia and bone loss." *Curr Opin Mol Ther* 12(5): 586-597.

Ran, F. A., P. D. Hsu, J. Wright, et al. (2013). "Genome engineering using the CRISPR-Cas9 system." *Nat Protoc* 8(11): 2281-2308.

Rouault, T. A. and W. H. Tong (2008). "Iron-sulfur cluster biogenesis and human disease." *Trends Genet* 24(8): 398-407.

Russler-Germain, D. A., D. H. Spencer, M. A. Young, et al. (2014). "The R882H DNMT3A mutation associated with AML dominantly inhibits wild-type DNMT3A by blocking its ability to form active tetramers." *Cancer Cell* 25(4): 442-454.

Saito, Y., T. Hayashi, A. Tanaka, et al. (1999). "Selenoprotein P in human plasma as an extracellular phospholipid hydroperoxide glutathione peroxidase. Isolation and enzymatic characterization of human selenoprotein p." *J Biol Chem* 274(5): 2866-2871.

Salmon, P., V. Kindler, O. Ducrey, et al. (2000). "High-level transgene expression in human hematopoietic progenitors and differentiated blood lineages after transduction with improved lentiviral vectors." *Blood* 96(10): 3392-3398.

Santambrogio, P., B. G. Erba, A. Campanella, et al. (2011). "Over-expression of mitochondrial ferritin affects the JAK2/STAT5 pathway in K562 cells and causes mitochondrial iron accumulation." *Haematologica* 96(10): 1424-1432.

Santini, V. (2015). "Anemia as the Main Manifestation of Myelodysplastic Syndromes." *Semin Hematol* 52(4): 348-356.

Savage, K. I., J. J. Gorski, E. M. Barros, et al. (2014). "Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability." *Mol Cell* 54(3): 445-459.

Schanz, J., H. Tuchler, F. Sole, et al. (2012). "New comprehensive cytogenetic scoring system for primary myelodysplastic syndromes (MDS) and oligoblastic acute myeloid

leukemia after MDS derived from an international database merge." *J Clin Oncol* 30(8): 820-829.

Schellenberg, M. J., E. L. Dul and A. M. MacMillan (2011). "Structural model of the p14/SF3b155 . branch duplex complex." *RNA* 17(1): 155-165.

Schmid, M. (2009). "Iron chelation therapy in MDS: what have we learnt recently?" *Blood Rev* 23 Suppl 1: S21-25.

Schmittgen, T. D. and K. J. Livak (2008). "Analyzing real-time PCR data by the comparative C(T) method." *Nat Protoc* 3(6): 1101-1108.

Shen, T., M. Zheng, C. Cao, et al. (2007). "Mitofusin-2 is a major determinant of oxidative stress-mediated heart muscle cell apoptosis." *J Biol Chem* 282(32): 23354-23361.

Shiozawa, Y., L. Malcovati, A. Sato-Otsubo, et al. (2015). "Different Mutant Splicing Factors Cause Distinct Missplicing Events and Give Rise to Different Clinical Phenotypes in Myelodysplastic Syndromes." *Blood* 126(23): 139-139.

Steensma, D. P., K. A. Hecksel, J. C. Porcher, et al. (2007). "Candidate gene mutation analysis in idiopathic acquired sideroblastic anemia (refractory anemia with ringed sideroblasts)." *Leuk Res* 31(5): 623-628.

Steensma, D. P. (2011). "The role of iron chelation therapy for patients with myelodysplastic syndromes." *J Natl Compr Canc Netw* 9(1): 65-75.

Steensma, D. P. and N. Gattermann (2013). "When is iron overload deleterious, and when and how should iron chelation therapy be administered in myelodysplastic syndromes?" *Best Pract Res Clin Haematol* 26(4): 431-444.

Strom, S. S., Y. Gu, S. K. Gruschkus, et al. (2005). "Risk factors of myelodysplastic syndromes: a case-control study." *Leukemia* 19(11): 1912-1918.

Sugioka, R., S. Shimizu and Y. Tsujimoto (2004). "Fzo1, a protein involved in mitochondrial fusion, inhibits apoptosis." *J Biol Chem* 279(50): 52726-52734.

Suragani, R. N., S. M. Cadena, S. M. Cawley, et al. (2014). "Transforming growth factor-beta superfamily ligand trap ACE-536 corrects anemia by promoting late-stage erythropoiesis." *Nat Med* 20(4): 408-414.

Tehranchi, R., B. Fadeel, A. M. Forsblom, et al. (2003). "Granulocyte colony-stimulating factor inhibits spontaneous cytochrome c release and mitochondria-dependent apoptosis of myelodysplastic syndrome hematopoietic progenitors." *Blood* 101(3): 1080-1086.

Tehranchi, R., B. Fadeel, J. Schmidt-Mende, et al. (2005a). "Antiapoptotic role of growth factors in the myelodysplastic syndromes: concordance between in vitro and in vivo observations." *Clin Cancer Res* 11(17): 6291-6299.

Tehranchi, R., R. Invernizzi, A. Grandien, et al. (2005b). "Aberrant mitochondrial iron distribution and maturation arrest characterize early erythroid precursors in low-risk myelodysplastic syndromes." *Blood* 106(1): 247-253.

Testa, U. (2004). "Apoptotic mechanisms in the control of erythropoiesis." *Leukemia* 18(7): 1176-1199.

Thol, F., S. Kade, C. Schlarmann, et al. (2012). "Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes." *Blood* 119(15): 3578-3584.

Visconte, V., H. J. Rogers, J. Singh, et al. (2012). "SF3B1 haploinsufficiency leads to formation of ring sideroblasts in myelodysplastic syndromes." *Blood* 120(16): 3173-3186.

Visconte, V., A. Tabarroki, L. Zhang, et al. (2014). "Splicing factor 3b subunit 1 (Sf3b1) haploinsufficient mice display features of low risk Myelodysplastic syndromes with ring sideroblasts." *J Hematol Oncol* 7: 89.

Visconte, V., N. Avishai, R. Mahfouz, et al. (2015). "Distinct iron architecture in SF3B1-mutant myelodysplastic syndrome patients is linked to an SLC25A37 splice variant with a retained intron." *Leukemia* 29(1): 188-195.

von Lindern, M. (2006). "Cell-cycle control in erythropoiesis." *Blood* 108(3): 781-782.

Wang, C., G. Sashida, A. Saraya, et al. (2014). "Depletion of Sf3b1 impairs proliferative capacity of hematopoietic stem cells but is not sufficient to induce myelodysplasia." *Blood* 123(21): 3336-3343.

Wang, Z., M. Gerstein and M. Snyder (2009). "RNA-Seq: a revolutionary tool for transcriptomics." *Nat Rev Genet* 10(1): 57-63.

Webb, T. R., A. S. Joyner and P. M. Potter (2013). "The development and application of small molecule modulators of SF3b as therapeutic agents for cancer." *Drug Discov Today* 18(1-2): 43-49.

West, A. H., L. A. Godley and J. E. Churpek (2014). "Familial myelodysplastic syndrome/acute leukemia syndromes: a review and utility for translational investigations." *Ann N Y Acad Sci* 1310: 111-118.

Willekens, C., F. Dumezy, T. Boyer, et al. (2013). "Linezolid induces ring sideroblasts." *Haematologica* 98(11): e138-140.

Wojda, U., P. Noel and J. L. Miller (2002). "Fetal and adult hemoglobin production during adult erythropoiesis: coordinate expression correlates with cell proliferation." *Blood* 99(8): 3005-3013.

Yang, Z., S. B. Keel, A. Shimamura, et al. (2016). "Delayed globin synthesis leads to excess heme and the macrocytic anemia of Diamond Blackfan anemia and del(5q) myelodysplastic syndrome." *Sci Transl Med* 8(338): 338ra367.

Ye, H., S. Y. Jeong, M. C. Ghosh, et al. (2010). "Glutaredoxin 5 deficiency causes sideroblastic anemia by specifically impairing heme biosynthesis and depleting cytosolic iron in human erythroblasts." *J Clin Invest* 120(5): 1749-1761.

Yien, Y. Y., R. F. Robledo, I. J. Schultz, et al. (2014). "TMEM14C is required for erythroid mitochondrial heme metabolism." *J Clin Invest* 124(10): 4294-4304.

Yien, Y. Y., A. R. Ringel and B. H. Paw (2015). "Mitochondrial transport of protoporphyrinogen IX in erythroid cells." *Oncotarget* 6(25): 20742-20743.

Yoshida, K., M. Sanada, Y. Shiraishi, et al. (2011). "Frequent pathway mutations of splicing machinery in myelodysplasia." *Nature* 478(7367): 64-69.

Yoshida, K. and S. Ogawa (2014). "Splicing factor mutations and cancer." *Wiley Interdiscip Rev RNA* 5(4): 445-459.

Yusuke, S. M., L.; Sato-Otsubo, A.; Gallì, A.; Yoshida,K.; Yoshizato, T.; Sato, Y.; Kataoka, K.; Sanada, M.; Makishima, H.; Shiraishi, Y.; Chiba, K.; Miyano, S.; Hellström Lindberg, E.; Ogawa, S.; Cazzola, M. (2015). "Different Mutant Splicing Factors Cause Distinct Missplicing Events and Give Rise to Different Clinical Phenotypes in Myelodysplastic Syndromes." *Blood* 126

Zhou, Q., A. Derti, D. Ruddy, et al. (2015). "A chemical genetics approach for the functional assessment of novel cancer genes." *Cancer Res* 75(10): 1949-1958.