bjh research paper

Aberrant splicing of genes involved in haemoglobin synthesis and impaired terminal erythroid maturation in *SF3B1* mutated refractory anaemia with ring sideroblasts

Simona Conte,[1] Shintaro Katayama,[2] Liselotte Vesterlund,[2] Mohsen Karimi,[1] Marios Dimitriou,[1] Monika Jansson,[1] Teresa Mortera-Blanco,[1] Per Unneberg,[3] Elli Papaemmanuil,[4] Birgitta Sander,[5] Tiina Skoog,[2] Peter Campbell,[4] Julian Walfridsson,[1] Juha Kere[2] and Eva Hellström-Lindberg[1]

[1]Karolinska Institutet, Department of Medicine (Huddinge), Centre for Hematology and Regenerative Medicine, Stockholm, Sweden, [2]Karolinska Institutet, Department of Biosciences and Nutrition and Center for Innovative Medicine, Stockholm, Sweden, [3]Department of Biochemistry and Biophysics, Science for Life Laboratory, Stockholm, Sweden, [4]Cancer Genetics & Genomics, The Wellcome Trust Sanger Institute, Hinxton, Cambridge, UK and [5]Karolinska Institutet, Department of Laboratory Medicine, Division of Pathology, Stockholm, Sweden

Received 22 April 2015; accepted for publication 25 June 2015

Correspondence: Eva Hellström-Lindberg, Karolinska Institutet, Department of Medicine, Center for Hematology and Regenerative Medicine, Karolinska University Hospital Huddinge, 141 86 Stockholm, Sweden.
E-mail: eva.hellstrom-lindberg@ki.se

Summary

Refractory anaemia with ring sideroblasts (RARS) is distinguished by hyperplastic inefficient erythropoiesis, aberrant mitochondrial ferritin accumulation and anaemia. Heterozygous mutations in the spliceosome gene *SF3B1* are found in a majority of RARS cases. To explore the link between *SF3B1* mutations and anaemia, we studied mutated RARS $CD34^+$ marrow cells with regard to transcriptome sequencing, splice patterns and mutational allele burden during erythroid differentiation. Transcriptome profiling during early erythroid differentiation revealed a marked up-regulation of genes involved in haemoglobin synthesis and in the oxidative phosphorylation process, and down-regulation of mitochondrial ABC transporters compared to normal bone marrow. Moreover, mis-splicing of genes involved in transcription regulation, particularly haemoglobin synthesis, was confirmed, indicating a compromised haemoglobinization during RARS erythropoiesis. In order to define the phase during which erythroid maturation of *SF3B1* mutated cells is most affected, we assessed allele burden during erythroid differentiation *in vitro* and *in vivo* and found that *SF3B1* mutated erythroblasts showed stable expansion until late erythroblast stage but that terminal maturation to reticulocytes was significantly reduced. In conclusion, *SF3B1* mutated RARS progenitors display impaired splicing with potential downstream consequences for genes of key importance for haemoglobin synthesis and terminal erythroid differentiation.

Keywords: myelodysplastic syndromes, erythropoiesis, refractory anaemia with ring sideroblasts, *SF3B1*.

Myelodysplastic syndromes (MDS) constitute a group of clonal stem cell disorders distinguished by pancytopenia, ineffective haematopoiesis and a predisposition to acute myeloid leukaemia (AML) (Hellstrom-Lindberg, 2005). By contrast, the MDS subgroup refractory anaemia with ring sideroblasts (RARS) is characterized by a stable clinical course, isolated anaemia, expanded erythropoiesis and the morphological presence of ring sideroblasts (RS) (Mufti *et al*, 2008). Sideroblast formation starts during early erythroid maturation, when iron accumulates in mitochondria in the form of aberrant mitochondrial ferritin (FTMT) (Cazzola *et al*, 2003; Tehranchi *et al*, 2005a). Somatic mutations in *SF3B1*, a core component of the RNA splicing machinery were initially found in more than 70% of RARS patients (Papaemmanuil *et al*, 2011). In parallel, Yoshida *et al* (2011) reported additional mutations affecting multiple components of the RNA splicing machinery in MDS (*U2AF35*, *ZRSR2*, *SRSF2* and *SF3B1*). Of these, *SF3B1* mutations alone had a positive predictive value for disease phenotypes with ring sideroblasts of 97·7% (Malcovati *et al*, 2011). Interestingly, *SF3B1* mutated samples show under-expressed mRNA profiles with emphasis on mitochondrial pathways (Papaemmanuil *et al*, 2011).

More recently we reported mutational data from a cohort of 159 morphologically well-defined patients with RARS and

First published online 10 August 2015
doi: 10.1111/bjh.13610

refractory cytopenia with multilineage dysplasia and ring sideroblasts (RCMD-RS) and showed (81/90) 90% *SF3B1* mutations in the former category and (48/67) 72% in the latter, if patients with *TP53* mutations were removed. Moreover, no differences in survival or disease progression were observed between RARS or RCMD-RS patients with sole *SF3B1* mutations, and those with additional mutations in the two main other genes *TET2* and *DNMT3A*. *SF3B1* wild type patients showed a varying mutational pattern; 10/33 (33%) had other splice factor mutations (*SRSF2, U2AF1, ZRSR2*), and 11/30 (37%) had epigenetic mutations (*TET2, DNMT3A, ASXL1*) (Malcovati *et al*, 2015).

We and others have shown that the markedly under-expressed iron transporter gene *ABCB7* is an important mediator of the RARS phenotype (Nikpour *et al*, 2013; Dolatshad *et al*, 2015).

Albeit these findings indicate that abnormal RNA is a common consequence of *SF3B1* mutations, downstream targets associated with the development of ineffective erythropoiesis have not yet been identified, nor the maturation phase during which erythropoiesis is affected. Splicing patterns and gene expression undergo major changes during normal cell differentiation (An *et al*, 2014; Cheng *et al*, 2014; Pimentel *et al*, 2014) making longitudinal studies during erythroid maturation highly relevant (Nikpour *et al*, 2010, 2013). We speculated that *SF3B1* mutations could alter splicing and expression patterns of genes essential for normal erythroid maturation.

Hereby, we report altered splicing of genes involved in haemoglobin synthesis, and a failure to up-regulate essential cell growth signalling pathways during erythroid differentiation. Importantly, we present evidence that *SF3B1* mutation does not confer a growth or maturation disadvantage to erythroid cells until terminal maturation into reticulocytes, providing a novel understanding about the mechanism for anaemia in *SF3B1* mutated RARS.

Material and methods

Sample collection, erythroid and myeloid cell cultures

Patients with MDS, <5% marrow blasts, and >15% ring sideroblasts [RARS, RCMD-RS, RARS associated with marked thrombocytosis (RARS-T)] were diagnosed using World Health Organization (WHO) criteria and a multi-professional conference approach (Howe *et al*, 2004; Vardiman *et al*, 2009). The study was approved by the Ethics Research Committee at Karolinska Institutet (2010/427-31/1 and 2011/1257-31) and informed consent was obtained from patients and healthy donors. RARS and normal $CD34^+$ cells from bone marrow (BM) were cultured for 14 days, as previously described in detail (Tehranchi *et al*, 2003; Pellagatti *et al*, 2007; Nikpour *et al*, 2010, 2013). In order to obtain fresh mature erythroid progenitors, Glycophorin A (GPA) selection (Miltenyi Biotech, Bergisch Gladbach, Germany) was performed on freshly isolated mononuclear cells (MNCs) with a purity of >95% as previously reported (Tehranchi *et al*, 2003, 2005a; Nikpour *et al*, 2010). To induce myeloid differentiation, cells were treated with hydrocortisone (0·01M, StemCell Technology, Grenoble, France) and granulocyte colony-stimulating factor (20 ng/ml, PeproTech, Rocky Hill, NJ, USA) at day 7 in Myelocult medium. Myeloid purity was evaluated by flow cytometry (CD66b >90%) at day 14 of culture. Reticulocytes were isolated from peripheral blood (PB) of 5 patients maintaining stable anaemia without treatment or transfusions, as described by Petruzzelli *et al* (2009). Genomic DNA was extracted from isolated $CD34^+$, $CD34^-$, PB $CD3^+$ cells and granulocytes using GenElute™ Mammalian DNA Miniprep Kit (Sigma Aldrich, St. Louis, MO, USA) and total RNA using Trizol (Invitrogen, Paisley, UK).

Targeted sequencing of candidate genes

Patient's DNA from PB granulocytes or BM MNCs was screened for recurrently mutated genes in myeloid disorders either by SureSelect (Agilent Technologies, Berkshire, UK) as previously reported (Papaemmanuil *et al*, 2013) or using high-throughput sequencing by Haloplex™ target enrichment technology (Agilent Technologies). Haloplex™ target enrichment kit G9901A/B was designed for 42 genes (Table SI) using Halodesign wizard available via Agilent (http://www.halogenomics.com). For details on sequencing data analysis, see Supporting Information.

RNA Sequencing

Total RNA (between 177 ng and 3 μg) from BM $CD34^+$ cells of 3 patients (MDS 12, 13 and 14) and 3 normal bone marrows (NBM) during erythroid differentiation was depleted of ribosomal RNA using RiboMinus Eukaryote kit for RNA-Seq (LifeTechnologies, Carlsbad, CA, USA). Strand-specific libraries were constructed using SOLiD™ Total RNA-Seq Kit according to the manufacturer's instructions (LifeTechnologies). For details on libraries sequencing, see Supporting Information. Samples were sequenced as single-reads on the SOLiD 5500 Wildfire system. An average of approximately 246 million reads were generated per sample and out of these, an average of 141 million reads were mapped to the human reference genome (hg19). For differential expression and splicing analysis, Cufflinks version 2·1·1 (http://cole-trapnell-lab.github.io/cufflinks/) was used. The alignment of sequence reads in the same sample type were assembled into transcripts by Cufflinks, using RefSeq transcripts as reference.

Assembled transcripts were grouped by Cuffmerge and the merged transcripts assembly was compared with the reference RefSeq transcripts by Cuffcompare using default setting (Trapnell *et al*, 2012). Relative abundance (fragments per kilobase of exon per million of fragments mapped, FPKM) was calculated. Differential expression and exon-usage were

estimated by Cuffdiff using default setting (Trapnell *et al*, 2012). An arbitrary cutoff of FPKM>1 was applied, thus genes were considered as expressed if FPKM>1.

Gene set enrichment analysis

The input data for the Gene Set Enrichment Analysis (GSEA) (Subramanian *et al*, 2005) were complete lists of genes differentially expressed between RARS and NBM at day 0 and at day 4. The lists were pre-ranked according to the $\log_{10}$ (*P* value) of each gene. Probesets (12038 annotated transcripts at day 0 and 9186 at day 4) were collapsed into gene symbol. Inclusion gene set size was set between 15 and 500 and the phenotype was permutated 1000 times. Gene sets that met a false discovery rate (FDR) less than 0·05 were considered. For this analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG; Kanehisa & Goto, 2000) was used as gene matrix.

Statistical analysis

Statistical analysis was performed using GraphPad Prism 6 software (La Jolla, CA, USA). Gene expression and splicing isoform validation data by quantiative polymerase chain reaction (qPCR) were analysed using the Comparative Ct method according to User Bulletin #2 (Applied Biosystems, Foster City, CA, USA). Due to small sample size and skewed distribution, gene expression in patient samples and healthy controls was compared using the Mann–Whitney *U* test. Statistical analysis for allele burden assessment during erythroid maturation was performed using paired Student *t*-test.

Taqman low density array (TLDA)

Gene expression profiles were investigated using TDLA. BM $CD34^+$ cells at day 0 and 4 from 7 RARS and 4 NBM erythroid cultures were run in triplicate. Each TLDA card allowed simultaneous quantification of 61 target genes per sample. TLDA data was analysed using Comparative Ct method according to User Bulletin #2 (Applied Biosystems). For details on TLDA analysis see Supporting Information. Principle Component Analysis (PCA) was performed using R statistics package "prcomp" (http://www.r-project.org/).

Pyrosequencing

Pyrosequencing (PyroMark Q24 system, Qiagen, Düsseldorf, Germany) was applied to follow the fate of the *SF3B1* clone during erythroid differentiation. DNA from cultured BM $CD34^+$, freshly isolated BM MNCs and GPA^+ cells of RARS patients was used with DNA from NBM, as wild type control. cDNA synthetized using High Capacity cDNA Reverse Transcription kit (Applied Biosystems) was used to assess allele burden in reticulocytes. Mutation-specific assays (Table SII) were designed using PyroMark assay design software and performed according to the PyroMark Q24 user manual (Qiagen).

Results

Patients

Eighteen patients with *SF3B1* mutations and 2 *SF3B1* wild type patients had material available for this study. Median age was 74 years (49–91 years) and final diagnosis was established at a multi-professional conference. Clinical and morphological characteristics as well as the molecular profile of the cohort are presented in Table SIII. Several of the samples had additional recurrent mutations, such as *JAK2, TET2* and *DNMT3A* (Papaemmanuil *et al*, 2013).

Marked differences in gene expression during early erythroid differentiation

To examine if splicing factor mutations may contribute to dissimilarities between *SF3B1* mutated RARS and NBM, we assessed loci-based differential expressions by RNA sequencing during early erythroid differentiation at two time-points (day 0 and 4). We detected 5292 loci differentially expressed (FDR<5% and FPKM>1) at day 0 and 2329 loci at day 4 (Fig 1A–B). Next, we investigated the genes showing different development of expression between RARS and NBM over the 4 days in culture. We observed marked alterations during RARS differentiation (Figure 1C-D). In particular, 953 genes were significantly (FDR<5% and fold change of ≥1·5) up-regulated during normal erythropoiesis, compared to only 333 genes in the RARS cultures. Hence, the expression of 710 genes failed to increase during RARS erythropoiesis. The functional annotation of these genes was analysed using DAVID (Database for Annotation, Visualization and Integrated Discovery) (Huang da *et al*, 2009) and the most enriched functional clusters were mainly involved in the cell cycle process including mitosis, spindle organization and chromosome segregation (Table SIV). Furthermore, we identified 511 down-regulated genes during normal erythropoiesis compared to 893 in RARS, i.e. 222 genes failed to decrease during RARS erythropoiesis. Interestingly, the functional annotation of these 222 genes failed to identify any significant enriched clusters (Table SV). Overall, these data suggest that *SF3B1* mutated RARS is associated with decreased gene expression resulting from a failure to up-regulate genes regulating the cell cycle process during early erythroid differentiation.

Gene Set Enrichment Analysis identifies the key pathways of RARS pathogenesis

We used GSEA (Subramanian *et al*, 2005) to identify biological pathways that showed coordinated up-regulation or down-regulation in RARS compared to NBM. With a FDR

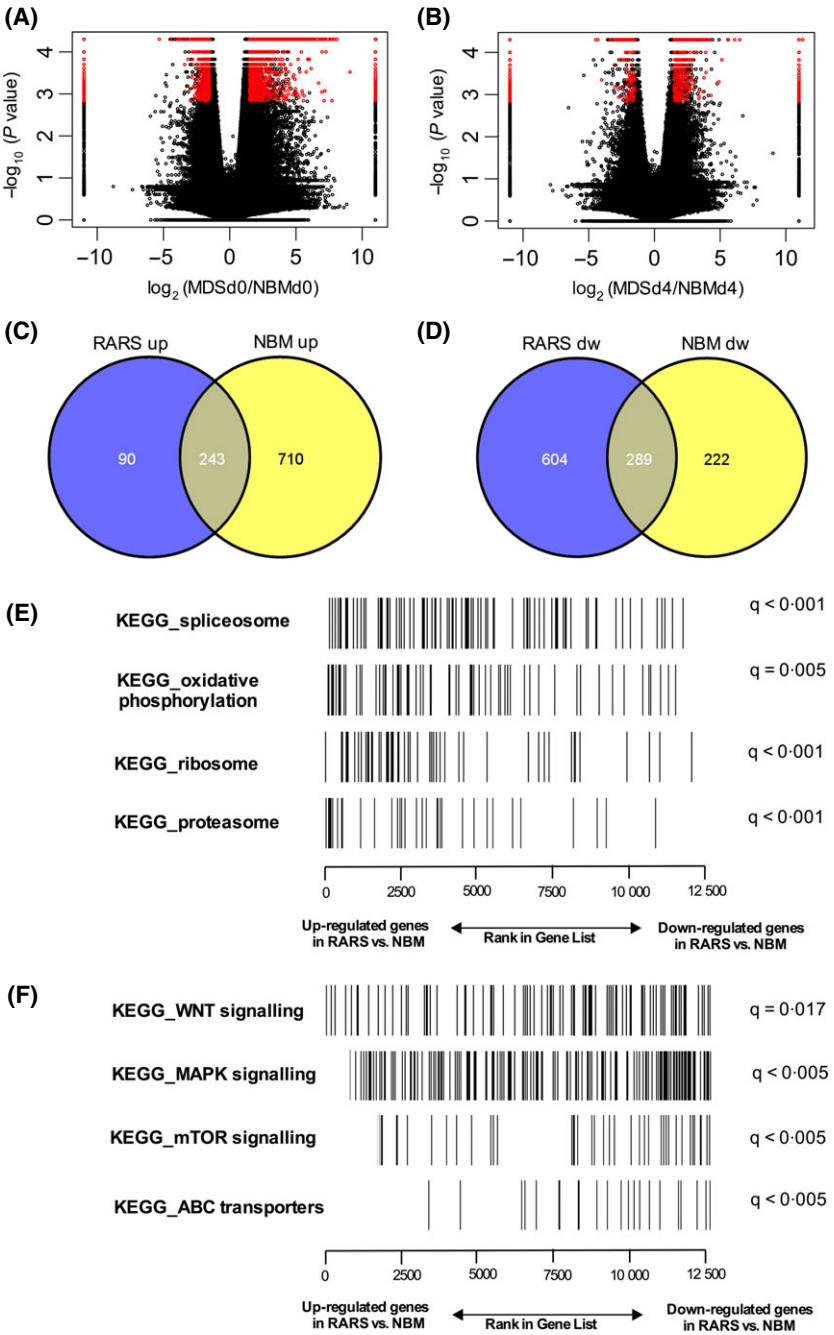

Fig 1. RNA sequencing reveals gene expression differences between RARS and NBM during early differentiation. Bone marrow $CD34^+$ cells were isolated from 3 RARS patients (MDS12, 13, 14) and 3 NBM controls and cultured in the erythroblast system for 4 days. Samples for RNA were taken at day 0 and day 4. (A) Volcano plots show 5292 loci (both annotated and un-annotated) up-regulated at day 0 and (B) 2329 down-regulated at day 4 (FDR<5% and FPKM>1). Each graph is constructed by plotting the negative log of the *P*-value (*Y*-axis) *versus* the log of the fold change (1·5-fold change) (*X*-axis) between RARS and NBM. Red circles denote statistical significance (FDR <5%) defined loci by Cufflinks. (C–D) Venn diagrams display distributions of differentially expressed transcripts between NBM and RARS during early differentiation (0–4 days) based on FDR <5% and a fold change of 1·5. Only annotated transcripts were included for these diagrams. During erythropoiesis, 333 (90 + 243) genes were up-regulated in RARS and 953 (710 + 243) genes in NBM. 243 genes were up-regulated both in RARS and NBM. During early RARS erythropoiesis, 893 genes were down-regulated, compared to 511 in NBM. A total of 289 genes were down-regulated in both RARS and NBM. (E) Gene Set Enrichment Analysis (GSEA) of RARS compared to NBM at day 0 reveals positive enrichment plots for canonical pathways of RARS such as spliceosome, oxidative phosphorylation, ribosome and proteasome. (F) Conversely, at day 0 the GSEA analysis shows a negative enrichment of signalling pathways such as: WNT, MAPK and mTOR and ABC transporters. Each vertical stripe represents the rank of a gene within the gene set among all (12038 at day 0 and 9186 at day 4). Name of the gene set is shown to the left, and q value to the right. FPKM, fragments per kilobase of exon per million of fragments mapped; FDR, false discovery rate; MDS, myelodysplastic syndrome; RARS, refractory anaemia with ring sideroblasts; NBM, normal bone marrow; dw, down-regulated; KEGG, Kyoto Encyclopedia of Genes and Genomes.

of less than 0·05, we identified 14 gene sets (out of 151 screened) positively enriched, and 12 gene sets negatively enriched at day 0, whereas 18 and 1 gene sets showed a positive and negative enrichment at day 4, respectively. GSEA based on KEGG database revealed an up-regulation at both time points for canonical pathways of RARS pathogenesis, such as spliceosome, but also of the oxidative phosphorylation, ribosome and proteasome pathways (Fig 1E). The oxidative phosphorylation enrichment, a possible consequence of the mitochondrial ferritin accumulation or compromised mitochondrial respiratory chain activity, mainly included genes encoding subunits of the Complex I (NADH dehydrogenase) and V (ATP synthase) of the respiratory chain. Ribosomal abnormalities have been previously described in MDS pathophysiology in conjunction with deletion of 5q (Ebert *et al*, 2008). Studies have shown that haploinsufficiency of ribosomal protein genes causes selective activation of TP53 (p53) in human erythroid progenitor cells (Dutt *et al*, 2011). The positive enrichment of the proteasome has been reported only in high risk MDS, and often associated with apoptosis after treatment with proteasome inhibitors (Alimena *et al*, 2011; Liesveld *et al*, 2011).

Conversely, GSEA showed down-regulation of WNT, MAPK and mTOR signalling pathways involved in apoptosis and transcription regulation, as well as ATP-binding cassette mitochondrial transporters (ABC transporters) (Figure 1F). The latter gene set included *ABCB7*, a key mediator of the RARS phenotype (Nikpour *et al*, 2013) involved in iron metabolism and transport of Fe/S protein precursors.

RARS progenitors up-regulate genes involved in haemoglobin synthesis

Compared to NBM, several of the top 20 up-regulated genes in RARS at day 0 were involved in the haem and haemoglobin synthesis pathway (*HBM, HBA2, ALAS2, HBA1*) (Table I). In particular, *HBA1* and *HBA2* encode the α-globin protein and both genes are located close together in the α-globin locus on chromosome 16. Mutations of these genes are the most common cause of α-thalassemia (Murru *et al*, 1991); however, *HBA2* was also reported down-regulated in differentiating erythroblasts from MDS patients with del(5q) after treatment with lenalidomide (Pellagatti *et al*, 2007). Interestingly, *HBM* was initially identified as a pseudogene and later reported as part of the transcriptome profile of human reticulocytes and transcribed in a highly regulated pattern in erythroid cells with its maximal expression during the erythroblast terminal differentiation (Goh *et al*, 2005). We also confirmed a strong up-regulation of *ALAS2* previously reported in RARS (Pellagatti *et al*, 2006). The expression levels of *ALAS2, HBM*, *HBA2* and *HBA1* were similar in *SF3B1* mutated ($n = 6$) and wild type ($n = 2$) patients (data not shown).

In contrast to the up-regulation of haemoglobin synthesis related genes, we found down-regulation of tumour suppressors and apoptosis-promoting genes, such as *FHIT* and *CAMK1D* (Table I) suggesting a mechanism for RARS progenitors to survive until a more mature stage of differentiation and a potential contribution to erythroid hyperplasia.

We validated the expression of 5 of the top 20 up-regulated and down-regulated in 3 NBM and 6 RARS using qPCR. Indeed, *ALAS2, HBM* and *SLC4A1* were significantly up-regulated compared to NBM ($P = 0{\cdot}023$ for all tested genes) (Fig 2A). In contrast, *CAMK1D* and *FHIT* were down-regulated versus NBM ($P = 0{\cdot}047$, $P = 0{\cdot}023$ respectively) (Fig 2B).

Finally, at day 4 of erythroid culture, we found up-regulation of genes with a scavenger receptor role, such as *CD5L* and *CD163*. *CD163* is an haemoglobin scavenger receptor that may protect from free haemoglobin-mediated oxidative damage (Moestrup & Moller, 2004) (Table SVI).

SF3B1 mutated RARS progenitors show alternative splicing of genes involved in haemoglobin synthesis

Alternative splicing can generate multiple mRNA isoforms and proteins from a single locus. Changes in the splicing process are important for cell differentiation and may contribute to the difference in gene expression between RARS and NBM (Pellagatti *et al*, 2006; Nikpour *et al*, 2010). RNA sequencing revealed 18 mis-spliced genes at day 0, none of which had been reported before to have a role in MDS (Table SVII). By contrast, 2 genes (*ANKHD1* and *TMEM14C*) showed altered splicing at day 4 and only *TMEM14C* was differentially spliced at both time points. Interestingly, the function of *TMEM14C* and *TFCP2* is closely related to mitochondria and to haem biosynthesis in particular (Yien *et al*, 2014). *TMEM14C* is a mitochondrial transporter and the knockdown of this gene in zebrafish led to profound anaemia without affecting erythroid lineage specification (Nilsson *et al*, 2009). Two isoforms of *TMEM14C* were detected in our material: TCONS_00409015 was 3·5 times up-regulated in RARS compared to NBM, while the expression of TCONS_00409016 was significantly down-regulated in RARS versus NBM (Fig 3A). Notably, the shorter isoform TCONS_00409016 differs from the other variant in that it has 14 bases (328-341) missing in the 5′UTR (Figure 3B). An internal ribosome entry site (IRES), a translation initiation alternative to the conventional 5′-cap translation mechanism, was predicted by UTRdb (Grillo *et al*, 2010) at position 295-385 of *TMEM14C*.

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 (Fig 3A), TCONS_0012739 showed a skipping of 153 bases (1286–1438) in the sixth exon of the mRNA corresponding to the CP2 DNA binding domain that binds the–α-globin promoter. 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 & Kim, 2003).

We selected *TMEM14C* and *TFCP2* for further splice-form validation due to their potential relevance to the disease phenotype. We measured the abundance of each splicing isoform in 3 NBM and 5 RARS patients using qPCR. Primers specificity and amplicons identity were validated using Sanger Sequencing (Table SVIII).

In regard to *TMEM14C*, TCONS_00409015 (referred as TCONS15 in the figure) was significantly up-regulated in RARS compared to NBM ($P = 0{\cdot}035$). By contrast TCONS_00409016 (referred as TCONS16 in the figure) was decreased 1·3-fold in RARS versus NBM ($P = 0{\cdot}035$) (Figure 3B).

For *TFCP2* we measured the abundance of transcripts containing exon 6 and interestingly, we observed a 1·3 fold-decrease in RARS *versus* NBM ($P = 0{\cdot}035$); whereas the expression of TCONS00112740 (referred as TCONS40 in the figure, Taqman probe: Hs_01117310_m1) showed the same trend as in the RNA Sequencing, however inter-patient variation was high and the comparison was not statistically significant (Fig 3C).

Table I. Top 20 up and down-regulated genes between RARS and normal bone marrow at day 0.

Gene	Locus	FPKM NBM	FPKM MDS	log2(MDS/NBM)	FDR	Function
(A) Up-regulated genes						
XIST	chrX:73012039-73072588	0·026	54·783	11·032	0·004	X-Inactivation that silences one of the pair of X chromosomes
**HBM*	chr16:215972-216767	0·316	170·276	9·075	0·016	Haemoglobin μ, iron ion binding and oxygen binding
FCGBP	chr19:40353962-40440533	0·064	17·509	8·095	0·004	Maintenance of the mucosal structure
TIMD4	chr5:156345573-156345938	0·996	237·879	7·900	0·004	Regulation of T-cell proliferation
HBA2	chr16:222845-223709	1·300	301·970	7·859	0·004	Haemoglobin α-2, ion and oxygen biniding
ALAS2	chrX:55035487-55057497	0·415	87·171	7·716	0·004	Haem biosynthetic pathway
HBA1	chr16:226643-228694	1·097	222·552	7·665	0·004	Haemoglobin α-1, release of oxygen by erythrocytes and scavenging of haem
HRH1	chr3:11178778-11304939	0·040	6·833	7·416	0·004	Histamine receptor activity
**VCAM1*	chr1:101204878-101205214	0·695	107·010	7·267	0·007	Integrin binding and cell adhesion molecule binding
EDNRB	chr13:78469615-78549664	0·102	15·320	7·237	0·004	Peptide hormone binding and endothelin receptor activity
ATP6V0D2	chr8:87111077-87166454	0·112	16·466	7·199	0·004	Providing energy required for transport processes in the vacuolar system
CERS3	chr15:100940599-101084925	0·022	3·160	7·181	0·034	Sphingosine N-acyltransferase activity and transcription factor activity
**CD5L*	chr1:157800703-157811634	3·677	782·091	7·172	0·004	Scavenger receptor activity, possible role as an inhibitor of apoptosis
RND3	chr2:151324709-151344221	0·180	25·674	7·154	0·004	GTP binding and GTPase activity
LILRB5	chr19:54753206-54753669	0·294	41·095	7·127	0·049	Transmembrane signaling receptor activity
LYVE1	chr11:10562786-10715535	0·480	60·302	6·972	0·007	Autocrine regulation of cell growth
CCL18	chr17:34399750-34402443	0·033	4·096	6·954	0·035	Immunoregulation and inflammation, chemotactic activity for T cells
MSR1	chr8:15965386-16050319	0·444	54·932	6·951	0·004	Macrophge scavenger receptor activity
SLC4A1	chr17:42325757-42345502	0·250	30·605	6·937	0·004	Uptake of carbon dioxide and release of oxygen by erythrocytes
CCL8	chr17:32646065-32648421	0·295	34·326	6·862	0·004	Protein kinase activity and chemokine activity
(B) Down-regulated genes						
**ULK4P3*	chr15:32698800-32727250	4·611	0·449	−3·360	0·004	*ULK4* pseudogene 3, unknown function
WWOX	chr16:78760629-78761056	5·973	0·834	−2·840	0·007	Oxidoreductase that acts as a tumor suppressor and plays a role in apoptosis
CAMK1D	chr10:12466821-12467259	5·474	0·852	−2·684	0·007	Calcium/calmodulin-dependent kinase, role in apoptosis of erythroleukemia cells
PRKCE	chr2:45997530-45997967	3·401	0·536	−2·665	0·029	Protein kinase activity
**RNF220*	chr1:45030902-45031276	6·598	1·154	−2·516	0·007	Ubiquitination and proteasomal degradation of SIN3B
PRKCH	chr14:61933032-61933557	2·599	0·457	−2·509	0·037	Regulation of cell differentiation in keratinocytes and pre-B cell receptor
PLCB1	chr20:8291278-8291669	3·810	0·736	−2·371	0·041	Hydrolysis of phospholipids into fatty acids
GNLY	chr2:85921413-85925977	2·085	0·419	−2·315	0·004	Antimicrobial activity
BCR	chr22:23563521-23563723	33·971	6·880	−2·304	0·018	GTPase and serine/threonine kinase activity
LRIG1	chr3:66488799-66489239	3·708	0·767	−2·273	0·048	Negative regulator of signaling by receptor tyrosine kinases
INPP5A	chr10:134526579-134527222	1·983	0·424	−2·225	0·047	Inositol-polyphosphate 5-phosphatase activity
PRKG1	chr10:53142352-53143618	1·053	0·242	−2·123	0·021	Mediator of the nitric oxide (NO)/cGMP signaling pathway
AOX2P	chr2:201560443-201660451	32·892	7·630	−2·108	0·004	Aldehyde oxidase 2, pseudogene affiliated with the lncRNA class
SLC16A14	chr2:230899676-230933816	1·483	0·347	−2·094	0·028	Proton-linked monocarboxylate transporter
BRF1	chr14:105726207-105726823	4·543	1·124	−2·015	0·004	RNA Polymerase III Transcription Initiation and Transcription of tRNA

Table I. (*Continued*)

Gene	Locus	FPKM NBM	FPKM MDS	log2(MDS/NBM)	FDR	Function
FOXP1	chr3:71532609-71532947	8·119	2·079	−1·965	0·018	Transcriptional repressor, essential transcriptional regulator of B-cell
TMEM163	chr2:135360910-135361983	1·279	0·329	−1·957	0·016	Zinc ion binding
NCK2	chr2:106380446-106380910	6·308	1·720	−1·874	0·012	Regulation of receptor protein tyrosine kinases and cytoskeletal adaptor activity
FHIT	chr3:59928983-59929714	2·466	0·675	−1·870	0·031	Pyrimidine metabolism and inducer of apoptosis via SRC and AKT1 signaling
STK32B	chr4:5202129-5202796	2·610	0·724	−1·851	0·028	Serine/threonine kinase activity and metal ion binding

FPKM, fragments per kilobase of exon per million of fragments mapped; FDR, false discovery rate; MDS, myelodysplastic syndrome; RARS, refractory anaemia with ring sideroblasts; NBM, normal bone marrow.
*Genes commonly deregulated both at day 0 and day 4. Notably, the up-regulation of *TIMD4, ALAS2, CD5L, LYVE1, CCL18* and *MSR1* detected by RNA sequencing was confirmed by Taqman Low Density Array.

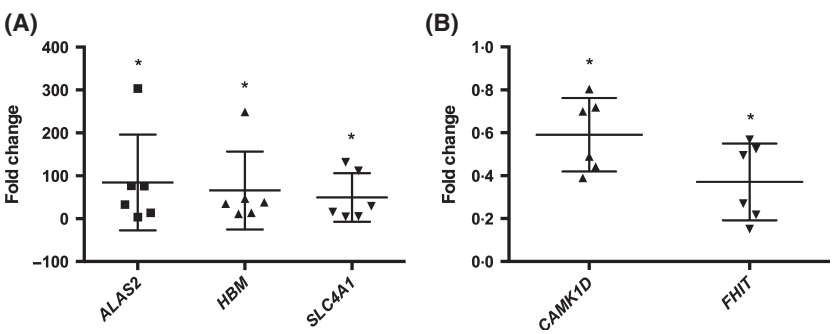

Fig 2. Validation of gene expression profiles. Quantiative polymerase chain reaction analysis of RARS ($n = 6$) *versus* NBM ($n = 3$) confirms (A) the up-regulation of *ALAS2, HBM* and *SLC4A1* and (B) the down-regulation of *CAMK1D* and *FHIT* observed in the RNA sequencing. Level of statistical significance was assessed using Mann-Whitney test (*$P < 0.05$).

In order to assess the impact of *SF3B1* mutations on these erythroid-related genes, we measured the expression of *TMEM14C* and *TFCP2* splice variants in two additional *SF3B1* wild type patients analysed by exome sequencing (Table SIII, data from exome sequencing not shown). With the caveat of the statistical limitation, we observed that *TMEM14C* TCONS15 was significantly down-regulated in *SF3B1* wild type compared to mutated patients ($P = 0·001$) (Fig. S1). This indicates that *SF3B1* mutations might contribute to ineffective erythropoiesis by altering the splicing of genes involved in haemoglobin syntesis. Functional studies will be needed to establish how these genes cooperate in generating the RARS phenotype.

The additional 16 mis-spliced genes (Table SVII) were related to various functions, such as DNA replication (*DUT*), MAPK signalling pathway (*MBIP*), transcription regulation (*ZNF518B, EIF2B4, TGIF2*), blood coagulation (*PROCR*), intracellular trafficking and energy metabolism (*ACBD3, SNX3, CYP2S1, NUDT7*). Of note, the coverage of *ABCB7* for which we previously showed altered exon usage at day 7 and 14 of erythroid culture was suboptimal for analysis of splicing patterns. Taken together, these results show that the inappropriate splicing seems to be closely associated not only with aberrant expression, but also with a dys-regulation of downstream target genes important for erythroid differentiation.

TLDA analysis confirms data from RNA sequencing and highlights a clear heterogeneity of the RARS cohort

We then used a gene expression approach to study selected genes during early erythroid differentiation (day 0–4) from 7 RARS patients and 4 NBM. The assay included 61 genes that were selected based on pilot RNA sequencing results (data not shown) and previous publications (Pellagatti *et al*, 2006; Nikpour *et al*, 2010). TLDA data revealed 20 genes differentially expressed at day 0 and 17 genes at day 4 ($P < 0·05$) between RARS and NBM (Table SIX). The up-regulation of *TIMD4, ALAS2, CD5L LYVE1, CCL18* and *MSR1,* observed in the RNA sequencing (Table I), and *CXCL12* (Table SVI) was validated and confirmed by TLDA (Table SIX).

Importantly, several of the genes selected for this assay (*ABCB7, ALAS2, APOE, HBB, MFN2, SEPP1*) were differentially expressed between RARS and NBM and involved in the defence against oxidative stress (Miyata & Smith, 1996; Saito *et al*, 1999; Shen *et al*, 2007; Liesa *et al*, 2012; Khechaduri *et al*, 2013). These data corroborate with the up-regulation of the oxidative phosphorylation pathway observed in the GSEA and suggest that RARS progenitors may be challenged by oxidative stress at an early phase of differentiation.

Principle Component Analysis was performed to identify patterns and variance in the gene expression data obtained by TLDA, which included more RARS patients than in the

(A)

Gene symbol	Gene name	Similar RefSeq accession	Transcript ID	FPKM NBM	FPKM MDS	log2 (MDS/NBM)	FDR
TMEM14C	transmembrane protein 14C	NM_001165258	TCONS_00409015	5·218	61·997	3·571	0·016
TMEM14C	transmembrane protein 14C	NM_016462	TCONS_00409016	18·005	12·151	−0·567	
TFCP2	transcription factor CP2	NM_001173452	TCONS_00112738	5·769	2·334	−1·305	
TFCP2	transcription factor CP2	NM_001173453	TCONS_00112739	0·498	0·142	−1·809	0·024
TFCP2	transcription factor CP2	NM_005653	TCONS_00112740	0·075	8·235	6·770	

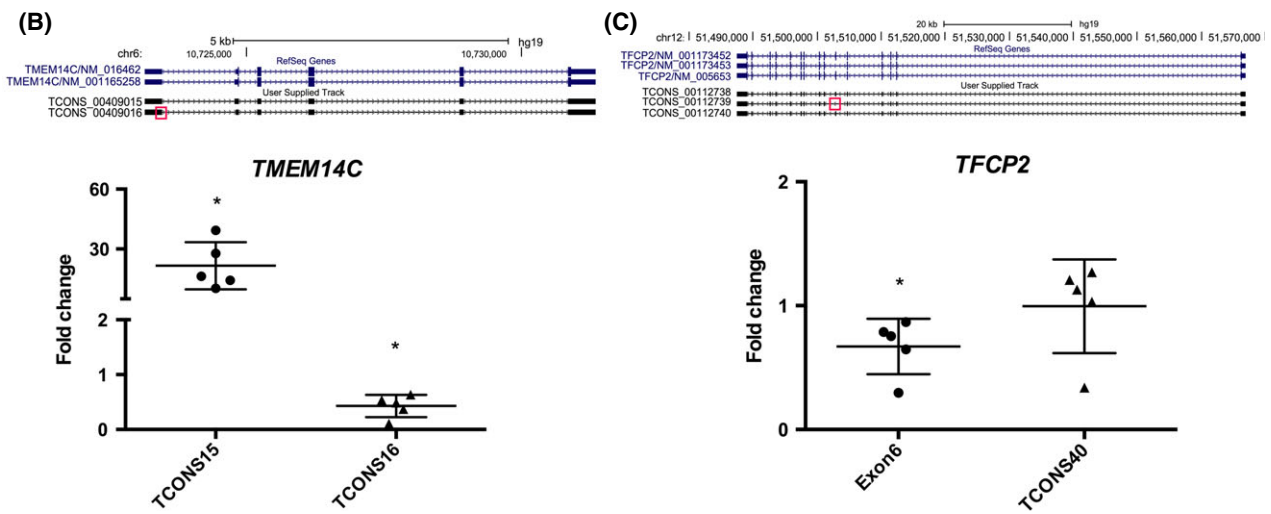

Fig 3. Quantification of *TMEM14C* and *TFCP2* transcript splice variants in RARS (n = 5) *versus* NBM (n = 3). (A) RNA sequencing based analysis shows differential expression of different TCONS corresponding to *TMEM14C* and *TFCP2*. (B–C) Top panels show browser view of the exon-intron structure for *TMEM14C* and *TFCP2*, respectively (the red square indicating the differential splicing site). Bottom panel, quantiative polymerase chain reaction detected fold change of the TCONS is shown. (B) In NBM, the *TMEM14C* isoform TCONS15 was significantly up-regulated compared to RARS (P = 0·035) while the shorter isoform TCONS16 (missing 14 bp in the 5′UTR) was decreased 1·3 fold in RARS versus NBM (P = 0·035). (C) The levels of *TFCP2* transcripts containing exon 6 were significantly lower (P = 0·035) in RARS than NBM, while the isoform TCONS40 was higher expressed in RARS compared to NBM. Delta CT were normalized using *GAPDH* as reference gene, NBM samples (n = 3) were used as calibrator. Level of statistical significance was assessed using Mann–Whitney test (*P < 0·05). FPKM, fragments per kilobase of exon per million of fragments mapped; FDR, false discovery rate; MDS, myelodysplastic syndrome; RARS, refractory anaemia with ring sideroblasts; NBM, normal bone marrow.

RNA sequencing analysis. The RARS population showed a heterogeneity at both time points, while NBMs were much more uniform in gene expression, especially at day 0. All samples were *SF3B1* mutated (Table SI) and clustered independently from the amino acid substitutions (Fig 4A–B). Furthermore, another PCA highlighted the target genes that contribute to distinguish NBM from RARS both at day 0 (Fig 4C) and day 4 (Fig 4D) which include *SEPP1, APOE, HBB and ALAS2*.

Selection against SF3B1 occurs first during terminal maturation into erythrocytes

Mechanisms behind the low reticulocyte count and surviving erythrocytes in RARS have, to our knowledge, not been studied. To test whether *SF3B1* mutated erythroid progenitors have a growth disadvantage compared to *SF3B1* wildtype cells, we analysed the allele burden of *SF3B1* (Table SII) in RARS samples during erythroid differentiation (n = 5) (days 0, 4, 7, 11 and 14) (Nikpour *et al*, 2010, 2013), two of which were also exposed to myeloid maturation. For these experiments we utilized samples from *SF3B1* mutated RARS without need for transfusions or growth factor treatment.

Interestingly, the *SF3B1* allele burden was stable (± 10%) during both erythroid and myeloid differentiation (Fig 4E). The pattern was similar in patients with and without transfusion need, i.e. the production or not of erythrocytes did not influence the results. Furthermore, to exclude the possibility that the culture system promoted survival of mutated cells, glycophorin A was used to select GPA^+ cells from freshly isolated BM $CD34^-$ cells of 7 RARS patients carrying different *SF3B1* mutations (H662Q, K700E, N626D, E622D, K666R, T663I). allele burden in the GPA^+ fraction was similar (41% ± 5·40SE) to the $BMCD34^+$ and erythroblast D14 fractions previously evaluated (46% ± 4·09SE and 43% ± 6·66SE, respectively) (Fig 4F).

We finally assessed *SF3B1* allele burden in reticulocytes obtained from non-transfused RARS subjects with stable anaemia and a median of 47% (33–70) RS of total erythroblasts. We obtained fresh blood from 4 patients (MDS 8, 12, 13, 14) and one healthy control, isolated RNA from reticulocytes and performed *SF3B1* mutational analysis using cDNA.

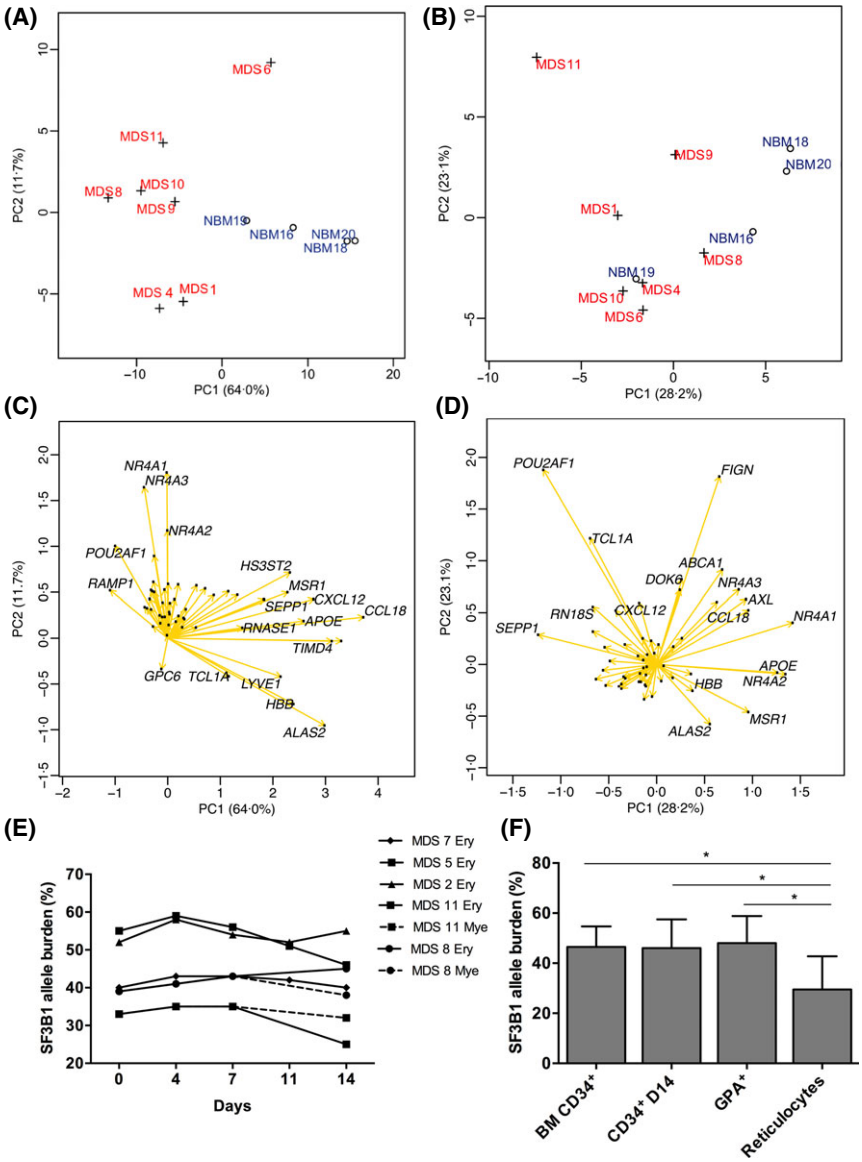

Figure 4. The RARS cohort is heterogeneous and shows defective terminal maturation into erythrocytes. RARS patients ($n = 7$) and NBM ($n = 4$) were analysed at day 0 and 4 of erythroid differentiation by Taqman Low Density Array. (A) Principle Component Analysis (PCA) shows a marked heterogeneity in the gene expression of RARS samples (MDS 1, 4, 6, 8, 9, 10, 11) compared to only a minimal inter-individual variation in the NBM population (NBM 16, 18, 19, 20) both at day 0 (B) and at day 4. (C) Genes that contribute to distinguish NBM from RARS are displayed at day 0 (D) and day 4. (E) *SF3B1* allele burden was quantified by pyrosequencing at different time points in 5 RARS samples cultured in erythroblast system (Ery). Two patients (MDS 8 and MDS 11) were also cultured in a myeloid culture system (Mye) between days 7–14. Allele burden was stable (± 10%) both during erythroid and myeloid differentiation. (F) *SF3B1* allele burden of different cell fractions: BM CD34$^+$, cultured erythroblasts at day 14, GPA$^+$ and reticulocytes from RARS patients ($n = 4$). There is a 1·5-fold decrease of the allele burden in the RARS reticulocytes compared to CD34$^+$ cells ($P = 0·017$) and erythroblast at day 14 ($P = 0·033$) in addition to a 1·4-fold decrease when comparing to GPA$^+$ cells ($P = 0·011$). Level of statistical significance was assessed using paired Student *t*-test (*$P < 0·05$). MDS, myelodysplastic syndrome; RARS, refractory anaemia with ring sideroblasts; NBM, normal bone marrow

Indeed, the mutant cDNA was expressed only in the reticulocytes from the patients and not in the control (Supplementary Figure 2). The *SF3B1* allele burden was 1·5-fold lower in reticulocytes compared to corresponding fresh CD34$^+$ cells ($P = 0·017$) and cells cultured for 14 days ($P = 0·033$), and 1·4-fold lower than in freshly isolated GPA$^+$ cells ($P = 0·011$). These results imply reduced differentiation of *SF3B1* mutated erythroblasts into mature erythrocytes (Fig 4F).

Discussion

To test the hypothesis that *SF3B1* mutations in RARS (Papaemmanuil *et al*, 2011; Yoshida *et al*, 2011) implicate

mRNA abnormalities, which may alter gene function essential for erythroid maturation, we focused this study to *SF3B1* mutated RARS, RCMD-RS and RARS-T, with or without other candidate mutations (Papaemmanuil *et al*, 2012) but identified two *SF3B1* wild type patients for comparison. We adopted a genome wide approach to evaluate the impact of *SF3B1* mutations on the transcriptome and to depict dissimilarities between RARS and NBM that could add to the understanding of the disease phenotype.

RNA sequencing of RARS BM mononuclear (Visconte *et al*, 2012, 2015) and $CD34^+$ (Dolatshad *et al*, 2015) cells was previously used to study normal bone marrow, and *SF3B1* mutated and wild type patients from different WHO subgroups showing deregulated gene expression and splicing of genes involved in iron homeostasis and mitochondrial function.

By contrast, our study is the first to specifically address aberrancies during erythroid differentiation. Aberrant splicing is thought to play a pivotal role in the pathogenetic mechanism of myeloid malignancies carrying splice factor mutations, but little is known about *SF3B1* specific gene targets in RARS pathogenesis. A preliminary report revealed 230 splicing events significantly enriched in $CD34^+$ cells from *SF3B1* mutated RARS, of which 90% were caused by misrecognition of the 3′ splice site. Fifty per cent of these sites resulted in frameshift, indicating that *SF3B1* mutations simultaneously may cause detrimental consequences in many genes (Shiozawa *et al*, 2014).

In the present study, we report aberrant splicing of the mitochondrial transporter *TMEM14C* and the α-globin transcription factor *TFCP2*, both regulating erythroid gene expression and involved in haemoglobin synthesis. In particular, the skipping of 150 bases in the CP2 domain that binds the α-globin promoter may indicate that aberrant splicing affects key genes for erythroid differentiation, suggesting a possible link between splicing and transcription regulation. Furthermore, we detected two isoforms of *TMEM14C* differentially spliced between RARS and NBM. The shorter isoform, TCONS00409016, has 14 bases missing in the 5′UTR within the predicted internal ribosome entry site (IRES). IRES elements have been shown to be important for translation of specific mRNAs during mitosis, apoptosis and hypoxia (Hellen & Sarnow, 2001). Therefore, a sequence change in the 5′UTR and in the IRES element may affect the translation of *TMEM14C*, with consequences for mitochondrial function and haemoglobin synthesis (Yien *et al*, 2014). Interestingly, two *SF3B1* wild type patients showed a significantly different pattern with regard to *TMEM14C* TCONS15, a finding that warrants further investigation. We suggest that the aberrant expression of genes involved in haemoglobin synthesis and the splicing of erythroid-related genes may play a key role in the ineffective erythropoiesis of RARS. Therfore, it will be very valuable to further unravel the role of *TMEM14C* and *TFCP2* in the ineffective erythropoiesis of RARS.

Furthermore, gene expression data from RNA sequencing and TLDA revealed a marked up-regulation of haem and haemoglobin genes, a positive enrichment of the oxidative phosphorylation pathway and a down-regulation of ABC transporters, all essential for mitochondria to execute their functions through the different cellular compartments. These findings, together with *ABCB7* down-regulation during erythroid differentiation of RARS $CD34^+$ cells *versus* NBM (Boultwood *et al*, 2008; Nikpour *et al*, 2013), implicate a mitochondrial dysfunction associated with the accumulation of mitochondrial ferritin (Cazzola *et al*, 2003). Haem synthesis and mitochondrial respiration are intrinsically coupled by their shared metabolic pathways and mitochondrial location in the cell. The marked up-regulation of several haemoglobin genes in the RARS transcriptome may be viewed as a compensatory reaction to the defective terminal erythroid maturation with accumulation of iron in mitochondria instead of incorporation into haem. While haemoglobin genes were up-regulated in RARS $CD34^+$ freshly isolated progenitors, corresponding studies in NBM $CD34^+$ showed that the process of differentiation-associated haemoglobin accumulation occurred first at day 6 followed by a rapid raise in the ratio between foetal and adult haemoglobin production between day 7 and day 9 (Wojda *et al*, 2002). Therefore, the early up-regulation of haemoglobin genes observed in RARS may indicate a compromised haemoglobinization during the subsequent phases of RARS erythropoiesis. The ineffective erythropoiesis, a hallmark of RARS, depicts an erythroid defect associated with decreased erythrocytes production despite increased early erythropoiesis (Tanno & Miller, 2010) and up-regulation of haemoglobin genes.

RARS is characterized by a failure of erythroid maturation and anaemia (Corey *et al*, 2007) but the exact stage of maturation at which this defect occurs is poorly understood (Hattangadi *et al*, 2011). Previous investigations have described apoptotic features of RARS erythroblasts in patient biopsies (Hellstrom-Lindberg *et al*, 1997) as well as in erythroid culture systems (Tehranchi *et al*, 2003, 2005a,b). 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). To explore at which level of differentiation anaemia develops in RARS, we followed the fate of *SF3B1* mutated BM progenitors throughout differentiation. We showed that the allele burden of *SF3B1* mutated progenitors remained stable during marrow erythroid differentiation (14 days), suggesting that *SF3B1* mutations do not inhibit growth and maturation of marrow erythroblasts, in spite of mitochondrial iron accumulation. Conversely, the decreased allele burden in reticulocytes indicates that the real threat to erythropoiesis develops during terminal differentiation to erythrocytes.

We believe that these investigations offer novel insights into the erythroid differentiation process and the mechanisms of anaemia in *SF3B1* mutated RARS, providing

evidence that aberrant splicing of key genes involved in iron transport and haemoglobin synthesis may alter terminal erythroid maturation. Hence, drugs targeting this process may help to relieve the anaemia of these patients. A potential therapeutic candidate is ACE-536, which was recently shown to target late stage erythropoiesis in a mouse model of MDS (Suragani *et al*, 2014) and which preliminary study also has shown efficacy in an early phase 1 trial (Attie *et al*, 2014) enrolling anaemic RARS patients.

Acknowledgements

The authors would like to thank the National Genomics Infrastructure and SciLife Laboratory through Uppsala Genome Centre and UPPMAX (b2013263) for the RNA sequencing computation analysis. This study was supported by grants from the Swedish Cancer Society, the Cancer Society in Stockholm (EH-L) and by the Scientific Research Council (EH-L and JK). SC is the recipient of a PhD fellowship from Karolinska Institutet.

Authors contribution

SC performed experiments and wrote the manuscript. EH-L designed the study and supervised the writing of the manuscript. LV, SK, PU, JK, MK, TMB, JW provided input on experiments. LV, TS, MD, performed experiments. SK analysed RNA sequencing data, EP, PC, MK, targeted sequencing. MJ, BS, provided clinical and morphological data. All authors approved the final manuscript.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig S1. Differences between RARS patients *SF3B1* mutated and wild type.

Fig S2. *SF3B1* mutational analysis in reticulocytes.

Table SI. List of 42 genes selected for targeted sequencing by Haloplex.

Table SII. Pyrosequencing primers used to quantify *SF3B1* allelic burden.

Table SIII. Clinical characteristics of patients at diagnosis and sampling.

Table SIV. Gene annotation cluster of the 710 genes that failed to up-regulate in RARS.

Table SV. Gene annotation cluster of the 222 genes that failed to down-regualte in RARS.

Table SVI. Top 20 up and down-regualted genes between RARS and NBM at day 4.

Table SVII. Mis-spliced genes in RARS versus NBM at day 0 of the erythoid culture.

Table SVIII. Primers designed to validate splice isoforms by qPCR.

Table SIX. Differentially expressed genes between RARS and NBM by TLDA.

References

Alimena, G., Breccia, M., Musto, P., Cilloni, D., D'Auria, F., Latagliata, R., Sanpaolo, G., Gottardi, E., Saglio, G. & Mandelli, F. (2011) Erythroid response and decrease of WT1 expression after proteasome inhibition by bortezomib in myelodysplastic syndromes. *Leukemia Research*, **35**, 504–507.

An, X., Schulz, V.P., Li, J., Wu, K., Liu, J., Xue, F., Hu, J., Mohandas, N. & Gallagher, P.G. (2014) Global transcriptome analyses of human and murine terminal erythroid differentiation. *Blood*, **123**, 3466–3477.

Attie, K.M., Allison, M.J., McClure, T., Boyd, I.E., Wilson, D.M., Pearsall, A.E. & Sherman, M.L. (2014) A phase 1 study of ACE-536, a regulator of erythroid differentiation, in healthy volunteers. *American Journal of Hematology*, **89**, 766–770.

Boultwood, J., Pellagatti, A., Nikpour, M., Pushkaran, B., Fidler, C., Cattan, H., Littlewood, T.J., Malcovati, L., Della Porta, M.G., Jadersten, M., Killick, S., Giagounidis, A., Bowen, D., Hellstrom-Lindberg, E., Cazzola, M. & Wainscoat, J.S. (2008) The role of the iron transporter ABCB7 in refractory anemia with ring sideroblasts. *PLoS ONE*, **3**, e1970.

Cazzola, M., Invernizzi, R., Bergamaschi, G., Levi, S., Corsi, B., Travaglino, E., Rolandi, V., Biasiotto, G., Drysdale, J. & Arosio, P. (2003) Mitochondrial ferritin expression in erythroid cells from patients with sideroblastic anemia. *Blood*, **101**, 1996–2000.

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

Cheng, A.W., Shi, J., Wong, P., Luo, K.L., Trepman, P., Wang, E.T., Choi, H., Burge, C.B. & Lodish, H.F. (2014) Muscleblind-like 1 (Mbnl1) regulates pre-mRNA alternative splicing during terminal erythropoiesis. *Blood*, **124**, 598–610.

Corey, S.J., Minden, M.D., Barber, D.L., Kantarjian, H., Wang, J.C. & Schimmer, A.D. (2007) Myelodysplastic syndromes: the complexity of stem-cell diseases. *Nature Reviews Cancer*, **7**, 118–129.

Dolatshad, H., Pellagatti, A., Fernandez-Mercado, M., Yip, B.H., Malcovati, L., Attwood, M., Przychodzen, B., Sahgal, N., Kanapin, A.A., Lockstone, H., Scifo, L., Vandenberghe, P., Papaemmanuil, E., Smith, C.W., Campbell, P.J., Ogawa, S., Maciejewski, J.P., Cazzola, M., Savage, K.I. & Boultwood, J. (2015) Disruption of SF3B1 results in deregulated expression and splicing of key genes and pathways in myelodysplastic syndrome hematopoietic stem and progenitor cells. *Leukemia*, **29**, 1092–1103.

Dutt, S., Narla, A., Lin, K., Mullally, A., Abayasekara, N., Megerdichian, C., Wilson, F.H., Currie, T., Khanna-Gupta, A., Berliner, N., Kutok, J.L. & Ebert, B.L. (2011) Haploinsufficiency for ribosomal protein genes causes selective activation of p53 in human erythroid progenitor cells. *Blood*, **117**, 2567–2576.

Ebert, B.L., Pretz, J., Bosco, J., Chang, C.Y., Tamayo, P., Galili, N., Raza, A., Root, D.E., Attar, E., Ellis, S.R. & Golub, T.R. (2008) Identification of RPS14 as a 5q- syndrome gene by RNA interference screen. *Nature*, **451**, 335–339.

Goh, S.H., Lee, Y.T., Bhanu, N.V., Cam, M.C., Desper, R., Martin, B.M., Moharram, R., Gherman, R.B. & Miller, J.L. (2005) A newly discovered human alpha-globin gene. *Blood*, **106**, 1466–1472.

Grillo, G., Turi, A., Licciulli, F., Mignone, F., Liuni, S., Banfi, S., Gennarino, V.A., Horner, D.S., Pavesi, G., Picardi, E. & Pesole, G. (2010) UTRdb and UTRsite (RELEASE 2010): a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research*, **38**, D75–D80.

Hattangadi, S.M., Wong, P., Zhang, L., Flygare, J. & Lodish, H.F. (2011) From stem cell to red cell: regulation of erythropoiesis at multiple levels by multiple proteins, RNAs, and chromatin modifications. *Blood*, **118**, 6258–6268.

Hellen, C.U. & Sarnow, P. (2001) Internal ribosome entry sites in eukaryotic mRNA molecules. *Genes & Development*, **15**, 1593–1612.

Hellstrom-Lindberg, E. (2005) Update on supportive care and new therapies: immunomodulatory drugs, growth factors and epigenetic-acting agents. *American Society of Hematology Education Program*, **2005**, 161–166.

Hellstrom-Lindberg, E., Kanter-Lewensohn, L. & Ost, A. (1997) Morphological changes and apoptosis in bone marrow from patients with myelodysplastic syndromes treated with granulocyte-CSF and erythropoietin. *Leukemia Research*, **21**, 415–425.

Howe, R.B., Porwit-MacDonald, A., Wanat, R., Tehranchi, R. & Hellstrom-Lindberg, E. (2004) The WHO classification of MDS does make a difference. *Blood*, **103**, 3265–3270.

Huang da, W., Sherman, B.T. & Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols*, **4**, 44–57.

Kanehisa, M. & Goto, S. (2000) KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, **28**, 27–30.

Kang, H.C., Chae, J.H., Lee, Y.H., Park, M.A., Shin, J.H., Kim, S.H., Ye, S.K., Cho, Y.S., Fiering, S. & Kim, C.G. (2005) Erythroid cell-specific alpha-globin gene regulation by the CP2 transcription factor family. *Molecular and Cellular Biology*, **25**, 6005–6020.

Khechaduri, A., Bayeva, M., Chang, H.C. & Ardehali, H. (2013) Heme levels are increased in human failing hearts. *Journal of the American College of Cardiology*, **61**, 1884–1893.

Liesa, M., Qiu, W. & Shirihai, O.S. (2012) Mitochondrial ABC transporters function: the role of ABCB10 (ABC-me) as a novel player in cellular handling of reactive oxygen species. *Biochimica et Biophysica Acta*, **1823**, 1945–1957.

Liesveld, J.L., Rosell, K.E., Bechelli, J., Lu, C., Messina, P., Mulford, D., Ifthikharuddin, J.J., Jordan, C.T. & Phillips Ii, G.L. (2011) Proteasome inhibition in myelodysplastic syndromes and acute myelogenous leukemia cell lines. *Cancer Investigation*, **29**, 439–450.

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., Hellstrom-Lindberg, E. & Cazzola, M. (2011) Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. *Blood*, **118**, 6239–6246.

Malcovati, L., Karimi, M., Papaemmanuil, E., Ambaglio, I., Jadersten, M., Jansson, M., Elena, C., Galli, A., Walldin, G., Della Porta, M.G., Raaschou-Jensen, K., Travaglino, E., Kallenbach, K., Pietra, D., Ljungstrom, V., Conte, S., Boveri, E., Invernizzi, R., Rosenquist, R., Campbell, P.J., Cazzola, M. & Hellstrom Lindberg, E. (2015) SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood*, **126**. Epub ahead of print. doi: 10.1182/blood-2015-03-633537

Miyata, M. & Smith, J.D. (1996) Apolipoprotein E allele-specific antioxidant activity and effects on cytotoxicity by oxidative insults and beta-amyloid peptides. *Nature Genetics*, **14**, 55–61.

Moestrup, S.K. & Moller, H.J. (2004) CD163: a regulated hemoglobin scavenger receptor with a role in the anti-inflammatory response. *Annals of Medicine*, **36**, 347–354.

Mufti, G.J., Bennett, J.M., Goasguen, J., Bain, B.J., Baumann, I., Brunning, R., Cazzola, M., Fenaux, P., Germing, U., Hellstrom-Lindberg, E., Jinnai, I., Manabe, A., Matsuda, A., Niemeyer, C.M., Sanz, G., Tomonaga, M., Vallespi, T. & Yoshimi, A. (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**, 1712–1717.

Murru, S., Loudianos, G., Deiana, M., Camaschella, C., Sciarratta, G.V., Agosti, S., Parodi, M.I., Cerruti, P., Cao, A. & Pirastu, M. (1991) Molecular characterization of beta-thalassemia intermedia in patients of Italian descent and identification of three novel beta-thalassemia mutations. *Blood*, **77**, 1342–1347.

Nikpour, M., Pellagatti, A., Liu, A., Karimi, M., Malcovati, L., Gogvadze, V., Forsblom, A.M., Wainscoat, J.S., Cazzola, M., Zhivotovsky, B., Grandien, A., Boultwood, J. & Hellstrom-Lindberg, E. (2010) Gene expression profiling of erythroblasts from refractory anaemia with ring sideroblasts (RARS) and effects of G-CSF. *British Journal of Haematology*, **149**, 844–854.

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., Stephen Wainscoat, J., Campbell, P., Cazzola, M., Boultwood, J., Grandien, A. & Hellstrom-Lindberg, E. (2013) The transporter ABCB7 is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts. *Leukemia*, **27**, 889–896.

Nilsson, R., Schultz, I.J., Pierce, E.L., Soltis, K.A., Naranuntarat, A., Ward, D.M., Baughman, J.M., Paradkar, P.N., Kingsley, P.D., Culotta, V.C., Kaplan, J., Palis, J., Paw, B.H. & Mootha, V.K. (2009) Discovery of genes essential for heme biosynthesis through large-scale gene expression analysis. *Cell Metabolism*, **10**, 119–130.

Papaemmanuil, E., Cazzola, M., Boultwood, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J.S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., Godfrey, A.L., Rapado, I., Cvejic, A., Rance, R., McGee, C., Ellis, P., Mudie, L.J., Stephens, P.J., McLaren, S., Massie, C.E., Tarpey, P.S., Varela, I., Nik-Zainal, S., Davies, H.R., Shlien, A., Jones, D., Raine, K., Hinton, J., Butler, A.P., Teague, J.W., Baxter, E.J., Score, J., Galli, A., Della Porta, M.G., Travaglino, E., Groves, M., Tauro, S., Munshi, N.C., Anderson, K.C., El-Naggar, A., Fischer, A., Mustonen, V., Warren, A.J., Cross, N.C., Green, A.R., Futreal, P.A., Stratton, M.R. & Campbell, P.J. (2011) Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *New England Journal of Medicine*, **365**, 1384–1395.

Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., Pellagatti, A., Groves, M.J., Cross, N.C.P., Gambacorti-Passerini, C., Green, A.R., Boultwood, J., Vyas, P., Hellstrom-Lindberg, E., Bowen, D., Cazzola, M., Stratton, M.R. & Campbell, P. (2012) High Throughput Targeted Gene Sequencing in 738 Myelodysplastic Syndromes Patients Reveals Novel Oncogenic Genes, Rare Driver Mutations and Complex Molecular Signatures with Potential Impact for Patient Diagnosis and Prognosis in the Clinic. *Blood (ASH Annual Meeting Abstracts)*, **120**, LBA-5.

Papaemmanuil, E., Gerstung, M., Malcovati, L., Tauro, S., Gundem, G., Van Loo, P., Yoon, C.J., Ellis, P., Wedge, D.C., Pellagatti, A., Shlien, A., Groves, M.J., Forbes, S.A., Raine, K., Hinton, J., Mudie, L.J., McLaren, S., Hardy, C., Latimer, C., Della Porta, M.G., O'Meara, S., Ambaglio, I., Galli, A., Butler, A.P., Walldin, G., Teague, J.W., Quek, L., Sternberg, A., Gambacorti-Passerini, C., Cross, N.C., Green, A.R., Boultwood, J., Vyas, P., Hellstrom-Lindberg, E., Bowen, D., Cazzola, M., Stratton, M.R. & Campbell, P.J. & Chronic Myeloid Disorders Working Group of the International Cancer Genome Consortium. (2013) Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood*, **122**, 3616–3627; quiz 3699.

Pellagatti, A., Cazzola, M., Giagounidis, A.A., Malcovati, L., Porta, M.G., Killick, S., Campbell, L.J., Wang, L., Langford, C.F., Fidler, C., Oscier, D., Aul, C., Wainscoat, J.S. & Boultwood, J. (2006) Gene expression profiles of CD34 + cells in myelodysplastic syndromes: involvement of interferon-stimulated genes and correlation to FAB subtype and karyotype. *Blood*, **108**, 337–345.

Pellagatti, A., Jadersten, M., Forsblom, A.M., Cattan, H., Christensson, B., Emanuelsson, E.K., Merup, M., Nilsson, L., Samuelsson, J., Sander, B., Wainscoat, J.S., Boultwood, J. & Hellstrom-Lindberg, E. (2007) Lenalidomide inhibits the malignant clone and up-regulates the SPARC gene mapping to the commonly deleted region in 5q- syndrome patients. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 11406–11411.

Petruzzelli, R., Palumbo, I., Gaudino, S., Scalia, G., Del Vecchio, L., Izzo, P. & Grosso, M. (2009) An optimized method to separate reticulocytes from peripheral blood for molecular analysis. *International Journal of Laboratory Hematology*, **31**, 359–363.

Pimentel, H., Parra, M., Gee, S., Ghanem, D., An, X., Li, J., Mohandas, N., Pachter, L. & Conboy, J.G. (2014) A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. *Nucleic Acids Research*, **42**, 4031–4042.

Saito, Y., Hayashi, T., Tanaka, A., Watanabe, Y., Suzuki, M., Saito, E. & Takahashi, K. (1999)

Selenoprotein P in human plasma as an extracellular phospholipid hydroperoxide glutathione peroxidase. Isolation and enzymatic characterization of human selenoprotein p. *Journal of Biological Chemistry*, **274**, 2866–2871.

Shen, T., Zheng, M., Cao, C., Chen, C., Tang, J., Zhang, W., Cheng, H., Chen, K.H. & Xiao, R.P. (2007) Mitofusin-2 is a major determinant of oxidative stress-mediated heart muscle cell apoptosis. *Journal of Biological Chemistry*, **282**, 23354–23361.

Shiozawa, Y., Sato-Otsubo, S., Gallì, A., Yoshida, K., Yoshizato, T., Sato, Y., Kataoka, K., Sanada, M., Shiraishi, Y., Chiba, K., Miyano, S., Malcovati, L., Cazzola, M. & Ogawa, S. (2014) Comprehensive Analysis of Aberrant RNA Splicing in Myelodysplastic Syndromes. *Blood*, **124**, 826.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. & Mesirov, J.P. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 15545–15550.

Suragani, R.N., Cadena, S.M., Cawley, S.M., Sako, D., Mitchell, D., Li, R., Davies, M.V., Alexander, M.J., Devine, M., Loveday, K.S., Underwood, K.W., Grinberg, A.V., Quisel, J.D., Chopra, R., Pearsall, R.S., Seehra, J. & Kumar, R. (2014) Transforming growth factor-beta superfamily ligand trap ACE-536 corrects anemia by promoting late-stage erythropoiesis. *Nature Medicine*, **20**, 408–414.

Tanno, T. & Miller, J.L. (2010) Iron loading and overloading due to ineffective erythropoiesis. *Advances in Hematology*, **2010**, 358283.

Tehranchi, R., Fadeel, B., Forsblom, A.M., Christensson, B., Samuelsson, J., Zhivotovsky, B. & Hellstrom-Lindberg, E. (2003) Granulocyte colony-stimulating factor inhibits spontaneous cytochrome c release and mitochondria-dependent apoptosis of myelodysplastic syndrome hematopoietic progenitors. *Blood*, **101**, 1080–1086.

Tehranchi, R., Invernizzi, R., Grandien, A., Zhivotovsky, B., Fadeel, B., Forsblom, A.M., Travaglino, E., Samuelsson, J., Hast, R., Nilsson, L., Cazzola, M., Wibom, R. & Hellstrom-Lindberg, E. (2005a) Aberrant mitochondrial iron distribution and maturation arrest characterize early erythroid precursors in low-risk myelodysplastic syndromes. *Blood*, **106**, 247–253.

Tehranchi, R., Fadeel, B., Schmidt-Mende, J., Forsblom, A.M., Emanuelsson, E., Jadersten, M., Christensson, B., Hast, R., Howe, R.B., Samuelsson, J., Zhivotovsky, B. & Hellstrom-Lindberg, E. (2005b) Antiapoptotic role of growth factors in the myelodysplastic syndromes: concordance between in vitro and in vivo observations. *Clinical Cancer Research*, **11**, 6291–6299.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L. & Pachter, L. (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, **7**, 562–578.

Vardiman, J.W., Thiele, J., Arber, D.A., Brunning, R.D., Borowitz, M.J., Porwit, A., Harris, N.L., Le Beau, M.M., Hellstrom-Lindberg, E., Tefferi, A. & Bloomfield, C.D. (2009) The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood*, **114**, 937–951.

Visconte, V., Rogers, H.J., Singh, J., Barnard, J., Bupathi, M., Traina, F., McMahon, J., Makishima, H., Szpurka, H., Jankowska, A., Jerez, A., Sekeres, M.A., Sauntharajah, Y., Advani, A.S., Copelan, E., Koseki, H., Isono, K., Padgett, R.A., Osman, S., Koide, K., O'Keefe, C., Maciejewski, J.P. & Tiu, R.V. (2012) SF3B1 haploinsufficiency leads to formation of ring sideroblasts in myelodysplastic syndromes. *Blood*, **120**, 3173–3186.

Visconte, V., Avishai, N., Mahfouz, R., Tabarroki, A., Cowen, J., Sharghi-Moshtaghin, R., Hitomi, M., Rogers, H.J., Hasrouni, E., Phillips, J., Sekeres, M.A., Heuer, A.H., Sauntharajah, Y., Barnard, J. & Tiu, R.V. (2015) Distinct iron architecture in SF3B1-mutant myelodysplastic syndrome patients is linked to an SLC25A37 splice variant with a retained intron. *Leukemia*, **29**, 188–195.

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

Yien, Y.Y., Robledo, R.F., Schultz, I.J., Takahashi-Makise, N., Gwynn, B., Bauer, D.E., Dass, A., Yi, G., Li, L., Hildick-Smith, G.J., Cooney, J.D., Pierce, E.L., Mohler, K., Dailey, T.A., Miyata, N., Kingsley, P.D., Garone, C., Hattangadi, S.M., Huang, H., Chen, W., Keenan, E.M., Shah, D.I., Schlaeger, T.M., DiMauro, S., Orkin, S.H., Cantor, A.B., Palis, J., Koehler, C.M., Lodish, H.F., Kaplan, J., Ward, D.M., Dailey, H.A., Phillips, J.D., Peters, L.L. & Paw, B.H. (2014) TMEM14C is required for erythroid mitochondrial heme metabolism. *The Journal of Clinical Investigation*, **124**, 4294–4304.

Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., Chalkidis, G., Suzuki, Y., Shiosaka, M., Kawahata, R., Yamaguchi, T., Otsu, M., Obara, N., Sakata-Yanagimoto, M., Ishiyama, K., Mori, H., Nolte, F., Hofmann, W.K., Miyawaki, S., Sugano, S., Haferlach, C., Koeffler, H.P., Shih, L.Y., Haferlach, T., Chiba, S., Nakauchi, H., Miyano, S. & Ogawa, S. (2011) Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature*, **478**, 64–69.