



HAL
open science

Conséquences fonctionnelles des mutations du gène SF3B1 dans les syndromes myélodysplasiques avec sidéroblastes en couronne

Sabrina Bondu

► **To cite this version:**

Sabrina Bondu. Conséquences fonctionnelles des mutations du gène SF3B1 dans les syndromes myélodysplasiques avec sidéroblastes en couronne. Hématologie. Université Sorbonne Paris Cité, 2019. Français. NNT : 2019USPCB002 . tel-02495356

HAL Id: tel-02495356

<https://theses.hal.science/tel-02495356>

Submitted on 2 Mar 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Thèse de doctorat de l'Université Paris Descartes (Paris V)
École doctorale Hématologie, Oncogénèse et Biothérapies (HOB, ED 561)

Spécialité Hématologie

Présentée et soutenue publiquement le 28/02/2019 par
Sabrina Bondu

Conséquences fonctionnelles des mutations du gène *SF3B1* dans les syndromes myélodysplasiques avec sidéroblastes en couronne

Équipe « Hématopoïèse normale et pathologique »
Département Développement, Reproduction et Cancer
Institut Cochin, INSERM U1016, Paris

Directrice de thèse : Pr Michaela Fontenay

Jury

Pr Marie-Bérengère Troadec	Rapporteur
Dr Sarah Lambert	Rapporteur
Dr Sophie Vaulont	Examineur
Dr Marc-Henri Stern	Examineur
Dr Jean-Baptiste Micol	Examineur

Remerciements

Comme SF3B1 qui s'entoure d'une multitude de partenaires protéiques au sein de la cellule, de nombreuses personnes m'ont accompagnée pendant mon master 2 et ma thèse de doctorat. En espérant que vous trouverez dans les lignes suivantes tout le témoignage de ma gratitude.

Au Professeur Marie-Bérengère Troadec et au Docteur Sarah Lambert, rapporteurs de ma thèse. Je vous remercie d'avoir accepté d'évaluer mon travail.

Aux Docteurs Marc-Henri Stern, Sophie Vaulont et Jean-Baptiste Micol, examinateurs de ma thèse. Je vous remercie d'être présents dans mon jury et d'apporter votre expertise. Marc-Henri, merci d'avoir participé à mes comités de suivi de thèse au cours desquels vous avez exprimé vos remarques constructives.

Au Professeur Michaela Fontenay, ma directrice de thèse, pour son enthousiasme scientifique, ses encouragements et sa détermination. Je vous remercie d'avoir mis en place les collaborations nécessaires à l'avancée de ma thèse, et je vous remercie de votre aide tout au long de ces quatre dernières années.

Au Docteur Patrick Mayeux et au Professeur Didier Bouscary qui m'ont accueillie dans leur équipe. Patrick, merci de nous faire partager ta passion pour la protéomique et l'érythroïèse, merci pour ta disponibilité et le regard critique que tu portes sur chaque expérience.

À tous ceux qui m'ont quotidiennement aidée au laboratoire :

- ❖ À Olivier, pour son investissement considérable dans le projet « SF3B1 et ERFE », son humour et sa perspicacité.
- ❖ À Carine, pour sa bienveillance, sa sérénité et sa gentillesse. Merci d'avoir eu la patience de m'apprendre toutes les techniques de biologie cellulaire et biologie moléculaire, de répondre à chacune de mes questions, de relire chacun de mes rapports de master 2 et chacune de mes thèses toujours avec la même application. Ton calme et ton écoute ont dissipé bien des nuages.
- ❖ À Anne-Sophie, pour son assiduité (cf. les dizaines et dizaines de RT-qPCR, PCR et RT-PCR) et sa sympathie à mon égard. J'ai apprécié travailler à tes côtés.
- ❖ À Alice, pour son esprit créatif et pragmatique. Merci de m'avoir tant de fois prêté main-forte, notamment pour les extractions d'ADN/ARN et les Ficoll/tris CD34⁺.
- ❖ À Laila, qui a repris le flambeau d'Alice et des RT-qPCR. Merci pour ta bonté.
- ❖ À David, qui a cette merveilleuse qualité de rendre les inextricables données de bio-informatique claires et concises. Et à Abderrahmane, qui a déniché l'aiguille (ERFE) dans la botte de foin (les RNA-Seq).

À tous nos collaborateurs :

- ❖ Au Dr Léon Kautz – alias Monsieur Érythroferrone – qui nous a beaucoup éclairés au sujet de sa protéine fétiche. Merci pour ta grande lucidité et ta participation au projet « SF3B1 et ERFÉ » et à mes comités de suivi de thèse.
- ❖ Au Dr Samar Alsafadi et à Alexandre Houy avec qui j'ai fréquemment discuté pendant ma thèse. Samar, merci pour ta clairvoyance et tes explications sur la technique des minigènes. Merci d'avoir toi aussi été présente à mes comités de suivi de thèse. Alexandre, merci d'avoir réalisé l'analyse bio-informatique des jonctions alternatives dans les cellules mononucléées, analyse qui a servi de point de départ au projet « SF3B1 et ERFÉ ».
- ❖ À François Guillonnet et à l'ensemble de la plateforme de protéomique 3P5 qui ont eu la lourde tâche de chercher l'aiguille (ERFÉ, encore et toujours !) dans la botte de foin (les plasmas et lysats d'érythroblastes).
- ❖ À toutes les personnes qui ont contribué au projet « SF3B1 et ERFÉ » : les Prs Tomas Ganz et Elizabeta Nemeth ainsi que Grace Jung pour les dosages de l'érythroferrone plasmatique ; le Pr Hervé Puy ainsi que les Drs Zoubida Karim et Thibaud Lefebvre pour les dosages de l'hepcidine plasmatique ; le Pr Sophie Park pour les analyses statistiques multivariées ; Nicolas Cagnard pour l'analyse bio-informatique de l'expression génique dans les cellules mononucléées ; et, bien sûr, tous les cliniciens et biologistes – en France, en Allemagne et en Italie – qui nous ont fourni les précieux échantillons de moelle osseuse et de sang.
- ❖ Aux Drs Benoît Miotto et Anne Letessier, qui ont quant à eux contribué au projet « SF3B1 et réplication de l'ADN ». Benoît, je te remercie d'avoir pris part à mes comités de suivi de thèse et de nous avoir orientés vers Frédéric Chédin pour l'analyse des *R-loops*. Anne, un grand – que dis-je, un gigantesque – merci pour ton aide de A à Z sur la technique de peignage moléculaire. Je te remercie d'avoir été aussi accueillante et bienveillante avec moi.
- ❖ Au Dr Angelos Constantinou, pour nos échanges sur les protéines SF3B1, FANCD2 et FANCI.

Aux ingénieurs et techniciens des plateformes de cytométrie, de génomique et d'imagerie de l'Institut Cochin, grâce à qui nous disposons d'un matériel performant. À Lucie Sengmanivong et à l'ensemble du centre d'imagerie Nikon de l'Institut Curie qui m'ont si gentiment reçue pour l'acquisition des images de fibres d'ADN.

À l'Institut National du Cancer et à la Société Française d'Hématologie qui ont financé ma thèse de doctorat.

À tous les patients qui ont donné leur consentement pour l'utilisation des échantillons biologiques à des fins de recherche.

Aux étudiants de l'ancienne équipe Mayeux/Bouscary, avec qui j'ai partagé de longues et laborieuses journées au laboratoire mais aussi beaucoup de rires et de discussions animées. Votre amitié revigorante m'a souvent été d'un grand réconfort. Je vous souhaite bon vent, puisse-t-il régulièrement nous rassembler autour du couscous du Val Royal !

- ❖ À Salomé, ma délicate et sensible co-doctorante. Délicate parce que derrière une apparente discrétion se cache une jeune femme prévenante envers sa famille et ses amis. Sensible parce que tu discernes les nuances en chacun de nous, et il est toujours agréable de se confier à toi. De cette thèse – et de ces innombrables heures au L1, au L2, dans le bureau, mais aussi rue Descartes ou ailleurs – est née une vraie complicité. Je te remercie pour ta présence apaisante et réjouissante.
- ❖ À Fetta, dont la force n'a d'égale que la générosité. Déjà quatre années passées à tes côtés et je n'ai pourtant pas encore fait le tour de ton immense cœur. Merci pour ton authenticité, tes éclats de rire qui résonnent dans tout le 7^{ème} étage et (surtout) tes délicieuses bricks ! J'espère un jour contempler les montagnes de ta Kabylie natale de mes propres yeux.
- ❖ À Meriem, douce, curieuse et tolérante. Il me tarde de te revoir, il me tarde de repasser des heures à parler de tout et de rien, et rire aussi de tout et de rien. Ta famille est à l'image de la ville d'Oran : « Généreuse et spontanée, il n'était pas question, pour elle, de se découvrir une joie sans songer à la partager » (Yasmina Khadra).
- ❖ À Justine. Ta persévérance et ton empathie sont une richesse. Nos confidences au petit matin, ta grandeur d'âme et ta quête pour le beau – j'entends par là l'émotion que tu recherches dans chaque poésie, roman, musique, film, pièce de théâtre mais aussi dans chaque personne que tu côtoies – vont beaucoup me manquer au quotidien.
- ❖ À Adrien, friand des débats passionnés et fougueux. Merci de répandre autant de gaieté autour de toi. À Pierre, pour son rire communicatif. Le bureau est bien (trop) calme depuis votre départ.
- ❖ À Ismaël et Charles, toujours prêts à rendre service. À Rudy, pour ses farces récréatives, sa bonne humeur et son énergie à toute épreuve. À Johanna, solaire et paisible. Je vous remercie tous les quatre pour votre humour et votre bienveillance.
- ❖ À Godelieve, qui prend soin de tout le monde et ne juge personne. Merci pour tes jeux de mots et anecdotes amusantes. À Gabriel et Cyrielle, pour ces discussions empreintes d'évasion : quelque part dans un pays du Nord avec toi Gabriel, et quelque part dans notre contrée aux saveurs de crêpes et caramel au beurre salé avec toi Cyrielle.
- ❖ À Marjorie, Zubaidan et Samia, à qui je souhaite – ainsi qu'à Salomé, Justine, Ismaël, Charles, Rudy, Johanna et Cyrielle – une bonne fin de thèse.
- ❖ Aux stagiaires de master 2 qui, chaque début d'année, apportent fraîcheur et jovialité au 7^{ème} étage : Florian, Myriam, Sylvain, Sophie, Clarisse et tous les autres. Et à Camille également, qui a contribué au projet « SF3B1 et réplication de l'ADN » lors de son stage de master 1.

Aux « chefs » de l'ancienne équipe Mayeux/Bouscary : Frédérique Verdier – que je remercie tout particulièrement pour sa disponibilité et sa gentillesse – Yaël Zermati, Frédéric Pendino, Isabelle Dusanter, Evelyne Lauret, Nicolas Chapuis et Jérôme Tamburini. Merci pour votre aide et vos multiples conseils. À l'heure où vous découvrez quasiment tous de nouveaux horizons, je vous souhaite beaucoup de succès dans vos futurs projets.

À Maria-Lilia et Jaya, enthousiastes et avenantes.

À Émilie-Fleur et Sarah, chercheurs post-doctorantes. Merci de m'avoir consciencieusement initiée à la culture des érythroblastes primaires.

À Lamya et Daniela, qui ont en commun l'élégance et le sourire chaleureux des Méditerranéens. Du Maroc à l'Italie, en passant par la France, c'est finalement en Algérie que nos chemins se sont croisés. Lamya, merci de veiller sur moi telle une grande sœur, merci pour tes nombreuses attentions, ton écoute et tes conseils toujours très avisés. J'admire la femme que tu es, ton ouverture d'esprit, ta sagesse et ta générosité. Daniela, affable et respectueuse, tu aimes rassembler tes amis autour d'une belle table. Merci pour ces soirées enchantées que tu nous offres, à Lamya et moi.

À Shanthi, Christophe et Ségolène qui m'ont gentiment accueillie dans la pièce du groupe Filo lors de mes extractions d'ADN/ARN et RT-qPCR.

À mes amis : ceux de Paris qui, chacun à leur façon, rendent le ciel un peu moins gris et ceux de Nantes ou d'ailleurs qui, malgré la distance, n'ont jamais cessé de m'encourager.

- ❖ À Anaïs, qui a l'audace et le courage de réaliser ses rêves. Du bac à sable de la petite section jusqu'aux plages de sable australiennes, vingt-cinq années se sont écoulées dans le sablier. Vingt-cinq années auprès de ta sensibilité, ton imagination et ton humour.
- ❖ À Mélinda et Noémie, les années passent mais « on ne change pas ». Qu'il est apaisant de vous savoir présentes à mes côtés quoi qu'il arrive, avec votre authenticité et votre noblesse d'âme. À Fabian et Mario, mes joyeux lurons préférés. Merci à tous les quatre d'avoir accepté mes choix d'études, choix qui m'ont conduite à 400 km de votre bonne humeur et de celle d'Alisson, Sandie et Honorine.
- ❖ À Stéven, Chloé et Fleur (ou Fleur et Chloé, je ne sais plus), drolatiques, poétiques et attachants. Je vous remercie pour ces instants de détente pendant lesquels, entre deux éclats de rire, nous refaisons le monde et nous inventons des talents de critique littéraire, musicale et cinématographique. Vous êtes ma bouffée d'air frais à Paris. Stéven, mille mercis pour ton écoute, ton éclectisme et ta formidable capacité à saisir la nature profonde d'autrui. J'aurais aimé glisser ici une citation intelligente de Boris Vian ou Romain Gary, mais seules les paroles entêtantes d'une chanteuse oubliée de tous (sauf de toi) me viennent à l'esprit : « Ce soir on sort, on oublie nos galères ».
- ❖ À Caroline, juste et déterminée. Je vous souhaite, à Jean et toi, le plus beau des voyages en Amérique du Sud.
- ❖ À Claudine dont les confitures ont embelli bien des matins. À Julie qui, avec Claudine, m'a poussée à demander un stage à Gustave Roussy, stage à l'origine de mon aventure parisienne.

- ❖ À Marie, qui est venue du Sud-ouest avec son franc-parler, sa chaleur et sa passion pour les cultures orientales.
- ❖ À Célia, force de la nature, et à Sarah, touchante et égayante.

À ma famille, pour son soutien infailible.

- ❖ À Fabrice, mon truculent beau-frère qui donne des airs de fête à chaque repas.
- ❖ À Laurent et Laurence, ma fée marraine, pour nos conversations tantôt légères tantôt réfléchies. Merci de m'accueillir à bras ouverts dans votre maison de l'hédonisme.
- ❖ À Marjorie, avec qui j'ai parcouru des pays insolites avec joie et excitation. Merci pour ta spontanéité et ta curiosité intellectuelle.
- ❖ À Anthynéa, à qui je souhaite tout le bonheur et la réussite qu'elle mérite.
- ❖ À Jean-Pierre, qui nous gâte tant, Magalie et moi.
- ❖ À Estelle, mon phare (breton) de Montrouge. Merci de m'avoir initiée aux fabuleux cours de yoga d'Irène !

À ma grande sœur, ma meilleure amie, qui, depuis ma naissance, saupoudre bonté et bienveillance au-dessus de moi. De tes mains habiles, guidée par ton cœur fondant, tu fais lentement mijoter notre fraternité, l'assaisonne d'un zeste de complicité jusqu'à parfaite harmonie. Maman serait très fière de la jeune femme altruiste et tolérante que tu es devenue.

À mon père, pour la liberté et la confiance qu'il m'a toujours accordées. Et pour la liste infinie de tout ce qui m'a permis de me construire : nos vacances aux quatre coins de la France, tes encouragements à chaque compétition sportive, chaque examen, chaque départ vers une nouvelle destination, ta force, ton humilité, ton dévouement.

À ma mère, pour son amour qui traverse le temps sans jamais faiblir.

Mes derniers remerciements seront pour ma nièce et mon neveu, oasis de douceur dans le tumulte de la thèse. Pour vos rires d'enfants et vos petites mains chaudes qui s'agrippent à mes jambes, témoignages candides de votre tendresse. À l'aube d'une vie pleine de promesses, merci d'avoir déjà autant ensoleillé la mienne.

Résumé

Les syndromes myélodysplasiques avec sidéroblastes en couronne (SMD-SC) sont des pathologies clonales de la cellule souche hématopoïétique caractérisées par une érythropoïèse inefficace et une surcharge mitochondriale et systémique en fer. La médiane d'âge au diagnostic est de 70 ans. Le gène *SF3B1* est muté chez près de 90% des patients porteurs d'un SMD-SC. Il code pour un facteur d'épissage appartenant au complexe nucléaire ribonucléoprotéique U2 du spliceosome. L'interaction de SF3B1 avec l'intron de part et d'autre du site de branchement permet le recrutement du complexe U2 lors de l'épissage co-transcriptionnel des transcrits naissants. SF3B1 s'associe également avec la chromatine et sa dynamique de distribution dans le noyau est sous le contrôle de diverses protéines de réparation de l'ADN, telles que les « gardiennes » du réplisome FANCD2 et FANCI, et BRCA1. Les conséquences fonctionnelles des mutations du gène *SF3B1* (*SF3B1*^{MUT}) dans les SMD-SC sont mal connues.

Lors de ma thèse, je me suis dans un premier temps intéressée aux anomalies d'épissage générées dans les cellules des patients porteurs d'un syndrome myélodysplasique (SMD) avec mutation sur *SF3B1*. En accord avec les précédentes analyses de séquençage de l'ARN dans les SMD et autres pathologies *SF3B1*^{MUT}, les mutations de *SF3B1* sont associées à une utilisation préférentielle de sites accepteurs alternatifs dans les cellules mononucléées (CMN) des patients porteurs d'un SMD. Notre travail a permis d'identifier la surexpression d'un transcrit alternatif de l'érythroferrone (ERFE) dans les CMN et érythroblastes en culture *SF3B1*^{MUT}. ERFE est une hormone sécrétée par les érythroblastes qui régule l'expression hépatique de l'hepcidine, cette dernière limitant la biodisponibilité du fer dans le plasma. Le transcrit alternatif d'ERFE (ERFE⁺¹²) n'est retrouvé que dans les cellules érythroïdes *SF3B1*^{MUT} et le ratio d'expression de la jonction aberrante par rapport à la jonction totale reflète l'érythropoïèse clonale induite par les mutations de *SF3B1*. ERFE⁺¹² est traduit en un variant protéique fonctionnel. L'augmentation des concentrations des protéines ERFE canonique et alternative dans le plasma des patients SMD *SF3B1*^{MUT} est corrélée à la diminution des taux d'hepcidine, et explique le développement précoce d'une surcharge systémique en fer. Dans un deuxième temps, j'ai caractérisé – au cours de la différenciation érythrocytaire – les processus de réplication et de réparation de l'ADN dans les érythroblastes primaires humains *SF3B1*^{MUT} et dans la lignée érythroïde murine G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E}. Nos expériences de cycle cellulaire et de peignage moléculaire démontrent que les mutations de *SF3B1* confèrent un avantage de prolifération aux précurseurs érythroïdes immatures en augmentant la dynamique de réplication de l'ADN, *via* l'accélération de la progression des fourches. Il a récemment été montré que l'augmentation de la vitesse des fourches de réplication s'accompagne d'une exposition d'ADN simple-brin recouvert par la protéine p-RPA32, ce que nous confirmons. Il n'y a par contre pas de cassure double-brin de l'ADN. Les mutations de *SF3B1* contribuent donc à accroître de façon inadéquate le compartiment des cellules érythroïdes immatures dans la moelle osseuse des patients SMD-SC. Nos prochains travaux viseront à déterminer par quels mécanismes les fourches de réplication sont accélérées, notamment avec l'exploration des intermédiaires de transcription *R-loops*.

Abstract

Myelodysplastic syndromes with ring sideroblasts (MDS-RS) are clonal hematopoietic stem cell (HSC) disorders characterized by anemia, bone marrow (BM) erythroid dysplasia and systemic iron accumulation. Mitochondrial iron accumulation and apoptosis of mature erythroblasts cause ineffective erythropoiesis. Splicing factor gene *SF3B1* is mutated in ~90% of MDS-RS. *SF3B1* mediates U2 snRNP recruitment to the branch point (BP) by interacting with the intronic RNA on both sides of the BP. *SF3B1* also interacts with nucleosomes located at exonic positions. The nuclear dynamics of *SF3B1* is under the control of Fanconi anemia FANCD2 and FANCI proteins, suggesting coordination between DNA repair pathway and cotranscriptional splicing during S phase. In response to DNA damage, the BRCA1/*SF3B1* complex facilitates the splicing of DNA repair genes. Functional consequences of *SF3B1* mutations in MDS-RS are still unknown.

We first assessed the consequences of *SF3B1* gene mutations on gene expression and splicing by sequencing the transcriptome of the BM mononuclear cells (MNC). *SF3B1* hotspot mutations induce cryptic 3' splice site selection through use of a different BP, as previous studies in MDS and other *SF3B1*-driven cancers have already shown. We identified an alternative transcript of erythroferrone (ERFE) in MNC of MDS patients with *SF3B1* mutation. ERFE has recently been described as a major erythroid regulator of hepcidin, a key hormone that is involved in the control of iron homeostasis in the body. The expression of the variant *ERFE* transcript was restricted to *SF3B1*-mutated erythroblasts and allows the monitoring of clonal erythropoiesis. ERFE neo-isoform (ERFE⁺¹²) accounts for an overall increase of *ERFE* expression in *SF3B1*^{MUT} MDS. ERFE⁺¹² is translated into a functional variant protein in *SF3B1*^{MUT} erythroblasts. Plasma concentrations of both canonical and aberrant ERFE proteins are significantly more elevated in *SF3B1*^{MUT} MDS compared to *SF3B1*^{WT} MDS and inversely correlated to hepcidin/ferritin ratio. Thus, we identified a *SF3B1*-dependent and erythroid cell intrinsic mechanism of hyperferritinemia in MDS-RS.

We next analysed the DNA replication and DNA repair processes in human *SF3B1*^{MUT} primary erythroblasts and in murine CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 proerythroblastic cell line. We confirmed that *SF3B1*^{MUT} erythroblasts have an increased proliferation rate, but they enter into terminal erythroid differentiation more rapidly than healthy controls or *SF3B1*^{WT} MDS. We reported an increased DNA replication dynamics, since the rate of DNA fiber elongation is accelerated in *Sf3b1*^{K700E} G1E-ER4 cells. It has recently been suggested that high speed of fork progression induces DNA replication stress. Consistently, we observed the spontaneous accumulation of single-stranded DNA covered by replication protein A in actively dividing *SF3B1*^{MUT} proerythroblast and basophilic erythroblasts. No DNA-double strand break and no fork collapse were observed in *SF3B1*^{MUT} cells. Thus, *SF3B1* mutations induce a replication stress without stalling forks and contribute to the development of a clonal erythropoiesis. Mechanistically, we hypothesize that replication forks encounter less obstacles in *SF3B1*-mutated compared to *SF3B1*-wildtype erythroblasts. To address this point, we will investigate the presence of R-loops at the promoter regions of selected genes.

Sommaire

Liste des abréviations.....	11
Lexique.....	15
Liste des tableaux et figures.....	16
Préface.....	18
Chapitre I – Érythropoïèse normale et pathologique.....	19
I. L'érythropoïèse humaine normale.....	20
A. Généralités.....	20
1. L'hématopoïèse humaine.....	20
2. Ontogénie de l'érythropoïèse.....	23
B. Différenciation érythroïde.....	24
1. L'érythropoïèse précoce.....	26
2. L'érythropoïèse terminale.....	27
a. Caractéristiques morphologiques et immunophénotypiques des précurseurs érythroïdes.....	27
b. Mécanisme de condensation de la chromatine.....	29
C. Régulation de l'érythropoïèse.....	31
1. Régulation par les facteurs de croissance.....	31
a. Le SCF et son récepteur c-KIT.....	31
b. L'EPO et son récepteur EPOR.....	32
c. Voies de signalisation activées en aval de c-KIT et EPOR.....	34
2. Régulation transcriptionnelle.....	38
3. Régulation par les caspases.....	39
a. Rappels sur l'apoptose cellulaire.....	39
b. Rôle des caspases dans la différenciation érythroïde terminale.....	40
c. Rôle des caspases dans la régulation négative de l'érythropoïèse.....	42
D. Modulation de l'expression génique au cours de l'érythropoïèse.....	44
1. Au niveau transcriptionnel.....	44
2. Au niveau protéique.....	45
3. Régulation post-transcriptionnelle de l'expression génique : rôle de l'épissage	46
E. Érythropoïèse et fer.....	49
1. La distribution du fer dans l'organisme.....	49
2. Le métabolisme du fer.....	50
3. Les îlots érythroblastiques.....	54

4.	Régulation du métabolisme du fer.....	56
a.	Au niveau intracellulaire	56
b.	Au niveau systémique.....	57
II.	Un exemple d'érythropoïèse pathologique : les syndromes myélodysplasiques avec sidéroblastes en couronne.....	67
A.	Les anémies sidéroblastiques	67
1.	Les sidéroblastes en couronne	68
2.	Classification des anémies sidéroblastiques	68
a.	Les anémies sidéroblastiques congénitales.....	68
b.	Les anémies sidéroblastiques acquises	70
B.	Les syndromes myélodysplasiques	71
1.	Définition et épidémiologie des SMD	71
2.	Classifications diagnostique et pronostique des SMD	72
a.	Classification diagnostique OMS 2016	72
b.	Classification pronostique IPSS-R	75
C.	Physiopathologie des SMD-SC.....	77
1.	Dysérythropoïèse terminale	77
a.	Apoptose mitochondriale des érythroblastes	77
b.	Accumulation de fer dans les mitochondries des érythroblastes	78
2.	Surcharge systémique en fer	79
3.	Conclusion	80
D.	Principaux traitements des SMD-SC.....	82
1.	Traitements actuels	82
2.	Traitement en cours d'évaluation clinique	83
	Chapitre II – La protéine SF3B1	86
	Partie 1	87
I.	La protéine SF3B1 et l'épissage des ARN pré-messagers.....	88
A.	Épissage des ARN pré-messagers	88
B.	Structure de la protéine SF3B1	92
1.	Domaines protéiques de SF3B1	92
2.	Phosphorylation de SF3B1	93
C.	Localisations intranucléaires de SF3B1	94
1.	<i>Nuclear speckles</i>	94
2.	Nucléosomes.....	96
D.	Études d'invalidation du gène <i>Sf3b1</i>	98
II.	Mutations de <i>SF3B1</i> et dérégulation de l'épissage	99

A.	Description des mutations de <i>SF3B1</i> dans les SMD	99
B.	Caractéristiques clinico-biologiques des souris <i>knock-in Sf3b1</i> ^{MUT}	103
C.	Épissage alternatif en aval des mutations de <i>SF3B1</i>	104
1.	Définition de l'épissage alternatif.....	104
2.	Mécanisme de dérégulation de l'épissage dans les cancers <i>SF3B1</i> ^{MUT}	106
3.	Jonctions alternatives dans les SMD <i>SF3B1</i> ^{MUT}	111
4.	Épissage alternatif en aval des mutations des autres gènes du spliceosome	115
Partie 2	116
I.	La protéine SF3B1 et les processus de réplication et de réparation de l'ADN	117
A.	Stress réplicatif	117
1.	Causes du stress réplicatif.....	118
2.	Réponses cellulaires au stress réplicatif	120
a.	Voie de signalisation ATR/CHK1	121
b.	Voie de signalisation ATM/CHK2	124
3.	Stress réplicatif et vieillissement du système hématopoïétique.....	127
B.	Rôle de la protéine SF3B1 dans le contrôle de la réplication et de la réparation de l'ADN	131
II.	Mutations de <i>SF3B1</i> et dommages à l'ADN	135
Chapitre III - Résultats	141
Partie 1	142
I.	Problématique.....	143
II.	Objectif.....	143
III.	Résultats	144
Partie 2	190
I.	Problématique.....	191
II.	Objectif.....	191
III.	Résultats	192
Conclusion générale	233
Discussion et perspectives	235
Annexes	249
Bibliographie	260

Liste des abréviations

3'ss	Site accepteur d'épissage en 3'
5'ss	Site donneur d'épissage en 5'
53BP1	TP53-binding protein 1
A	Adénine
Acido-E	Erythroblastes acidophiles
ADN	Acide désoxyribonucléique
ADNc	ADN complémentaire
ADNg	ADN génomique
ADNsb	ADN simple-brin
AF	Anémie de Fanconi
AG	3'ss canonique (adénine-guanine)
AG'	3'ss alternatif (adénine-guanine)
AGM	Aorte-gonades-mésonephros
AKT	<i>Protein kinase B</i>
ALA	Acide δ -aminolévulinique
ALAS2	δ -aminolévulinate synthase 2
ARN	Acide ribonucléique
ARNm	Acide ribonucléique messager
ARNr	ARN ribosomique
ASE	Agents stimulants l'érythropoïèse
ATM	<i>Ataxia telangiectasia mutated</i>
ATR	<i>Ataxia telangiectasia and rad3-related protein</i>
ATRIP	<i>ATR-interacting protein</i>
Baso-E	Erythroblastes basophiles
BFU-E	<i>Burst forming unit-erythroid</i>
BMP	<i>Bone morphogenetic protein</i>
BrdU	Bromodésoxyuridine
BRCA1	<i>Breast cancer type 1 susceptibility protein</i>
BS	Site de branchement
BRE	<i>BMP-responsive elements</i>
CFU-E	<i>Colony forming unit-erythroid</i>
CD	Cluster de différenciation
CDC25	<i>Cell division cycle 25</i>
CDB	Cassures double-brin de l'ADN
CDK	Cycline-dépendante kinase
ChIP	Immunoprécipitation de la chromatine
CHX	Cycloheximide
c-KIT	Récepteur au SCF
CLIP-Seq	<i>UV-Cross-linking and RNA Immunoprecipitation-RNA sequencing</i>
CLP	Progéniteurs lymphoïdes communs
CMN	Cellules mononucléées
CMP	Progéniteurs myéloïdes communs
CRBN	Cereblon
CSH	Cellules souches hématopoïétiques

C-terminal	Extrémité carboxy-terminale d'une protéine
CV	Coefficient de variation de l'intensité moyenne de fluorescence
DMT1	<i>Divalent metal transporter 1</i>
DNase II	Désoxyribonucléase II
dNTPs	Désoxyribonucléotides
DRIP-qPCR	<i>DNA/RNA Immunoprecipitation-qPCR</i>
DT40-ASF	Lignée lymphoblastique (B) de poulet, invalidée pour le gène <i>ASF/SF2</i>
Edu	5-éthynyl-2'-désoxyuridine
Emp	<i>Erythroblast macrophage protein</i>
EPHB4	Récepteur B4 de l'éphrine
EPO	Erythropoïétine
EPOR	Récepteur de l'érythropoïétine
ERFE	Erythroferrone
ERFE ⁺¹²	Transcrit ERFE alternatif
ERFE ^{VPFQ}	Protéine ERFE alternative (valine-proline-phénylalanine-glutamine)
ERK	<i>Extracellular signal-regulated kinase</i>
Fc	Fragment cristallisable des immunoglobulines
Fe/S	Fer/Soufre
FECH	Ferrochélatase
G	Guanine
G4	G-quadruplex
GDF15	<i>Growth differentiation factor 15</i>
GDP	Guanosine diphosphate
γ-H2AX	Variant d'histone H2AX phosphorylé sur la sérine 139
GLRX5	Glutarédoxine-5
GM-CSF	<i>Granulocyte/macrophage colony-stimulating factor</i>
GMP	Progéniteurs granulo-monocytaires
GO	<i>Gene ontology</i>
GPA	Glycophorine A
Grb2	<i>Growth factor receptor-bound protein 2</i>
GTP	Guanosine triphosphate
HFE	<i>Homeostatic iron regulator</i>
HIF-1	<i>Hypoxia inducible factor-1</i>
HJV	Hémojuvéline
HRE	<i>Hypoxia response element</i>
HU	Hydroxyurée
IF	Immunofluorescence
IFN γ	Interféron γ
IL3	Interleukine 3
IL6	Interleukine 6
IL3R	Récepteur de l'interleukine 3
IPSS	<i>International prognostic scoring system</i>
IPSS-R	IPSS révisé
IRE	<i>Iron responsive element</i>

IRP	<i>Iron regulatory protein</i>
JAK2	<i>Janus kinase 2</i>
KI	<i>Knock-in</i>
KO	<i>Knock-out</i>
LAM	Leucémie aiguë myéloblastique
LLC	Leucémie lymphoïde chronique
LMPP	<i>Lymphoid-primed multipotent progenitors</i>
LNA	<i>Locked nucleic acid</i>
MAPK	<i>Mitogen-activated protein kinase</i>
MAPKK	<i>Mitogen-activated protein kinase kinase</i>
MCM	<i>Minichromosome maintenance</i>
MEP	Progéniteurs érythro-mégacaryocytaires
MFI	Intensité moyenne de fluorescence
MFRN1	Mitoferrine-1
MGG	May-Grünwald Giemsa
MLASA	Myopathie, acidose lactique et anémie sidéroblastique
MPP	Progéniteurs multipotents
mSCF	Forme membranaire du SCF
mTOR	<i>mammalian target of rapamycin</i>
MU	Mélanome de l'uvée
NHEJ	<i>Non-homologous end joining</i> , ligature d'extrémités non homologues
NMD	<i>Nonsense-mediated decay</i> , dégradation des ARNm non-sens
NTBI	<i>Non-transferrin bound iron</i> , fer non lié à la transferrine
N-terminal	Extrémité amino-terminale d'une protéine
O ₂	Dioxygène
OAS	Oligonucléotide anti-sens
OMS	Organisation mondiale de la santé
pb	Paire de bases
PCNA	Antigène nucléaire de prolifération cellulaire
PDK1	<i>Phosphoinositide-dependent kinase-1</i>
PEM	Progéniteurs érythro-myéloïdes
PHD	Prolyl-hydroxylase
PI3K	<i>Phosphatidylinositol-3-kinase</i>
PIP2	Phosphatidyl-inositol-diphosphate
PIP3	Phosphatidyl-inositol-triphosphate
Poly-E	Erythroblastes polychromatophiles
PPT	<i>Polypyrimidine tract</i>
Pré-ARNm	ARN pré-messager
Pro-E	Proérythroblastes
p-RPA32	Sous-unité RPA2 de RPA phosphorylée
PTEN	<i>Phosphatase and tensin homolog</i>
PUS1	Pseudo-uridylylate synthase 1

RFC	Facteur de réplication C
RH	Recombinaison homologue
RNA-Seq	Séquençage de l'ARN
RNR	Ribonucléotide réductase
ROS	<i>Reactive oxygen species</i> , espèces réactives à l'oxygène
RPA	Protéine de réplication A
RT-PCR	<i>Reverse Transcription-Polymerase Chain Reaction</i> , transcription inverse-réaction en chaîne par polymérase
RT-qPCR	RT-PCR quantitative
SC	Sidéroblastes en couronne
SCF	<i>Stem cell factor</i>
SF3B1	<i>Splicing factor 3B subunit 1</i>
<i>SF3B1</i> ^{MUT}	Cellules avec un gène <i>SF3B1</i> muté
<i>SF3B1</i> ^{WT}	Cellules avec un gène <i>SF3B1</i> non muté (<i>wild-type</i>)
SFC	Sites fragiles communs
SH2	<i>Src homology domain 2</i>
SH3	<i>Src homology domain 3</i>
shARN	<i>small hairpin</i> ARN
siARN	<i>small interfering</i> ARN
SMD	Syndromes myélodysplasiques
SM	Spectromètre de masse
SMD-SC	SMD avec sidéroblastes en couronne
snARN	Petit ARN nucléaire
snRNP	Petite particule nucléaire ribonucléoprotéique
Sos	<i>Son of sevenless</i>
SR	Domaines sérine-arginine
SRE	Système réticulo-endothélial
SSA	Spliceostatine A
sSCF	Forme soluble du SCF
STAT5	<i>Signal transducer and activator of transcription 5</i>
TFR1	Récepteur à la transferrine de type 1 (TFRC, CD71)
TFR2	Récepteur à la transferrine de type 2
TGFβ	<i>Transforming growth factor β</i>
TIPIN	<i>Timeless-interacting protein</i>
TopBP1	<i>DNA topoisomerase 2-binding protein 1</i>
TRAIL	<i>Tumor-necrosis-factor related apoptosis inducing ligand</i>
TSC2	<i>Tuberous sclerosis 2</i>
TWSG1	<i>Twisted gastrulation 1</i>
U2AF35	Facteur auxiliaire d'U2, sous-unité 1 (35 kDa)
U2AF65	Facteur auxiliaire d'U2, sous-unité 2 (65 kDa)
UV	Rayons ultraviolets
VGM	Volume globulaire moyen
VHL	Von Hippel-Lindau
XLSA	Anémie sidéroblastique liée à l'X
XLSA/A	XLSA avec ataxie

Lexique

Tout au long de cette thèse, les Pro-E et Baso-E sont considérés comme des précurseurs érythroïdes (érythroblastes) immatures, tandis que les Poly-E et Acido-E comme des précurseurs érythroïdes matures. La différenciation érythroïde terminale – ou maturation érythroïde terminale – correspond à la différenciation de la lignée rouge des stades Pro-E à Acido-E dans la moelle osseuse. Elle ne doit pas être confondue avec l'érythropoïèse définitive qui, chez les hommes, débute au cours de la vie fœtale et englobe l'ensemble des étapes de différenciation de la CSH jusqu'au globule rouge.

Liste des tableaux et figures

Tableaux

Tableau 1 : Principaux acteurs du métabolisme du fer	66
Tableau 2 : Principales mutations responsables du développement d'une anémie sidéroblastique congénitale	70
Tableau 3 : Classification diagnostique des SMD selon l'OMS	74
Tableau 4 : Évaluation pronostique des SMD selon l'IPSS-R.....	76

Figures

Figure 1 : Organisation hiérarchique classique de l'hématopoïèse humaine	21
Figure 2 : Organisation hiérarchique redéfinie de l'hématopoïèse humaine (d'après Notta et al., 2016).....	22
Figure 3 : Les différents sièges de l'érythropoïèse humaine au cours du développement embryonnaire et fœtal (d'après Nandakumar et al., 2016)	24
Figure 4 : L'érythropoïèse humaine	25
Figure 5 : Caractérisation et isolement des érythroblastes humains de la moelle osseuse par cytométrie en flux (d'après Hu et al., 2013).....	28
Figure 6 : Analyse par IF de la lamine B et de l'histone nucléaires pendant l'érythropoïèse murine terminale (Zhao et al., 2016).....	30
Figure 7 : Contrôle de l'expression du gène <i>EPO</i> par le facteur de transcription HIF-1	33
Figure 8 : Voies de signalisation activées par le couple EPO/EPOR.....	36
Figure 9 : Activation de la caspase-3 par dépolarisation de la membrane mitochondriale au cours de la maturation érythroïde terminale.....	41
Figure 10 : Régulation négative de l'érythropoïèse par les caspases	43
Figure 11 : Régulation de l'érythropoïèse par la protéine chaperonne HSP70 (d'après Vandekerckhove et al., 2008).....	43
Figure 12 : Dynamique de la rétention d'intron dans le transcrit SF3B1 au cours de l'érythropoïèse terminale (d'après Conboy, 2017)	48
Figure 13 : Métabolisme du fer	52

Figure 14 : Biosynthèse de l'hème	53
Figure 15 : Biosynthèse des protéines à clusters Fe/S	53
Figure 16 : L'îlot érythroblastique	54
Figure 17 : L'hepcidine régule l'homéostasie du fer	59
Figure 18 : Régulation de l'expression du gène <i>HAMP</i> dans les hépatocytes	64
Figure 19 : Mécanisme d'action de l'érythroferrone (d'après Kautz, 2014).....	65
Figure 20 : Sidéroblastes en couronne	68
Figure 21 : Mécanismes impliqués dans la dysérythropoïèse terminale des SMD	81
Figure 22 : Épissage des pré-ARNm par le spliceosome majeur	91
Figure 23 : Structure du complexe humain SF3b (d'après Cretu et al., 2016).....	93
Figure 24 : L'association de SF3B1 aux nucléosomes détermine l'épissage des pré-ARNm .	97
Figure 25 : Distribution des résidus mutés de la protéine SF3B1 dans les cancers (d'après Cretu et al., 2016).....	102
Figure 26 : Les différents types d'évènements d'épissage alternatif	105
Figure 27 : Modèle de dérégulation de l'épissage induit par les mutations <i>hotspot</i> de <i>SF3B1</i> (d'après Alsafadi et al., 2016).....	109
Figure 28 : Utilisation préférentielle d'un site de branchement alternatif dans les cellules <i>SF3B1</i> ^{MUT} et conséquences aux niveaux ARNm et protéines (d'après Darman et al., 2015)	110
Figure 29 : Utilisation préférentielle du site 3' cryptique d' <i>ABCB7</i> dans les cellules SMD <i>SF3B1</i> ^{MUT}	112
Figure 30 : Causes du stress répliatif (d'après Zeman et Cimprich, 2014)	120
Figure 31 : Réponses cellulaires aux dommages à l'ADN induits par le stress répliatif	126
Figure 32 : L'accumulation de foyers intranucléaires γ -H2AX est un marqueur de la perte de fonctionnalité des CSH murines âgées (d'après Flach et al., 2014)	128
Figure 33 : Les stress oxydatif et répliatif participent tous les deux à l'épuisement du système hématopoïétique de la souris (d'après Walter et al., 2015).....	130
Figure 34 : Rôle du complexe BRCA1/protéines d'épissage dans la maintenance de la stabilité génomique (d'après Savage et al., 2014)	133
Figure 35 : FANCI et FANCD2 contrôlent la dynamique de distribution de SF3B1 dans le noyau (d'après Moriel-Carretero et al., 2017)	134
Figure 36 : Le stress répliatif causé par les mutations des gènes <i>U2AF1</i> et <i>SRSF2</i> est responsable du défaut de prolifération des progéniteurs SMD (d'après Chen et al., 2018) ...	139
Figure 37 : Conséquences fonctionnelles des mutations du gène <i>SF3B1</i> dans les SMD-SC.	234

Préface

Les syndromes myélodysplasiques (SMD) constituent un ensemble hétérogène de pathologies clonales de la cellule souche hématopoïétique (CSH). Parmi les SMD, les SMD avec sidéroblastes en couronne (SMD-SC) se distinguent par une anémie macrocytaire profonde qui contraste avec une moelle osseuse riche à prédominance érythroblastique. Une partie des érythroblastes médullaires renferme des granules de fer précipités au sein des mitochondries : ce sont les sidéroblastes en couronne. Cette surcharge mitochondriale en fer s'accompagne d'une surcharge systémique en fer, et ce avant même le traitement de l'anémie par transfusions de culots globulaires. Environ 15% des patients atteints d'un SMD-SC évoluent vers une leucémie aigüe myéloblastique. Le chapitre I de ce manuscrit revient sur les processus d'érythropoïèse normale et pathologique.

Grâce aux techniques de séquençage d'exome, il est désormais établi que près de 90% des patients porteurs d'un SMD-SC ont une mutation somatique sur le gène *SF3B1*. *SF3B1* code pour une protéine du complexe ribonucléoprotéique U2 du spliceosome dont les fonctions sont multiples. En effet, en plus de participer à l'épissage des ARN pré-messagers, *SF3B1* contrôle – avec de nombreux partenaires protéiques – la réplication et la réparation de l'ADN. Le chapitre II de ce manuscrit explique les différents rôles connus de *SF3B1*. Le fait que les mutations de *SF3B1* dans les CSH soient quasiment toujours associées au développement d'une dysérythropoïèse avec SC suggère que la protéine *SF3B1* mutée pourrait être directement impliquée dans la pathogénèse des SMD-SC. L'objectif de ma thèse a donc consisté à étudier les conséquences fonctionnelles des mutations de *SF3B1* dans des cellules à différents stades de maturation érythrocytaire. J'ai dans un premier temps identifié un variant de l'érythroferrone (ERFE) formé par épissage aberrant uniquement dans les érythroblastes primaires issus de la moelle osseuse des patients atteints d'un SMD-SC avec mutation sur le gène *SF3B1*. ERFE est une hormone qui régule l'expression de l'hepcidine, chef d'orchestre de l'homéostasie du fer. Ce travail fait l'objet de la première partie des résultats. Dans un deuxième temps, j'ai caractérisé les processus de réplication et de réparation de l'ADN dans les lignées murines érythroïdes G1E-ER4 CRISPR-Cas9 *Sf3b1*^{WT} et *Sf3b1*^{K700E} établies au sein du laboratoire, et dans les érythroblastes primaires issus de la moelle osseuse des patients atteints d'un SMD-SC avec ou sans mutation sur le gène *SF3B1*. Ce travail fait l'objet de la deuxième partie des résultats. Une conclusion générale ainsi qu'une discussion et les perspectives d'expériences à mettre en place clôturent ce manuscrit.

Chapitre I – Érythropoïèse normale et pathologique

I. L'érythropoïèse humaine normale

Les syndromes myélodysplasiques avec sidéroblastes en couronne (SMD-SC) se caractérisent par une érythropoïèse inefficace. L'érythropoïèse est le processus par lequel l'organisme assure la production des **globules rouges**, également appelés érythrocytes ou hématies, à partir des **cellules souches hématopoïétiques** (CSH). D'une durée de vie de 120 jours, les érythrocytes garantissent le **transport de l'oxygène** des poumons aux tissus par l'intermédiaire de l'hémoglobine, protéine contenant du **fer**. Chez un adulte sain, près de 200 milliards de globules rouges sont produits chaque jour dans la moelle osseuse, ce qui requiert 20 à 25 mg de fer. En cas de besoins accrus, l'érythropoïèse peut s'adapter et produire 7 à 8 fois plus d'érythrocytes. La régulation fine de l'homéostasie du fer s'avère ainsi nécessaire à celle de l'érythropoïèse, et vice-versa.

Cette première partie présentera l'érythropoïèse physiologique et son interaction forte avec le métabolisme du fer, de manière à aborder, dans une deuxième partie, une situation pathologique particulière associant dysérythropoïèse et surcharge en fer : les syndromes myélodysplasiques avec sidéroblastes en couronne.

A. Généralités

1. L'hématopoïèse humaine

L'hématopoïèse est la production et la mise en circulation des éléments figurés du sang à partir des CSH multipotentes. Chez l'homme, ces dernières se trouvent majoritairement dans la **moelle osseuse** rouge, tissu spongieux présent au centre des os plats tels que le sternum et les os du bassin. Elles sont quantitativement peu nombreuses, car elles ne représentent environ que 0,05% des cellules de la moelle osseuse. Parallèlement à un auto-renouvellement constant qui permet le maintien d'une hématopoïèse fonctionnelle tout au long de la vie des individus, les CSH sont capables de se différencier en globules rouges, globules blancs ou plaquettes.

Lorsqu'elles ne sont pas quiescentes, les CSH donnent tout d'abord naissance à des progéniteurs multipotents (**MPP**) caractérisés par une réduction de leur potentiel d'auto-renouvellement. Puis, les MPP s'orientent soit vers la myélopoïèse en produisant des progéniteurs myéloïdes communs (**CMP**), soit vers la lymphopoïèse en produisant des progéniteurs lymphoïdes communs (**CLP**). Dans la lignée myéloïde, les CMP oligopotents se

différencient ensuite en progéniteurs bipotents granulo-monocytaires (**GMP**), à l'origine de la formation des **granulocytes** et **monocytes**, ou en progéniteurs bipotents érythro-mégacaryocytaires (**MEP**), à l'origine de la formation des **plaquettes** et **érythrocytes**. L'ensemble de ces différentes étapes de différenciation se déroule au sein de la moelle osseuse, puis les cellules matures sont libérées dans la circulation sanguine pour assurer leurs fonctions respectives (Figure 1, flèches traits pleins). Des voies alternatives à cette hématopoïèse classique sont possibles, telle que la différenciation des CSH directement en CMP ou MEP. De plus, des progéniteurs mixtes lympho-myéloïdes (LMPP pour *Lymphoid-primed multipotent progenitor*) dérivant des CSH ont été identifiés chez la souris (Adolfsson et al., 2005). Ces LMPP ont perdu leur potentiel de différenciation érythro-mégacaryocytaire (Figure 1, flèches en pointillés).

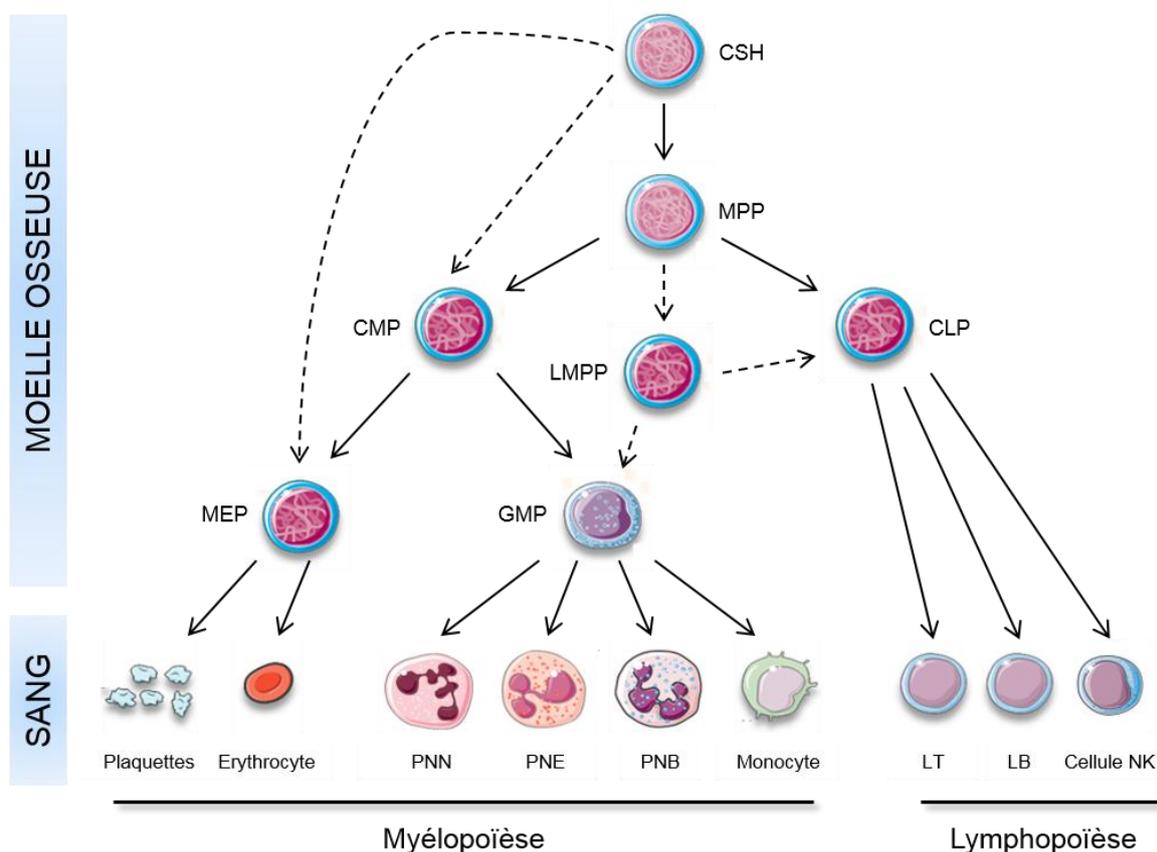


Figure 1 : Organisation hiérarchique classique de l'hématopoïèse humaine

Chez l'homme, les cellules sanguines proviennent de la maturation progressive des CSH au sein de la moelle osseuse. Il existe deux grandes voies de différenciation : la lymphopoïèse et la myélopoïèse. Les CSH peuvent parfois se différencier directement en CMP ou MEP, ou bien produire des LMPP. **PNN** = polynucléaires neutrophiles, **PNE** = polynucléaires éosinophiles, **PNB** = polynucléaires basophiles, **LT** = lymphocytes T, **LB** = lymphocytes B, **NK** = *Natural Killer*.

Depuis quelques années, de nombreuses études ont revisité l'organisation hiérarchique de l'hématopoïèse adulte, en insistant notamment sur le fait que le compartiment des CSH contiendrait des **progéniteurs myéloïdes unipotents** capables de s'auto-renouveler. En effet, en août 2013, il a été montré que les MEP des souris pouvaient directement dériver des CSH (Yamamoto et al., 2013). Trois années plus tard, l'équipe de John Dick a analysé le potentiel de différenciation en granulocytes, érythrocytes ou mégacaryocytes dans près de 3000 CSH humaines issues du foie fœtal, du sang de cordon ou de la moelle osseuse adulte (Notta et al., 2016) (Figure 2). La combinaison des techniques de phénotypage cellulaire par cytométrie en flux et de séquençage de l'ARN (RNA-Seq) à l'échelle unicellulaire a permis de confirmer l'hypothèse de Yamamoto *et al.* chez l'homme, puisque les MEP dérivent majoritairement des CSH et non des CMP, et ce quel que soit l'âge de l'individu. Ainsi, la spécificité des lignées érythroïde et mégacaryocytaire serait très précoce dans l'hématopoïèse, ce qui tend à dissocier la capacité d'auto-renouvellement d'une cellule de sa multipotence.

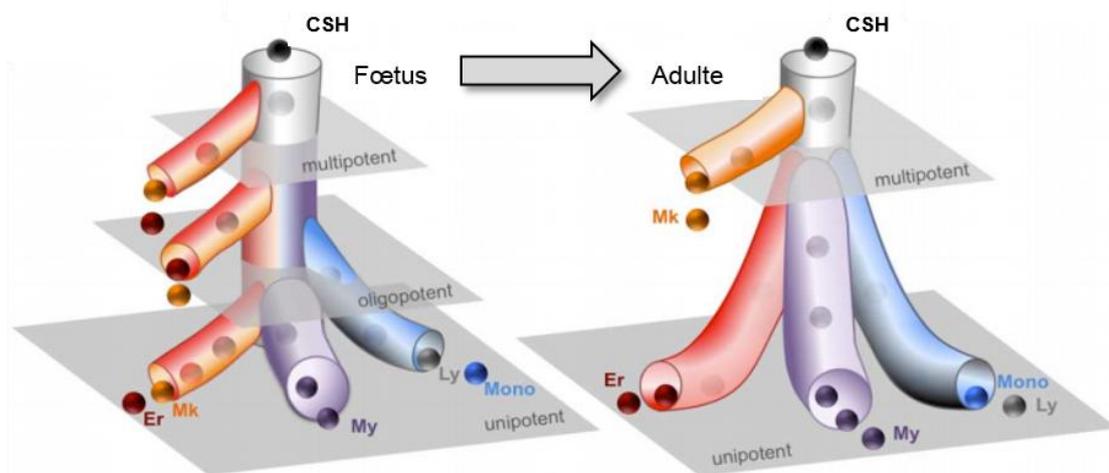


Figure 2 : Organisation hiérarchique redéfinie de l'hématopoïèse humaine (d'après Notta et al., 2016)

D'après les dernières études menées par Notta *et al.*, les progéniteurs hématopoïétiques perdent progressivement leur multipotence au cours de la vie d'un individu. L'hématopoïèse adulte est ainsi composée de deux compartiments cellulaires : les CSH multipotentes et les progéniteurs unipotents donnant naissance aux cellules matures. Les MEP dérivent quant à eux majoritairement des CSH. Cette connexion physiologique étroite entre les CSH et le lignage érythro-mégacaryocytaire peut expliquer la déplétion plus importante en MEP par rapport aux GMP lors d'une anémie aplasique. **Mk** = mégacaryocytes, **Er** = érythrocytes, **My** = granulocytes, **Ly** = lymphocytes, **Mono** = monocytes.

L'ensemble du processus d'hématopoïèse est finement organisé et régulé par de multiples facteurs intrinsèques et extrinsèques. Parmi les facteurs extrinsèques solubles, une grande partie est produite localement au sein d'une structure définie : la **niche hématopoïétique**. Cette dernière est constituée de divers types cellulaires, tels que les macrophages et les cellules stromales mésenchymateuses qui produisent des adipocytes, ostéoblastes, fibroblastes et chondrocytes. La niche hématopoïétique offre ainsi un microenvironnement adapté aux CSH, et chaque cytokine sécrétée peut orienter leur différenciation vers un lignage cellulaire particulier.

La suite des données présentées dans cette thèse s'intéressera à une branche spécifique de l'hématopoïèse humaine : l'**érythropoïèse**.

2. Ontogénie de l'érythropoïèse

Le siège de l'érythropoïèse se déplace au cours du développement de l'embryon et du fœtus (Figure 3). L'érythropoïèse a tout d'abord lieu dans des tissus extra-embryonnaires : les îlots sanguins du sac vitellin (Baron et al., 2012). Ces îlots renferment de grands progéniteurs érythroïdes nucléés, mais également des progéniteurs plaquettaires diploïdes et des macrophages (Potts et al., 2014). Chez l'homme, cette première vague d'érythropoïèse, appelée **érythropoïèse primitive**, commence entre les 18^{ème} et 20^{ème} jours de gestation et dure 3 semaines (Palis, 2014). Elle est nécessaire à la survie et la croissance de l'embryon. Puis, l'**érythropoïèse définitive** se met en place en deux étapes. Dans un premier temps, des progéniteurs multipotents érythro-myéloïdes (PEM) émergent dans le sac vitellin et colonisent le foie fœtal (Palis et al., 2001). Dans un deuxième temps, des CSH se forment au niveau de l'aorte dorsale de la région aorto-gonado-mésonephros (AGM), et au niveau d'autres sites tels que le placenta et l'artère ombilicale (Yoder, 2014). Ces CSH, capables de s'auto-renouveler et de donner naissance à l'ensemble des cellules sanguines dont les globules rouges, migrent rapidement vers le foie fœtal où elles s'expandent avant de rejoindre la moelle osseuse à partir des deuxième et troisième trimestres de grossesse. Contrairement aux érythrocytes primitifs qui peuvent conserver leur noyau dans la circulation sanguine, les érythrocytes définitifs sont anucléés. En outre, l'expression des gènes de la globine varie au cours du développement embryonnaire et fœtal, ce qui contribue à différencier les deux types d'érythrocytes (globines α , ζ , ϵ et γ dans les érythrocytes primitifs ; globines α , β , γ et δ dans les érythrocytes définitifs). À la naissance et durant toute la vie post-natale, l'érythropoïèse a lieu uniquement dans la moelle osseuse.

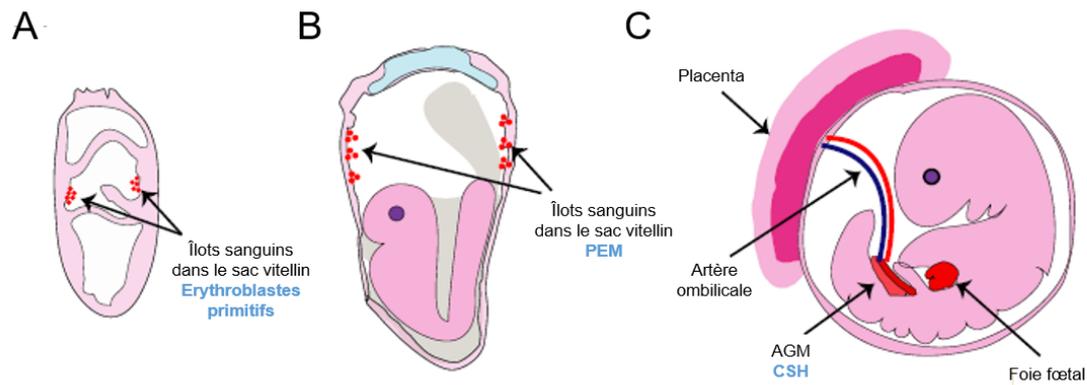


Figure 3 : Les différents sièges de l'érythropoïèse humaine au cours du développement embryonnaire et fœtal (d'après Nandakumar et al., 2016)

(A) Au début du développement embryonnaire, l'érythropoïèse siège dans les îlots sanguins du sac vitellin. Ces îlots sanguins renferment des érythroblastes primitifs mégalo-blastiques. (B) Puis, lors d'une deuxième vague d'érythropoïèse, des progéniteurs multipotents érythro-myéloïdes (PEM) apparaissent. (C) Enfin, les cellules souches hématopoïétiques (CSH) émergent au niveau de la région aorto-gonado-mésonephros (AGM). Tout comme les PEM, elles migrent vers le foie fœtal et donnent naissance à des érythrocytes définitifs anucléés.

B. Différenciation érythroïde

Chez l'homme, l'érythropoïèse médullaire se déroule en 8 jours et se compose de deux phases successives : une **phase précoce de prolifération des progéniteurs érythroïdes** et une **phase terminale de différenciation des précurseurs érythroïdes**. Durant la phase précoce, les progéniteurs peuvent s'expandre de façon conséquente, contrairement à la phase terminale où les précurseurs prolifèrent moins mais entrent dans un processus de maturation progressive. Ce passage de l'auto-renouvellement des progéniteurs à la différenciation érythrocytaire terminale des précurseurs est finement régulé, puisqu'il détermine *in fine* la production érythroïde. À la fin de l'érythropoïèse, les globules rouges matures anucléés sont libérés dans la circulation sanguine. Les progéniteurs morphologiquement identiques sont différenciés grâce au type de colonies qu'ils génèrent dans des systèmes de culture sur milieu semi-solide (méthylcellulose). En revanche, chacun des précurseurs érythroïdes a sa propre morphologie observable au microscope après coloration May-Grünwald Giemsa (MGG). De plus, durant l'ensemble de l'érythropoïèse, l'expression de certains marqueurs de surface varie. L'immuno-marquage avec analyse au cytomètre en flux permet ainsi de distinguer chaque stade de l'érythropoïèse, du progéniteur au globule rouge (Figure 4).

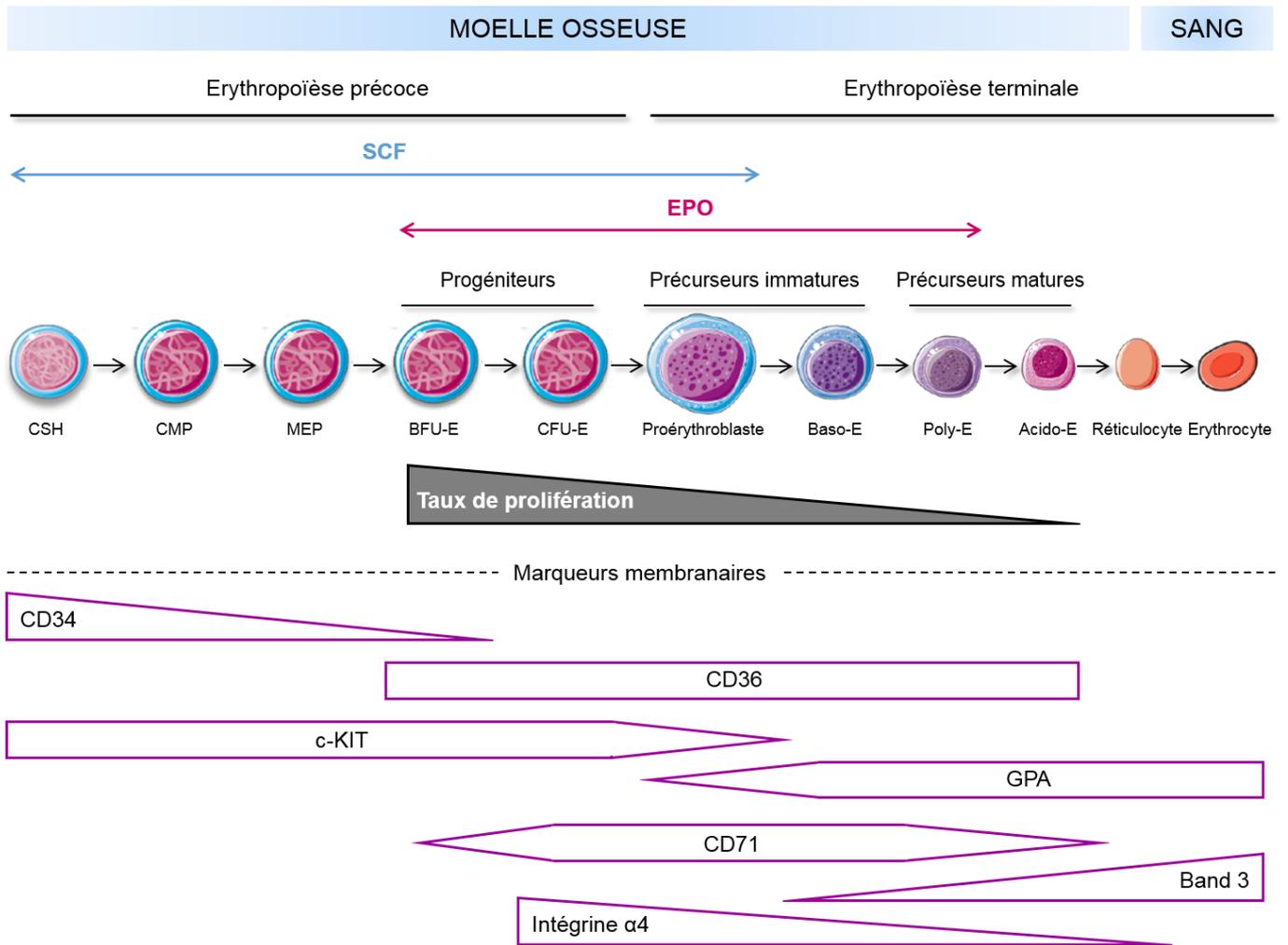


Figure 4 : L'érythropoïèse humaine

L'érythropoïèse humaine est un processus continu qui se divise en deux étapes : l'érythropoïèse précoce puis l'érythropoïèse terminale. L'ensemble de ce processus a lieu au sein de la moelle osseuse, puis les érythrocytes sont libérés dans la circulation sanguine pour assurer leur fonction. Deux principaux facteurs de croissance régulent la différenciation érythrocytaire : le SCF et l'EPO. Le taux de prolifération diminue constamment du stade BFU-E au stade Acido-E. L'expression des marqueurs membranaires varie au cours de l'érythropoïèse et leur analyse par immunophénotypage permet de distinguer chacun des progéniteurs et précurseurs érythroïdes. **CSH** = cellule souche hématopoïétique, **CMP** = progéniteur commun myéloïde, **MEP** = progéniteur érythro-mégacaryocytaire, **BFU-E** = *Burst forming unit-erythroid*, **CFU-E** = *Colony forming unit-erythroid*, **Baso-E** = érythroblaste basophile, **Poly-E** = érythroblaste polychromatophile, **Acido-E** = érythroblaste acidophile. D'un point de vue lexical, les BFU-E et CFUE sont des progéniteurs érythroïdes, les proérythroblastes et Baso-E des précurseurs érythroïdes immatures et les Poly-E et Acido-E des précurseurs érythroïdes matures.

1. L'érythropoïèse précoce

Durant l'érythropoïèse précoce, les progéniteurs myéloïdes érythro-mégacaryocytaires s'engagent définitivement vers le lignage érythroïde en se différenciant en **BFU-E** (pour *Burst forming unit-erythroid*), puis en **CFU-E** (pour *Colony forming unit-erythroid*).

Après 14 jours de culture *in vitro* en méthylcellulose, les BFU-E humaines forment de grosses colonies observables en microscopie optique. Chaque colonie contient plusieurs centaines d'érythroblastes hémoglobinisés. Les CFU-E donnent des colonies d'érythroblastes plus petites (< 100 cellules) après 7 jours de culture en milieu semi-solide.

Plusieurs facteurs de croissance sont nécessaires à la survie, la prolifération et la différenciation des progéniteurs érythroïdes, comme le *stem cell factor* (SCF), le *granulocyte/macrophage colony-stimulating factor* (GM-CSF) et l'interleukine 3 (IL3). Les BFU-E expriment peu de récepteurs de l'érythropoïétine (EPOR), contrairement aux CFU-E dont la survie est hautement dépendante de l'érythropoïétine (EPO).

Outre les études de clonogénicité, l'immunophénotypage cellulaire permet aujourd'hui de distinguer correctement les BFU-E des CFU-E. En 2014, l'expression de différents marqueurs de surface a en effet été étudiée dans les progéniteurs humains érythroïdes, dont les clusters de différenciation (CD) 45, 34 et 36 (Li et al., 2014). Le CD45 et le CD34 ne sont pas spécifiques de la lignée érythroïde, puisqu'ils s'expriment respectivement au niveau de la membrane plasmique des leucocytes et des CSH. Le CD235a ou glycophorine A (GPA), le récepteur à la transferrine CD71 et le récepteur de l'IL3 (IL3R) ont aussi été analysés. D'après cette étude de caractérisation immunophénotypique, les BFU-E sont CD45⁺, GPA⁻, IL3R⁻, CD34⁺, CD36⁻ et CD71^{low} ; tandis que les CFU-E sont CD45⁺, GPA⁻, IL3R⁻, CD34⁻, CD36⁺ et CD71^{high}. Ainsi, la transition des BFU-E vers les CFU-E s'accompagne d'une perte du CD34, d'une acquisition du CD36 et d'une augmentation de l'expression du CD71 à la surface des cellules érythroïdes (Figure 4). Les cellules exprimant l'IL3R correspondent quant à elles aux progéniteurs granulo-monocytaires.

Les CFU-E s'engagent ensuite dans la différenciation érythrocytaire terminale qui aboutit à la formation des cellules érythroïdes fonctionnelles : les globules rouges.

2. L'érythropoïèse terminale

Au cours de l'érythropoïèse terminale, les **proérythroblastes** (Pro-E) donnent tout d'abord naissance aux **érythroblastes basophiles** (Baso-E) précoces puis tardifs, qui se différencient ensuite en **érythroblastes polychromatophiles** (Poly-E) et en **érythroblastes acidophiles** (Acido-E) (Figure 4). L'ensemble des étapes de la maturation terminale a lieu dans les îlots érythroblastiques, structures qui seront décrites dans le Chapitre I-I-E-3. Les précurseurs érythroïdes immatures Pro-E et Baso-E prolifèrent rapidement. À partir du stade Poly-E, les érythroblastes continuent leur maturation mais la prolifération ralentit de façon drastique. La survie et la prolifération des précurseurs érythroïdes dépendent quasi-uniquement de l'EPO, principal régulateur de l'érythropoïèse.

a. Caractéristiques morphologiques et immunophénotypiques des précurseurs érythroïdes

Contrairement aux BFU-E et CFU-E, les précurseurs érythroïdes sont facilement reconnaissables après une coloration MGG, du fait des modifications morphologiques qu'ils subissent progressivement (Figures 4 et 5). Parmi ces modifications morphologiques, on observe une diminution de la taille des cellules et des noyaux, une condensation de la chromatine et une perte de la basophilie du cytoplasme due à l'augmentation de la synthèse de l'hémoglobine intracellulaire. Ainsi, au microscope, les proérythroblastes apparaissent comme de grosses cellules arrondies (20-30 µm de diamètre) avec un rapport nucléo-cytoplasmique élevé et un cytoplasme basophile bleu riche en acide ribonucléique (ARN), tandis que les érythroblastes acidophiles sont de petites cellules (10 µm) renfermant une chromatine très condensée.

En parallèle de ces changements morphologiques, l'analyse, par cytométrie en flux, de l'expression de certaines protéines de surface permet d'identifier les différents stades de la maturation érythroïde terminale. Le CD36, récepteur des acides gras présent au niveau de la membrane plasmique des CFU-E, continue à s'exprimer jusqu'au stade Acido-E. À partir du stade Pro-E, l'expression du CD45 et du récepteur du SCF (c-KIT ou CD117) disparaît. Les cellules ne sont alors plus dépendantes du SCF pour leur survie et leur différenciation. À l'inverse, les proérythroblastes acquièrent petit-à-petit l'antigène GPA, marqueur spécifique de la lignée érythroïde codé par le gène *GYP A*. Le récepteur de la transferrine (CD71) – permettant l'entrée du fer nécessaire à la synthèse de l'hème dans la cellule – voit son expression diminuer à partir du stade Poly-E, jusqu'à devenir nulle au stade réticulocyte

(Figure 4). La distinction immunophénotypique des derniers stades de l'érythropoïèse terminale a nettement été améliorée suite aux travaux de l'équipe de Mohandas Narla en 2013. Cette dernière a en effet montré une acquisition progressive du transporteur membranaire échangeur d'anions Band 3 à partir du stade Baso-E, associée à la perte de l'intégrine $\alpha 4$ ou CD49d à partir du stade Poly-E (Hu et al., 2013) (Figures 4 et 5).

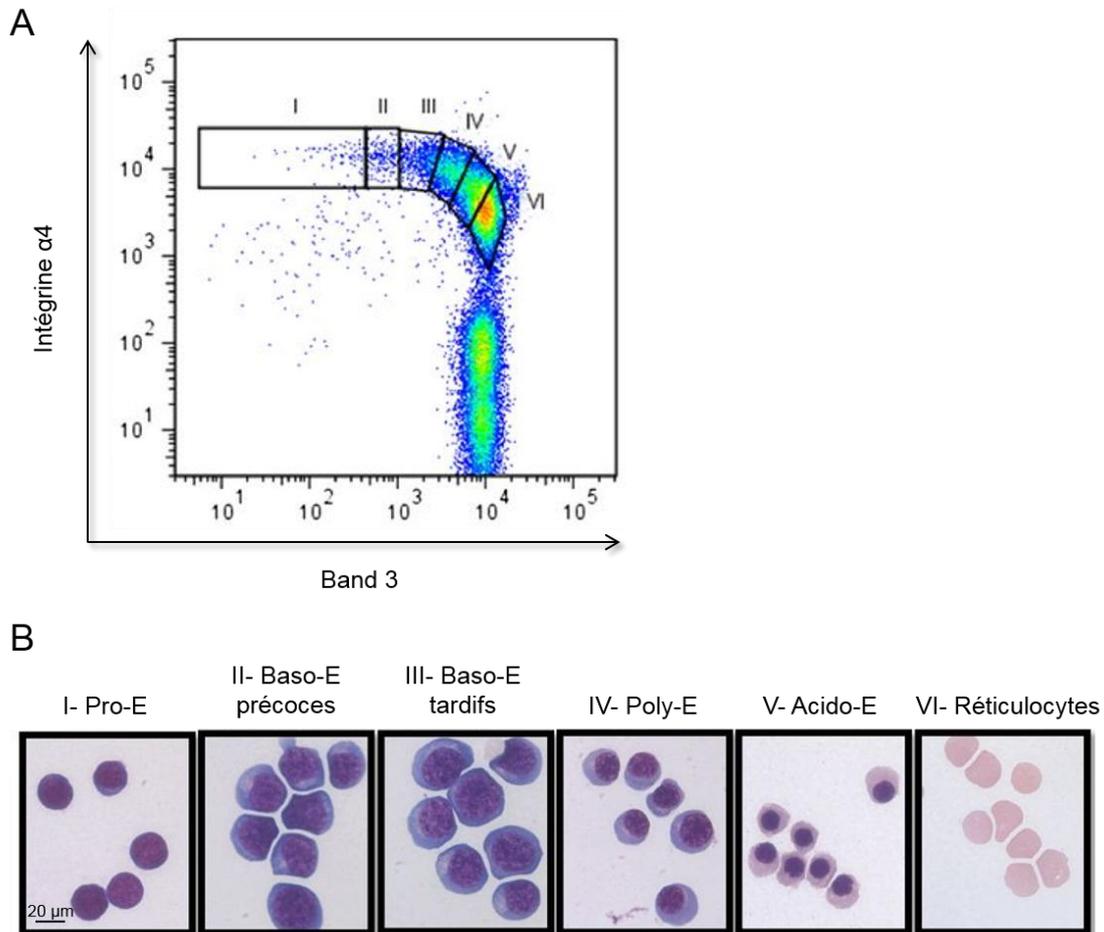


Figure 5 : Caractérisation et isolement des érythroblastes humains de la moelle osseuse par cytométrie en flux (d'après Hu et al., 2013)

(A) Les cellules humaines CD45⁻ ont été isolées à partir de la moelle osseuse de donneurs sains. Puis, l'analyse de l'expression des marqueurs membranaires Band 3 et intégrine $\alpha 4$ à la surface des cellules érythroïdes GPA⁺ a permis d'isoler six populations d'érythroblastes. (B) Après tri au cytomètre en flux, l'étude cytologique – par coloration MGG – a confirmé la pureté des populations isolées. **Pro-E** = proérythroblastes, **Baso-E** = érythroblastes basophiles, **Poly-E** = érythroblastes polychromatophiles, **Acido-E** = érythroblastes acidophiles. Échelle : 20 μ m.

b. Mécanisme de condensation de la chromatine

Comme dit précédemment, la chromatine des érythroblastes se condense progressivement au cours de l'érythropoïèse terminale. En 2016, Zhao *et al.* révèlent la formation d'un ou plusieurs pores dans l'enveloppe nucléaire des érythroblastes murins issus de foie fœtaux et cultivés *in vitro* en conditions de différenciation érythroïde, et ce dès le stade Pro-E (Zhao et al., 2016). L'analyse par immunofluorescence (IF) montre effectivement une rupture de la **lamine B**, protéine fibreuse qui participe à la structure de la lamina nucléaire (Figure 6). La lamina nucléaire constitue un maillage fibrillaire dense bordant la membrane interne des noyaux. Le pore créé par clivage de la lamine B est transitoire et dynamique : il s'ouvre lors de la phase G0/G1 du cycle cellulaire puis se referme, et ainsi de suite. Il s'accompagne d'une libération massive des **histones** H2A (Figure 6), H2AX, H1, H2B, H3 et H4 vers le cytoplasme où elles sont rapidement dégradées par le protéasome. Le mécanisme de sortie des histones n'est pas encore élucidé. Les exportines assurent normalement l'export des macromolécules à travers les pores nucléaires classiques. L'inhibition des exportines par la leptomycine B n'affectant pas la sortie des histones lors de l'ouverture transitoire du noyau, le processus de différenciation érythroïde terminale semble donc indépendant des exportines. De plus, l'histone H2AZ – tout comme d'autres protéines nucléaires (GATA1, Ezh2, HDAC2) – reste dans le noyau lorsque la lamine B est clivée, ce qui suggère un mécanisme sélectif d'export des histones. Pour rappel, les histones entrent dans la composition des nucléosomes, premier niveau de compaction de l'acide désoxyribonucléique (ADN). Il semblerait donc que l'histone H2AZ ait un rôle clé dans le remodelage de la chromatine au cours de l'érythropoïèse terminale, contrairement aux autres histones qui sont éliminées de façon à ne pas encombrer le noyau de plus en plus petit. La condensation de la chromatine lors de la maturation des érythroblastes implique également diverses modifications post-traductionnelles des histones, telles que leur désacétylation par l'histone désacétylase HDAC2 (Ji et al., 2010).

À la fin de l'érythropoïèse humaine terminale, les érythroblastes acidophiles expulsent leur noyau et deviennent alors des **réticulocytes** qui migrent dans la circulation sanguine (Figure 4). Riches en ARN messager (ARNm), les réticulocytes contiennent encore des organites comme les mitochondries, les ribosomes, le réticulum endoplasmique et l'appareil de Golgi. Les noyaux, entourés d'une petite quantité de cytoplasme et de membrane plasmique, forment quant à eux les pyrénocytes. Ces pyrénocytes possèdent une chromatine condensée au maximum. Ils sont rapidement phagocytés par les macrophages médullaires de

l'îlot érythroblastique qui reconnaissent la phosphatidylsérine, signal d'apoptose cellulaire exposée sur leur versant extracellulaire (Yoshida et al., 2005).

Enfin, deux à trois jours après l'énucléation, les ARNm contenus dans les réticulocytes sont dégradés et les organites restants éliminés par autophagie. Les réticulocytes se transforment alors en érythrocytes matures. Les **globules rouges** constituent plus de 80% des cellules du corps humain (Sender et al., 2016). Ils ont une forme de disque biconcave avec un diamètre de 8 μm environ. Ils assurent le transport de l'oxygène des poumons vers tous les tissus de l'organisme grâce à l'hémoglobine qui représente le tiers de leur masse. Ils possèdent une forte capacité de déformabilité et peuvent ainsi s'adapter aux contraintes de la microcirculation sanguine. Chez les mammifères, l'énucléation est un processus indispensable au développement des érythrocytes puisqu'elle permet *in fine* de : 1) libérer de l'espace pour l'hémoglobine, 2) augmenter leur flexibilité. Les érythrocytes vieillissants perdent leur flexibilité et ne peuvent donc plus passer dans les capillaires les plus fins. Ils sont éliminés par la rate.

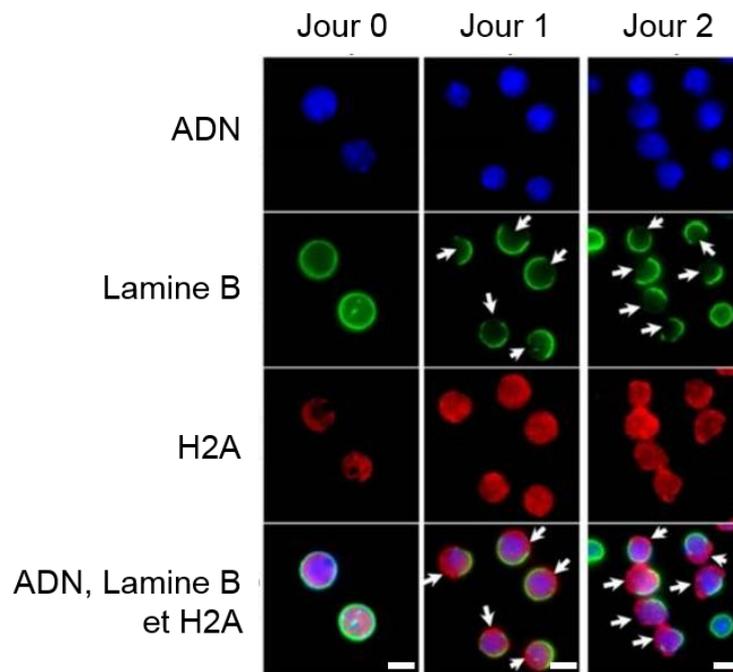


Figure 6 : Analyse par IF de la lamine B et de l'histone nucléaires pendant l'érythropoïèse murine terminale (Zhao et al., 2016)

Érythroblastes murins Ter119⁻ (équivalent de la GPA chez la souris) issus des foies fœtaux (souris C57BL/6, jour E13.5) et cultivés *in vitro* en conditions de différenciation érythroïde. Le clivage de la lamine B (flèches blanches) permet la sortie de l'histone H2A du noyau vers le cytosol. Échelle : 5 μm .

C. Régulation de l'érythropoïèse

L'érythropoïèse est finement régulée de manière à répondre aux besoins en oxygène de l'ensemble des tissus de l'organisme. Parmi les molécules impliquées dans cette régulation, on distingue trois sous-groupes : 1) les facteurs de croissance, 2) les facteurs de transcription, 3) les caspases.

1. Régulation par les facteurs de croissance

De nombreux facteurs de croissance assurent la survie et la prolifération des érythroblastes, les deux principaux étant le *stem cell factor* (SCF) et l'**érythropoïétine** (EPO). Le SCF contrôle la phase précoce de l'érythropoïèse (des CSH aux Pro-E), tandis que l'EPO intervient davantage dans la régulation de la maturation érythroïde terminale (des CFU-E aux Poly-E) (Figure 4).

a. Le SCF et son récepteur c-KIT

Le *stem cell factor* est une cytokine hématopoïétique qui, *via* sa liaison à son récepteur transmembranaire (c-KIT ou CD117) à activité tyrosine kinase, induit des signaux pour **la survie et la prolifération des progéniteurs érythroïdes BFU-E et CFU-E** (Ogawa et al., 1991).

Produit par les cellules stromales de la moelle osseuse (Heinrich et al., 1993), le SCF humain existe sous deux formes : une forme soluble (sSCF) et une forme transmembranaire (mSCF) (Broudy, 1997). Ces deux isoformes résultent d'un épissage alternatif au niveau de l'exon 6 de l'ARNm, exon codant notamment pour un site de clivage protéolytique. Le transcrit du mSCF perd l'exon 6 lors de l'épissage alternatif, il conserve donc son domaine transmembranaire situé en aval de l'exon 6, et génère une protéine de 220 acides aminés. En absence d'épissage alternatif, l'activation du site de clivage protéolytique rapidement après la traduction de la protéine entraîne la production du sSCF de 165 acides aminés (Flanagan et al.). Les souris knock-out (KO) pour *c-Kit* ou son ligand ne sont pas viables.

L'expression du c-KIT humain n'est pas spécifique de la lignée érythroïde puisqu'il est également retrouvé à la surface des progéniteurs hématopoïétiques myéloïdes, des mastocytes, des polynucléaires éosinophiles et des cellules *Natural Killer*. Au cours de l'érythropoïèse, c-KIT s'exprime fortement au niveau de la membrane des CSH, puis son expression diminue après le stade CFU-E jusqu'à devenir nulle après le stade Pro-E.

Lorsque le SCF se fixe à son récepteur, ce dernier s'active : il se dimérise et change de conformation, ce qui conduit à l'autophosphorylation de ses résidus tyrosine présents dans le domaine intracellulaire (Finger et al., 2009). Les tyrosines phosphorylées sont ensuite reconnues par des protéines à domaines de liaison SH2 (pour *Src homology domain 2*), et il s'ensuit l'activation de plusieurs voies de signalisation dont les voies **PI3K/AKT** et **MAPK/ERK**.

b. L'EPO et son récepteur EPOR

Comme son nom l'indique, l'érythropoïétine est l'hormone spécifique de l'érythropoïèse. Quand elle se lie à son récepteur EPOR, l'EPO active des voies de signalisation indispensables à **la survie et la prolifération des précurseurs érythroïdes**, favorisant ainsi leur différenciation.

L'EPO est une glycoprotéine de 34 kDa sécrétée par les cellules interstitielles péri-tubulaires du rein (Lacombe et al., 1988; Maxwell et al., 1990), en réponse aux variations de la pression partielle tissulaire en dioxygène (O₂). En effet, l'expression du gène codant pour l'EPO est sous le contrôle d'une région régulatrice de type HRE (pour *Hypoxia Response Element*) reconnue principalement par le facteur de transcription HIF-1 (pour *Hypoxia Inducible Factor-1*) qui est composé de deux sous-unités α et β (Pugh et al., 1991; Wang et al., 1995). Lorsque la pression partielle en O₂ diminue, la sous-unité HIF-1 α est stabilisée et transloquée dans le noyau où elle forme un hétérodimère avec la sous-unité HIF-1 β . Le dimère HIF-1 α /HIF-1 β interagit alors avec la région HRE, ce qui active la transcription de l'ARNm de l'EPO. Il s'ensuit une augmentation du taux d'EPO circulante, permettant ainsi de palier à l'hypoxie tissulaire par stimulation de l'érythropoïèse et, *in fine*, une augmentation de la production des érythrocytes. En cas de normoxie tissulaire, la sous-unité HIF-1 α est hydroxylée par des prolyl-hydroxylases (PHD) puis ubiquitinylée par le complexe VHL (pour Von Hippel-Lindau), une ubiquitine ligase E3. Elle est ultérieurement reconnue et dégradée par le protéasome, ce qui empêche la transcription du gène codant pour l'EPO (Figure 7).

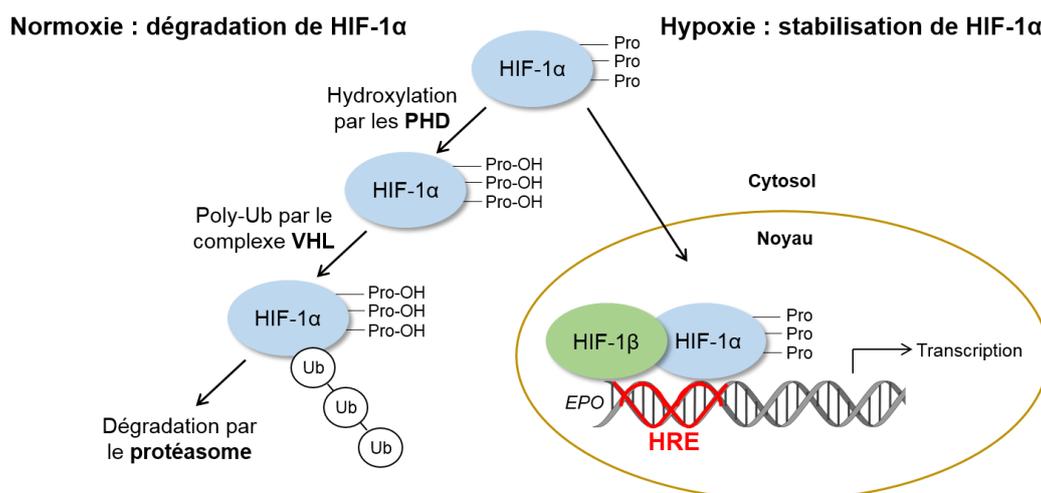


Figure 7 : Contrôle de l'expression du gène *EPO* par le facteur de transcription HIF-1

En situation d'hypoxie, HIF-1 α est stabilisé et transloqué dans le noyau où il régule, avec HIF-1 β , la transcription du gène de l'*EPO* en se fixant au niveau des régions régulatrices HRE. En situation de normoxie, HIF-1 α est dégradé par le protéasome après hydroxylation par les PHD et poly-ubiquitinylation par le complexe VHL. **Pro** = proline, **Pro-OH** = proline hydroxylée, **PHD** = prolyl-hydroxylases, **Poly-Ub** = poly-ubiquitinylation, **Ub** = ubiquitine.

Le récepteur à l'érythropoïétine se situe à la surface des érythroblastes (Broudy et al., 1991). Il est retrouvé dès le stade BFU-E, son expression étant maximale dans les CFU-E et Pro-E puis diminuant progressivement au cours de la différenciation érythrocytaire terminale (Figure 4). L'inactivation du gène *Epo* ou *EpoR* chez les souris est létale au stade embryonnaire (Wu et al., 1995). En effet, l'analyse histologique du foie fœtal des souris *Epo*^{-/-} ou *EpoR*^{-/-} révèle un défaut majeur de l'érythropoïèse définitive responsable du développement d'une anémie sévère. De plus, les CFU-E meurent par apoptose, ce qui souligne le rôle crucial de la signalisation EPO/EPOR dans la maturation érythroïde terminale.

L'EPOR est une glycoprotéine transmembranaire qui appartient à la famille des récepteurs de cytokines hématopoïétiques de classe I. Présent à la surface des cellules érythroïdes sous forme d'homodimère, il ne possède pas d'activité enzymatique intrinsèque, mais transduit un signal intracellulaire grâce à son interaction constitutive avec les tyrosines kinases JAK2 (pour *Janus kinases 2*) (Neubauer et al., 1998). Lorsque l'EPO se fixe à son récepteur, celui-ci change de conformation, ce qui entraîne le rapprochement et l'activation

par transphosphorylation des deux protéines JAK2 fixées sur le domaine intracellulaire d'EPOR (Remy et al., 1999). Les kinases JAK2 phosphorylent ensuite les résidus tyrosines également présents au niveau du domaine intracellulaire d'EPOR, de façon à initier différentes voies de signalisation, telles que les voies **STAT5**, **PI3K/AKT** et **MAPK/ERK** (Figure 8). Des mutations du gène codant pour l'EPOR ont déjà été rapportées. L'une d'entre elles induit la perte du domaine intracellulaire de liaison avec SHP-1, une phosphatase impliquée dans la déphosphorylation de JAK2. Les patients développent alors une polyglobulie du fait de l'absence de régulation négative de JAK2 (de la Chapelle et al., 1993).

c. Voies de signalisation activées en aval de c-KIT et EPOR

L'activation des voies de signalisation en aval des récepteurs c-KIT et EPOR a pour principale conséquence d'assurer la **survie** et la **prolifération** des cellules érythroïdes.

i. La voie STAT

La voie de signalisation STAT5 (pour *Signal transducer and activator of transcription* 5) est activée en aval du récepteur EPOR. Présents à l'état latent dans le cytoplasme, les facteurs de transcription STAT5A et STAT5B possèdent des domaines SH2 qui reconnaissent les tyrosines phosphorylées de l'EPOR. Ils sont alors à leur tour phosphorylés par les kinases JAK2, ce qui entraîne leur homodimérisation et leur translocation dans le noyau où ils se lient à l'ADN de façon à réguler la transcription de divers gènes impliqués dans le cycle cellulaire ou l'apoptose (Figure 8). STAT5 inhibe par exemple l'expression de gènes pro-apoptotiques tels que *BIM* et *FOXO3A*, tandis qu'il induit l'expression de Bcl-xL, un facteur anti-apoptotique (Socolovsky et al., 1999).

Il a récemment été montré que STAT1 et STAT3 étaient aussi phosphorylées par JAK2 suite à la liaison de l'EPO avec son récepteur (Moraga et al., 2015). En 2017, l'équipe de Vijay Sankaran a décrit un rôle essentiel de l'activation de STAT1 et STAT3 pour la prolifération et la différenciation érythrocytaires (Kim et al., 2017). En effet, une mutation du gène de l'*EPO* de type homozygote (R150Q) a été retrouvée pour la première fois chez un jeune garçon présentant une aplasie médullaire. L'EPO possède deux sites de fixation à l'EPOR, dont un site de haute affinité où se trouve l'arginine en position 150 (Syed et al., 1998). La mutation R150Q affecte considérablement la cinétique d'association du ligand EPO à son récepteur, ce qui, en aval, diminue la phosphorylation de STAT1 et STAT3 par JAK2 sans pour autant modifier la phosphorylation de STAT5. *In vitro*, les cellules

hématopoïétiques CD34⁺ – isolées par cytophérèse puis transduites avec l'EPO recombinante mutée R150Q et cultivées en conditions de différenciation érythroïde – arrêtent leur maturation au stade Baso-E et prolifèrent très peu. Ces travaux démontrent donc l'importance des voies de signalisation STAT1 et STAT3 dans l'érythropoïèse humaine.

ii. La voie PI3K/AKT

La voie de signalisation PI3K/AKT (pour *Phosphatidylinositol-3-kinase/Protein kinase B*) est activée en aval des récepteurs c-KIT et EPOR. L'enzyme PI3K est composée de deux sous-unités : une sous-unité régulatrice p85 qui – grâce à ses domaines SH2 – reconnaît les tyrosines phosphorylées dans le domaine intracellulaire de c-KIT et EPOR ; et une sous-unité catalytique p110 activée après fixation de p85 avec c-KIT ou EPOR. L'hétérodimère p85/p110 phosphoryle le phosphatidyl-inositol-diphosphate (PIP2) en phosphatidyl-inositol-triphosphate (PIP3) au niveau de la membrane plasmique, ce qui permet le recrutement de la sérine/thréonine kinase AKT et son activation par phosphorylation de sa thréonine 308 et de sa sérine 473 par la kinase PDK1 (pour *Phosphoinositide-dependent kinase-1*) et par mTORC2 (pour *Mammalian target of rapamycin complex 2*), respectivement. La phosphatase PTEN (pour *Phosphatase and tensin homolog*) déphosphoryle le PIP3 en PIP2, et assure donc la régulation négative de l'axe PI3K/AKT (Yu and Cui, 2016).

Lorsqu'elle est activée, la kinase AKT inhibe d'une part l'action de facteurs pro-apoptiques, comme BAD et FOXO3A et d'autre part, active indirectement mTORC1 (pour *Mammalian target of rapamycin complex 1*), en phosphorylant TSC2 (pour *Tuberous sclerosis 2*) au sein du complexe TSC1-2, un régulateur négatif de l'activité catalytique de mTORC1. Le complexe protéique mTORC1 contrôle de nombreuses voies de signalisation indispensables à la prolifération cellulaire, telles que la biogenèse du ribosome, la traduction et la synthèse des nucléotides et lipides (Figure 8).

iii. La voie MAPK/ERK

La voie de signalisation MAPK/ERK (pour *Mitogen-activated protein kinase/extracellular signal-regulated kinase*) est activée en aval des récepteurs c-KIT et EPOR (Gobert et al., 1995). La protéine Grb2 (pour *Growth factor receptor-bound protein 2*) se lie – via son domaine SH2 – aux récepteurs c-KIT et EPOR phosphorylés. Grb2 forme un complexe membranaire – via cette fois-ci ses domaines SH3 (pour *Src homology domain 3*) – avec le facteur d'échange Sos (pour *Son of sevenless*), qui assure la conversion de la forme

inactive de Ras liée au guanosine diphosphate (GDP) en sa forme active liée au guanosine triphosphate (GTP). Il s'ensuit alors l'activation en cascade de la GTPase Ras, la sérine/thréonine kinase Raf-1 puis les sérine/thréonine/tyrosine kinases MEK1 et MEK2 (ou MAPKK pour *Mitogen-activated protein kinase kinase*) qui phosphorylent les protéines effectrices MAPK (ERK1 et ERK2). Cette voie favorise la prolifération et inhibe la différenciation des progéniteurs érythroïdes immatures (Figure 8).

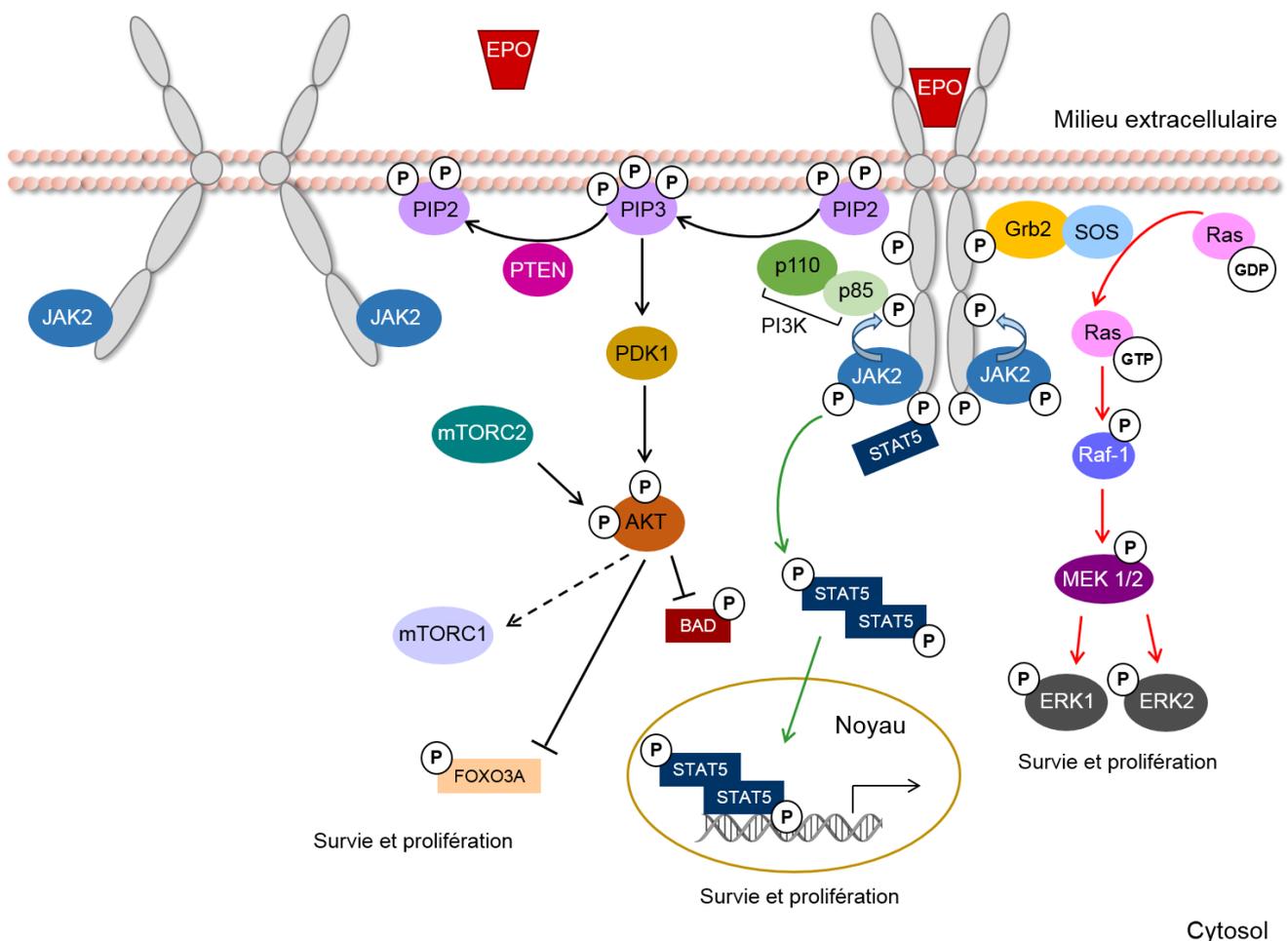


Figure 8 : Voies de signalisation activées par le couple EPO/EPOR

La liaison de l'EPO à son récepteur EPOR conduit à l'activation par autophosphorylation des kinases JAK2. Ces dernières phosphorylent ensuite EPOR (flèches bleues), ce qui active trois principales voies de signalisation : la voie PI3K/AKT (flèches noires), la voie MAPK/ERK (flèches rouges) et la voie STAT5 (flèches vertes). Chacun des effecteurs finaux de ces trois voies de signalisation est impliqué dans la survie et la prolifération des progéniteurs et précurseurs érythroïdes. **P** = acide aminé phosphorylé.

2. Régulation transcriptionnelle

Parallèlement aux voies de signalisation activées en aval de c-KIT et EPOR, de nombreux **facteurs de transcription** contrôlent, d'une part, l'engagement des progéniteurs multipotents vers la voie érythroïde et, d'autre part, leur différenciation en érythrocytes matures.

Parmi ces facteurs de transcription, **GATA1** est le plus important pour l'érythropoïèse. L'équipe de Stuart H. Orkin a tout d'abord montré que l'inactivation du gène *Gata1* dans les cellules souches embryonnaires murines, cultivées *in vitro*, conduit à l'arrêt de la différenciation érythroïde au stade Pro-E (Weiss et al., 1994). Par la suite, les travaux *in vivo* de cette même équipe ont confirmé le rôle majeur de GATA1 dans l'érythropoïèse (Fujiwara et al., 1996). Le gène *Gata1* étant situé sur le chromosome X, son inactivation entraîne effectivement une mort *in utero* des embryons mâles dès le 10^{ème} jour de gestation. Les souris femelles hétérozygotes *Gata1*^{+/-} naissent quant à elles avec une anémie profonde. D'après les études fonctionnelles subséquentes, GATA1 régule la différenciation et la survie érythroïdes. En inhibant le facteur de transcription SPI1 (également nommé PU.1) qui favorise la granulopoïèse et le facteur de transcription GATA2 exprimé dans les CSH, il oriente en effet les cellules progénitrices multipotentes vers le lignage érythroïde (Chou et al., 2009; Suzuki et al., 2013). De plus, il active la transcription de différents gènes spécifiquement exprimés au cours de la maturation érythrocytaire, tels que l'*EPOR*, le *GYP A*, ou encore les gènes codant pour les globines et les enzymes participant à la biosynthèse de l'hème (comme la porphobilinogène-désaminase) (Orkin, 1992). Au contraire, il régule négativement la transcription de c-KIT à partir du stade CFU-E, ce qui limite la prolifération des cellules érythroïdes par blocage en phase G1 du cycle cellulaire, mais permet leur différenciation (Munugalavadla et al., 2005). Enfin, GATA1 coopère avec l'EPO pour promouvoir la survie des précurseurs érythroïdes *via* l'induction de l'expression du facteur anti-apoptotique Bcl-xL (Gregory et al., 1999).

Durant la même période, l'équipe de Stuart H. Orkin a analysé les conséquences *in vivo* de l'inactivation des gènes codant pour les facteurs de transcription **TAL1** et **KLF1**. En plus de l'absence d'érythrocytes primitifs nucléés, les souris *Tal1*^{-/-} présentent une réduction majeure des progéniteurs multipotents myéloïdes, comme en témoignent les études de clonogénicité réalisées à partir des cellules hématopoïétiques du foie foetal (Shivdasani et al., 1995). Ces défauts d'érythropoïèses primitive et définitive sont létaux au stade embryonnaire.

En outre, le facteur de transcription KLF1 se fixant au niveau de la région promotrice du gène codant pour la β -globine, les souris *Klf1*^{-/-} se caractérisent quant à elles par une absence d'expression de la β -globine conduisant à un arrêt de la maturation érythrocytaire terminale (Perkins et al., 1995). Elles meurent au stade fœtal.

GATA1, TAL1 et KLF1 peuvent s'associer avec d'autres facteurs de transcription et ainsi former des complexes nucléaires activateurs ou répresseurs de l'expression des gènes érythroïdes. Par exemple, lorsque les cellules sont engagées dans la différenciation érythroïde terminale, le complexe GATA1/SCL/NFE2 induit l'expression du gène codant pour la globine α après remodelage de la chromatine – notamment hyperacétylation des histones – et recrutement de l'ARN polymérase II (Anguita et al., 2004). En revanche, le complexe GATA1/GFI1B inhibe l'expression des gènes érythroïdes *via* le recrutement d'histone désacétylases ou d'histone méthyltransférases (Rodriguez et al., 2005). Le facteur de transcription **GFI1B** a été étudié plus récemment. Le KO *Gfi1b*^{-/-} chez la souris entraîne un arrêt de la différenciation érythro-mégacaryocytaire dans le foie fœtal – avec notamment une incapacité à produire des érythrocytes définitifs énucléés –, ce qui cause la mort *in utero* de ces souris (Saleque et al., 2002). La régulation de l'érythropoïèse au niveau transcriptionnel fait donc intervenir de nombreux facteurs de transcription, chacun d'eux étant essentiel à la **survie** et/ou à la **différenciation érythroïdes**.

3. Régulation par les caspases

Les processus d'**apoptose cellulaire** et de **maturation érythroïde terminale** partagent des changements morphologiques identiques, tels que la condensation du noyau et de la chromatine. Au cours de l'érythropoïèse physiologique, on observe effectivement une activation transitoire des caspases, protéases effectrices du programme de mort cellulaire. Selon les cytokines présentes, l'activation des caspases induit la différenciation des érythroblastes ou leur apoptose assurant ainsi une **régulation négative de l'érythropoïèse**.

a. Rappels sur l'apoptose cellulaire

Il existe deux voies majeures d'activation des caspases : la **voie extrinsèque** et la **voie intrinsèque**. La voie extrinsèque fait intervenir les récepteurs membranaires à domaine de mort, comme le récepteur Fas qui appartient à la superfamille des récepteurs au TNF (pour *Tumor Necrosis Factor*). Lorsque le ligand FasL se lie à son récepteur Fas, il y a activation de la caspase-8. La voie intrinsèque est quant à elle initiée par dépolarisation puis

perméabilisation de la membrane mitochondriale. Des facteurs pro-apoptotiques sont alors libérés de l'espace intermembranaire mitochondrial vers le cytosol, dont le cytochrome *c* qui, avec l'APAF-1 et la pro-caspase-9, forme un complexe d'activation de la caspase-9 : l'apoptosome. Les caspases-8 et -9 activent ensuite la caspase-3, cette dernière clivant divers substrats responsables de la condensation du noyau et de la chromatine, ainsi que de la fragmentation de l'ADN.

b. Rôle des caspases dans la différenciation érythroïde terminale

En 2001, l'équipe d'Olivier Hermine a démontré que la transition entre les stades Baso-E et Poly-E s'accompagne d'une réduction du potentiel transmembranaire mitochondrial aboutissant, entre autres, à l'activation de la caspase-9 puis -3 (Zermati et al., 2001). La **caspase-3** activée clive la **lamine B** et la protéine **acinus**, ce qui entraîne une **condensation du noyau et de la chromatine**. De ce fait, l'inhibition des caspases par le ZVAD – un pan-inhibiteur synthétique des caspases – induit l'arrêt de la maturation des érythroblastes au stade Baso-E chez l'homme (Ekert et al., 1999; Zermati et al., 2001). En revanche, contrairement à l'apoptose cellulaire, il n'y a pas de clivage de la DFF45/ICAD. L'endonucléase DFF40/CAD – réprimée par la DFF45/ICAD – n'est donc pas activée, et l'ADN conserve son intégrité en fin d'érythropoïèse. De même, le facteur de transcription GATA1 n'est pas clivé par la caspase-3 activée (Figure 9). En 2016, Zhao *et al.* confirment que le clivage de la lamine B par la caspase-3 est essentiel à la différenciation érythroïde terminale. En effet, le traitement des érythroblastes immatures murins cultivés *in vitro* avec un inhibiteur spécifique de la caspase-3 ou un shARN (sh pour *small hairpin*, petit ARN en épingle à cheveux) dirigé contre la caspase-3 empêche l'ouverture transitoire de l'enveloppe nucléaire des érythroblastes, et donc la fuite des histones nucléaires vers le cytosol (Zhao et al., 2016) (cf. Chapitre I-I-B-2-b). À noter que dans les érythroblastes en cours de maturation, l'enveloppe nucléaire conserve quasi-entièrement son intégrité – excepté au niveau de l'ouverture –, tandis que l'enveloppe nucléaire des cellules apoptotiques est morcelée de façon irrégulière.

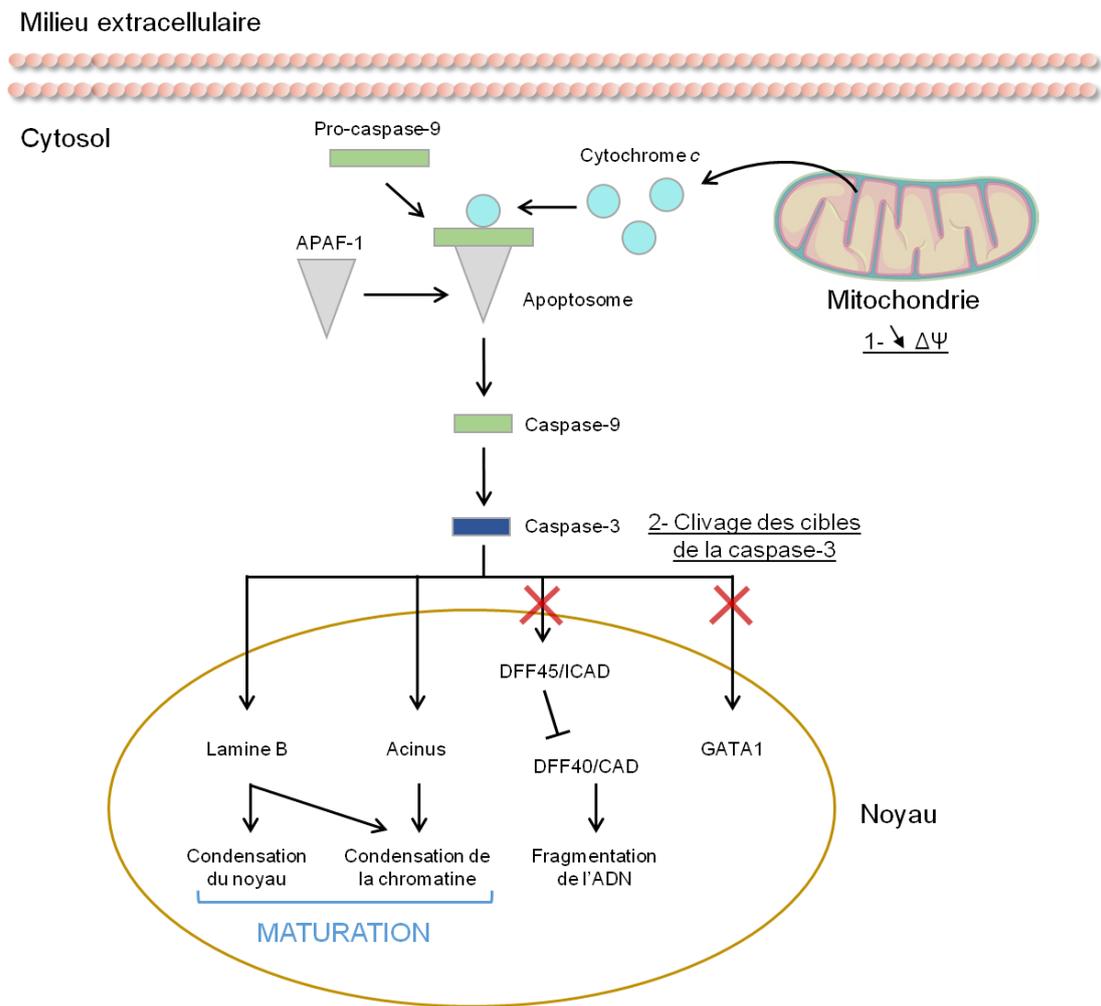


Figure 9 : Activation de la caspase-3 par dépolérisation de la membrane mitochondriale au cours de la maturation érythroïde terminale

(1) Entre les stades Baso-E et Poly-E, le potentiel transmembranaire mitochondrial ($\Delta\Psi$) diminue, et le cytochrome *c* est libéré dans le cytosol. (2) La voie intrinsèque d'activation des caspases aboutit au clivage de la caspase-3, qui clive ensuite divers substrats dont la lamine B et la protéine acinus, responsables de la condensation du noyau et de la chromatine. Contrairement au processus d'apoptose cellulaire, la sous-unité inhibitrice DFF45 de l'hétérodimère DFF45/DFF40 et GATA1 ne sont pas clivés durant la maturation érythroïde terminale (Zermati et al., 2001).

c. Rôle des caspases dans la régulation négative de l'érythropoïèse

La régulation négative de l'érythropoïèse est essentielle pour limiter une production trop importante de globules rouges. Elle a lieu au niveau rénal – par diminution de l'expression et de la sécrétion d'EPO (cf. Chapitre I-I-C-1-b) – mais aussi au niveau de l'îlot érythroblastique où elle est alors de type paracrine.

La liaison de l'EPO à son récepteur a deux conséquences majeures : l'activation de la transcription de **Bcl-xL** *via* la voie STAT5, et l'inhibition de son partenaire pro-apoptotique BAD *via* la voie PI3K/AKT (cf. Chapitre I-I-C-1-c). Lorsque les taux d'EPO sont faibles, les érythroblastes entrent en apoptose par défaut d'expression et/ou blocage de Bcl-xL, et dépolérisation de la membrane mitochondriale (Figure 10-1).

De plus, au sein de l'îlot érythroblastique, les érythroblastes immatures et matures interagissent autour d'un macrophage central (cf. Chapitre I-I-E-3). En absence d'EPO, les **érythroblastes immatures** (Baso-E) expriment fortement le récepteur membranaire à domaine de mort **Fas** (De Maria et al., 1999). Fas s'associe alors avec son ligand **FasL** présenté à la surface des **érythroblastes matures** (Poly-E), ce qui déclenche la voie extrinsèque d'activation de la caspase-3 qui clive **GATA1**. Les érythroblastes immatures entrent en apoptose (Figure 10-2). Par contre, en présence d'EPO, la protéine chaperonne **HSP70** colocalise avec GATA1 dans le noyau et le protège du clivage (Ribeil et al., 2007) (Figure 11). GATA1 peut alors réguler la transcription de ses gènes cibles, et assurer la maturation érythroïde terminale ainsi que la survie des érythroblastes en induisant la transcription de *Bcl-xL*.

En résumé, les facteurs de croissance, les caspases et le facteur de transcription GATA1 en association avec HSP70 assurent l'équilibre entre apoptose, survie et différenciation terminale des cellules érythroblastiques immatures.

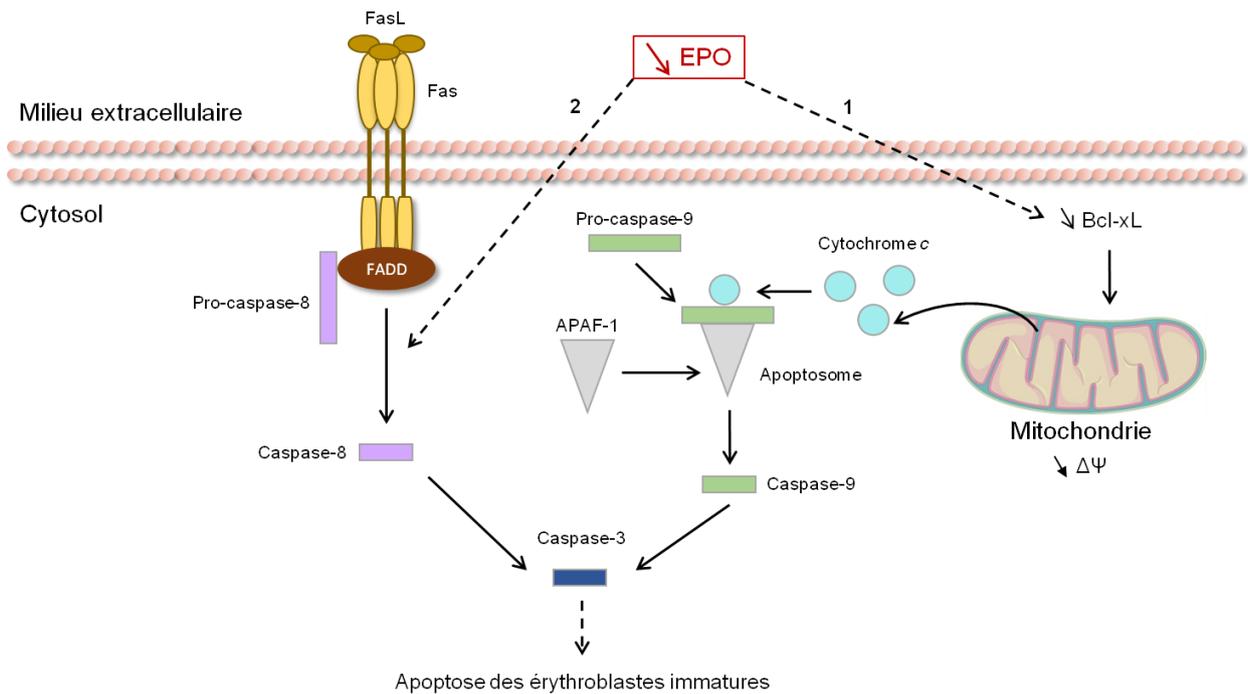


Figure 10 : Régulation négative de l'érythropoïèse par les caspases

Lorsque les concentrations en EPO diminuent, la production érythropoïétique est limitée par activation de la caspase-3 puis apoptose des cellules les plus immatures au niveau de l'îlot érythroblastique. **(1)** Voie intrinsèque d'activation des caspases : Bcl-xL n'assure plus son rôle anti-apoptotique, d'où la dépolarisation de la membrane mitochondriale et l'activation de la caspase-3. **(2)** Voie extrinsèque d'activation des caspases : FasL exprimé à la surface des érythroblastes matures se lie à son récepteur Fas exprimé à la surface des érythroblastes immatures, ce qui active la caspase-3.

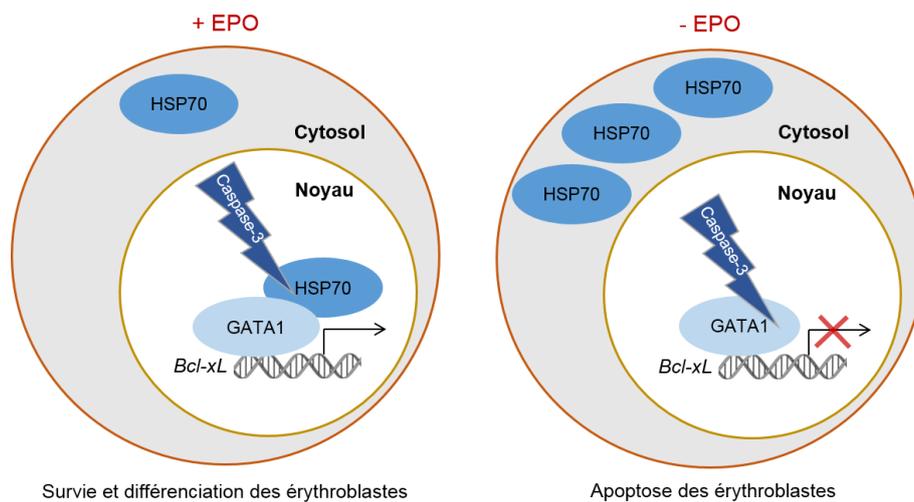


Figure 11 : Régulation de l'érythropoïèse par la protéine chaperonne HSP70 (d'après Vandekerckhove et al., 2008)

En présence d'EPO, la protéine chaperonne HSP70 est localisée dans le noyau et protège ainsi GATA1 du clivage par la caspase-3, ce qui assure la survie et la différenciation des érythroblastes par transcription, notamment, de *Bcl-xL*. En absence d'EPO, HSP70 est localisée dans le cytoplasme, d'où le clivage de GATA1 et l'apoptose des érythroblastes.

D. Modulation de l'expression génique au cours de l'érythropoïèse

Malgré la définition de stades de maturation érythroïdes, l'érythropoïèse est un processus biologique continu au cours duquel l'expression génique varie aux niveaux transcriptionnel, post-transcriptionnel et protéique.

1. Au niveau transcriptionnel

En 2014, l'équipe de Mohandas Narla a publié l'analyse transcriptomique globale des érythroblastes humains du stade Pro-E au stade Acido-E (An et al., 2014). Les CSH CD34⁺ isolées à partir de sang de cordon ont tout d'abord été cultivées *in vitro* en conditions de différenciation érythrocytaire puis, d'après leurs précédents travaux, l'analyse des marqueurs membranaires GPA, Band 3 et intégrine $\alpha 4$ a permis le tri des diverses populations érythroïdes au cytomètre en flux (Hu et al., 2013). L'ARN de chacune de ces populations érythroïdes a ensuite été séquencé. Cette étude démontre que chaque mitose donne naissance à deux cellules filles dont le profil d'expression génique diffère de celui de leur cellule mère. De manière générale, les gènes différentiellement exprimés sont classés en six groupes. Les deux premiers groupes, constitués de gènes impliqués notamment dans les processus de catabolisme et d'apoptose cellulaires, voient leur expression augmenter continuellement au cours de l'érythropoïèse terminale. Au contraire, les groupes 3 et 4 renferment des gènes dont l'expression maximale au stade Pro-E décroît durant la différenciation érythroïde terminale. Parmi les fonctions biologiques des gènes appartenant à ces deux groupes, on retrouve la traduction, la biogenèse du ribosome, la réplication et la réparation de l'ADN ainsi que le cycle cellulaire. Enfin, les groupes 5 et 6 contiennent des gènes – comme les gènes de contrôle de la mitose – ayant soit un pic soit un nadir d'expression au stade Poly-E. Quel que soit le stade de différenciation, les transcrits les plus exprimés sont essentiels à la synthèse de l'hémoglobine, tels que *HBB* et *HBG2* qui codent pour les globines β et γ respectivement. L'expression des gènes impliqués dans le métabolisme du fer – dont *TFRC* (ou *TFRI*) codant pour le récepteur à la transferrine de type 1 – est également très importante, de même que celle des gènes codant pour les protéines ribosomiques dans les Pro-E et Baso-E précoces très proliférants. Parmi les facteurs de transcription, *GATA1*, *KLF1* et *GFI1B* font partie des transcrits les plus exprimés.

En conclusion, la **différenciation érythroïde terminale** s'accompagne d'une évolution du transcriptome, sachant que les **changements d'expression génique les plus drastiques apparaissent lors des transitions Baso-E tardifs/Poly-E et Poly-E/Acido-E**. Les variations du transcriptome sont propres à chaque espèce, les travaux de cette équipe ayant montré une diminution globale de l'expression des gènes durant la différenciation érythroïde murine, ce qui n'est pas observé chez l'homme.

2. Au niveau protéique

Récemment, notre équipe a réalisé la quantification absolue des protéines présentes à chaque étape de la différenciation érythroïde humaine, du stade BFU-E au stade Acido-E (Gautier et al., 2016). Pour cela, les protéines des érythroblastes cultivés *in vitro* à partir des progéniteurs hématopoïétiques CD34⁺ ont été digérées à la trypsine puis analysées au spectromètre de masse (SM).

Cette étude a permis d'identifier 6 130 protéines, dont l'expression a été évaluée en nombre de copies par cellule. **Beaucoup de protéines voient leur expression diminuer au cours de l'érythropoïèse**. C'est le cas notamment des protéines impliquées dans la maturation des ARNm, la traduction protéique, le métabolisme cellulaire, la réplication et la réparation de l'ADN, mais aussi les trois principaux facteurs de transcription érythroïde : GATA1, TAL1 et KLF1. Seulement une quarantaine de protéines sont plus exprimées dans les érythroblastes matures par rapport aux érythroblastes immatures, comme l'hémoglobine et la Band 3 qui sont des constituants majeurs du futur globule rouge.

De façon étonnante, **les profils d'expression des ARNm sont faiblement corrélés aux profils d'expression des protéines**. Au total, 302 gènes présentent une augmentation de leur expression au niveau transcriptionnel, mais une diminution de leur expression au niveau protéique. Parmi ces gènes, un grand nombre d'entre eux subissent une rétention d'intron en fin d'érythropoïèse, ce qui explique cette discordance (cf. Chapitre I-I-D-1) (Pimentel et al., 2016). L'analyse d'enrichissement, à partir de la base de données GO (GO pour *Gene ontology*), souligne une surreprésentation des gènes contrôlant la maturation des ARNm, dont *SF3B1*.

La dichotomie entre le transcriptome et le protéome apparaît au stade Baso-E tardif. **La transition Baso-E précoce/Baso-E tardif** constitue en effet un point central dans

l'érythropoïèse, au cours duquel la cellule érythroblastique subit diverses modifications : l'expression protéique des globines, de la Band 3 et de la GPA commence à augmenter, tandis que celle du c-KIT et des protéines d'épissage baisse de façon drastique. De même, les cibles des voies de signalisation activées par le couple EPO/EPOR – telles que STAT5A – ont une expression protéique réduite après le stade Baso-E précoce. La diminution de la signalisation intracellulaire en aval des récepteurs c-KIT et EPOR intervient donc concomitamment à l'activation transitoire des caspases, processus indispensable à la différenciation érythrocytaire terminale (Zermati et al., 2001).

3. Régulation post-transcriptionnelle de l'expression génique : rôle de l'épissage

Par des études de RNA-Seq, John G. Conboy a mis en évidence un **programme dynamique d'épissage alternatif au cours de l'érythropoïèse terminale humaine** (pour les définitions d'épissage et d'épissage alternatif, se référer au Chapitre II Partie 1, I-A et II-C-1). L'évolution de l'épissage à chaque stade de différenciation érythroïde peut générer des modifications dans la structure et l'abondance des isoformes protéiques (pour revue Conboy, 2017). L'épissage alternatif concerne particulièrement les gènes contrôlant le cycle cellulaire, l'organisation des organites, la structure de la chromatine et la maturation des ARNm (Pimentel et al., 2014). Cet épissage alternatif peut être responsable de l'ajout d'exons codant pour un codon stop prématuré ou d'un décalage du cadre de lecture (avec, également, formation d'un codon stop prématuré), ces deux phénomènes ayant pour conséquence principale la dégradation du transcrit différenciellement épissé par la voie **NMD** (pour *Nonsense-mediated decay*, dégradation des ARNm non-sens). Le NMD est un « contrôle qualité » des ARNm. Il a lieu dans le cytosol et vise à éliminer les ARNm comportant des erreurs de transcription et/ou d'épissage. Il évite ainsi la production de protéines anormales qui pourraient avoir des conséquences délétères pour la cellule. En contrepartie, l'expression du gène diminue. De manière intéressante, une augmentation significative des transcrits sujets à une dégradation par la voie NMD a été mise en évidence dans les érythroblastes matures Poly-E et Acido-E, ce qui représente un mécanisme majeur de régulation post-transcriptionnelle de l'expression des gènes. En effet, à partir du stade Poly-E, certaines isoformes des facteurs d'épissage SRSF2, SRSF3, SRSF6 et SRSF7 présentent une inclusion alternative d'exon responsable de la dégradation du transcrit, et donc une diminution de

l'expression génique. Par exemple, alors que le pourcentage d'expression du transcrit alternatif de SRSF2 n'est que de 7% dans les Pro-E, il atteint 54% dans les Acido-E.

En 2016, les nouveaux travaux de cette même équipe ont souligné le rôle de la **réention d'intron** dans la régulation de l'expression génique au sein des érythroblastes, et plus particulièrement au sein des érythroblastes matures (Pimentel et al., 2016). En effet, alors que peu de transcrits dans les Pro-E et Baso-E présentent une réention d'intron, il y a jusqu'à 20 à 50% de transcrits conservant un intron après épissage en fin d'érythropoïèse. Parmi les transcrits les plus enrichis en réention d'intron dans les stades Poly-E et Acido-E, on trouve les transcrits de maturation des ARNm, dont le facteur d'épissage SF3B1 qui conserve l'intron 4 après épissage dans quasiment 50% des transcrits dans les Acido-E contre 15% dans les Pro-E (Figure 12). Je reviendrai sur la protéine SF3B1 tout au long du chapitre II. Contrairement à *SF3B1*, d'autres gènes se distinguent par un fort taux de réention d'intron durant l'ensemble du processus d'érythropoïèse terminale, c'est-à-dire du stade Pro-E au stade Acido-E. C'est le cas par exemple des gènes *SLC25A37* et *SLC25A28*, codant respectivement pour les importateurs de fer mitochondriaux mitoferrine-1 et mitoferrine-2, essentiels à la synthèse de l'hème (cf. Chapitre I-I-E-2).

Les transcrits avec réention d'intron peuvent être divisés en deux classes : l'une regroupant les transcrits avec introduction d'un codon stop prématuré et, donc, dégradation dans le cytosol par le NMD, et l'autre regroupant les transcrits bloqués dans le noyau et, donc, résistants au NMD. Les transcrits bloqués dans le noyau sont communément appelés transcrits avec déention d'intron (pour revue Fu, 2017). Ce mode de régulation post-transcriptionnel peut introduire une discordance entre l'abondance des ARNm et celle des protéines, puisque les transcrits avec déention d'intron ne sont pas traduits en protéines (l'expression du gène au niveau ARNm reste stable et/ou augmente lors de l'érythropoïèse à l'inverse de l'expression au niveau protéique qui, elle, diminue). Les fonctions physiologiques du mécanisme de déention d'intron sont à ce jour peu connues, outre le fait qu'il permette de fournir un « réservoir » pour la production de nouveaux ARNm en réponse à des signaux appropriés.

Ainsi, l'analyse des variations d'expression génique aux niveaux transcriptionnel, post-transcriptionnel et protéique a nettement amélioré, d'un point de vue moléculaire, la caractérisation de chaque stade de maturation érythroïde terminale.

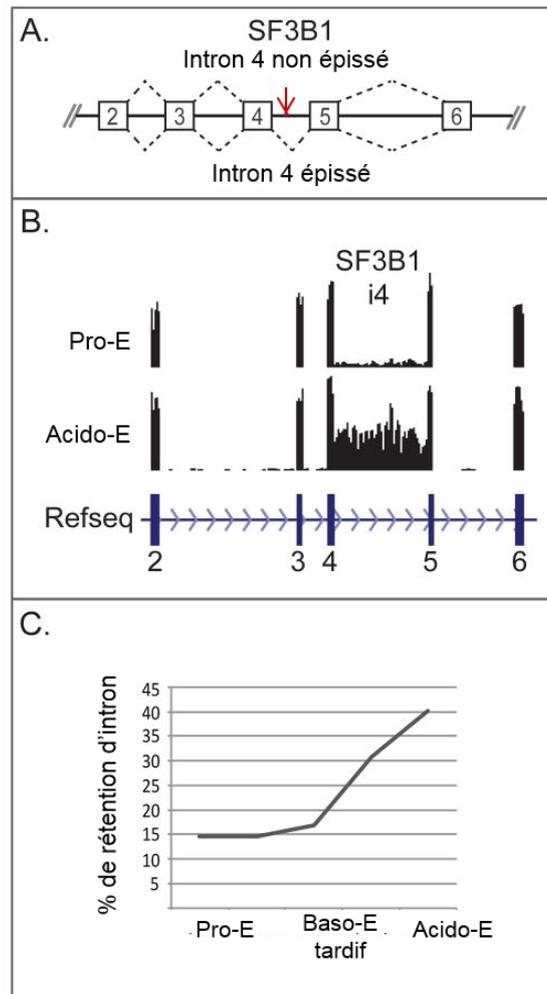


Figure 12 : Dynamique de la rétention d'intron dans le transcript SF3B1 au cours de l'érythropoïèse terminale (d'après Conboy, 2017)

(A) Le gène *SF3B1* peut être transcrit en plusieurs isoformes. L'une de ces isoformes conserve l'intron 4 après épissage, ce qui introduit un codon stop prématuré (flèche rouge) et empêche alors la traduction protéique. (B) Les profils RNA-Seq de *SF3B1*, annotés selon la base de données Refseq, montrent que la rétention de l'intron 4 est plus importante en fin de maturation érythroïde (stade Acido-E). (C) Le pourcentage de transcrits *SF3B1* avec rétention d'intron augmente progressivement au cours de l'érythropoïèse jusqu'à atteindre 40% en fin de différenciation. **i4** = intron 4.

E. Érythropoïèse et fer

Le fer est un élément essentiel à de nombreuses fonctions biologiques, telles que la respiration cellulaire, la réparation de l'ADN et surtout le transport de l'oxygène *via* l'hémoglobine contenue dans les globules rouges. Le processus d'érythropoïèse est de loin le plus consommateur de fer dans le corps humain. En effet, les érythrocytes renferment les deux tiers du fer total. Comme signalé au début de ce chapitre, la **production quotidienne de 200 milliards de globules rouges** – soit 2,4 millions par seconde – requiert **20 à 25 mg de fer**. En cas d'augmentation de l'activité érythropoïétique – par exemple en réponse à un saignement ou à l'hypoxie – il est donc indispensable pour l'organisme de rapidement mobiliser du fer.

La distribution du fer dans le corps humain ainsi que le métabolisme du fer seront dans un premier temps détaillés. Nous verrons qu'une partie de ce fer est stockée au sein des macrophages médullaires dans des structures organisées : les îlots érythroblastiques. Ces îlots érythroblastiques seront dans un deuxième temps développés. Enfin, dans un troisième temps, la régulation du métabolisme du fer, notamment par des facteurs sécrétés par les érythroblastes, sera expliquée. Les principaux acteurs du métabolisme du fer sont indiqués dans le tableau 1.

1. La distribution du fer dans l'organisme

L'organisme d'un adulte contient environ 4 g de fer, existant sous les formes de fer ferreux (Fe^{2+}) et de fer ferrique (Fe^{3+}). Ces derniers sont quasiment toujours inclus dans des complexes hétéroprotéiniques. On distingue trois compartiments : le fer fonctionnel, le fer transporté et le fer de réserve.

Le fer fonctionnel entre dans la composition de l'**hémoglobine** (70-80% du fer total), mais aussi de la **myoglobine** – métalloprotéine impliquée dans l'apport de l' O_2 aux muscles – et de diverses **protéines à clusters Fer/Soufre** (Fe/S) comme les complexes I à III de la chaîne respiratoire, les ADN polymérases répliquatives α , δ et ϵ et diverses protéines du métabolisme du fer détaillées par la suite. L'hémoglobine et la myoglobine sont toutes deux constituées d'un groupement prosthétique – l'hème contenant un atome de fer – et de molécules de globines. Contrairement à l'hémoglobine qui, chez l'adulte, se compose le plus souvent de deux chaînes de globines α et de deux chaînes de globines β (Perutz, 1963), la myoglobine est une protéine monomérique formée d'une seule sous-unité de globine.

Le transport plasmatique du fer est assuré par la **transferrine** (Schade and Caroline, 1946). Cette dernière assure un rôle essentiel, puisque le fer à l'état libre, c'est-à-dire non lié à la transferrine (NTBI pour *Non-transferrin bound iron*) entraîne la formation d'espèces réactives à l'oxygène (ROS pour *Reactive oxygen species*) (Brissot et al., 2012). Ces ROS peuvent oxyder de nombreux constituants cellulaires tels que les lipides et les acides nucléiques, et sont particulièrement toxiques au niveau des tissus hépatiques et cardiaques.

Le fer de réserve est quant à lui principalement couplé à la **ferritine** dans le cytoplasme des hépatocytes et des macrophages du système réticulo-endothélial (SRE). Chaque molécule de ferritine peut contenir jusqu'à 4 500 atomes de fer (Harrison et al., 1967). Lorsque le fer est associé à la transferrine et à la ferritine, il est sous forme de Fe^{3+} .

2. Le métabolisme du fer

Les différentes étapes du métabolisme du fer sont illustrées dans la figure 13. Le fer alimentaire est absorbé au niveau du pôle apical des entérocytes duodénaux, *via* le transporteur **DMT1** (pour *Divalent metal transporter 1*, également nommé DCT1 ou NRAMP2) (Canonne-Hergaux et al., 1999; Fleming et al., 1998; Gunshin et al., 1997). Pour cela, il doit tout d'abord, si nécessaire, être réduit en Fe^{2+} par une réductase ferrique – la Dcytb – associée à DMT1 (McKie et al., 2001). C'est par exemple le cas du fer « végétal », c'est-à-dire le fer non héminique Fe^{3+} . Au contraire, le fer « animal », c'est-à-dire le fer héminique Fe^{2+} n'a pas besoin de subir cette étape de réduction. Dans l'entérocyte, le fer est soit stocké sous forme de ferritine, soit exporté à la surface basolatérale où il est ensuite libéré dans la circulation sanguine par la **ferroportine** (Abboud and Haile, 2000; Donovan et al., 2000; McKie et al., 2000). Celle-ci est associée à une ferroxidase – l'héphaestine – qui oxyde le Fe^{2+} en Fe^{3+} , lui permettant alors d'être rapidement pris en charge par la transferrine plasmatique (Vulpe et al., 1999).

La transferrine est une glycoprotéine de 80 kDa produite par les hépatocytes. Elle possède deux domaines capables de fixer le fer et existe sous deux formes : l'holotransferrine, qui séquestre deux atomes de Fe^{3+} , et l'apotransferrine, qui est dépourvue de fer. Les adultes sains ont un coefficient de saturation de la transferrine en fer compris entre 20 et 30%. L'holotransferrine délivre le fer aux cellules en interagissant avec son récepteur de type 1 (**TFR1**, également nommé TFRC ou CD71) (pour revue Richardson and Ponka, 1997; pour revue Testa et al., 1993) ou de type 2 (**TFR2**) (Kawabata et al., 1999). Bien que TFR1 soit un

récepteur membranaire quasi-ubiquitaire, environ 70 à 80% des TFR1 sont situés à la surface des précurseurs médullaires érythroïdes. TFR2 est quant à lui principalement retrouvé au niveau de la membrane des hépatocytes. L'affinité du TFR2 pour la transferrine est très inférieure à celle du TFR1, mais il est essentiel à l'internalisation et au stockage du fer dans le foie. Le complexe holotransferrine/TFR1 est intégré dans une vésicule endosomale formée par invagination de la membrane plasmique des érythroblastes. L'acidification de cet organite permet la libération du fer de la transferrine, qui, après réduction en Fe^{2+} , entre alors dans le cytoplasme de la cellule érythroïde par la protéine DMT1. L'apotransferrine est quant à elle recyclée et libérée dans la circulation sanguine.

Au sein des érythroblastes, le fer est transporté dans la matrice mitochondriale *via* la **mitoferrine-1** (MFRN1 ou SLC25A37) (Shaw et al., 2006) où il sert, d'une part, à la synthèse de l'hème et, d'autre part, à la synthèse des protéines à clusters Fe/S. L'expression érythroïde du transcrite MFRN1 est contrôlée par les facteurs de transcription GATA1 et GATA2 (Amigo et al., 2011). La synthèse de l'hème est catalysée par une succession d'enzymes cytoplasmiques et mitochondriales, la δ -aminolévulinate synthase 2 (ALAS2) étant l'enzyme limitante (Figure 14) (pour revue Layer et al., 2010). La synthèse des protéines à clusters Fe/S se déroule en trois étapes exclusivement mitochondriales : 1) formation *de novo* de clusters [2Fe-2S] à partir du Fe^{2+} importé *via* la mitoferrine-1 et des ions sulfures produits lors de la conversion de la cystéine en alanine ; et assemblage sur la protéine d'échafaudage ISCU, 2) transfert du cluster [2Fe-2S] sur la glutarédoxine-5 (GLRX5) (Bandyopadhyay et al., 2008) *via* la protéine chaperonne HSPA9 et la co-chaperonne HSC20, 3) conversion du cluster [2Fe-2S] en cluster [4Fe-4S] et liaison définitive à l'apoprotéine cible (pour revue Stehling et al., 2014). Les protéines à clusters Fe/S n'étant pas toutes mitochondriales, le transporteur ABCB7 – situé sur la membrane mitochondriale interne – assure l'export des clusters matures vers le cytosol (Figure 15) (Csere et al., 1998; Pondarré et al., 2006). Plusieurs notions ne sont pas encore clairement définies : quel est le rôle exact de la protéine GLRX5 dans la maturation des clusters Fe/S ? Quelle est la nature du substrat véhiculé par le transporteur ABCB7 ? Comment est régulé le métabolisme du fer mitochondrial ? Des mutations constitutionnelles ont été identifiées sur plusieurs gènes impliqués dans la biosynthèse de l'hème ou des protéines à clusters Fe/S. Elles causent, entre autres, le développement d'anémies sidéroblastiques, hémopathies abordées dans le Chapitre I-II-A-2-a. Il est important de noter qu'à l'état physiologique, le fer stocké dans les mitochondries sous forme de ferritine mitochondriale est quasiment absent puisqu'il est normalement systématiquement mobilisé.

Les érythrocytes vieillissants sont phagocytés par les macrophages du SRE – surtout présents dans la rate – qui assurent le recyclage du fer. Ces macrophages expriment à leur surface la ferroportine associée à une ferroxidase – la céruléoplasmine (Osaki et al., 1966) – et peuvent donc libérer à nouveau du fer dans le plasma en cas de besoin, ou le stocker sous forme de ferritine. La majorité du fer de l'organisme provient en fait de l'érythrophagocytose, l'absorption intestinale n'apportant qu'1 mg de fer par jour.

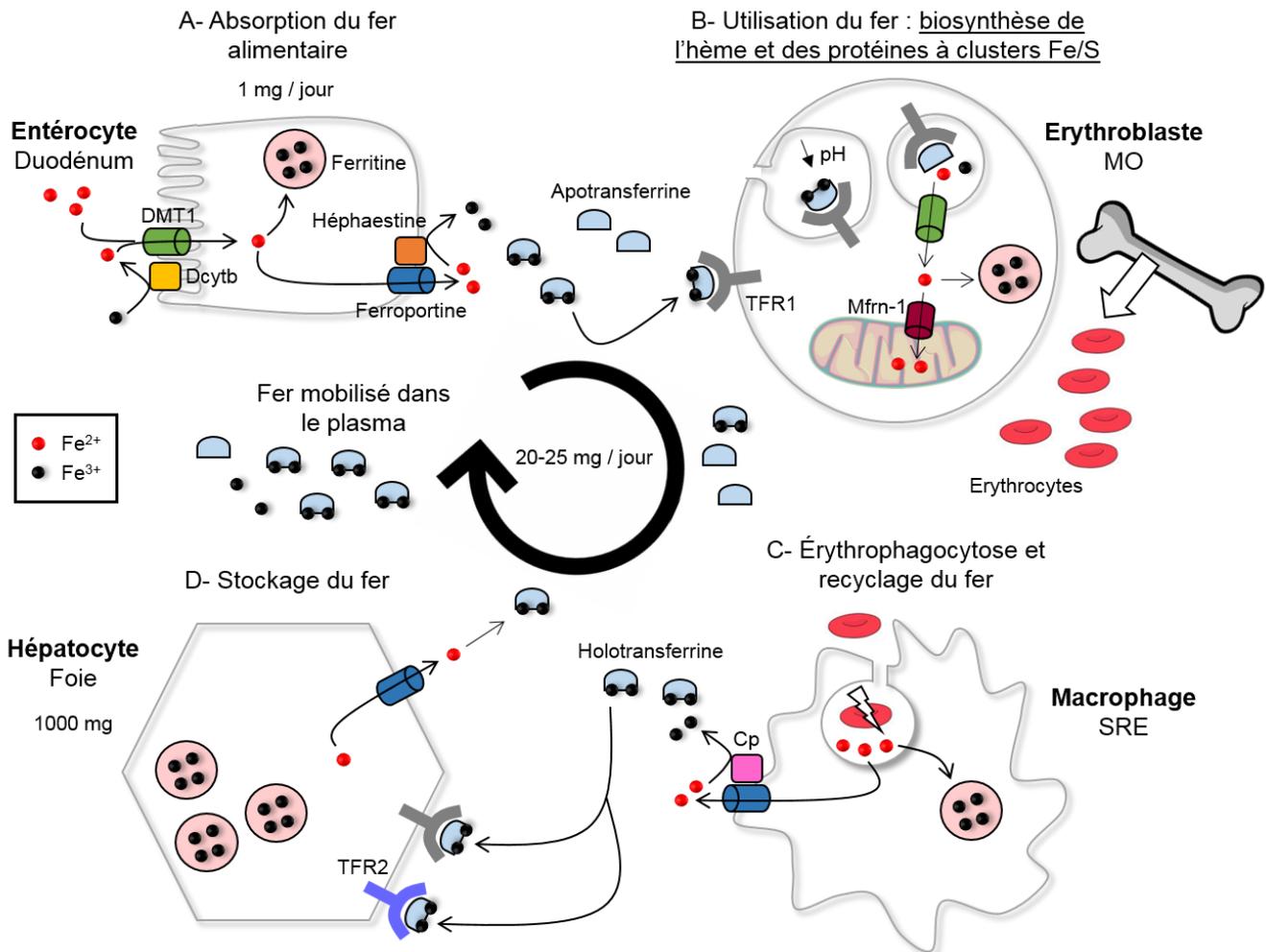
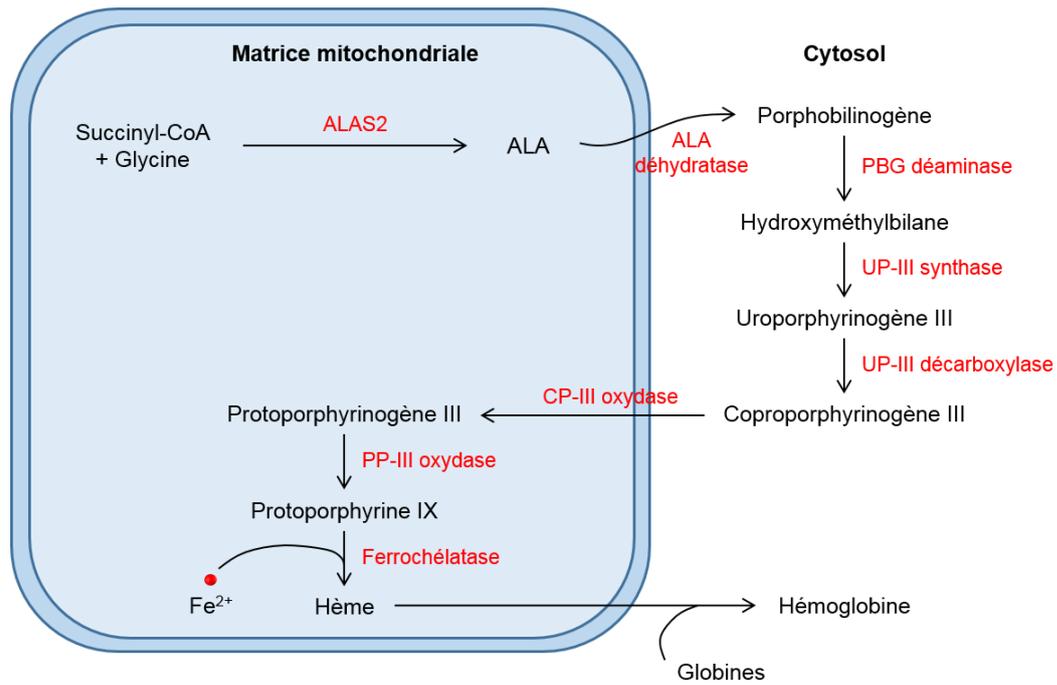


Figure 13 : Métabolisme du fer

(A) Chaque jour, environ 1 mg de fer alimentaire est absorbé au niveau des entérocytes duodénaux sous forme de fer ferreux (Fe^{2+}), de façon à compenser les pertes physiologiques (desquamation, menstruations...). (B) Le fer est majoritairement utilisé par les érythroblastes de la moelle osseuse (MO) pour la synthèse de l'hème et des protéines à clusters Fe/S (cf. Figures 14 et 15). (C) Les érythrocytes sénescents sont phagocytés par les macrophages du système réticulo-endothélial (SRE), qui assurent le recyclage du fer. La majeure partie du fer ferrique (Fe^{3+}) mobilisé dans le plasma sanguin provient en fait de l'érythrophagocytose. (D) En plus des macrophages du SRE et des entérocytes duodénaux, les hépatocytes stockent le fer et sont capables de le libérer du fer en cas de besoin. **Mfrn-1** = mitoferrine-1, **Cp** = céruléoplasmine.



Figure

14 : Biosynthèse de l'hème

La biosynthèse de l'hème a lieu principalement dans la matrice mitochondriale et le cytosol des précurseurs érythroïdes. Sept enzymes catalysent à tour de rôle la formation du noyau porphyrique. Finalement, la huitième et dernière enzyme – la ferrochélatase – incorpore un atome Fe^{2+} dans la protoporphyrine IX ce qui produit l'hème. **ALAS2** = δ -aminolévulinate synthase 2, **ALA** = acide δ -aminolévulinique, **PBG** = porphobilinogène, **UP-III** = uroporphyrinogène III, **CP-III** = coproporphyrinogène III, **PP-III** = protoporphyrinogène III.

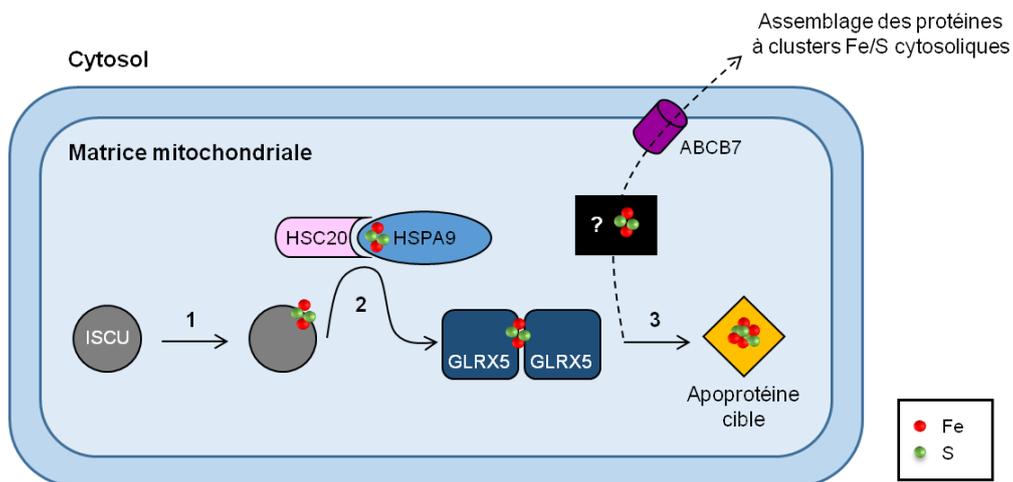


Figure 15 : Biosynthèse des protéines à clusters Fe/S

(1) Le cluster $[2Fe-2S]$ est assemblé sur la protéine d'échafaudage ISCU. (2) Des protéines chaperonnes aident au transfert du cluster sur les GLRX5. (3) Le cluster $[2Fe-2S]$ est converti en cluster $[4Fe-4S]$, ce dernier étant inséré dans l'apoprotéine cible. Le transporteur ABCB7 permet l'export des clusters Fe/S matures de la matrice mitochondriale vers le cytoplasme, mais la nature du substrat transporté n'est pas connue.

3. Les îlots érythroblastiques

Dans la moelle osseuse, les érythroblastes – grands consommateurs de fer – sont directement en contact avec des macrophages dans les îlots érythroblastiques. Les îlots érythroblastiques ont été décrits pour la première fois par Marcel Bessis en 1958. Ils se composent d'un **macrophage** central autour duquel sont rassemblés des **érythroblastes** à différents stades de maturation (Figure 16). Chez l'homme, on compte entre 5 et 30 cellules érythroïdes par îlot. Comme précédemment indiqué dans ce manuscrit, l'ensemble du processus d'érythropoïèse terminale a lieu dans ces compartiments médullaires, qui font office de « niches érythroïdes » (de Back et al., 2014; Manwani and Bieker, 2008).

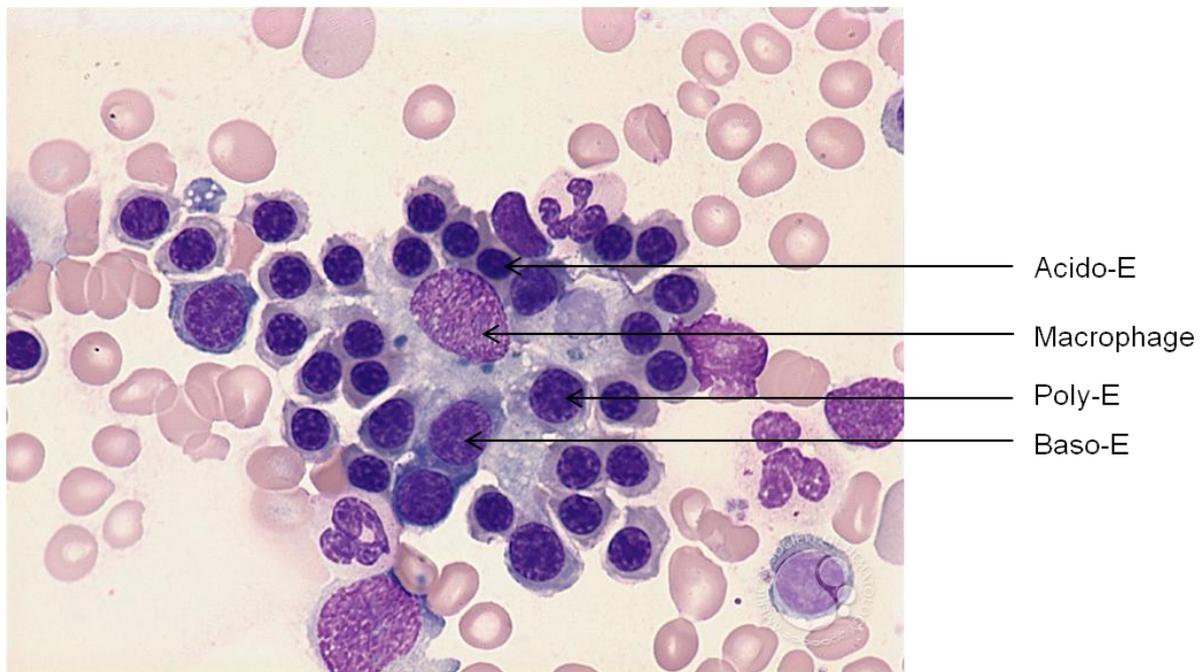


Figure 16 : L'îlot érythroblastique

Dans la moelle osseuse, les érythroblastes basophiles (**Baso-E**), polychromatophiles (**Poly-E**) et acidophiles (**Acido-E**) entourent le macrophage central. L'ensemble constitue l'îlot érythroblastique. Frottis médullaire photographié au microscope après coloration MGG (d'après <https://imagebank.hematology.org/imageset/708/erythroblastic-island>).

Au sein des îlots érythroblastiques, les interactions entre les cellules sont possibles grâce à diverses **molécules d'adhésion**, comme Emp (pour *Erythroblast macrophage protein*). Emp s'exprime à la surface des macrophages et des érythroblastes. En 1998, il a été démontré que l'ajout d'un anticorps anti-Emp dans les cultures *in vitro* d'érythroblastes primaires induit une réduction majeure de la prolifération, maturation et énucléation érythroïdes accompagnée d'une forte augmentation de l'apoptose (Hanspal et al., 1998). De plus, les souris invalidées pour le gène *Emp* développent une anémie sévère létale au stade fœtal, ce qui souligne le rôle crucial d'Emp durant l'érythropoïèse définitive (Soni et al., 2006). Les macrophages de ces souris sont phénotypiquement immatures, avec une petite taille, une forme ronde et peu d'extensions cytoplasmiques. D'autres molécules d'adhésion ont été identifiées. Par exemple, l'intégrine $\alpha 4\beta 1$ et le récepteur B4 de l'éphrine (EPHB4), présents sur la membrane plasmique des érythroblastes, interagissent respectivement avec le récepteur VCAM-1 et l'éphrine 2, exprimés à la surface des macrophages. Le rôle précis de ces molécules d'adhésion n'a pas encore été clairement élucidé. Il est probable qu'elles déclenchent des voies de signalisation intracellulaire régulant l'expression des gènes.

Grâce à son contact direct avec les érythroblastes, le macrophage soutient activement l'érythropoïèse terminale. Il **phagocyte** d'une part le **noyau expulsé** des érythroblastes acidophiles (le pyrénocyte) et dégrade l'ADN à l'aide de la **désoxyribonucléase II** (DNase II). Les souris KO pour la DNase II présentent une anémie sévère car il n'y a pas d'énucléation des érythroblastes dans le foie fœtal, et donc pas de production de globules rouges anucléés. Elles meurent *in utero* (Kawane et al., 2001). La DNase II est donc indispensable à l'érythropoïèse définitive. D'après des travaux plus récents, il est désormais établi que les érythroblastes primitifs peuvent être énucléés dans la circulation sanguine, c'est-à-dire en dehors de la niche érythroïde médullaire (Kingsley et al., 2004). Les macrophages du placenta seraient dans ce cas impliqués dans le processus d'énucléation des érythroblastes primitifs (Van Handel et al., 2010).

D'autre part, le macrophage sécrète des **cytokines pro-inflammatoires** capables de **réguler l'érythropoïèse**. Parmi ces cytokines, le **TNF α** et l'**interféron γ** (IFN γ) induisent tous les deux l'apoptose des cellules érythroïdes et limitent ainsi leur prolifération. En effet, l'IFN γ augmente entre autres l'expression du facteur soluble TRAIL (pour *Tumor-necrosis-factor Related Apoptosis Inducing Ligand*) qui appartient à la superfamille du TNF α (Felli et al., 2005). En absence d'EPO, la liaison du TNF α et du TRAIL à leurs récepteurs présents à la

surface des érythroblastes immatures (Pro-E et Baso-E) entraîne l'activation des caspases responsables du clivage de GATA1 (De Maria et al., 1999; Zamai et al., 2000). Pour rappel, au sein des îlots érythroblastiques, l'interaction entre FasL exprimé par les érythroblastes matures (Poly-E et Acido-E) et Fas exprimé par les érythroblastes immatures aboutit également au clivage de GATA1 lorsque les concentrations en EPO sont faibles (De Maria et al., 1999) (cf. Chapitre I-I-C-3-c).

Enfin, le macrophage participe au **maintien de l'homéostasie du fer**. Il sécrète l'**interleukine 6** (IL6), une cytokine pro-inflammatoire qui diminue la quantité de fer disponible (cf. Chapitre I-I-E-4-b-ii). Le macrophage peut aussi, à l'inverse, apporter du fer aux cellules érythroïdes en cours de maturation. Comme je l'ai précédemment expliqué, les macrophages de la rate phagocytent les érythrocytes sénescents. À l'intérieur du phagolysosome, le fer est dissocié de l'hémoglobine pour être stocké dans la ferritine cytoplasmique des macrophages. En cas de besoin, le fer est libéré dans la circulation sanguine grâce à la ferroportine et pris en charge par la transferrine. Toutefois, certains travaux suggèrent que le macrophage présent au centre de l'îlot érythroblastique serait capable de **transférer directement du fer sous forme de ferritine** aux précurseurs érythroïdes et non *via* la ferroportine (Bessis and Breton-Gorius, 1962). L'exocytose de la ferritine à partir du macrophage de l'îlot, suivie de sa captation par endocytose par les érythroblastes, a été mise en évidence en 2008. Au sein de la vésicule d'endocytose, l'acidification du milieu et l'action des enzymes protéolytiques assurent la dissociation du fer de la ferritine (Leimberg et al., 2008). Le fer peut alors être utilisé pour la synthèse de l'hème.

4. Régulation du métabolisme du fer

Le métabolisme du fer est régulé à deux niveaux : au **niveau intracellulaire** grâce au système **IRP/IRE** (pour *Iron regulatory protein/Iron responsive element*, protéine régulatrice du fer/élément de réponse au fer), et au **niveau systémique** sous l'action de l'**hepcidine**, une hormone produite par le foie.

a. Au niveau intracellulaire

Le système IRP/IRE permet à chaque cellule de moduler l'entrée, l'utilisation, le stockage et/ou l'export de son fer selon ses propres besoins. Dans le cytoplasme, les IRP (majoritairement IRP1) contrôlent la synthèse protéique de certains gènes impliqués dans

l'homéostasie du fer. Elles reconnaissent puis se fixent sur des structures en tige-boucle : les IRE, situés dans les régions 5' ou 3' UTR des ARNm des gènes cibles.

La structure et la fonction des IRP dépendent de la teneur intracellulaire en fer. Lorsque le pool labile en fer est conséquent, les clusters Fe/S se lient aux IRP ce qui leur confère une activité aconitase (enzyme du cycle de Krebs qui convertit le citrate en isocitrate). À l'inverse, lorsque le pool labile en fer est faible, les IRP perdent leur centre Fe/S et donc, par la même occasion, leur activité aconitase (Haile et al., 1992). Elles peuvent alors se fixer sur les motifs IRE, de façon à rétablir des concentrations optimales en fer. Les IRP se fixent d'une part sur les IRE localisés dans la région **5' UTR** de l'**ALAS2**, la **ferritine** (Rouault et al., 1988) et la **ferroportine**, **bloquant ainsi leur traduction** par encombrement stérique (Muckenthaler et al., 1998). L'utilisation, le stockage et la sortie extracellulaire du fer sont ainsi réduits. D'autre part, les IRP se fixent sur les IRE localisés dans la région **3' UTR** de **TFRC** (Casey et al., 1989) et **DMT1** (Hubert and Hentze, 2002), ce qui protège leur ARNm d'une éventuelle dégradation par les ribonucléases cytosoliques et **favorise leur traduction** (Leipuviene and Theil, 2007). Il en résulte une augmentation de l'entrée du fer dans le cytosol des cellules (pour revue Omar et al., 2006).

b. Au niveau systémique

La surcharge martiale étant toxique pour l'organisme, un processus de régulation négative des concentrations plasmatiques en fer est nécessaire. À ce jour, aucune voie physiologique d'excrétion du fer n'est connue. Plutôt que par élimination, le taux de fer est contrôlé au niveau de sa libération dans la circulation sanguine grâce à une hormone produite par le foie, l'hépcidine.

i. Mécanisme d'action de l'hépcidine

L'hépcidine est un peptide antimicrobien de 25 acides aminés codé par le gène *HAMP* (Park et al., 2001). La première connexion entre l'hépcidine et le métabolisme du fer a été établie suite aux travaux de Pigeon *et al.*, qui recherchaient les gènes surexprimés dans le foie des souris avec surcharge en fer (Pigeon et al., 2001). Ils ont démontré que l'expression du transcrite de l'hépcidine est induite par le fer mais également par l'inflammation. Peu de temps après, l'équipe de Sophie Vaulont confirme la fonction de l'hépcidine en tant que régulateur du fer dans l'organisme, puisque la diminution de l'expression de l'hépcidine chez les souris KO pour *Usp2* cause une surcharge en fer (Nicolas et al., 2001). Le gène *Usp2* est adjacent au

gène *Hamp*, et l'inhibition de l'expression d'*Usp2* inhibe également celle de *Hamp*. Cette surcharge en fer, principalement localisée dans les cellules parenchymateuses du foie et du pancréas, résulte d'une importante absorption intestinale associée à une libération massive du fer stocké dans les macrophages du SRE. Par la suite, cette même équipe a montré que le croisement de souris atteintes d'une surcharge martiale – par invalidation du gène *Hfe* (pour *homeostatic iron regulator*) – avec des souris atteintes d'une carence martiale – par surexpression du gène *Hamp* – permet de rétablir des taux physiologiques de fer (Nicolas et al., 2002, 2003). Le gène *HFE* code pour une protéine membranaire qui interagit, entre autres, avec TFR1 à la surface du pôle basolatéral des entérocytes immatures – dits cryptiques –, et avec TFR1 ou TFR2 à la surface des hépatocytes.

L'hepcidine réduit l'export cellulaire du fer vers le plasma sanguin. Elle se lie à la ferroportine et contrôle l'homéostasie du fer selon deux mécanismes distincts élucidés par l'équipe de Tomas Ganz et Elizabeta Nemeth. Elle induit d'une part un changement conformationnel de la ferroportine, ce qui entraîne son **ubiquitylation** puis son internalisation et sa dégradation (Nemeth et al., 2004a; Qiao et al., 2012). D'autre part, la liaison de l'hepcidine à la ferroportine obstrue directement le passage du fer (Aschemeyer et al., 2018). Pour rappel, la ferroportine est exprimée à la surface des macrophages du SRE, des hépatocytes et au niveau du pôle basolatéral des entérocytes duodénaux. Sous l'action de l'hepcidine, le fer reste donc stocké dans les cellules et la concentration plasmatique en holotransferrine, reflétant celle du fer, diminue (Figure 17).

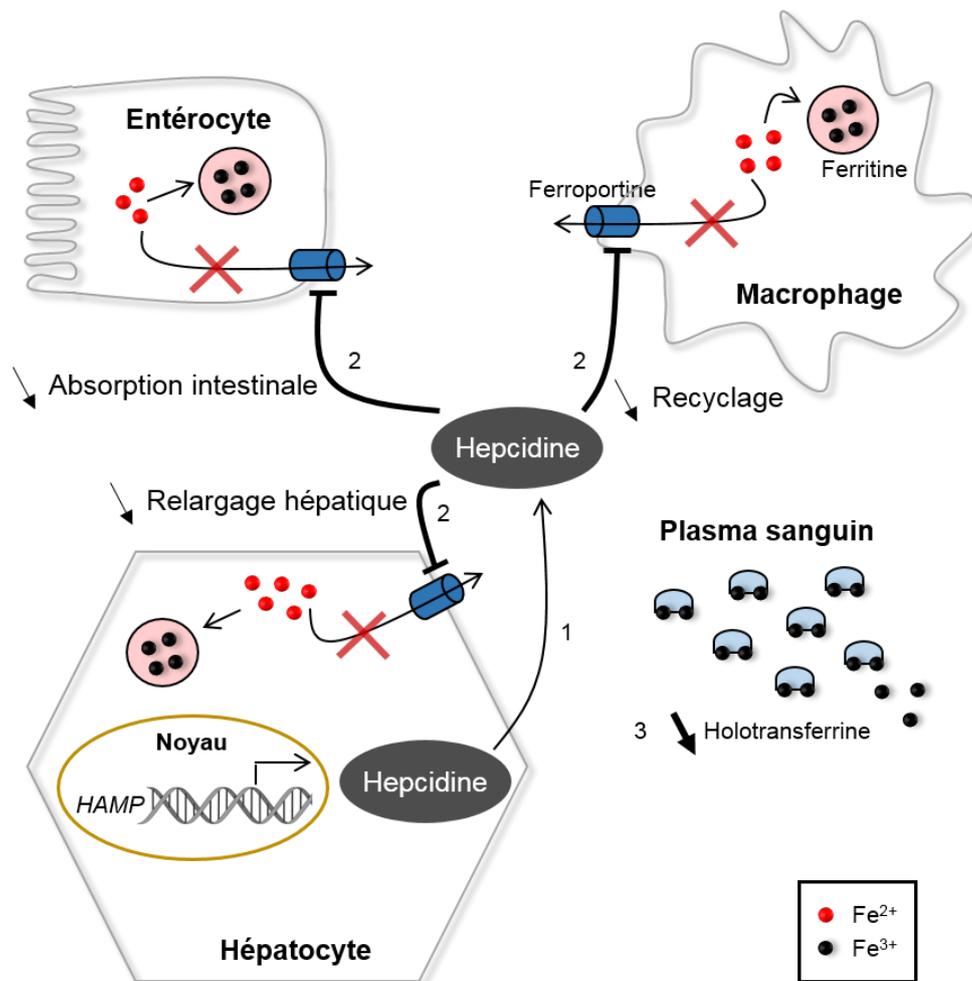


Figure 17 : L'hepcidine régule l'homéostasie du fer

(1) L'hepcidine est sécrétée dans le sang par les hépatocytes. (2) Elle se fixe à la ferroportine et limite la sortie extracellulaire du fer à trois niveaux différents : l'absorption intestinale, le relargage hépatique et le recyclage par les macrophages du SRE. (3) La biodisponibilité du fer diminue.

ii. Régulation de l'expression de l'hepcidine

Le mécanisme exact d'activation des voies de signalisation aboutissant à la transcription du gène *HAMP* n'est pas encore clairement compris. Au cours de la dernière décennie, de nombreux travaux ont permis de mieux expliquer le contrôle de l'expression de l'hepcidine par le **fer** et l'**inflammation**. De plus, très récemment, des **régulateurs érythroïdes** produits par les érythroblastes médullaires ont été identifiés. En fonction de la demande érythropoïétique, ces protéines modulent l'homéostasie du fer.

- Régulateurs non érythroïdes : le fer et l'inflammation

L'expression de l'hepcidine dépend tout d'abord du statut en fer de l'organisme : elle augmente lorsque les concentrations plasmatiques et les réserves hépatiques en fer sont élevées, elle diminue en cas de déficit en fer. Il existe deux senseurs des taux de fer dans l'organisme : l'**holotransferrine** – senseur systémique – et les **BMPs** (pour *Bone morphogenetic proteins*) – senseurs locaux – qui sont sécrétés sous forme d'hétérodimères BMP2/BMP6, ou d'homodimère BMP2/BMP2 et BMP6/BMP6 par les cellules endothéliales sinusoïdales du foie (Canali et al., 2017; Koch et al., 2017; Lin et al., 2007). L'apport alimentaire en fer entraîne une hausse des taux d'holotransferrine circulante ainsi qu'une hausse de la transcription hépatique des BMPs. À la surface de l'hépatocyte, l'holotransferrine se lie à ses récepteurs **TFR1** et **TFR2** tandis que les BMPs se fixent sur un complexe formé de deux sous-unités : les **récepteurs aux BMPs** de type I ou II et le co-récepteur **hémoujuvéline** (HJV). Comme vu précédemment, l'affinité de TFR1 pour l'holotransferrine est nettement supérieure à celle de TFR2 (cf. Chapitre I-I-E-2). De plus, la protéine membranaire **HFE** et l'holotransferrine possèdent le même site de fixation sur TFR1, contrairement à TFR2 où elles se fixent sur deux sites distincts. Ainsi, lorsque l'holotransferrine se lie à TFR1, HFE est dissociée de TFR1 et la liaison entre HFE et TFR2 est potentialisée (Goswami and Andrews, 2006). Finalement, les interactions entre l'holotransferrine, HFE et TFR2 ainsi qu'entre les BMPs et le complexe récepteur aux BMPs permettraient l'activation de voies de signalisation convergeant vers la phosphorylation des **SMAD1/5/8**. Dans le cas de TFR2, la phosphorylation des SMAD interviendrait en aval de la voie ERK1/2 (Ramey et al., 2009). SMAD1/5/8 phosphorylé forme ensuite, avec **SMAD4**, un dimère transloqué dans le noyau de l'hépatocyte (Babitt et al., 2006; Ramey et al., 2009). Ce dimère agit comme facteur de transcription car il induit l'expression de l'hepcidine en se fixant au niveau des régions promotrices de type BRE (pour *BMP-responsive elements*) du gène *HAMP* (Figure 18).

D'autres protéines modulent la transcription de l'hepcidine au niveau de l'hépatocyte. Par exemple, le gène *TMPRSS6* code pour une sérine protéase transmembranaire – la **matriptase-2** – ayant pour cible l'HJV (Figure 18). Grâce à un test luciférase, l'équipe de Clara Camaschella a montré une inhibition de l'activation de l'hepcidine sous l'action de la matriptase-2 (Silvestri et al., 2008a). Cette dernière est donc un potentiel régulateur négatif de l'expression de l'hepcidine.

Enfin, en cas d'inflammation, diverses cytokines sont libérées et interagissent avec leur récepteur exprimé à la surface des hépatocytes. D'une part, les macrophages sécrètent l'**IL6** qui, après liaison à son récepteur associé à une kinase JAK, active le facteur de transcription STAT3. Après translocation dans le noyau, STAT3 se fixe sur le promoteur du gène *HAMP*, ce qui initie sa transcription (Wrighting and Andrews, 2006). D'autre part, l'**activine B**, un membre de la superfamille du TGF β (pour *Transforming growth factor β*) produite dans le foie, active la transcription de *HAMP* via la phosphorylation et la translocation nucléaire de SMAD1/5/8 (Besson-Fournier et al., 2012). Dans ces conditions, l'activine B ne se fixe pas sur son récepteur conventionnel mais sur le récepteur aux BMP de type I (Figure 18). Sous l'action de l'IL6 et/ou de l'activine B, les patients atteints d'un syndrome inflammatoire développent donc une anémie hyposidérémique par séquestration intramacrophagique du fer, puisque l'expression de l'hepcidine est nettement augmentée (Nemeth et al., 2004b).

- Régulateurs érythroïdes

Lorsque l'activité érythropoïétique augmente, par exemple à la suite d'un saignement, il est indispensable pour l'organisme de rapidement mobiliser du fer et donc de diminuer l'expression de l'hepcidine. Divers acteurs pourraient alors moduler la transcription du gène *HAMP* : le facteur de transcription HIF stabilisé en cas d'hypoxie tissulaire, l'EPO ou des molécules produites par les érythroblastes. Plusieurs études démontrent que la concentration plasmatique en EPO, principal régulateur de l'érythropoïèse, est inversement corrélée à l'expression de l'hepcidine (Ashby et al., 2010; Origa et al., 2007). Cependant, l'EPO, de même qu'HIF, ne supprime pas directement la production hépatique de l'hepcidine. En effet, l'injection d'EPO dans des souris irradiées n'abolit pas la transcription de *HAMP* (Pak et al., 2006a; Vokurka et al., 2006). De plus, lors de l'invalidation *in vivo* de *VHL* – le gène codant pour le complexe responsable de l'ubiquitinylation et la dégradation d'HIF –, les concentrations plasmatiques en EPO augmentent et les taux d'hepcidine baissent uniquement chez les souris dont l'érythropoïèse est préservée (Liu et al., 2012). Après administration de carboplatine – un composé chimique qui, comme les rayons γ , entraîne une ablation de la moelle osseuse – aux souris *Vhl*^{-/-}, l'expression de l'hepcidine n'est pas diminuée malgré l'élévation drastique des concentrations en EPO. L'ensemble de ces travaux indique donc que les régulateurs érythroïdes de l'hepcidine sont en fait des facteurs circulants sécrétés par les érythroblastes en réponse à l'EPO.

Deux membres de la superfamille du TGF β – GDF15 (pour *Growth differentiation factor 15*) et TWSG1 (pour *Twisted gastrulation 1*) – ont dans un premier temps été proposés comme les potentiels régulateurs érythroïdes de l'expression de l'hepcidine. Synthétisés par de nombreux tissus, dont les érythroblastes médullaires, GDF15 et TWSG1 suppriment la transcription du gène *HAMP* dans les cultures *in vitro* d'hépatocytes humains primaires (Tanno et al., 2007, 2009). Toutefois, le rôle *in vivo* de ces deux protéines est beaucoup moins clair. D'après les résultats de RT-qPCR, l'expression de *Twsg1* dans la moelle osseuse des souris n'est pas augmentée à la suite d'une saignée par phlébotomie (Casanovas et al., 2013). La transcription de *Gdf15* est quant à elle induite, mais les niveaux d'expression de *Hamp* sont identiques entre les souris sauvages et KO pour *Gdf15*. Ainsi, lors d'un stress érythropoïétique, GDF15 ne régule pas directement l'homéostasie du fer.

En 2014, l'équipe de Tomas Ganz a identifié un nouveau régulateur érythroïde de l'expression de l'hepcidine : l'**érythroferrone (ERFE)**, une hormone de 37 kDa et 354 acides aminés codée par le gène *FAM132B* (ou *ERFE*) qui, chez l'homme, se situe sur le chromosome 2 (Kautz et al., 2014). ERFE appartient à la famille des *C1q/TNF-related proteins* (CTRPs), un ensemble de protéines dont le domaine globulaire en C-terminal présente une homologie de séquence avec la protéine du complément C1q (Seldin et al., 2014). Outre ce domaine C-terminal, les CTRPs partagent diverses caractéristiques structurales : un peptide signal de sécrétion, un domaine N-terminal composé de résidus cystéine hautement conservés (Cys) et un domaine semblable au collagène avec des répétitions Glycine-X-Y. Grâce à leurs résidus Cys, les CTRPs forment des complexes multimériques par l'intermédiaire de ponts disulfures. De plus, les résidus proline et lysine au sein du domaine N-terminal subissent des modifications post-traductionnelles de type hydroxylation et glycosylation respectivement.

D'après les travaux *in vivo* de Kautz *et al.* (2014), ERFE s'exprime à l'état basal dans plusieurs tissus tels que le côlon, les muscles et la moelle osseuse. Toutefois, après administration d'**EPO** ou induction d'un **stress érythropoïétique** par saignée, la transcription d'*Erfe* augmente uniquement dans la **moelle osseuse** et la rate des souris, qui deviennent alors les organes majeurs de production d'ERFE. Pour rappel, chez la souris, la rate est le siège de l'érythropoïèse de stress, ce qui suggère donc qu'ERFE ait une spécificité de fonction érythroïde. Plus précisément, les quantités les plus abondantes d'ARNm de *Fam132b* sont retrouvées dans les précurseurs érythroïdes **Baso-E** et **Poly-E** quinze heures après saignée.

Étant donné la présence d'éléments de réponse à STAT5 au niveau des régions promotrices du gène *Fam132b*, il est probable que la transcription de ce dernier soit initiée en aval du récepteur **EPOR** via la voie de signalisation **JAK2/STAT5**. Le traitement *in vitro* des érythroblastes murins avec un inhibiteur de STAT5 confirme cette hypothèse, car la synthèse d'*Erfe* est alors abolie après stimulation à l'EPO.

ERFE agit à distance sur les **hépatocytes** et réprime la transcription de *Hamp* en cas de stress érythropoïétique (Figure 18). En effet, les souris invalidées pour *Fam132b* sont incapables de réduire leurs taux d'hepcidine en réponse à la saignée. Au contraire, l'injection de la protéine ERFE recombinante – produite dans la lignée HEK 293T – dans des souris sauvages entraîne une diminution rapide et importante de l'expression de *Hamp*, qui est indépendante d'un potentiel effet inflammatoire puisque la voie des BMP n'est pas activée. Enfin, après traitement *in vitro* des hépatocytes murins primaires avec le surnageant des HEK 293T surexprimant ERFE recombinante, la transcription de *Hamp* est là encore réduite. ERFE module donc les taux d'hepcidine directement au niveau du foie. **L'étude de Kautz et al. (2014) souligne ainsi le rôle fondamental d'ERFE dans le contrôle de l'expression de l'hepcidine – et donc de l'homéostasie du fer – lorsque la demande érythropoïétique est accrue** (Figure 19). Les paramètres hématologiques des souris *Fam132b*^{-/-} étant quasi-identiques à ceux des souris sauvages, l'action d'ERFE serait restreinte aux situations d'érythropoïèse de stress. Existe-t-il un même régulateur érythroïde de l'expression de l'hepcidine dans les situations physiologiques et pathologiques ou plusieurs régulateurs selon les situations ?

Il a récemment été montré qu'ERFE diminue l'expression de l'hepcidine par l'intermédiaire de la voie de signalisation **BMP/SMAD** (Arezes et al., 2018). ERFE se lie en fait à diverses BMPs, dont BMP6 qui ne peut alors plus se fixer sur son récepteur. *In fine* la phosphorylation des SMAD1/5/8 est réduite. Toutefois, divers points restent encore à éclaircir : ERFE n'inhibant pas BMP2, l'abolition de la transcription du gène *HAMP* nécessite-t-elle d'autres partenaires ? La protéine ERFE a-t-elle un récepteur au niveau hépatique ? Si oui, quel est-il ? Son activation induit-elle des voies de signalisation indépendantes de celles activées par les BMPs ? À noter que les conclusions des études de Kautz et al. et Arezes et al. sont tirées de modèles murins.

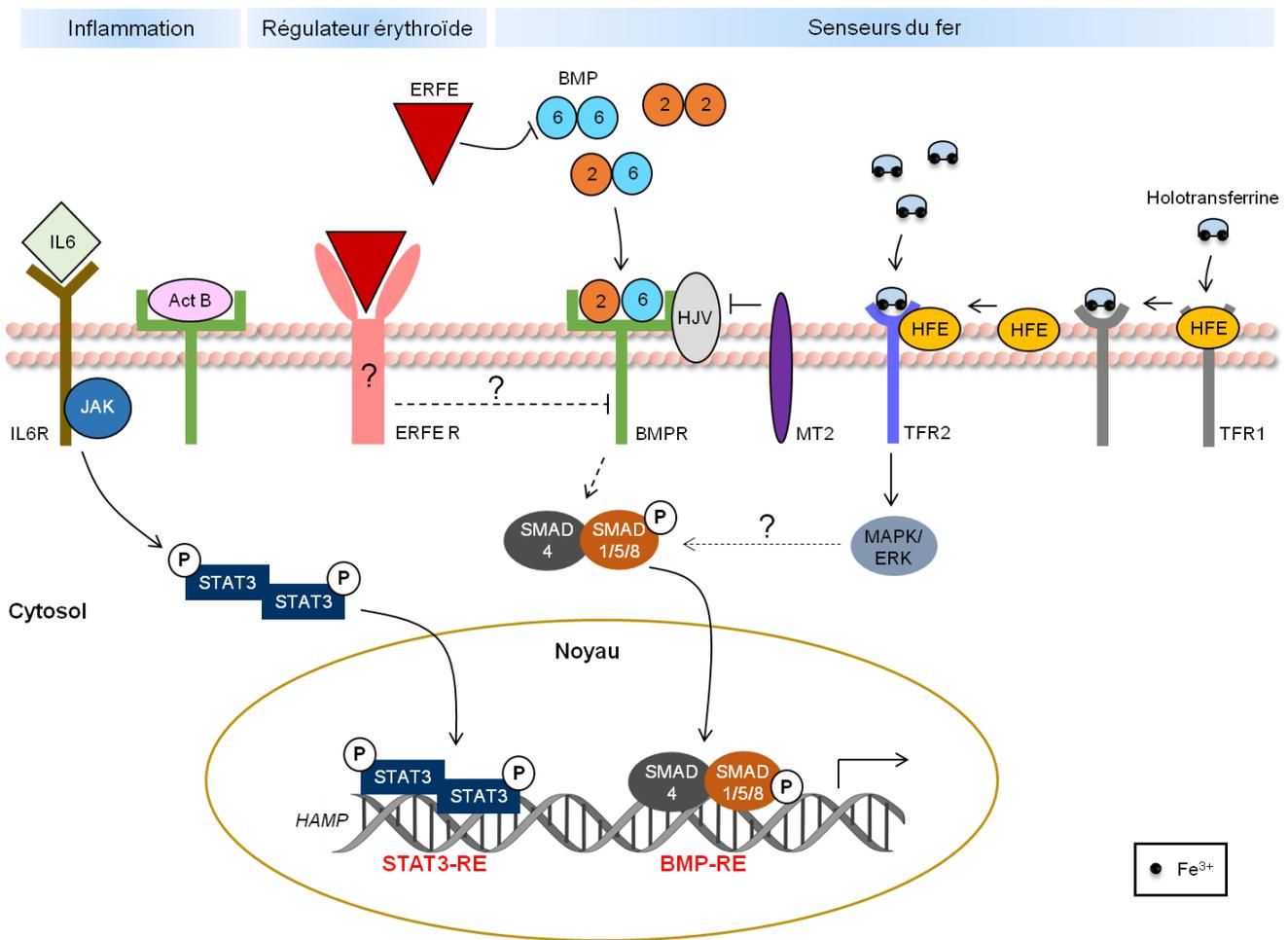


Figure 18 : Régulation de l'expression du gène *HAMP* dans les hépatocytes

La transcription du gène *HAMP* codant pour l'hepcidine est activée par divers stimuli : les taux élevés de fer dans l'organisme qui induisent d'une part la sécrétion des BMP2 et BMP6 par les cellules endothéliales sinusoidales, et d'autre part la saturation de la transferrine alors capable de se fixer sur ses récepteurs TFR1 et TFR2 ; l'inflammation responsable de la production des cytokines IL6 et activine B ; et l'induction d'un stress érythropoïétique qui s'accompagne d'une augmentation de la libération d'ERFE dans la circulation sanguine. L'holotransferrine, les BMPs et les cytokines pro-inflammatoires possèdent toutes un récepteur hépatocytaire identifié. La liaison de chacune de ces molécules à son récepteur active des voies de signalisation qui initient, *in fine*, la transcription de *HAMP*. Concernant ERFE, bien que son action directe sur l'expression de l'hepcidine soit probablement en partie due à la rétention des BMP6, son récepteur au niveau du foie et les potentielles voies de signalisation induites en aval ne sont à ce jour pas connus. **Act B** = Activine B, **BMPR** = récepteur aux BMP de type I, **MT2** = matriptase-2, **RE** = élément de réponse.

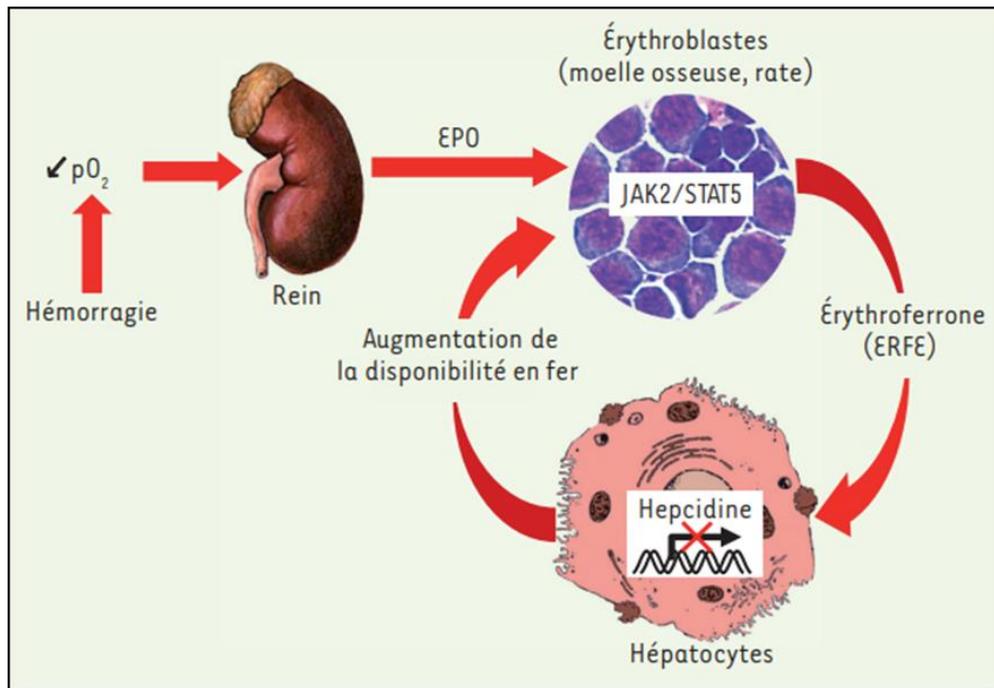


Figure 19 : Mécanisme d'action de l'érythroferrone (d'après Kautz, 2014)

Lors d'un stress érythropoïétique – par exemple une hémorragie – la pression partielle en O₂ diminue et la synthèse d'EPO est alors stimulée au niveau rénal. L'EPO se fixe à son récepteur EPOR présent à la surface des érythroblastes de la moelle osseuse (et de la rate chez la souris), ce qui active la voie de signalisation JAK2/STAT5 et, *in fine*, la production d'ERFE. ERFE est une hormone qui agit sur les hépatocytes en réprimant la transcription du gène codant pour l'hépcidine. En conséquence, la biodisponibilité du fer augmente de façon à répondre à l'intensification de l'érythropoïèse. **pO₂** = pression partielle en dioxygène, **EPO** = érythropoïétine.

Pour conclure, la régulation de l'expression de l'hépcidine – chef d'orchestre de l'homéostasie du fer – fait intervenir de nombreux facteurs, pouvant probablement agir simultanément sur l'hépatocyte en fonction des circonstances physio-pathologiques. Parmi ces facteurs, l'érythroferrone apparaît aujourd'hui comme le meilleur candidat érythroïde qui fait le lien entre deux processus complexes : l'érythropoïèse et le métabolisme du fer. Dans la ligne des travaux récents de l'équipe de Tomas Ganz, l'hypothèse du rôle d'ERFE dans les pathologies associant dysérythropoïèse et surcharge martiale, telles que les syndromes myélodysplasiques avec sidéroblastes en couronne, fait sens.

Tableau 1 : Principaux acteurs du métabolisme du fer

Protéine	Répartition tissulaire et/ou cellulaire	Répartition sub-cellulaire	Fonction
DMT1	Intestins (pôle apical des entérocytes duodénaux)	Membrane plasmique	Absorption intestinale du fer ferreux
	Ubiquitaire (vacuole d'endocytose)	Membrane plasmique	Transfert du fer ferreux de la vacuole d'endocytose au cytosol
Erythroferrone	Plasma sanguin, produite surtout par les érythroblastes médullaires lors d'un stress érythropoïétique		Régulation de l' expression de l' hepcidine en réponse à l'EPO
Ferritine	Ubiquitaire, surtout dans les macrophages du SRE, le foie (hépatocytes) et les intestins	Cytoplasme	Stockage intracellulaire du fer ferrique
Ferroportine	Intestins (pôle basal des entérocytes duodénaux), macrophages du SRE, foie (hépatocytes)	Membrane plasmique	Export du fer des cellules vers le plasma sanguin
Hémojuvéline	Muscles, foie, cœur	Membrane plasmique	Co-récepteur du récepteur aux BMPs de type I : régulation de l' expression de l' hepcidine en fonction des taux de fer dans l'organisme
HFE	Foie (hépatocytes)	Membrane plasmique	Association avec TFR2 : régulation de l'expression de l' hepcidine en fonction des taux de fer dans l'organisme
	Intestins (pôle basal des cellules cryptiques progénitrices des entérocytes)	Membrane plasmique	Programmation de l'absorption digestive du fer*
Hepcidine	Plasma sanguin, produite par le foie (hépatocytes)		Liaison à la ferroportine : inhibition de l' absorption digestive du fer et de son relargage macrophagique et/ou hépatique
Mitoferrine-1	MO, foie, rate	Membrane mitochondriale interne	Entrée du fer ferreux dans la matrice mitochondriale
Récepteur aux BMPs de type I	Foie (hépatocytes)	Membrane plasmique	Récepteur aux BMPs produits par les cellules sinusoidales du foie : régulation de l' expression de l' hepcidine en fonction des taux de fer dans l'organisme
TFR1	Ubiquitaire, 70 à 80 % dans la MO	Membrane plasmique	Récepteur à l'holotransferrine : entrée du fer lié à la transferrine dans les cellules , surtout dans les érythroblastes médullaires
TFR2	Foie (hépatocytes)	Membrane plasmique	Récepteur à l'holotransferrine : entrée du fer lié à la transferrine dans les hépatocytes
Transferrine	Plasma sanguin, produite par le foie (hépatocytes)		Transport plasmatique et distribution du fer ferrique aux tissus, surtout à la MO

* Il a été proposé que la protéine HFE intervienne dans la régulation de l'entrée du fer au niveau de la cellule cryptique intestinale, ce qui expliquerait alors pourquoi les patients mutés sur *HFE* développent une hyperabsorption pathologique de fer : l'hémochromatose (Roy and Enns, 2000). Cependant, ce modèle a été remis en question depuis la mise en évidence du rôle de la protéine HFE dans l'expression de l'hepcidine.

II. Un exemple d'érythropoïèse pathologique : les syndromes myélodysplasiques avec sidéroblastes en couronne

Les syndromes myélodysplasiques avec sidéroblastes en couronne (SMD-SC) – auparavant nommés anémies sidéroblastiques idiopathiques acquises ou anémies réfractaires avec sidéroblastes en couronne – sont des **hémopathies** du **sujet âgé** associant le dérèglement de deux processus précédemment expliqués et intrinsèquement liés : l'érythropoïèse et le métabolisme du fer au niveau systémique et mitochondrial. En effet, dans la **moelle osseuse**, coexistent une **érythropoïèse inefficace**, c'est-à-dire une incapacité à produire suffisamment de globules rouges malgré une augmentation du nombre de progéniteurs et précurseurs érythroïdes immatures, et une **utilisation inadéquate du fer** pour la synthèse de l'hémoglobine bien que les réserves soient suffisantes voire supérieures aux besoins. Les patients porteurs d'un SMD-SC développent donc progressivement une **anémie** profonde, illustrée par une baisse anormale de l'hémoglobine circulante, qui s'accompagne d'une **surcharge en fer**. Les valeurs standards minimales de l'hémoglobine chez l'adulte sont de 14 g/dL chez l'homme et 12 g/dL chez la femme. Parmi les autres hémopathies caractérisées par une érythropoïèse inefficace et une surcharge systémique en fer, les thalassémies font office d'exemples paradigmatiques. La surcharge en fer dans les mitochondries n'est néanmoins pas observée dans ces anémies héréditaires dues à un déficit de production des chaînes α ou, surtout, β de l'hémoglobine.

Dans cette partie, nous replacerons, dans un premier temps, les SMD-SC au sein des deux grands groupes de pathologies sanguines auxquelles ils appartiennent : les anémies sidéroblastiques et les syndromes myélodysplasiques (SMD). Puis, dans un second temps, la physiopathologie des SMD-SC sera expliquée avant d'aborder les principaux traitements actuellement utilisés en clinique.

A. Les anémies sidéroblastiques

Les anémies sidéroblastiques constituent un ensemble hétérogène de pathologies de la moelle osseuse congénitales ou acquises ayant pour caractéristique commune un défaut d'utilisation du fer mitochondrial malgré des **taux de fer plasmatiques normaux ou élevés**. Finalement, la **synthèse de l'hémoglobine est diminuée** et le fer s'accumule dans les mitochondries des érythroblastes, ce qui forme les **sidéroblastes en couronne**.

1. Les sidéroblastes en couronne

Les sidéroblastes en couronne (SC) sont des érythroblastes médullaires porteurs d'**inclusions ferriques non hémoglobiniques dans leurs mitochondries**. Ils sont exclusivement retrouvés dans des situations pathologiques. Au laboratoire, les SC sont visualisés grâce à la coloration de Perls qui repose sur une réaction cytochimique : sous l'action du ferrocyanure de potassium, le Fe^{3+} anormalement accumulé dans les érythroblastes précipite sous forme de granules de ferrocyanure ferrique bleu-vert. D'après un consensus international, on parle de SC dès lors qu'il y a au moins cinq granules contenant du fer dispersés tout autour du noyau érythroblastique ou plus de cinq granules recouvrant au moins le tiers de la circonférence nucléaire (Figure 20) (Mufti et al., 2008).



Figure 20 : Sidéroblastes en couronne

Frottis médullaire photographié au microscope après coloration de Perls (d'après Malcovati and Cazzola, 2016). Dans les sidéroblastes en couronne, le fer est anormalement stocké sous forme de ferritine mitochondriale.

2. Classification des anémies sidéroblastiques

a. Les anémies sidéroblastiques congénitales

Les anémies sidéroblastiques congénitales sont des pathologies très rares. Elles sont causées par des mutations constitutionnelles touchant des gènes du métabolisme du fer mitochondrial (cf. Chapitre I-I-E-2). Le fer précipite alors dans les mitochondries des érythroblastes et n'est plus disponible pour la synthèse de l'hème au cours de l'érythropoïèse. En conséquence, les patients présentent une anémie microcytaire, c'est-à-dire une anémie avec diminution du volume globulaire moyen (VGM) ($\text{VGM} < 80 \text{ fL}$ chez l'adulte). On distingue deux sous-types d'anémies sidéroblastiques congénitales : les formes syndromiques – qui s'accompagnent le plus souvent d'anomalies du développement, de défauts neurosensoriels et/ou métaboliques – et les formes isolées (Tableau 2).

Parmi les formes non syndromiques, la mieux connue est l'anémie sidéroblastique liée à l'X (XLSA), qui résulte en fait d'une mutation récessive sur le gène *ALAS2* (Prades et al., 1995). Étant donné que ce gène se situe sur le chromosome X, la XLSA touche principalement les garçons. D'autres gènes peuvent également entraîner le développement d'une anémie sidéroblastique isolée lorsqu'ils sont mutés. Par exemple, en 2007, une mutation homozygote du gène *GLRX5* a été rapportée chez un adulte (Camaschella et al., 2007). L'année suivante, une équipe américaine a identifié une mutation dans le promoteur du gène *FECH* codant pour la ferrochélatase, dernière enzyme impliquée dans la synthèse de l'hème, chez un enfant de presque trois ans (Caudill et al., 2008). Ces deux cas restent uniques à ce jour. De plus, en 2009, des mutations non-sens du gène *SLC25A38* ont été mises en évidence au Canada (Guernsey et al., 2009). *SLC25A38* code pour un transporteur d'acides aminés ancré dans la membrane mitochondriale interne. La diminution de l'expression de *SLC25A38* dans les embryons du poisson zèbre – suite à l'injection d'un ARN antisens – induit le développement d'une anémie et suggère ainsi un rôle primordial de ce transporteur dans l'érythropoïèse. Dans cette même étude, il a été montré que l'inactivation de l'orthologue de *SLC25A38* dans la levure entraîne un défaut de la biosynthèse de l'hème corrigé par l'addition d'acide δ-aminolévulinique (ALA) ou de glycine au milieu de culture. Il est donc probable que cette protéine intervienne plus précisément soit dans l'import de la glycine au sein de la mitochondrie, soit dans l'export de l'ALA vers le cytoplasme (ou les deux dans l'hypothèse d'un échange). Enfin, récemment, plusieurs cas de mutations de *HSPA9* – homologue mitochondrial de *HSP70* – ont été identifiés. Pour rappel, la protéine chaperonne HSPA9 intervient dans le transfert à la GLRX5 des clusters Fe/S formés *de novo* (cf. Figure 15) (Schmitz-Abe et al., 2015).

Parmi les formes syndromiques, il existe également une anémie sidéroblastique liée à l'X mais qui, cette fois-ci, est due à une mutation du gène *ABCB7* responsable du développement d'un trouble hématologique et neurologique : la XLSA avec ataxie (XLSA/A) (Allikmets et al., 1999). D'autres pathologies plus complexes associent des déficits neuromusculaires et métaboliques aux anémies sidéroblastiques, telles que la myopathie, acidose lactique et anémie sidéroblastique (MLASA) provoquée par une mutation du gène *PUS1* codant pour la pseudo-uridylylase 1 dont la fonction dans le métabolisme du fer n'est pas encore connue.

Tableau 2 : Principales mutations responsables du développement d'une anémie sidéroblastique congénitale

Forme	Gène muté	Année d'identification de la mutation	Fonction physiologique de la protéine codée par le gène muté
Non syndromique	<i>ALAS2</i>	1995	Catalyse la première étape de la biosynthèse de l'hème
	<i>GLRX5</i>	2007	Probablement impliquée dans les dernières étapes de la maturation des clusters Fe/S
	<i>FECH</i>	2008	Catalyse la dernière étape de la biosynthèse de l'hème
	<i>SLC25A38</i>	2009	Probablement impliquée dans le transport de la glycine et/ou de l'ALA à travers la membrane mitochondriale interne
	<i>HSPA9</i>	2015	Participe au transfert des clusters Fe/S sur GLRX5
Syndromique	<i>ABCB7</i>	1999	Probablement impliquée dans l'export des clusters Fe/S à travers la membrane mitochondriale interne
	<i>PUS1</i>	2004	Assure la trans-glycosylation post-transcriptionnelle des ARNs ribosomiques et des ARNs de transfert, fonction dans le métabolisme du fer mitochondrial non connue

L'ensemble des mutations actuellement identifiées chez les patients porteurs d'une anémie sidéroblastique congénitale touche des gènes du métabolisme du fer mitochondrial, dont la fonction s'avère essentielle à l'érythropoïèse. Les gènes écrits en **bleu** correspondent à ceux codant pour des protéines impliquées dans la **biosynthèse de l'hème**, les gènes écrits en **rouge** correspondent à ceux codant pour des protéines impliquées dans la **biosynthèse des protéines à clusters Fe/S**. **ALA** = acide δ -aminolévulinique.

b. Les anémies sidéroblastiques acquises

Les anémies sidéroblastiques acquises font généralement partie d'un **syndrome myélodysplasique**, hémopathie clonale détaillée dans le chapitre I-II-B. Certaines d'entre elles sont parfois secondaires à une hypothermie, la prise d'alcool ou d'un médicament, tels que le chloramphénicol et des antituberculeux comme la cyclosérine, l'isoniazide et le pyrazinamide. Elles sont alors réversibles, contrairement aux SMD.

Aucune mutation des gènes *ALAS2*, *GLRX5*, *FECH*, *SLC25A38*, *HSPA9*, *ABCB7* et *PUS1* n'a été mise en évidence dans les SMD. Toutefois, les cellules CD34⁺ isolées à partir de la moelle osseuse des patients porteurs d'une anémie sidéroblastique acquise présentent un profil d'expression génique différent des CD34⁺ de témoins sains, puisque certains gènes impliqués dans le métabolisme de l'hème – dont *ALAS2* et *FECH* – ont une expression augmentée alors que l'expression d'*ABCB7* est quant à elle diminuée (Boulwood et al., 2008; Pellagatti et al., 2006). De plus, le gène *HSPA9* situé en position 5q31.2 est délété dans les SMD avec del(5q) (cf. Chapitre I-II-B-2-a) (Horrigan et al., 2000). En 2011, Chen *et al.* ont

transduit des progéniteurs hématopoïétiques humains CD34⁺ avec un shHSPA9. Les cellules ont ensuite été cultivées *in vitro* en conditions de différenciation érythroïde, granuleuse ou mégacaryocytaire. Seule l'érythropoïèse est affectée par la diminution de l'expression d'*HSPA9*, avec une réduction drastique du nombre de BFU-E et de précurseurs érythroïdes matures CD71⁺/GPA⁺ (Chen et al., 2011). L'enjeu actuel consiste donc à identifier les mutations somatiques potentiellement responsables de la dérégulation du métabolisme du fer mitochondrial dans les SMD-SC.

B. Les syndromes myélodysplasiques

1. Définition et épidémiologie des SMD

Les SMD regroupent des **hémopathies myéloïdes hétérogènes** qui se caractérisent par une ou plusieurs **cytopénies** contrastant avec une **moelle osseuse normo- ou hypercellulaire et dysplasique**. En effet, bien que les progéniteurs myéloïdes prolifèrent excessivement, l'apoptose des précurseurs aboutit à un défaut de production des cellules matures, d'où la qualification de pathologies avec une « hématopoïèse inefficace ». Les SMD sont des **maladies clonales acquises de la cellule souche hématopoïétique** : le clone porteur des mutations principales présent dans le compartiment médullaire a un avantage de croissance et étouffe petit à petit l'hématopoïèse normale avant d'être retrouvé dans le compartiment périphérique.

L'évolution des SMD est variable d'un sujet à l'autre, les cytopénies pouvant être responsables de complications comme, par exemple, les infections chez les personnes neutropéniques. Entre 30 et 40% des patients progressent vers une **leucémie aiguë myéloblastique (LAM)** secondaire, ce qui fait de ces hémopathies les plus fréquents des états pré-leucémiques de l'adulte. Pour rappel, une LAM est définie par un pourcentage de myéloblastes intramédullaires > 19%. L'espérance de vie d'une LAM secondaire à un SMD est inférieure à six mois.

La médiane d'âge au diagnostic est de l'ordre de **70 ans**. Le taux d'incidence des SMD en Europe se situe environ à 4 personnes pour 100 000 par an, mais il augmente fortement avec l'âge. En effet, ce taux atteint 40 à 50 nouveaux cas sur 100 000 par an chez la population âgée de plus de 70 ans (Neukirchen et al., 2011). À l'inverse, le taux d'incidence des myélodysplasies infantiles est de l'ordre de 1 à 4 enfants sur 1 million.

La grande majorité des SMD sont des **pathologies idiopathiques** acquises *de novo*. Seuls 10 à 20% d'entre eux sont secondaires à l'exposition à des chimiothérapies de type agents alkylants ou analogues de la purine, à des radiations ionisantes et/ou à des toxiques environnementaux comme le tabac ou le benzène. Chez l'enfant, les SMD peuvent être secondaires à une maladie génétique telle que l'anémie de Fanconi ou le syndrome de Noonan.

2. Classifications diagnostique et pronostique des SMD

Les classifications diagnostique et pronostique des SMD reposent toutes deux sur des critères biologiques – comme le nombre ou l'intensité des cytopénies et le pourcentage de blastes médullaires – et cytogénétiques. La classification diagnostique a été établie par un groupe d'experts scientifiques de l'Organisation mondiale de la Santé (OMS) en 2001, puis réétudiée en 2008 et 2016. L'évaluation pronostique a quant à elle été mise en place par une équipe américaine et correspond à un système de score : l'IPSS (pour *International prognostic scoring system*), qui a lui aussi été révisé en 2012 (IPSS-R).

a. Classification diagnostique OMS 2016

Les SMD sont des néoplasmes myéloïdes hétérogènes. L'OMS a défini cinq critères permettant la classification diagnostique de ces pathologies (Arber et al., 2016) (Tableau 3) :

- **le nombre de dysplasies intramédullaires,**
- **le nombre de cytopénies périphériques,**
- **le pourcentage de blastes intramédullaires et périphériques,**
- **le pourcentage de sidéroblastes en couronne intramédullaires,**
- **les anomalies du caryotype.**

Le myélogramme constitue l'examen biologique de référence. Il est systématiquement réalisé dans un contexte de suspicion d'un SMD ou d'une autre hémopathie maligne. Il comprend l'évaluation de la richesse médullaire et de la dysmyélopoïèse (nucléaire et cytoplasmique) après coloration MGG, le décompte des blastes également après coloration MGG, et le décompte des SC après coloration de Perls. L'hémogramme détermine quant à lui le nombre et la morphologie des cellules myéloïdes matures circulantes, grâce à une numération formule sanguine et à une coloration MGG respectivement. Les anomalies cytologiques – dans la moelle osseuse et/ou le sang périphérique – peuvent toucher l'une ou

plusieurs des trois lignées myéloïdes : la lignée érythroïde, la lignée granuleuse et la lignée mégacaryocytaire.

Les patients atteints d'un **SMD-SC** présentent une **anémie profonde** (hémoglobine < 8 g/dL), malgré une moelle érythroblastique très riche. Plus de 10% des cellules érythroïdes sont dysplasiques dans la moelle osseuse, ce qui signe la **dysérythroïèse**. Outre le frottis médullaire, l'immunophénotypage peut également – par exemple en cas de morphologie incertaine des cellules – aider au diagnostic de dysérythroïèse. En effet, en 2013, les travaux du service d'Hématologie Biologique de l'Hôpital Cochin ont permis d'établir un score de dysérythroïèse, également appelé RED-score, à partir du taux d'hémoglobine et de deux paramètres cytométriques : le coefficient de variation de l'intensité moyenne de fluorescence du CD36 (CD36 CV) et celui du CD71 (CD71 CV) (Mathis et al., 2013). La baisse du taux d'hémoglobine, associée à l'augmentation du CD36 CV et/ou du CD71 CV, est en faveur d'une dysérythroïèse marquée, caractéristique majeure des SMD. Le pourcentage de **blastés** dans la moelle osseuse reste inférieur à 5%. Les **SC** représentent **15% ou plus des érythroblastés intramédullaires**. La présence de SC est fortement corrélée à la présence d'une mutation somatique hétérozygote du gène *SF3B1* (Papaemmanuil et al., 2011; Yoshida et al., 2011). Selon la classification OMS 2016, le diagnostic des SMD-SC est désormais établi dès lors qu'un individu est porteur d'une **mutation sur SF3B1** avec, dans ce cas, **au moins 5% de SC médullaires** (Arber et al., 2016). Enfin, les anomalies cytogénétiques sont rares.

Tableau 3 : Classification diagnostique des SMD selon l'OMS

Nom	Nombre de dysplasies médullaires*	Nombre de cytopénies périphériques**	% de SC	% de blastes médullaires (BM) et périphériques (BP)	Cytogénétique
SMD avec dysplasie unilignée (SMD-DUL)	1	1 ou 2	<15%/<5% †	BM <5%, BP <1%, pas de corps d'Auer***	Tous les caryotypes, sauf del(5q)
SMD avec dysplasie multilignée (SMD-DML)	2 ou 3	1-3	<15%/<5% †	BM <5%, BP <1%, pas de corps d'Auer	Tous les caryotypes, sauf del(5q)
SMD avec sidéroblastes en couronne (SMD-SC)					
SMD-SC-DUL	1	1 ou 2	≥15%/≥5% †	BM <5%, BP <1%, pas de corps d'Auer***	Tous les caryotypes, sauf del(5q)
SMD-SC-DML	2 ou 3	1-3	≥15%/≥5% †	BM <5%, BP <1%, pas de corps d'Auer	Tous les caryotypes, sauf del(5q)
SMD avec del(5q) isolée	1-3	1-2	-	BM <5%, BP <1%, pas de corps d'Auer	del(5q) isolée ou avec 1 autre anomalie, sauf -7 ou del(7q)
SMD avec excès de blastes (SMD-EB)					
SMD-EB1	0-3	1-3	-	BM 5-9%, BP 2-4%, pas de corps d'Auer	Tous les caryotypes
SMD-EB2	0-3	1-3	-	BM 10-19%, BP 5-19% ou corps d'Auer	Tous les caryotypes
SMD inclassables (SMD-I)					
Avec 1% de BP	1-3	1-3	-	BM <5%, BP = 1%, pas de corps d'Auer	Tous les caryotypes
Avec DUL et pancytopenie	1	3	-	BM <5%, BP <1%, pas de corps d'Auer	Tous les caryotypes
Avec anomalie du caryotype	0	1-3	<15%	BM <5%, BP <1%, pas de corps d'Auer	Anomalies du caryotype caractéristiques des SMD****
Cytopénies réfractaires de l'enfant	1-3	1-3	-	BM <5%, BP <2%	Tous les caryotypes

* Dysplasie médullaire : ≥ 10% des cellules médullaires d'une ou de plusieurs lignée(s) myéloïde(s) sont dysplasiques,

** Cytopénie périphérique : hémoglobine < 10 g/dL ; plaquettes < 100 G/L ; polynucléaires neutrophiles < 1,8 G/L,

*** Corps d'Auer : structures en aiguille allongées (3-10 µm) et fines (< 0,5 µm), provenant de la fusion de plusieurs granulations azurophiles et observables dans les blastes de LAM,

**** Anomalies du caryotype caractéristiques des SMD : cf. Tableau 4 (Évaluation pronostique des SMD selon l'IPSS-R).

† Si le patient est porteur d'une mutation sur le gène *SF3B1*.

del(5q) = délétion du bras long du chromosome 5, **-7** = monosomie 7, **del(7q)** = délétion du bras long du chromosome 7.

b. Classification pronostique IPSS-R

À la fin des années 90, une équipe américaine a effectué une analyse multivariée sur une cohorte de plusieurs centaines de malades atteints d'un SMD primaire ou ayant évolué en LAM (Greenberg et al., 1997). Cette évaluation a permis l'élaboration de l'IPSS et a identifié les principales variables influençant le pronostic des SMD :

- **le pourcentage de blastes intramédullaires,**
- **les anomalies du caryotype,**
- **les cytopénies.**

En 2012, la même équipe dévoile une classification IPSS révisée (IPSS-R), qui prend notamment en compte un nouveau regroupement cytogénétique (Tableau 4) (Greenberg et al., 2012; Schanz et al., 2012). L'IPSS-R sépare les SMD en sept catégories de risque, du plus faible (score 0) au plus élevé (score ≥ 4.5). L'objectif majeur de ce travail consiste à identifier rapidement les SMD non traités de mauvais pronostic – c'est-à-dire à haut risque de transformation en LAM – afin de guider la prise en charge thérapeutique. Il est important de préciser que la présence d'une ou plusieurs mutations somatiques n'est pas prise en considération dans le score pronostique. Pourtant, en 2015, il a été montré que certaines d'entre elles induisent un changement de l'expression des gènes au niveau de l'ARNm dans les cellules CD34⁺ de patients. Lorsque les données du transcriptome sont corrélées aux paramètres clinico-biologiques du patient, leur analyse statistique multivariée permet, *in fine*, de prédire la survie sans LAM (Gerstung et al., 2015).

Les **SMD-SC** sont classiquement des pathologies de **bon pronostic** : le caryotype ne présente généralement pas d'anomalie et le pourcentage de blastes intramédullaires reste faible (Tableau 3). Le risque de **progression des SMD-SC vers la LAM** n'est que de **15%**.

Tableau 4 : Évaluation pronostique des SMD selon l'IPSS-R

Variable pronostique	Score						
	0	0.5	1.0	1.5	2.0	3.0	4.0
Cytogénétique*	Très favorable	-	Favorable	-	Intermédiaire	Défavorable	Très défavorable
% de blastes médullaires	≤ 2	-	2-5	5-10	5-10	> 10	-
Hémoglobine (g/dl)	≥ 10	-	8-10	< 8	-	-	-
Plaquettes (G/l)	≥ 100	50-100	< 50	-	-	-	-
PNN (G/l)	≥ 0,8	< 0,8	-	-	-	-	-

*Cytogénétique :

- Très favorable : perte du chromosome Y, délétion du bras long du chromosome 11
- Favorable : caryotype normal, délétion du bras long du chromosome 5 = del(5q), délétion du bras court du chromosome 12, délétion du bras long du chromosome 20, deux anomalies avec del(5q)
- Intermédiaire : délétion du bras long du chromosome 7 = del(7q), gain d'un chromosome 8 (trisomie 8), = gain d'un chromosome 19 (trisomie 19), inversion du bras long du chromosome 17, autre anomalie simple ou double
- Défavorable : perte d'un chromosome 7 (monosomie 7) = -7, inversion du chromosome 3, translocation du bras long du chromosome 3, délétion du bras long du chromosome 3, deux anomalies avec -7/del(7q), caryotype complexe avec 3 anomalies
- Très défavorable : caryotype complexe avec > 3 anomalies

Les groupes de risque sont les suivants : ≤ 1.5 = très faible ; 1.5-3 = faible ; 3-4.5 = intermédiaire ; 4.5-6 = élevé ; > 6 = très élevé. **PNN** = polynucléaires neutrophiles.

C. Physiopathologie des SMD-SC

Au cours des vingt dernières années, de nombreuses études ont permis de mieux comprendre la physiopathologie des SMD-SC. Dans ce chapitre, nous reviendrons dans un premier temps sur les différents facteurs potentiellement responsables de la **dysérythroïèse terminale**, tels que l'**apoptose mitochondriale** et l'**accumulation de fer dans les mitochondries**. Puis, dans un second temps, la **surcharge en fer**, au niveau **systémique** cette fois-ci, sera expliquée.

1. Dysérythroïèse terminale

a. Apoptose mitochondriale des érythroblastes

La dysérythroïèse terminale correspond à une anomalie de maturation des érythroblastes dans la moelle osseuse, aboutissant à une insuffisance de production des globules rouges. Les **précurseurs érythroïdes matures** des patients porteurs d'un SMD-SC subissent en fait une apoptose mitochondriale massive dont l'origine exacte n'est pas encore parfaitement connue. En effet, en 2000, il a tout d'abord été montré une augmentation de l'activité de la caspase-3 dans les cellules mononucléées (CMN) de la moelle osseuse des patients atteints d'une anémie sidéroblastique acquise (Matthes et al., 2000). Cette suractivité s'accompagne d'une diminution du potentiel transmembranaire mitochondrial et d'une augmentation du nombre de CMN marquées par l'annexine V, protéine à forte affinité pour la phosphatidylsérine externalisée à la surface des cellules en début d'apoptose. Il est intéressant de souligner que la diminution du potentiel transmembranaire mitochondrial a aussi été observée dans les CMN des patients atteints d'une anémie sidéroblastique congénitale. L'équipe d'Eva Hellström-Lindberg a révélé, grâce à une technique d'IF, une augmentation de la libération spontanée du cytochrome *c* de l'espace intermembranaire mitochondrial vers le cytosol dans les cellules des patients porteurs d'un SMD-SC (Tehranchi et al., 2003, 2005). Cette augmentation est visible dès le stade immature CD34⁺ et s'accroît tout au long de la culture en conditions de différenciation érythroïde. Pour rappel, le cytochrome *c* participe à la formation de l'apoptosome qui active ensuite la caspase-9. La caspase-9 active *in fine* la caspase-3, enzyme effectrice de l'apoptose. Les RT-qPCR, à partir des ARN isolés des progéniteurs CD34⁺ de SMD-SC, ont permis de confirmer la surexpression du cytochrome *c* mais également d'autres facteurs pro-apoptotiques, dont Bax et Bid. Ces trois molécules sont très peu exprimées dans les cellules CD34⁺ isolées à partir de témoins sains.

Les caspases sont également activées *via* la voie extrinsèque de l'apoptose. Notre équipe a mis en évidence une augmentation de l'expression du récepteur à domaine de mort Fas et de son ligand FasL dans les cellules cultivées *in vitro* après tri des progéniteurs multipotents CD34⁺ de la moelle osseuse des patients (Claessens et al., 2002). L'expression de Fas est plus importante dans les progéniteurs et précurseurs immatures, tandis que celle de FasL augmente de façon conséquente en fin de différenciation érythrocytaire. Enfin, toujours dans les érythroblastes des patients en culture, il a été montré une absence de colocalisation nucléaire de la protéine chaperonne HSP70 avec le facteur de transcription GATA1 (Frisan et al., 2012). Ce dernier n'est donc plus protégé du clivage par la caspase-3, ce qui participe là encore à la dysérythropoïèse terminale. En 2005, Tehranchi *et al.* avaient quant à eux souligné une expression plus faible du transcrit GATA1 dans les cellules en culture des patients atteints d'un SMD par rapport aux cellules saines (Tehranchi et al., 2005). Les multiples mécanismes entraînant l'apoptose des érythroblastes pathologiques sont illustrés dans la figure 21.

b. Accumulation de fer dans les mitochondries des érythroblastes

Parallèlement à l'augmentation de la libération du cytochrome *c* au cours de la différenciation érythroïde, la ferritine s'accumule dans les mitochondries des érythroblastes en culture issus des SMD-SC (Tehranchi et al., 2005) (Figure 21). Étant donné la forte corrélation d'expression entre la ferritine mitochondriale et le cytochrome *c*, il a été suggéré que l'accumulation de fer ait un effet pro-apoptotique, *via* par exemple la production d'espèces réactives à l'oxygène ou l'altération des fonctions mitochondriales. Toutefois, le lien exact entre la surcharge en fer dans les mitochondries des cellules érythroïdes et l'apoptose mitochondriale n'est à ce jour pas clarifié. Concernant les mécanismes de surcharge mitochondriale en fer, plusieurs hypothèses autour de la protéine ABCB7 ont été émises. Taketani *et al.*, ont d'une part démontré que l'interaction du domaine C-terminal de FECH avec ABCB7 est indispensable au maintien de son centre Fe/S, et donc à la catalyse du transfert du Fe²⁺ sur l'hème (Taketani et al., 2003). Si l'expression d'ABCB7 diminue (cf. Chapitre I-II-A-2-b), la synthèse de l'hémoglobine diminue et les érythroblastes accumulent du fer dans leurs mitochondries. D'autre part, le déficit d'ABCB7 favoriserait la perte du centre Fe/S de la protéine cytosolique IRP1 (Pondarré et al., 2006), qui, je le rappelle, joue un rôle majeur dans la régulation de l'homéostasie du fer intracellulaire (cf. Chapitre I-I-E-4-a). Lorsqu'IRP1 est dépourvue de son centre Fe/S, elle se fixe sur les motifs IRE, ce qui induit entre autres une augmentation de l'entrée du fer dans les cellules par stabilisation de la région

3'UTR des transcrits TFRC et DMT1. Des travaux supplémentaires sont nécessaires à la compréhension du processus de régulation du métabolisme du fer mitochondrial et, donc, au mécanisme d'accumulation de fer dans les mitochondries des érythroblastes des patients porteurs d'un SMD-SC.

2. Surcharge systémique en fer

Outre l'accumulation de fer dans les mitochondries, les patients atteints d'un SMD présentent également une surcharge systémique en fer dès les stades précoces de leur maladie, c'est-à-dire avant traitement de leur anémie par transfusions sanguines (Cortezzi et al., 2000). L'érythropoïèse inefficace s'accompagne d'une augmentation significative des taux de NTBI, le fer sérique non lié à la transferrine. Les anémies sidérolastiques représentent le sous-groupe de SMD le plus touché par cette surcharge en fer. D'après les travaux de Valeria Santini *et al.*, le NTBI mais aussi la ferritine sérique et le coefficient de saturation de la transferrine sont plus élevés par rapport aux autres sous-groupes de SMD et aux contrôles sains (Santini et al., 2011). Parallèlement à ces paramètres biochimiques, les concentrations en hepcidine sont de l'ordre d'1,4 nmol/L dans le sérum des patients porteurs d'un SMD-SC, alors qu'elles dépassent systématiquement 3,5 nmol/L dans le sérum des autres patients SMD et des témoins. En conséquence, **le rapport hepcidine/ferritine diminue** de façon drastique **chez les patients atteints d'une anémie sidérolastique acquise** (en moyenne 1,97 *versus* 52,94 chez les témoins). De manière intéressante, **les taux anormalement bas d'hepcidine sérique sont très souvent associés à la présence d'une mutation somatique sur le gène SF3B1** (Ambaglio et al., 2013).

Parmi les potentiels régulateurs érythroïdes de l'expression de l'hepcidine, GDF15 est dosé dans le sérum des patients et des témoins grâce à un test immuno-enzymatique de type ELISA. Aucune corrélation entre les taux de GDF15 et d'hepcidine n'a été mise en évidence dans l'étude de Santini *et al.* (Santini et al., 2011). En 2007 déjà, Tanno *et al.* n'avaient pas trouvé d'augmentation significative des concentrations de GDF15 dans le sérum des patients atteints d'un SMD (Tanno et al., 2007). En revanche, d'après l'étude d'Ambaglio *et al.*, les patients atteints d'un SMD avec mutation sur le gène *SF3B1* ont des concentrations sériques de GDF15 plus importantes que les patients non mutés. Dans ce contexte, les taux de GDF15 pourraient refléter l'érythropoïèse inefficace caractéristique des patients myélodysplasiques, sans pour autant directement réguler l'expression de l'hepcidine au niveau hépatique. Il s'agit alors d'un épiphénomène. Pour rappel, tout comme les anémies sidérolastiques, les

thalassémies sont des pathologies avec dysérythropoïèse, surcharge systémique en fer et diminution des concentrations sériques d'hepcidine (Nemeth and Ganz, 2006). En ce qui concerne TWSG1, bien que l'ARNm de *TwsG1* soit légèrement augmenté dans la moelle osseuse des souris *Hbb^{th3/th3}* (un modèle de thalassémie), la corrélation avec la diminution de l'expression de l'hepcidine n'a pas été prouvée (Tanno et al., 2009). Le taux d'ARNm d'*Erfe* est lui aussi significativement augmenté dans la moelle osseuse et la rate des souris thalassémiques (*Hbb^{th3/+}*) (Kautz et al., 2014). Il existe cette fois-ci une corrélation entre l'érythroferrone et l'hepcidine, puisque l'inactivation du gène *Erfe* chez ces souris permet de rétablir le niveau physiologique de l'hepcidine et de diminuer la surcharge martiale. Le dosage immuno-enzymatique de l'érythroferrone humaine dans le sérum, récemment mis au point par l'équipe de Tomas Ganz, a été effectué chez les patients thalassémiques (Ganz et al., 2017). Il montre une augmentation significative des concentrations d'ERFE dans le sérum des patients porteurs d'une β -thalassémie et non transfusés, par rapport aux contrôles sains donneurs de sang. L'hepcidine est inversement corrélée à ERFE ($R^2 = 0,7$; p-value < 0,001).

Au moment où débute ce travail, la protéine ERFE n'a pas encore été dosée dans le plasma ou le sérum des patients porteurs d'un SMD.

3. Conclusion

La dysérythropoïèse terminale est responsable du développement de l'anémie dans les SMD-SC et est associée à une surcharge systémique en fer. Les études précédemment détaillées n'abordent cependant pas tous les aspects de la pathologie. La plupart des patients atteints d'un SMD-SC présentent une mutation somatique sur le gène *SF3B1* (Papaemmanuil et al., 2011; Yoshida et al., 2011). Quelle est la relation entre le génotype et le phénotype ? La présence d'une mutation du gène *SF3B1* peut-elle contribuer à la fois à l'expansion du compartiment des progéniteurs et précurseurs érythroïdes immatures et à l'apoptose des précurseurs matures ? Pour tenter de répondre à cette question, j'exposerai dans le chapitre II, les processus biologiques dans lesquels la protéine codée par *SF3B1* est impliquée, ainsi que les conséquences fonctionnelles des mutations de ce gène.

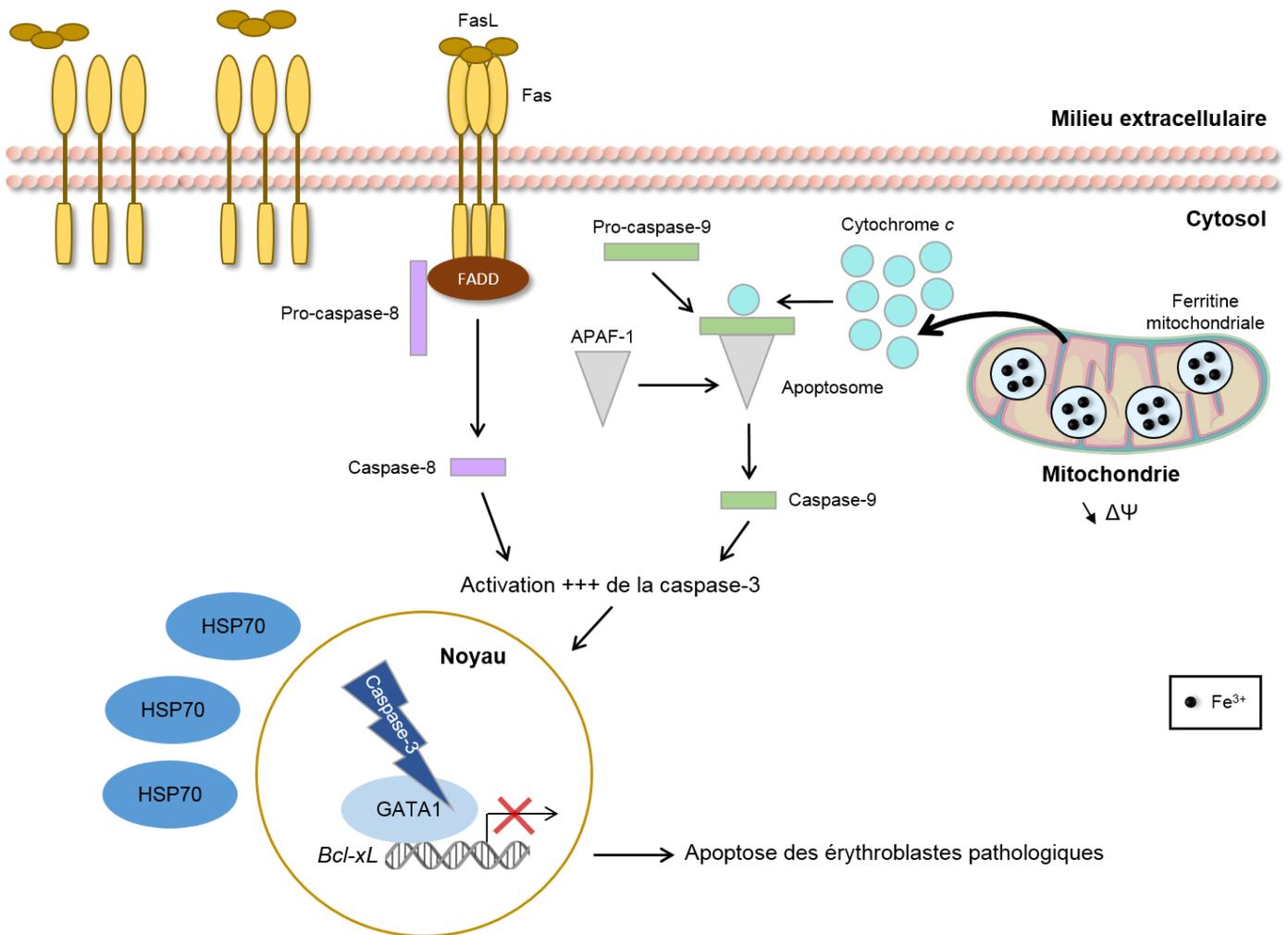


Figure 21 : Mécanismes impliqués dans la dysérythropoïèse terminale des SMD

Différents mécanismes participent à la mise en place de la dysérythropoïèse terminale caractéristique des SMD-SC. L'accumulation de fer – sous forme de ferritine mitochondriale – dans les mitochondries des érythroblastes est associée à une libération massive de cytochrome *c*. La caspase-3 est donc davantage activée dans les érythroblastes pathologiques par rapport aux cellules saines. La surexpression du récepteur Fas et de son ligand FasL à la surface des érythroblastes entraîne elle aussi une sur-activation de la caspase-3. Enfin, chez les patients, la protéine chaperonne HSP70 ne colocalise plus avec GATA1 dans le noyau des érythroblastes. La transcription des gènes anti-apoptotiques, comme *Bcl-xL*, n'est donc plus stimulée. Finalement, les cellules érythroïdes des patients porteurs d'un SMD-SC subissent une apoptose majeure, d'où le développement de l'anémie.

D. Principaux traitements des SMD-SC

Étant donné le faible risque de transformation en LAM des SMD-SC, la prise en charge thérapeutique a pour objectif principal de **corriger l'anémie, conséquence directe de l'érythropoïèse inefficace**.

1. Traitements actuels

Dès lors que l'anémie devient symptomatique – en pratique, un taux d'hémoglobine inférieur à 9-10 g/dL avec une mauvaise tolérance clinique –, des **agents stimulants l'érythropoïèse** (ASE) tels que l'EPO recombinante sont administrés. En cas d'anémie inférieure à 8 g/dL, les **transfusions érythrocytaires** sont mises en place. Il est important de signaler que 50% des patients porteurs d'un SMD sont résistants aux ASE et la prise en charge de leur anémie nécessite donc la transfusion régulière de culots de globules rouges (Park et al., 2008). La plupart des patients atteints d'un SMD ont effectivement des concentrations sériques en EPO plus élevées que la normale avant injection des ASE, ce qui suppose que l'altération de l'érythropoïèse intervient en aval des étapes régulées par l'EPO (Aul et al., 1991). De plus, l'EPO agit principalement sur les précurseurs érythroïdes immatures (ProE et Baso E), mais quasiment pas sur les progéniteurs de type BFU-E ou les précurseurs érythroïdes matures (Poly E et Acido E) qui n'expriment pas EPOR (cf. Chapitre I-I-C-1-b). L'apoptose massive des précurseurs matures ne peut donc être prévenue par l'EPO. Enfin, chez les patients répondeurs, la durée médiane de réponse aux ASE est de deux ans environ, suggérant l'émergence de mécanismes de résistance intrinsèques tels que l'accumulation de mutations ou extrinsèques tels que l'interaction avec un microenvironnement inflammatoire (Fenaux and Ades, 2013). À noter que le lénalidomide et les agents hypométhylants (azacitidine) ont été proposés aux patients porteurs d'un SMD de faible risque en cas de résistance aux ASE (Santini et al., 2016; Thépot et al., 2016). Le lénalidomide active le complexe E3-ubiquitine ligase CRL4-CRBN par liaison à la protéine cereblon (CRBN) (Krönke et al., 2015). Il en résulte la poly-ubiquitinylation puis la dégradation par le protéasome de protéines cibles, dont la caséine kinase 1A codée par le gène *CSNK1A1* situé en 5q32. La dégradation de la protéine codée par l'allèle restant expliquerait la sensibilité des cellules porteuses de la délétion 5q au lénalidomide. Toutefois, le lénalidomide induit également une indépendance transfusionnelle chez 30% à 50% des patients porteurs d'un SMD de bas risque non del(5q) (Toma et al., 2016). Notre équipe a montré, d'une part, une diminution de la charge allélique suggérant le ciblage du clone muté

et, d'autre part, une réponse lymphocytaire T suggérant un mécanisme immunomodulateur (Chesnais et al., 2017; Kerdivel et al., 2018).

Sur le long terme, les transfusions érythrocytaires accentuent la surcharge martiale. Le fer en excès s'accumule dans les cellules parenchymateuses et devient toxique pour de nombreux organes, tels que le foie, le cœur et les glandes endocrines. Des études ont ailleurs que la surcharge systémique en fer est associée à une diminution de la survie et à une progression vers la LAM chez les patients atteints d'un SMD (Jabbour et al., 2008). Pour éviter ces complications, des **traitements chélateurs du fer** sont administrés aux patients régulièrement transfusés. Actuellement, le traitement le plus utilisé est le déférasirox, chélateur du fer par voie orale. Pour obtenir une réduction rapide de la surcharge en fer, la déféroxamine parentérale peut néanmoins être préférée.

2. Traitement en cours d'évaluation clinique

Du fait des résistances à l'EPO et des surcharges en fer secondaires aux transfusions érythrocytaires, l'enjeu majeur de la prise en charge thérapeutique des SMD-SC consiste aujourd'hui à cibler l'érythropoïèse inefficace indépendamment des étapes régulées par l'EPO. Parmi les molécules en cours de développement, l'une d'entre elles fait l'objet d'essais cliniques de phase 3 depuis 2016 : le **luspatercept** (ou ACE-536). Il s'agit d'une protéine recombinante de fusion formée du domaine extracellulaire modifié du récepteur à l'activine de type IIB lié au fragment cristallisable (Fc) de l'immunoglobuline humaine G1. Ainsi, le luspatercept piège les ligands de la superfamille du TGF β , dont le GDF11, le GDF8 et l'activine B (Dussiot et al., 2014; Suragani et al., 2014).

De nombreuses études *in vitro* soulignent un rôle inhibiteur du TGF β sur l'érythropoïèse. En effet, ce dernier, sécrété au sein de la niche hématopoïétique notamment par les cellules stromales mésenchymateuses, induit d'une part une accélération de la conversion des BFU-E en CFU-E et, d'autre part, une accélération de la différenciation érythroïde terminale (Krystal, 1994; Zermati et al., 2000). En outre, le TGF β provoque l'arrêt du cycle cellulaire dans les progéniteurs érythroïdes immatures (Zermati et al., 2000). À cause de cette inhibition majeure de la prolifération cellulaire, le nombre d'érythrocytes produits *in fine* diminue drastiquement. Ces mécanismes d'inhibition de l'érythropoïèse sont particuliers et différents de ceux observés pour d'autres cytokines, telles que l'IFN γ et le TNF α qui ont des effets pro-apoptotiques (cf. Chapitre I-I-E-3). En ce qui concerne plus spécifiquement les

ligands du récepteur à l'activine de type IIB, il a été montré que l'expression de **GDF11** diminue progressivement au cours de l'érythropoïèse, suggérant un **effet inhibiteur** de ce dernier **sur les dernières étapes de la maturation érythrocytaire** (Dussiot et al., 2014; Suragani et al., 2014). GDF11 est un facteur circulant connu pour ses effets « anti-vieillesse » – notamment sur le muscle squelettique – mais son rôle exact dans le tissu érythroïde n'est à ce jour pas défini (Sinha et al., 2014). D'un point de vue mécanistique, la liaison de GDF11 à son récepteur entraîne la phosphorylation des SMAD2/3. SMAD2/3 phosphorylé forme ensuite, avec SMAD4, un dimère transloqué dans le noyau. Ce dimère se fixe sur des régions spécifiques des gènes cibles et régule ainsi leur transcription (Shi and Massagué, 2003).

En 2008, Zhou *et al.* ont démontré, par des études d'immunohistochimie, que la protéine SMAD2 était surexprimée dans les cellules CD34⁺ des patients atteints d'un SMD (Zhou et al., 2008). Après culture *in vitro* de ces cellules CD34⁺, il a aussi été mis en évidence une activation constitutive de SMAD2 dans les précurseurs érythroïdes pathologiques. Enfin, les **concentrations plasmatiques de GDF11**, déterminées par dosage ELISA, semblent **plus élevées** – bien que très variables – **chez les patients myélodysplasiques** par rapport aux témoins sains (Suragani et al., 2014). Cependant, le dosage de GDF11 dans la niche hématopoïétique serait plus informatif que la concentration plasmatique, puisque cette cytokine s'exprime dans de nombreux tissus.

Un essai clinique de phase 2 a été mené en Allemagne sur une cohorte de 51 patients porteurs d'un SMD de bas risque (IPSS-R de très faible à intermédiaire) (Platzbecker et al., 2017). Après trois mois de traitement (0,75-1,75 mg/kg administrés en sous-cutané tous les 21 jours), le luspatercept corrige l'anémie en augmentant les taux d'hémoglobine et/ou en diminuant le nombre de transfusions sanguines chez 63% des patients. 69% des patients atteints d'un SMD-SC présentent une réponse érythroïde au luspatercept, contre seulement 43% des patients atteints d'un SMD non sidéroblastique. De même, 77% des patients porteurs d'une mutation sur le gène *SF3B1* présentent une réponse érythroïde, contre seulement 40% des patients non mutés sur le gène *SF3B1*. Les résultats de l'étude de phase 3 MEDALIST ont été exposés au congrès de l'*American Society of Hematology* en décembre 2018. 58 des 153 (37,9%) patients ayant reçu le luspatercept ont atteint le critère principal d'évaluation – à savoir une absence de besoin transfusionnel pendant au moins 8 semaines – contre 10 des 76 (13,2%) patients ayant reçu le placebo. La médiane de réponse au traitement est de 30,6

semaines. Au total, 90% des patients de l'essai clinique MEDALIST étaient porteurs d'une mutation sur le gène *SF3B1* au moment de l'inclusion (List, 2018).

Le mécanisme d'action du luspatercept n'est pas encore parfaitement élucidé : quelle est l'action exacte du GDF11 sur la différenciation érythrocytaire terminale ? Quel est le rôle du microenvironnement dans la production de GDF11 dans la niche hématopoïétique ? Est-ce que le GDF8 et l'activine B – qui sont normalement attrapés par le luspatercept – ont également un rôle au cours de l'érythropoïèse ? Le luspatercept fixe-t-il d'autres molécules de la superfamille du TGF β ? Pourquoi certains patients ne sont pas répondeurs au luspatercept et, surtout, pourquoi les SMD-SC sont-elles plus sensibles au traitement que les autres sous-types de SMD ?

Chapitre II – La protéine SF3B1

Partie 1

I. La protéine SF3B1 et l'épissage des ARN pré-messagers

Chez l'homme, le gène *SF3B1* (pour *Splicing factor 3B subunit 1*) est situé sur le bras long du chromosome 2 (2q33.1). Il code pour la sous-unité 1 du facteur d'épissage 3B, une protéine de 155 kDa qui participe à la formation du **complexe ribonucléoprotéique nucléaire U2 du spliceosome**. Le gène *SF3B1* s'exprime dans l'ensemble des tissus de l'organisme. En ce qui concerne les érythroblastes humains, l'expression de *SF3B1* au niveau transcriptionnel et protéique est maximale dans les Pro-E et Baso-E. Comme indiqué dans le chapitre I-I-D-3, le pourcentage de transcrits SF3B1 avec rétention d'intron augmente progressivement au cours de l'érythropoïèse terminale. Ces transcrits restent dans le noyau et ne sont donc pas traduits ultérieurement en protéines.

Les approches de séquençage à haut débit ont permis d'identifier des mutations somatiques dans la moelle osseuse des patients porteurs d'un SMD. Il a ainsi été mis en évidence la présence récurrente de mutations touchant des gènes impliqués dans les processus de maturation des ARNm, dont notamment les gènes d'épissage *SF3B1*, *SRSF2*, *U2AF1* et *ZRSR2* (Papaemmanuil et al., 2011; Yoshida et al., 2011). De manière inattendue, **les mutations du gène *SF3B1* sont présentes chez 28% des patients atteints d'un SMD et, plus particulièrement, chez plus de 80% des patients atteints d'un SMD avec SC** (Malcovati et al., 2011, 2015). *SF3B1* est également muté dans 15% des leucémies lymphoïdes chroniques (LLC) (Quesada et al., 2012), dans les mélanomes de l'uvée (MU) (Martin et al., 2013), et plus rarement dans les cancers du sein (Stephens et al., 2012) et du pancréas (Biankin et al., 2012).

Cette première partie du chapitre II détaillera le rôle de la protéine SF3B1 dans l'épissage des ARN pré-messagers, ainsi que sa structure et ses diverses localisations intranucléaires. Je reviendrai également sur les études *in vivo* et *in vitro* d'inactivation du gène *SF3B1*.

A. Épissage des ARN pré-messagers

Chez les eucaryotes, l'épissage est un processus biologique fondamental par lequel les ARN pré-messagers (pré-ARNm) – transcrits à partir de l'ADN génomique – subissent des étapes de coupure et ligature qui conduisent à l'élimination des séquences non codantes (les

introns) de l'ARN final. Les séquences codantes qui se retrouvent dans l'ARNm final sont les exons. L'épissage est l'une des trois étapes nécessaires à la maturation du pré-ARNm, les deux autres étant l'ajout de la coiffe à l'extrémité 5' et la polyadénylation à l'extrémité 3'. Il est réalisé dans le noyau sous l'action du spliceosome (ou machinerie d'épissage), un complexe dynamique constitué de particules nucléaires ribonucléoprotéiques (snRNP) – elles-mêmes formées d'un petit ARN nucléaire (snARN) et de protéines – et de facteurs protéiques additionnels (Wahl et al., 2009).

Le spliceosome majeur assure l'épissage de la quasi-totalité des pré-ARNm humains. Il se compose des snRNP U1, U2, U4/U6 et U5 s'assemblant de manière séquentielle au niveau des introns. Ces derniers contiennent en effet des séquences consensus définissant le site donneur d'épissage en 5' (5'ss), le site de branchement (BS), le *polypyrimidine tract* (PPT) et le site accepteur d'épissage en 3' (3'ss), qui sont respectivement et dans un premier temps reconnus par la snRNP U1, le facteur d'épissage SF1, et les facteurs auxiliaires de la snRNP U2 U2AF65 (également nommé U2AF2) et U2AF35 (également nommé U2AF1) (Figure 22a). Le complexe ainsi généré est le **complexe E** d'initiation de l'épissage (Figure 22b). Il joue un rôle crucial dans la définition des 5'ss et 3'ss aux extrémités de l'intron.

Dans un second temps, le snARN U2 reconnaît le BS de l'intron par appariement de bases (Wahl et al., 2009). La snRNP U2 déplace alors SF1, ce qui entraîne la formation du **complexe A** ou « pré-spliceosome » (Figure 22b). La séquence nucléotidique du BS étant courte et dégénérée, l'interaction seule entre le snARN U2 et le BS ne suffit pas pour assurer la liaison spécifique de la snRNP U2. Cette interaction est donc stabilisée par des protéines additionnelles, comme les complexes hétéromériques SF3a et SF3b, les protéines riches en domaines sérine-arginine (SR) SRSF1 et SRSF2, mais aussi par le domaine SR d'U2AF65. SF3b comprend plusieurs sous-unités, dont le facteur d'épissage p14/SF3b14a qui s'associe avec l'adénosine du BS, et le facteur d'épissage SF3B1 (Zhang et al., 2018). **SF3B1 interagit avec l'ARN intronique de part et d'autre du BS, ainsi qu'avec le domaine C-terminal d'U2AF65** (Figure 22b). En 1998, il a été démontré que **SF3B1 permet le recrutement de l'ensemble de la snRNP U2 au niveau du BS** (Gozani et al., 1998).

Les snRNPs U4/U6 préassemblées et la snRNP U5 sont ensuite recrutées au niveau de l'intron, de façon à former le **complexe B** (Figure 22b). Bien que les cinq snRNPs soient présentes dans le complexe B, celui-ci reste inactif : il s'agit du complexe pré-catalytique. L'activation du spliceosome nécessite divers réarrangements conformationnels et

compositionnels, tels que la libération des snRNPs U1 et U4 qui donne ainsi lieu au **complexe B*** (Wahl et al., 2009) (Figure 22b). À noter que les réarrangements entre les différentes protéines du spliceosome sont mieux caractérisés pour les premières étapes de l'épissage (complexes E et A), par rapport aux dernières étapes.

L'épissage du pré-ARNm à proprement parler se fait grâce à deux réactions successives de transestérification. Tout d'abord, le 5'ss est attaqué par le groupement hydroxyle de l'adénosine du BS : l'extrémité 5' de l'intron est coupée et ce dernier forme un lasso. Le **complexe C** du spliceosome subit à nouveau des réarrangements. Puis, le 3'ss est attaqué par le groupement hydroxyle du 5'ss cette fois-ci : le lasso est excisé et les deux exons adjacents ligaturés (Wahl et al., 2009). Enfin, les protéines du spliceosome sont dissociées et recyclées, libérant ainsi l'ARNm (Figure 22b).

L'épissage peut être co-transcriptionnel, c'est-à-dire que les introns sont excisés au cours de la synthèse du pré-ARNm. Cette notion a été introduite dès la fin des années 1980, suite aux travaux de microscopie électronique de Beyer *et al.* sur des embryons de *Drosophile* (Beyer and Osheim, 1988). Les ARN naissants – encore attachés à la chromatine décondensée (euchromatine) – sont effectivement recouverts de snRNP. Plus récemment, il a été montré que plus de 80% des pré-ARNm humains sont épissés de façon co-transcriptionnelle au niveau des fibrilles de périchromatine, des régions riches en ARN situées en bordure de la chromatine (Fakan, 1994; Girard et al., 2012). Il est important de signaler que certains facteurs d'épissage, comme par exemple SRSF2 (également nommé SC35), sont directement impliqués dans la régulation de la transcription de l'ADN, et ce indépendamment de leur fonction dans l'épissage. En effet, lorsque la protéine SRSF2 est recrutée au niveau des transcrits naissants associés aux séquences promotrices des gènes, la kinase cycline-dépendante P-TEFb – séquestrée au sein du complexe 7SK – est libérée. P-TEFb agit alors comme un facteur d'élongation de la transcription : il phosphoryle l'ARN polymérase II dans son domaine C-terminal, ce qui active la progression de cette dernière sur l'ADN (Ji et al., 2013). En conclusion, la transcription de l'ADN et l'épissage des pré-ARNm sont deux processus distincts mais simultanés, et leur régulation peut faire intervenir les mêmes protéines.

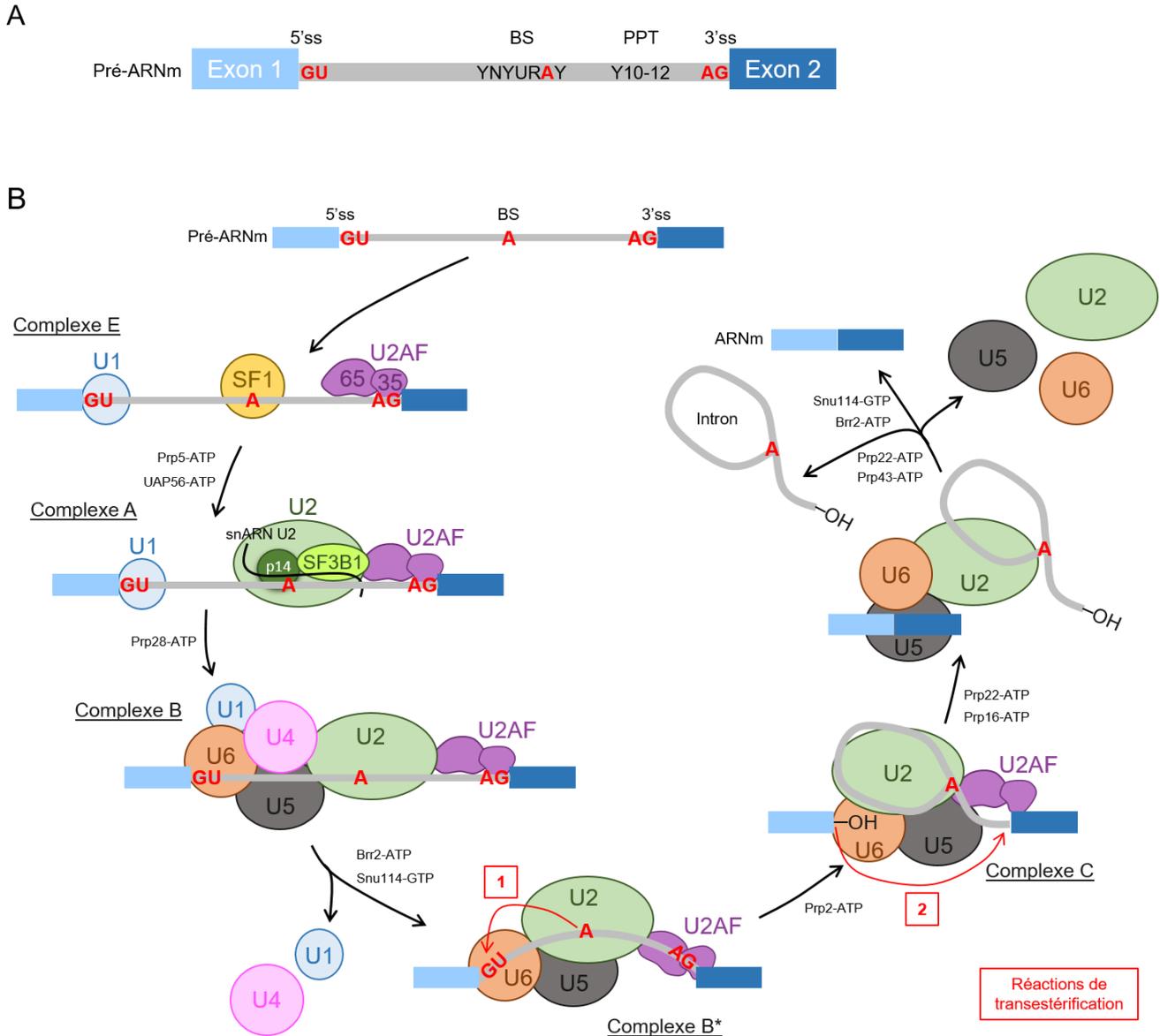


Figure 22 : Épissage des pré-ARNm par le spliceosome majeur

(A) Dans les introns des ARN pré-messagers (**pré-ARNm**), des séquences consensus permettent de définir le site donneur d'épissage (**5'ss**), le site de branchement (**BS**), le *polypyrimidine tract* (**PPT**) et le site accepteur d'épissage (**3'ss**). Le PPT est une région intronique riche en pyrimidines reconnue par la sous-unité U2AF65 d'U2AF. Gris = intron, bleu clair = exon 1, bleu foncé = exon 2, **G** = guanine, **U** = uridine, **A** = adénosine, **Y** = pyrimidine, **R** = purine. (B) Les snRNPs U1, U2, U4/U6 et U5 s'assemblent au niveau de l'intron du pré-ARNm de manière séquentielle. La liaison de la snRNP U2 sur le BS génère le complexe A du spliceosome. La protéine SF3B1 intervient dans la stabilisation de l'interaction entre U2 et le BS. Son extrémité N-terminale s'associe avec U2AF65. Huit hélicases ATP-dépendantes (pour adénosine triphosphate) aident au remodelage de l'ARN au cours de l'épissage : Prp5, UAP56, Prp28, Brr2, Prp2, Prp22 et Prp16 et Prp43. La GTPase Snu114 catalyse également les divers réarrangements. Suite aux deux réactions de transestérification, l'intron est excisé puis dégradé par des nucléases. L'ARNm composé d'exons est exporté dans le cytoplasme où il sera traduit en protéines par les ribosomes.

B. Structure de la protéine SF3B1

1. Domaines protéiques de SF3B1

Chez l'homme, le facteur d'épissage SF3B1 (également nommé SF3b155 ou SAP155) représente la plus grande sous-unité du complexe multi-protéique **SF3b**. En plus de SF3B1, six autres protéines entrent dans la formation de SF3b : SF3b130, SF3b145, SF3b49, SF3b14b, p14/SF3b14a et SF3b10.

D'après les études de cristallographie, le domaine N-terminal (NTD) hydrophile de SF3B1 est non structuré (Cretu et al., 2016). Il joue un rôle important lors du processus d'épissage des pré-ARNm, puisqu'il est directement impliqué dans l'interaction de SF3B1 avec U2AF65. La région adjacente au NTD s'associe quant à elle avec la sous-unité p14 du complexe SF3b (Figure 23a).

À l'opposé du NTD, le **domaine C-terminal** (CTD) de SF3B1 s'organise en **20 répétitions HEAT** (H1 à H20), chaque répétition étant formée d'une paire de deux hélices α anti-parallèles reliées par une petite boucle (Figure 23b). Le nom « HEAT » fait référence à quatre protéines connues pour posséder ce type de motifs : l'huntingtine (*huntingtin*), le facteur d'élongation 3 (*elongation factor 3*), la protéine phosphatase 2A (*protein phosphatase 2A*), et la kinase TOR1 (*target of rapamycin 1*). Les séquences des répétitions HEAT varient en longueur entre 37 et 47 acides aminés (Kennedy et al., 2009). Elles varient également en composition. Toutefois, certains résidus hydrophobes sont hautement conservés d'une séquence à l'autre, tels que la proline en position 11. De même, les résidus aspartate et arginine – aux positions 19 et 25 respectivement – sont retrouvés dans l'ensemble des répétitions HEAT, ce qui suggère une base structurale commune (Groves et al., 1999). SF3B1 interagit avec les sous-unités SF3b130, SF3b10 et SF3b14b de SF3b par l'intermédiaire de son domaine C-terminal (Cretu et al., 2016) (Figure 23a). Le CTD est également impliqué dans l'interaction de SF3B1 avec les introns du pré-ARNm, comme le soulignent les travaux de cryo-microscopie électronique chez la Levure (*Saccharomyces cerevisiae*) (Yan et al., 2016). **Les répétitions HEAT chargées positivement s'assemblent aux séquences introniques chargées négativement des deux côtés du BS.** Les premières études de cristallographie du spliceosome indiquent que les motifs hélicoïdaux dans le CTD de SF3B1 obstruent le motif de reconnaissance de l'intron situé sur p14 (Schellenberg et al., 2006). En

fait, le repliement tridimensionnel du complexe p14-SF3B1 permettrait la formation d'une interface de reconnaissance du BS, interface située en plein cœur du spliceosome.

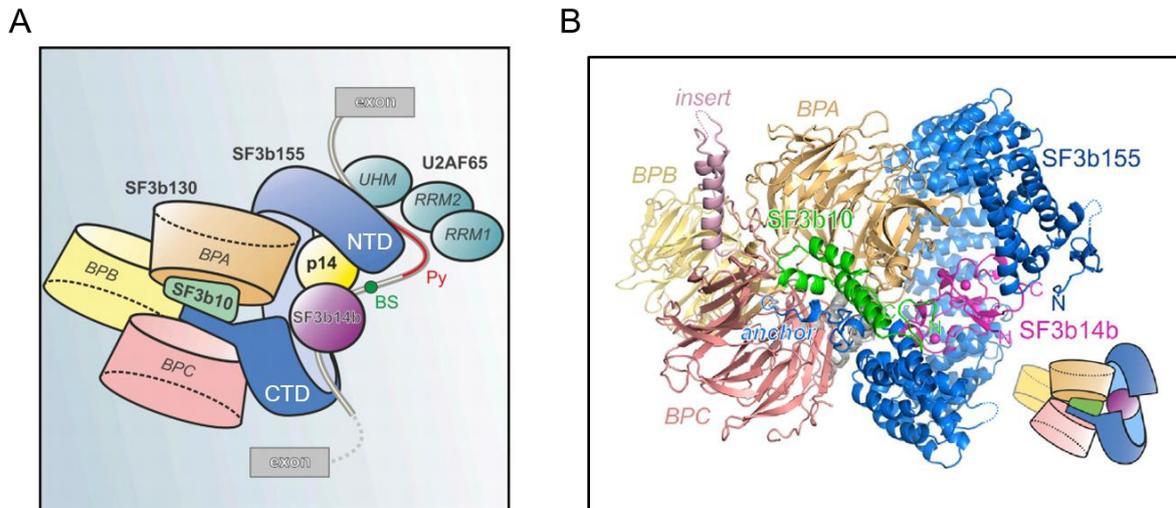


Figure 23 : Structure du complexe humain SF3b (d'après Cretu et al., 2016)

(A) Le complexe SF3b se compose de sept sous-unités : SF3b130, SF3b145, SF3b49, SF3b14b, p14, SF3b10 et SF3b155. La liaison de SF3b avec le site de branchement (BS) de l'intron se fait par l'intermédiaire de la sous-unité p14. Le domaine N-terminal (NTD) de SF3b155 (ou SF3B1) interagit avec U2AF65, notamment au niveau des régions UHM (pour *U2AF-homology motif*), RPM2 et RPM1 de ce dernier. Pour rappels, U2AF65 se lie à une séquence intronique riche en pyrimidines (Py). Le domaine C-terminal (CTD) de SF3b155 s'associe quant à lui avec SF3b10, SF3b14b et les régions BPA, BPB et BPC de SF3b130. (B) D'un point de vue structural, vingt répétitions HEAT forment le CTD de SF3b155. Insert = élément en épingle à cheveux qui stabilise l'interaction entre SF3b10 et la région BPA de SF3b130, *anchor* (pour ancre) = région terminale du CTD de SF3b155 qui interagit avec SF3b10 et la région BPC de SF3b130. Contrairement au reste du CTD de SF3b155, l'ancre n'a pas une structure en hélice α .

2. Phosphorylation de SF3B1

La protéine SF3B1 est phosphorylée lors de l'activation du spliceosome (complexe B*), et reste hyperphosphorylée jusqu'à la fin de la deuxième réaction de transestérification (Bessonov et al., 2010; Wang et al., 1998a). Les sites de phosphorylation sont situés dans le NTD de SF3B1, une région notamment composée d'un grand nombre de dipeptides thréonine-proline (Boudrez et al., 2002). Les thréonines T211, T235, T313 et T328 ainsi que les sérines 129 et 322 figurent parmi les nombreux résidus phosphorylables actuellement identifiés (Girard et al., 2012). Diverses enzymes phosphorylent SF3B1 *in vitro*, comme les kinases cycline-dépendantes **CDK1** et **CDK2** (Boudrez et al., 2002). Pour rappel, le cycle cellulaire se déroule en quatre phases : la phase G1 de préparation à la réplication de l'ADN,

la phase S de synthèse de l'ADN, la phase G2 de préparation à la division cellulaire et, enfin, la phase M de mitose. Le couple cycline B/CDK1 assure l'entrée en phase M, tandis que le couple cycline E/CDK2 est responsable de la transition entre les phases G1 et S. Des travaux d'immunoprécipitation ont mis en évidence une interaction entre SF3B1 et la cycline E (Seghezzi et al., 1998). Ainsi, l'épissage des pré-ARNm dépend de la machinerie du cycle cellulaire. SF3B1 est également phosphorylé *in vitro* et *in vivo* par la kinase **DYRK1A** (de Graaf et al., 2006). En plus d'être indispensable à l'activation du spliceosome, la phosphorylation de l'extrémité N-terminale de SF3B1 garantit l'association de ce dernier avec d'autres protéines, telles que l'inhibiteur nucléaire de la protéine phosphatase 1 (NIPP1) (Boudrez et al., 2002).

C. Localisations intranucléaires de SF3B1

1. *Nuclear speckles*

Dans les cellules des mammifères, les éléments de la machinerie d'épissage – dont les snRNP et protéine additionnelles telles que SF3B1 et SRSF2 – sont stockés et assemblés dans des structures particulières nommées « *nuclear speckles* » (ou domaines SC35/SRSF2). Les *nuclear speckles* sont des compartiments intranucléaires localisés dans les espaces interchromatiniens du nucléoplasme (Lamond and Spector, 2003). Dépourvus de membrane, la taille et la morphologie des *nuclear speckles* résultent de l'échange dynamique des facteurs d'épissage entre ces derniers et le nucléoplasme. Leur formation évolue au cours du cycle cellulaire. En effet, le contenu des *nuclear speckles* se répartit de manière diffuse dans le cytoplasme lorsque les cellules entrent en mitose (prophase). Durant la métaphase, l'anaphase et le début de la télophase, des granules interchromatiniens apparaissent petit à petit dans le cytoplasme. À la fin de la télophase, ces granules migrent vers le noyau des futures cellules filles. Les *nuclear speckles* sont à nouveau assemblés en phase G1 du cycle cellulaire et le resteront tout au long de l'interphase (Lamond and Spector, 2003).

Certains introns – notamment ceux localisés à l'extrémité 3' des pré-ARNm – sont plutôt excisés après la transcription (Wetterberg et al., 1996). Les transcrits naissants sont alors libérés de leur matrice d'ADN par clivage de l'extrémité 3', puis épissés. Contrairement à l'épissage co-transcriptionnel qui a lieu – comme dit précédemment – au niveau des fibrilles de périchromatine en périphérie des *nuclear speckles*, l'épissage post-transcriptionnel se déroule au sein même des *nuclear speckles* (Girard et al., 2012).

2. Nucléosomes

SF3B1 peut s'associer à la chromatine, en particulier avec les **mononucléosomes positionnés sur les régions exoniques** (Kfir et al., 2015). Pour rappels, chez les eucaryotes, les nucléosomes représentent l'unité de base de la chromatine et le premier niveau de compaction de l'ADN. Ils se composent d'un octamère d'histones autour duquel s'enroule un fragment d'ADN (Figure 24). D'après les travaux de Kfir *et al.*, SF3B1 serait pré-chargé sur les nucléosomes exoniques avant l'initiation de la transcription de l'ADN, ce qui assurerait par la suite la reconnaissance de ces exons par la machinerie d'épissage et leur inclusion dans l'ARNm (Figure 24). Ainsi, le couplage de SF3B1 avec les nucléosomes exoniques est important pour la régulation de l'épissage de ces exons. Puis, lorsque le gène est transcrit, SF3B1 se détache du nucléosome et est recruté sur les ARN naissants. Pour confirmer cette hypothèse, Kfir *et al.* ont traité la lignée cellulaire HeLa avec des composés chimiques séparant SF3B1 des nucléosomes exoniques. Les exons normalement associés avec SF3B1 au sein des nucléosomes sont dans ce cas exclus de l'ARNm lors de l'épissage (Figure 24). **SF3B1 véhicule finalement les informations contenues dans la structure de la chromatine – notamment les sites d'épissage – à la machinerie d'épissage.** L'interaction de SF3B1 avec les nucléosomes est maximale pendant les phases G1 et S du cycle cellulaire (Murthy et al., 2018). Elle est alors sous le contrôle de la kinase CDK2 (cf. Chapitre II-Partie 1-I-B-2). Au contraire, pendant les phases G2 et M du cycle cellulaire, SF3B1 est phosphorylé par la kinase CDK1 et se dissocie des nucléosomes.

En conclusion, SF3B1 est stocké dans les *nuclear speckles*. Lorsque l'ADN est transcrit, il est mobilisé au niveau des ARN naissants puis phosphorylé au sein du complexe B* du spliceosome. Le spliceosome activé épisse les pré-ARNm nouvellement synthétisés. SF3B1 peut aussi être recruté directement au niveau de la chromatine, et ce indépendamment de l'activation de la transcription. L'association de SF3B1 phosphorylé aux nucléosomes positionnés sur les séquences exoniques de l'ADN est garante du bon épissage de ces exons. Seule une petite portion (15 à 20%) de phospho-SF3B1 est retrouvée dans les *nuclear speckles*. Elle assure l'épissage post-transcriptionnel. Les fonctions biologiques du facteur SF3B1 sont donc multiples et complexes puisqu'il se situe à la frontière entre l'organisation de la chromatine, la transcription de l'ADN et l'épissage des pré-ARNm. Son activité est modulée tout au long du cycle cellulaire.

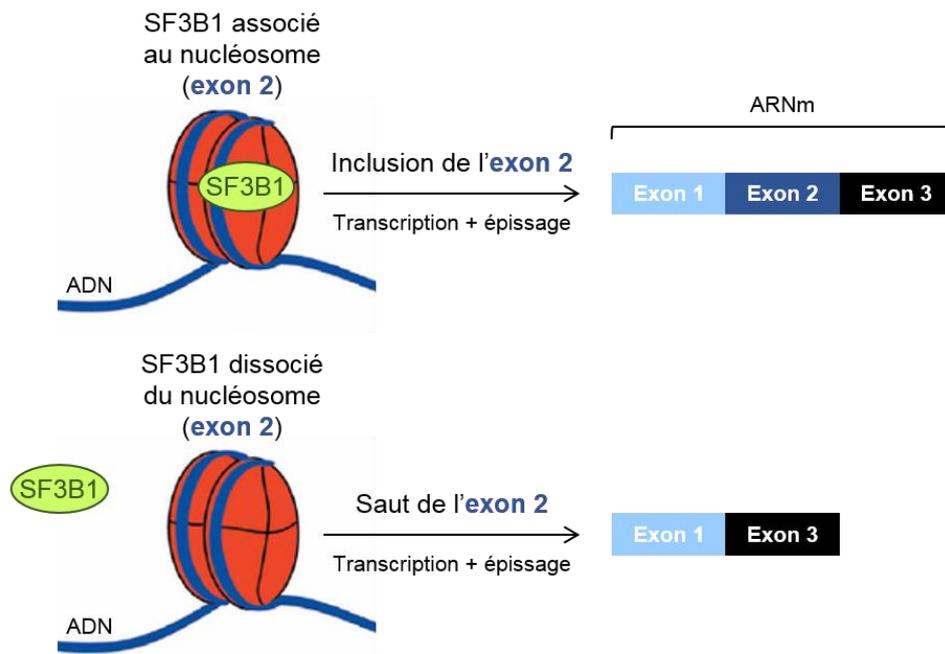


Figure 24 : L'association de SF3B1 aux nucléosomes détermine l'épissage des pré-ARNm

SF3B1 s'associe aux nucléosomes (orange) localisés sur les séquences exoniques, ce qui permet la reconnaissance de ces exons par la machinerie d'épissage et leur inclusion dans l'ARNm. En cas de rupture de l'association de SF3B1 aux nucléosomes, les exons normalement liés à SF3B1 ne sont plus reconnus par la machinerie d'épissage et sont donc exclus de l'ARNm.

D. Études d'invalidation du gène *Sf3b1*

D'après les études *in vivo*, **l'invalidation totale (KO) du gène *Sf3b1* est létale au stade embryonnaire** (Isono, 2005). Les souris hétérozygotes *Sf3b1*^{+/-} naissent quant à elles avec de légères modifications squelettiques dues à la dérégulation des gènes *Hox*. En effet, Sf3b1 interagit avec les protéines PcG (pour *Polycomb Group*), comme Zfp144 et Rnf2. Ces protéines s'associent à la chromatine et répriment la transcription des gènes *Hox* lors du développement embryonnaire. L'haplo-insuffisance de *Sf3b1* a donc les mêmes conséquences que les mutations de *PcG* au niveau du squelette (Isono, 2005). En ce qui concerne le tissu hématopoïétique plus spécifiquement, l'invalidation partielle de *Sf3b1* chez les souris entraîne une réduction du nombre de CSH ainsi qu'une altération de leur capacité à reconstituer l'hématopoïèse suite à une transplantation compétitive (*Sf3b1*^{+/-} vs *Sf3b1*^{+/+}) (Wang et al., 2014). Les paramètres hématologiques – tels que le taux d'hémoglobine – ne sont quant à eux pas modifiés et le nombre de SC médullaires est identique entre les souris *Sf3b1*^{+/-} et les souris WT. En conclusion, bien que le niveau d'expression de *Sf3b1* semble critique pour la capacité de prolifération des CSH, **l'haplo-insuffisance de Sf3b1 ne permet pas de reproduire le phénotype des anémies sidéroblastiques**. Enfin, en 2018, Huang *et al.* ont exploré avec un shSF3B1 les effets de la diminution de l'expression de *SF3B1* dans les progéniteurs hématopoïétiques humains CD34⁺ cultivés *in vitro* en conditions de différenciation érythroïde (Huang et al., 2018). Ils observent d'une part une diminution drastique de la prolifération cellulaire due à une augmentation de l'apoptose et un arrêt du cycle cellulaire en phase G0/G1 dans les BFU-E et CFU-E. Les expériences d'immuno-empreinte montrent une induction de la voie p53. Les lignées cellulaires myéloïdes TF1, K562, HEL et SKM1 transduites avec un siSF3B1 sont elles aussi stoppées dans le cycle cellulaire en phase G0/G1 (Dolatshad et al., 2015). D'autre part, d'après Huang *et al.*, la diminution de l'expression de *SF3B1* retarde la différenciation érythroïde terminale et change la morphologie des noyaux des Poly-E et Acido-E (gros noyaux polylobés). Les fonctions assurées par la protéine SF3B1 sont donc nécessaires au bon déroulement de l'érythropoïèse. En revanche, le pourcentage de SC dans les cellules traitées avec un shSF3B1 est identique à celui des cellules contrôles.

II. Mutations de *SF3B1* et dérégulation de l'épissage

Le gène *SF3B1* est muté dans la quasi-totalité des SMD-SC. L'association d'une mutation génétique spécifique avec un caractère morphologique distinct – les SC – suggère que la protéine SF3B1 mutée pourrait avoir un rôle unique dans la pathogénèse des SMD.

Lors de l'analyse de l'hémogramme au diagnostic d'un SMD, le taux d'hémoglobine est généralement plus bas si les cellules hématopoïétiques du patient sont mutées sur le gène *SF3B1* (8,4 g/dL *versus* 9,8 g/dL chez les patients *SF3B1*^{WT}) (Damm et al., 2012). L'analyse morphologique du frottis médullaire révèle quant à elle une dysplasie spécifique de la lignée érythroïde chez 79% des patients porteurs d'un SMD *SF3B1*^{MUT} contre 46% des patients *SF3B1*^{WT}. Les mutations de *SF3B1* sont donc très souvent associées à une **dysérythropoïèse** et, par conséquent, au développement d'une **anémie profonde**. Enfin, les patients SMD *SF3B1*^{MUT} ont un **meilleur pronostic** que les patients SMD *SF3B1*^{WT}, puisque la survie globale et la survie sans transformation en LAM sont allongées (médiane de survie après diagnostic de 90 mois dans le groupe *SF3B1*^{MUT} *versus* 50 mois dans le groupe *SF3B1*^{WT}) (Malcovati et al., 2011, 2015; Papaemmanuil et al., 2011).

Dans cette partie, je décrirai dans un premier temps les mutations du gène *SF3B1* dans les SMD. Puis, dans un second temps, les caractéristiques clinico-biologiques des souris dans lesquelles la mutation du gène *Sf3b1* a été introduite par mutagenèse dirigée seront expliquées. Enfin, les conséquences fonctionnelles des mutations de *SF3B1* au niveau de l'épissage des pré-ARNm seront détaillées.

A. Description des mutations de *SF3B1* dans les SMD

Chez les patients porteurs d'un SMD, les mutations de *SF3B1* sont :

- **Hétérozygotes** : l'expression de l'allèle sauvage en regard de l'allèle mutant est indispensable à la survie de la cellule *in vitro* (Papaemmanuil et al., 2011; Yoshida et al., 2011; Zhou et al., 2015),
- De type **faux-sens** : le nucléotide d'un codon est modifié, ce qui entraîne le changement de l'acide aminé correspondant,

- **Somatiques** : elles ne sont pas présentes initialement dans le génome mais acquises au cours de la vie du patient dans une cellule hématopoïétique (Papaemmanuil et al., 2011),
- La plupart du temps **initiatrices** et **drivers** de l'oncogenèse : elles apparaissent très précocement dans le développement d'un SMD et confèrent un avantage de prolifération à la cellule mutée, d'où l'apparition d'une population clonale de cellules transformées (Chesnais et al., 2017).

D'après les travaux de thèse d'une ancienne doctorante de mon équipe, les mutations de *SF3B1* se produisent dans le compartiment CD34⁺ CD38⁻, c'est-à-dire dans les **CSH et/ou les progéniteurs multipotents** (MPP) (Chesnais et al., 2017). Elles sont ensuite détectées dans les lignées myéloïdes et, parfois, dans les cellules CD19⁺ représentant la population des lymphocytes B. Les mutations présentes uniquement dans les cellules myéloïdes sont probablement à l'origine d'un **biais de différenciation vers la lignée myéloïde** de la cellule atteinte. Les mutations de *SF3B1* sont souvent associées à des mutations touchant des gènes de l'épigénétique comme *TET2*, *DNMT3A* et *ASXL1*. Par contre, elles sont exclusives des mutations touchant les autres gènes d'épissage tels que *SRSF2*, *U2AF1* et *ZRSR2* (Chesnais et al., 2016).

Dans les cellules hématopoïétiques, l'acquisition d'une mutation somatique sur le gène *SF3B1* n'est pas forcément corrélée à un néoplasme myéloïde. En effet, des analyses de séquençage à haut débit montrent, qu'à partir de 65 ans environ, la fréquence de mutation de certains gènes – dont *TET2*, *DNMT3A*, *ASXL1* et *SF3B1* – augmente (Genovese et al., 2014; Jaiswal et al., 2014; Steensma et al., 2015). Ces mutations confèrent un avantage de prolifération aux cellules et entraînent donc le développement d'une **hématopoïèse clonale**, qui, dans de très rares cas, peut évoluer en SMD, voire se transformer secondairement en LAM. Par exemple, d'après une étude portant sur plus de 12 000 échantillons d'ADN de sang périphérique, l'hématopoïèse clonale associée à ce type de mutation est retrouvée chez 10% des individus de plus de 65 ans (contre seulement 1% chez les moins de 50 ans) (Genovese et al., 2014). L'hématopoïèse clonale représente ici un facteur de risque pour le développement d'un cancer hématologique, puisque le *hazard ratio* est de l'ordre de 12,9% (soit 12,9 fois plus de risque de développer un cancer hématologique chez un sujet avec une hématopoïèse clonale dans les 2 à 7 années suivantes). De plus, 42% des patients avec un cancer hématologique présentaient une hématopoïèse clonale plus de 6 mois avant le diagnostic de

leur cancer. Toutefois, en terme de risque absolu, seulement 4% des individus porteurs d'une mutation conférant une hématoïèse clonale développent un cancer hématologique dans les 150 mois suivants (Jaiswal et al., 2014).

Quelle que soit la pathologie (SMD, LLC, MU...), les mutations de *SF3B1* sont généralement localisées dans la région C-terminale de la protéine, et plus particulièrement dans les quatre **répétitions HEAT** H4 à H7 (Cretu et al., 2016) (Figure 25). Curieusement, la substitution **K700E** – qui correspond au remplacement d'une lysine par un acide glutamique au sein de l'exon 15 du gène *SF3B1* – est de loin la plus fréquente dans les hémopathies malignes (Papaemmanuil et al., 2011), tandis que la majorité des patients atteints d'un MU présente une altération de l'arginine 625 (R625) (Martin et al., 2013). Aucune mutation ne modifie la structure de la protéine de façon drastique, de même qu'aucune mutation n'est située aux interfaces entre SF3B1 et les autres composants du complexe SF3b. Par conséquent, les mutations de *SF3B1* n'induisent pas une déstabilisation complète du spliceosome, déstabilisation qui conduirait alors à un arrêt partiel de l'épissage. La plupart des mutations affectent en fait des résidus impliqués dans la structure tertiaire de la protéine (ponts hydrogènes...) ou dans l'interaction de la protéine SF3B1 avec l'intron du pré-ARNm et/ou d'autres composants du spliceosome (Cretu et al., 2016). Par exemple, la lysine K700 se trouve dans une région de SF3B1 chargée positivement, l'intron du pré-ARNm étant quant à lui chargé négativement. De plus, l'analyse du repliement de la protéine SF3B1 par cristallographie indique que cette région se situe à proximité de la sous-unité U2AF65 de l'hétérodimère U2AF. La mutation K700E change la charge du résidu, il est donc probable qu'elle affecte – par la même occasion – l'interaction de SF3B1 avec le pré-ARNm et/ou U2AF65. Toutefois, d'après les expériences de Cretu *et al.*, l'affinité des protéines SF3B1 sauvage et mutée K700E envers le pré-ARNm et U2AF65 semblent identiques. La dynamique d'assemblage et de désassemblage du spliceosome n'étant pas encore parfaitement connue, il n'est pas exclu que la protéine SF3B1 mutée présente une moindre affinité envers d'autres facteurs d'épissage.

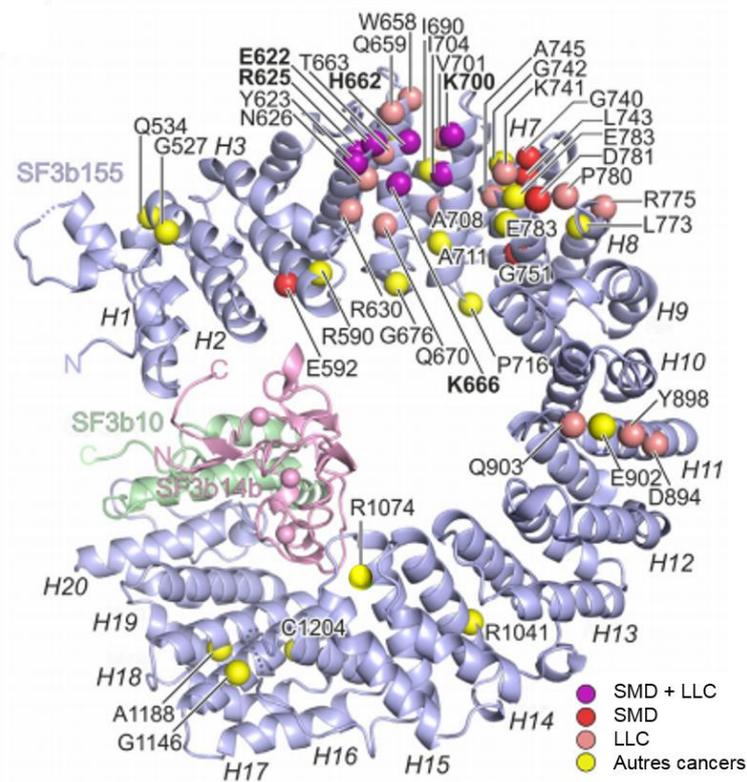


Figure 25 : Distribution des résidus mutés de la protéine SF3B1 dans les cancers (d'après Cretu et al., 2016)

De nombreux résidus sont susceptibles d'être mutés dans la région C-terminale de la protéine SF3B1 qui se compose de 20 répétitions HEAT (H1 à H20). Les mutations fréquemment associées à un SMD sont indiquées en rouge, les mutations fréquemment associées à une LLC sont indiquées en rose, les mutations autant associées à un SMD qu'à une LLC sont indiquées en violet et les mutations associées aux autres cancers sont indiquées en jaune. Les *hotspots* sont indiqués en gras.

En conclusion, l'absence de mutation non-sens, d'insertion ou de délétion sur le gène *SF3B1* ainsi que l'absence de mutation faux-sens au niveau d'un nucléotide associé à un acide aminé fondamental pour la structure de la protéine laissent penser que la protéine SF3B1 mutée conserve son intégrité structurale et, donc, une certaine fonctionnalité dans l'épissage des pré-ARNm.

B. Caractéristiques clinico-biologiques des souris *knock-in Sf3b1^{MUT}*

Des souris transgéniques exprimant la mutation hétérozygote *Sf3b1^{K700E}* ont été obtenues dans un laboratoire américain grâce à un *knock-in* (KI) conditionnel (Obeng et al., 2016). Durant la période d'observation (64 semaines), les souris *Sf3b1^{K700E}* développent progressivement une **anémie macrocytaire**, c'est-à-dire que leur taux d'hémoglobine diminue alors que leur VGM augmente. Aucune anomalie plaquettaire et granuleuse ainsi qu'aucune transformation leucémique n'ont été relevées. L'analyse des marqueurs membranaires CD71 et Ter119 (équivalent de la GPA chez la souris) par cytométrie en flux a permis de quantifier les différents précurseurs érythroïdes dans la moelle osseuse et la rate. De façon intéressante, le nombre de précurseurs érythroïdes immatures (ProE et Baso-E) est augmenté chez les souris *Sf3b1^{K700E}*. En revanche, le nombre de précurseurs érythroïdes matures (Poly-E) est diminué chez les souris *Sf3b1^{K700E}*; d'où finalement un **blocage de la maturation érythroïde terminale**, ce qui – je le rappelle – est l'une des principales caractéristiques des patients atteints d'une anémie sidérolastique. Cette érythropoïèse inefficace s'accompagne d'une **dysplasie** touchant là encore spécifiquement la **lignée érythroïde**. Toutefois, la coloration de Perls ne montre pas d'augmentation significative des SC dans la moelle osseuse des souris *Sf3b1^{K700E}*. Enfin, il est important de noter que le blocage de la différenciation érythroïde terminale n'a pas été mis en évidence chez les jeunes souris (12 semaines), aussi bien dans la rate que dans la moelle osseuse. Ainsi, l'âge semble jouer un rôle prépondérant dans le développement des troubles hématologiques, ce qui est là encore plutôt cohérent avec la pathologie humaine puisque la moyenne d'âge au diagnostic d'un SMD est de l'ordre de 70 ans.

C. Épissage alternatif en aval des mutations de *SF3B1*

1. Définition de l'épissage alternatif

Comme dit précédemment, l'épissage est un mécanisme de maturation de l'ARN qui permet à un ARN transcrit à partir d'un gène (pré-ARNm), de se débarrasser des introns, pour former un ARNm final – composé d'exons –, qui sera ensuite traduit en protéine dans le cytoplasme de la cellule. L'épissage qui consiste à garder tous les exons et à exciser tous les introns est appelé épissage constitutif. En réalité il est possible, à partir d'**un même pré-ARNm**, d'obtenir **différents ARNm** – et donc **différentes protéines** –, selon les séquences conservées ou supprimées dans la version finale : c'est le principe de l'épissage alternatif. Ainsi, à partir d'un nombre limité de gènes (environ 22 000 chez l'homme), une grande diversité d'isoformes protéiques sont potentiellement générées.

À sa découverte, l'épissage alternatif était considéré comme un processus marginal de régulation post-transcriptionnelle de l'expression des gènes (Berget et al., 1977; Chow et al., 1977). Grâce aux études récentes de RNA-Seq, il est désormais établi, chez l'homme, que plus de **95% des pré-ARNm multi-exoniques subissent une étape d'épissage alternatif** au cours de leur maturation (Pan et al., 2008).

L'épissage alternatif implique l'utilisation différentielle de sites d'épissage. Il existe cinq types d'événements d'épissage alternatif : 1) **sauts d'exon** ou « exon cassette », 2) **exons mutuellement exclusifs**, 3) **sites donneurs alternatifs**, 4) **sites accepteurs alternatifs** et 5) **réentions d'intron** (Wang et al., 2008) (Figure 26). Plusieurs types d'événements d'épissage alternatif peuvent parfois être couplés, ce qui crée des événements d'épissage complexes (par exemple un site accepteur alternatif et un saut d'exon). La diversité des ARNm produits s'explique aussi par l'utilisation de promoteurs ou sites de polyadénylation alternatifs, respectivement responsables de l'inclusion (ou non) d'un premier ou dernier exon (Figure 26). Ces deux derniers événements sont dus à un changement d'initiation ou de terminaison de la transcription, et ne sont donc pas à proprement parler des événements d'épissage alternatif.

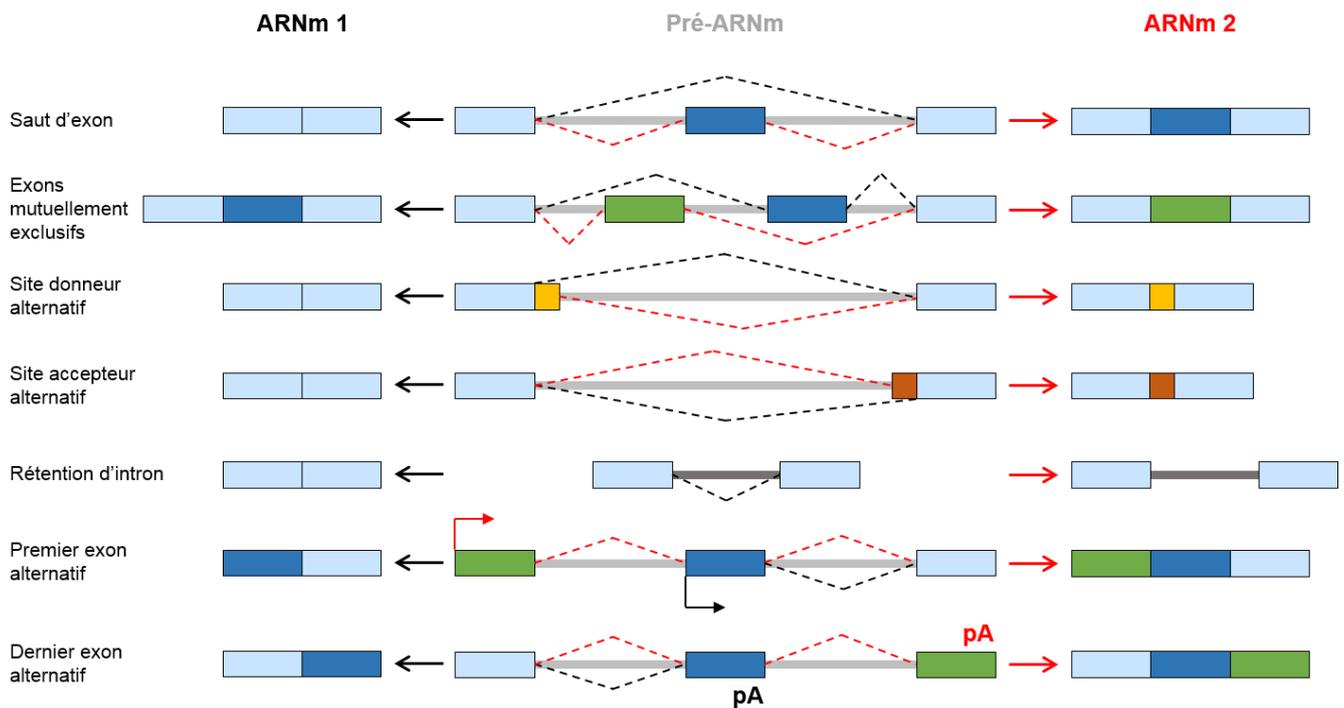


Figure 26 : Les différents types d'évènements d'épissage alternatif

La diversité des évènements d'épissage alternatif conduit à la formation de nombreux variants d'ARNm (**noir et rouge**) à partir d'un même **pré-ARNm**. L'évènement d'épissage alternatif le plus courant est le saut d'exon : il correspond à l'exclusion d'un exon dans l'ARNm final. Deux exons consécutifs sont dits mutuellement exclusifs lorsque l'inclusion d'un exon entraîne l'exclusion de l'autre. Le décalage en aval du site donneur d'épissage en 5' ou en amont du site accepteur d'épissage en 3' permet d'inclure des portions normalement non codantes dans l'ARNm final, ce qui crée des sites donneurs alternatifs ou accepteurs alternatifs respectivement. Il est possible de générer des sites ambigus, à mi-chemin entre le 5'ss et le 3'ss. Enfin, en cas de rétention d'intron, l'intron tout entier est inclus dans le transcrit mature. Les promoteurs (premier exon) et sites de polyadénylation (**pA**) (dernier exon) alternatifs sont très souvent rattachés aux évènements d'épissage alternatif.

Normalement considéré comme un évènement physiologique, l'épissage alternatif peut aussi être causé et/ou potentialisé par des mutations spécifiques conduisant à la formation et/ou à l'augmentation de l'utilisation de transcrits aberrants. C'est le cas des mutations *hotspot* du gène *SF3B1*.

2. Mécanisme de dérégulation de l'épissage dans les cancers *SF3B1*^{MUT}

Pour rappel, les protéines SF3B1 et p14 font partie du complexe SF3b qui stabilise la fixation de la snRNP U2 au niveau du BS de l'intron. SF3B1 interagit avec la sous-unité U2AF65 de l'hétérodimère U2AF, la sous-unité U2AF35 reconnaissant quant à elle le site accepteur en 3' composé d'une séquence adénine-guanine (AG).

En 2015 et 2016, plusieurs équipes ont démontré – grâce à des analyses de RNA-Seq – que les mutations *hotspot* du gène *SF3B1* entraînent l'**utilisation préférentielle d'un site accepteur alternatif** (ou cryptique) **localisé en amont du site accepteur constitutif** (ou canonique). Il est important de signaler que les travaux présentés dans ce paragraphe ne portent que sur l'analyse des jonctions différentiellement exprimées entre les échantillons *SF3B1*^{MUT} et *SF3B1*^{WT}, et ne s'intéressent donc qu'aux 5'ss et 3'ss alternatifs. En effet, DeBoever *et al.* ont tout d'abord étudié les données transcriptomiques obtenues à partir d'échantillons de LLC, MU et cancers du sein mutés ou non sur le gène *SF3B1* (DeBoever *et al.*, 2015). Parmi les 1 749 jonctions 5'ss et 3'ss différentiellement exprimées entre les échantillons *SF3B1*^{MUT} et *SF3B1*^{WT} (p-value corrigée Benjamini-Hochberg < 0,1), ils ont identifié 1 117 (64%) 3'ss cryptiques (nommés AG'). Plus de la moitié (619) de ces AG' se situent entre 10 et 30 paires de bases (pb) en amont de leur 3'ss canonique (nommé AG) correspondant. Ils sont tous davantage utilisés dans les échantillons *SF3B1*^{MUT}. En outre, la majorité de ces AG' sont retrouvés dans les trois différents types de tumeurs, suggérant un mécanisme identique de sélection de l'AG' par la protéine SF3B1 mutée quel que soit le tissu. Curieusement, seules les mutations *hotspot* de *SF3B1* localisées dans les répétitions HEAT H4 à H7 sont associées à ce type d'épissage aberrant. Pour rappel, les *hotspot* de *SF3B1* affectent des résidus proches sur le plan spatial et situés à l'intérieur de la protéine repliée, d'où probablement le même impact fonctionnel (Quesada *et al.*, 2012), contrairement aux mutations « non-*hotspot* ». Par exemple, la mutation *SF3B1*^{E902K} positionnée dans la région H9-H12 – éloignée des domaines de liaison au pré-ARNm – induit l'utilisation préférentielle d'un AG' situé en aval (et non en amont) de l'AG dans les cellules cancéreuses de la vessie (Seiler *et al.*, 2018a). D'après DeBoever *et al.*, les AG' sont positionnés entre le BS et l'AG, c'est-à-dire entre une région d'encombrement stérique et une région compétitive pour l'AG'. Toutefois, certains potentiels AG' répondant à ces critères ne sont pas utilisés dans les échantillons *SF3B1*^{MUT}.

Au cours de la même période, deux autres études de RNA-Seq portant sur des échantillons primaires de LLC, MU, mélanomes cutanés et cancers du sein (Darman et al., 2015) ou sur des échantillons primaires de LLC uniquement (Wang et al., 2016) confirment que les cancers avec des mutations *hotspots* de *SF3B1* sont associés à la sélection d'un AG' cryptique.

En 2015 et 2016, les équipes de Marc-Henri Stern à l'Institut Curie et Silvia Buonamici proposent un mécanisme plus détaillé de l'épissage alternatif induit en aval des mutations de *SF3B1* (Alsafadi et al., 2016; Darman et al., 2015). L'analyse des jonctions différentiellement exprimées à partir du RNA-Seq de 74 échantillons de MU confirme dans un premier temps les résultats de DeBoever *et al.*, soit l'utilisation préférentielle des AG' dans les tumeurs *SF3B1*^{MUT} (1 124 AG'/1 469 jonctions avec une p-value corrigée Benjamini-Hochberg $\leq 10^{-5}$, soit 76,5%) (Alsafadi et al., 2016). **La plupart des AG' (725) sont regroupés entre 12 et 24 pb en amont de l'AG.** Si on ne s'intéresse qu'aux AG' localisés à proximité de l'AG associé, 53% des AG' identifiés dans l'étude de DeBoever *et al.* sont également identifiés dans l'étude d'Alsafadi *et al.*, d'où la grande robustesse de ce pattern d'épissage malgré des échantillons et des pipelines d'analyse bio-informatique différents. Contrairement à DeBoever *et al.* qui ont restreint leurs travaux à des études *in silico*, Alsafadi *et al.* ont validé certains des AG' en transfectant notamment des lignées de MU *SF3B1*^{WT} ou *SF3B1*^{MUT} avec des minigènes. Comme attendu, l'analyse de fragments à partir des échantillons de RT-PCR montre bien l'utilisation préférentielle des AG' dans les cellules *SF3B1*^{MUT}. À noter que l'AG' est détecté à des niveaux très bas mais significatifs dans les cellules *SF3B1*^{WT}, ce qui indique que **les mutations de *SF3B1* n'induisent pas de novo l'utilisation d'un AG' mais renforcent son utilisation.** La transfection des lignées de MU avec un siSF3B1 n'ayant aucun impact sur le ratio d'utilisation de la jonction alternative, les mutations de *SF3B1* ne sont pas considérées comme des pertes de fonction de la protéine. De même, les mutations de *SF3B1* n'entraînent pas un gain de fonction de la protéine, puisque la transfection d'un vecteur d'expression *SF3B1*^{K700E} augmente fortement le ratio d'utilisation de la jonction alternative dans les cellules *SF3B1*^{WT}, tandis que la transfection d'un vecteur d'expression *SF3B1*^{WT} n'a aucun effet sur ce ratio. La surexpression de *SF3B1*^{K700E} dans les cellules *SF3B1*^{MUT} augmente quant à elle très peu le ratio d'utilisation de la jonction alternative, suggérant un effet saturant des mutations de *SF3B1*. Ainsi, **les mutations de *SF3B1* sont des mutations de type « changement de fonction »** (ou mutations « néomorphiques ») : la protéine SF3B1 mutée est capable de participer au processus

d'épissage du pré-ARNm comme la protéine SF3B1 sauvage, mais l'utilisation préférentielle des AG' génère des ARNm différents. Dans un deuxième temps, Alsafadi *et al.* ont démontré – grâce aux données générées *in silico* – une diminution de la fréquence du nucléotide G situé directement en aval de l'AG'. Cette guanine faisant partie du motif de reconnaissance pour U2AF35, il est probable que l'AG' reconnu dans le contexte *SF3B1*^{MUT} soit **moins dépendant d'U2AF** que l'AG' reconnu dans le contexte *SF3B1*^{WT}. Cette hypothèse a été vérifiée par transfection d'un siU2AF35 et d'un siU2AF65 dans des lignées cellulaires *SF3B1*^{WT} ou *SF3B1*^{MUT}; le ratio d'utilisation de la jonction alternative étant augmenté uniquement dans les cellules *SF3B1*^{WT}. Enfin, dans un dernier temps, les travaux d'Alsafadi *et al.* révèlent l'**utilisation d'un BS alternatif** (nommé BS') **dans les échantillons *SF3B1*^{MUT}**. Ce BS' se situe environ entre 10 et 14 pb en amont de l'AG' cryptique. De façon intéressante, l'appariement de bases entre le BS' et le snARN U2 est plus fort qu'entre le BS canonique et le snARN U2.

Pour résumer, l'utilisation préférentielle de l'AG' dans les cellules *SF3B1*^{MUT} repose sur trois propriétés : 1) un AG' cryptique avec une plus faible affinité pour U2AF que l'AG canonique, 2) la présence d'un BS' alternatif en amont de l'AG', 3) un BS' avec une plus grande affinité pour le snARN U2 que le BS canonique. Alsafadi *et al.* proposent finalement que les mutations *hotspot* de *SF3B1* pourraient induire un changement conformationnel de la snRNP U2, et donc altérer sa sélectivité pour le BS. Dans les conditions physiologiques, la snRNP U2 est recrutée au niveau du BS après liaison de la protéine U2AF35 au site accepteur. La fixation de la snRNP U2 contenant une protéine SF3B1 mutée n'est quant à elle plus dépendante de la liaison d'U2AF35 sur le site accepteur. Pour compenser ce manque d'interaction avec l'intron, la snRNP U2 contenant une protéine SF3B1 mutée recherche un BS' dit « fort », c'est-à-dire ayant une forte affinité pour le snARN U2. Ce BS' alternatif est situé en amont du BS canonique (Figure 27).

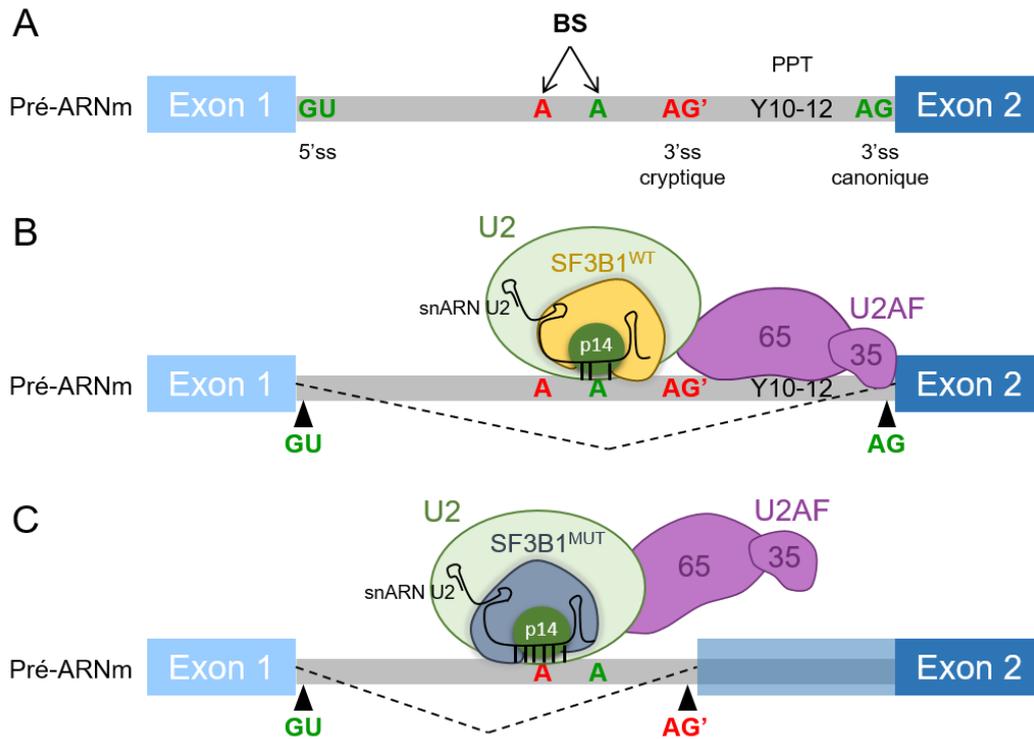


Figure 27 : Modèle de dérégulation de l'épissage induit par les mutations *hotspot* de *SF3B1* (d'après Alsafadi et al., 2016)

(A) et (B) Le 3'ss canonique (AG) est reconnu par U2AF35 et le *polypyrimidine tract* (PPT) par U2AF65. La snRNP U2 est ensuite recrutée au niveau du site de branchement (BS) canonique (A) grâce à l'interaction par appariement de bases entre le snARN U2 et le BS, et grâce aux interactions de SF3B1^{WT} et p14 de part et d'autre du BS. (C) Les mutations *hotspot* de *SF3B1* modifient la structure de la snRNP U2. Le complexe U2 contenant une protéine SF3B1^{MUT} est recruté sur l'intron indépendamment de la liaison d'U2AF35 au 3'ss cryptique (AG'), au niveau d'un BS cryptique (A). L'affinité du BS cryptique vis-à-vis du snARN U2 est plus forte que celle du BS canonique. 5'ss = site donneur d'épissage en 5', 3'ss = site accepteur d'épissage en 3'.

Tandis que la sous-unité U2AF35 interagit avec le site accepteur en 3', la sous-unité U2AF65 interagit avec le PPT. Darman *et al.* ont démontré que la séquence constituant le potentiel PPT alternatif en amont de l'AG' est courte et de faible affinité vis-à-vis d'U2AF65 (Darman et al., 2015). Ce PPT alternatif ne serait en fait pas capable de recruter le complexe U2AF. Ainsi, l'AG' est dépendant du PPT canonique situé en aval pour induire l'épissage alternatif. De plus, d'après les analyses bio-informatiques, la majorité des AG' ne sont pas localisés à une distance nucléotidique multiple de 3 de l'AG correspondant. **Le décalage du cadre de lecture suggère donc que de nombreux ARNm aberrants (≈ 50%) générés dans les cellules SF3B1^{MUT} sont dégradés par la voie NMD, d'où finalement une diminution de l'expression du gène canonique** (Figure 28). Cette hypothèse a été validée après

traitement de la lignée leucémique B Nalm-6 – transduite avec un vecteur $SF3B1^{WT}$ ou $SF3B1^{K700E}$ – avec du cycloheximide (CHX), un inhibiteur connu du NMD. Sous CHX, l'expression des gènes cibles du NMD est ré-augmentée au niveau ARNm dans les cellules $SF3B1^{K700E}$. **Les transcrits aberrants dont l'AG' est situé à une distance nucléotidique multiple de 3 de l'AG correspondant peuvent potentiellement être traduits en protéines aberrantes s'il n'y a pas d'inclusion d'exon stop (Figure 28).**

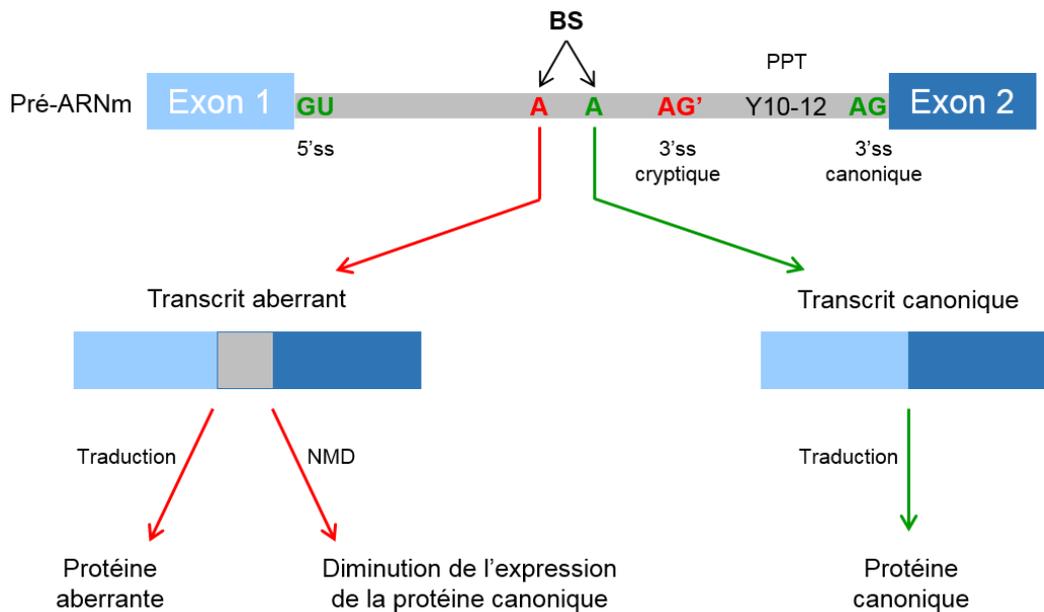


Figure 28 : Utilisation préférentielle d'un site de branchement alternatif dans les cellules $SF3B1^{MUT}$ et conséquences aux niveaux ARNm et protéines (d'après Darman et al., 2015)

Le décalage du site de branchement (**BS**) alternatif en amont du BS canonique induit un ajout de nucléotides dans l'exon 2 du transcrit aberrant. Si l'ajout de nucléotides est un multiple de 3 et/ou n'entraîne pas la formation d'un codon stop prématuré, le transcrit aberrant est potentiellement traduit en une protéine aberrante. Si l'ajout de nucléotides n'est pas un multiple de 3 et/ou entraîne la formation d'un codon stop prématuré, le transcrit aberrant est majoritairement dégradé par la voie NMD. **5'ss** = site donneur d'épissage en 5', **A** = BS alternatif, **A** = BS canonique, **PPT** = *polypyrimidine tract*, **3'ss** = site accepteur d'épissage en 3'.

Pour finir, il est important de signaler que la plupart des AG' cryptiques – bien que plus utilisés dans les cellules *SF3B1*^{MUT} comparées aux cellules *SF3B1*^{WT} – sont moins utilisés que l'AG canonique associé, y compris dans les cellules *SF3B1*^{MUT}. Est-ce que l'usage d'un nouveau 3'ss à un faible niveau peut avoir des conséquences biologiques significatives et, donc, expliquer les effets oncogéniques des mutations de *SF3B1* ? Une attention particulière devra également être portée à la petite proportion de gènes dont l'AG' devient le 3'ss principal.

3. Jonctions alternatives dans les SMD *SF3B1*^{MUT}

Tout comme dans les LLC, MU, mélanomes cutanés et cancers du sein, les mutations *hotspot* de *SF3B1* dans les SMD conduisent à l'usage préférentiel d'un AG' cryptique situé en amont de l'AG canonique (Dolatshad et al., 2016; Obeng et al., 2016). Dolatshad *et al.* ont analysé le RNA-Seq obtenu à partir de l'ARN des cellules hématopoïétiques immatures CD34⁺ (Dolatshad et al., 2016). Ils ont identifié 126 événements d'épissage alternatif significatifs (échantillons *SF3B1*^{MUT} *versus* *SF3B1*^{WT}; p-value $\leq 2.10^{-4}$), dont 42 AG'. Obeng *et al.* ont analysé le RNA-Seq obtenu à partir de l'ARN de l'ensemble des CMN de la moelle osseuse (Obeng et al., 2016). Ils ont identifié 134 événements d'épissage alternatif significatifs (échantillons *SF3B1*^{MUT} *versus* *SF3B1*^{WT}; q-value < 0,05), dont 91 AG'. Dans les deux études, les AG' les plus significatifs sont par exemple retrouvés dans les transcrits des gènes *ORAI2* (un échangeur transmembranaire de Ca²⁺), *TMEM14C* (une protéine de la membrane mitochondriale interne), *ENOSF1* (une enzyme du métabolisme glucidique) et *MAP3K7* (une sérine/thréonine kinase de la voie des MAPK).

De façon intéressante, une jonction alternative a été trouvée dans le transcrit du gène *ABCB7* (p-value < 2.10^{-4} ; *false discovery rate* = 0,006) (Dolatshad et al., 2016). Ce gène code pour une protéine de la membrane mitochondriale interne chargée d'exporter les clusters Fe/S vers le cytosol (cf. Chapitre I-I-E-2). Pour rappel, les mutations « perte de fonction » du gène *ABCB7* sont responsables du développement d'une anémie sidéroblastique congénitale (XLSA/A, cf. Chapitre I-II-A-2-a). L'AG' se situe entre les exons 8 et 9, à 21 pb en amont de l'AG. Il se positionne donc à une distance nucléotidique multiple de 3. Toutefois, l'un des sept codons intégrés dans le transcrit alternatif est un codon stop (uracile-adénine-guanine, UAG) (Figure 29). Par conséquent, le transcrit *ABCB7* alternatif est dégradé par la voie NMD, ce qui explique la diminution de l'expression du gène *ABCB7* au niveau ARNm dans les cellules CD34⁺ des patients *SF3B1*^{MUT} par rapport aux cellules CD34⁺ des patients

SF3B1^{WT} (Dolatshad et al., 2015). Les RT-PCR – à partir des ARN d'érythroblastes primaires cultivés *in vitro* – montrent en effet une augmentation de l'expression du transcrit aberrant dans les cellules *SF3B1*^{MUT} traitées avec du CHX par rapport aux cellules *SF3B1*^{MUT} non traitées (Dolatshad et al., 2016). En 2008, Boulwood *et al.* avaient déjà mis en évidence une expression diminuée du gène *ABCB7* au niveau ARNm dans les cellules CD34⁺ et les érythroblastes en culture des patients porteurs d'un SMD-SC par rapport aux patients porteurs d'un SMD sans SC ou aux témoins sains (cf. Chapitre I-II-A-2-b) (Boulwood et al., 2008). Ces résultats avaient été obtenus grâce à l'analyse des données issues de puces à ADN. Le même 3'ss cryptique d'*ABCB7* est retrouvé dans les CMN *SF3B1*^{MUT} de la moelle osseuse des patients atteints d'un SMD (q-value ≤ 0,05) (Obeng et al., 2016).



Figure 29 : Utilisation préférentielle du site 3' cryptique d'*ABCB7* dans les cellules SMD *SF3B1*^{MUT}

Le site 3' cryptique (**ag**) d'*ABCB7* se situe à 21 pb en amont du site 3' canonique (**ag**), entre les exons 8 et 9. Sept codons sont donc ajoutés dans le transcrit aberrant, dont un codon stop (**UAG**) (Dolatshad et al., 2016).

En résumé, les mutations du gène *SF3B1* entraînent l'utilisation préférentielle d'un AG' dans le transcrit du gène *ABCB7*. Le transcrit aberrant formé est dégradé par la voie NMD du fait de l'introduction d'un codon stop prématuré. L'expression du gène *ABCB7* est donc diminuée dans les cellules *SF3B1*^{MUT} par rapport aux cellules *SF3B1*^{WT}. Les clusters Fe/S seraient moins exportés vers le cytosol et s'accumuleraient alors dans les mitochondries des cellules *SF3B1*^{MUT}. Dolatshad *et al.* suggèrent que la diminution de l'expression d'*ABCB7* serait responsable du développement de la surcharge mitochondriale en fer chez les patients porteurs d'un SMD-SC.

Plusieurs points nuancent tout de même ce propos : 1) que ce soit dans les travaux de Dolatshad *et al.* ou Obeng *et al.*, l'AG' d'ABCB7 ne fait pas partie de la liste des jonctions alternatives les plus significatives ; 2) l'épissage aberrant d'ABCB7 n'est pas spécifique des SMD puisque le même AG' est retrouvé de façon significative dans les MU (p-value corrigée Benjamini-Hochberg = $4,70 \cdot 10^{-10}$; seuil de significativité $\leq 10^{-5}$) (Alsafadi *et al.*, 2016) et dans les LLC (q-value = 0,0247 ; seuil de significativité $\leq 0,05$) (Darman *et al.*, 2015) ; 3) le rôle exact d'ABCB7 dans la biosynthèse des clusters Fe/S n'est pas encore parfaitement connu. Des études fonctionnelles supplémentaires sont donc nécessaires pour appuyer l'éventuelle contribution du 3'ss cryptique d'ABCB7 à la physiopathologie des SMD-SC. De plus, l'expression de nombreux gènes impliqués dans la biosynthèse de l'hème – dont *FECH* et *ALAS2* – est modifiée dans les cellules *SF3B1*^{MUT}, sans qu'il y ait pour autant une jonction alternative identifiée. Est-ce que l'AG' d'ABCB7 peut expliquer à lui seul l'accumulation de fer dans les mitochondries et la formation de SC ? Ou est-ce que l'AG' d'ABCB7 reflète un phénomène sous-jacent, responsable de la diminution de l'expression d'ABCB7 et de la perturbation de l'ensemble du métabolisme du fer mitochondrial ? L'expression de la mutation hétérozygote *Sf3b1*^{K700E} chez la souris induit – tout comme chez l'homme – l'utilisation préférentielle de nouveaux AG' (48 AG'/72 jonctions alternatives ; q-value < 0,1) (Obeng *et al.*, 2016). Cependant, d'après les données de RNA-Seq, les souris KI *Sf3b1*^{K700E} n'expriment pas l'AG' d'ABCB7. En fait, très peu de jonctions alternatives identifiées dans les échantillons humains *SF3B1*^{MUT} ont également été identifiées dans les échantillons murins *Sf3b1*^{K700E}, puisque les **séquences introniques** sont **faiblement conservées entre les différentes espèces** (William Roy and Gilbert, 2006). Entre 20 et 68% des introns d'une espèce sont spécifiques de cette espèce.

En 2018, Shiozawa *et al.* se sont intéressés à l'ensemble des événements d'épissage alternatif induits en aval des mutations de *SF3B1* dans les SMD, contrairement aux précédentes équipes qui ne se focalisaient que sur les 5'ss et 3'ss cryptiques (Shiozawa *et al.*, 2018). Leur étude de RNA-Seq porte sur 165 échantillons d'ARN issus de CMN de la moelle osseuse et sur 100 échantillons d'ARN issus de cellules CD34⁺. La comparaison des échantillons *SF3B1*^{MUT} et *SF3B1*^{WT} révèle en tout 3 831 événements d'épissage alternatif, dont 538 communs entre les CMN et les cellules CD34⁺ (q-value < 0,01). Comme attendu, les mutations *hotspot* de *SF3B1* sont d'une part associées à une augmentation de l'utilisation des 3'ss cryptiques. Ces 3'ss cryptiques entraînant très souvent un décalage du cadre de lecture (56%), la moitié des transcrits aberrants (50%) sont dégradés par la voie NMD et – par

conséquent – l'expression du gène diminue. D'autre part, Shiozawa *et al.* montrent pour la première fois une **réduction de la rétention d'intron dans les cellules de SMD *SF3B1*^{MUT}**. Cette réduction de la rétention d'intron a été confirmée dans la lignée cellulaire HEK 293T CRISPR-Cas9 *SF3B1*^{K700E}, les transcrits conservant leur intron se situant majoritairement dans le noyau mais aussi dans le cytoplasme. Pour rappel, seuls les transcrits présents dans le cytoplasme peuvent être dégradés par la voie NMD. Le traitement des cellules HEK 293T CRISPR-Cas9 *SF3B1*^{K700E} avec du CHX ne conduit pas à une augmentation de l'expression des transcrits avec rétention d'intron. Ainsi, la réduction des transcrits avec rétention d'intron observée dans les cellules *SF3B1*^{MUT} n'est pas causée par leur dégradation par la voie NMD, mais est plutôt due à une potentialisation de l'excision des introns et/ou à un défaut d'export des transcrits avec rétention d'intron hors du noyau.

Enfin, les travaux de Huang *et al.* montrent que le profil d'épissage des progéniteurs érythroïdes CFU-E – cultivés *in vitro* à partir des cellules CD34⁺ transfectées avec un shSF3B1 – est différent de celui des cellules primaires de patients (Huang et al., 2018). La diminution de l'expression de *SF3B1* est principalement associée à des sauts d'exon et des sites de terminaison de la transcription alternatifs.

4. Épissage alternatif en aval des mutations des autres gènes du spliceosome

Chez les patients porteurs d'un SMD, les mutations des gènes *SRSF2*, *U2AF1* et *ZRSR2* causent elles aussi l'épissage alternatif de nombreux pré-ARNm.

Les mutations des gènes *SRSF2* et *U2AF1* modifient toutes deux l'inclusion des exons. En effet, *SRSF2* est une protéine à domaines SR impliquée dans la reconnaissance de l'exon *via* sa fixation à une séquence ESE (pour *exonic splicing enhancer*) (Liu et al., 2000). Elle est recrutée sur le pré-ARNm dès les premières étapes de l'épissage (complexe E du spliceosome). Les mutations *hotspot* de *SRSF2* entraînent un changement de son affinité et de sa spécificité de liaison à l'exon : des exons normalement inclus dans l'ARNm mature en sont exclus, ou inversement (Kim et al., 2015). Les mutations *hotspot* d'*U2AF1* S34 et Q157 altèrent quant à elles la reconnaissance du nucléotide situé en position -3 ou +1 du 3'ss respectivement (Ilagan et al., 2015). Les phénotypes d'épissage diffèrent donc en fonction du type de mutation. Enfin les mutations « perte de fonction » de *ZRSR2* affectent spécifiquement l'assemblage du spliceosome mineur, un complexe ribonucléoprotéique qui assure l'épissage des introns de type U12 (les introns de type U2 étant excisés par le spliceosome majeur) (Madan et al., 2015). Bien que représentant moins de 1% de l'ensemble des introns chez l'homme, les introns de type U12 sont souvent retrouvés dans des gènes impliqués dans des fonctions cellulaires fondamentales (réplication et réparation de l'ADN, transcription...) (Burge et al., 1998).

Partie 2

I. La protéine SF3B1 et les processus de réplication et de réparation de l'ADN

L'ADN est le support de l'information génétique. Il est pourtant soumis à de nombreux stress au cours de la vie de la cellule, qu'ils soient endogènes (ROS, erreurs des ADN polymérases lors de la réplication de l'ADN...) ou exogènes (radiations ionisantes, rayonnements ultraviolets (UV), agents alkylants, hydrocarbures aromatiques polycycliques...). Chaque jour, près de 10^5 lésions sont ainsi générées dans le génome des mammifères (Collins, 1999; Gaillard et al., 2015). Un programme de réparation est systématiquement mis en place dans les cellules, chaque type de lésion ayant son mécanisme de réparation spécifique. En cas de dommages à l'ADN persistants, les voies d'apoptose ou de sénescence cellulaires sont activées. Toutefois, certaines cellules échappent à la réponse aux dommages à l'ADN : des mutations et/ou des réarrangements chromosomiques surviennent alors. Ils sont sources d'instabilité génomique.

Plusieurs papiers parus ces dernières années révèlent que la protéine SF3B1 participe au **contrôle de la réplication et de la réparation de l'ADN**. Cette partie reviendra donc, dans un premier temps, sur les causes et les mécanismes de résolution du stress réplicatif. Nous verrons que les « senseurs » des dommages à l'ADN et les « effecteurs » de la réparation de l'ADN sont très nombreux lors de la phase S du cycle cellulaire. L'objectif ici n'est pas d'en dresser la liste exhaustive, mais de replacer les principales protéines mises en jeu dans ces processus cellulaires complexes, de façon à mieux comprendre, dans un second temps, le rôle de SF3B1 dans le maintien de la stabilité génomique. Une attention particulière sera portée à l'implication du stress réplicatif dans le changement de fonctionnalité des CSH.

A. Stress réplicatif

Chaque fois qu'une cellule de mammifère se divise, des milliards de nucléotides doivent auparavant être dupliqués avec précision. La **réplication de l'ADN** désigne le processus au cours duquel la molécule d'ADN parent est copiée en deux molécules d'ADN fille identiques. Elle est initiée au niveau des origines de réplication, sur lesquelles ont été chargés les complexes hélicases MCM2-7 (pour *minichromosome maintenance*) lors de la phase G1 du cycle cellulaire. Sous l'action des kinases CDK et DDK (pour *Dbf4-dependent kinase*), les origines de réplication sont activées (ou « allumées ») pendant la phase S du cycle

cellulaire et les deux brins d'ADN parent s'ouvrent. Dans les conditions physiologiques, seules 10% des origines de réplication associées à MCM2-7 sont activées, les autres restant « en sommeil » (Ge and Blow, 2010). Puis, le réplisome – un complexe multiprotéique contenant notamment l'ADN polymérase – progresse dans les deux sens à partir de l'origine créant ainsi deux fourches de réplication. La réplication de l'ADN se termine lors de la rencontre des fourches de réplication (pour revue Gaillard et al., 2015; pour revue Zeman and Cimprich, 2014).

Bien que fondamentale pour la cellule, la réplication de l'ADN est un véritable challenge puisque l'avancée des ADN polymérases répliquatives (α , δ et ϵ) peut être perturbée par de nombreux obstacles. En cas d'**altération de la dynamique de réplication de l'ADN** – c'est-à-dire en cas de ralentissement voire d'arrêt des fourches de réplication – la cellule subit un **stress répliquatif** et accumule de l'ADN simple-brin (ADNsb) protégé par la protéine de réplication A (RPA) (pour revue Zeman and Cimprich, 2014). De façon à éviter des dommages secondaires de type cassures double-brin (CDB), des mécanismes de résolution du stress répliquatif doivent rapidement être mis en place.

1. Causes du stress répliquatif

Les causes du stress répliquatif sont variées. Cinq d'entre elles seront développées : la quantité limitante de désoxyribonucléotides (dNTPs), l'interférence entre la réplication et la transcription de l'ADN, les structures secondaires complexes, les lésions de l'ADN et les sites fragiles communs (pour revue Zeman and Cimprich, 2014) (Figure 30).

Les dNTPs représentent l'unité de construction des acides nucléiques. La **diminution du pool de dNTPs** dans une cellule – par exemple après traitement à l'hydroxyurée (HU), une drogue qui inhibe la ribonucléotide réductase (RNR) responsable de la transformation des ribonucléotides en dNTPs – empêche donc logiquement la synthèse d'un nouvel ADN.

Outre la diminution des éléments composant l'ADN, il existe des barrières physiques à la progression des fourches de réplication, tels que les **R-loops** (ou boucles R). Les *R-loops* sont des structures nucléiques à trois brins : un hybride ADN/ARN, et un ADNsb déplacé identique à la molécule d'ARN (pour revue Aguilera and García-Muse, 2012). Ils sont formés lors de la transcription de l'ADN par l'ARN polymérase II, le transcrit naissant s'hybridant avec la matrice d'ADN. Bien qu'étant impliqués dans des processus physiologiques – comme la commutation de classe des immunoglobulines (Yu et al., 2003) –, les *R-loops* en excès

entraînent une instabilité génomique puisqu'ils favorisent la collision entre les machineries de réplication et de transcription de l'ADN, ce qui bloque l'avancée des ADN polymérases. Les cellules saines utilisent donc plusieurs mécanismes pour réguler la formation des *R-loops* : les ribonucléases H (RNases H1 et H2) qui hydrolysent spécifiquement le fragment d'ARN (Cerritelli and Crouch, 2009) et les hélicases – comme l'hélicase Aquarius (AQR) – qui déroulent les hybrides (Sollier et al., 2014). Chez les mammifères, la RNase H1 hydrolyse les longs fragments d'ARN produits lors de l'élongation de l'ARNm par l'ARN polymérase II, tandis que la RNase H2 hydrolyse chaque ribonucléotide anormalement incorporé à l'ADN (Parajuli et al., 2017; Reijns et al., 2012). La RNase H1 est localisée dans les mitochondries et dans le noyau. Il a de plus récemment été montré que la PrimPol (pour *primase and DNA directed polymerase*) limite la formation des *R-loops* pendant la phase S (Šviković et al., 2018). Dans les cellules eucaryotes, la PrimPol possède une activité primase et polymérase (García-Gómez et al., 2013), notamment lors de la synthèse translésionnelle (cf. Chapitre II-Partie 2-2-a).

L'ADN n'adopte pas toujours sa conformation la plus courante de double hélice mais peut former des motifs plus complexes que les fourches de réplication ne franchissent pas. Par exemple, les **G-quadruplex** (G4), structures secondaires à quatre brins parfois retrouvées dans les ADN riches en guanine (G), sont déroulés en amont des fourches de réplication par des hélicases, comme l'hélicase de la voie « anémie de Fanconi » (AF) FANCD1 (Castillo Bosch et al., 2014). Si le G4 n'est pas déroulé, la réplication de l'ADN est arrêtée. Les *R-loops* se forment surtout au niveau des promoteurs de gènes riches en bases G/C (guanine/cytosine) et lorsque l'ADNsb déplacé (donc non matrice) se replie en G4 (Chen et al., 2017). La PrimPol initie le réamorçage de la réplication de part et d'autre des structures G4, de façon à poursuivre la synthèse de l'ADN et ainsi réduire la formation des *R-loops* (Šviković et al., 2018)

Les dommages à l'ADN, de type **adduits à l'ADN, ponts inter- et/ou intra-brins et cassures simple- et/ou double-brin**, sont également l'une des causes majeures de stress répliatif en cas de non réparation avant l'arrivée des fourches de réplication. Comme précisé auparavant, des agents endogènes et exogènes peuvent créer ces dommages à l'ADN.

Il existe enfin des régions au sein du génome où surviennent très souvent des cassures : les **sites fragiles communs** (SFC). Ces SFC enrichis en dinucléotides adénine-thymidine sont, par contre, appauvris en origines de réplication (pour revue Le Tallec et al.,

2014). Lorsqu'une fourche de réplication est stoppée au niveau des SFC à cause d'une collision avec la machinerie de transcription des ARN, d'une structure secondaire complexe et/ou d'une lésion de l'ADN, l'initiation d'une nouvelle fourche n'est donc pas possible.

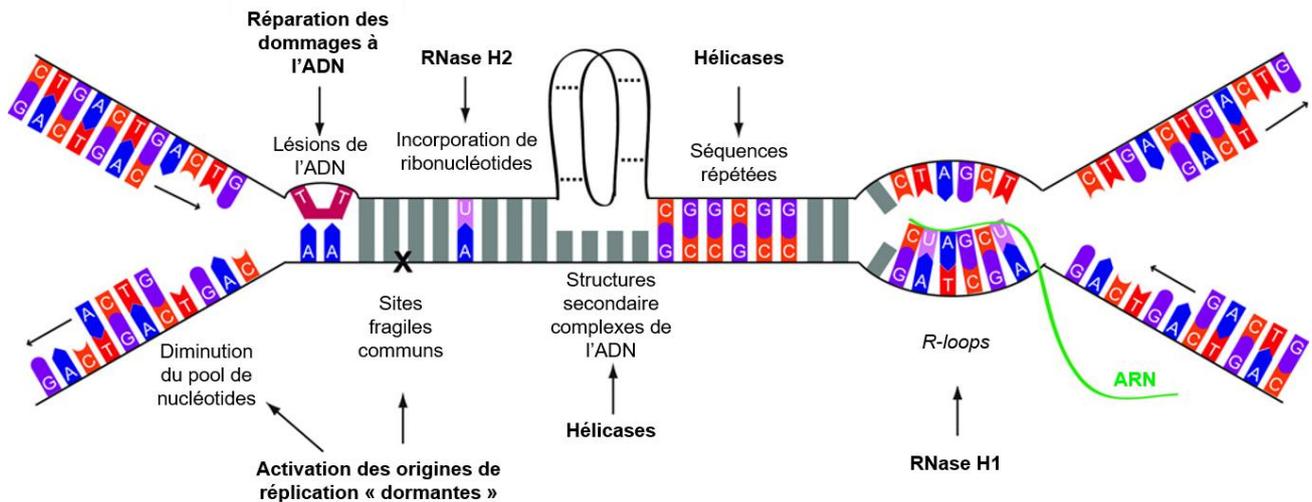


Figure 30 : Causes du stress réplcatif (d'après Zeman et Cimprich, 2014)

De nombreux facteurs peuvent ralentir la progression des fourches de réplication : une quantité limitante de nucléotides, des lésions de l'ADN, des sites fragiles communs, des structures secondaires complexes de l'ADN, des séquences répétées au sein de l'ADN, une incorporation de ribonucléotides à la place des désoxyribonucléotides, des *R-loops* etc. Les voies de résolution connues pour chaque cause de stress réplcatif sont indiquées en gras.

2. Réponses cellulaires au stress réplcatif

Le **checkpoint** (ou contrôle) **de la phase S** est activé en réponse aux dommages à l'ADN en cours de réplication. Il comprend principalement les voies **ATR/CHK1** et **ATM/CHK2**. ATR (pour *Ataxia telangiectasia and rad3-related protein*) et ATM (pour *Ataxia telangiectasia mutated*) font partie de la superfamille des kinases apparentées aux PI3K. Le checkpoint de la phase S permet, entre autres, l'arrêt réversible du cycle cellulaire, la stabilisation des fourches de réplication et la régulation de l'activation des origines de réplication. L'objectif est d'assurer rapidement le redémarrage des fourches de réplication bloquées. En cas de lésions de l'ADN trop importantes, le facteur de transcription p53 est activé par phosphorylation et induit l'apoptose et/ou la sénescence des cellules.

a. Voie de signalisation ATR/CHK1

La voie de signalisation ATR/CHK1 peut être activée en fin de phase G1 et en phase S du cycle cellulaire. Lorsque la fourche de réplication est ralentie ou bloquée, de longues séquences d'**ADNsb** apparaissent du fait, notamment, du découplage entre l'ADN polymérase et les hélicases MCM2-7. L'hétérotrimère **RPA** (RPA32, RPA70 et RPA14) recouvre l'ADNsb, ce qui permet le recrutement de deux complexes au niveau des foyers de dommages à l'ADN : le complexe **ATR/ATRIP** (pour *ATR-interacting protein*) via l'interaction entre ATRIP et l'ADNsb chargé de RPA (pour revue Zou and Elledge, 2003), et l'**hétérotrimère 9/1/1** (RAD9/HUS1/RAD1) par l'intermédiaire du complexe RAD17/facteur de réplication C (RFC) via, cette fois-ci, l'interaction entre RAD17/RFC2-5 et l'ADNsb chargé de RPA (Zou et al., 2003). La protéine RAD9 du complexe 9/1/1 étant associée à la protéine **TopBP1** (pour *DNA topoisomerase 2-binding protein 1*), cette dernière est à son tour recrutée sur les sites d'ADN endommagé. TopBP1 stimule l'activation de la sérine-thréonine kinase ATR (Kumagai et al., 2006). Parmi les cibles d'ATR figurent **RAD9, RAD17, TopBP1, RPA, FANCI, FANCD2, FANCM, MCM2 et CHK1** (Figure 31). Les conséquences de la phosphorylation des protéines effectrices du *checkpoint* de la phase S vont être détaillées dans les paragraphes ci-dessous.

La phosphorylation des trois protéines RAD9, RAD17, TopBP1 permet d'augmenter la fixation de TopBP1 et de 9/1/1 à la fourche de réplication bloquée, et ainsi de créer une boucle positive d'activation d'ATR. Concernant RPA, il a été montré en 2006 que la phosphorylation des sérines 4, 8 et 33 de sa sous-unité RPA32 (ou RPA2) était diminuée lorsque l'activité d'ATR était inhibée (Olson et al., 2006). RPA phosphorylé favorise le ralentissement de la réplication de l'ADN en cas de stress. Il limite également la formation des *R-loops* et participe à leur résolution. En effet, les travaux d'une équipe américaine révèlent que l'association de phospho-RPA avec le fragment d'ADNsb des *R-loops* (Leong et al., 2017; Nguyen et al., 2017) assure le recrutement et l'activation de la RNase H1, l'enzyme responsable de la dégradation du fragment ARN. En outre, la sous-unité RPA70 de RPA est capable de mobiliser la PrimPol au niveau des fourches de réplication bloquées, puis d'activer sa fonction primase (Guilliam et al., 2017).

ATR phosphoryle de plus des protéines de la voie AF : FANCI, FANCD2 et FANCM (une des huit sous-unités du complexe ubiquitine ligase E3 FANCA-M). Les protéines de la voie AF sont activées en cas de blocage de la fourche de réplication par des pontages

intercaténaire. La phosphorylation de FANCI induit la mono-ubiquitinylation de FANCD2 et lui-même par FANCA-M (Ishiai et al., 2008). Les protéines FANCI et FANCD2 mono-ubiquitinylées sont alors mobilisées sur la chromatine endommagée et facilitent la réparation des ponts inter-brins en recrutant des endo- et exo-nucléases. En outre, en 2013, une équipe collaboratrice de Montpellier a montré une interaction transitoire entre FANCD2 phosphorylé – mais non mono-ubiquitinylé – et les hélicases MCM2-7 (Lossaint et al., 2013). Lorsque le pool de nucléotides diminue au sein de la cellule, FANCD2 restreint la synthèse de l'ADN de façon à limiter la formation d'ADNsb. FANCD2 est donc directement impliqué dans la surveillance du réplisome. L'hélicase MCM2 du complexe MCM2-7 fait aussi partie des cibles d'ATR, sa phosphorylation participant *in fine* à la stabilisation des fourches de réplication bloquées (Cortez et al., 2004).

Enfin, la sérine-thréonine kinase CHK1 est l'une des principales protéines « effectrices » de la réponse aux dommages à l'ADN. Son activation nécessite – en plus de la phosphorylation de ses sérines 317 et 345 par ATR – le recrutement de trois médiateurs sur l'ADNsb : le complexe **TIMELESS/TIPIN** (pour *Timeless-interacting protein*) et la **claspine** (Ünsal-Kaçmaz et al., 2007; Yoshizawa-Sugata and Masai, 2007). Le complexe TIMELESS/TIPIN – qui interagit avec RPA – facilite le chargement de la claspine sur la chromatine endommagée. La claspine interagit directement avec CHK1 et favorise sa phosphorylation par ATR. (Kemp et al., 2010). Au cours de la phase M du cycle cellulaire, la claspine poly-ubiquitinylée est dégradée et CHK1 est inactivée (Freire et al., 2006). D'ailleurs, le blocage de la claspine – notamment par des anticorps anti-claspine – entraîne un défaut de phosphorylation de CHK1, et donc un checkpoint non fonctionnel (Kumagai and Dunphy, 2000; Mailand et al., 2006). À noter que le complexe TIMELESS/TIPIN intervient aussi dans la stabilisation des fourches de réplication bloquées.

La voie de signalisation ATR/CHK1 conduit d'une part à l'arrêt du cycle cellulaire par phosphorylation des phosphatases **CDC25** (pour *Cell division cycle 25*) et du facteur de transcription **p53**. CDC25A déphosphoryle CDK1 et CDK2 tandis que CDC25B et CDC25C déphosphorylent uniquement CDK1. La phosphorylation des CDC25 par CHK1 inhibe l'activation des complexes cycline B/CDK1 et cycline E/CDK2 et, par conséquent, les cellules sont bloquées en phases G2 et G1 respectivement (Donzelli and Draetta, 2003; Patil et al., 2013; Uto et al., 2004). La phosphorylation de p53 aboutit elle aussi à l'inhibition des CDK, mais par l'intermédiaire de l'induction de l'expression de p21 (Dulić et al., 1994). La

voie de signalisation ATR/CHK1 régule d'autre part l'activation des origines de réplication dites « dormantes » situées à proximité du site de blocage de la fourche active. Par contre, elle inhibe l'initiation de nouvelles fourches de réplication, de façon à limiter les dommages à l'ADN (Ge and Blow, 2010).

La résolution du stress réplicatif nécessite donc la collaboration de dizaines de protéines. Les cellules ont par ailleurs développé des mécanismes de tolérance aux dommages à l'ADN orchestrés par l'ubiquitinylation de l'antigène nucléaire de prolifération cellulaire (PCNA). Lorsqu'il est mono-ubiquitinylé par le complexe RAD6/RAD18, PCNA est mobilisé sur l'ADNsb recouvert de RPA, ce qui permet le recrutement des ADN polymérases translésionnelles (Davies et al., 2008). Ces dernières ajoutent des nucléotides – corrects ou incorrects – en face de la lésion de l'ADN : c'est la synthèse translésionnelle. Cette voie de tolérance qui assure la complétion de la réplication de l'ADN en présence de dommages peut être source d'erreurs. En 2008, il a été montré que CHK1, TIMELESS et la claspine contrôlent la mono-ubiquitinylation de PCNA indépendamment d'ATR (Yang et al., 2008). Lorsqu'il est poly-ubiquitinylé, PCNA recrute l'ADN translocase ZRANB3 de façon à initier l'inversion de la fourche de réplication, c'est-à-dire la conversion d'une fourche à trois voies à une fourche à quatre voies, la 4^{ème} voie étant le brin en cours d'inversion (Vujanovic et al., 2017).

L'arrêt prolongé des fourches de réplication peut conduire à l'effondrement de ces dernières, ce qui génère des CDB de l'ADN. La voie de signalisation ATM/CHK2 est alors stimulée. Le mécanisme de l'effondrement des fourches de réplication et de la formation des CDB secondaires n'est pas encore clairement élucidé. Néanmoins, il existe diverses hypothèses, non exclusives l'une de l'autre : 1) la dissociation des composants du réplisome ; 2) la digestion des fourches de réplication bloquées ou inversées par des endonucléases, comme par exemple l'endonucléase MUS81 qui convertit les fourches bloquées en des intermédiaires plus aptes à la réparation par recombinaison homologue (RH, cf. paragraphe ci-dessous) (Hanada et al., 2007) et 3) la digestion de l'ADNsb par des endonucléases, ou sa rupture passive (pour revue Zeman and Cimprich, 2014). L'ADNsb est présent au niveau de la fourche de réplication bloquée, mais également au niveau des interstices formés après redémarrage de la machinerie de réplication en aval de la lésion (Lopes et al., 2006).

b. Voie de signalisation ATM/CHK2

La voie de signalisation ATM/CHK2 peut être activée tout au long du cycle cellulaire suite aux CDB de l'ADN. Tout comme RPA qui se fixe à l'ADNsb, il existe des « senseurs » des CDB de l'ADN, dont le complexe MRE11/RAD50/NBS1 (MRN) (Lamarche et al., 2010). Le complexe **MRN** est rapidement recruté au niveau du site d'ADN endommagé – de façon à stabiliser les extrémités cassées de l'ADN –, puis recrute à son tour la kinase multimérique **ATM**. Lors de son activation, ATM autophosphoryle sa sérine 1981 (So et al., 2009). ATM activé phosphoryle ensuite les trois membres du complexe MRN (Lavin et al., 2015) ainsi que le variant d'histone **H2AX** sur sa sérine 139, un autre senseur non spécifique des CDB de l'ADN (Burma et al., 2001). La phosphorylation de H2AX est nécessaire au recrutement d'une protéine médiatrice : **MDC1** (Stucki et al., 2005). Enfin, H2AX phosphorylé (γ -H2AX) et MDC1 collaborent toutes les deux pour mobiliser les protéines effectrices de la réparation des CDB de l'ADN, dont **BRCA1** (pour *Breast cancer type 1 susceptibility protein*), **53BP1** (pour *TP53-binding protein 1*) et **CHK2**. Ces trois protéines sont phosphorylées et activées par ATM (Figure 31).

Deux mécanismes principaux entrent en compétition pour la réparation des CDB de l'ADN : la RH ayant lieu uniquement pendant les phases S et G2 du cycle cellulaire, et la ligature d'extrémités non homologues (NHEJ pour *non-homologous end joining*) surtout employée en phase G0/G1 et en début de phase S du cycle cellulaire. La RH utilise une séquence homologue intacte – présente sur la chromatide sœur – pour réparer la CDB et est initiée par une résection d'ADN. L'ADNsb formé est protégé par RPA, ensuite remplacé par la recombinaise RAD51. Brièvement, l'endo- et exonucléase MRE11 coopère avec l'endonucléase CtIP (ou RBBP8 pour *RB binding protein 8*) pour éliminer quelques dizaines de nucléotides à l'extrémité de la CDB. Après phosphorylation par les CDK, CtIP est en effet recrutée sur l'ADN endommagé grâce à son interaction avec BRCA1 et le complexe MRN phosphorylés (Sartori et al., 2007). Au contraire, la NHEJ ne nécessite pas d'homologie de séquence et permet la ligature des extrémités d'ADN sans étape préalable de résection de l'ADN. Deux substrats d'ATM jouent un rôle important dans le choix du mécanisme de réparation des CDB : 53BP1 qui bloque la résection de l'ADN, et BRCA1 qui élimine 53BP1 des extrémités cassées de l'ADN (Hustedt and Durocher, 2016). À noter que BRCA1 peut également être phosphorylée par la kinase ATR.

Il existe en fait de nombreux parallèles entre les voies de signalisation ATM/CHK2 et ATR/CHK1, puisqu'elles sont par exemple toutes les deux capables d'induire l'arrêt du cycle cellulaire par phosphorylation des phosphatases CDC25 et/ou par phosphorylation de p53. De même, la kinase CHK1 est indispensable pour la résolution du stress répliatif mais aussi pour la RH. En effet, la phosphorylation de **RAD51** sur la thréonine 309 dépend de CHK1 (Sørensen et al., 2005). RAD51 catalyse l'invasion de brin lors du processus de RH, c'est-à-dire qu'elle permet la recherche de la séquence homologue dans le noyau et l'appariement des brins d'ADN homologues (Krishna et al., 2007). La lignée de mammifère CHO exprimant un dominant négatif de RAD51 et soumise à des rayonnements UV de type C est particulièrement sujette à la mutagenèse (Lambert and Lopez, 2002). Ceci montre que la fonction de RAD51 est essentielle au processus de RH. RAD51 participe de plus à l'inversion des fourches de réplication, et ce indépendamment de PCNA poly-ubiquitinylé (cf. Chapitre II-Partie 2-I-A-2-a) (Zellweger et al., 2015). Le brin en cours d'inversion est ici protégé par BRCA2 d'une éventuelle résection par MRE11 (Kolinjivadi et al., 2017). Enfin, RAD51 assure le redémarrage des fourches de réplication bloquées (Petermann et al., 2010).

En conclusion, la dynamique de réplication de l'ADN doit être finement régulée par les *checkpoints* de la phase S, puisque l'effondrement des fourches de réplication peut avoir des conséquences désastreuses pour la cellule. Le stress répliatif fait d'ailleurs partie des « marqueurs précoces » de la tumorigenèse (Gorgoulis et al., 2005). Par exemple, l'activation des proto-oncogènes ou l'inhibition des gènes suppresseurs de tumeurs peut entraîner la dérégulation des kinases CDK, et donc l'activation anormale d'un grand nombre d'origines de réplication de l'ADN (pour revue Macheret and Halazonetis, 2015). Les réplisomes entrant en compétition vis-à-vis d'une quantité limitante de dNTPs, les fourches de réplication ralentissent, ce qui favorise l'instabilité génomique. En outre, plusieurs équipes ont récemment pointé le rôle du **stress répliatif** dans le **déclin progressif des CSH** (Flach et al., 2014; Walter et al., 2015). Cette notion est d'autant plus intéressante que les **SMD** sont des **pathologies du sujet âgé**.

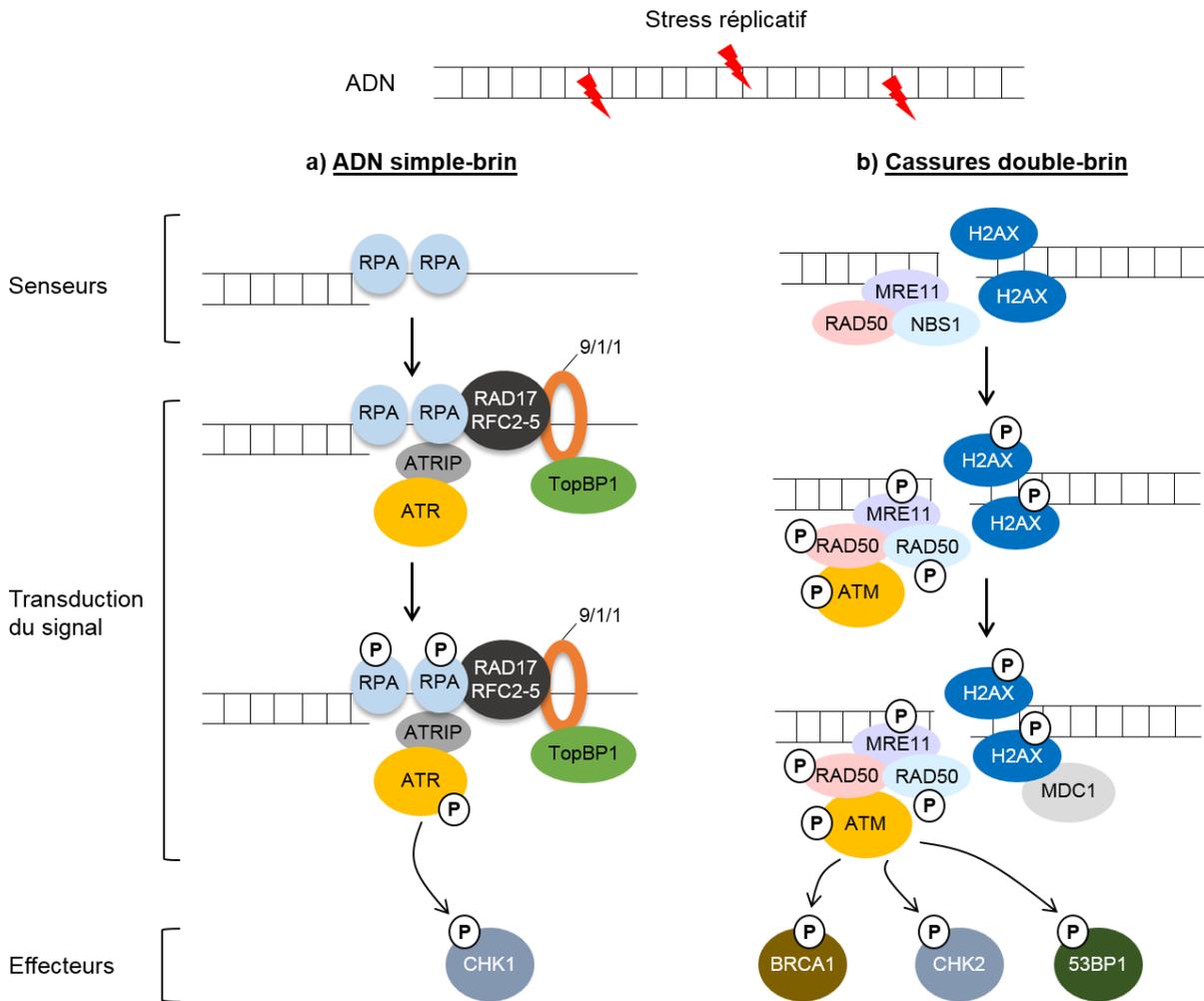


Figure 31 : Réponses cellulaires aux dommages à l'ADN induits par le stress réplcatif

En cas de stress réplcatif, les voies ATR/CHK1 et/ou ATM/CHK2 sont activées. Elles assurent le *checkpoint* de la phase S. **(a)** Le ralentissement ou l'arrêt des fourches de réplcation entraîne la formation d'ADN simple-brin (ADNsb) reconnu par la protéine RPA. Les complexes ATR/ATRIP, RAD17/RFC2-5 et 9/1/1 ainsi que la protéine TopBP1 colocalisent avec l'ADNsb recouvert de RPA. TopBP1 participe à l'activation de la kinase ATR. ATR activée phosphoryle de nombreuses protéines effectrices, dont la kinase CHK1. **(b)** L'arrêt prolongé des fourches de réplcation peut conduire à la formation de cassures double-brin de l'ADN (CDB). Le complexe MRN (pour MRE11/RAD50/NBS1) et l'histone H2AX sont mobilisés au niveau des sites de CDB, ce qui permet le recrutement de la kinase ATM. ATM activée phosphoryle le complexe MRN et H2AX. La protéine médiatrice MDC1 est alors à son tour mobilisée sur l'ADN endommagé, ce qui permet le recrutement des protéines effectrices CHK2, BRCA1 et 53BP1. CHK2, BRCA1 et 53BP1 sont toutes trois des cibles de la kinase ATM. *In fine*, le *checkpoint* de la phase S contrôle l'arrêt réversible du cycle cellulaire, la stabilisation puis la réparation des fourches de réplcation ou l'apoptose des cellules trop endommagées.

3. Stress répliatif et vieillissement du système hématopoïétique

En 2014, l'équipe d'Emmanuelle Passegué a étudié la dynamique de réplication de l'ADN dans les CSH des souris à différents âges (6-12 semaines ou 22-30 mois) (Flach et al., 2014). Il s'avère que les CSH isolées à partir des souris les plus âgées accumulent des foyers intranucléaires de γ -H2AX, RPA et ATRIP visibles par IF. De plus, une analyse transcriptomique – validée par RT-qPCR à partir de l'ARN des CSH – révèle une diminution avec l'âge de l'expression des gènes *Mcm4* et *Mcm6* qui, comme chez l'homme, codent pour des hélicases nécessaires à l'activation des origines de réplication et au déroulement de l'ADN en amont des fourches de réplication. Enfin, les CSH jeunes traitées *in vitro* pendant 36 h avec un inhibiteur des ADN polymérase répliatives – l'aphidicoline –, puis greffées dans des souris présentent un défaut majeur de reconstitution de la moelle osseuse au même niveau que les CSH âgées traitées ou non avec l'aphidicoline. Pour résumer, un **déficit en protéines MCM** s'installe progressivement dans les CSH des souris. L'initiation et l'élongation de la réplication de l'ADN sont alors perturbées, d'où la perte de fonctionnalité des CSH au cours du temps (Figure 32). Les raisons de la diminution de l'expression des gènes *Mcm* ne sont à ce jour pas connues. Curieusement, les CSH âgées accumulent des foyers γ -H2AX en cours de réplication mais aussi à l'état quiescent. Le marquage γ -H2AX à l'état quiescent se localise dans les nucléoles – usines de transcription des ARN ribosomiques (ARNr) – et est justement associé à une diminution de la biogenèse des ribosomes dans les CSH âgées (Figure 32). Des études supplémentaires devront être réalisées pour montrer une éventuelle implication de l'inactivation de la transcription des ARNr dans la défaillance de la moelle osseuse. Toutefois, ce point permet d'insister sur le fait que les modifications épigénétiques de l'histone H2AX ne sont pas exclusivement des marqueurs des dommages à l'ADN. Chez l'homme, le vieillissement semble également affecter la réponse aux dommages à l'ADN des CSH, puisque les CSH âgées (moelles osseuses, > 70 ans) accumulent plus de foyers intranucléaires de γ -H2AX après irradiation que les CSH jeunes (cordons) (Rübe et al., 2011). La forte variabilité interindividuelle chez la population d'âge moyen (< 50 ans) laisse toutefois penser que des mécanismes supplémentaires intrinsèques aux CSH et/ou des facteurs extrinsèques contribuent à l'accumulation de dommages à l'ADN associés à l'âge.

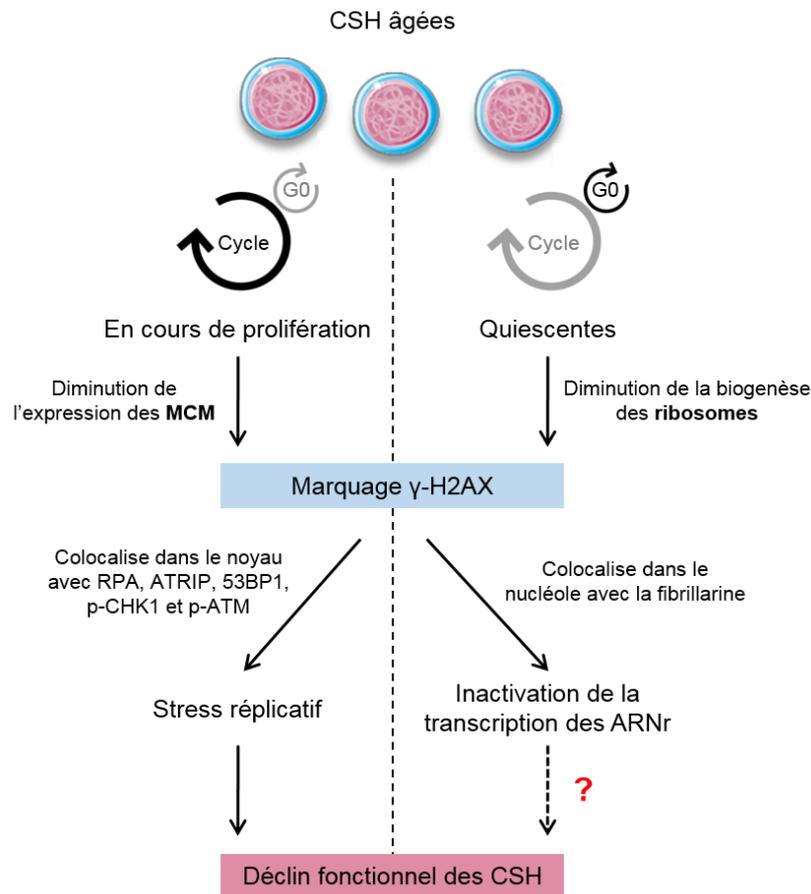


Figure 32 : L'accumulation de foyers intranucléaires γ -H2AX est un marqueur de la perte de fonctionnalité des CSH murines âgées (d'après Flach et al., 2014)

Lorsque les CSH âgées sont engagées dans le cycle cellulaire, les foyers intranucléaires de γ -H2AX colocalisent avec les foyers intranucléaires de RPA, ATRIP, 53BP1, p-CHK1 et p-ATM. Les CSH âgées subissent un stress réplcatif du fait de la diminution de l'expression des hélicases MCM au cours du temps. Ce stress réplcatif contribue au déclin fonctionnel des CSH murines au cours du temps. Lorsque les CSH âgées sont quiescentes, les foyers intranucléaires de γ -H2AX colocalisent cette fois-ci avec les foyers intranucléaires de fibrillarine, protéine du nucléole. Le marquage γ -H2AX dans les nucléoles des CSH âgées quiescentes est dans ce cas associé à l'inactivation de la transcription des ARNr.

En plus du stress réplicatif, le stress oxydatif dans la niche médullaire conduit petit à petit au vieillissement du système hématopoïétique. Pour rappel, à l'état physiologique, la plupart des CSH sont quiescentes dans la moelle osseuse. Leur ADN n'est donc pas dupliqué et leurs besoins énergétiques sont faibles, d'où une utilisation minimale de la chaîne respiratoire mitochondriale. Lorsque les CSH entrent dans le cycle cellulaire en réponse à un quelconque stress – par exemple des infections, des saignements – les processus de réplication de l'ADN et de production d'énergie par phosphorylation oxydative sont initiés de manière coordonnée. En 2015, les études *in vivo* de Walter *et al.* montrent que l'utilisation accrue de la chaîne de transport d'électrons mitochondriale est corrélée à une **surproduction de ROS dans les CSH qui prolifèrent** (Walter et al., 2015). Produits en excès, ces ROS peuvent former des **adduits à l'ADN** – tels que la 8-oxo-2'-désoxyguanosine – et donc bloquer l'avancée des fourches de réplication. Les expériences d'IF révèlent effectivement l'augmentation significative des foyers intranucléaires de γ -H2AX et 53BP1 – marqueurs des dommages à l'ADN – dans les CSH qui prolifèrent par rapport aux CSH quiescentes. En cas de réparation non optimale, les CSH accumulent des mutations ou sont éliminées par apoptose. Seul un faible pourcentage de CSH ne bénéficie pas d'une réparation adéquate à chaque cycle d'activation. Néanmoins, au cours de la vie d'un organisme, l'accumulation de lésions de l'ADN dans les CSH conduit à l'épuisement progressif du système hématopoïétique (Figure 33).

La sensibilité des progéniteurs hématopoïétiques au **stress réplicatif** induit par diminution du pool de nucléotides varie d'une lignée à l'autre. De manière intéressante, la **lignée la plus affectée est la lignée érythroïde**, suivie des lignées lymphoïdes T et B puis de la lignée granuleuse (Austin et al., 2012).

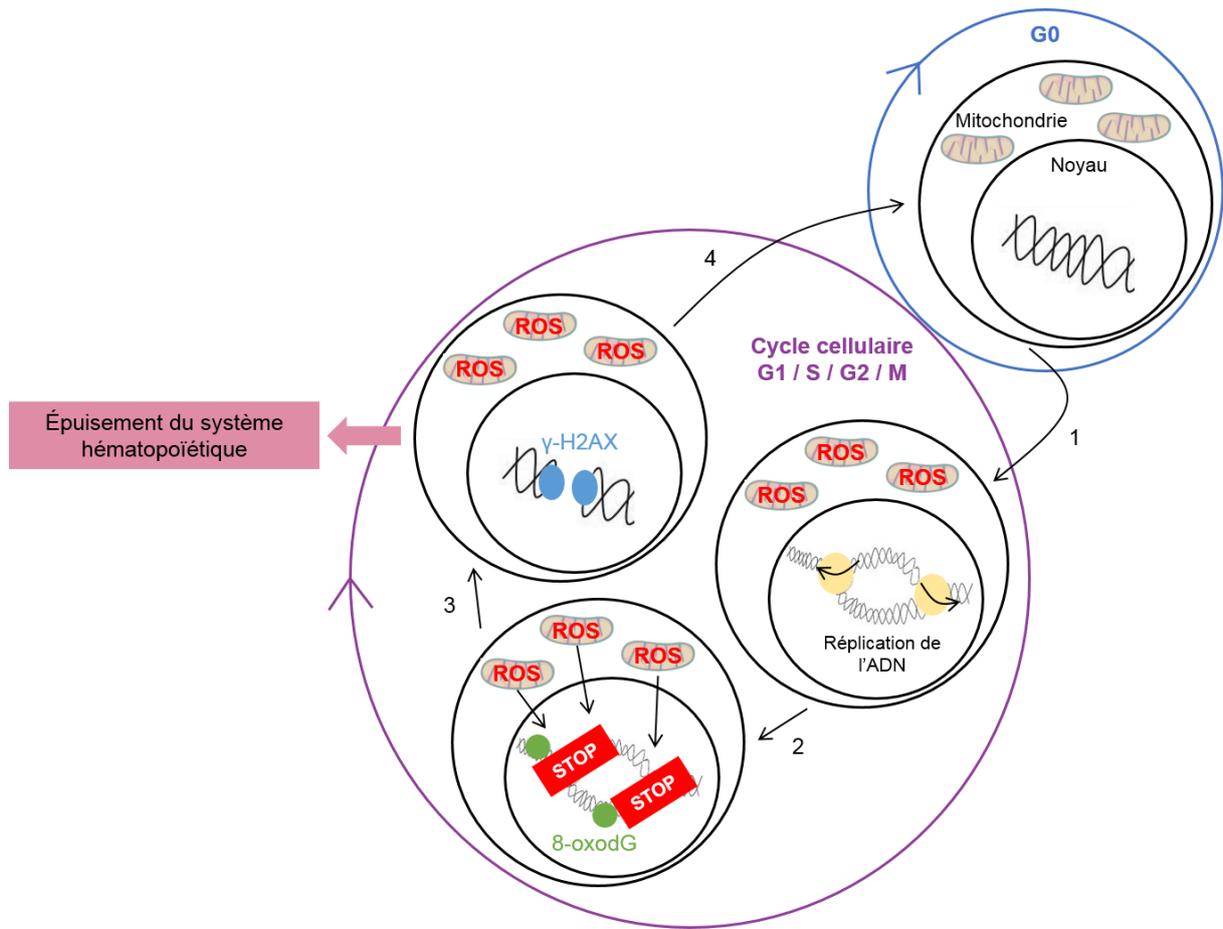


Figure 33 : Les stress oxydatif et réplcatif participent tous les deux à l'épuisement du système hématopoïétique de la souris (d'après Walter et al., 2015)

(1) Les CSH quiescentes s'engagent dans le cycle cellulaire en réponse à un stress (saignement, infections etc.). Au cours de leur prolifération, les CSH consomment de l'énergie produite par phosphorylation oxydative, un processus cellulaire qui génère des ROS. (2) Les ROS forment du 8-oxo-2'-désoxyguanosine (**8-oxodG**), un adduit à l'ADN qui bloque l'avancée des fourches de réplication. (3) L'effondrement des fourches de réplication induit des CDB de l'ADN. L'histone γ -H2AX colocalise avec les sites de CDB. (4) Les dommages à l'ADN sont réparés. Si la réparation n'est pas efficace, les CSH accumulent des mutations ou sont éliminées par apoptose, ce qui entraîne petit à petit un épuisement du système hématopoïétique.

B. Rôle de la protéine SF3B1 dans le contrôle de la réplication et de la réparation de l'ADN

Chez les mammifères, les kinases ATR et ATM du *checkpoint* de la phase S phosphorylent plus de 700 protéines en réponse aux dommages à l'ADN (Matsuoka et al., 2007). Outre les protéines classiquement connues pour participer au maintien de la stabilité génomique, des protéines impliquées dans la maturation des pré-ARNm possèdent des sites de phosphorylation reconnus par ATR et ATM. Depuis le début des années 2000, un large panel d'études a confirmé le rôle de certains facteurs d'épissage – dont SF3B1 – dans le contrôle de la réplication et de la réparation de l'ADN.

En 2009, un screen par siARN (si pour *small interfering*, petit ARN interférent) du génome (\approx 21 000 gènes ciblés dans l'étude, 4 siARN par gène) a permis d'identifier les principaux gènes participant à la réponse aux dommages à l'ADN (Paulsen et al., 2009). D'un point de vue technique, chaque siARN a été transfecté dans la lignée cellulaire humaine HeLa. Puis, les cellules ont été marquées avec un anticorps dirigé contre γ -H2AX. Enfin, l'intensité de fluorescence a été quantifiée par cytométrie. De façon étonnante, la **diminution de l'expression des gènes de maturation des pré-ARNm** – dont SF3B1 – est associée à une **augmentation significative de l'intensité de fluorescence de γ -H2AX**. Ainsi, en plus de sa fonction canonique d'épissage, SF3B1 semble avoir un rôle dans le maintien de la stabilité génomique. Quelques années plus tard, une autre équipe a réalisé un screen par siARN dans la lignée d'ostéosarcome U2OS (Kavanaugh et al., 2015). Brièvement, 72h après la transfection des siARN, les cellules U2OS ont été incubées pendant 30 min avec un analogue halogéné de la thymidine (la bromodésoxyuridine, BrdU). Puis, les cellules ont été lavées et traitées pendant 24h avec l'HU. Enfin, les cellules ont été relavées et incubées 4h avec la 5-éthynyl-2'-désoxyuridine (EdU) avant d'être fixées. Les analyses d'IF ont été faites avant et après traitement des cellules avec l'HU, donc en conditions basales ou en conditions de stress réplcatif. La mesure de l'intensité de fluorescence de l'anticorps dirigé contre le BrdU permet d'illustrer la synthèse de l'ADN en conditions basales. Kavanaugh *et al.* montrent ainsi que les protéines impliquées dans les processus de transcription et de maturation des pré-ARNm – en plus de celles classiquement impliquées dans la réplication de l'ADN – sont nécessaires à la synthèse optimale de l'ADN, puisque la diminution de leur expression conduit à une baisse de l'incorporation de la BrdU. La mesure de l'intensité de fluorescence de l'anticorps dirigé contre l'EdU permet d'illustrer la reprise de la synthèse de l'ADN après exposition à l'HU. Là encore, les protéines impliquées dans les processus de transcription et de maturation des pré-

ARNm semblent indispensables à la reprise rapide de la synthèse de l'ADN en réponse au stress répliatif, puisque la diminution de leur expression conduit à une baisse de l'incorporation de l'EdU associée à une augmentation de l'intensité de fluorescence de γ -H2AX. Par conséquent, les protéines de la transcription et de la maturation des pré-ARNm semblent participer au maintien de la stabilité génomique.

En 2014, une équipe britannique a montré que la protéine de réparation BRCA1 phosphorylée co-immunoprécipite avec des protéines de la machinerie d'épissage uniquement lorsque la cellule est soumise à des dommages à l'ADN de type blocage des fourches de réplication ou CDB (Savage et al., 2014). Les expériences de microscopie confocale et d'immunoprécipitation de la chromatine (ChIP) confirment qu'en cas de dommages à l'ADN, BRCA1 peut être mobilisée et phosphorylée au niveau de la chromatine à différents sites, dont les sites de CDB (cf. Chapitre II-Partie 2-I-A-2-b) et les promoteurs de certains gènes de la réparation de l'ADN. **Sous le contrôle d'ATR et/ou ATM, BRCA1 forme en effet un complexe avec BCLAF1** – une protéine du spliceosome à domaine SR –, l'hétérodimère **U2AF** et **SF3B1**. Ce complexe se lie préférentiellement aux promoteurs des gènes impliqués dans la réparation de l'ADN, comme ATRIP, et favorise ainsi leur épissage co-transcriptionnel et leur expression (Figure 34).

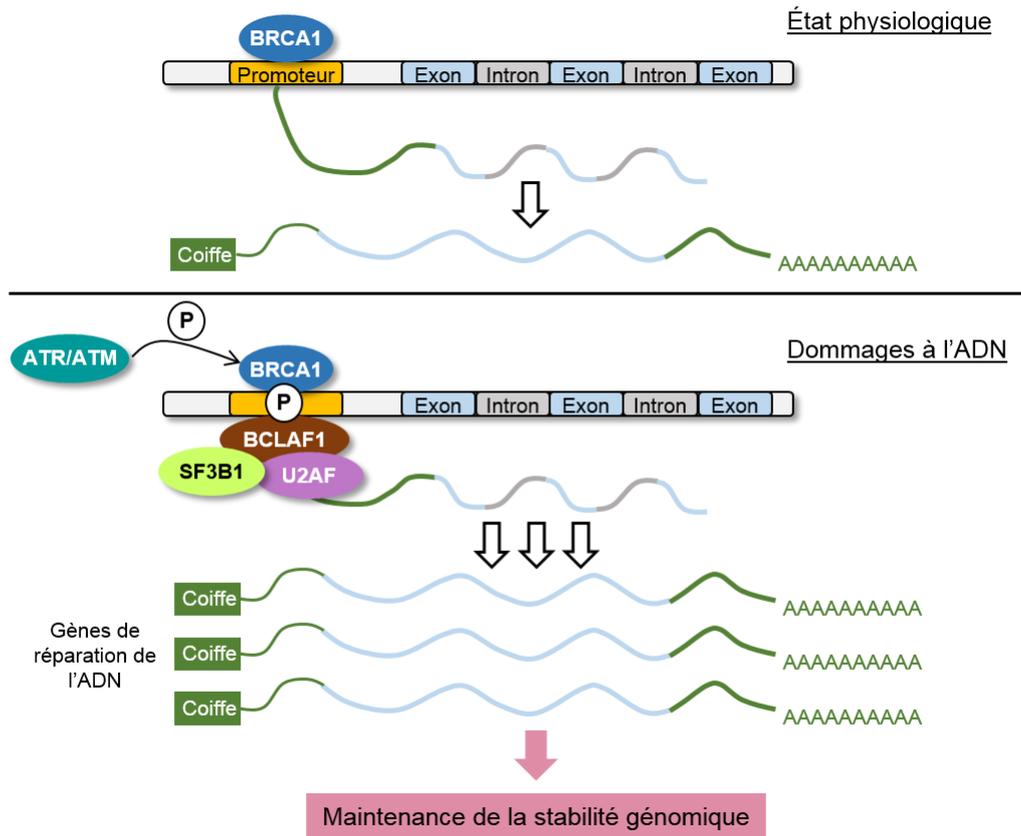


Figure 34 : Rôle du complexe BRCA1/protéines d'épissage dans la maintenance de la stabilité génomique (d'après Savage et al., 2014)

En plus des sites de CDB de l'ADN, BRCA1 se lie aux régions promotrices de certains gènes de réparation de l'ADN. Lorsque la cellule est soumise à des dommages à l'ADN, les kinases ATR et/ou ATM phosphorylent BRCA1 et les protéines d'épissage SF3B1, BCLAF1 et U2AF sont recrutées au niveau de ces régions promotrices. L'épissage co-transcriptionnel des gènes de réparation de l'ADN est alors favorisé, ce qui participe à la maintenance de la stabilité génomique. **P** = acide aminé phosphorylé.

Enfin, plus récemment, l'équipe d'Angelos Constantinou à Montpellier a analysé le protéome associé aux protéines de la voie AF sur la chromatine (Moriel-Carretero et al., 2017). Elle a ainsi mis en évidence la colocalisation de FANCI et FANCD2 avec SF3B1. Comme dit précédemment, le facteur d'épissage SF3B1 est stocké dans les *nuclear speckles*, puis s'associe sur la chromatine aux nucléosomes exoniques, et est enfin engagé dans l'épissage co-transcriptionnel des transcrits naissants. D'après les travaux de Moriel-Carretero et al., SF3B1 est libéré des *nuclear speckles* et recruté au niveau des fibrilles de périchromatine sous le contrôle de FANCI. La protéine FANCD2 facilite quant à elle le désassemblage du spliceosome puis la mobilisation des facteurs d'épissage de la chromatine

vers les *nuclear speckles*. Par conséquent, **FANCI et FANCD2 régulent la dynamique de distribution de SF3B1 dans le noyau**, y compris lorsque la cellule est soumise à un stress répliatif. L'activation de la kinase ATR entraîne effectivement l'accumulation de FANCI et FANCD2 sur la chromatine endommagée et le relargage de SF3B1 des *nuclear speckles*, un processus auquel FANCI contribue (Figure 35). Pour rappel, FANCI et FANCD2 sont recrutées sur les fourches de réplication bloquées, et contribuent à assurer la maintenance de la stabilité génomique pendant la phase S. FANCD2 phosphorylé s'associe notamment avec les hélicases MCM, ce qui permet *in fine* de réduire la vitesse de progression des fourches soumises à un stress (Lossaint et al., 2013). La protéine SF3B1 est-elle aussi impliquée dans la surveillance du réplisome ?

L'ensemble de ces études souligne que le maintien de l'intégrité du génome dépend de la coordination entre trois activités cellulaires fondamentales : la réplication, la transcription et l'épissage. SF3B1 se situe à l'interface de toutes ces activités.

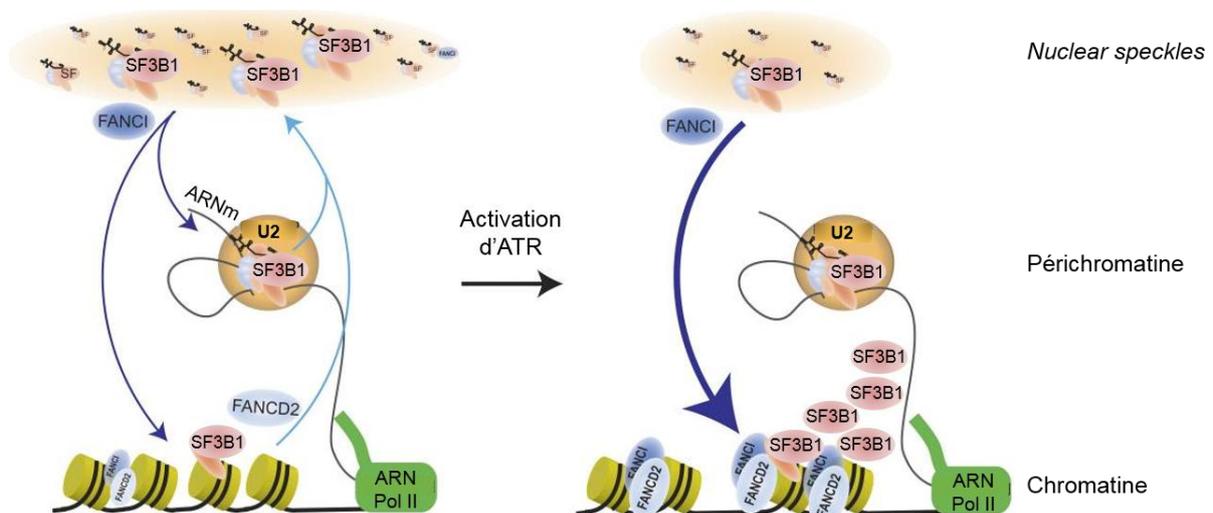


Figure 35 : FANCI et FANCD2 contrôlent la dynamique de distribution de SF3B1 dans le noyau (d'après Moriel-Carretero et al., 2017)

Le facteur d'épissage SF3B1 est stocké dans les *nuclear speckles*. Il est recruté au niveau de la chromatine et des fibrilles de périchromatine sous le contrôle de FANCI. Puis, lorsque le pré-ARNm est épissé, le spliceosome se dissocie et SF3B1 est de nouveau stocké dans les *nuclear speckles* sous le contrôle de FANCD2 cette fois-ci. En cas de stress répliatif et d'activation d'ATR, FANCI et FANCD2 sont mobilisées sur l'ADN endommagé. SF3B1 est alors massivement libéré des *nuclear speckles*.

II. Mutations de *SF3B1* et dommages à l'ADN

La majorité des patients porteurs d'un SMD a plus de 65 ans. Il est donc probable que cette pathologie soit liée au vieillissement du système hématopoïétique et, donc, au déclin de la détection et/ou de la réparation de l'ADN endommagé (Flach et al., 2014; Walter et al., 2015). Pour rappel, 10 à 20% des SMD sont secondaires à l'exposition à des chimiothérapies, des radiations ionisantes et/ou des toxiques environnementaux. De plus, les enfants qui ont une maladie génétique associée à un défaut de réparation de l'ADN peuvent développer un SMD. Ainsi, depuis quelques années, le concept selon lequel les SMD pourraient être causés par un stress génotoxique et un défaut de réparation des dommages à l'ADN a émergé. Avant tout, il est important de préciser que le nombre de mutations somatiques dans les cellules de SMD – comme l'ensemble des hémopathies malignes – reste très faible (≈ 10) par rapport aux tumeurs solides de l'adulte (≈ 150 dans les cancers du poumon et les mélanomes cutanés), ce qui suggère que les voies majeures de réparation de l'ADN sont au moins en partie préservées (Vogelstein et al., 2013).

Dans cette partie, plusieurs études soulignant une association entre les facteurs d'épissage mutés et la présence de dommages à l'ADN dans les cellules pré-cancéreuses ou cancéreuses seront développées.

En 2015, Te Raa *et al.* ont déterminé – grâce à des expériences de cytométrie en flux – l'intensité moyenne de fluorescence (MFI) de γ -H2AX dans des cellules primaires de LLC CD19⁺/CD5⁺ mutées ou non sur le gène *SF3B1* (6 échantillons *SF3B1*^{WT}, 6 échantillons *SF3B1*^{MUT}) (Te Raa et al., 2015). Ils montrent une **augmentation significative de la MFI de γ -H2AX dans les cellules de LLC *SF3B1*^{MUT} par rapport aux cellules *SF3B1*^{WT}**, en conditions basales et 2h post-irradiation 5 Gy. L'IF a permis d'appuyer ces résultats, puisque les foyers intranucléaires de γ -H2AX sont plus nombreux dans les cellules *SF3B1*^{MUT} que dans les cellules *SF3B1*^{WT}. Pour rappel, la phosphorylation de l'histone H2AX en position γ est une modification de la chromatine classiquement associée aux CDB de l'ADN (cf. Chapitre II-Partie 2-I-A-2-b). Te Raa *et al.* proposent donc que les mutations du gène *SF3B1* contribuent soit à une augmentation des dommages à l'ADN soit à un défaut de réponse aux dommages à l'ADN dans les cellules de LLC. Ils n'ont cependant pas réalisé d'étude fonctionnelle permettant de valider leur hypothèse et ainsi confirmer le lien de causalité entre la protéine SF3B1 et le maintien de la stabilité génomique.

En ce qui concerne les SMD, dès 2005, Li et Manley indiquent que la déplétion de la protéine ASF/SF2 dans la lignée lymphoblastique de poulet DT40 aboutit à l'accumulation de *R-loops* et à la formation de CDB de l'ADN (Li and Manley, 2005). *In fine*, les cellules meurent par apoptose. ASF/SF2 est une protéine à domaines SR qui participe à l'interconnexion entre les processus de transcription de l'ADN et de maturation des pré-ARNm. En effet, elle peut d'une part interagir avec le domaine C-terminal de l'ARN polymérase II et ainsi être recrutée sur les sites de transcription des gènes activés (Misteli and Spector, 1999). D'autre part, elle assure la mobilisation de la snRNP U1 au niveau du site donneur d'épissage (5'ss) des pré-ARNm (Kohtz et al., 1994). Les précédentes études de James L. Manley avaient mis en évidence l'utilisation préférentielle de 5'ss alternatifs dans les cellules DT40-ASF invalidées pour le gène *ASF/SF2* (DT40-ASF) (Wang et al., 1996, 1998b). Les travaux publiés dans Cell en 2005 indiquent que **le recrutement d'ASF/SF2 par l'ARN polymérase II sur les néo-transcrits empêche la formation des hybrides ARN/ADN mutagènes**. La surexpression de la RNase H dans les cellules DT40-ASF permet de supprimer les *R-loops* et, donc, de limiter les dommages à l'ADN et l'apoptose.

Récemment, deux autres équipes américaines ont confirmé la contribution de deux facteurs d'épissage mutés à l'accumulation des *R-loops* et, donc, au stress réplcatif par arrêt des fourches de réplication. Nguyen *et al.* se sont tout d'abord intéressés à la protéine U2AF1 (ou U2AF35) (Nguyen et al., 2017, 2018). Ils ont transduit les lignées cellulaires HeLa et K562 avec des vecteurs viraux *U2AF1*^{WT} ou *U2AF1*^{MUT} (*U2AF1*^{S34F}). L'expression de la protéine U2AF1 mutée S34F génère des *R-loops* visualisés par IF (Nguyen et al., 2017). L'anticorps utilisé dans cette étude est le S9.6, un anticorps monoclonal spécifique des hybrides ARN/ADN et qui ne reconnaît pas l'ADN simple-brin. De plus, des foyers intranucléaires de phospho-RPA32 (p-RPA32, sérine 33) colocalisent avec les *R-loops* dans les cellules HeLa et K562 *U2AF1*^{MUT} (Nguyen et al., 2018). Nguyen *et al.* ont ensuite établi des lignées cellulaires HeLa exprimant une RNase H1 sauvage ou mutée et couplée au fluorochrome GFP. Le vecteur RNase H1-GFP est inductible à la doxycycline. Les *R-loops* persistent dans la lignée HeLa *U2AF1*^{MUT} qui exprime une RNase H1 mutée non fonctionnelle. Au contraire, l'induction de l'expression de la RNase H1 sauvage réduit l'intensité de fluorescence des anticorps S9.6 et anti-p-RPA32 dans les cellules *U2AF1*^{MUT}. Le nombre de foyers intranucléaires p-RPA32 est également diminué après traitement des cellules *U2AF1*^{MUT} avec un inhibiteur d'ATR (VE-821), mais pas avec un inhibiteur d'ATM (KU55933). Ainsi, **les mutations d'*U2AF1* sont associées à la formation de *R-loops* et à**

l'activation de la kinase ATR du *checkpoint* de la phase S. L'activation de la voie ATR/CHK1 assure la stabilisation et le redémarrage des fourches de réplication bloquées. Sous inhibiteur d'ATR, les cellules HeLa *U2AF1*^{MUT} accumulent donc des CDB de l'ADN mises en évidence par IF (anticorps anti- γ -H2AX), puis meurent par apoptose. Il en est de même pour les cellules primaires CD34⁺ *U2AF1*^{MUT} issues de patients SMD. De façon intéressante, l'induction de l'expression de la RNase H1 sauvage dans les cellules *U2AF1*^{MUT} supprime les *R-loops* mais ne réduit pas les niveaux de transcrits alternatifs. **La protéine U2AF1 mutée semble induire un stress réplcatif indépendamment de l'épissage alternatif.** Cependant, si l'accumulation de *R-loops* ne contribue pas à l'altération de l'épissage, il n'est pas exclu que l'épissage aberrant participe à la formation des *R-loops* ou empêche leur résolution et contribue à la déstabilisation des processus de réplication et de réparation de l'ADN.

L'accumulation des *R-loops* au niveau des promoteurs des gènes et l'activation de la voie ATR/CHK1 en aval des mutations d'*U2AF1* a été confirmée par Chen et al. (Chen et al., 2018). Ils observent les mêmes perturbations en aval des mutations de *SRSF2*. Chen *et al.* ont transduit les cellules HEK 293T avec des vecteurs pTRIPZ inductibles à la doxycycline. Ces vecteurs contiennent un domaine shARN dirigé contre le transcrit *U2AF1* (ou le transcrit *SRSF2*) endogène et un domaine ADN complémentaire (ADNc) permettant l'expression d'un gène *U2AF1* (ou *SRSF2*) exogène sauvage ou muté. Comme attendu, les analyses RNA-Seq indiquent une augmentation du nombre de transcrits alternatifs avec sauts d'exon dans les clones mutés. Deux techniques ont été utilisées pour mettre en évidence l'accumulation de *R-loops* dans les cellules HEK 293T *U2AF1*^{MUT} ou *SRSF2*^{MUT} : l'IF S9.6 et la R-ChIP. La R-ChIP permet de détecter les *R-loops* à l'échelle du génome. Il s'agit d'une adaptation du protocole standard de ChIP : 1) les cellules HEK 293T sont transfectées avec un plasmide humain RNase H1 (ADNc) étiqueté V5 en C-terminal ; 2) la chromatine est immunoprécipitée grâce à un anticorps dirigé contre le peptide V5 ; 3) l'ADN des hybrides ARN/ADN est converti en ADN double-brin ; 4) l'ADN est amplifié par PCR puis séquencé (séquençage à haut débit) (Chen et al., 2017). Le plasmide code pour une forme mutée de la RNase H1 (D210N). Cette RNase H1 est capable de se lier aux hybrides ARN/ADN nucléaires mais – son activité catalytique étant abolie – elle ne peut pas hydrolyser les fragments ARN de ces hybrides. Selon Chen *et al.*, la technique de R-ChIP est plus résolutive que l'IF S9.6. Cependant, on assume que la RNase H1 est présente à tous les sites de *R-loops*, ce qui n'est pas vrai puisque la RNase H2 est également impliquée dans l'hydrolyse des ARN

anormalement hybridés à l'ADN (cf. Chapitre II-Partie 2-I-A-1). Les expériences d'immunoempreinte révèlent une phosphorylation de la kinase CHK1 (p-CHK1, sérine 345) exclusivement dans les cellules *U2AF1*^{MUT} ou *SRSF2*^{MUT}. Les kinases ATM et CHK2 ne sont par contre pas phosphorylées, aussi bien dans les cellules WT que dans les cellules mutées. Le nombre de foyers intranucléaires γ -H2AX visualisés par IF est significativement augmenté dans les cellules HEK 293T *U2AF1*^{MUT} ou *SRSF2*^{MUT} par rapport aux cellules WT. La littérature n'est pas très claire à ce sujet, mais il semblerait que la kinase ATR soit également capable de phosphoryler l'histone H2AX en réponse au stress réplcatif (Ward and Chen, 2001). Ainsi, **les mutations des gènes *U2AF1* et *SRSF2* causent un stress réplcatif mais a priori pas de CDB de l'ADN**. Elles induisent en outre un arrêt du cycle cellulaire en phase G2/M et une diminution de la prolifération cellulaire due à un ralentissement de la progression des fourches de réplication. L'expression constitutive de la RNase H1 dans les cellules HEK 293T *U2AF1*^{MUT} ou *SRSF2*^{MUT} supprime les *R-loops*, désactive la voie de signalisation ATR/CHK1 et réduit le nombre de foyers intranucléaires γ -H2AX. Enfin, Chen *et al.* ont exploré les effets des mutations de *SRSF2* dans les progéniteurs hématopoïétiques c-KIT⁺ des souris *Mx1-Cre* KI hétérozygotes *Srsf2*^{P95H}. Les analyses d'IF montrent des foyers intranucléaires S9.6 et γ -H2AX uniquement dans les progéniteurs mutés. Les tests de clonogénicité confirment quant à eux le défaut de prolifération des progéniteurs mutés par rapport aux progéniteurs WT, défaut de prolifération partiellement corrigé après surexpression d'une RNase H1 exogène. Pour conclure, Chen *et al.* suggèrent que **l'accumulation de *R-loops* en aval des mutations *U2AF1* et *SRSF2* contribue au dysfonctionnement du système hématopoïétique des patients porteurs d'un SMD**. Le modèle proposé s'appuie sur le rôle connu de certains facteurs d'épissage – dont *SRSF2* – dans la transcription de l'ADN (cf. Chapitre II-Partie 1-I-A) (Ji et al., 2013). Pour rappel, l'ARN polymérase II est très souvent stoppée de façon temporaire à proximité des transcrits naissants encore associés aux promoteurs des gènes. La reprise de l'élongation de l'ARN n'est possible qu'après phosphorylation du CTD de l'ARN polymérase II par la kinase P-TEFb. Contrairement à la protéine *SRSF2* sauvage, la protéine *SRSF2* mutée est incapable d'extraire P-TEFb de son complexe inhibiteur 7SK. L'ARN polymérase II n'est donc pas phosphorylée, et l'élongation de la transcription est définitivement interrompue. Les *R-loops* s'accumulent au niveau des promoteurs des gènes. *In fine*, le stress réplcatif diminue la capacité de prolifération des progéniteurs hématopoïétiques des patients atteints d'un SMD (Figure 36). La protéine *U2AF1* semble elle aussi impliquée dans le contrôle de la transcription de l'ADN, mais le mécanisme n'est pas encore connu.

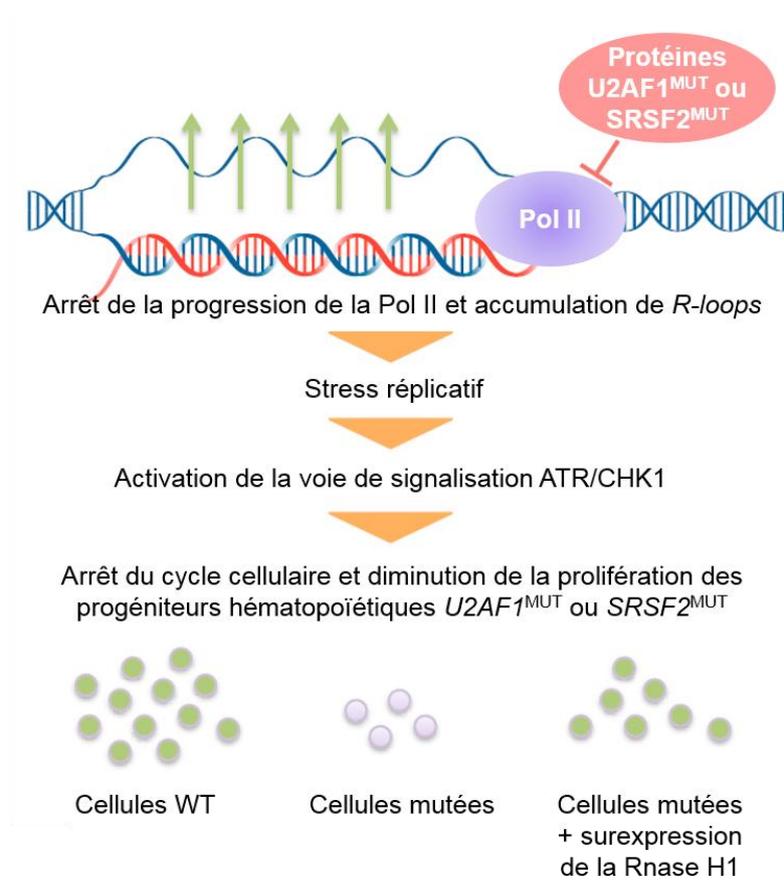


Figure 36 : Le stress réplcatif causé par les mutations des gènes *U2AF1* et *SRSF2* est responsable du défaut de prolifération des progéniteurs SMD (d'après Chen et al., 2018)

D'après Chen *et al.*, l'accumulation de *R-loops* dans les progéniteurs hématopoïétiques des patients porteurs d'un SMD avec mutation *U2AF1* ou *SRSF2* est une conséquence directe du blocage de l'ARN polymérase II (Pol II) au niveau des séquences promotrices des gènes.

En conclusion, plusieurs points sont à relever dans les études précédemment citées : 1) les modifications épigénétiques de l'histone H2AX ne sont pas exclusivement des marqueurs des CDB de l'ADN (cf. Chapitre II-Partie 2-I-A-3), et la mise en évidence d'un défaut de réponse aux dommages à l'ADN ne peut donc pas se limiter à l'analyse par IF ou cytométrie en flux de γ -H2AX ; 2) les techniques utilisées pour détecter les *R-loops* dans les cellules *U2AF1*^{MUT} ou *SRSF2*^{MUT} sont aujourd'hui remises en question puisque l'anticorps S9.6 reconnaît – en plus des hybrides ARN/ADN – les ARN double-brin (Hartono et al., 2018) et la R-ChIP repose sur l'hypothèse selon laquelle la RNase H1 se lie aux *R-loops in vitro* mais aussi *in vivo*, ce qui n'a pas été clairement démontré ; 3) le choix des lignées cellulaires (HeLa et HEK 293T non hématopoïétiques) et du mode d'expression des mutations *U2AF1* et *SRSF2* (vecteurs viraux surexprimant les formes mutées) n'est pas le plus pertinent. À ce jour, le stress répliatif – et notamment l'accumulation de *R-loops* – n'a pas été étudié dans les cellules *SF3B1*^{MUT}. En outre, la protéine SF3B1 n'est pas connue pour posséder une fonction dans la transcription de l'ADN.

Chapitre III - Résultats

Partie 1

I. Problématique

Les SMD-SC sont des pathologies myéloïdes qui se caractérisent par une anémie, une dysérythropoïèse médullaire et une accumulation de fer dans les mitochondries des érythroblastes polychromatophiles. 90% des patients atteints d'un SMD-SC sont porteurs d'une mutation sur le gène d'épissage *SF3B1*. Contrairement aux autres sous-types de SMD, l'érythropoïèse inefficace des SMD-SC s'accompagne d'une surcharge systémique en fer qui se manifeste par une augmentation des taux plasmatiques de ferritine et de NTBI, et ce avant même la correction de l'anémie par transfusion sanguine. **Toutefois, les mécanismes d'augmentation du fer disponible ne sont à ce jour pas connus.**

II. Objectif

L'hépcidine est une hormone chargée de réguler la biodisponibilité du fer dans l'organisme et dont l'expression dépend, entre autres, d'une autre hormone : l'érythroferrone (ERFE). Les travaux récents de RNA-Seq ont permis de souligner la dérégulation importante de l'épissage dans les cellules SMD *SF3B1*^{MUT}. En effet, la protéine *SF3B1*^{MUT} reconnaît un BS cryptique situé en amont du BS canonique, ce qui entraîne – *in fine* – l'utilisation préférentielle d'un AG' cryptique situé lui aussi en amont de l'AG canonique. Du fait de l'introduction d'un codon stop prématuré, la plupart des transcrits aberrants ainsi générés sont ensuite dégradés dans le cytosol par la voie NMD, d'où la diminution de l'expression du gène. C'est notamment le cas du gène *ABCB7*, qui code pour une protéine d'export des clusters Fe/S du noyau vers le cytosol. La diminution de son expression participerait à l'accumulation de fer dans les mitochondries des érythroblastes. Les transcrits aberrants pour lesquels l'ajout de nucléotides n'induit pas la formation d'un codon stop prématuré pourraient quant à eux échapper à la voie NMD et être traduits en variants protéiques. Selon le domaine protéique dans lequel s'ajoutent les nouveaux acides aminés, la fonction et/ou la stabilité des protéines aberrantes sont potentiellement modifiées. Notre travail a permis d'identifier un variant protéique d'ERFE dans les érythroblastes *SF3B1*^{MUT}. Notre objectif a été de montrer **la contribution de ce variant protéique à la physiopathologie des SMD-SC et, plus particulièrement, à la surcharge systémique en fer.** Ces résultats font l'objet d'un article en cours de révisions à *Science Translational Medicine* (STM). Les données complémentaires de cet article figurent en annexes 1-4. Les tableaux S2 et S3 (données de RNA-Seq) sont disponibles dans la base de données GEO (numéro d'accès : GSE113433).

III. Résultats

Identification d'un variant de l'érythroferrone dans les syndromes myélodysplasiques avec mutation sur le gène *SF3B1*

A variant erythroferrone induced by clonal erythropoiesis disrupts iron homeostasis in *SF3B1*-mutated myelodysplastic syndrome

Sabrina Bondu,^{1,2,3,4,*} Anne-Sophie Alary,^{1,2,3,4,5,*} Carine Lefèvre,^{1,2,3,4,6} Alexandre Houy,⁷ Grace Jung,⁸ Thibaud Lefebvre,^{6,9} Abderrahmane Bousta,^{1,2,3,4} François Guillonnet,^{1,2,3,4,10} Benjamin Billoré,¹¹ Samar Alsafadi,¹² Michel Wassef,¹³ Raphaël Margueron,¹³ Alice Rousseau,^{1,2,3,4} Nathalie Droin,¹⁴ Nicolas Cagnard,¹⁵ Ismael Boussaid,^{1,2,3,4} Susann Winter,¹⁶ Anne-Sophie Kubasch,¹⁶ Didier Bouscary,¹⁷ Valeria Santini,¹⁸ Andrea Toma,¹⁹ Mathilde Hunault,²⁰ Aspasia Stamatoullas,²¹ Emmanuel Gyan,²² Thomas Cluzeau,²³ Uwe Platzbecker,²⁴ Lionel Adès,²⁵ Hervé Puy,^{6,9} Marc-Henri Stern,²⁶ Zoubida Karim,^{6,9} Patrick Mayeux,^{1,2,3,4,6,10} Elizabetha Nemeth,⁸ Sophie Park,²⁷ Tomas Ganz,⁸ Léon Kautz,¹¹ Olivier Kosmider,^{1,2,3,4,5,6,†} Michaëla Fontenay^{1,2,3,4,5,6,†}

¹ Institut Cochin, Department Development, Reproduction and Cancer, Paris, France

² Institut National de la Santé et de la Recherche médicale (INSERM) U1016, Paris, France

³ Centre National de la Recherche Scientifique (CNRS) Unité Mixte de recherche (UMR) 8104, Paris, France

⁴ Université Paris Descartes, Sorbonne-Paris Cité, Paris, France

⁵ Assistance Publique-Hôpitaux de Paris, Hôpitaux Universitaires Paris Centre – Cochin, Service d'hématologie biologique, Paris, France

⁶ Laboratoire d'excellence du Globule Rouge GR-Ex, Université Paris Descartes, Paris, France

⁷ Department of Human Genetics and Oncogenesis, Institut Curie, Paris, France

⁸ Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, CA

⁹ INSERM, UMR 1149/ERL CNRS 8252, Centre de Recherches sur l'inflammation, Université Paris Diderot, Paris, France

¹⁰ Proteomic platform 3P5, Université Paris Descartes, Paris, France

¹¹ Institut de Recherche en Santé Digestive (IRSD), Université de Toulouse, INSERM U1220, Institut National de la Recherche Agronomique U1416, Ecole Nationale Vétérinaire de Toulouse, Université Paul Sabatier, Toulouse, France

¹² Institut Curie, PSL Research University, Department of Translational Research, Paris, France

¹³ Institut Curie, PSL Research University, INSERM 934/UMR 3215, Genetics and biology of Development, Paris, France

¹⁴ Institut Gustave Roussy, Genomic platform, Villejuif, France

¹⁵ Platform Bioinformatics, Paris Descartes University, Paris, France

¹⁶ Medizinische Klinik und Poliklinik I, Dresden, Germany

¹⁷ Assistance Publique-Hôpitaux de Paris, Hematology Department, Hôpital Cochin, and Paris Descartes University, Paris, France

¹⁸ MDS unit, Hematology, AOU Careggi, University of Florence, Italy

¹⁹ Assistance Publique-Hôpitaux de Paris, Hematology Department, Hôpital Henri-Mondor, and Paris 12 University, Créteil, France

²⁰ Service des Maladies du Sang, Centre hospitalo-universitaires, Angers, France

²¹ Département d'hématologie, Centre Henri-Becquerel, Rouen, France

²² Service d'hématologie et thérapie cellulaire, Centre hospitalo-universitaire, CNRS ERL 7001 LNOx, Université de Tours, Tours, France

²³ Côte d'Azur University, CHU of Nice, Hematology department and INSERM U1065, Mediterranean Center of Molecular Medicine, Nice, France

²⁴ Medizinische Klinik und Poliklinik 1, Hämatologie und Zelltherapie, Universitätsklinikum, Leipzig, Germany

²⁵ Assistance Publique-Hôpitaux de Paris, Service d'Hématologie Senior, Hôpital Saint-Louis, Paris, France

²⁶ Institut Curie, PSL Research University, Inserm U830, DNA repair and uveal melanoma (D.R.U.M.) and Département de génétique, Équipe labellisée par la Ligue nationale contre le cancer, Paris, France.

²⁷ Department of Hematology, Centre Hospitalier Universitaire, Université de Grenoble Alpes, Grenoble, France

* These authors equally contributed to the work.

† To whom correspondence should be addressed: michaela.fontenay@inserm.fr; olivier.kosmider@aphp.fr

One Sentence Summary: The expression of a variant erythroferrone specifically induced in mutant erythroblasts contributes to hepcidin modulation and systemic iron accumulation in patients with *SF3B1*-mutated myelodysplastic syndrome.

Abstract

Myelodysplastic syndromes (MDS) with ring sideroblasts are hematopoietic stem cell disorders with erythroid dysplasia and mutations in the *SF3B1* splicing factor gene. MDS patients with *SF3B1* mutations present with an altered iron homeostasis, even in the absence of transfusions, but the mechanisms that are responsible for their parenchymal iron overload are unknown. Under physiological conditions, high iron levels are regulated by hepcidin, a hormone that prevents iron efflux, and hepcidin transcription is suppressed by erythroferrone (ERFE), an erythropoietin-responsive gene. Here, we identified an alternative ERFE transcript in MDS patients with *SF3B1* mutation. Induction of this ERFE transcript in primary *SF3B1*-mutated bone marrow erythroblasts

resulted in the generation of a variant protein that maintained the capacity to suppress hepcidin transcription. Plasma concentrations of ERFE were more elevated in MDS patients with a *SF3B1* gene mutation compared to patients with *SF3B1* wild-type MDS. Thus, hepcidin reduction by a variant erythroferrone is likely responsible for the increased ferritin and plasma iron levels in patients with *SF3B1*-mutated MDS, revealing a way to prevent iron-mediated organ toxicity. The expression of the variant ERFE transcript was restricted to *SF3B1*-mutated erythroblasts and decreased in lenalidomide-responsive anemic patients, identifying a specific biomarker of clonal erythropoiesis.

Introduction

Myelodysplastic syndromes with ring sideroblasts (MDS-RS) are clonal hematopoietic stem cell (HSC) disorders with a prominent erythroid dysplasia of the bone marrow (BM) responsible for a macrocytic anemia. Mitochondrial iron accumulation and apoptosis of mature erythroblasts cause ineffective erythropoiesis (Gyan et al., 2008; Tehranchi et al., 2005). In contrast to other MDS subtypes, patients with MDS-RS exhibit signs of systemic iron accumulation that is reflected by increased ferritin and non-transferrin bound iron levels even before they become transfusion-dependent and develop parenchymal iron overload thereafter (Swart et al., 2018; Zhu et al., 2016).

Splicing factor gene *SF3B1* is mutated in ~90% of MDS-RS and the diagnosis is considered whenever the gene is mutated, even if the percentage of RS is relatively low between 5 and 15% (Damm et al., 2012; Papaemmanuil et al., 2011; Yoshida et al., 2011). Mutations arise in the HSC (Chesnais et al., 2017; Mian et al., 2015; Mortera-Blanco et al., 2017). Aberrant splicing events are reported in MDS and other *SF3B1*-driven cancers including uveal melanoma and chronic lymphocytic leukemia (CLL) (Alsafadi et al., 2016; Darman et al., 2015; Obeng et al., 2016; Wang et al., 2016). The selection of an alternative branch site (BS) resulting in the use of a cryptic 3' splice site (ss) is the most frequently detected abnormality.

Computational analysis revealed that the majority of cryptic 3'ss are located upstream of canonical 3'ss at nucleotide distances that are not multiple of 3 suggesting that the aberrant transcripts would likely contain a premature termination codon (PTC) and be degraded by the non-sense mediated decay (NMD). It has been predicted that half of the aberrantly spliced transcripts in *SF3B1*-mutated cells are NMD-sensitive and the canonical isoforms are down-regulated. For instance, ABCB7 transcript encoding a mitochondrial transporter involved in the export of Fe-S clusters is aberrantly spliced and undergoes NMD (Allikmets et al., 1999; Darman et al., 2015; Dolatshad et al., 2016; Nikpour et al., 2013). It is also possible that NMD-insensitive aberrant transcripts are translated into proteins with altered function (Darman et al., 2015). How these aberrant splicing events contribute to the disease phenotype and in particular to systemic iron overload is unclear.

In contrast with other MDS subtypes, MDS-RS are associated with lower levels of the iron homeostasis regulator, hepcidin, and as a consequence, the absorption of iron by duodenal enterocytes and its release from erythrophagocytic macrophages may be increased (Ambaglio et al., 2013; Nemeth et al., 2003; Nicolas et al., 2002; Pak et al., 2006b; Santini et al.,

2011; Weinstein et al., 2002). Inappropriately low hepcidin levels in MDS-RS could depend on the degree of ineffective erythropoiesis linked to impaired iron incorporation into heme because of mitochondrial iron trapping, or to increased expression of hepcidin repressor (Malcovati et al., 2011). Growth differentiation factor 15 (GDF-15) and twisted gastrulation (TWSG1), two members of the transforming growth factor- β superfamily have been proposed as pathological suppressors of hepcidin in ineffective erythropoiesis (Tanno et al., 2007, 2009). More recently, erythroferrone (ERFE), a C1q-tumor necrosis factor-related family of proteins (CTRP) member has been described as a major erythroid regulator of hepcidin and involved in the pathological suppression of hepcidin in patients with β -thalassemia (Ganz et al., 2017; Kautz et al., 2014). In the present study, we identify a variant transcript of ERFE specific of *SF3B1*-mutated MDS which expression is restricted to the erythroid lineage and that contributes to increased concentration of ERFE protein leading to hepcidin suppression and systemic iron accumulation in patients. This variant appeared as a pertinent biomarker of clonal erythropoiesis including for monitoring treatment response of anemic patients with *SF3B1*-mutated MDS.

Results

Upregulation of ERFE by the use of an alternative 3' splice site in SF3B1-mutated MDS

To investigate the mechanism of systemic iron overload, a cohort of 156 patients with lower risk MDS including 60 MDS-RS with single lineage dysplasia (SLD), 17 MDS-RS with multilineage dysplasia (MLD), 2 5q-syndrome, 17 MDS-SLD, 42 MDS-MLD, and 18 MDS with type 1

excess of blasts (EB1) was established (**table S1**). The Revised-International Prognosis Scoring System score was very low in 21, low in 98, intermediate in 29, and high in 2 patients. Twenty-six genes commonly mutated in MDS were sequenced in the BM mononuclear cell (MNC) fraction. *SF3B1* gene was mutated in 94 MDS including 63 (67%) cases affected by a *SF3B1*^{K700E} mutation. Among the 62 patients with no *SF3B1* mutation,

other splicing genes, *SRSF2*, *U2AF1* or *ZRSR2* were mutated in 27 cases, and no splicing gene mutation was observed in 35 cases. Some cases presented two splicing mutations (**fig. S1**). We evaluated the consequences of *SF3B1* gene mutations on gene expression and splicing by sequencing the transcriptome of the BM MNCs isolated from 27 patients of the cohort including 21 *SF3B1*^{MUT} MDS, 6 *SF3B1*^{WT} MDS, and 5 healthy controls. Differential analyses of gene expression and splice junctions were conducted using DESeq2 (Love et al., 2014). We detected 6,343 genes as differentially expressed between *SF3B1*^{MUT} MDS and *SF3B1*^{WT} MDS with a P-value <0.05 (**table S2**). Principal component analysis (PCA) of gene expression profiles separated *SF3B1*^{MUT} and *SF3B1*^{WT} MDS (**fig. 1A**). The differentially expressed genes were enriched in 73 specific GO terms with an absolute log₂ fold-change (FC) >1 and a Benjamini-Hochberg (BH)-adjusted P-value <0.05 (**fig. 1B**). Genes with a log₂(FC) <-1 were involved in serine/threonine kinase signaling, apoptosis, myeloid differentiation, inflammation, cell-cell adhesion while those with a log₂(FC) >1 were involved in heme metabolism, erythrocyte differentiation, and iron homeostasis (**fig. 1C**). We plotted the log₂(FC) of 16 differentially expressed genes belonging to the IRON_ION_HOMEOSTASIS gene set GO:0055072 (<http://amigo.geneontology.org>) and showed that the *FAM132B/ERFE* transcript encoding erythroferrone was up-regulated (**fig. 1D**).

We then identified 1,528 differentially expressed 5' and 3' junctions including annotated 5' donor and 3' acceptor ss, ambiguous junctions and canonical junctions with BH-adjusted P-values ≤10⁻⁵ and absolute log₂(FC) ≥1 (**table S3**). These junctions allowed the hierarchical clustering of the 21 *SF3B1*^{MUT} and 6 *SF3B1*^{WT} MDS samples (**fig. 2A**). After excluding differentially expressed

canonical junctions, we then considered the 1,147 alternative junctions among which we identified 786 3' acceptor junctions (68.5%), 176 5' donor junctions (15.3%) and 185 ambiguous junctions (16.1%) attributed to either the alternative 5' or 3'ss or both alternative 5' and 3'ss (**fig. 2B**). The analysis of distances between the alternative and canonical 3'ss showed that the majority of alternative 3'ss (AG') was located between -24 and -9 nucleotides preceding the canonical 3'ss (AG) (**fig. 2C**). The proportion of alternative AG' junctions in these novel splice variants was generally increased in *SF3B1* mutants and less frequently in *SF3B1*^{WT} samples (**fig. 2D**). To detect alternative transcripts that were likely to generate significant amounts of variant proteins with a modified length, we applied a filter selecting transcripts with an additional stretch of nucleotides numbering multiples of 3, a percentage of alternative junction coverage (AG') to alternative and canonical junctions coverage (AG'+AG) over 0.1 and an expression ratio of the alternative junction in *SF3B1*^{MUT} versus *SF3B1*^{WT} samples over 10 (**fig. S2A**). We obtained 66 alternative junctions in 63 genes (**table S4 & fig. S2B**). These genes were involved in the epigenetic regulation and transcription, mRNA processing and translation, intracellular transport, cell cycle and migration, signaling and apoptosis, mitochondrial metabolism and iron homeostasis (**fig. 2E**). Among the 66 alternative junctions, 29 were related to an in-frame insertion of 9 to 27 nucleotides and 26 of them were differentially expressed between *SF3B1*^{MUT} and *SF3B1*^{WT} MDS with an absolute log₂(FC) >0.3 (**fig. 2F**). *FAM132B/ERFE* gene was identified among the upregulated genes with an alternative junction due to the use of a cryptic 3'ss located between the exons 2 and 3 (chr2: 239,070,357 – 239,071,364) and no PTC. The aberrant transcript contained 12 additional nucleotides in the open reading frame and is referred to as ERFE⁺¹² from here on (**fig. 2G**). It was

systematically detected in all *SF3B1*^{MUT} MDS samples and represented a mean of 24.8% of FAM132B/ERFE transcripts in *SF3B1*^{MUT} MDS versus 0.2% in *SF3B1*^{WT} MDS. This indicates that the expression of ERFE⁺¹² was linked to the presence of a mutation in *SF3B1* gene.

SF3B1^{MUT}-restricted expression of alternative ERFE⁺¹² transcript

To ascertain that expression of the ERFE⁺¹² transcript is dependent on the presence of a *SF3B1* mutation, we transfected the erythro-megakaryocytic cell line UT7/EPO with a pLVX plasmid encoding a synthetic full-length *SF3B1*^{WT} or mutant *SF3B1*^{K700E} cDNA. We then designed a sensitive fluorescent RT-PCR allowing the detection of ERFE⁺¹² and ERFE^{WT} transcripts as 162 nucleotides (nt) and 150 nt fragments by capillary electrophoresis (**table S5**). Twenty-four hours post-transfection, ERFE⁺¹² transcript was detected in *SF3B1*^{K700E} transfected cell line, but not in *SF3B1*^{WT} cell line (**fig. 3A**). To validate the splice pattern induced in the context of *SF3B1* mutation in ERFE pre-mRNA, we performed a minigene splicing assay. An *ERFE* sequence of about 200 nucleotides located on both sides of the cryptic 3'ss (AG') was cloned in an ExonTrap vector. The alternative junction in *ENOSF1* gene (chr18: 683,395-685,920) cloned in the same vector was used as a control (Alsafadi et al., 2016). These minigenes were transfected into the murine G1E-ER4 cell line in which the *Sf3b1*^{K700E} mutation was edited by CRISPR-Cas9 technology. As expected due to the lack of intron homology between species, sequencing the transcriptome of G1E-ER4 *Sf3b1*^{K700E} cell line or isogenic *Sf3b1*^{WT} cell line demonstrated that endogenous murine ERFE and *ENOSF1* did not undergo alternative splicing (not shown). After transfection, the alternative 3'ss AG' ERFE⁺¹² was detected on the capillary electrophoresis in G1E-ER4 *Sf3b1*^{K700E}

cells, but not in the isogenic G1E-ER4 *Sf3b1*^{WT} cells (**fig. 3B**). The usage of alternative 3'ss AG' was detectable for *ENOSF1* gene in *Sf3b1*^{WT} cell line and became predominant in *Sf3b1*^{K700E} cell line suggesting that the mutation favored the usage of alternative AG'.

We then investigated the expression of ERFE⁺¹² in primary BM samples of MDS patients using the fluorescent PCR. In 46 lower risk MDS patients, *SF3B1*^{MUT} was present in 25 cases including 20 MDS-SLD-RS/MDS-MLD-RS, 3 MDS-MLD, and 2 MDS-EB1. ERFE⁺¹² was detected in the BM MNC fractions of all *SF3B1*^{MUT} MDS, and none of the *SF3B1*^{WT} MDS. ERFE⁺¹² was not detected in any other cases of MDS with mutations in *SRSF2* (n=10), *U2AF1* (n=1), *ZRSR2* (n=1), or in 21 MDS patients with mutations in other genes (**fig. 3C**). In one patient with an *SF3B1* monoallelic deletion and a substitution G742D on the remaining allele, the expression of alternative ERFE⁺¹² exceeded that of the canonical ERFE^{WT} (**fig. S3A-C**). We then amplified primary erythroblasts from the BM CD34⁺ cells of two *SF3B1*^{MUT} and one *SF3B1*^{WT} MDS and showed that ERFE⁺¹² was expressed in *SF3B1*^{MUT} but not in *SF3B1*^{WT} erythroid cells (**fig. 3D**). This further supports that ERFE⁺¹² is related to *SF3B1* mutation. *SF3B1*^{MUT} MDS are characterized by the enrichment of the BM with erythroid cells. To investigate whether ERFE⁺¹² could be detected in cells deriving from erythroblastic BMs with another genetic background, we collected samples from one patient with an acquired sideroblastic anemia with 38% of BM erythroblasts, more than 15% of ring sideroblasts, a normal karyotype, no *SF3B1* mutation but one *SRSF2* and one *TET2* mutation, three patients with a congenital sideroblastic anemia due to an *ALAS2* mutation/deletion or a *GLRX5* mutation, one patient with a severe β -thalassemia and one patient with *SF3B1*^{K700E} MDS. Using fluorescent PCR, ERFE^{WT} was present while ERFE⁺¹² was

not detectable in any case except the *SF3BI*^{K700E} MDS sample (**fig. 3E**). This confirms that the onset of an aberrant ERFE⁺¹² is not dependent on the amplification of erythroid compartment or the presence of ring sideroblasts, but depends on the presence of a mutant *SF3BI*.

Translation of ERFE⁺¹² into a variant protein that represses hepcidin

Human *ERFE* encodes a 354-aminoacid polypeptide. The addition of 12 nucleotides in *ERFE* mRNA generates an *ERFE* variant (further referred to as *ERFE*^{VPFQ}) containing a valine-proline-phenylalanine-glutamine (VPFQ) insertion immediately upstream of the collagen domain (**fig. 4A & fig. S4A**). To investigate whether the mutant *ERFE*^{VPFQ} protein was produced in vivo, we prepared cell lysates of erythroblasts derived in culture from the BM CD34⁺ cells of patients with *SF3BI*^{MUT} MDS. Through LC MS/MS protein identification, we obtained several peptide-matching propositions including the ALHELGVYYLPDAEGAFR peptide (**fig. S4B**) already reported in public databases (www.proteomicsdb.org/) and we identified a cryptic peptide VPFQFGLPGPPGPPGPPGQGGP GPIIPEALLK corresponding to the VPFQ insertion at position 108 of *ERFE* (**fig. 4B**). This confirms that alternative *ERFE*⁺¹² transcript is translated into *ERFE*^{VPFQ} in *SF3BI*^{MUT} erythroblasts.

ERFE is known to repress hepcidin in mice and contribute to pathological hepcidin suppression in patients with transfused and non-transfused β -thalassemia (Ganz et al., 2017; Kautz et al., 2014). Whether *ERFE*^{VPFQ} also repressed hepcidin was tested. For this purpose, recombinant proteins *ERFE*^{WT} and *ERFE*^{VPFQ} were produced in HEK293F cells. SDS-PAGE analysis of *ERFE*^{WT} and *ERFE*^{VPFQ} in reducing and non-reducing conditions demonstrated a comparable molecular

weight and a multimerization pattern as both proteins were predominantly found in a trimeric form of approximately 130 kDa (**fig. 4C**). Then, Hep3B hepatoma cells were treated for 16 hours with HEK293F cell supernatants overexpressing *ERFE*^{VPFQ} or *ERFE*^{WT}. Both proteins caused a similar 5-fold reduction in the expression of HAMP mRNA encoding hepcidin compared to controls (**fig. 4D**). These data indicate that insertion of 4 amino-acids upstream of the collagen domain did not affect the bioactivity of the protein.

Overproduction of erythroferrone is predictive of hyperferritinemia in SF3BI^{MUT} MDS

We then measured the level of plasma *ERFE* in the learning cohort of 156 MDS patients and 20 healthy non-blood donor controls using a validated immunoassay (Ganz et al., 2017). We first verified that the *ERFE* immunoassay was able to detect both *ERFE*^{WT} and *ERFE*^{VPFQ}. Indeed, human *ERFE* ELISA detected similar amounts (1 μ g/mL) of each recombinant protein in the supernatants of HEK293F cells transiently transfected with *ERFE*^{WT} and *ERFE*^{VPFQ} expression vectors. This establishes that both isoforms were detectable by ELISA. The mean concentration of *ERFE* in *SF3BI*^{MUT} or *SF3BI*^{WT} MDS was higher than in controls ($P < 0.0001$; **fig. 5A**). Among the MDS samples, the *ERFE* concentration was higher in *SF3BI*^{MUT} (135.0 ± 72.5 ng/mL) compared to *SF3BI*^{WT} (62.1 ± 36.7 ng/mL) MDS ($P < 0.0001$; **fig. 5A**). Consistently, the *ERFE* concentration was also higher in MDS-RS compared to all other WHO MDS subtypes (**fig. S5A, S5B**). Ferritin levels were significantly higher in *SF3BI*^{MUT} patients compared to *SF3BI*^{WT} patients ($P < 0.0001$; **fig. 5B**). We also measured plasma hepcidin levels and confirmed that the concentration of hepcidin in *SF3BI*^{MUT} MDS was significantly lower compared to *SF3BI*^{WT} MDS ($P = 0.0309$; **fig. 5C**). We found no

difference between $SF3BI^{MUT}$ MDS and healthy controls. The hepcidin/ferritin ratio was even more decreased in $SF3BI^{MUT}$ versus $SF3BI^{WT}$ MDS ($P < 0.0001$; **fig. 5D**) and also in MDS-RS compared to other WHO subtypes (**fig. S5C**), and this was due to both a lower level of hepcidin and a higher concentration of ferritin in $SF3BI^{MUT}$ patients. The plasma concentration of ERFE was inversely correlated to the hepcidin/ferritin ratio (Pearson test; $P < 0.0001$; $r^2 = 0.360$). Our analysis also highlights that an ERFE concentration above a threshold of 100 ng/mL repressed hepcidin more efficiently (**fig. S6A**). The elevated concentration of plasma ERFE was associated with a more pronounced degree of ineffective erythropoiesis as assessed by a significant increase of plasma concentration of soluble transferrin receptor (sTfR) in $SF3BI^{MUT}$ compared to $SF3BI^{WT}$ MDS patients (**fig. S6B**). Serum EPO level was equally increased in $SF3BI^{MUT}$ and $SF3BI^{WT}$ patients (**table S1**), and, although ERFE is regulated by erythropoietin in mice, we did not find any correlation between serum EPO level and ERFE concentration suggesting that ERFE production was not restricted to the stimulation of erythroid cells by EPO (**fig. S6C**) (Kautz et al., 2014).

To validate these findings, we prospectively enrolled an external cohort of lower risk MDS patients in our study until the number of $SF3BI^{MUT}$ MDS was comparable to that of the learning cohort (**table S6**). This validation cohort included 55 MDS patients with 42 (78%) of MDS with $SF3BI$ mutation, and notably, a similar mean number of RBC units/8 weeks between $SF3BI^{MUT}$ and $SF3BI^{WT}$ patients. The mean concentration of ERFE and ferritin were significantly increased ($P = 0.0005$ and $P = 0.0488$, respectively; **fig. 5E, 5F**) while hepcidin and hepcidin/ferritin ratio were significantly decreased in $SF3BI^{MUT}$ patients ($P = 0.0038$ and $P = 0.0002$, respectively; **fig. 5G, 5H**). This confirms the results of the

learning cohort and suggests that the increase of ERFE concentration in $SF3BI^{MUT}$ MDS patients is independent of RBC transfusions.

Iron homeostasis changes upon RBC transfusions which diminish the degree of anemia and increase the level of circulating iron, both effects resulting in the up-regulation of hepcidin. To assess the influence of RBC transfusion in our analysis, we delineated a subset of 61 MDS in the learning cohort including 25 $SF3BI^{MUT}$ and 36 $SF3BI^{WT}$ patients with a low transfusion burden before inclusion (< 4 RBC units per 8 weeks). In this subset, patients with $SF3BI^{MUT}$ or $SF3BI^{WT}$ MDS were equally transfused (mean 0.5 RBC unit/8 weeks), but ferritin and plasma iron levels remained higher in $SF3BI^{MUT}$ patients (**fig. S7A, S7B**). The hepcidin/ferritin or hepcidin/plasma iron ratios were significantly lower in $SF3BI^{MUT}$ patients (**fig. S7C, S7D**) and the circulating ERFE concentration remained significantly more elevated in low transfusion burden $SF3BI^{MUT}$ patients compared to $SF3BI^{WT}$ patients (**fig. S7E**).

We then explored the determinants of hyperferritinemia >300 $\mu\text{g/mL}$ in low transfusion burden patients. By univariate analysis, ERFE, hepcidin and $SF3BI$ mutation, but not sTfR level and the number of transfused RBC units, were significantly linked to the concentration of serum ferritin. By multivariate analysis, ERFE, hepcidin and $SF3BI$ mutation remained independent predictors of high ferritin level (**table 1**). Altogether, these results indicate that, before patients reached a critical threshold of transfusion dependence, hyperferritinemia is caused by $SF3BI^{MUT}$ -induced expression of ERFE, which in turn lowers hepcidin.

Erythroid lineage-restricted expression of ERFE⁺¹²

In mice, ERFE mRNA expression in the BM is regulated by erythropoietin (EPO)

and is predominant in basophilic and polychromatic erythroblasts (Kautz et al., 2014). To investigate whether ERFE and ERFE⁺¹² expression is restricted to the erythroid lineage, we amplified in parallel the erythroid or granulocytic precursors deriving from the BM CD34⁺ cells of 1 *SF3B1*^{MUT} and 1 *SF3B1*^{WT} sample. The purity of each lineage was assessed by the cytological examination of May-Grünwald-Giemsa-stained cytopins and the quantification of lineage-restricted markers by RT-qPCR (**fig. 6A & fig. S8**). To compare the amount of each transcript isoform at the different stages of differentiation, we quantified the canonical ERFE^{WT} and the aberrant ERFE⁺¹² by RT-qPCR (**table S5**). The expression of the canonical transcript ERFE^{WT} expressed as Normalized Ratio Quantities (NRQ) increased in both *SF3B1*^{MUT} and *SF3B1*^{WT} MDS erythroblasts. In granulocytes, the expression of ERFE^{WT} was close to the limit of detection (**fig. 6B**). The expression of ERFE⁺¹² was restricted to the erythroid lineage, increased along the differentiation of *SF3B1*^{MUT} erythroblasts and was higher in *SF3B1*^{MUT} compared to *SF3B1*^{WT} erythroblasts (**fig. 6B**). These results indicate that ERFE⁺¹² expression is specific to the erythroid lineage.

SF3B1 gene is mutated in 15% of patients suffering from CLL and present in CD19⁺ lymphocytes (Damm et al., 2014; Wang et al., 2016). To investigate whether ERFE⁺¹² was detectable in *SF3B1*^{MUT} CLL, we collected one sample from a patient with a CLL followed by a MDS which BM MNC expressed a clonal *SF3B1*^{K700E} mutation, and 3 peripheral blood (PB) MNC samples from 2 patients with a CLL, one of them harboring a clonal *SF3B1*^{T663I} mutation with a variant allele frequency over 40% and the second with no mutation of *SF3B1* gene and the third from 1 MDS patient with a *SF3B1*^{K700E} mutation. The MAP3K7 transcript, already known to be alternatively spliced in *SF3B1*^{MUT} CLL or MDS by the use of a cryptic 3'ss, was analyzed (Dolatshad et al., 2016; Wang et

al., 2016). A 170 nt fragment corresponding to the alternative MAP3K7 transcript was enriched in all *SF3B1*^{MUT} samples compared to the *SF3B1*^{WT} CLL (**fig. 6C**). ERFE⁺¹² was detected in the PB MNC of the *SF3B1*^{K700E} MDS, but not in the *SF3B1*^{T663I} or *SF3B1*^{WT} CLL. ERFE⁺¹² was not detected in BM MNC of the *SF3B1*^{K700E} CLL+MDS sample. To get further insights on the cell types expressing ERFE⁺¹², we sorted myeloid cells containing erythroblasts, CD19⁺CD5⁻ B cells, CD19⁺CD5⁺ pathological B cells and CD3⁺ T cells from the BM MNC fraction of the *SF3B1*^{K700E} CLL+MDS, one *SF3B1*^{K700E} MDS and one *SF3B1*^{WT} MDS, and from the PB MNC of the *SF3B1*^{T663I} CLL. The number of cells in the myeloid fraction of *SF3B1*^{T663I} CLL was too small for further studies (**fig. S9**). The 170 nt fragment of MAP3K7 was detected in CD19⁺CD5⁺ pathological B cells of *SF3B1*^{K700E} CLL+MDS and *SF3B1*^{T663I} CLL and also in the myeloid fraction, but not in the CD19⁺CD5⁻ B cells of *SF3B1*^{K700E} MDS (**fig. 6D**). By sequencing *SF3B1* RNA, we demonstrated that the mutation was present in the cell populations in which the alternative MAP3K7 transcript was detected (**fig. 6E**). The alternative ERFE⁺¹² transcript was not detected in the CD19⁺CD5⁺ pathological B cells of *SF3B1*^{K700E} MDS+CLL or *SF3B1*^{T663I} CLL and its expression was restricted to *SF3B1*^{K700E} myeloid MDS cells. Altogether, these results indicate that ERFE is expressed in erythroid cells and ERFE⁺¹² is restricted to *SF3B1*^{MUT} MDS erythroid cells.

Changes in ERFE⁺¹² expression correlated with the response to lenalidomide

Fifty percent of lower risk MDS patients, including patients with MDS-RS experience primary resistance or secondary failure to treatments they receive to cure their anemia. Whether the mechanism of resistance involves the persistence of clonal erythropoiesis is always unknown.

We have previously shown that lenalidomide administered to ESA-resistant non-del(5q) MDS patients targets the malignant clone and in some cases, eliminates the dominant *SF3BI*^{MUT} clone for the duration of response (Chesnais et al., 2016). However, the frequency of the *SF3BI* variant allele, which is expressed in erythroid and myeloid cells does not only reflect the abundance of clonal erythroblasts. Here, we monitored the expression of erythroid-specific ERFE⁺¹² transcript for the follow-up of patients included in the GFM-LenEpo-08 clinical trial (NCT01718379 at clinicaltrials.gov). For this purpose, we performed a fluorescent PCR and integrated ERFE⁺¹² and ERFE^{WT} peak heights as a ratio $ERFE^{+12}/ERFE^{+12}+ERFE^{WT}$ in 14 *SF3BI*^{MUT} MDS patients for whom paired

RNA samples were available. The ratio $ERFE^{+12}/ERFE^{+12}+ERFE^{WT}$ decreased in 6 responding patients while it remained always stable in 8 non-responding patients (**fig. 7A, 7B**). The percent variation of ratios at screening and at 4 cycles of treatment was significantly different between responding and non-responding patients (Mann-Whitney test $P=0.0013$; **fig. 7C, left**). By contrast, the percent variation of *SF3BI* variant allele frequency in BM mononuclear cells between screening and evaluation at 4 cycles was less variable between 6/8 non-responding and 4/6 responding patients (**fig. 7C, right**). Altogether, our results demonstrate that ERFE⁺¹² expression allows the evaluation of clonal erythropoiesis under therapy.

Discussion

In this study, we show that in patients with *SF3BI*^{MUT} MDS, an alternative FAM132B/ERFE⁺¹² transcript is translated into an ERFE^{VPFQ} protein and, together with the canonical transcript, contributes to the overexpression of ERFE. Similarly to ERFE, the ERFE^{VPFQ} protein efficiently represses hepcidin. ERFE⁺¹² is specifically induced by the clonal erythropoiesis.

In *SF3BI*^{MUT} cancers, alternative 3'ss usages are the most frequent splicing aberrancies (Alsafadi et al., 2016; Darman et al., 2015; Shiozawa et al., 2014; Wang et al., 2016). The recently resolved crystal structure of SF3b complex helps understanding this feature (Cretu et al., 2016). In the spliceosome, SF3b complex interacts with the pre-mRNA, and is involved in the BS selection during splicing. Mutations in *SF3BI* affect the structure of SF3b RNA binding platform, leading to the selection of alternative BS and to alternative transcripts of which more than 50% are subjected to NMD (Alsafadi et al., 2016; Cretu et al., 2016;

Darman et al., 2015). Based on these findings, we focused our attention on the transcripts generated by the use of an alternative AG'. The analysis of the distance separating AG from AG' revealed strong peaks at 15, 18 and 21 nucleotides showing that inserts could be multiple of 3 nucleotides. Among alternative in-frame junctions, we identified two regulators of iron homeostasis, ABCB7 and ERFE. As already reported, the alternative ABCB7 mRNA results from the addition of 21 nucleotides between exon 8 and 9 and, in our sample set, was slightly down-regulated in *SF3BI*^{MUT} compared to *SF3BI*^{WT} MDS (Darman et al., 2015; Dolatshad et al., 2016; Nikpour et al., 2013). By contrast, the aberrant ERFE transcript modified by the addition of 12 nucleotides was up-regulated 2.3 times in *SF3BI*^{MUT} MDS. ERFE⁺¹² has never been reported in other *SF3BI*^{MUT} cancers despite some overlap between the sets of alternative 3'ss transcripts in *SF3BI*^{MUT} uveal melanoma, CLL and MDS y upstream of the collagen domain putatively

involved in protein-protein interactions in other CTRP family members (Seldin et al., 2014). VPFQ may not influence protein conformation as the apparent molecular weight in reducing and non-reducing conditions was similar for recombinant ERFE^{WT} and ERFE^{VPFQ} proteins that both repressed *HAMP* gene expression. However, we could not exclude an effect of the VPFQ insertion on protein stability and/or still unknown functions since ERFE mRNA is also expressed in non-hematopoietic tissues (Kautz et al., 2014).

In MDS, serum EPO is elevated but ineffective in stimulating erythropoiesis because of maturing erythroid cell apoptosis in terminal phases of differentiation (Gyan et al., 2008; Tehranchi et al., 2005). Elevated EPO and sTfR concentrations have been inversely correlated with hepcidin levels in MDS and β -thalassemia (Ambaglio et al., 2013; Brada et al., 1998; Metzgeroth Georgia et al., 2007; Origa et al., 2007; Santini et al., 2011). Our study indicates that, hepcidin in *SF3BI*^{MUT} MDS was similar to healthy non-blood donor controls and inappropriately low compared to *SF3BI*^{WT} MDS suggesting that its production could be dysregulated by a high degree of ineffective erythropoiesis and/or a more specific mechanism (Ambaglio et al., 2013). We show here the *SF3BI*^{MUT}-dependent expression of ERFE⁺¹² that is cell autonomous, and a role for increased levels of circulating ERFE in hepcidin reduction. This confirms the involvement of ERFE in human pathologies (Ganz et al., 2017).

ERFE is the major erythroid regulator of hepcidin (Kautz et al., 2014). Other mechanisms of hepcidin regulation have been ruled out. Indeed, the rapid decrease of hepcidin elicited by EPO administration is independent of increased iron utilization as changes in transferrin saturation are delayed, and does not depend on growth differentiation factor-15 (GDF-15), which is poorly induced by EPO (Ashby et al.,

2010). Although GDF-15 was increased in β -thalassemia or congenital dyserythropoietic anemia, it was not inversely correlated with hepcidin in MDS (Santini et al., 2011; Tamary et al., 2008; Tanno et al., 2007).

Low hepcidin preserves ferroportin on enterocytes and macrophages, causing increased intestinal iron absorption and macrophage iron release (Nemeth et al., 2004; Nicolas et al., 2003). Compared to *SF3BI*^{WT} MDS, the lower levels of hepcidin in *SF3BI*^{MUT} MDS may explain early iron overload, before patients receive erythrocyte transfusions (Ambaglio et al., 2013; Santini et al., 2011). In low transfusion burden patients who had received a mean of 0.5 RBC unit, hepcidin together with ERFE, and *SF3BI* mutation, were independent predictors of hyperferritinemia. This demonstrates that iron overload is strongly related to the control of hepcidin through the *SF3BI*-regulated production of ERFE. Hepcidin increases with transfusion dependency in patients with MDS or β -thalassemia because of exogenous iron loading and transient suppression of ineffective erythropoiesis (Cui et al., 2014; Ganz et al., 2017; Pasricha et al., 2013). Here, in regularly transfused *SF3BI*^{MUT} MDS patients, hepcidin remained lower and ERFE higher than in *SF3BI*^{WT} MDS patients. This clearly establishes a driver role for *SF3BI* mutation in the stimulation of ERFE expression and systemic iron overload that characterized patients.

The *SF3BI* mutation confers a good prognosis in MDS (Cazzola et al., 1988; Malcovati et al., 2015). However, when patients are transfusion-dependent, iron overload in cardiac and liver tissues becomes clinically evident and may impair life expectancy (Malcovati et al., 2005). Chelation efficiently reduces iron burden in regularly transfused MDS patients and represents adjuvant therapy of anemia, as deferasirox may improve hematopoiesis by protecting against oxidative stress

(Gattermann et al., 2012; Leitch, 2007; List et al., 2012; Meunier et al., 2017). Promising substitutes for erythropoiesis-stimulating agents in resistant MDS patients include lenalidomide that may transiently reduce *SF3B1* mutant allele burden, and activin receptor ligand traps such as luspatercept that rescues late-stage erythropoiesis in mouse models of β -thalassemia and MDS and diminishes iron overload in β -thalassemia mice (Chesnais et al., 2016; Dussiot et al., 2014; Platzbecker et al., 2017; Suragani et al., 2014). In lenalidomide-treated patients, the kinetics of ERFE⁺¹² expression correlated with the response to treatment, indicating that lenalidomide targeted clonal erythroid precursors either directly or indirectly in responding patients. Therefore, ERFE⁺¹² expression could be useful for disease surveillance of patients with MDS-RS

treated with other drugs, such as luspatercept. Further studies in clinical trials are required to validate ERFE⁺¹² expression as a theranostic biomarker of ineffective erythropoiesis of patients with *SF3B1*^{MUT} MDS.

Finally, our findings also open new therapeutic avenues for preventing iron accumulation in MDS-RS patients. Indeed, increasing hepcidin could produce therapeutic benefits in *SF3B1*^{MUT} MDS, as a moderate increase in hepcidin decreased iron and improved anemia in β -thalassemia mice (Gardenghi et al., 2010). Therefore, either the administration of a hepcidin agonist or the targeting of erythroferrone overexpression may provide a potential new strategy for preventing iron overload and improving erythropoiesis in *SF3B1*^{MUT} MDS patients.

Material and Methods

Patients

For the learning cohort, lower risk MDS patients (n=156) were recorded between 2008 and 2017. Bone marrow (BM) aspirates and peripheral blood (PB) plasmas were collected after each patient gave his informed consent for biological investigations according to the recommendations of institutional review board and local ethics committee (Institutional review board (IRB) numbers: IdF X GFM-LenEpo-08 EudraCT 2008-008262-12; IdFII 2010-A00033-36; IdFIII: 2010-2753; IdFV 212-A01395-38 EudraCT 2012-002990-7338; OncoCCH 2015-08-11-DC). For the validation cohort, lower risk MDS patients (n=55) were recorded prospectively after patient gave his consent for plasma collection in France (5 centres; IRB Onco-CCH 2015-08-11DC) and Germany (one centre; IRB Ethikkommission an der TU Dresden; EK 115032015) between 2016 and 2018. BM samples from 5 age-matched controls and

PB plasmas from 20 healthy controls were also collected. BM mononuclear cells were purified on Ficoll gradient. Cell pellets were processed for DNA extraction and further put in Trizol (Thermo Fisher Scientific, Waltham, MA) before RNA extraction. The patient characteristics including gender, age, WHO classification, hemoglobin (Hb) level, neutrophils, platelet counts, BM blast percentage, BM erythroblast percentage, karyotype, IPSS or IPSS-R, serum EPO level, transfusion burden according to *SF3B1* gene status are indicated in the tables S1 and S6. Low transfusion burden was defined as < 4 RBC units per 8 weeks.

Genomic studies

BM mononuclear cells were purified on Ficoll gradient and were processed for DNA extraction using the DNA/RNA Kit (Qiagen, Hilden, Germany). Mutations in a selected panel of 26 genes (*ASXLI*, *CBL*,

DNMT3A, ETV6, EZH2, FLT3, IDH1, IDH2, JAK2, KIT, KRAS, NRAS, MPL, NPM1, PHF6, PTPN11, RIT1, RUNX1, SETBP1, SF3B1, SRSF2, TET2, TP53, U2AF1, WT1 and *ZRSR2*) were screened by a Next-Generation Sequencing (NGS) assay using the Ion AmpliSeq™ library kit2 384 (Life Technologies, Chicago, IL). Multiplex PCR amplifications (233 primer pairs) were performed from 2x10 ng of genomic DNA. After amplification, barcodes and adaptors were added to amplicons by ligation. Products were subjected to a selective purification on AMPure beads (Life Technologies). Emulsion PCR (emPCR) was performed using the OneTouchV2 (Life Technologies) instrument. Sequencing was performed on Ion PGM™ (Life Technologies) onto the 318 V2 chip (15 samples per chip). All the samples were also screened for *ASXL1* (including c.1934dupG; p.G646WfsX12) and *SRSF2* mutations by Sanger sequencing. *JAK2, NPM1*, and *FLT3-ITD* mutations were also investigated by qPCR and fluorescent PCR. For analysis, base calls were generated by the Torrent Browser software using the included variant caller with an additional plug-in (Life Technologies). The .bam and .vcf files were used for further analysis. The .vcf files were annotated with the Ion reporter software (Life Technologies) and processed for a second analysis of the indexed files using the NextGENe software (Softgenetics, State College, PA). Results were compared to select abnormalities that will be further considered. Mutation patterning is shown in fig. S1.

RNA-sequencing

cDNA synthesis was conducted with MuLV Reverse Transcriptase (Invitrogen, Carlsbad, CA), with quality controls conducted on an Agilent 2100 Bioanalyzer (Agilent, Les Ulis, France). Libraries were constructed using the TruSeq Stranded mRNA Sample Preparation Kit (Illumina,

San Diego, CA) and sequenced on an Illumina HiSeq 2500 platform using a 100-bp paired-end sequencing strategy. An average depth of global sequence coverage of 114 million and a median coverage of 112 million was attained. For analysis, TopHat (v2.0.6) was used to align the reads against the human reference genome Hg19 RefSeq (RNA sequences, GRCh37) downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>).

Bioinformatic analysis of RNA-sequencing

FASTQ files were mapped using TopHat (v2.0.6)28 to align the reads against the human reference genome Hg19 RefSeq (RNA sequences, GRCh37) downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>). Read count normalizations and groups comparisons were performed using DESeq2, which converts the counts to log-counts per million with associated precision weights and flags were computed using a custom algorithm within R (The R project for Statistical Computing [<http://www.r-project.org/>]). Assuming that a maximum of 80% of genes are expressed we selected the 20% lowest counts for each sample as background. A threshold is fixed at two standard deviations over the mean of the background. All transcripts for which normalized counts were lower than the computed threshold were designated as background for each array. When comparing normalized counts between two groups, a transcript was included in the analysis if its counts exceeded the background in at least 80% of the sample from at least 1 group. To identify differentially expressed transcripts, we used empirical Bayes estimation. For differential expression study, the results obtained after Deseq2 comparison were selected for further analysis and filtered at P -value ≤ 0.05 and fold-changes of 1.2, 1.5 or 2. Principal component analysis calculated from averaged RefSeq exons robust multiarray average (RMA) values

by RefSeq gene symbol was generated using R and the GGplot2 package. The bivariate normal density is a function of the means and standard deviations of the X and Y variables, PC 1 and PC2. The ellipsoid was computed from the bivariate normal distribution fit to each group. Read counts for splicing junctions from junctions.bed TopHat output were considered. Differential analysis was performed on junction read counts using DESeq2 (Love et al., 2014). Only alternative acceptor splice sites (two or more 30ss with junctions to the same 50ss) and alternative donor splice sites (two or more 50ss with junctions to the same 30ss) were considered for this analysis. According to Alsafadi et al., predictions and analysis of branch point sequence analysis were done with the online tools SVM-BPfinder (http://regulatorygenomics.upf.edu/Software/SVM_BP/) and the Human Splicing Finder (<http://www.umd.be/HSF/>) (Alsafadi et al., 2016; Corvelo et al., 2010; Desmet et al., 2009). The SVM_BP code was altered to allow for BP six base pairs from the 3'ss by setting `minidist3ss=6` in `svm_getfeat.py`.

Fluorescent PCR and quantitative PCR

RNA was extracted using RNeasy Mini Kit (Qiagen) and used as a template for cDNA synthesis with the Maxima First Strand cDNA Synthesis Kit (Thermo Fisher Scientific). cDNA was then amplified by PCR (Phire Hot Start II DNA Polymerase, Thermo Fisher Scientific) using specific 5'FAM-forward or 5'FAM-reverse primers (table S5). Finally, DNA fragment analysis was performed by capillary electrophoresis: 1 µl of diluted PCR product was added to 0.2 µl of GeneScan 500 ROX dye Size Standard (Thermo Fisher Scientific) and 18 µl of RNase-free water. The DNA was denatured 5 minutes at 95°C and placed at 4°C. Fragments were separated using a 3730xl DNA Analyzer and data were analyzed using GeneMapper

Software 5 (Thermo Fisher Scientific). For RT-qPCR, the mean amplification efficiency of each PCR was 95%. According to the MIQE guidelines, the expression of each gene of interest was normalized on the expression of two housekeeping genes *B2M* and *ACTB*. The list of primers is given in table S5.

Transfection

UT-7/EPO cell line was transfected with a synthetic full length *SF3B1*^{WT} or mutant *SF3B1*^{K700E} cDNA provided by Dr S. Buonamici (H3 Biomedicine Inc, Cambridge, MA) using the Amaxa program T-024 in R solution (Lonza, Basel, Switzerland).

*CRISPR/Cas9 generation of isogenic *Sf3b1*^{K700E} and *Sf3b1*^{WT} cell lines*

The murine erythroid cell line G1E-ER4 was used to generate isogenic *Sf3b1*^{K700E} and *Sf3b1*^{WT} cell lines using CRISPR/Cas9-stimulated homology-mediated repair (Welch et al., 2004). Cells were transfected using Amaxa Nucleofector™ 2b Device program A-024 and Nucleofector kit L (Lonza) with Cas9 (Ref. 41815, Addgene, Cambridge, MA), a *mSf3b1*-specific gRNA (built from gRNA cloning vector, Ref. 41824, Addgene) and a donor template encoding a hygromycin selection cassette at a 1:1:1 ratio. After two days, hygromycin was added to the culture at 800 µg/mL for 72h, and maintained at 400 µg/mL for another 72h. Hygromycin-resistant cells were then single cell-sorted using the FACS Aria III (Becton Dickinson Biosciences, Franklin Lakes, NJ) and clones were analyzed 3 weeks later by DNA sequencing to assess the presence of the *Sf3b1*^{K700E} mutation. The selection cassette was removed by flippase-mediated excision. *Sf3b1*^{K700E} and *Sf3b1*^{WT} hygromycin-resistant clones were transfected with a flippase-encoding vector and single cell-sorted. Cassette-excised clones were assessed by PCR and

Sf3b1^{K700E} and *Sf3b1*^{WT} expression was verified by RT-PCR and cDNA Sanger sequencing.

ERFE and ENOSF1 minigenes

The *ERFE* minigene was synthesized from the linearized pET01 Exontrap vector (Mobictec, Göttingen, Germany) in which we insert an *ERFE* sequence centered on alternative 3' splice sites (AG^{3'}) flanked upstream of 100 bp in intron 2-3 and downstream of 100 bp in exon 3. The *ERFE* insert was PCR-amplified from the genomic DNA of MOLM-14 cells using Phusion High-Fidelity DNA Polymerase 2 U/μl (Thermo Fisher Scientific), and purified with the QIAquick PCR Purification Kit (Qiagen). The primer sequence information is provided in table S5. We introduced 16 bp of homology with the ends of the linearized vector at 5'-end of the forward and reverse primers. Using In-Fusion HD Cloning Kit (Clontech, Mountain View, CA), we cloned the purified amplicon between the XhoI and XbaI restriction sites of Exontrap vector containing a functional splice donor site. A validated *ENOSF1* minigene was used as a control (Alsafadi et al., 2016). G1E-ER4 9.2 (*Sf3b1*^{WT}) and G1E-ER4 5.13H (*Sf3b1*^{K700E}) cells were transfected by electroporation (Amaxa NucleofectorTM 2b, Lonza) with 400 ng of *ERFE* or *ENOSF1* minigene in L solution. After 24 hours, total RNA was extracted and processed for fluorescent PCR.

Primary cell culture

For erythroid and granulocytic cell expansion, CD34⁺ cells were isolated from the mononuclear cell fraction of bone marrow samples using the MidiMacs system (Miltenyi Biotec, Bergisch Gladbach, Germany). For erythroid differentiation, CD34⁺ cells, which purity was higher than 80%, were cultured at 0.8 × 10⁶ per mL for 4 days in Iscove's modification of Dulbecco medium containing 15% BIT9500 (Miltenyi

Biotech), penicillin/streptomycin, L-Glutamine, Epo 2 UI/mL (Roche, Basel, Switzerland), SCF 100 ng/mL (Miltenyi Biotec), IL6 10 ng/mL (Miltenyi Biotec) and dexamethasone 2.10⁻⁷ M (Sigma Merck, Darmstadt, Germany). Cells were diluted every day in the same medium until day 4. From day 4, IL6 was removed. From days 10 to 16 cells were switched to Epo (2 UI/mL) to obtain terminal erythroid differentiation followed by flow cytometry (GPA, CD71, CD49d and Band3). For granulocytic differentiation, CD34⁺ cells were cultured in the same medium with SCF 100 ng/mL, IL3 10 ng/mL (Miltenyi Biotec) and G-CSF 20 ng/mL (Sandoz, Holzkirchen, Germany). Every 2 days, MGG staining and RT-qPCR for erythroid genes (*GATA1*, *HBB*) or granulocytic genes (*MPO*, *SPI1*) were performed.

Cell sorting

Cells were sorted using CD45-PE, CD3-FITC, CD19-ECD and CD5-AA700 antibodies (Beckman Coulter, Brea, USA) from the mononuclear cell fraction of bone marrow or blood samples using FACS ARIA3 device (Becton Dickinson, Franklin Lakes, USA). CD45⁺CD19⁻CD3⁺CD5⁻ T lymphoid cells, CD45⁺CD19⁺CD3⁻CD5⁻ healthy B cells, CD45⁺CD19⁺CD3⁻CD5⁺ pathologic B CLL cells and myeloid cells were gated as described in fig. S9. In sorted cell populations, RNA was extracted and *SF3B1* gene was sequenced by Sanger method.

Mass spectrometry analysis

Sequencing grade trypsin (50ng) was used to digest proteins of erythroblast lysates in solution at 37°C overnight (Promega, Fichburg, WI). Resulting products were analyzed by nano liquid chromatography hyphenated to a Q-Exactive Plus mass spectrometer (Thermo) operating in data dependent scheme: peptides loaded on the chromatography were trapped and washed

with 0.1% TFA 2% Acetonitrile in milliQH₂O on a C18 reverse phase precolumn (Acclaim Pepmap 100 Angstroms pores, 5µm particles, 2cm long, 75µm inner diameter) for 3 minutes at 5µL/min. Trapped peptides were then separated on a C18 reverse phase analytical column (2 µm particle size, 100 Å pore size, 75 µm inner diameter, 25 cm length) with a 1.5 hours gradient starting from 99% of solvent A containing 0.1% formic acid in milliQH₂O and ending in 40% of solvent B containing 80% ACN and 0.085% formic acid in milliQH₂O. The mass spectrometer acquired data throughout the elution process. The MS scans spanned from 350 to 1500 Th with AGC target 1.106 with 60ms MIIT and resolution of 70 000. HCD fragmentations were performed on the 10 most abundant ions with a dynamic exclusion time of 30 s. Precursor selection window was set at 2Th. HCD Normalized Collision Energy (NCE) was set at 27% and MS/MS scan resolution was set at 17,500 with AGC target 1.105 within 60ms MIIT. Spectra were recorded in profile mode. MS data were analyzed using Mascot 2.5.1 (www.matrixscience.com). The database used was a concatenation of human sequences from the Uniprot-Swissprot database (release 2017-10, 20314 sequences) and a customized database of modified sequences including the ERFE^{VPFQ} sequence, and a list of in-house common contaminant sequences. Cystein carbamidomethylation was set as constant modification, methionine oxidation and proline hydroxylation were set as variable modifications. Spectrum annotation was done according to nomenclature of peptide fragmentation proposed by Roepstorff and Fohlman (Roepstorff and Fohlman, 1984).

Recombinant ERFE^{WT} and ERFE⁺¹² production and purification

Human ERFE^{WT} and ERFE⁺¹² cDNA sequences were cloned into pcDNA3.1 with the following modifications: vector

signal sequence (interleukin-2) was used instead of the native, followed by a spacer and a FLAG tag. Recombinant proteins were produced in suspension culture in Freestyle 293F cells (Life Technologies) transiently transfected using Freestyle F-MAX reagent (Life Technologies). Supernatants from cells overexpressing FLAG-tagged ERFE proteins were collected after 5 days and supplemented with protease inhibitor cocktail. Recombinant proteins were purified using an anti-FLAG affinity gel according to the manufacturer's protocol (Sigma-Aldrich Chimie, Saint-Quentin-Fallavier, France) and eluted with 150 µg/mL FLAG peptide (Sigma). FLAG peptide was eliminated by filtering preparations through an Amicon Ultra 30K device (Merck Millipore, Guyancourt, France), and recombinant ERFE proteins were suspended in a saline solution (0.9% NaCl). Protein concentration was determined using Coomassie Imperial Protein Stain and Pierce bicinchoninic acid protein assay (Thermo Fisher Scientific).

Western blot analysis

Supernatants from 293F cells secreting ERFE^{WT} or ERFE^{VPFQ} cells were subjected to SDS-PAGE in reducing (diluted in Laemmli buffer 10% 2-mercaptoethanol and incubated for 5 min at 95° C) and non-reducing (diluted in Laemmli buffer without 2-mercaptoethanol) conditions and electroblotted to nitrocellulose membranes. Membranes were blocked with 5% of dry milk in TBS-T buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.15% Tween 20) and incubated 2h at RT with an anti-DDK monoclonal antibody conjugated to horseradish peroxidase (HRP) (86861, Cell Signaling), which recognizes DYKDDDDK epitope (FLAG-Tag) or with rabbit polyclonal antibodies to ERFE (produced by YenZym) overnight at 4° C before incubation with a goat anti-rabbit HRP-linked antibody (7074S, Cell signaling). Enzyme activity was visualized

by an ECL-based detection system (GE Lifesciences). Blot imaging and analysis was performed on a Chemidoc MP Imaging System (Bio-Rad) with the Image Lab software.

Hepcidin repression in Hep3B hepatoma cell line

Hep3B hepatoma cell line were incubated overnight (16 hours) in D-MEM Glutamax (Life Technologies) supplemented with 10% fetal bovine serum and 50% (v/v) supernatants from control freestyle 293-F cells or 293F cells transfected with a plasmid encoding *ERFE*^{WT} or *ERFE*⁺¹² cDNA. Total RNA from Hep3B cells was extracted using UPzol Total RNA Isolation Reagent (Biotech Rabbit). cDNA was synthesized using MMLV-RT (Promega). Quantitative PCR (qPCR) reactions were prepared with LightCycler® 480 DNA SYBR Green I Master reaction mix (Roche Diagnostics) and run in duplicate on a LightCycler® 480 System (Roche Diagnostics). HAMP mRNA transcript abundance was normalized to the reference gene *HPRT*. Primers are listed in table S5.

Human hepcidin and soluble transferrin receptor dosages

Plasma soluble transferrin receptor assays were performed on Dimension VISTA 1500 (Siemens Healthcare, Saint Denis, France). Circulating hepcidin was quantified from EDTA-plasma samples by the previously described liquid chromatography coupled with tandem mass spectrometry (LC-MS/MS) method (Lefebvre et al., 2015)

Human erythroferrone immunoassay

Plasma erythroferrone concentration was determined as previously described²⁸. Briefly, high binding 96-well plates

(Corning) were coated with overnight at 4°C with 100 µL/well of 1.0 µg/mL capture antibody in 50 mM sodium carbonate buffer (pH 9.6). Plates were washed (TBS, 0.5% Tween-20) and blocked for 1h at RT with 200 µL/well blocking buffer (PBS, 0.2% Na casein, 0.05% Tween 20, 0.1M NaCl). Recombinant human ERFE standard was serially diluted to 10, 5, 2.5, 1.25, and 0.625 ng/mL. After a 1-hour incubation at room temperature, the plate was washed and incubated for 1 hour with 100 µL per well biotinylated rabbit monoclonal antibody to human erythroferrone (1 µg/mL). The plate was then washed, incubated for 45 minutes with Neutravidin-horseradish peroxidase conjugate (ThermoScientific #31030) 1/5000 (100 µL/well), washed 3 times and developed with 100 µL TMB Substrate System for ELISA (ThermoScientific #34028) at RT in the dark for 10 min. The reaction was stopped by adding 50 µl of 2N sulfuric acid and plates were read on a Spectramax 250 (Molecular Devices) at 450 nm.

Statistical analysis

For quantitative variables, values were expressed as median and interquartile range (IQR) or means and standard error of the mean (SEM) and compared using the Mann-Whitney test. Categorical variables reported as counts or percentages were compared using Chi-square or Fisher's exact tests. For transcript quantification, the Mann-Whitney test was used to assign a statistical significance for each group comparison. *P*-values <0.05 were considered significant. A multivariate logistic regression analysis was adjusted for selected variables chosen with a *P*-value <0.1 in univariate analysis (JMP version 10.0.2, SAS Institute Inc, Cary, NC).

Supplemental figures and tables

Fig. S1: Related to Table S1: **Genomic characteristics of the studied MDS cohort.**

Fig. S2: Related to Fig. 2: **Identification of in-frame alternative transcripts in $SF3B1^{MUT}$ MDS samples.**

Fig. S3: Related to Fig. 3: **Massive use of FAM132B/ERFE cryptic junction in one MDS-RS with a bi-allelic alteration of $SF3B1$ gene.**

Fig. S4: Related to Fig. 4: **Identification of ERFE peptide by LC tandem mass spectrometry.**

Fig. S5: Related to Fig. 5: **Erythroferrone and hepcidin/ferritin ratio according to the WHO classification 2016 of the learning cohort.**

Fig. S6: Related to Fig. 5: **Biological parameters of 156 MDS patients of the learning cohort.**

Fig. S7: Related to Fig. 5: **Biological parameters of low transfusion burden MDS patients of the learning cohort.**

Fig. S8: Related to Fig. 6: **Detection of ERFE⁺¹² in $SF3B1^{MUT}$ erythroblasts.**

Fig. S9: Related to Fig. 6: **Gating strategy for myeloid and lymphoid cell sorting.**

Table S1: **Clinical and biological characteristics of the learning cohort of 156 MDS patients according to $SF3B1$ status.**

Table S2: **Differentially expressed transcripts between $SF3B1^{MUT}$ and $SF3B1^{WT}$ MDS patients by RNA-seq.**

Table S3: **Differentially expressed 5' and 3' junctions between $SF3B1^{MUT}$ and $SF3B1^{WT}$ MDS patients by RNA-seq.**

Table S4: **In-frame and differentially expressed 3' cryptic splice site junctions in $SF3B1^{MUT}$ MDS.**

Table S5: **Primer sequences used for fluorescent RT-PCR and non-fluorescent RT-qPCR.**

Table S6: **Clinical and biological characteristics of the validation cohort of 55 MDS patients according to $SF3B1$ status.**

Acknowledgements

The authors want to thank Dr Marc-Henri Stern from INSERM U830, Institut Curie (Paris), Dr Françoise Pflumio from INSERM UMR967, CEA/DSV/iRCM (Fontenay-aux-Roses), Dr Naomi Taylor for very helpful discussions, Dr Michel Wassef and Dr Raphaël Margueron for very helpful contribution to CRISPR-Cas9 construction, Dr Emilie-Fleur Gautier from the 3P5 proteomic platform and Dr Joelle Vinh from the Ecole Supérieure de Physique et de Chimie Industrielles

(ESPCI), CNRS USR3149, Paris for fruitful discussions; Dr Jean-Benoît Arlet, Service de Médecine Interne, Hôpital Européen G. Pompidou, Université Paris Descartes, Paris, Dr Caroline Kannengiesser, Service de Génétique, Hôpital Bichat, Université Paris Diderot for providing patient samples; Dr Sophie Kaltenbach, Laboratoire de Génétique, Hôpital Necker; Dr Franck Letourneur from the Genom'IC platform, Dr Sylvain Baulande from the Next generation

sequencing platform of Institut Curie, Paris, Ms Evangeline Benana for proteomics sample processing; Benjamin Saint-Pierre, bioinformatician funded by the Site de Recherche Intégrée sur le Cancer (SIRIC) CANcer Research for Personalized Medicine CARPEM, Damien Drubay, biostatistician, Institut Gustave Roussy, Victoria Verjus-Lisfranc for assistance in recording clinical data at

Hôpital Cochin; Pr Mitchell Weiss, St Jude Children's research hospital (Memphis, TN) and Dr Vikram Paralkar, Department of Medicine, University of Pennsylvania (Philadelphia, PA) for providing the G1E-ER4 cell line and expert comments; Dr Silvia Buonamici, H3Biomedicine (Cambridge, MA) for providing plasmids.

Funding

This study was funded by INSERM, by the Institut National du Cancer INCa PLBio 2015 (INCa_9290) and the Direction Générale de l'Offre de Soins (DGOS) of the French Ministry of Social Affairs and Health through the Programme Hospitalier de Recherche Clinique (PHRC MDS-04; INCa-DGOS_5480). CL is the recipient of a post-doc funding from the Laboratory of Excellence on red cells GR-Ex. The Orbitrap Fusion mass spectrometer was acquired with funds from the FEDER through the "Operational Program for Competitiveness Factors and employment 2007-2013" and from the "Canceropole Ile de France". LK received a support from ANR-16-ACHN-0002-01 and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant agreement N0 715491). TG received funding from NIH by R01 DK 065029 (and the UCLA Center for Accelerated Innovation, under NIH grant U54HL119893 (Palazzolo). Author contributions: SB, A-SA, CL, GJ, FG, TL, BB, AR and LK performed the

experiments and analyzed the data; AR, SW, A-SK, recorded patient clinical data; AH, AB, NC and IB performed bioinformatics analyses; AT, SP, DB, HP, LA, UP, TC, MH, AS, and EG provided samples; SP performed statistical analysis and reviewed the manuscript; FG, ND, SA, ZK, HP, PM, EN, TG, LK analyzed the data and reviewed the manuscript; MF and OK designed the study, analyzed the data, supervised the work; MF wrote the manuscript. Competing interest: T.G., L.K., and E.N. are inventors on a patent application on ERFE. T.G. and E.N. are scientific founders of Intrinsic LifeSciences and Silarus Pharma, companies that have interests related to ERFE. The other authors have no competing financial interest to disclose. Data and material availability: RNA-seq data (tables S2 and S3) are available in the Gene Expression Omnibus (GEO) is a database repository (accession number GSE113433); Material transfer agreement with H3 Biomedicine for synthetic full length *SF3B1*^{WT} or mutant *SF3B1*^{K700E} cDNAs.

Figures

Figure 1: Differential gene expression in *SF3B1*^{MUT} MDSs

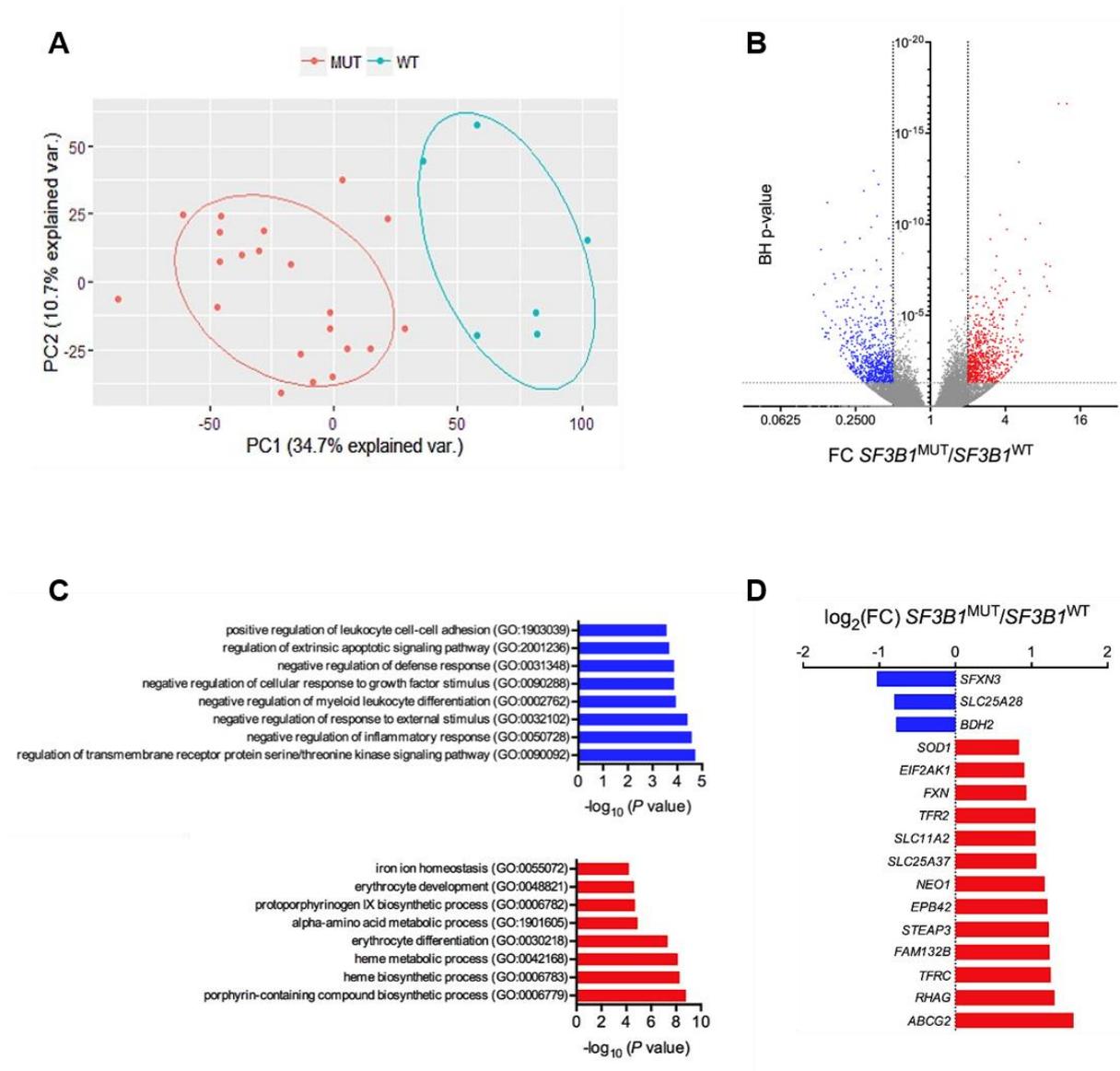


Fig. 1: Differential gene expression in *SF3BI*^{MUT} MDSs. (A) Principal component analysis (PCA). Bi-plot of the first 2 principal components (PC) showing 45.4% of the variability within the data (PC1, x-axis; PC2, y-axis). Most of the variability in the data represented by PC1 (34.7%) shows *SF3BI*^{MUT} MDS separated from *SF3BI*^{WT} MDSs. (B) Volcano plot showing differentially expressed transcripts in *SF3BI*^{MUT} bone marrow (BM) mononuclear cells (MNC) compared to *SF3BI*^{WT} BM MNCs. The x-axis specifies the fold-changes (FC) and the y-axis specifies the negative log₁₀ of Benjamini-Hochberg (BH) corrected P-values. Dashed vertical and horizontal lines reflect the filtering criteria (FC < 0.5 or > 2.0 and BH-corrected *P-value* < 0.05). The red dots represent transcripts differentially up regulated, the blue dots represent transcripts differentially down regulated and the gray dots represent transcripts without differential expression. (C) GO enrichment analysis of 6,343 differentially expressed genes with an absolute log₂ fold-change (FC) > 1 and a *P-value* < 0.05 in *SF3BI*^{MUT} MDS. The most down-regulated and up-regulated gene sets are represented. (D) Expression level in log₂(FC) of 19 genes among the 96 genes of the IRON_ION_HOMEOSTASIS gene set (GO:005072; <http://amigo.geneontology.org>) deregulated in *SF3BI*^{MUT} MDS.

Figure 2: Differential splice junctions in *SF3B1*^{MUT} MDSs

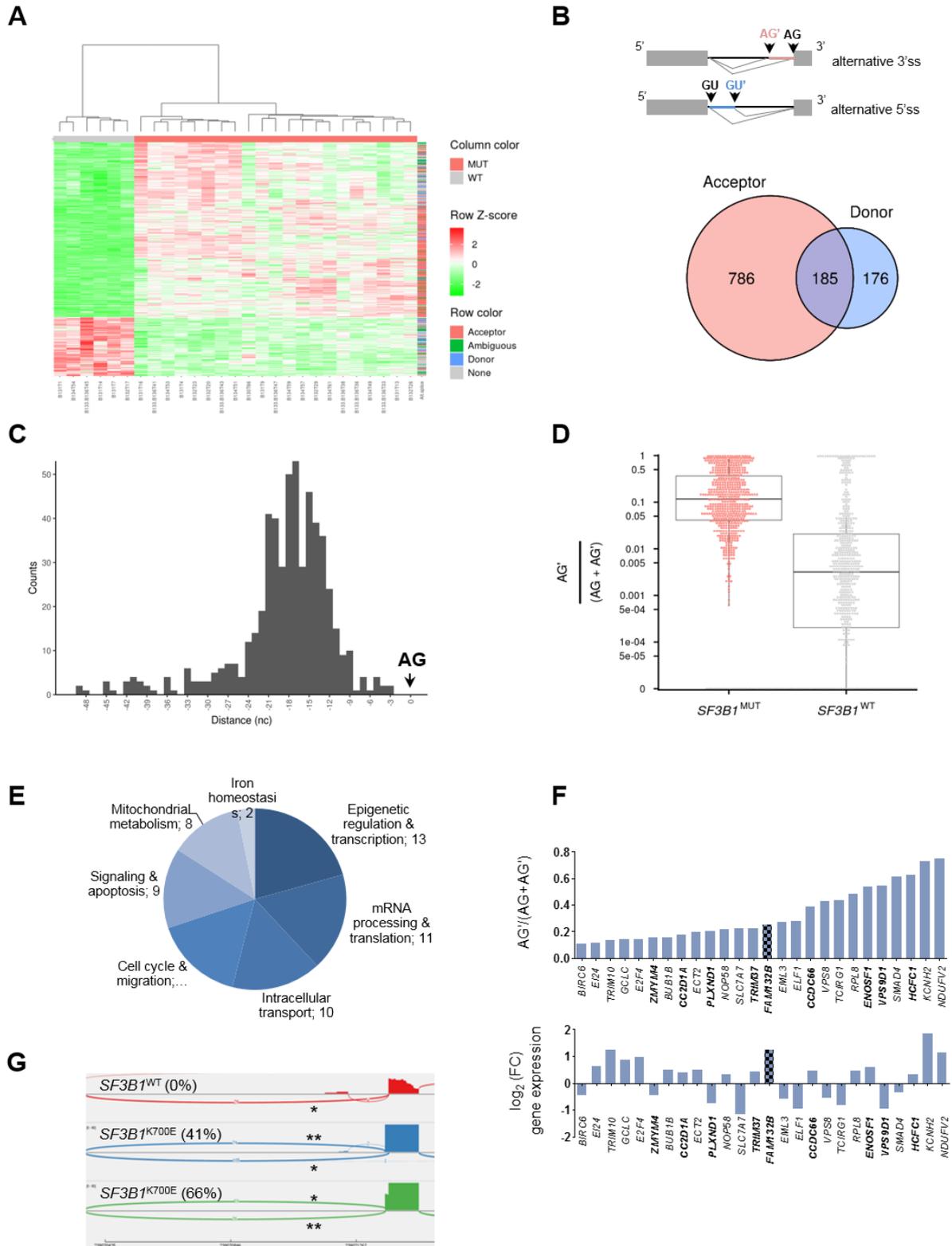


Fig. 2: Differential splice junctions in *SF3B1*^{MUT} MDSs. (A) Hierarchical clustering and heat-map of 1,528 differential splice junctions between *SF3B1*^{MUT} and *SF3B1*^{WT} MDS samples. Values indicating percent usage of the differential splice junction versus all other junctions sharing the same splice site are normalized as Z scores across patients and limited to a maximum of $|Z| = 2$. Rows are splice junctions and columns are patients. For each row, the type of splice junctions is indicated by a distinct color in the right column: 3'ss (red), 5'ss (blue), ambiguous (green) and differentially expressed canonical junction (gray). The ambiguous junctions are those that are attributed to an alternative donor or acceptor site, or both. The complete list of up- and down-regulated splice junctions is provided in table S3. (B) Schematic representation of alternative 5' donor splice site (ss) and 3' acceptor splice site. Venn diagram of differential splice junctions in *SF3B1*^{MUT} compared to *SF3B1*^{WT} MDS. The overlapping area represents ambiguous junctions (n = 185). (C) Distances between the alternative and canonical 3'ss. For the 870 alternative 3'ss (AG') located within the 50 nucleotides upstream of the canonical 3'ss (AG), the distance between AG' and AG was plotted as a histogram. (D) Comparison of the expression level of alternative junctions. The ratio AG'/AG'+AG for each junction in *SF3B1*^{MUT} (red) versus *SF3B1*^{WT} (gray) MDS is shown. (E) Repartition of the biological functions of 63 genes affected by one or two aberrant 3'ss junctions (AG') located at <50 bases of the canonical 3'ss (AG), an additional sequence multiple of 3 nucleotides, a ratio AG'/(AG'+AG) in *SF3B1*^{MUT} samples >0.1 and a fold-change >10. (F) Expression level of 3'ss aberrant junction shown as a ratio of AG'/(AG'+AG) in 26 genes which expression was globally up or down-regulated in *SF3B1*^{MUT} patients. Bold font indicates transcripts without a premature termination codon (G) Sashimi plot of 3'ss canonical (*) and aberrant (**) junctions between exon 2 and 3 in *FAM132B/ERFE* gene located at chr2: 239,070,357 – 239,071,364 in three bone marrow mononuclear cell samples, 1 *SF3B1*^{WT} and 2 *SF3B1*^{K700E} MDS. The aberrant 3'ss junction is defined by the addition of 12 nucleotides. The percentage of aberrant junction AG'/(AG'+AG) is indicated within brackets.

Figure 3: *SF3B1*-dependent expression of 3'ss aberrant ERFE⁺¹²

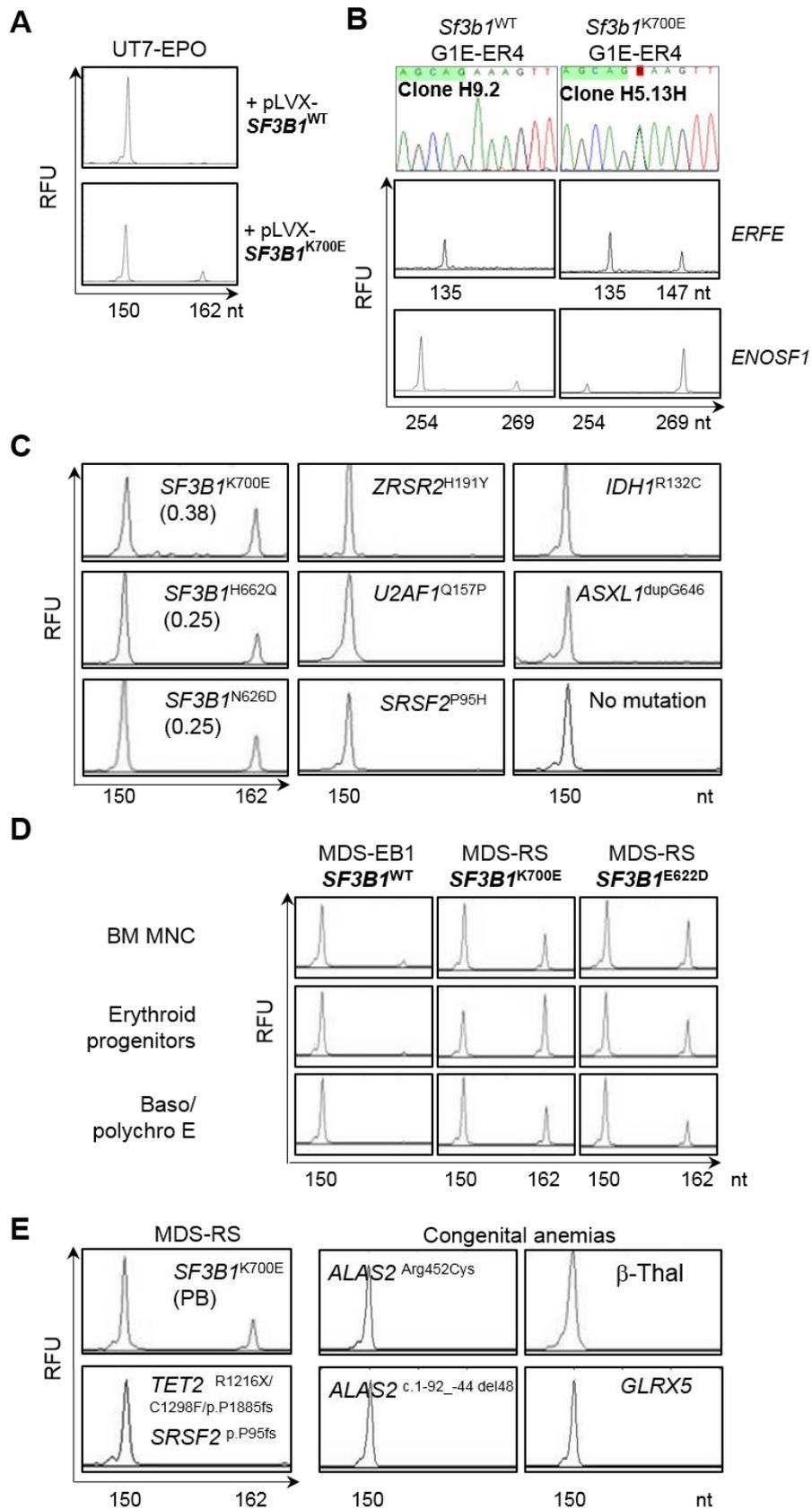


Fig. 3: *SF3B1*-dependent expression of 3'ss aberrant $ERFE^{+12}$. (A) Induction of $ERFE^{+12}$ by expressing $SF3B1^{K700E}$ in human $SF3B1^{WT}$ erythro-megakaryocytic UT-7/EPO cell line. Cells were transfected using a pLVX plasmid encoding a synthetic $SF3B1^{WT}$ or $SF3B1^{K700E}$ cDNA. The canonical $ERFE$ and aberrant $ERFE^{+12}$ transcripts were detected by capillary electrophoresis of fluorescent RT-PCR products. The X-axis represents molecular size (nt for nucleotides) of PCR amplicons, and the Y-axis represents relative fluorescent units (RFU). The peak of 150 nt corresponds to the canonical transcript while the peak of 162 nt refers to the alternative transcript due to cryptic AG' usage. (B) Minigene assay. Analysis of alternative AG' and canonical AG usage of $ERFE$ and $ENOSF1$ minigenes in murine CRISPR-Cas9 $Sf3b1^{WT}$ G1E-ER4 cells (clone H9.2) and CRISPR-Cas9 $Sf3b1^{K700E}$ G1E-ER4 cells (clone H5.13H) by capillary electrophoresis of RT-PCR products. The peak of 135 nt corresponds to the transcript generated by a canonical AG usage while the peak of 147 nt refers to the alternative transcript due to cryptic AG' usage. The alternative $ENOSF1$ transcript is much more used in the H5.13H clone. (C) Detection of $ERFE^{+12}$ depends on the presence of a $SF3B1$ mutant in MDS. Bone marrow mononuclear cells (BM MNC) RNAs from $SF3B1^{MUT}$ or $SF3B1^{WT}$ MDS were collected. Three patients had a $SF3B1$ mutation ($SF3B1^{K700E}$, $SF3B1^{H622Q}$, $SF3B1^{N626D}$), 3 had a mutation in splice genes ($ZRSR2^{H191Y}$, $U2AF1^{Q157P}$, $SRSF2^{P95H}$) other than $SF3B1$ and 3 had no mutation in splice genes. The ratio ($ERFE^{+12}$)/($ERFE^{+12}$ + $ERFE^{WT}$) is indicated in brackets. (D) Detection of $ERFE^{WT}$ and $ERFE^{+12}$ transcripts in $SF3B1^{MUT}$ erythroblasts in culture. (E) Absence of $ERFE^{+12}$ in $SF3B1^{WT}$ diseases with ineffective erythropoiesis. Peripheral blood (PB) mononuclear cells from 1 patient with an acquired sideroblastic anemia with no $SF3B1$ mutation but $TET2$ and $SRSF2$ mutations, and samples from 4 patients with a congenital anemia: 3 sideroblastic anemias (2 BM samples with an $ALAS2$ mutation and 1 PB sample with a $GLRX5$ mutation) and one PB sample for a patient with a severe β -thalassemia patient were analyzed by capillary electrophoresis of $ERFE$ RT-PCR products. In this series, one PB sample from a $SF3B1^{K700E}$ MDS patient was used as a positive control.

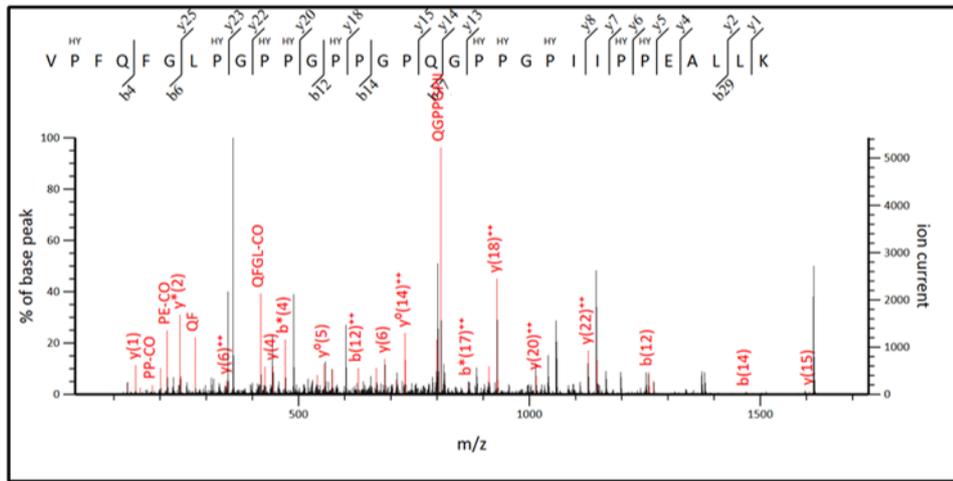
Figure 4: Identification of ERFE^{VPFQ} peptide by mass spectrometry and hepcidin repression by recombinant variant ERFE^{VPFQ} protein

A

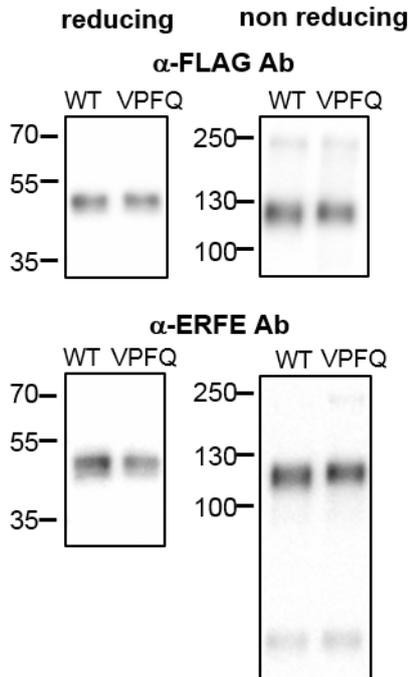
N ERFE^{WT} C
MLFVQRQSDKGVNGKKRSRGKAKKLLK FGLPGPPGPPGPPGPPGIIPPEALLKEFQLLLKG

N ERFE^{VPFQ} C
MLFVQRQSDKGVNGKKRSRGKAKKLLK VPFQ FGLPGPPGPPGPPGPPGIIPPEALLKEFQLLLKG

B



C



D

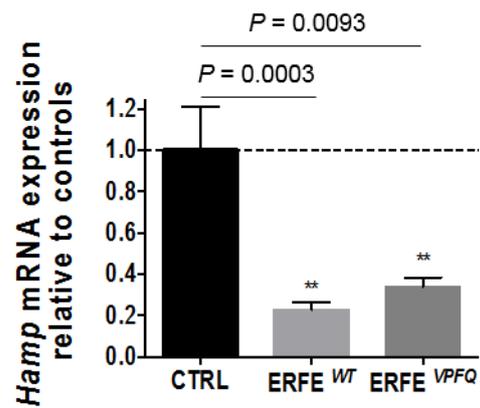


Fig. 4: Identification of ERFE^{VPFQ} peptide by mass spectrometry and hepcidin repression by recombinant variant ERFE^{VPFQ} protein. (A) Amino-acid sequence of ERFE^{WT} and ERFE^{VPFQ} peptides. VPFQ (red) is located immediately upstream of the collagen domain (blue). (B) Identification of a specific ERFE^{VPFQ} peptide by mass spectrometry. Peptides were prepared from erythroblast cell lysates and analyzed by nano liquid chromatography hyphenated to a Q-Exactive Plus mass spectrometer. For peptide identification, MS data were analyzed using Mascot 2.5.1. The calculated peptide mass was 3227.650782 from 3+ ion observed at M/Z 1076.890870 with measured delta = 3.6 ppm, Mascot Score = 17, expectation value = 0.069. The sequence is a proline-rich region with a high content of predicted hydroxylated (HY) proline residues (HY=+15.9945 amu). (C) ERFE^{WT} and ERFE^{VPFQ} exhibit a similar molecular weight. Immunoblot analysis of ERFE^{WT} and ERFE^{VPFQ} using anti-FLAG and anti-ERFE antibodies in reducing and non-reducing conditions showed that the insertion of 4 aminoacids does not change the molecular weight of the protein or its conformation. Both proteins are mostly found in a trimeric conformation exhibiting a molecular weight of approximately 130 kDa. (D) ERFE^{VPFQ} is a potent suppressor of hepcidin. Recombinant human ERFE^{WT} and ERFE^{VPFQ} were produced in 293F cells. Hep3B hepatoma cells were treated with supernatants of 293F cells overexpressing ERFE^{WT} and ERFE^{VPFQ} for 16 hours. *HAMP* expression was quantified by RT-qPCR and its expression was normalized to *HPRT* quantities. Data shown are means \pm SEM of four independent experiments and represent a fold change of hepcidin mRNA expression in ERFE treated compared to untreated (CTRL) cells. Two-tailed Student t-test for *P*-values.

Figure 5: Increased plasma concentration of ERFE in *SF3B1*^{MUT} MDS patients

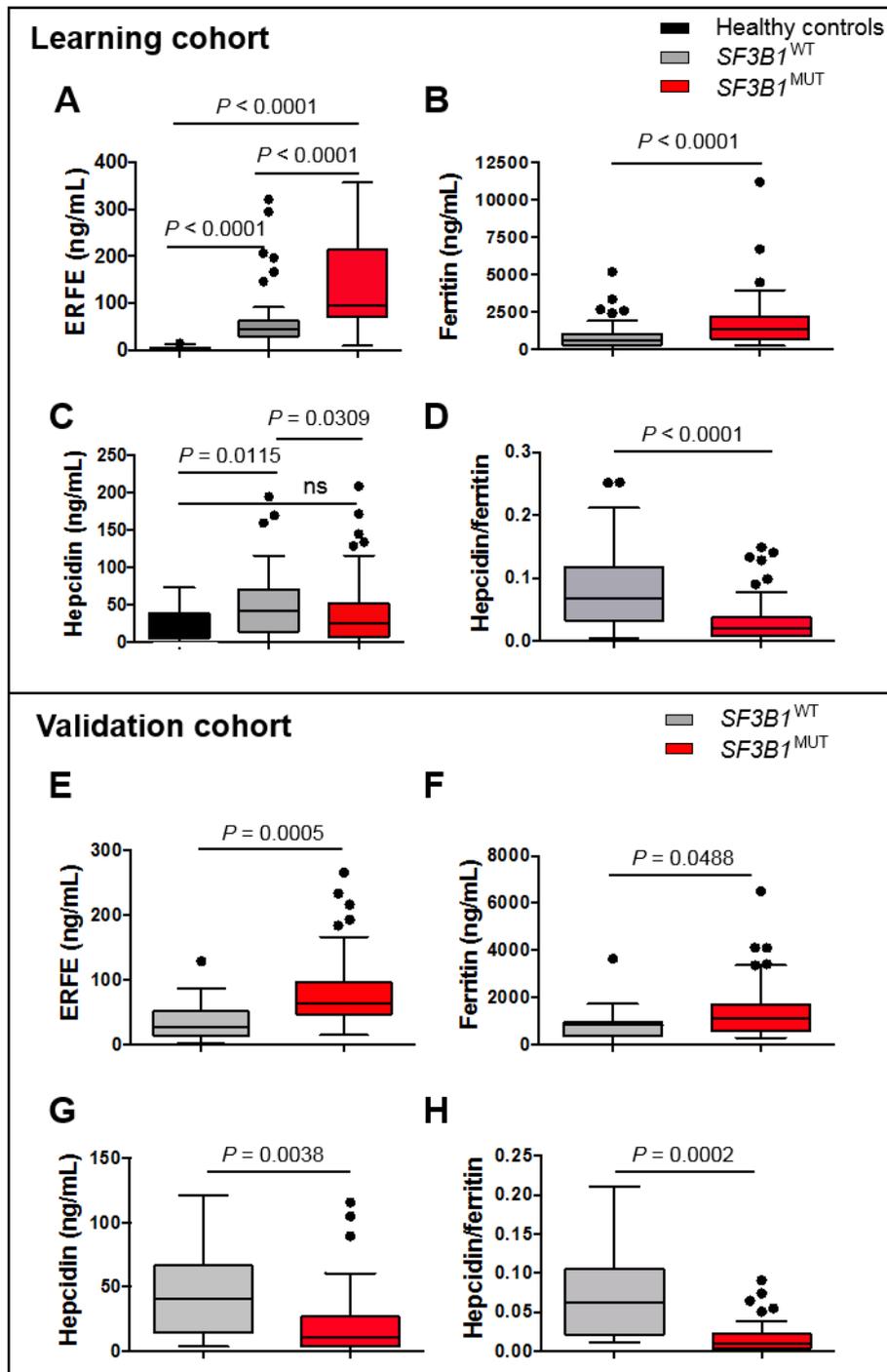


Fig. 5: Increased plasma concentration of ERFE in $SF3BI^{MUT}$ MDS patients. Dosages were performed in plasma collected from 20 non-blood donor healthy volunteers (black boxes), 156 MDS patients including 94 $SF3BI^{MUT}$ (red boxes) and 62 $SF3BI^{WT}$ (gray boxes) representing the learning cohort (A-D) and from 55 MDS patients including 42 $SF3BI^{MUT}$ (red boxes) and 13 $SF3BI^{WT}$ (gray boxes) representing the validation cohort (E-H). Quantification of erythroferrone (A, E), ferritin (B, F), hepcidin (C, G) and hepcidin/ferritin ratio (D, H). Results are expressed as medians \pm interquartile ranges (IQR). The boxplots represent the median and the first and third quartiles and the whiskers represent the lowest value still within the 1.5 IQR of the lower quartile and the highest value still within the 1.5 IQR of the upper quartile. Mann-Whitney for P -values.

Figure 6: Erythroid cell-restricted expression of ERFE⁺¹²

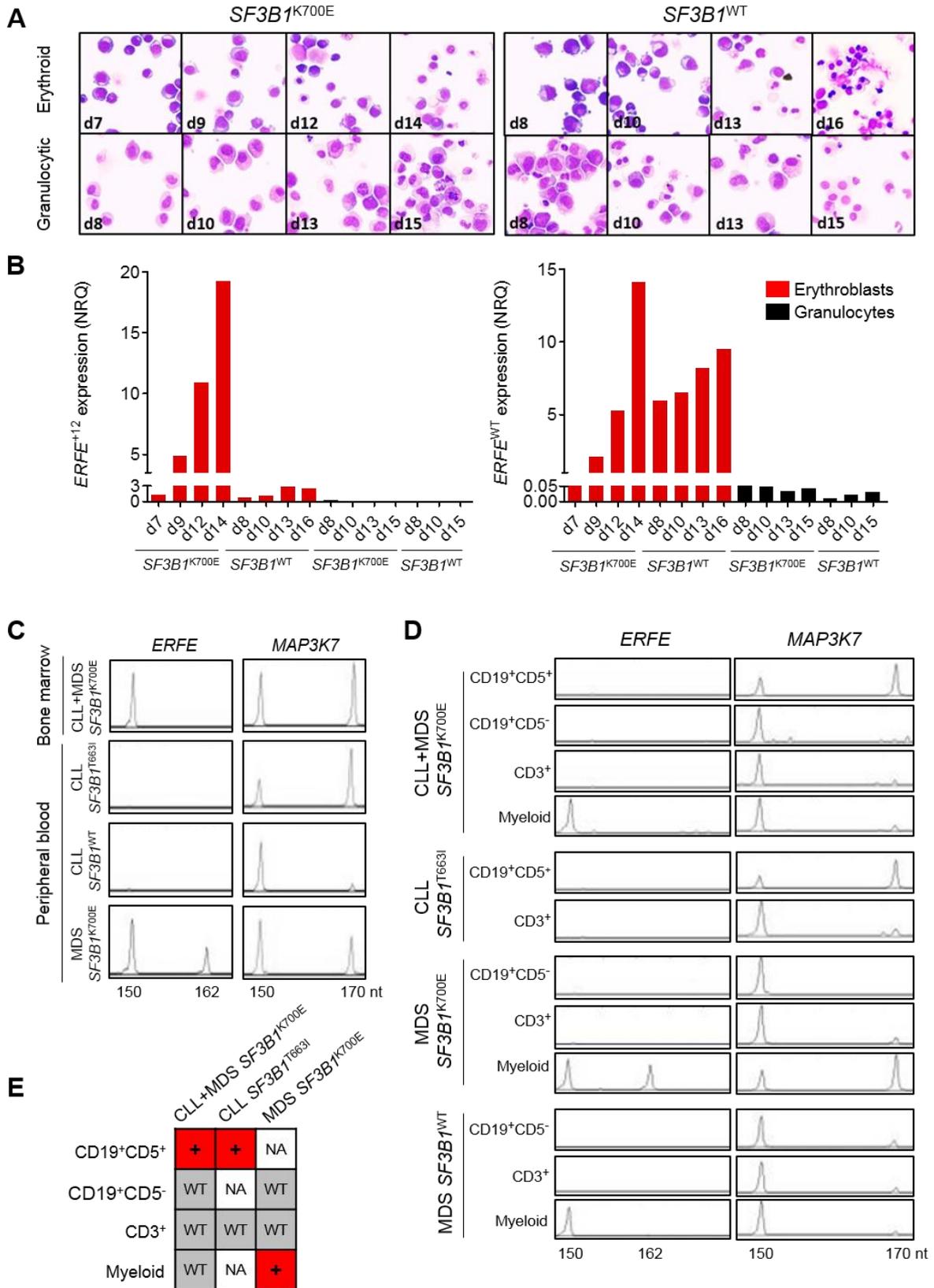


Fig. 6: Erythroid cell-restricted expression of ERFE⁺¹². (A) May-Grünwald-Giemsa-stained cytopins. Erythroid and granulocytic precursors were obtained from the bone marrow (BM) CD34⁺ progenitors of one *SF3B1*^{WT} and one *SF3B1*^{K700E} patient in liquid culture and collected at indicated timepoints. (B) Quantification of ERFE^{WT} and ERFE⁺¹² in erythroid and granulocytic precursors by RT-qPCR. Results are expression as normalized ratio quantities (NRQ) ± standard error of the mean (SEM) to *ACTB* and *B2M* housekeeping genes. (C) ERFE +12 is absent in *SF3B1*^{MUT} chronic lymphocytic leukemia (CLL). Bone marrow (BM) mononuclear cells (MNCs) from a *SF3B1*^{K700E} CLL+MDS or peripheral blood (PB) MNCs from a *SF3B1*^{T663I} CLL, *SF3B1*^{WT} CLL and *SF3B1*^{K700E} MDS were collected for analysis by capillary electrophoresis of ERFE and MAP3K7 fluorescent RT-PCR products. ERFE⁺¹² was detected as a 162-nucleotide (nt) fragment and MAP3K7+20 as a 170-nt fragment. Peripheral blood samples from one patient with an *SF3B1*^{WT} CLL and one patient with an *SF3B1*^{MUT} MDS are shown as controls. (D) ERFE⁺¹² expression is restricted to *SF3B1*^{MUT} myeloid lineage. CD19⁺CD5⁻ B cells, CD19⁺CD5⁺ B CLL cells, CD3⁺ T cells and myeloid cells were sorted from the BM MNCs' fraction of a *SF3B1*^{K700E} CLL+MDS, a *SF3B1*^{K700E} MDS and a *SF3B1*^{WT} MDS, and from the PB MNCs of a *SF3B1*^{T663I} CLL, by ARIA FACS cell sorter (fig. S9). The myeloid cell fraction of *SF3B1*^{T663I} CLL was not analyzable. RNA was extracted from each cell population for the analysis of ERFE, MAP3K7 RT-PCR fluorescent products. (E) *SF3B1* mutational status of each cell fraction. The sequencing of *SF3B1* was performed on the cDNA. +: mutated; WT: wild type; NA: not available.

Figure 7: ERFE⁺¹² expression is a biomarker of clonal erythropoiesis

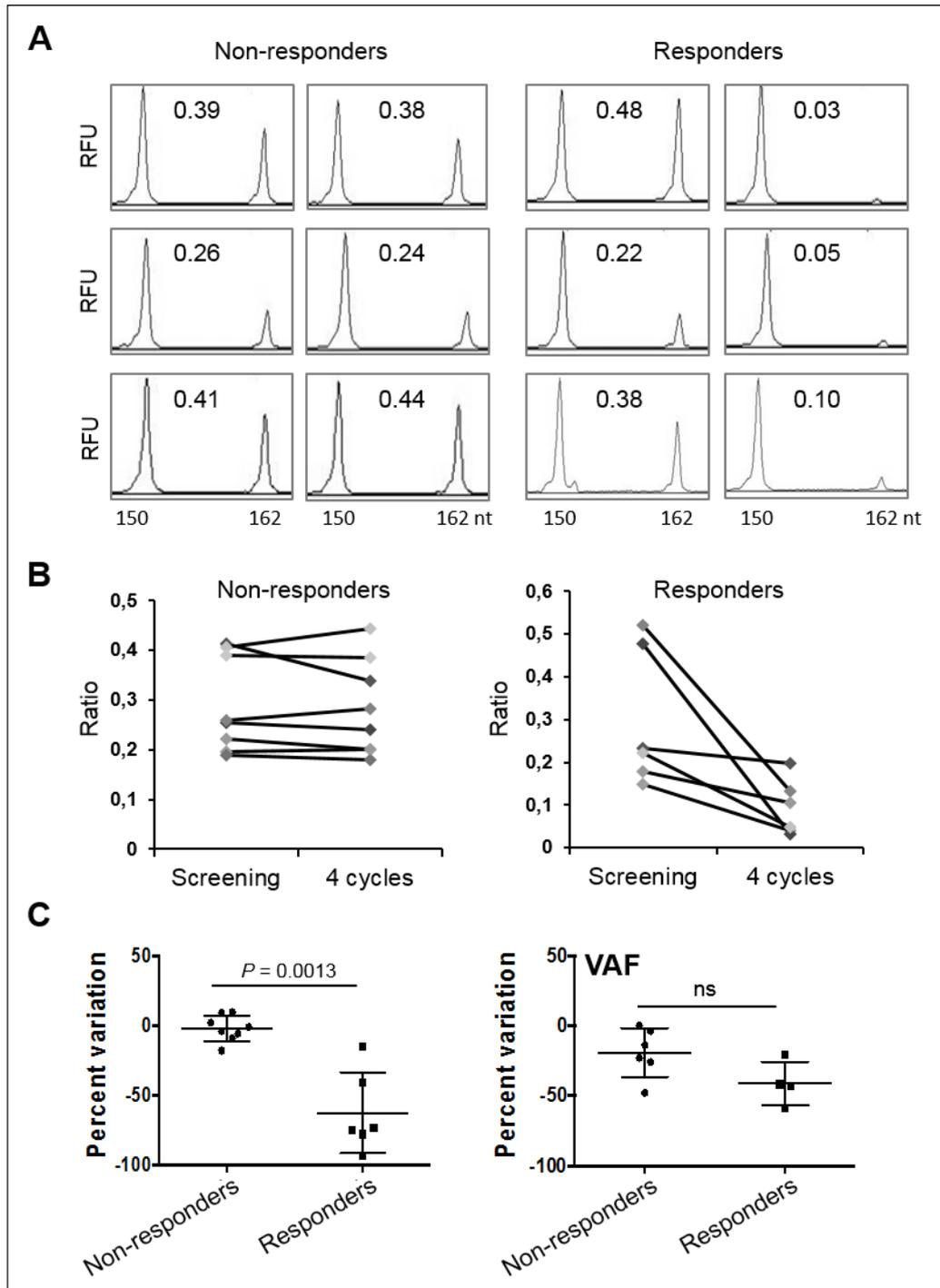


Fig. 7: ERFE⁺¹² expression is a biomarker of clonal erythropoiesis. (A) Fluorescent PCR was performed in 14 paired samples from *SF3BI*^{MUT} MDS patients enrolled in the GFM-LenEpo-08 clinical trial (8 non-responding and 6 responding patients). Three examples of each category are shown. Peak heights of ERFE⁺¹² and ERFE^{WT} signals were integrated as ERFE⁺¹²/ERFE⁺¹²+ERFE^{WT} ratios at screening and evaluation after 4 cycles of treatment. Ratios are indicated. (B) Evolution of the ERFE⁺¹²/ERFE⁺¹²+ERFE^{WT} ratio between screening and evaluation at 4 cycles of treatment in 8 non-responding patients (left) and 6 responding patients (right). (C) Percent variation of ERFE⁺¹²/ERFE⁺¹²+ERFE^{WT} ratios in 8 responding and 6 non-responding patients (left). Percent variation of *SF3BI* variant allele frequency (VAF) available for 6/8 non-responding and 4/6 responding patients is also shown (right). Results are expressed as medians ± interquartile ranges (first and third quartiles). Mann-Whitney test for *P*-values.

Table

Parameters	Ferritin <300 ng/mL	Ferritin ≥300 ng/mL	Univariate	Multivariate
	n=19	n=41	<i>P</i> -value	<i>P</i> -value
Erythroferrone (ng/mL) mean (95% CI)	36.7 (26.8 - 46.6)	72.7 (55.4 - 90.0)	0.005	0.002
Hepcidin (ng/mL) mean (95% CI)	17.2 (11.2 - 23.2)	35.9 (25.9 - 45.9)	0.013	<0.0001
sTfR (ng/mL) mean (95% CI)	1.21 (0.99 - 1.43)	1.42 (1.18 - 1.65)	0.484	
Number of RBC unit per 8 weeks, mean (range)	0.2 (0 - 2)	0.6 (0 - 3)	0.104	
<i>SF3BI</i> mutation yes, n (%)	3 (15.8%)	22 (53.6%)	0.006	0.023

Table 1: *SF3BI* mutation, erythroferrone and hepcidin levels are independent predictors of hyperferritinemia in low transfusion burden MDS patients. Erythroferrone, hepcidin, soluble transferrin receptor (sTfR), number of red blood cell (RBC) unit per 8 weeks and *SF3BI* mutation were considered in the group of 60 patients with less than 4 RBC units per 8 weeks according to ferritin with a cut-off value of 300 ng/mL. Parameters are indicated as means and 95% confidence intervals (95% CI) or range. For univariate analysis, Mann-Whitney and Chi-square tests were used to compare the variables between *SF3BI*^{MUT} and *SF3BI*^{WT} MDS. Multivariate logistic regression analysis was performed using variables with a *P*-value < 0.100 in univariate analysis.

Supplemental Figures and Tables

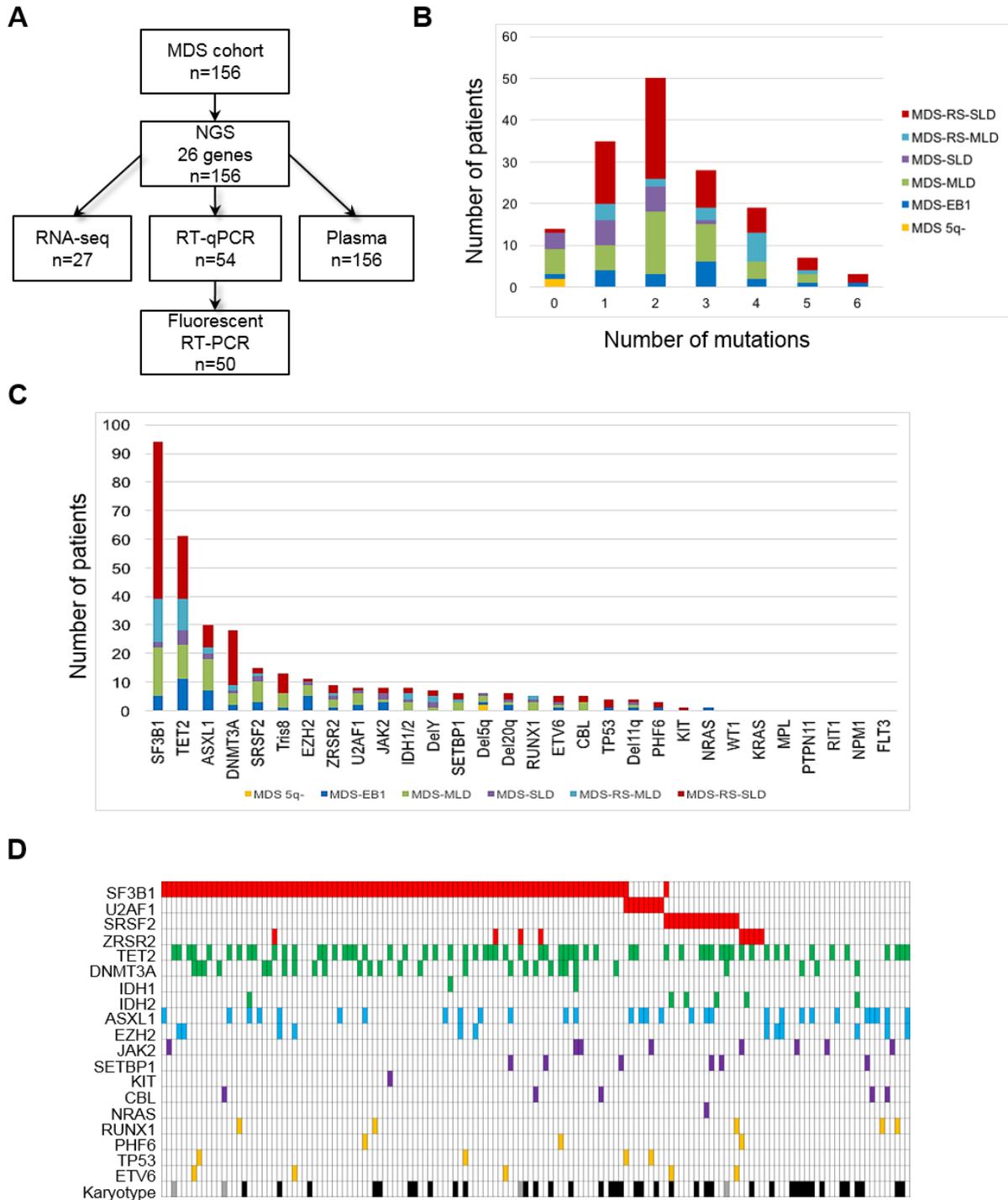
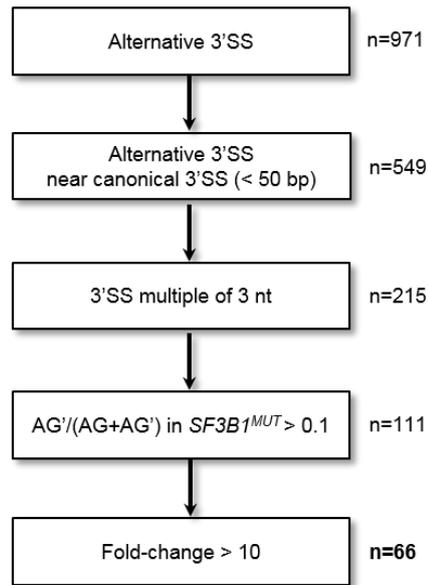


Fig. S1: Genomic characteristics of the studied MDS cohort. (A) Workflow of the study. **(B-D)** Next generation sequencing of 26 recurrently mutated genes in 156 patients. **(B)** Number of patients with 0 to 6 recurrent mutations. **(C)** Number of patients for each chromosomal or genetic abnormality. **(D)** Barcode representing the 149 patients (column) with mutations and/or karyotype abnormalities shown as normal (white boxes), abnormal (black boxes) or not done (gray boxes).

A



B

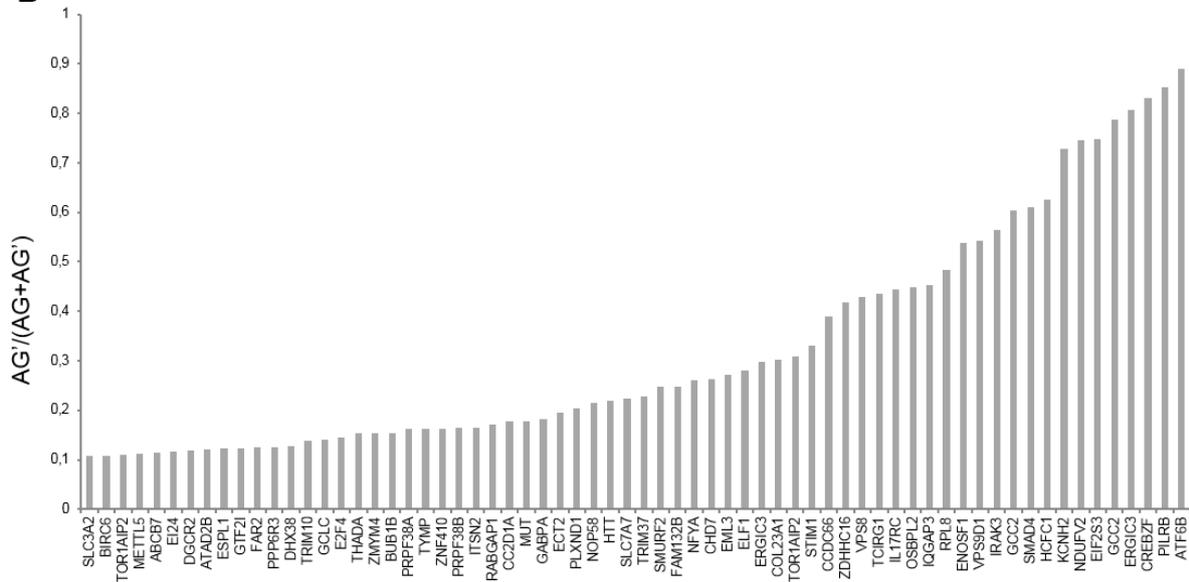


Fig. S2: Identification of in-frame alternative transcripts in *SF3B1*^{MUT} MDS samples. (A) Workflow for the identification of transcripts with 3' aberrant junction and in frame additional coding sequence. The aberrant 3' acceptor and ambiguous junctions, which were present in $\geq 80\%$ of *SF3B1*^{MUT} samples with a P -value $\leq 10^{-5}$ were considered ($n=971$). Among the 549 cryptic 3'ss located near the corresponding canonical 3'ss, 215 had an additional multiple of 3 nucleotides sequence, 111 had an $AG'/(AG+AG')$ ratio > 0.1 and 66 were expressed with a fold-change *ie.* an expression ratio of the alternative junction in *SF3B1*^{MUT} versus *SF3B1*^{WT} samples over 10. **(B)** Percent expression of the 66 aberrant junctions in 63 genes as $AG'/(AG+AG')$ ratio.

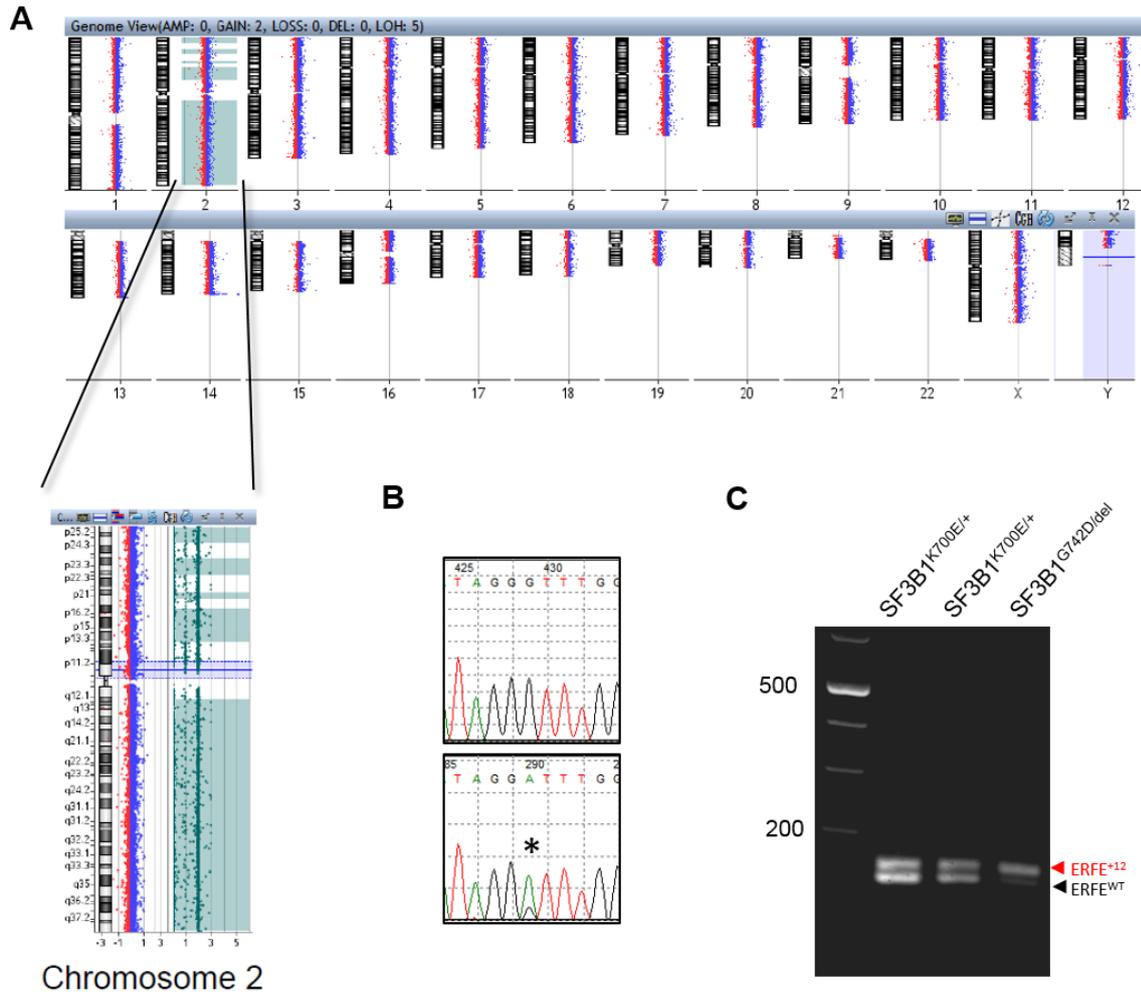


Fig. S3: Massive use of *FAM132B/ERFE* cryptic junction in one MDS-RS with a bi-allelic alteration of *SF3B1* gene. (A) CGH array analysis confirms the deletion (B) Sanger detection of *SF3B1*^{G742D} mutation. (C) *ERFE*^{WT} and *ERFE*⁺¹² RT-PCR products migration on a 3% agarose gel showing the major expression of alternative junction in *SF3B1*^{G742D/del} MDS-RS patient. Two patients with *SF3B1*^{K700E/+} have an equal amount of *ERFE*^{WT} and *ERFE*⁺¹².

A

Human ERFE **FLAG** collagen domain anti-ERFE antibody epitope
 MYRMQLLSICIALSLALVTNSISAMVRS**DYKDDDDK**SP EPPPPGNE LPRGPGES RAGPAARPPPEPTAERAHSV DPRDAW MLFVRQSD
 KGVNGKKRSRGKAKK**LK FGLPGPPGPPGPPGPPG**IPPEALLK EFQLLLK GAVRQRERAEP ECTCGPAGPVAAS LAPVSATAG EDD
 DDVVGDV LALLAAP LAPGPRAPRVEA AFLCRLRRDALVER RALHELGVVYLPDAEGAFRRGPG LNLTSGQYRAP VAGFYALAATLHVA
LGEPPRRGPPRRDHLRLLICIQSRCQRNASLEAIMGLESSSELTISVNGVLYLQMGQWTSVFLDNASGCSLTVRS GSHFS AVLLGV

Human ERFE^{VPFQ} **FLAG** tetrapeptide collagen domain anti-ERFE antibody epitope
 MYRMQLLSICIALSLALVTNSISAMVRS**DYKDDDDK**SP EPPPPGNE LPRGPGES RAGPAARPPPEPTAERAHSV DPRDAW MLFVRQSD
 KGVNGKKRSRGKAKK**LK VPFQ FGLPGPPGPPGPPGPPG**IPPEALLKEFQLLLK GAVRQRERAEP ECTCGPAGPVAAS LAPVSATAG
 EDDDDVWG DVLLAAP LAPGPRAPRVEA FLCRLRRDALVERRALHELGVVYLPDAEGAFRRGPG LNLTSGQYRAP VAGFYALAATL
 HVA**LGEPPRRGPPRRDHLR**LLICIQSRCQRNASLEAIMGLESSSELTISVNGVLYLQMGQWTSVFLDNASGCSLTVRS GSHFS AVLL
 GV

B

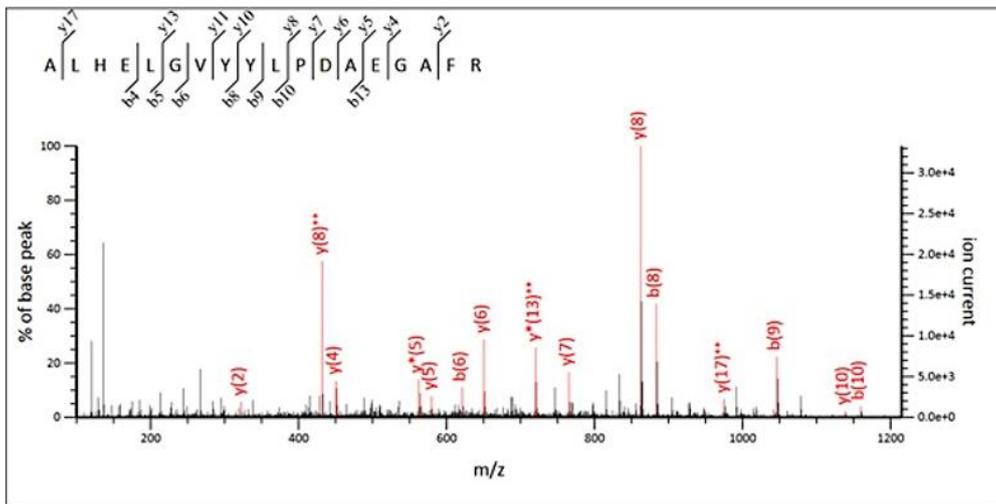


Fig. S4: Identification of ERFE peptide by LC tandem mass spectrometry (A) Peptidic sequences of human ERFE^{WT} and ERFE^{VPFQ}: FLAG (green), VPFQ (magenta), collagen domain (blue), anti-ERFE antibody epitope (yellow). **(B)** LC MS/MS identification. Peptides were prepared from *SF3B1*^{MUT} erythroblast lysates and analyzed by nano liquid chromatography hyphenated to a Q-Exacte Plus mass spectrometer. MS data were analyzed using Mascot 2.5.1 for confident identification of the peptide ALHELGVVYLPDAEGAFR for aminoacid residues 224 to 241 found in the Uniprot database (accession number: Q4G0M1, ERFE_HUMAN). The calculated peptide mass was 2020.0127 from 3+ ion observed at M/Z 674.344850, with measured delta = 3.62 ppm, Mascot Score = 39, *P*-value = 0.00024. This FAM132B/ERFE peptide is shown at www.proteomicsdb.org/proteomicsdb/#human/proteinDetails/62110/peptides.

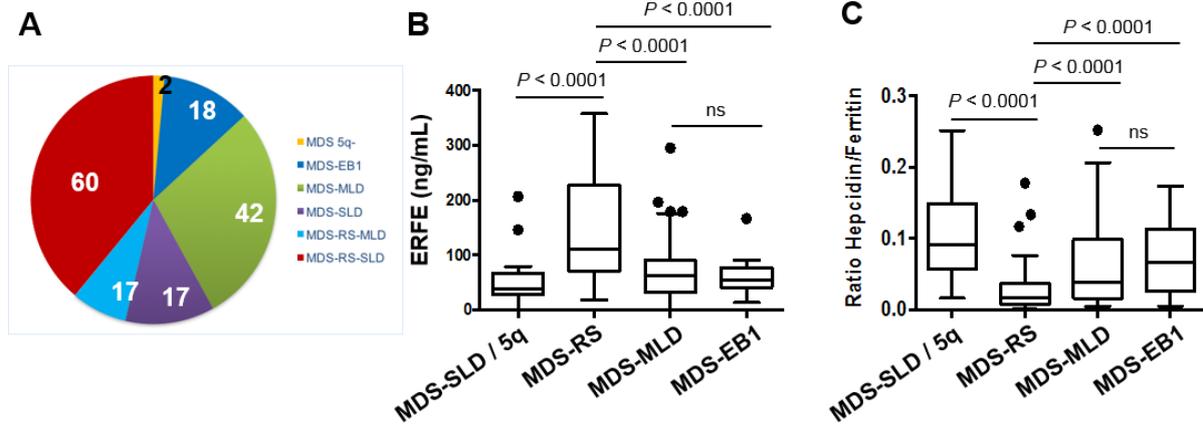


Fig. S5: Erythroferrone and hepcidin/ferritin ratio according to the WHO classification 2016 of the learning cohort. (A) Repartition of the 156 MDS samples. **(B).** Quantification of plasma erythroferrone in different MDS subtypes. 5q- and MDS with single lineage dysplasia (SLD) were pooled. MDS-MLD: MDS with multilineage dysplasia. MDS-EB1: MDS with excess of blasts 1. MDS-RS term includes MDS-SLD-RS and MDS-MLD-RS. **(C)** Ratio of hepcidin to ferritin expressed in the same categories. Box plots represent the median and the first and third quartiles, and the whiskers represent the lowest value still within the 1.5 interquartile range (IQR) of the lower quartile and the highest value still within the 1.5 IQR of the upper quartile. Mann-Whitney test for *P*-values. ns: not significant.

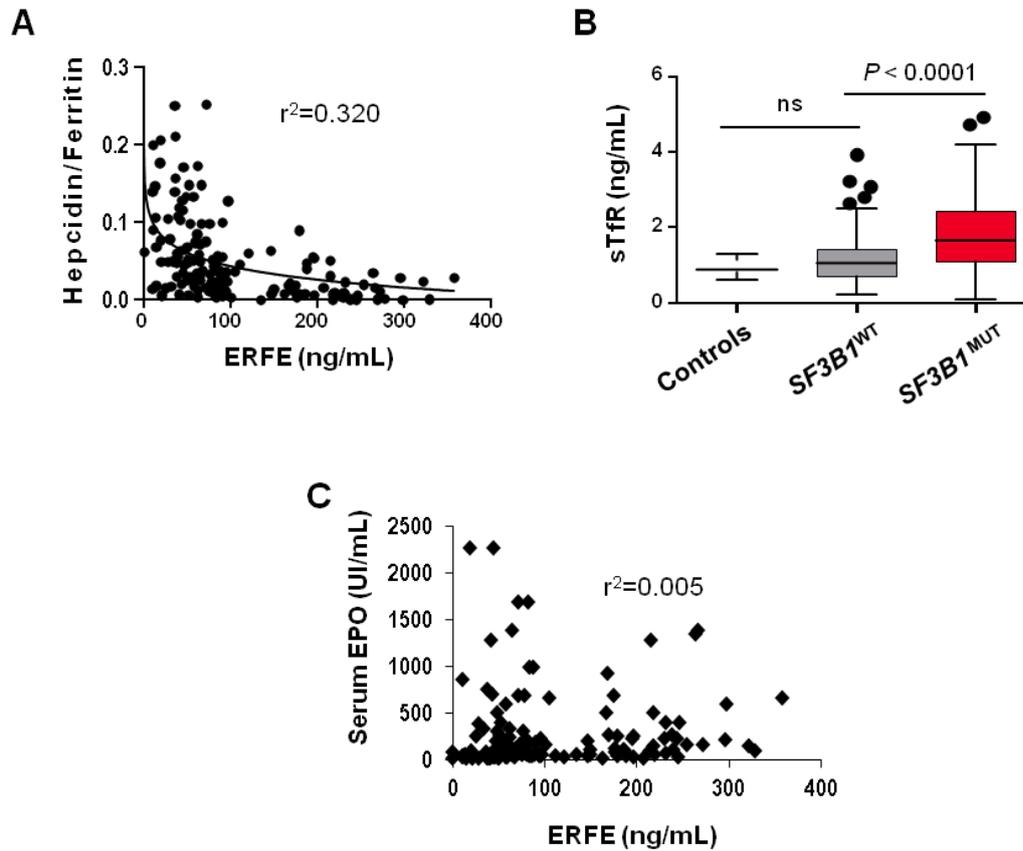


Fig. S6: Biological parameters of 156 MDS patients of the learning cohort. (A) Inverse relationship between ERFE and hepcidin/ferritin ratio. Pearson test for correlation; $P < 0.0001$. **(B)** Comparison of soluble transferrin receptor (sTfR) level between 94 $SF3B1^{MUT}$, 61 $SF3B1^{WT}$ and 20 healthy non-blood donor controls. Results are expressed as medians \pm IQR. Mann-Whitney test for P -values. ns: not significant. **(C)** Serum EPO level according to the concentration of ERFE. Pearson test for correlation; $P=0.403$.

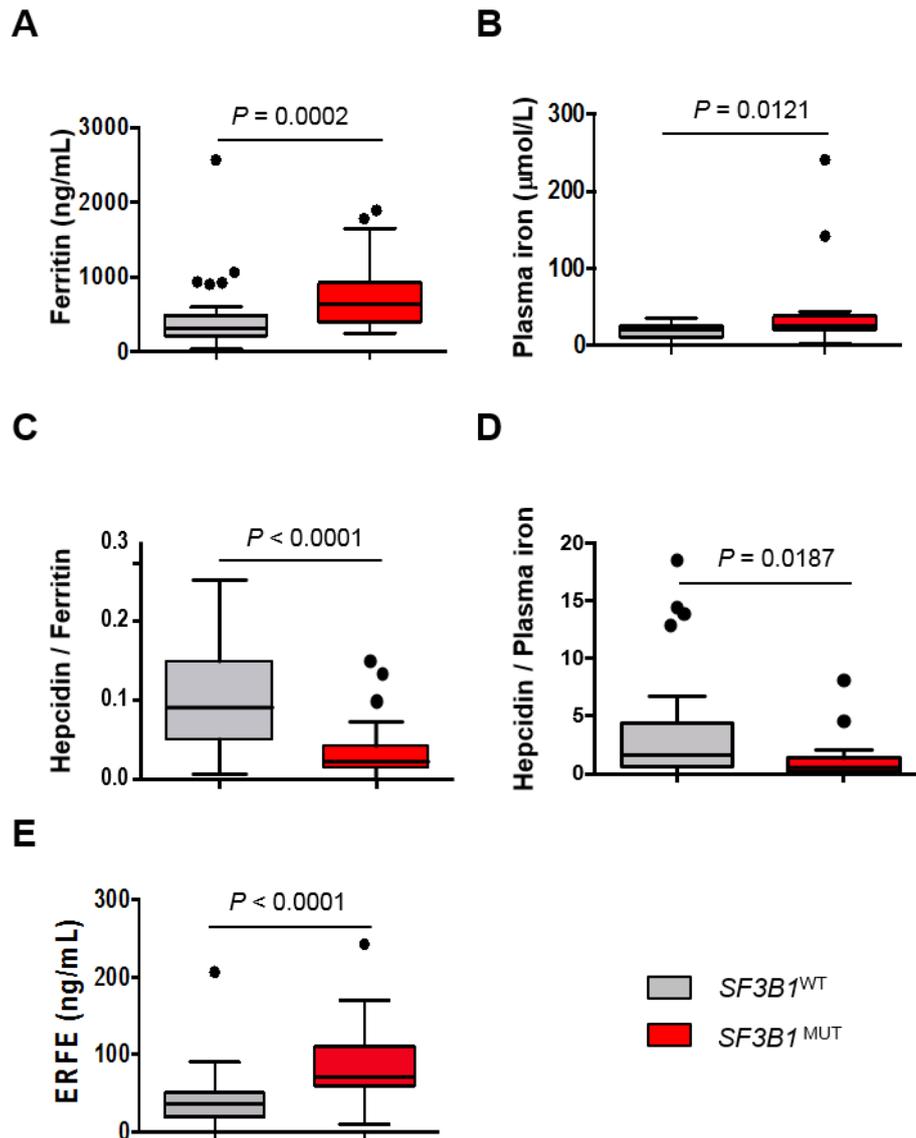


Fig. S7: Biological parameters of low transfusion burden MDS patients of the learning cohort. Comparison of ferritin (A), plasma iron (B), hepcidin/ferritin ratio (C), hepcidin/plasma iron ratio (D) and ERFE (E) between 25 $SF3B1^{MUT}$ and 36 $SF3B1^{WT}$ patients. Results are expressed as medians \pm IQR. Mann-Whitney test for P -values.

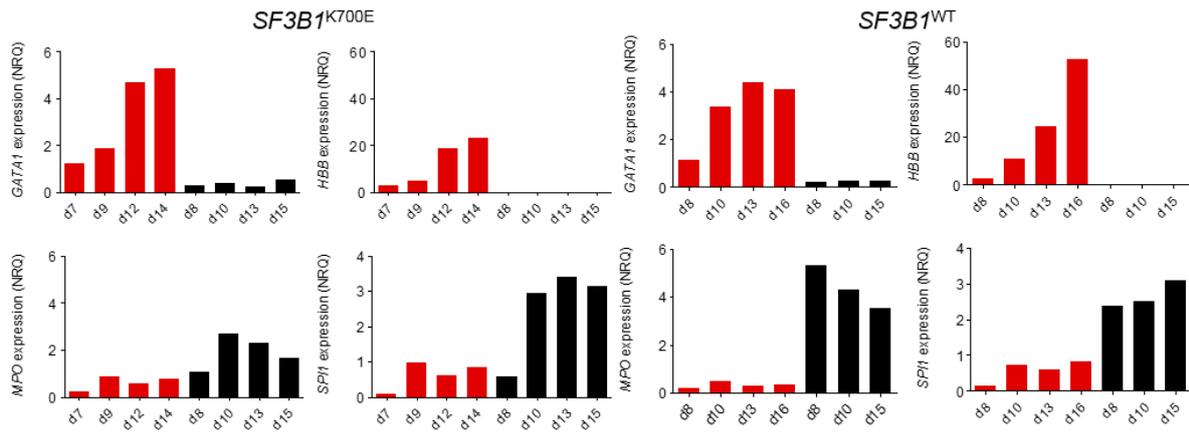


Fig. S8: Erythroid lineage-restricted expression of *ERFE*⁺¹². Quantification of lineage-restricted markers, *GATA1* and *HBB* for erythroid differentiation and *MPO* and *SPI1* for the granulocytic differentiation by RT-qPCR. Red bars for erythroid cells and black bars for granulocytes at indicated days of culture post-CD34⁺ cell sorting.

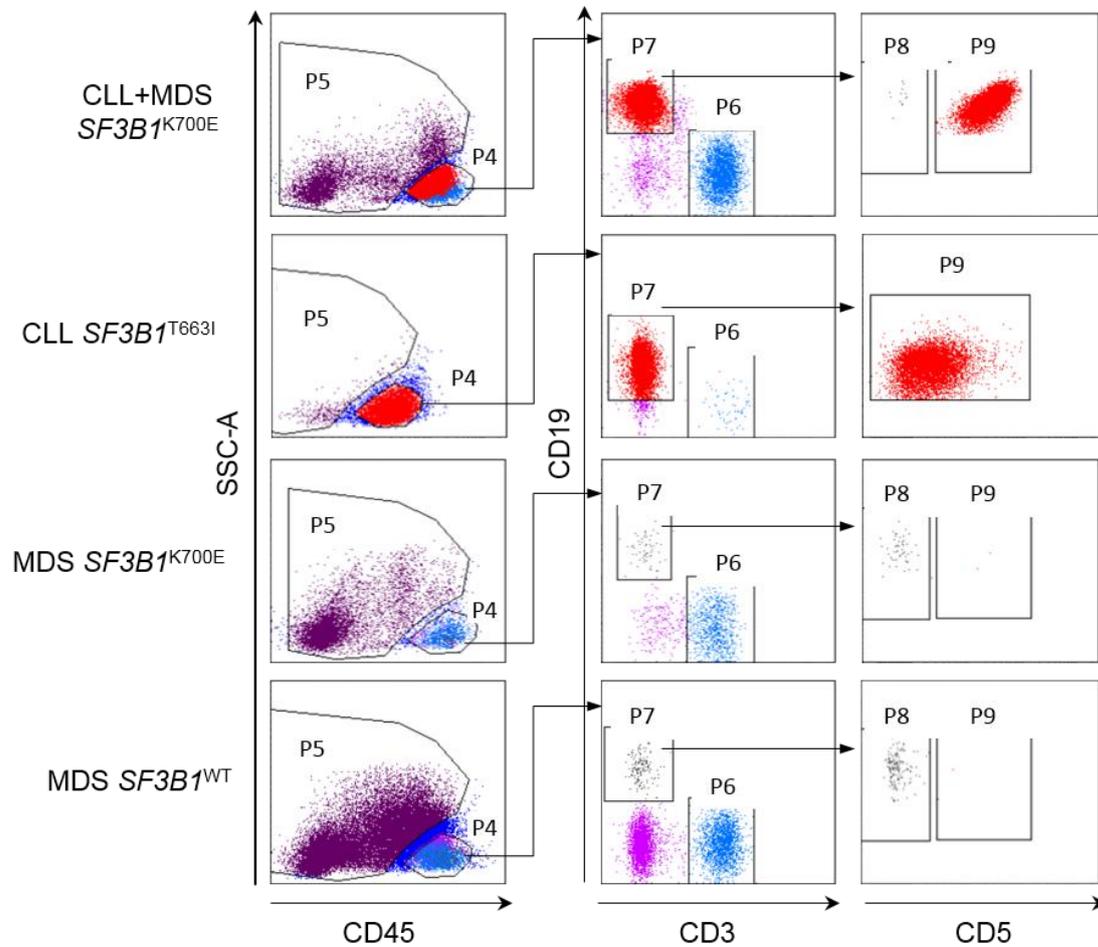


Fig. S9: Gating strategy for myeloid and lymphoid cell sorting. CD19⁺CD5⁻ B cells, CD19⁺CD5⁺ B CLL cells, CD3⁺ T cells and myeloid cells were sorted from the BM MNCs' fraction of a *SF3B1^{K700E}* CLL+MDS, a *SF3B1^{K700E}* MDS and a *SF3B1^{WT}* MDS, and from the PB MNCs of a *SF3B1^{T663I}* CLL, by ARIA FACS cell sorter. Lymphoid cells are gated in P4, myeloid cells are shown in purple (P5), CD45^{high}CD19⁻CD3⁺ T cells are shown in blue (P6), CD45^{high}CD19⁺CD3⁻CD5⁻ healthy B cells are shown in gray (P8), CD45^{high}CD19⁺CD3⁻CD5⁺ pathological B cells are shown in red (P9).

Table S1: Clinical and biological characteristics of the learning cohort of 156 MDS patients according to *SF3B1* status

Parameters	Learning cohort			P-value
	All	<i>SF3B1</i> ^{WT}	<i>SF3B1</i> ^{MUT}	
	n=156	n=62	n=94	
Age				
Median, years (range)	74 (46 - 91)	74 (54 - 88)	74 (46 - 91)	0.371
Sex				
male, no. (%)	92 (59)	38 (61)	54 (57)	0.740
WHO-Subtype 2016				
5q/MDS-SLD no. (%)	19 (12.2)	17 (27.4)	2 (2.1)	<0.0001
MDS-SLD-RS/MDS-MLD-RS, no. (%)	77 (49.4)	7 (11.3)	70 (74.5)	
MDS-MLD, no. (%)	42 (26.9)	25 (40.3)	17 (18.1)	
MDS-EB1, no. (%)	18 (11.5)	13 (21.0)	5 (5.3)	
Hemoglobin (G/L)				
Median (IQR)	8.6 (7.8-9.4)	9.2 (8.3-9.7)	8.2 (7.7-9.0)	<0.001
Neutrophil count (G/L)				
Median (IQR)	2.6 (1.6-3.7)	2.2 (1.3-3.7)	2.8 (1.6-3.8)	0.502
Platelet count (G/L)				
Median (IQR)	234 (164- 325)	172 (103-308)	255 (195-329)	0.008
Karyotype IPSS				
Good, no. (%)	125 (80.2)	48 (77.4)	77 (81.9)	0.133
Intermediate, no. (%)	23 (14.7)	10 (16.1)	13 (13.8)	
Poor, no. (%)	3 (1.9)	3 (4.9)	0 (0.0)	
Unavailable, no. (%)	5 (3.2)	1 (1.6)	4 (4.3)	
Karyotype IPSS-R				
Very good, no. (%)	7 (4.5)	4 (6.4)	3 (3.2)	0.270
Good, no. (%)	122 (78.2)	46 (74.2)	76 (80.9)	
Intermediate, no. (%)	20 (12.8)	9 (14.5)	11 (11.7)	
Poor, no. (%)	2 (1.3)	2 (3.2)	0 (0.0)	
Unavailable, no. (%)	5 (3.2)	1 (1.6)	4 (4.2)	
BM blasts (%)				
Median (IQR)	2.0 (1.0-3.5)	2.0 (1.0-4.0)	2.0 (1.0-3.0)	0.267
BM erythroblasts (%)				
Median (IQR)	25 (11-38)	15 (3-27)	32 (19-45)	<0.0001
IPSS				
Low risk, no. (%)	85 (54.5)	27 (43.6)	58 (61.7)	0.082
Int-1, no. (%)	64 (41.0)	32 (51.6)	32 (34.0)	
Int-2, no. (%)	1 (0.6)	1 (1.6)	0 (0.0)	
Unavailable, no. (%)	6 (3.9)	2 (3.2)	4 (4.3)	
IPSS-R				
Very low risk, no. (%)	15 (9.6)	9 (14.5)	6 (6.4)	0.107
Low, no. (%)	95 (60.9)	31 (50.0)	64 (68.1)	
Int, no. (%)	36 (23.1)	17 (27.4)	19 (20.2)	
High, no. (%)	4 (2.6)	3 (4.9)	1 (1.1)	
Unavailable, no. (%)	6 (3.8)	2 (3.2)	4 (4.2)	
Serum EPO level (U/L)				
Median (IQR)	99 (39-249)	84 (30-200)	101 (47-312)	0.657
Red blood cells unit, no				
Median (range)	4 (0-17)	2 (0-15)	5.5 (0-17)	<0.0001

Table S4: In-frame and differentially expressed 3' cryptic splice site junctions in *SF3BI*^{MUT} MDS

Gene identification		Gene expression			Alternative 3'SS			
ENSG ID	Gene symbol	p-val	Fold-change MUT vs WT	Log ₂ FC	AG'/(AG+AG')	Position 1	Position 2	Distance
ENSG000A2:C35001312	ABC7	0,234478098	0,836623791	-0,257349071	0,113226978	74291539	74293523	-21
ENSG00000119778	ATAD2B	0,967797591	0,995353649	-0,006718888	0,121236091	23977668	23980287	-24
ENSG00000168468	ATF6B	1	0,988766869	-0,016297692	0,88905072	32095539	32095893	-36
ENSG00000115760	BIRC6	0,030984221	0,735761363	-0,442690176	0,107816237	32614011	32617091	-15
ENSG00000156970	BUB1B	0,00565255	1,408008762	0,493656312	0,153490504	40493181	40494590	-15
ENSG00000132024	CC2D1A	0,021628922	1,315797877	0,395937891	0,176804545	14031735	14034130	-15
ENSG00000180376	CCDC66	0,015911994	1,377690965	0,462252308	0,389353302	56649300	56649931	-18
ENSG00000171316	CHD7	0,272340048	0,778221834	-0,361746637	0,261798276	61741365	61742868	-12
ENSG00000050767	COL23A1	0,949323505	1,016797085	0,0240318	0,30257093	177684583	177686721	-15
ENSG00000137504	CREBZF	0,07078856	1,430291248	-0,435199943	0,831568341	85374321	85374567	-15
ENSG00000070413	DGCR2	0,490313191	0,883826921	-0,17816422	0,117757003	19044699	19050714	-24
ENSG00000140829	DHX38	0,221776299	0,874697601	-0,193143758	0,127567984	72139523	72139882	-21
ENSG00000205250	E2F4	6,10E-07	1,967492092	0,976357837	0,144054606	67226773	67226963	-6
ENSG00000114346	ECT2	0,048336137	1,402074466	0,487562975	0,195410425	172479473	172480187	-18
ENSG00000149547	EI24	0,001810133	1,537129098	0,620238337	0,115338565	125442465	125445146	-12
ENSG00000130741	EIF2S3	0,324891032	1,088267388	0,122033072	0,747119175	24091380	24092454	-12
ENSG00000120690	ELF1	0,000212259	0,521015912	-0,940600662	0,2802663	41524061	41525464	-24
ENSG00000149499	EML3	0,003970792	0,660393885	-0,598601333	0,271507929	62376298	62376433	-21
ENSG00000132199	ENOSF1	0,003106653	1,520822096	0,604851398	0,537277657	683395	685920	-15
ENSG00000125991	ERIC3	0,698692334	1,040872987	0,057794034	0,298348149	34144042	34144725	-18
ENSG00000125991	ERIC3	0,698692334	1,040872987	0,057794034	0,806064772	34144519	34144725	-18
ENSG00000135476	ESPL1	0,05694348	1,36522197	0,449135536	0,122495295	53666641	53668573	-27
ENSG00000178752	FAM132B/ERFE	0,000811123	2,353774102	1,234975868	0,248374407	239070357	239071364	-12
ENSG00000064763	FAR2	0,430783166	0,782277897	-0,354246892	0,125531568	29450133	29460566	-24
ENSG00000154727	GABPA	0,52879223	0,913676045	-0,130245364	0,182604015	27107889	27113871	-12
ENSG00000135968	GCC2	0,688016585	0,923100258	-0,115440747	0,603623023	109102323	109102954	-12
ENSG00000135968	GCC2	0,688016585	0,923100258	-0,115440747	0,786999338	109102364	109102954	-12
ENSG00000001084	GCLC	0,001515057	1,805783753	0,852625137	0,140561765	53381038	53385575	-18
ENSG00000077809	GTF2I	0,341817864	1,2566444	0,32957646	0,122600817	74131270	74133179	-18
ENSG00000172534	HCFC1	0,042692379	1,247112795	0,318591956	0,625171398	153229750	153230028	-15
ENSG00000197386	HTT	0,103890796	0,811010018	-0,302208359	0,219075289	3142383	3144471	-21
ENSG00000163702	IL17RC	0,454467796	1,096646119	0,133098051	0,445227199	9960293	9962150	-24
ENSG00000183856	IQGAP3	0,398796864	1,102148373	0,140318455	0,453337207	156532466	156532925	-9
ENSG00000090376	IRAK3	0,064636315	0,64053351	-0,642654046	0,564852165	66597673	66601997	-18
ENSG00000198399	ITSN2	0,725932741	0,840226068	-0,251150548	0,165175455	24538114	24550920	-21
ENSG00000055118	KCNH2	1,93E-09	3,594114688	1,845636446	0,728594188	150643694	150643964	-9
ENSG00000138382	METTL5	0,060137138	1,281766904	0,334133924	0,111245037	170669034	170671986	-18
ENSG00000146085	MUT	0,280100924	1,158399659	0,212133082	0,178171448	49416664	49419178	-24
ENSG00000178127	NDUFV2	0,032841655	2,196660036	1,13531161	0,745212261	9133520	9136361	-12
ENSG00000001167	NFYA	0,511849577	1,098690416	0,135784927	0,261463522	41040823	41046743	-24
ENSG00000055044	NO58	0,048222426	1,260751177	0,334284251	0,214724772	203157626	203162080	-21
ENSG00000130703	OSBPL2	0,499497702	1,115985749	0,158318605	0,447927882	60859225	60860573	-18
ENSG00000121716	PILLR	0,069778268	0,663982662	-0,590782524	0,852938571	99954506	99955849	-15
ENSG00000004399	PLXND1	0,001427806	0,596023064	-0,746559935	0,203894971	129284872	129285369	-12
ENSG00000110075	PPP6R3	0,177765328	1,150162954	0,201838276	0,125722364	68331900	68334466	-15
ENSG00000134748	PRPF38A	0,743287097	0,970043911	-0,04387804	0,162776992	52880319	52880412	-21
ENSG00000134186	PRPF38B	0,290638706	0,865667533	-0,208115043	0,164828521	109235489	109236193	-18
ENSG00000011454	RABGAP1	0,263522987	0,897698386	-0,155697293	0,170615188	125759640	125760854	-21
ENSG00000161016	RPL8	0,008191813	1,36849251	0,452587538	0,484222007	146015873	146016661	-33
ENSG00000168003	SLC3A2	0,252217217	1,181715093	0,240882249	0,107457739	62648919	62649352	-12
ENSG00000155465	SLC7A7	0,004913795	0,451239266	-1,148035482	0,22391134	23242937	23243141	-12
ENSG00000141646	SMAD4	0,046292984	0,782263923	-0,354272664	0,609782472	48575694	48577707	-39
ENSG00000262909	SMURF2	1	0,735129317	-0,443930038	0,246815781	62574712	62576906	-18
ENSG00000167323	STIM1	0,886811502	1,019834908	0,028335626	0,330225534	4104212	4104471	-21
ENSG00000110719	TCIRG1	0,004471249	0,559359677	-0,838151837	0,434836836	67811120	67811222	-42
ENSG00000115970	THADA	0,905368114	0,982188947	-0,025927507	0,152717484	43819570	43823069	-15
ENSG00000169905	TOR1AIP2	0,14536373	1,342453412	0,424872023	0,108933847	179835004	179846373	-15
ENSG00000169905	TOR1AIP2	0,14536373	1,342453412	0,424872023	0,308152276	179835004	179846827	-15
ENSG00000204613	TRIM10	0,00020475	2,340318095	1,226704634	0,138147311	30124849	30125145	-18
ENSG00000108395	TRIM37	0,036769414	1,332718	0,414371541	0,227284463	57148329	57153007	-21
ENSG00000025708	TYMP	0,577207539	0,875872087	-0,191207902	0,162926621	50966161	50966940	-15
ENSG00000156931	VPS8	0,039439599	0,688611042	-0,538238779	0,428030267	184566983	184567705	-18
ENSG00000075399	VPS9D1	0,000596535	0,517219377	-0,951151769	0,543024896	89778630	89779081	-15
ENSG00000171307	ZDHHC16	0,969685981	1,008091021	0,011625907	0,41766861	99214556	99215395	-21
ENSG00000146463	ZMYM4	0,012209163	0,727176177	-0,459623158	0,152817946	35871069	35873587	-21
ENSG00000119725	ZNF410	0,78357134	1,064306999	0,089914355	0,163196578	74358911	74360478	-21

Table S5: Primer sequences used for fluorescent RT-PCR and non-fluorescent RT-qPCR

List of primers (5'->3')	Forward primer (5'->3')	Reverse primer (5'->3')
qPCR primers		
<i>FAM132B/ERFE</i> ^{WT} isoform	ACGCCTGGATGCTCTTCGTC	CCTGGCAA GCCGAACTT
<i>FAM132B/ERFE</i> ⁺¹² isoform	ACGCCTGGATGCTCTTCGTC	GGCAA GCCGAACTGAAAGGG
<i>ACT</i>	GGGTCA GAAGGATTCCTATG	GGTCTCAAACATGATCTGGG
<i>B2M</i>	CCA CTGAAAAAGATGAGTATGCCT	CAAACCTCCA TGATGCTGCTTACA
<i>EPOR</i>	CCCCAAGTTGAGAGCAAAG	GGTAGGAGAA GCTGTAGTTGC
<i>GATA1</i>	AGACGACCA CCA CGACACTG	AAGCCACCA GCTGGTCCTTC
<i>HAMP</i>	CCAGCTGGATGCCATGTT	GCCGCAGCAGAAAATGCA
<i>HPRT</i>	TGCTTTCCTGGTCAGGCAG	AAGCTTGCACCTTGACCAT
<i>HBB</i>	GGCCCA CAAGTATCACTAAGC	AGCCCTTCA TAATATCCC
<i>MPO</i>	CATGTCCA ACTCATATCCCGG	AGCCAGATATACCCCTCAC
<i>SPI1</i>	CGTGCAAAA TGGAAGGGTTTC	GTAATGGTCGCTATGGCTCTC
Fragment analysis		
<i>FAM132B/ERFE</i>	AGGGTGCAATGGCAAGAAG	FAM-TTTCAGCAGCAGCTGGA ACT
<i>MAP3K7</i>	GTGATGGAAATGCTGAAGGGG	FAM-AATTAGCGCTTTGGGTTGCA
Minigene primers		
<i>ERFE</i> In-Fusion	CGGGCCCCCCTCGAGGCTCTGTTGCCTGCCTCTTA	TGGCGGCGCTCTAGACAGCTGGAACTCCTTCA GCA
<i>ERFE</i> minigene sequence checking	GCTAGGCCCAAGGATAGGTA	AATGGGTACCGAGCTCCAC
	GCTAGGCCCAAGGATAGGTA	CCGCTCTAGACAGCTGGAAC
<i>ERFE</i> minigene	FAM-GATTCTTCTACACCCCATGTCC	CAGCTGGAACTCCTTCA GCA
<i>ENOSF1</i> minigene	FAM-GAGGGA TCCGCTTCCTGCCCC	TAGAACTAGTGGATCAGCTTGGACATCCACTCCAC
House keeping gene		
<i>ABL</i>	ACCTCCAGGAGAGCTGCAGAG	AGGAA GCGCTTGCTGGAGCTG
<i>SF3B1</i> status on c.DNA	CTCGTGGTCA TTGAACCGCT	TGGCTGCTTCTTGAAGGCT

Table S6: Clinical and biological characteristics of the validation cohort of 55 MDS patients according to *SF3B1* status

Parameters	Validation cohort			P-value
	All	<i>SF3B1</i> ^{WT}	<i>SF3B1</i> ^{MUT}	
	n=55	n=13	n=42	
Age				
Median, years (range)	74 (28 - 94)	76 (28 - 94)	74 (30 - 90)	0.374
Sex				
male, no. (%)	31 (56)	8 (62)	23 (55)	0.750
WHO-Subtype 2016				
5q/MDS-SLD, no. (%)	4 (7.3)	4 (30.8)	0 (0.0)	<0.0001
MDS-SLD-RS/MDS-MLD-RS/MDSMPN-RS-T no. (%)	41 (74.5)	4 (30.8)	37 (88.1)	
MDS-MLD, no. (%)	6 (10.9)	3 (23.1)	3 (7.1)	
MDS-EB1, no. (%)	4 (7.3)	2 (15.3)	2 (4.8)	
Hemoglobin (G/L)				
Median (IQR)	8.0 (6.2 - 8.8)	7.8 (5.5 - 8.2)	8.1 (7.3 - 9.0)	0.097
Neutrophil count (G/L)				
Median (IQR)	3.0 (1.8 - 3.8)	3.3 (2.1 - 3.9)	2.7 (1.6 - 3.8)	0.435
Platelet count (G/L)				
Median (IQR)	293 (222 - 388)	232 (113 - 399)	298 (227 - 381)	0.561
Karyotype IPSS				
Good, no. (%)	46 (83.6)	10 (76.9)	36 (85.7)	0.031
Intermediate, no. (%)	5 (9.2)	0 (0)	5 (11.9)	
Poor, no. (%)	2 (3.6)	2 (15.4)	0 (0)	
Unavailable, no. (%)	2 (3.6)	1 (7.7)	1 (2.4)	
Karyotype IPSS-R				
Very good, no. (%)	2 (3.6)	1 (7.7)	1 (2.4)	0.535
Good, no. (%)	45 (81.8)	9 (69.2)	36 (85.7)	
Intermediate, no. (%)	6 (11.0)	2 (15.4)	4 (9.5)	
Unavailable, no. (%)	2 (3.6)	1 (7.7)	1 (2.4)	
BM blasts (%)				
Median (IQR)	2.0 (1.0 - 3.0)	3.0 (1.8 - 4.0)	2.0 (1.0 - 3.0)	0.070
BM erythroblasts (%)				
Median (IQR)	34 (26 - 42)	36 (16 - 42)	34 (27 - 44)	0.788
IPSS				
Low risk, no. (%)	37 (67.3)	7 (53.8)	30 (71.4)	0.419
Int-1, no. (%)	16 (29.1)	5 (38.5)	11 (26.2)	
Unavailable, no. (%)	2 (3.6)	1 (7.7)	1 (2.4)	
IPSS-R				
Very low risk, no. (%)	8 (14.6)	1 (7.7)	7 (16.7)	0.106
Low, no. (%)	32 (58.2)	5 (38.5)	27 (64.2)	
Int, no. (%)	13 (23.6)	6 (46.1)	7 (16.7)	
Unavailable, no. (%)	2 (3.6)	1 (7.7)	1 (2.4)	
Red blood cells unit, no				
Median (range)	4 (0 - 9)	4 (0 - 8)	4 (0 - 9)	0.879

Partie 2

I. Problématique

Les travaux actuels se sont concentrés sur l'association entre les mutations des gènes *U2AF1* ou *SRSF2* et la réplication de l'ADN. Même si le gène *SF3B1* code aussi pour un facteur d'épissage, on ne peut pas extrapoler les défauts de réplication et de réparation de l'ADN aux cellules *SF3B1*^{MUT}, surtout au regard des différences clinico-biologiques majeures entre les patients atteints d'un SMD-SC *SF3B1*^{MUT} et les patients atteints d'un SMD sans SC *U2AF1*^{MUT} ou *SRSF2*^{MUT}. En effet, contrairement aux mutations de *SF3B1* qui sont de bon pronostic chez les patients porteurs d'un SMD, les mutations d'*U2AF1* et de *SRSF2* sont toutes deux associées à un risque plus élevé de transformation en LAM. Par exemple, 69% des patients *SRSF2*^{MUT} versus 39% des patients *SRSF2*^{WT} évoluent vers une LAM sur une période de cinq ans (Graubert et al., 2011; Papaemmanuil et al., 2013; Thol et al., 2012).

Les moelles osseuses des patients atteints d'un SMD-SC sont riches en progéniteurs et précurseurs érythroïdes immatures, ce qui suggère une prolifération accrue de ces derniers. Les expériences menées au sein du laboratoire montrent que : 1) plus de 30% des CMN de SMD-SC sont en phase S du cycle cellulaire, contre seulement 15% des CMN des autres sous-types de SMD (Fontenay-Roupie et al., 1999), 2) les mutations de *SF3B1* sont acquises de façon précoce dans les CSH et/ou les MPP et participent à la mise en place d'une hématopoïèse clonale (Chesnais et al., 2017) et 3) l'expression de 48 gènes de réplication et de réparation de l'ADN est augmentée dans les CMN des patients porteurs d'un SMD-SC *SF3B1*^{K700E} (Annexe 5, données du Dr Viginie Chesnais). **Les mutations de *SF3B1* sont-elles responsables d'une perte du contrôle de la réplication de l'ADN ? Les cellules *SF3B1*^{MUT} accumulent-elles des dommages à l'ADN ?**

II. Objectif

Mon travail a consisté à **caractériser les processus de réplication et de réparation de l'ADN dans les cellules érythroïdes mutées ou non sur *SF3B1***, grâce notamment à des expériences d'IF. Nous bénéficions de deux modèles cellulaires : les érythroblastes primaires des patients porteurs d'un SMD et les cellules murines G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E}. Ces deux modèles permettent de reproduire la différenciation érythrocytaire terminale *in vitro*.

III. Résultats

Caractérisation des processus de réplication et de réparation de l'ADN dans les cellules érythroïdes mutées sur le gène *SF3B1*

High speed of fork progression induces DNA replication stress in *SF3B1*-mutated erythroblasts

Sabrina Bondu,^{1,2,3,*} Carine Lefèvre,^{1,2,3,4,*} Anne Letessier,^{2,3,5} David Rombaut,^{1,2,3} Virginie Chesnais,^{1,2,3} Alice Rousseau,^{1,2,3,6} Olivier Kosmider,^{1,2,3,6} Sophie Park,⁷ Emmanuelle Clappier,⁸ Lionel Adès,⁹ Michel Wassef,¹⁰ Raphael Margueron,¹⁰ Angelos Constantinou,¹¹ Marc-Henri Stern,¹² Frédéric Chédin,¹³ Benoit Miotto,^{2,3,5} Michaela Fontenay^{1,2,3,6,¶}

¹ Institut Cochin, Department Development, Reproduction, Cancer, Team Normal and pathological hematopoiesis, Paris, France

² Institut National de la Santé et de la Recherche Médicale U1016, Paris, France

³ Université Paris Descartes, Paris, France

⁴ Laboratoire d'excellence du Globule Rouge GR-Ex, Université Paris Descartes, Paris, France

⁵ Institut Cochin, Department Development, Reproduction, Cancer, Team Epigenetics, DNA replication and Cancer, Paris, France

⁶ Assistance Publique-Hôpitaux de Paris, Hôpitaux Universitaires Paris Centre, Service d'hématologie biologique, Paris, France

⁷ Department of Hematology, Centre Hospitalier Universitaire, Université de Grenoble Alpes, Grenoble, France

⁸ Assistance Publique-Hôpitaux de Paris, Saint-Louis Hospital, Hematology Laboratory, University Paris Diderot, Paris, France.

⁹ Assistance Publique-Hôpitaux de Paris, Saint-Louis Hospital, Service Hematologie Séniors, University Paris Diderot, Paris, France

¹⁰ Institut Curie, PSL Research University, INSERM U934, UMR3215, Genetics and biology of development, Paris, France

¹¹ Institut de Génétique Humaine, Centre National de la Recherche Scientifique, Université de Montpellier, Montpellier, France

¹² Institut Curie, PSL Research University, INSERM U830, DNA repair and uveal melanoma and 11 Department of Genetics, Equipe labellisée par la Ligue National contre le Cancer, Paris, France

¹³ Department of Molecular and Cellular Biology and Genome Center, University of California, Davis, USA

* Equal contributors

¶ Corresponding author

Introduction

Myelodysplastic syndromes (MDS) with ring sideroblast (RS) represent a subtype of hematopoietic stem cell disorders characterized by a profound anemia, a bone marrow (BM) erythroid cell dysplasia, an excessive apoptosis of the mature precursors and a *SF3B1* gene mutation in 90% of cases (Papaemmanuil et al., 2011; Yoshida et al., 2011). Patients suffer from anemia-related morbidities and more specifically from systemic iron overload already present before any red blood cell transfusion. *SF3B1* mutants like other mutant splicing factors *SRSF2* or *U2AF1* may act as oncogenes by inducing subtle changes in mRNA maturation in approximately 1% of the whole transcriptome. Diseases development has been functionally linked to the inclusion of a toxic exon in the *Ezh2* gene in *Srsf2*^{P95H} knock-in mice (Kim et al., 2015) or a polyadenylation switch event in the autophagy gene *Atg7* in murine pro-B cell model overexpressing *U2af1*^{S34F} (Park et al., 2016). In *SF3B1*-mutated tumors, usage of alternative 3' acceptor splice site or 5' donor splice site generates aberrant transcripts mostly degraded by the nonsense-mediated decay (NMD) (Alsafadi et al., 2016; Darman et al., 2015; Dolatshad et al., 2016). In particular, the downregulation of the Fe-S cluster transporter *ABCB7* due to the degradation of the aberrant isoform of the transcript may contribute to the accumulation of iron in the mitochondria. Recently, we have demonstrated that hyperferritinemia is due to the transcriptional repression of hepcidin by overproduction of its erythroid regulator erythroferrone (ERFE). The canonical *ERFE* transcripts are overexpressed together with a mutant transcript isoform that is translated into a variant protein (Bondu, Alary et al., in revision). In addition, intron retention has been reported less frequent than in *SF3B1*-wild type MDS (Shiozawa et al., 2018).

As splicing usually occurs co-transcriptionally, mutant splicing factors may affect the transcriptional checkpoint, thus participating in a selective regulation of transcription. This would lead to specific patterns of gene expression contributing to the disease phenotypes. Mutant splicing factors may also support defects in the control of DNA replication and repair related to genomic instability. In particular, *SF3B1* has been identified in a screen for DNA repair pathway (Paulsen et al., 2009). More recently, SF3B1 proteins have been shown to interact with BRCA1 and facilitate the splicing of DNA repair genes in stress conditions (Savage et al., 2014). Conversely, the Fanconi anemia proteins FANCD2/I are involved in the control of the dynamic of nuclear distribution of SF3B1 (Moriel-Carretero et al., 2017).

We hypothesized that cells addicted to *SF3B1* mutations endure a chronic DNA replication stress and that *SF3B1* is therefore an oncogene. In the present paper, we report that DNA replication was increased in human *SF3B1*-mutated primary erythroblasts and murine CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 proerythroblastic cell line. Consistently, the rate of DNA fiber elongation was accelerated in *Sf3b1*^{K700E} G1E-ER4 cells. Single-stranded DNA (ssDNA) covered by replication protein A (RPA) spontaneously accumulated in actively dividing proerythroblast (Pro-E) and basophilic erythroblasts (Baso-E). Mechanistically, we hypothesize that replication forks encounter less obstacles in *SF3B1*-mutated compared to *SF3B1*-wildtype erythroblasts. To address this point, we will investigate the presence of R-loops at the promoter regions of genes that lose intron retention in the context of mutant *SF3B1*.

Results

Active DNA synthesis in SF3BI^{MUT} erythroblasts

We first defined the proliferative characteristics of *SF3BI^{MUT}* erythroblasts. To do so CD34⁺ hematopoietic stem and progenitor cells (HSPC) were isolated from the BM of patients with *SF3BI^{MUT}* or *SF3BI^{WT}* MDS and from femoral head BM samples of healthy donors. Cells were cultured in erythroid conditions as previously described (Bondu, Alary et al., in revision). Despite excessive apoptosis of *SF3BI^{MUT}* compared to *SF3BI^{WT}* precursors at day 7 and 10 of culture (**Supplementary Fig. 1A**), proliferation curves were superimposed until day 14 suggesting that the proliferation rate could be higher in *SF3BI^{MUT}* versus *SF3BI^{WT}* MDS or healthy controls (**Fig. 1A**). As shown by May-Grünwald-Giemsa staining, the differentiation was not abrogated in these cells (**Supplementary Fig. 1B**). Interestingly, the transition between proerythroblasts (Pro-E) and basophilic erythroblasts (Baso-E) was accelerated at day 9 of *SF3BI^{MUT}* erythroblast culture. By contrast, the percentage of polychromatophilic (Poly-E) and orthochromatic erythroblasts (Ortho-E) was lower in MDS compared to healthy donor cultures due to their high sensitivity of mature erythroblasts to apoptosis (**Fig. 1B**). According to Mohandas' team, CD49d and Band3 surface markers allow assessing differentiation by flow cytometry (Hu et al., 2013). Five subpopulations of erythroid precursors were distinguished among GPA⁺ cells. Using this method, we confirmed that *SF3BI^{MUT}* erythroblasts entered terminal differentiation more rapidly than healthy controls or *SF3BI^{WT}* MDS (**Fig. 1C & 1D**).

We directly investigated the rate of proliferation by labelling DNA during a 30-min pulse of BrdU at different steps of erythroid maturation. As expected, the mean fluorescence intensity (MFI) of labelling with FITC-anti-BrdU antibody decreased with time meaning that the replication rate slowed down during the maturation of normal erythroblasts from Pro-E, Baso-E to Poly-E (**Supplementary Fig. 1C**). Then we compared the proliferation rate of normal and *SF3BI^{MUT}* Baso-E. As shown in **Fig. 1E**, the FITC-anti-BrdU MFI was always increased in *SF3BI^{MUT}* cells compared to normal and the difference expressed as a ratio of fluorescence intensity (RFI) to isotypic antibody was statistically significant ($P < 0.001$), suggesting that the replication rate was higher in mutant Baso-E.

Figure 1: Active DNA synthesis in human *SF3B1*^{MUT} primary erythroblasts

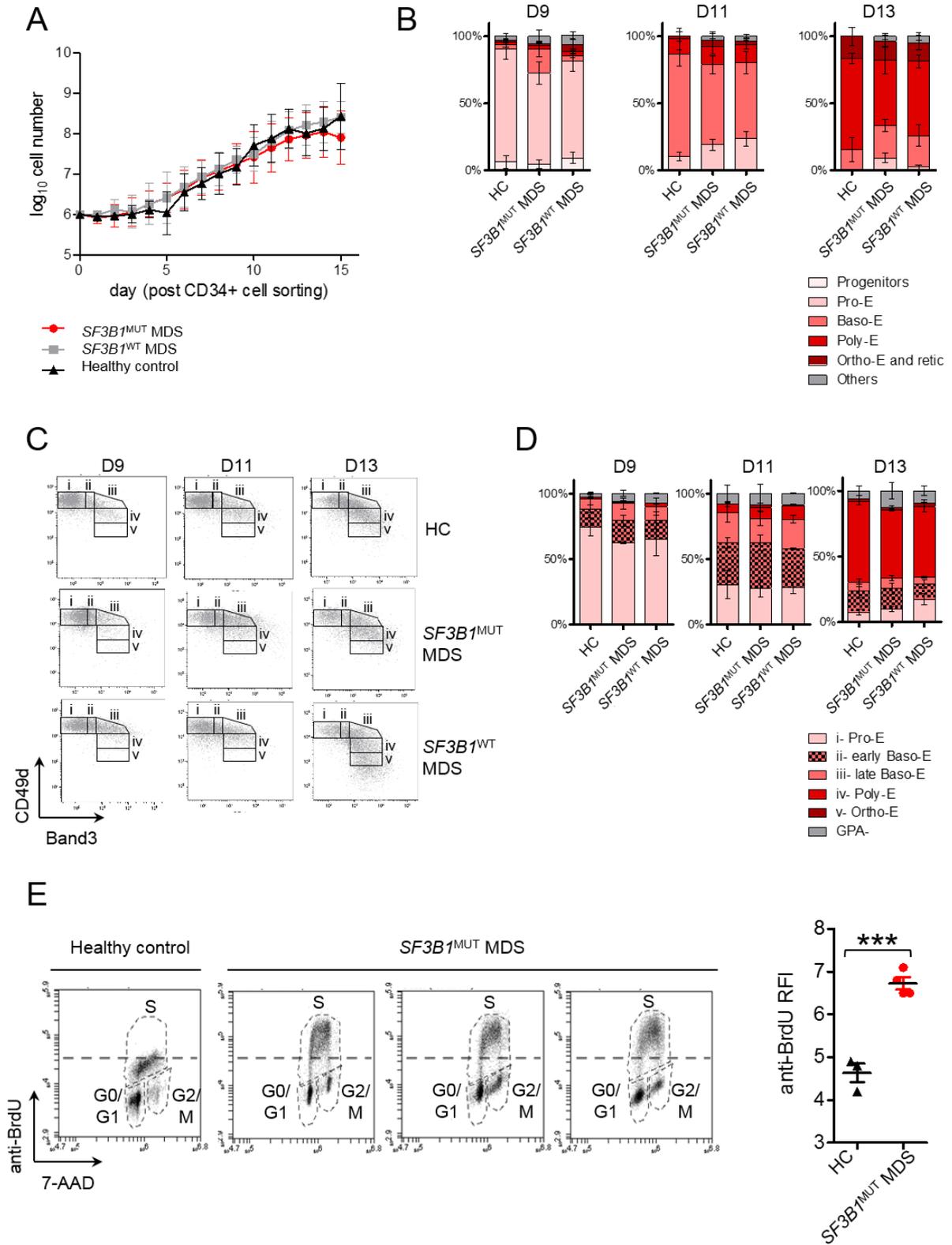


Figure 1: Active DNA synthesis in human *SF3BI*^{MUT} primary erythroblasts. (A) Human erythroblasts cell proliferation in culture. Erythroid cells were counted every day and proliferation curves were plotted in log₁₀ (cell number) for healthy controls (n=7), *SF3BI*^{MUT} MDS (n=13) and *SF3BI*^{WT} MDS (n=10). (B) Cytological analysis of human erythroblasts in culture. The cell composition in progenitors, Pro-E (proerythroblasts), Baso-E (basophilic erythroblasts), Poly-E (polychromatophilic erythroblasts), Ortho-E (orthochromatic erythroblasts), retic (reticulocytes) and others (granulocytes and/or monocytes) was determined every two days after MGG staining of cytopins. The mean percentages ± standard error of the mean (sem) are indicated for healthy controls (HC, n=3), *SF3BI*^{MUT} MDS (n=6) and *SF3BI*^{WT} MDS (n=3) at day 9, 11 and 13 post CD34⁺ cell sorting. (C) and (D) Flow cytometric analysis of human erythroblasts in culture. (C) Cultured erythroblasts at different days were stained with antibodies to GPA, Band3 and α4 integrin. Representative plots of α4 integrin versus Band3 expression in GPA⁺ cells are shown for 1 healthy control, 1 *SF3BI*^{MUT} patient and 1 *SF3BI*^{WT} patient. Precursors are separated as Pro-E (i), early Baso-E (ii), late Baso-E (iii), Poly-E (iv) and Ortho-E (v) (D) The proportion of each precursor was determined every 2 days according to Hu et al, Blood 2013. Mean percentages ± sem in healthy controls (n=3), *SF3BI*^{MUT} MDS (n=3) and *SF3BI*^{WT} MDS (n=3) are shown. (E) Cell cycle analysis in human Baso-E in culture. Cell cycle was analyzed by flow cytometry using a fluorescent anti-BrdU antibody and 7-aminoactinomycin D (7-AAD). Left: Representative plots of anti-BrdU versus 7-AAD of Baso-E are shown for 1 healthy control and 3 different *SF3BI*^{MUT} MDS. Right: Relative fluorescence intensity (RFI) is calculated as a ratio between the mean fluorescence intensity (MFI) of FITC anti-BrdU to FITC isotype control in S phase. Student t-test for statistical analysis (***) $P < 0.001$).

Mutant Sf3b1^{K700E} drives active DNA synthesis in murine erythroblasts

Epigenetic mutations (either *TET2* or *DNMT3A*) usually co-occur with *SF3B1* mutation in MDS-RS. To investigate whether active DNA synthesis depends on the sole presence of mutant *SF3B1*, we used CRISPR-Cas9 technology to edit the most frequent K700E aminoacid substitution in the murine G1E-R4 erythroblastic cell line deriving from embryonic stem cells invalidated for GATA1. These cells arrested at the proerythroblastic stage can be driven to terminal differentiation within 3 days by expressing a GATA1 transgene in fusion with estrogen receptor upon treatment with estradiol (Weiss et al., 1997; Welch et al., 2004) (**Supplementary Fig. 2A & 2B**). We obtained several clones expressing the *Sf3b1^{K700E}* mutation (clones H5.6, H5.13H and H10.26) or *Sf3b1^{WT}* (clones H9.2 and H9.82). Using the BrdU assay, we observed that the RFI of APC-anti-BrdU labelling tended to be higher in *Sf3b1^{K700E}* cell lines H5.6 and H5.13H compared to isogenic *Sf3b1^{WT}* H9.82 cell line and the mean percentage of cells in S phase also increased at the expense of G2-phase in *Sf3b1^{K700E}* cell lines H5.6 and H5.13H compared to isogenic *Sf3b1^{WT}* H9.82 cell line (n=3; 64.5% ± .4 in H5.6 and 64.9% ± 4.9 in H5.13H *versus* 56.0% ± 5.6 in H9.82 clone). The differences did not reach the statistical significance. In *Sf3b1^{K700E}* Baso-E obtained 24h after estradiol treatment, the percentage of S-phase cells was significantly increased between H5.6 and H9.82 clones ($P < 0.05$) and between H5.13H and H9.82 clones ($P < 0.0001$) (n=3; 61.0% ± 3.1 in H5.6 and 70.7% ± 0.3 in H5.13H *versus* 50.3% ± 0.9 in H9.82 clone) (**Fig. 2A, upper panel**).

We then incubated the cells with aphidicolin for 16h to block the cell cycle in G1/S phase without or with a 24-hour estradiol treatment. After wash-out, the percentage of S-phase Pro-E was equivalent in *Sf3b1^{K700E}* and *Sf3b1^{WT}* conditions. By contrast, the percentage of S-phase cells after release from G1/S blockade is increased in *Sf3b1^{K700E}* Baso-E compared to *Sf3b1^{WT}* Baso-E, especially in H5.13H clone in which the mean percentage of S-phase cells was 82.3% ± 1.8 *versus* 70.3% ± 0.9 in H9.82 ($P < 0.01$). *Sf3b1*-mutated Baso-E recover more efficiently from the G1/S blockade than *Sf3b1*-non mutated Baso-E, maybe due to an increase of DNA replication origin firing or fork speed. (**Fig. 2A, bottom panel**).

Altogether, our results demonstrate an increased rate of DNA replication in human primary *SF3B1^{MUT}* Baso-E and in murine CRISPR-Cas9 *Sf3b1^{K700E}* Pro-E and Baso-E.

Combing

To investigate whether the increased rate of DNA replication was related to a faster speed of replication forks in cells harbouring a mutant *SF3B1* allele, we performed DNA fiber combing after two consecutive pulses with DNA labelling probes (**Fig. 2B**). The experiment was first performed in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 cell line or isogenic CRISPR-Cas9 *Sf3b1*^{WT} G1E-ER4 maintained in a proerythroblastic stage. As shown by measuring the length of labelled fibers, the replication fork speed was significantly accelerated by 17% in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 cells (clone H5.6; 1.87 kb/min \pm 0.04) above the velocity in CRISPR-Cas9 *Sf3b1*^{WT} G1E-ER4 cells (clone H9.82; 1.60 kb/min \pm 0.04; $P < 0.0001$), showing an increase of replication fork speed (**Fig. 2C and 2D, upper panels**). To definitively rule out the possibility of firing of new origins, we need to analyse more DNA fibers. These experiments are ongoing.

Second, the experiment was conducted in CRISPR-Cas9 *Sf3b1*^{K700E/WT} G1E-ER4 cell lines entering the differentiation by the addition of estradiol. At the basophilic stage, the fork speed was also significantly increased by 14% in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 (clone H5.6; 1.66 kb/min \pm 0.05) compared to CRISPR-Cas9 *Sf3b1*^{WT} G1E-ER4 cells (clone H9.82; 1.39 kb/min \pm 0.1; $P < 0.001$) (**Fig. 2C and 2D, bottom panels**).

We next estimated the fork symmetry from the calculation of the IdU/CldU ratio. In *Sf3b1*^{WT} and *Sf3b1*^{K700E} Pro-E G1E-ER4 cells, the forks were symmetric as shown by the ratio median values of around 1. In *Sf3b1*^{WT} and *Sf3b1*^{K700E} Baso-E G1E-ER4 cells, the forks were also symmetric, suggesting no fork stalling and/or fork collapse. (**Fig. 2E**).

Altogether, these data show that in the context of *SF3B1* mutation, the increased rate of DNA replication is related to a faster speed of fork progression. *SF3B1* mutation induces a replication stress without stalling forks.

Figure 2: Enhanced speed of replication fork progression in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 cells

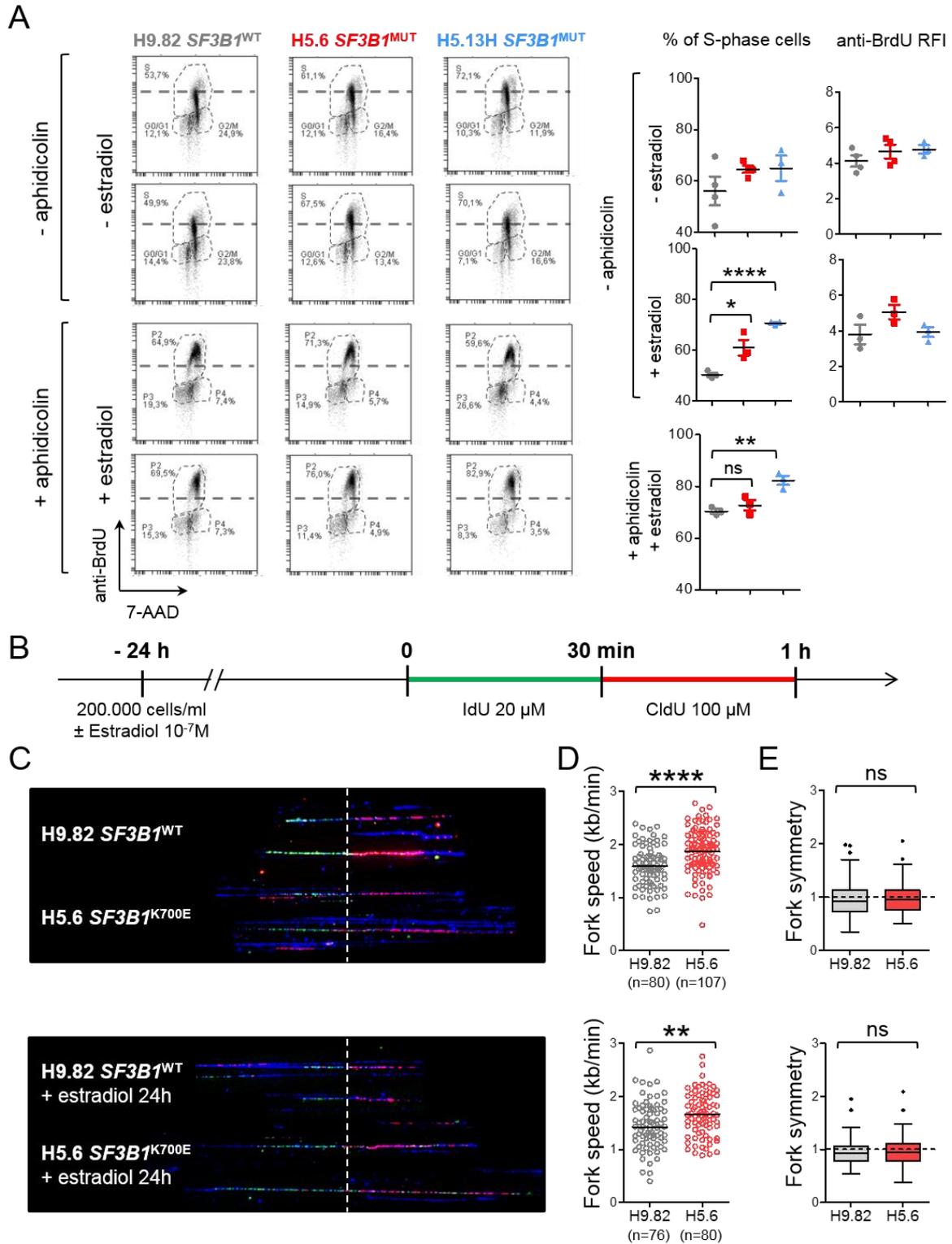


Figure 2: Enhanced speed of replication fork progression in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 cells. (A) Cell cycle analysis in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 Pro-E and Baso-E. Cell cycle was analyzed by flow cytometry using a fluorescent anti-BrdU antibody and 7-aminoactinomycine D (7-AAD) in CRISPR-Cas9 *Sf3b1*^{WT} (H9.82) and *Sf3b1*^{K700E} (H5.6 and H5.13H) clones. Left: Anti-BrdU *versus* 7-AAD plots representative of at least 3 independent experiments are shown. Estradiol treatment for 24h (10⁻⁷ M) induced the erythroid differentiation from Pro-E to Baso-E. Aphidicolin treatment for 16h (0.6 µg/mL) blocked the cells in G1/S phase. Right: Statistical analysis of the ratio of APC anti-BrdU mean fluorescence intensity (RFI) and percentage of S-phase cells. Right, upper panel: Pro-E. Middle panel: Baso-E after estradiol. Bottom panel: G1/S blockade in Baso-E by aphidicolin treatment and release. Student t-test for *P*-values; **** *P* < 0.0001; ** *P* < 0.01; * *P* < 0.05. (B) Schematic representation of molecular combing technique. *Sf3b1*^{WT} H9.82 and *Sf3b1*^{K700E} H5.6 clones ± estradiol were successively labelled with IdU and CldU. (C) Images representative of one experiment. Upper panel: *Sf3b1*^{WT} H9.82 and *Sf3b1*^{K700E} H5.6 clones without estradiol. Bottom panel: *Sf3b1*^{WT} H9.82 and *Sf3b1*^{K700E} H5.6 clones with estradiol. (D) Measurement of fork speed in *Sf3b1*^{WT} H9.82 and *Sf3b1*^{K700E} H5.6 clones. Results from one experiment as medians +/- interquartile ranges. Mann-Whitney test for *P*-values; **** *P* < 0.0001; ** *P* < 0.01. No estradiol, median length of fibers: H9.82 = 253.77; H5.6 = 239.13. With estradiol, median length of fibers: H9.82 = 228.71; H5.6 = 264.56. (E) Fork symmetry. Repartition of IdU/CldU ratios as medians +/- interquartiles ranges. Mann-Whitney test for *P*-values. The biological duplicate will be analyzed in March 2019.

Single-stranded DNA exposure, but no double stranded breaks in SF3BI^{MUT} erythroblasts

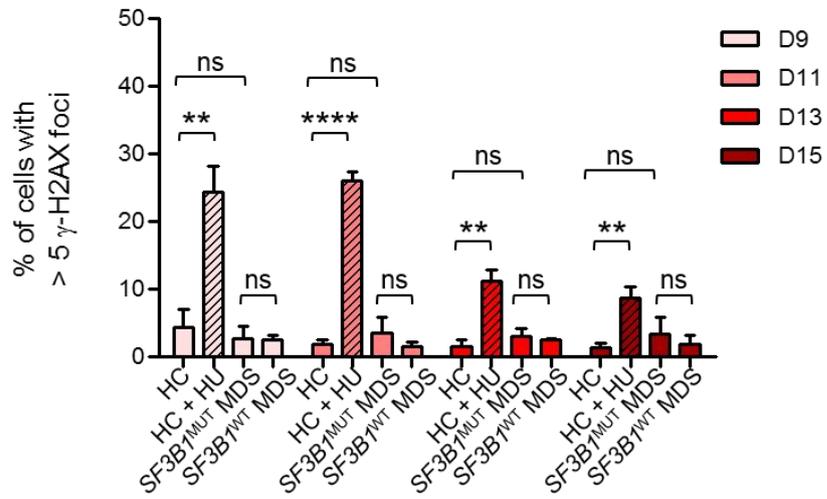
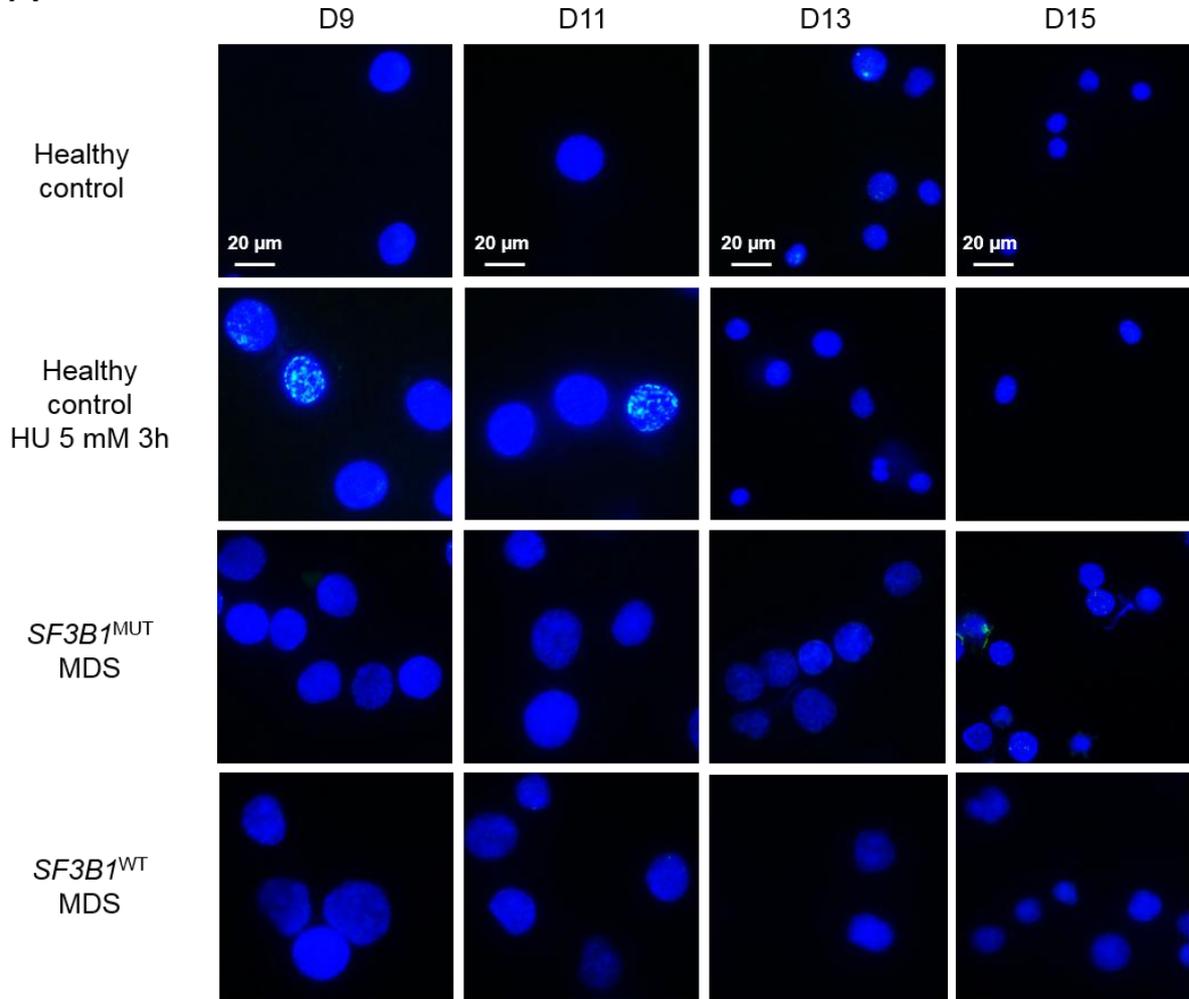
To test the hypothesis that fast fork progression induces DNA damage, we investigated the hallmarks of the DNA damage response (DDR). Immunofluorescence labelling of p-RPA32 serine4/8, γ -H2AX or 53BP1 was performed in human *SF3BI*-mutated or wild type primary erythroblasts. As a positive control, DNA damage was induced in healthy donor erythroblasts by depleting the pool of nucleotides with hydroxyurea (**Supplementary Fig. 3A**). Experiments did not show evidences of γ -H2AX or 53BP1 foci at any stages of the differentiation of primary erythroblasts derived from *SF3BI^{MUT}* or *SF3BI^{WT}* bone marrows (**Fig. 3A and 3B**). By contrast, p-RPA32 foci were spontaneously detected in *SF3BI^{MUT}* Pro-E or Baso-E but neither in healthy controls nor in *SF3BI^{WT}* erythroblasts. These foci were no longer observed in more mature *SF3BI^{MUT}* erythroblasts at Poly-E and Ortho-E stages (**Fig. 3C**).

Western blot experiments also showed that p-RPA32 s4/s8 was spontaneously expressed in *Sf3b1^{K700E}* G1E-ER4 H5.13H and H10.26 clones at Pro-E stage (**Supplementary Fig. 3B**). The *Sf3b1^{K700E}* H5.13H clone was also more sensitive to low doses HU than isogenic *Sf3b1^{WT}* G1E-ER4 cells (clone H9.2) as shown by a higher percentage of p-RPA32 s4/s8-positive cells both at Pro-E and Baso-E stages by immunofluorescence experiments (**Supplementary Fig. 3C**).

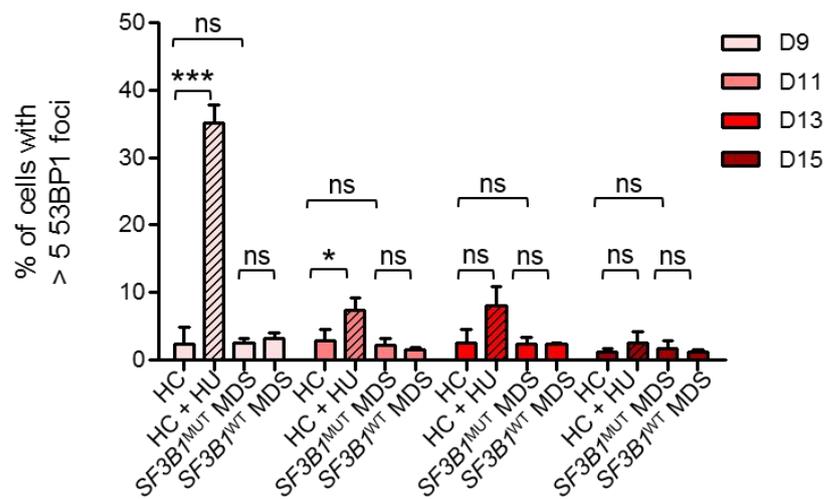
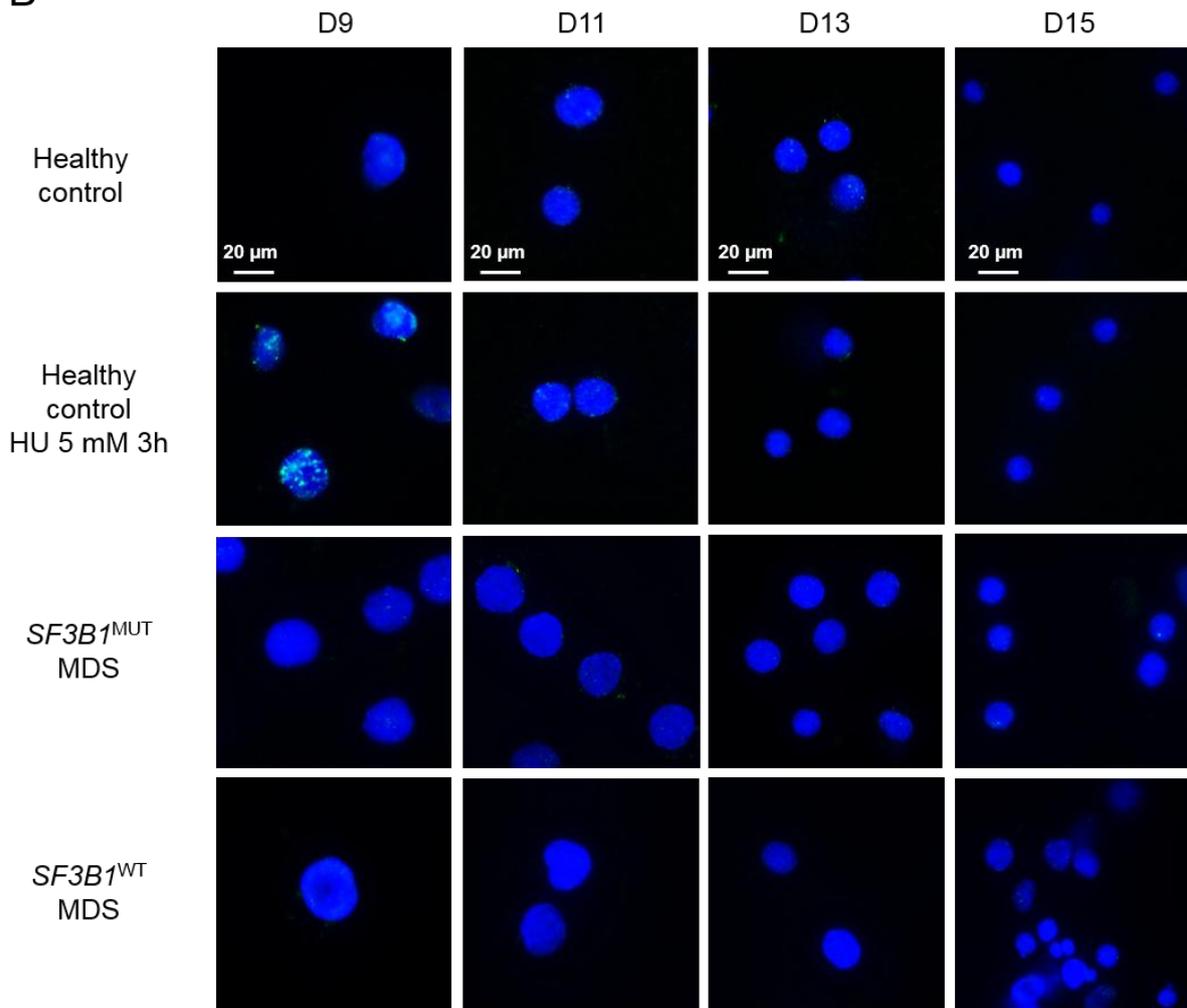
Altogether these data indicate that fast fork progression did not produce DNA double strand breaks but rather an exposure of ssDNA in actively proliferating Pro-E and Baso-E.

Figure 3: Single-stranded DNA exposure, but no double strand break in human *SF3B1*^{MUT} immature erythroblasts

A



B



C

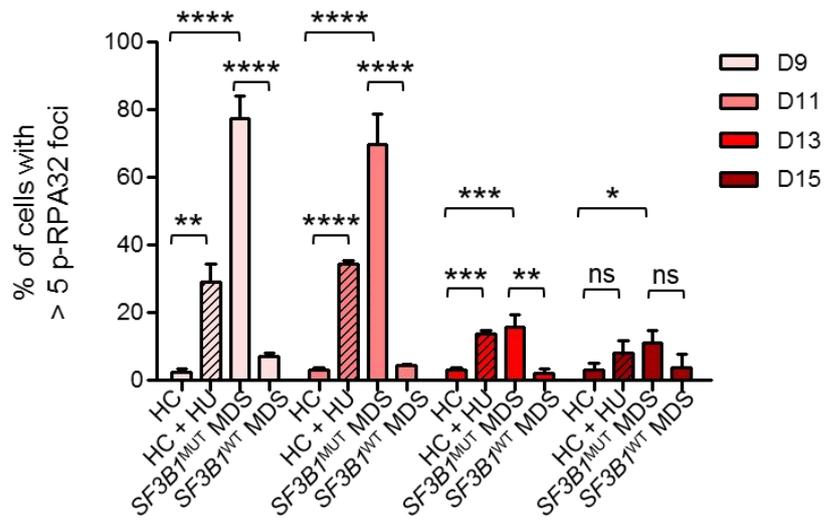
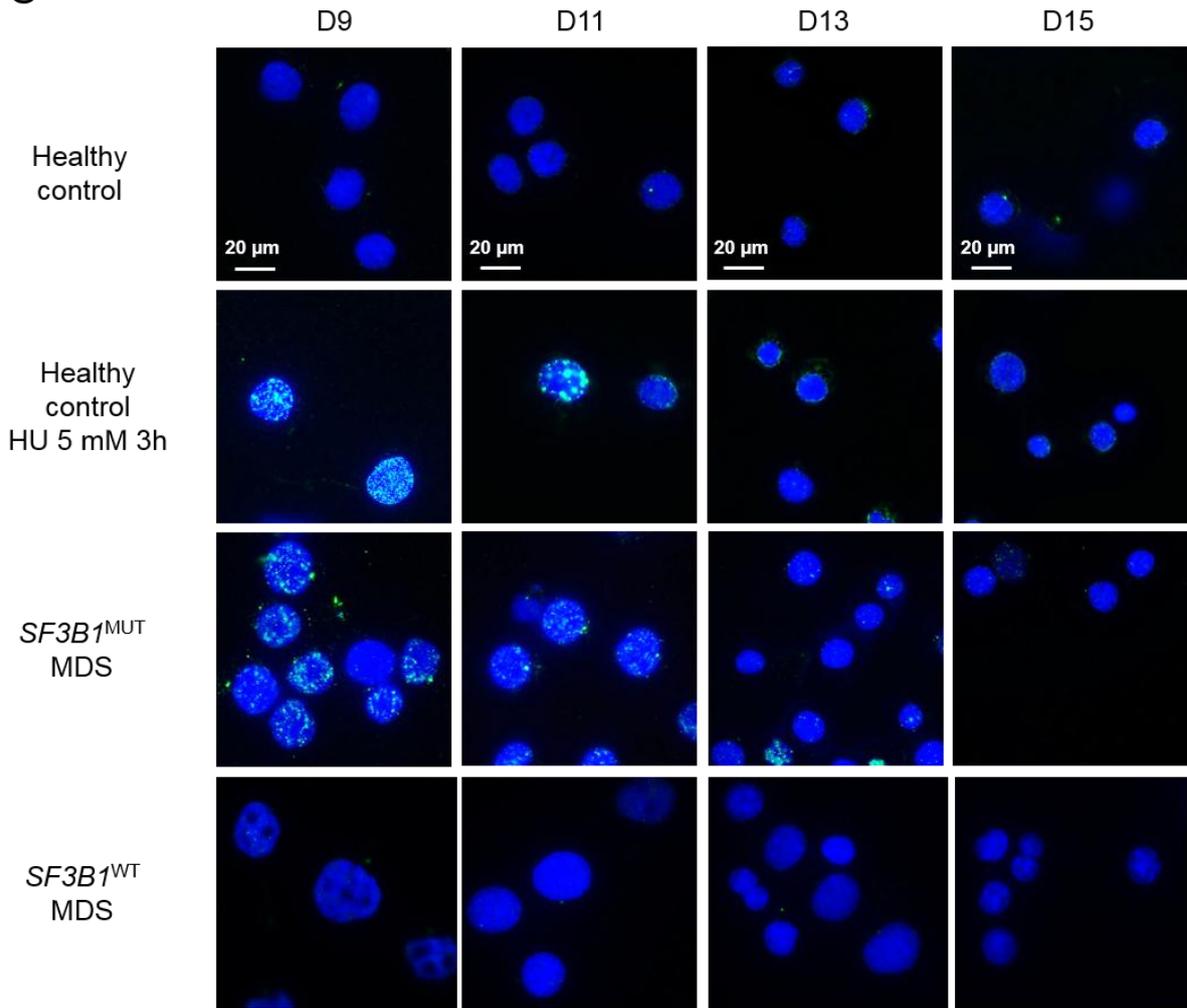


Figure 3: Single-stranded DNA exposure, but no double strand break in human *SF3B1*^{MUT} immature erythroblasts. Immunofluorescence analysis of DNA damage response (DDR) in human primary erythroblasts of healthy control samples (n=3), healthy control samples treated with HU (n=3), *SF3B1*^{MUT} MDS samples (n=3) and *SF3B1*^{WT} MDS samples (n=2) at day 9, 11, 13, 15 of *in vitro* erythroid differentiation. Upper panels: Representative DyLight 488-staining of DDR markers (A) γ H2AX s139, (B) 53BP1 and (C) p-RPA32 s4/s8. DNA was stained with DAPI (blue). Healthy control (HC) cells treated with 5 mM hydroxyurea (HU) for 3h were used as positive controls. Images were acquired using an inverted microscope at 100X magnification. Scale: 20 μ m. Bottom panels: Cells with > 5 intranuclear foci were considered positive and ≥ 100 cells were counted by slide. Student t-test for statistical analysis; **** $P < 0.0001$, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

Identification of loss of intron retention (LIR) in SF3B1^{MUT} cells by RNA-sequencing

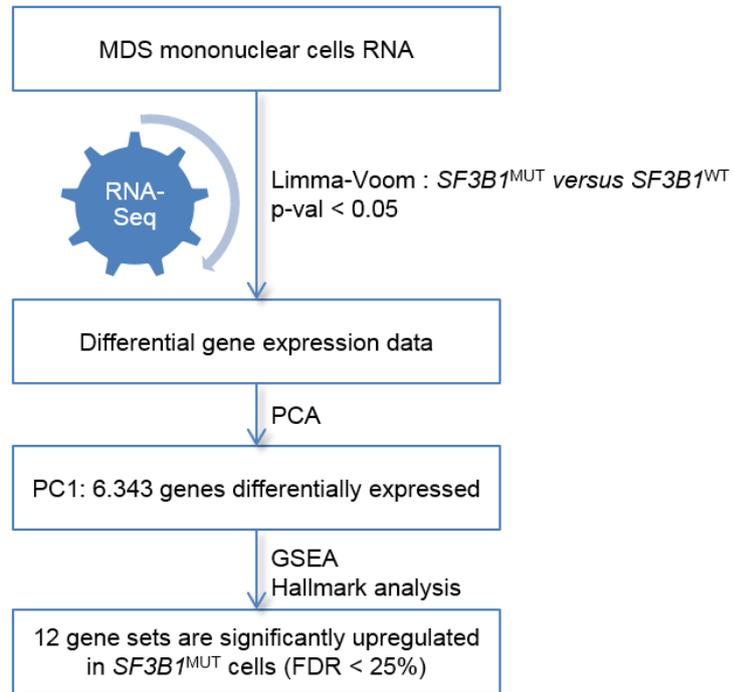
RNA splicing perturbation by either pharmacologic modulation or expression of the *U2AF1*^{S34F} or *SRSF2*^{P95H} mutant increased levels of R-loops, a transcription intermediate containing an RNA:DNA hybrid and displaced ssDNA across the genome, leading to a marked reduction in replicative DNA fiber length and compromised proliferation of bone-marrow-derived blood progenitors (Chen et al., 2018; Nguyen et al., 2017, 2018). Intron retention could cause R-loops formation because such mRNA retains the capacity to hybridize with the template DNA. Because DNA replication forks fastly progresses in *SF3B1*^{K700E} erythroblasts, we hypothesize that these mutant cells may have less intron retention events than normal erythroblasts. To explore this hypothesis, we performed deep RNA-sequencing both in human BM mononuclear cells from 21 patients with an *SF3B1* mutation and 6 patients without an *SF3B1* mutation, in human primary *SF3B1*^{MUT} or *SF3B1*^{WT} Baso-E and murine CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 cells. A differential analysis was performed to identify loss of intron retention (LIR) in mutant samples. Preliminary results of RNA-seq showed that a set of 6,343 genes discriminated *SF3B1*^{MUT} from *SF3B1*^{WT} BM mononuclear cell samples. A gene set enrichment analysis demonstrated that MYC targets, heme metabolism, G2/M checkpoint and DNA repair were among the 12 most upregulated gene sets (**Fig. 4**). Using SpliceSeq (Ryan et al., 2012), we found 615 transcripts with intron retention in all the *SF3B1*^{MUT} samples *versus* 725 transcripts with intron retention in all the *SF3B1*^{WT} samples (**Supplementary tables 7 and 8**). Among these intron retentions, 529 were found in both *SF3B1*^{MUT} and *SF3B1*^{WT} samples. We will focus on the transcripts with LIR in *SF3B1*^{MUT} samples.

DRIP-qPCR

R-loops will be investigated at promoter regions of genes identified with loss of intron retention in mutant cells.

Figure 4: Differential gene expression in human *SF3B1*^{MUT} bone marrow mononuclear cells by RNA-seq

A



B

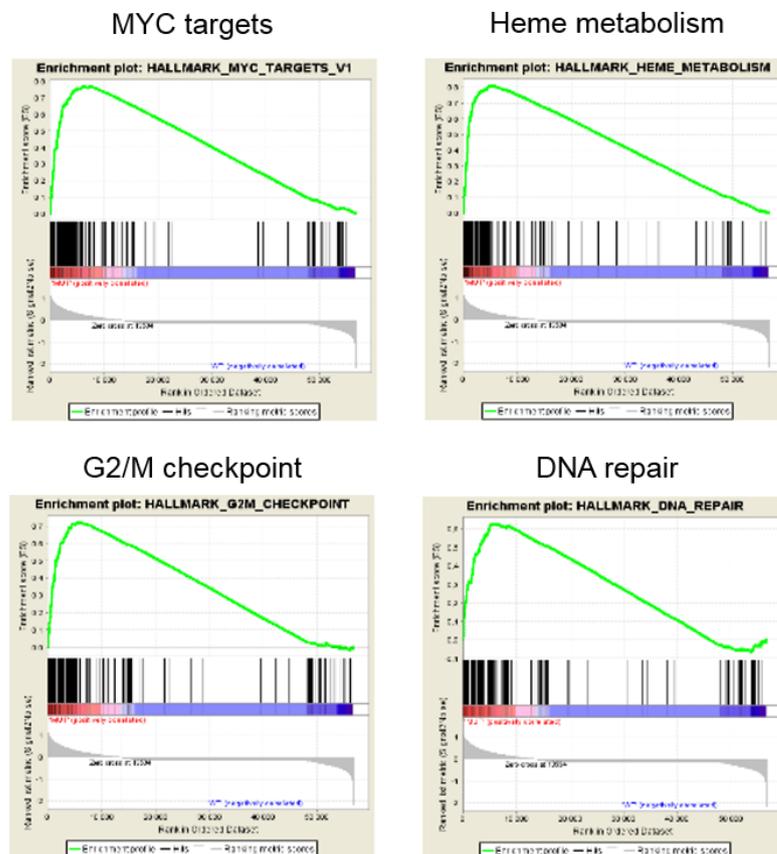
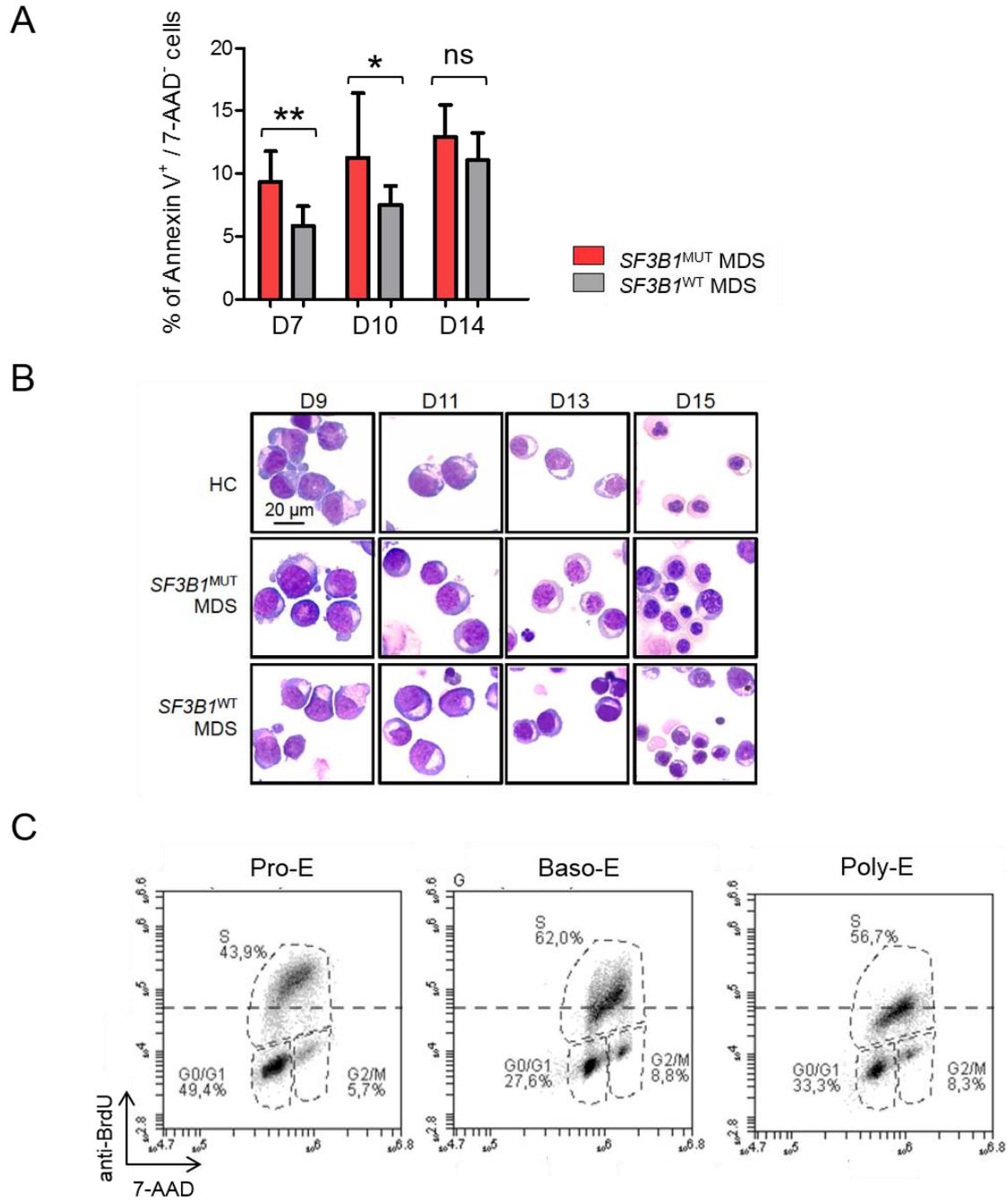


Figure 4: Differential gene expression in human *SF3B1*^{MUT} bone marrow mononuclear cells by RNA-seq. (A) RNA-Seq from 21 *SF3B1*^{MUT} BM mononuclear cell samples and 6 *SF3B1*^{WT} mononuclear cell samples. Principal component analysis and gene set enrichment analysis (GSEA) of the PC1. (B) Most upregulated gene sets panels in *SF3B1*^{MUT} samples.

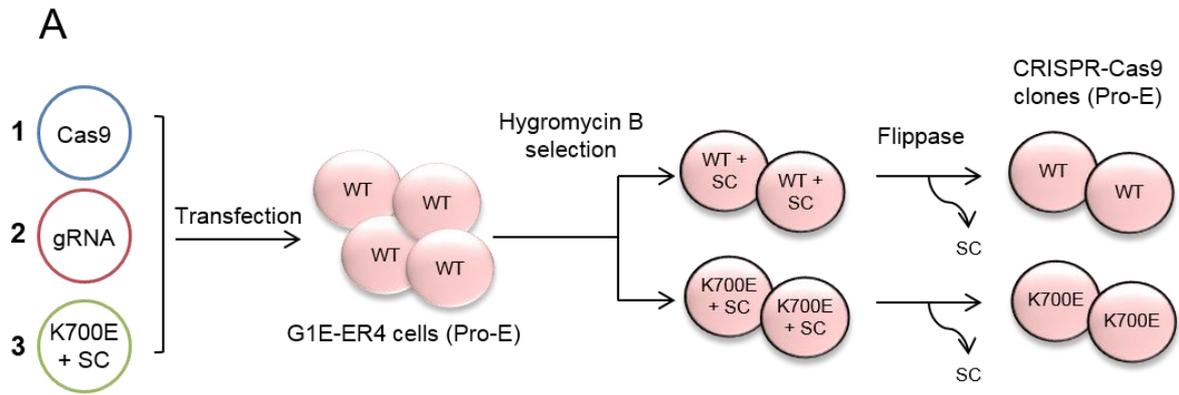
Supplemental Figures

Supplemental Figure 1: Human primary erythroblasts

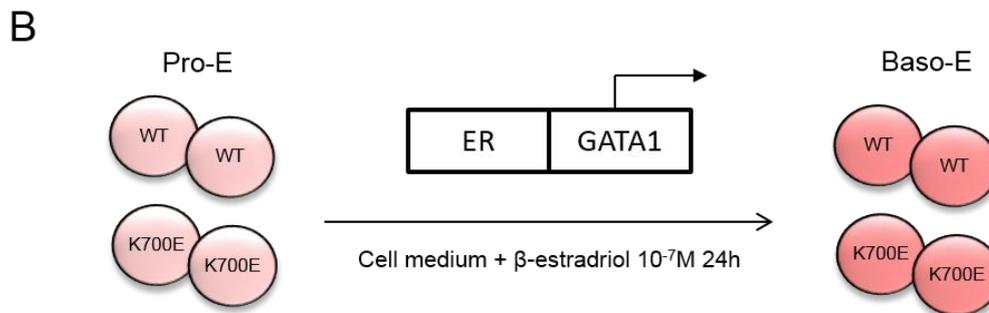


Erythroid precursors	anti-BrdU RFI
Pro-E	9,3
Baso-E	4,8
Poly-E	4,1

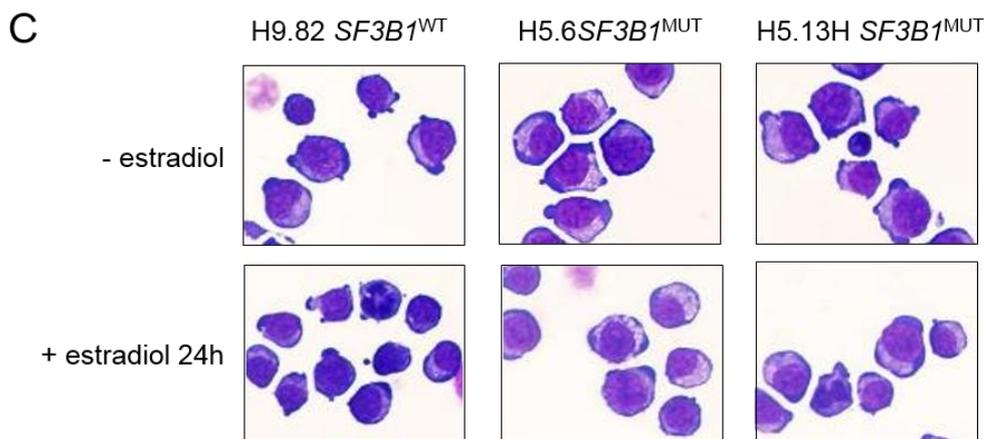
Supplemental Figure 2: G1E-ER4



1, 2 and 3 = vectors
3 : donor template (SC = hygromycin Selection Cassette)

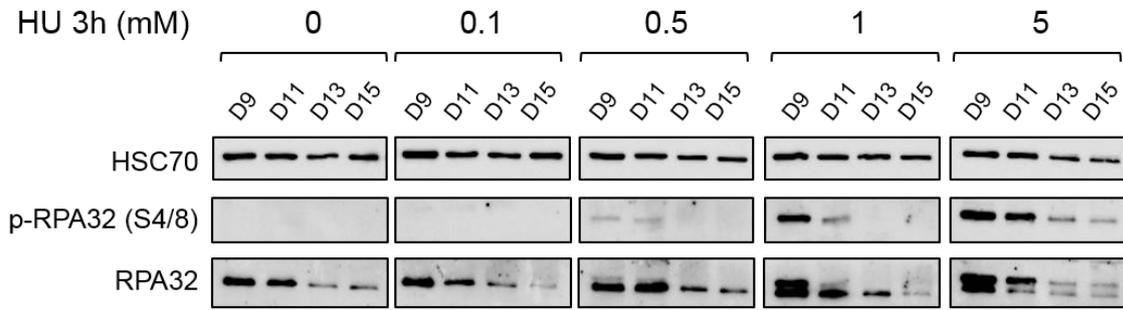


ER = estrogen receptor

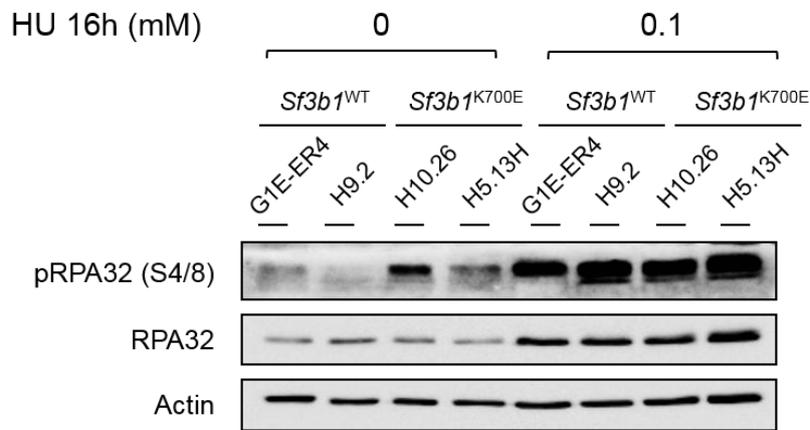


Supplemental Figure 3: Single-stranded DNA

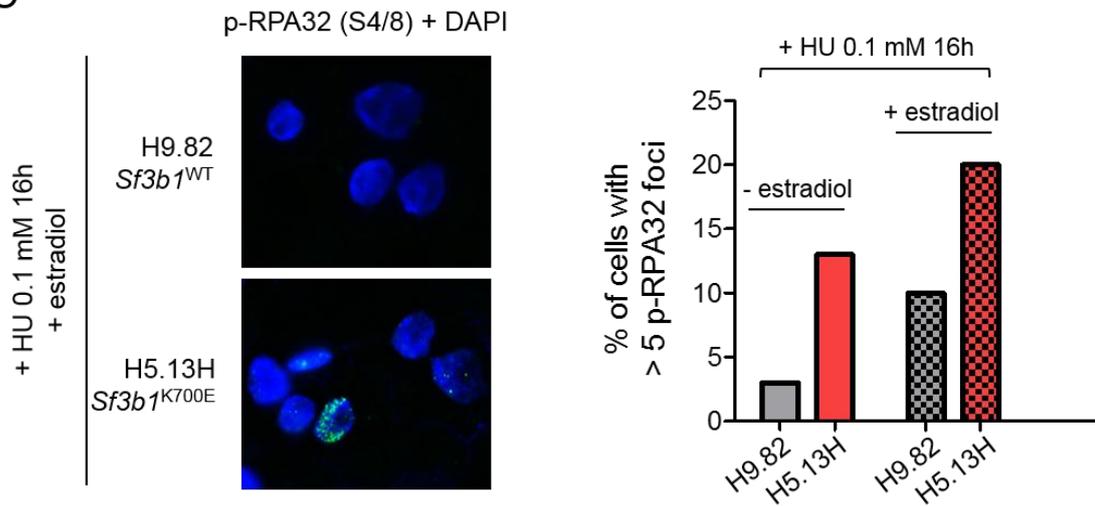
A



B



C



Supplemental Figure 1: Human primary erythroblasts. (A) Apoptosis is increased in *SF3BI*^{MUT} versus *SF3BI*^{WT} MDS erythroblasts in culture. Cell apoptosis was analysed by flow cytometry using a fluorescent annexin V and propidium iodide (PI) in 6 *SF3BI*^{MUT} MDS and 10 *SF3BI*^{WT} MDS cultures. Student t-test for statistical analysis; ** $P < 0.01$, * $P < 0.05$. (B) Cytological characteristics of human primary erythroblasts in culture. Representative images of erythroblast morphology on May-Grünwald-Giemsa-stained cytopins are shown. D9 corresponds to proerythroblasts, D11 to basophilic erythroblasts, D13 to polychromatophilic erythroblasts and D15 to orthochromatic erythroblasts. (C) Cell cycle analysis during *in vitro* normal erythroid differentiation. Cell cycle was analyzed by flow cytometry using a fluorescent anti-BrdU antibody and 7-AAD. Anti-BrdU RFI in S phase progressively decreases with erythroid differentiation.

Supplemental Figure 2: G1E-ER4. (A) Design of CRISPR-Cas9 constructs. The murine G1E-ER4 cell line was used to generate isogenic *Sf3b1*^{K700E} and *Sf3b1*^{WT} cell lines by CRISPR-Cas9 technology. Cells were transfected with the Cas9 endonuclease vector, the m*Sf3b1*-specific gRNA vector and a donor template encoding a hygromycin selection cassette (SC) and the *Sf3b1*^{K700E} mutation. Hygromycin was added to the culture medium to select clones with SC. Hygromycin-resistant cells were single cell-sorted and analysed by DNA sequencing to assess the presence of the *SF3BI*^{K700E} mutation. The SC was finally removed by flippase-mediated excision. Cassette-excised clones were assessed by PCR and *SF3BI*^{K700E} and *SF3BI*^{WT} expression was verified by RT-PCR and cDNA Sanger sequencing. (B) Schematic representation of the induction of erythroid differentiation in G1E-ER4 cells. G1E-ER4 murine Pro-E are invalidated for GATA1. A GATA1 transgene in fusion with estrogen receptor (ER) was induced upon treatment with β -estradiol (10^{-7} M), to force the transition from Pro-E to Baso-E within 24h. (C) Cytology of *Sf3b1*^{WT} and *Sf3b1*^{MUT} clones with or without induction of differentiation. β -estradiol treatment induced a reduction in cell size and condensation of chromatin, assessed by MGG staining.

Supplemental Figure 3: Single-stranded DNA. (A) Analysis of DNA replicative stress in normal erythroblasts in culture by western blot. Phosphorylation of RPA32 (s4/s8) in primary erythroblasts, at days 9 and 11 was induced by hydroxyurea (HU). (B) p-RPA32 s4/s8 expression by western blot. *Sf3b1*^{K700E} H10.26 and H5.13H clones were compared to parental G1E-ER4 cells and H9.2 clone (*Sf3b1*^{WT}). (C) ssDNA exposure in CRISPR-Cas9 *Sf3b1*^{K700E} G1E-ER4 treated with low doses of HU. Cells were treated with 0.2 mM HU during 16h and

RPA32 phosphorylation (s4/s8) was analysed by immunofluorescence. Right: IF representative images of H9.82 (*Sf3b1*^{WT}) and H5.13H (*Sf3b1*^{K700E}) Baso-E. Left: Percentage of cells with > 5 p-RPA32 foci in one experiment with estradiol and one experiment without estradiol.

Material and Methods

Patients

Lower risk MDS patients were recorded between 2014 and 2018. Bone marrow (BM) aspirates were collected after each patient gave his informed consent for biological investigations according to the recommendations of institutional review board (IRB) and ethics committee (IRB numbers: IdFV 212-A01395-38 EudraCT 2012-002990-7338; OncoCCH 2015-08-11-DC). Femoral head BM samples were collected during hip surgery from 7 healthy donors. Patient characteristics including gender, age, WHO, hemogram, BM blast and erythroblast percentages, karyotype, IPSS-R are indicated in **Supplementary table 1**.

Genomic studies

BM mononuclear cells were purified on Ficoll gradient and were processed for DNA extraction using the DNA/RNA Kit (Qiagen, Hilden, Germany). Mutations in *SF3B1* were screened by Sanger sequencing, as previously described (Damm et al., 2012).

Supplementary Table 1: Characteristics of patients according to *SF3BI* status

Parameters	<i>All</i>	<i>SF3BI</i> ^{WT}	<i>SF3BI</i> ^{MUT}	<i>P-value</i>
	n=19	n=6	n=13	
Age				
Median, years (range)	74 (44-86.6)	73.4 (44-79)	74.4 (59.4-86.6)	0.164
Sex				
male, no. (%)	7 (37)	3 (50)	4 (31)	ns
WHO-Subtype 2016				
5q/MDS-SLD no. (%)	3 (16)	3 (50)	0	< 0.001
MDS-SLD-RS/MDS-MLD-RS, no. (%)	13 (68)	0	13 (100)	
MDS-MLD, no. (%)	2 (11)	2 (33)	0	
MDS-EB1, no. (%)	1 (5)	1 (17)	0	
Hemoglobin (G/L)				
Median (IQR)	9.2 (6.2-12.9)	10.4 (9.2-12.9)	8.8 (6.2-10.9)	0.010
Neutrophil count (G/L)				
Median (IQR)	3 (0.8-5.8)	2.0 (1.3-4.1)	3.5 (0.8-5.8)	0.216
Platelet count (G/L)				
Median (IQR)	240 (46-577)	97.5 (46-393)	241 (175-577)	0.080
Karyotype IPSS-R				
Very good, no. (%)	1 (5)	0	1 (8)	0.672
Good, no. (%)	14 (74)	4 (67)	10 (77)	
Intermediate, no. (%)	4 (21)	2 (33)	2 (33)	
Poor, no. (%)	0	0	0	
BM blasts (%)				
Median (IQR)	2 (0-7)	1.5 (0-7)	2 (0-4)	0.451
BM erythroblasts (%)				
Median (IQR)	28.25 (1-58)	18 (1-24)	36.5 (25-58)	0.001
IPSS-R				
Very low risk, no. (%)	2 (11)	1 (17)	1 (8)	0.502
Low, no. (%)	13 (68)	3 (50)	10 (77)	
Int, no. (%)	4 (21)	2 (33)	2 (15)	
High, no. (%)		0	0	

Cell cultures

For erythroid cell expansion, CD34⁺ cells were isolated from the mononuclear cell fraction of bone marrow samples using the MidiMacs system (Miltenyi Biotec, Bergisch Gladbach, Germany). CD34⁺ cells, which purity was higher than 80%, were then cultured at 0.8×10^6 per mL for 4 days in Iscove's modification of Dulbecco medium (IMDM; Thermo Fisher Scientific) containing 15% BIT9500 (Miltenyi Biotec), 100 U/mL penicillin (Thermo Fisher Scientific), 100 µg/mL streptomycin (Thermo Fisher Scientific), 2 mM L-glutamine (Thermo Fisher Scientific), 2 UI/mL recombinant human erythropoietin (rHu Epo; Roche, Basel, Switzerland), 100 ng/mL stem cell factor (SCF; Miltenyi Biotec), 10 ng/mL interleukin 6 (IL6; Miltenyi Biotec) and 2.10^{-7} M dexamethasone (Sigma Merck, Darmstadt, Germany). Cells were diluted every day in the same medium until day 4. From day 4, IL6 was removed. From days 10 to 16 cells were switched to rHu Epo (2 UI/mL) to obtain terminal erythroid differentiation assessed by flow cytometry (GPA, CD71, CD49d and Band3).

Murine G1E-ER4 proerythroblastic cell line expressing a GATA1-estrogen receptor fusion gene (Weiss et al., 1997) is cultured in IMDM, containing 20% fetal calf serum (FCS; GE Healthcare, Chicago, IL), 100 U/mL penicillin, 100 µg/mL streptomycin, 2 mM L-glutamine, 2 U/mL rHU Epo, SCF in Chinese Hamster Ovary cell conditioned medium, monothioglycerol and 0,5 µg/ml puromycin (Sigma Merck, Darmstadt, Germany) to select cells expressing the GATA1-ER fusion gene.

CRISPR-Cas9 generation of isogenic $Sf3b1^{K700E}$ and $Sf3b1^{WT}$ cell lines

The murine G1E-ER4 cell line was used to generate isogenic $Sf3b1^{K700E}$ and $Sf3b1^{WT}$ cell lines using CRISPR-Cas9-stimulated homology-mediated repair (Welch 2004). Cells were transfected using Amaxa Nucleofector™ 2b Device program A-024 and Nucleofector kit L (Lonza) with Cas9 (Ref. 41815, Addgene, Cambridge, MA), a m*Sf3b1*-specific gRNA (built from gRNA cloning vector, Ref. 41824, Addgene) and a donor template encoding a hygromycin selection cassette at a 1:1:1 ratio. After two days, hygromycin was added to the culture at 800µg/mL for 72h, and maintained at 400µg/mL for another 72h. Hygromycin-resistant cells were then single cell-sorted using the FACS Aria III (Becton Dickinson Biosciences, Franklin Lakes, NJ) and clones were analyzed 3 weeks later by DNA sequencing to assess the presence of the $Sf3b1^{K700E}$ mutation. The selection cassette was removed by flippase-mediated excision. $Sf3b1^{K700E}$ and $Sf3b1^{WT}$ hygromycin-resistant clones were

transfected with a flippase-encoding vector and single cell-sorted. Cassette-excised clones were assessed by PCR and *Sf3b1*^{K700E} and *Sf3b1*^{WT} expression was verified by RT-PCR and cDNA Sanger sequencing.

May-Grunwald Giemsa staining and flow cytometry

Erythroid differentiation of murine G1E-ER4 cell line and human primary erythroblasts was followed by May-Grünwald Giemsa staining of cytopsins every 2 days and by flow cytometry. Briefly, 10⁵ cells are washed in 1X phosphate buffered saline (PBS) supplemented with 2% FCS and incubated for 20 min at 4°C with fluorescent antibodies to GPA, CD71, CD49d and Band3 (**Supplementary table 2**) and analyzed on LSRFortessa (Becton-Dickinson, Franklin Lakes, NJ) with Kaluza software (Beckman Coulter, Brea, CA).

Supplementary Table 2: Flow cytometry antibodies

Species reactivity	Antibody	Fluorochrome	Clone and Immunoglobulin class	Reference	Dilution
Human	GPA (CD235a)	PE-Cy7	11E4B-7-6, IgG1 (mouse)	A71564, Beckman Coulter	1/50
	CD71	FITC	YDJ1.2.2, IgG1 (mouse)	IM0483, Beckman Coulter	1/50
	Band 3 (CD233)	PE	BRIC 6, IgG3 (mouse)	9439PE, IBGRL	1/200
	α 4-integrin (CD49d)	APC	HP2/1, IgG1 (mouse)	B01682, Beckman Coulter	1/10
	isotype control	PE-Cy7	679.1Mc7, IgG1 (mouse)	IM2475, Beckman Coulter	1/50
	isotype control	FITC	679.1Mc7, IgG1 (mouse)	A07795, Beckman Coulter	1/50
	isotype control	APC	679.1Mc7, IgG1 (mouse)	IM2475, Beckman Coulter	1/10
Mouse	c-KIT	APC	2B8, IgG2b/ κ (rat)	553356, BD Biosciences	1/250
	Ter119	PE	Ter-119, IgG2b/ κ (rat)	553673, BD Biosciences	1/200
	CD44	PE-Cy7	IM7, IgG1/ κ (rat)	560569, BD Biosciences	1/400
	isotype control	APC	A95-1, IgG2b/ κ (rat)	553991, BD Biosciences	1/250
	isotype control	PE	A95-1, IgG2b/ κ (rat)	553989, BD Biosciences	1/200
	isotype control	PE-Cy7	A95-1, IgG1/ κ (rat)	552849, BD Biosciences	1/400

Cell cycle analysis

Cell cycle was analysed by double labelling using a fluorescent anti-BrdU antibody and 7-aminoactinomycin D (7-AAD, BD Biosciences). In some experiments, cells were arrested in G1/S phase with aphidicolin (Sigma Merck, 0.6 μ g/mL) for 16h before labelling. Cells were incubated with 10 μ M BrdU for 30 min at 37°C (BD Biosciences). Then cells were pelleted, fixed in 500 μ l 70% ethanol for 20 min at room temperature and washed in 1X PBS supplemented with 0.5% bovine serum albumin (BSA). Cell pellet is resuspended in 2M chlorhydric acid for 20 min at room temperature and washed in 1X PBS. Acid was neutralized by a 0.1M borate solution pH 8.5 for 2 min (Borax Na₂B₄O₇, Merck). Cells were washed 3 times and transferred to 96-well plates for incubation with 10 μ l anti-BrdU antibody or isotype control for 20 min at room temperature (FITC-anti-BrdU 556028, BD Biosciences or

APC-anti-BrdU 51-23619L, BD Biosciences). Cells were washed 3 times and incubated with 50 μ L of a 10 μ g/mL 7-AAD and 50 μ g/mL RNase A (Macherey Nagel, Hœrdt, France) solution for 30 min at room temperature. Finally, cells were transferred into Eppendorf vials containing 100 μ L 1X PBS and analysed on a BD Accuri C6 flow cytometer (Becton-Dickinson) with the CFlow Plus software (BD Biosciences).

Immunofluorescence experiments

Immunofluorescence was used to detect phospho-RPA32 (p-RPA32, serine 4/8), phospho-H2AX (γ -H2AX, serine 139) and 53BP1 nuclear foci (**Supplementary table 3**). Cells treated with 5 mM hydroxyurea (HU) for 3h were used as positive controls (Merck). Briefly, cytopins were prepared with 10^5 cells in 1X PBS, fixed in 1X PBS/2% paraformaldehyde (Santa Cruz Biotechnologies, Santa Cruz, CA) for 20 min at room temperature and washed. Cells were permeabilized in 1X PBS/0.5% Triton X-100 for 10 min and washed in cold 1X PBS. Saturation was performed using 1X PBS/3% BSA (Euromedex, Souffelweyersheim, France) for 30 min and cells were incubated with primary antibodies for 1h at 37°C. After washing, cells were incubated with secondary antibody for 30 min at 37°C in the dark. DNA was counterstained with 4',6-diamidino-2-phenylindole (DAPI) for 5 min and cytopins were rinsed in 1X PBS and mounted with Fluoromount-G (Clinisciences, Nanterre, France). Analysis was conducted using an inverted DMI600 microscope at 100X magnification (Leica, Wetzlar, Germany). Images were analysed using the ImageJ software (NIH, Bethesda, MD).

Supplementary Table 3: IF antibodies and reagents

Antibody	Fluorochrome	Characteristics	Reference	Dilution
p-RPA32, S4/8		Polyclonal IgG (rabbit)	A300-245A, Bethyl Laboratories	1/250
γ -H2AX, S139		Polyclonal IgG (rabbit)	25775, Cell Signaling Technology	1/250
53BP1		Polyclonal IgG (rabbit)	NB100-904, Novus Biologicals	1/500
Anti-rabbit	DyLight 488	Polyclonal IgG (donkey)	A120-108D2, Bethyl Laboratories	1/200
DAPI		DNA staining	D1306, Thermo Fisher Scientific	1/1000

RNA-sequencing

cDNA synthesis was conducted with MuLV Reverse Transcriptase (Invitrogen, Carlsbad, CA), with quality controls conducted on an Agilent 2100 Bioanalyzer (Agilent, Les Ulis, France). Libraries were constructed using the TruSeq Stranded mRNA Sample Preparation Kit (Illumina, San Diego, CA) and sequenced on an Illumina HiSeq 2500 platform using a 100-bp paired-end sequencing strategy. An average depth of global sequence coverage of 114 million and a median coverage of 112 million was attained. For analysis, TopHat (v2.0.6) was used to align the reads against the human reference genome Hg19 RefSeq (RNA sequences, GRCh37) downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>).

Bioinformatic analysis of RNA-sequencing

FASTQ files were mapped using TopHat (v2.0.6)²⁸ to align the reads against the human reference genome Hg19 RefSeq (RNA sequences, GRCh37) downloaded from the UCSC Genome Browser (<http://genome.ucsc.edu>). Read count normalizations and groups comparisons were performed using DESeq2, which converts the counts to log-counts per million with associated precision weights and flags were computed using a custom algorithm within R (The R project for Statistical Computing [<http://www.r-project.org/>]). Assuming that a maximum of 80% of genes are expressed we selected the 20% lowest counts for each sample as background. A threshold is fixed at two standard deviations over the mean of the background. All transcripts for which normalized counts were lower than the computed

threshold were designated as background for each array. When comparing normalized counts between two groups, a transcript was included in the analysis if its counts exceeded the background in at least 80% of the sample from at least 1 group. To identify differentially expressed transcripts, we used empirical Bayes estimation. For differential expression study, the results obtained after Deseq2 comparison were selected for further analysis and filtered at P -value ≤ 0.05 and fold-changes of 1.2, 1.5 or 2. Principal component analysis calculated from averaged RefSeq exons robust multiarray average (RMA) values by RefSeq gene symbol was generated using R and the GGplot2 package. The bivariate normal density is a function of the means and standard deviations of the X and Y variables, PC 1 and PC2. The ellipsoid was computed from the bivariate normal distribution fit to each group. To determine transcript with intron retention, we used SpliceSeq (Ryan et al., 2012). We selected introns retained in all *SF3BI*^{WT} samples or in all *SF3BI*^{MUT} samples.

Western blot analysis

Cell lysates were solubilized for 5 minutes at 95°C in Laemmli buffer (65 mM Tris [pH 6.8], 20% glycerol, 5% β -mercaptoethanol, 0.01% bromophenol blue, and 2% sodium dodecyl sulfate [SDS]) Proteins were separated by SDS–polyacrylamide gel electrophoresis and transferred to a nitrocellulose membrane (VWR, Radnor, PA). Membranes were blocked with 5% of dry milk in TBS-T buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.15% Tween 20) for 1h and incubated in specific antibody overnight at 4°C (**Supplementary table 4**). Membranes were washed in TBS-T buffer and incubated for 1h at RT with secondary horseradish peroxidase (HRP)-linked antibody (horse anti-mouse HRP-linked antibody 7076S or goat anti-rabbit HRP-linked antibody 7074V, Cell Signaling, Danvers, MA). Enzyme activity was visualized by an ECL-based detection system (10600002, VWR). Blot imaging was performed on the Fujifilm LAS-3000 Imager (Fujifilm, Tokyo, Japan) and images were analysed using the Multi Gauge software (Fujifilm).

Supplementary Table 4: Primary antibodies for Western blot

Antibody	Characteristics	Reference	Dilution
β -actin	IgG1 (mouse)	A1978, Sigma Merck	1/10.000
HSC70	IgG2a/ κ (mouse)	sc-7298, Santa Cruz Biotechnologies	1/5.000
p-CHK1 (S345)	Polyclonal IgG (rabbit)	2341, Cell Signaling Technology	1/1.000
p-RPA32 (S4/8)	Polyclonal IgG (rabbit)	A300-245A, Bethyl Laboratories	1/1.000
RPA32	12F3.3, IgG2a/ κ (mouse)	GTX70243, GeneTex	1/1.000
γ -H2AX, S139	Polyclonal IgG (rabbit)	25775, Cell Signaling Technology	1/1.000

DNA fiber combing

DNA replication was analyzed by pulse labelling with fluorescent thymidine analogs in CRISPR-Cas9 *Sf3b1*^{WT} (H9.82) and *Sf3b1*^{K700E} (H5.6) G1E-ER4 without or with β -estradiol 10^{-7} M (24h) for differentiation induction. Cells were first labelled with 20 μ M iododesoxyuridine (IdU) for 30 min at 37°C and then with 100 μ M chlorodesoxyuridine (CldU) for 30 min at 37°C. DNA replication was blocked by the addition of an excess of thymidine (300 μ M) on ice and cells were washed in 1X PBS and counted.

DNA combing was performed after DNA extraction in 1% low melting agarose plugs (0.3.10⁶ cells in 90 μ l per plug). Briefly, low melting agarose was maintained at 45°C. Cells were resuspended at 6.66x10⁶ in one volume of 1X PBS, mixed in the same volume of 2% agarose and distributed in plug mold. Then proteins in agarose plugs were digested with 1 mg/mL proteinase K in 0.25 M EDTA pH 8/ 1% SDS at 42°C for 48h. Finally, plugs were washed 3 times in 10 mM Tris-HCl pH 8 / 1 mM EDTA (TE), and agarose was eliminated by digestion using β -agarase for 48h. For DNA fiber stretching, extracted DNA was resuspended in 0.25M 2-(N-morpholino)-ethanesulfonic acid (MES) pH 5,5 in Eppendorf vials for 30 min at 65°C, 30 min at room temperature and 2 weeks at 4°C. Before combing, vials were placed at room temperature for 30 min. DNA was then placed in FiberComb reservoir (Genomic Vision, Bagneux, France) and stretched on silane treated coverslips. Coverslips were stuck on glass slides and incubated for 2h at 60°C and stored at -20°C. For hybridization, DNA was denatured in a 1N NaOH solution, rinsed in cold 1X PBS and dehydrated in 70%, 85% and

100% ethanol. Aspecific sites were blocked for 30 min at 37°C. DNA was first incubated with primary anti-IdU (mouse) and anti-CldU (rat) antibodies, and then with fluorescent secondary antibodies. Finally, DNA was labelled with a primary anti-ssDNA antibody (mouse), and with a first fluorescent secondary anti-mouse antibody (goat), and with a second fluorescent secondary anti-goat antibody (donkey). Slides were mounted in VectaShield medium (Vector Laboratories, Burlingame, CA). All antibodies and reagents were described in **Supplementary tables 5 & 6.**

Supplementary Table 5: DNA fiber combing antibodies

Antibody	Fluorochrome	Clone and Immunoglobulin class	Reference	Dilution
anti-BrdU (anti-IdU)		B44, IgG1/κ (mouse)	347580, BD Biosciences	1/5
anti-BrdU (anti-CldU)		BU1/75, IgG2a/κ (rat)	ab6326, Abcam	1/50
Anti-mouse (anti-IdU)	Alexa 488	Polyclonal IgG (goat)	A11029, Thermo Fisher Scientific	1/50
Anti-rat (anti-CldU)	Alexa 555	Polyclonal IgG (goat)	A21434, Thermo Fisher Scientific	1/50
anti-ssDNA		16-19, IgG2a/κ (souris)	MAB3034, Merck	1/25
Anti-mouse anti-ssDNA	Cy5*	Polyclonal IgG (goat)	A10524, Thermo Fisher Scientific	1/100
Anti-goat anti-ssDNA	Cy5	Polyclonal IgG (donkey)	ab6566, Abcam	1/100

Supplementary Table 6: DNA fiber combing reagents

Product	Reference
HU	H8627, Merck
IdU	I7125, Merck
CldU	C6891, Merck
Thymidine	89270, Merck
Agarose	A4018, Merck
50-Well Disposable Plug Molds	1703713, Bio-Rad
PK	3115844001, Merck
β -agarase	M0392, New England BioLabs
MES	M8250, Merck
Silane treated coverslips	COV-002-RUO, Genomic Vision
Blocking solution	B10710, Thermo Fisher Scientific

R-ChIP qPCR

These experiments will be performed in collaboration with Dr F. Chedin, UC Davis, CA.

Statistical analysis

For quantitative variables, values were expressed as median and interquartile range (IQR) or means and standard error of the mean (SEM) and compared using the Mann-Whitney test or the Student t-test. For transcript quantification, the Mann-Whitney test was used to assign a statistical significance for each group comparison. *P*-values <0.05 were considered significant (JMP version 10.0.2, SAS Institute Inc, Cary, NC).

Supplementary Table 7: Transcripts with one or more intron retention in *SF3B1*^{WT} MNC samples

ABCA2	ATP5G3	CCPG1	CYP4F12	ERRF1	GOLGA4L4	IREB2	MEIS1	NISCH	PHTF1	OSOX1	RTEL1-TNFRSF6B	STRADA	TRMT1	WDR62
ABC8	ATP1F1	CD7	DALRD3	EXOC3	GOLGA8M	IRF7	METTL7	NG7	PIDD	QTRT1	RUNX1	STX10	TRMT2A	WRAP53
ABC3	ATXN2L	CDCL6	DAPK2	EXOSC10	GPAM	ISCU	METTL23	NLR3	PIGO	RHDM4	RUSC1	SUGP2	TROAP	WRAP73
ABCD4	AURKAIP1	CD37	DCAF1	EXOSC9	GPS1	ISC20L2	METTL5	NME1-NME2	PIGQ	RAB24	SCAND1	SULT1A3	TSPAN17	XPO1
ABTB1	BHGA1T4	CD42SE2	DDIT3	EZH2	GSS	JMD17	MFSD10	NMRAL1	PIG3	RAB24	SCNN1D	SUPT7L	TTC14	YIP3
ACAA1	BAX	CDCA2	DDX11	FAH	GSTK1	JTB	MFSD11	NOD1	PIK3CD	RABGGTA	SCNNE1	TTC31	TTC31	YPEL3
ACBD4	BCKDK	CDCA5	DDX20	FAM111A	HAGHL	JUP	MGRN1	NOL12	PHRB	RANBP1	TAB3	TTC9C	TTC9C	YPEL4
ACBD5	BCL2	CDK10	DERL3	FAM114A2	HARS2	KAT5	MIEN1	NOP2	PLAGL1	RANGRF	TAF1C	TTC2	ZBTB11	ZBTB17
ACIN1	BCL2L1	CDK11A	DHRSL	HAUS5	KAT8	KAT8	MOAP1	NP1A5	PLCB2	RASGRP2	SEC24C	TBC1D3F	TTL3	ZBTB17
ACPI	BEST1	CDK2	DIAALO	FAM131A	HCF1R1	KCNAB2	MRPL23	NP1PB3	PLSRI	RASGRP4	SELPLG	TCEA2	TXNRD2	ZBTB44
ACTG1	BETL	CDK5RAP3	DIS3L2	FAM195B	HDAC10	KDM3A	MRPL27	NP1PB4	PML	RBM19	SERPINA1	TCF3	UBALDI	ZCCHC8
ADCY4	BSDC1	CDKN1C	DMAPI1	FAM214B	HDAC6	KIAA0907	MRPL35	NP1PB5	PMPCB	RBM48	SETD4	TEFM	UBB	ZCCHC9
ADPRH	BTBD3	CDKN2D	DNAJB12	FAM219B	HDDC3	KLC2	MRPL52	NP1L2	PNISR	RBM5	SCSM2	TET2	UBBP4	ZDHHC4
AKAP17A	C11orf48	CELF2	DNAJB2	FAM43A	HERPUD2	KLF7	MRPL55	NR1H3	POUT2	RBM6	SCSM3	TEX30	UBC	ZFAND2B
AKT1	C12orf73	CGBP1	DNAJC4	FAM50B	HES6	KLHDC2	MRPS12	NSUN5	POCK	RBMX	TTEC	TTEC	UBE2N	ZFC3H1
AMBR1	C16orf58	CHCHD1	DNPH1	FANCA	HEXDC	KLHL21	MRP57	NUDT16	POLL	RECQL	SIGIRR	THAP7	UBE3B	ZMYM2
AMDHD2	C16orf59	CHCHD7	DPAGT1	FASTK	HINT1	L2HGDH	MS4A6A	NUDT18	POLM	REM2	SLC15A4	THYNI	UBXN1	ZNF131
ANAPCS	C16orf91	CHORDC1	DRG2	FBLX12	HUTF	LAMTOR5	MSRB1	NUDT22	POLR2G	REPS1	SLC25A14	TM2D3	UBXN4	ZNF142
ANKLE2	C16orf93	CHTOP	DTX3	FBLX8	HMBS	LARP1B	MSTO1	NUDT8	POLR2J2	REVI	SLC25A29	TMC6	UCKL1	ZNF266
ANKRD13D	C18orf21	CIRBP	DUS4L	FBXO3	HNRNPA1	LCK	MTERFD3	NUPI60	POLR2J3	RHOT2	SLC25A3	TMCO6	ULK3	ZNF276
ANKRD49	C19orf54	CLDN15	DUSP22	FBXW7	HNRNPA2B1	LEPRE1	MTMR10	NUP62	POMGN1	RIFI	SLC39A1	TMEM138	UNC119	ZNF286A
ANKRD9	C19orf66	CLEC2D	DUSP28	FGD3	HNRNPAB	LEIM2	MTRF1	NUP85	POR	RIFI	SLC39A13	TMEM144	UNC50	ZNF426
AP1G2	C1orf50	CLN3	DYNLL1	FGD3	HNRNPH1	LIN37	MXD3	OGFR	PPAPDC1B	RNASEH2C	SMARCD3	TMEM161B	UNC93B1	ZNF511
AP4B1	C21orf59	CNIH2	DYNLT1	FLAD1	HNRNPL	LPCAT3	MYL6	OGG1	PPIL2	RNF166	SMC4	TMEM205	UOQRB	ZNF576
APH1A	C22orf29	CNTRL	DZIP3	FNI	HNRNPM	LPCAT3	NAAL6	ORMDL1	PPM1M	ROMO1	SMN1	TMEM234	UOQCRQ	ZNF646
APOBR	C2orf49	COA6	EDF1	FOXJ3	HPS4	LRSAM1	NAA60	OS9	PPP1R27	ROPN1L	SMOX	TMEM256-PLSCR3	USP20	ZNF692
AQR	C2orf68	COG8	EED	FXYD5	HRA5	LSMD1	NABP2	OSGEP	PPP2R3C	RPA2	SNX13	TMEM41B	USP36	ZNF711
ARFRP1	C9orf89	COMMD7	EEF1B2	GABARA1	HSH2D	LTBR2	NAGA	PABPN1	PKRARI	RPCR	SNX20	TMEM45B	VAMP2	ZNF76
ARHGAP32	CALCOCO1	COQ6	EJF2B1	GBA2	HSPB1	LUC7L	NAPLL1	PAFAH1B3	PRMT2	RPL10	SNX5	TMEM91	VAMP4	ZSCAN32
ARHGAP33	CARKD	CORO7	EJF3C	GGA3	HTATIP2	LYPLA2	NAT14	PAPOLA	PSMC3IP	RPL29	SOD2	TMUB2	VIM	ZSWIM7
ARM5	CARKD	COX7C	EJF4A1	GNS4	HUS1	MAF	NAT9	PAQR6	PSMC5	RPL32	SPI40L	TNFRSF10B	VIPAS39	ZWINT
ARMCX5	CASP1	CPT1B	ELMOD3	GLJ4	ICOSLG	MAFI	NBP11	PARP2	PSMDJ1	RPL36	SRSF1	TOP3B	VPS16	
ARPP19	CBWD1	CREB3L4	ELOVL6	GLYR1	IDI	MAP7D1	NCDN	PCGF3	PSME1	RPRD2	SRSF11	TOR2A	VPS28	
ARRDC1	CC2D1B	CSAD	ELF5	GMPFB	IDH3G	MARS	NCF4	PCK2	PSMG4	RPS15	SRSF2	TP53I3	VPS8	
ARS8	CCAR2	CSTF5	EMC9	GMPPA	IFRD2	MAX	NDUF52	PSCK7	PSRC1	RPS2	SRSF7	TRAPPC1	WASH4P	
ASB8	CCDC107	CTNNB1	ENDOV	GMPPB	IL32	MBD3	NDUF57	PDCD2	PTAR1	RPS20	SSH2	TREX1	WDR11	
ATF4	CCDC125	CXorf40A	ENG	GMPR2	ILK	MBTD1	NEK1	PFKM	PTOVI	RPS6K1	SSH3	TRIM13	WDR13	
ATG16L2	CCDC14	CYB561D2	EP400NL	GMPR2	ING3	CMC7	NEU3	PCM2	PTRH2	RPS9	SSSCA1	TRIM16	WDR33	
ATOX1	CCDC28B	CYB561D2	ERC5	ING5	INPP5F	MDM1	NEF2	PHF12	PTTGI	RPS9	ST3GAL1	TRIM41	WDR55	
ATP5D	CCNT2	CYP4F12	ERLN2	GNPDA1	IP6K2	MED17	NFKB1B	PHF6	QKI	RRP8	STRK40	TRIM5	WDR6	

Supplementary Table 8: Transcripts with one or more intron retention in *SF3B1*^{MUT} MNC samples

ABCA2	BETIL	CGGBP1	DUSP22	GAS2L1	ICOSLG	MED19	NMRAL1	PIGS	RASSF7	SMC4	TMEM41B	WDR6
ABCC5	BTBD6	CHCHD7	DYNLT1	GBA2	IDI	METTL7	NOL2	PIK3CD	RBM19	SMN1	TMEM55B	WDR62
ACA1	C11orf24	CHORDC1	EDF1	GGA3	IDH3G	METTL23	NOP2	PLACL1	RBM25	SMUG1	TMEM91	WDR62
ACBD5	C11orf48	CHTOP	EED	GINS4	IL18BP	METTL5	NPIPA5	PLCB2	RBM48	SNAPC3	TNFRSF10B	WRAP53
ACIN1	C12orf73	CIRBP	EEF1B2	GL4	IL32	MFSDD10	NPIPB3	PLSCR1	RBM5	SNW1	TOP3B	WRAP73
ACP1	C16orf58	CLDN15	EEF2B1	GLYR1	ILK	MFSD11	NPIPB4	PML	RBM6	SNX13	TOR2A	WSBI
ADCY4	C16orf59	CLEC2D	EIF3C	GMFB	ING3	MEN1	NPIPB5	PNPCB	RBMX	SNX5	TP53	YME1L1
ADKAP1A	C16orf91	CLN3	ELMOD3	GMPPA	INPP5F	MOA1	NPRL2	PNISR	REC8	SOD2	TRAPPC1	YPEL4
AMBRA1	C16orf93	CNTRL	ELOVL6	GMPPB	IREB2	MIRPL23	NSD1	POFUT2	RECQL	SP140L	TRIM41	YY1A1
AMDHD2	C17orf58	COASY	ELP5	GMPR2	IRF3	MIRPL35	NSUN5	POFUT2	REVI	SPATA20	TRIM5	ZBTB11
AMT	C18orf21	COCH	ELP5	GMPR2	IRF7	MIRPL52	NUDT16	POGK	RHOT1	SPRTN	TRMT2A	ZBTB44
ANKLE2	C21orf58	COMMID7	EMC9	GNG5	ISCU	MIRPL55	NUDT16L1	POLL	RHOT2	SRSF1	TROAP	ZCCHC8
ANKRD13D	C22orf29	CPT1B	ENG	GOLGA6L4	ISC20L2	MIRPS12	NUDT22	POLR2G	RIF1	SRSF11	TSPAN17	ZCCHC9
ANKRD49	C2orf49	CRELD1	ERCC5	GFAM	JMJD7	MS4A6A	NUDT8	POLR2J2	RNASEH2C	SRSF2	TTC14	ZDHHC4
ANKRD9	C2orf68	CSAD	ERLIN1	GFS1	JUP	MSTO1	NUPT60	POLR2J3	RNF166	SRSF7	TTC31	ZFAND2B
APIG2	C5orf24	CTNNB1	ERLIN2	GSS	KAT5	MTMR10	NUP62	POR	RPA2	SSH2	TTC9C	ZFC3H1
AP4B1	C9orf89	CXorf38	EXOC3	GSTK1	KAT8	MTRF1	NUP85	PPAPDC1B	RPGR	SSH3	TTI2	ZMYM2
APBB3	CALCOCO1	CXorf40A	EXOSC10	HAGHL	KCNAB2	MXD3	NXF1	PPIL2	RPL32	SSSCA1	JTLL3	ZNF131
APH1A	CARKD	CYB561D2	EXOSC9	HARS2	KDM3A	MYL6	OGFR	PPM1M	RPL36	ST3GAL1	UBB	ZNF131
APOBR	CBWD1	CYP4F12	EZH2	HAU5	KIAA0141	NA-A60	OGGI	PPP2R3C	RPL7L1	STRADA	UBBP4	ZNF266
APOL3	CC2D1B	CYP4F12	FAH	HCFICR1	KIAA0907	NABP2	ORMDL1	PRDM15	RPRD2	SUGP2	UBC	ZNF276
AQR	CCAR2	D2HGDH	FAM111A	HDAC10	KLC2	NA-GA	OSBP2	PRMT2	RPS15	SULT1A3	UBE2N	ZNF426
ARHGAP33	CCDC107	DALRD3	FAM114A2	HDAC6	KLHDC2	NAP1L1	OSGEP	PRRC1	RPS20	SUPT7L	UBE3B	ZNF511
ARMC5	CCDC125	DCAF11	FAM122C	HERPUD2	KLHL21	NAPRT1	PABPN1	PSMC3IP	RPS6K1	SYNE1	UBXN4	ZNF576
ARPP19	CCDC14	DDIT3	FAM214B	HES6	LANCL1	NAT14	PAFAH1B3	PSMC5	RPS9	TAB3	UCKL1	ZNF646
ARRDC1	CCDC28B	DDX11	FAM219B	HEXDC	LARPIB	NAT9	PAPOLA	PSMD11	RRNAD1	TAF1C	UNC119	ZNF692
ARSA	CCNT2	DDX20	FAM32A	HINT1	LEPRE1	NBPF11	PARP2	PSME1	RUSC1	TCEA2	UNC50	ZNF76
ASB8	CCPG1	DERL3	FAM3A	HLTF	LIN37	NCAHP2	PCGF3	PSRC1	SCRN2	TCF3	UQCRB	ZSCAN12
ATG16L2	CD7	DHRS1	FAM50B	HMSB	LPCAT3	NCDN	PCGF3	PSTK	SCYL1	TCIRG1	USP20	ZSCAN32
ATP5D	CDC16	DHRS12	FANCA	HNRNPA1	LSMD1	NCF4	PCID2	PTAR1	SEC16A	TEFM	USP36	ZSCAN32
ATP5G3	CDC37	DHRS13	FASTK	HNRNPA2B1	LTB4R2	NDUFS7	PCK2	PTOVI	SEC24C	TET2	VAMP2	ZSCAN32
ATXN2L	CDC42SE2	DIABLO	FBXO3	HNRNPH1	LUC7L	NECAB3	PCSK7	PTRH2	SGSM2	THAP7	VAMP4	ZSCAN32
AURKAIP1	CDCA2	DIS3L2	FBXW7	HNRNPL	LYPLA2	NEK1	PDCD2	PTTG1	SGSM3	THPA	VPS16	ZSCAN32
B4GALT4	CDCA5	DIS3L2	FGD3	HNRNPM	MAF1	NEU3	PFKM	QKI	SH2B1	THYNI	VPS28	ZSCAN32
BAX	CDK10	DMAPI1	FGD3	HPS4	MAGED2	NFE2	PGM2	QKI	SIDT2	TM2D3	VPS8	ZSCAN32
BCKDHB	CDK11A	DMPK	FIGLN1	HRA5	MARS	NFKB1B	PHF12	QSOX1	SIGIRR	TM6	WASH4P	ZSCAN32
BCKDK	CDK2	DNAJB12	FLAD1	HSH2D	MAX	NGDN	PHF6	QTRT1	SLC25A14	TMCO6	WDR11	ZSCAN32
BCL2	CDK5RAP3	DNAJB2	FOS	HSPB1	MBD1	NISCH	PIDD	R3HDM4	SLC25A3	TMEM138	WDR13	ZSCAN32
BCL2L1	CDKN1C	DNAJC4	FOXJ3	HITATIP2	MBTD1	NKG7	PIGO	RABGGTA	SLC39A13	TMEM161B	WDR33	ZSCAN32
BCS1L	CEP57	DRG2	FXYS5	HUS1	MDM1	NME1-NME2	PIGQ	RANGRF	SMARCD3	TMEM205	WDR55	ZSCAN32

Discussion

Les mutations de *SF3BI* confèrent un caractère oncogénique aux CSH des patients porteurs d'un SMD et participent à la mise en place d'une érythropoïèse clonale. Nos premiers résultats démontrent que les érythroblastes primaires humains *SF3BI*^{K700E} et les lignées érythroblastiques murines G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E} ont une dynamique de réplication de l'ADN plus intense que les cellules non mutées, ce qui se traduit par une augmentation de la prolifération cellulaire. L'accélération de la progression des fourches de réplication dans les Pro-E et Baso-E murins *Sf3b1*^{K700E} s'accompagne d'une exposition d'ADNsb comme en témoigne la détection de foyers de protéine RPA32 phosphorylée (p-RPA32). Les érythroblastes primaires humains *SF3BI*^{MUT} accumulent eux-aussi des foyers intranucléaires de p-RPA32, et ce stress réplicatif n'aboutit pas à la génération de cassures double brin de l'ADN.

1. Quel est le mécanisme permettant une accélération de la vitesse des fourches ?

Nos résultats sont en contradiction avec les travaux portant sur les facteurs d'épissage SRSF2 et U2AF1. En effet, les cellules HEK 293T *SRSF2*^{MUT} ou *U2AF1*^{MUT} prolifèrent moins que les cellules sans mutation, et ont une diminution marquée de la longueur des fibres liée, selon les données R-ChIP, à l'accumulation de *R-loops*. Pour rappel, les *R-loops* sont des intermédiaires de transcription contenant un hybride ADN/ARN et de l'ADN simple brin. (cf. Chapitre II-Introduction 2-II) (Chen et al., 2018).

Dans notre modèle, l'accélération de la vitesse des fourches suggère que ces dernières progressent en rencontrant peu d'obstacles ou que la cellule mutée est capable de résoudre ou contourner de potentiels obstacles. Cette observation nous amène à poser l'hypothèse qu'à l'inverse des cellules *U2AF1*^{MUT} et *SRSF2*^{MUT}, les cellules *SF3BI*^{MUT} n'accumulent pas de *R-loops*, puisque les *R-loops* en excès induisent un ralentissement et un arrêt des fourches de réplication dans les cellules des organismes procaryotes et eucaryotes (Gan et al., 2011; Parajuli et al., 2017).

Quels sont les arguments en faveur d'une telle hypothèse ? Dans une cellule normale, l'expression génique peut être régulée négativement au niveau post-transcriptionnel par rétention d'intron, et notamment par détention d'intron lorsque les transcrits restent dans le noyau après épissage (pour revue Fu, 2017). En particulier, le programme de détention

d'intron est différent dans les lignages érythroïde et mégacaryocytaire et entre ces lignages et le progéniteur commun érythro-mégacaryocytaire (MEP) suggérant un rôle majeur de la régulation post-transcriptionnelle d'expression génique par épissage alternatif pour l'engagement vers l'une ou l'autre des voies de différenciation (Edwards et al., 2016).

Ces observations de la littérature posent les questions suivantes : 1) La protéine SF3B1 a-t-elle, comme SRSF2, un rôle dans le contrôle de la transcription de l'ADN ? ; 2) Si tel est le cas, quel est le niveau de transcription de l'ADN dans les cellules *SF3B1*^{MUT} ? ; 3) Quel est l'impact des mutations de *SF3B1* sur l'expression des transcrits à introns détenus ? Des données massives de RNA-Seq obtenues à partir de cellules mononucléées de la moelle osseuse ou de progéniteurs CD34⁺ montre que l'épissage est accru avec une diminution de la fréquence des rétentions d'intron dans le contexte des mutations de *SF3B1* (Shiozawa et al., 2018). Nous supposons donc que, dans les cellules *SF3B1*^{MUT}, les transcrits épissés sans rétention d'intron devraient avoir une moindre capacité à s'hybrider avec la matrice d'ADN et à former des *R-loops*, obstacles naturels à la réplication.

Pour explorer cette hypothèse, deux types d'expérience sont nécessaires : 1) L'évaluation du niveau de transcription de l'ADN qui devra être comparé aux différents stades de la maturation érythroïde des échantillons mutés et non mutés, puisqu'une diminution globale de l'expression génique est observée au cours de l'érythropoïèse physiologique (cf. Chapitre I-I-D-1). Pour cela, la détection de l'ARN nouvellement synthétisé sera réalisée par incorporation dans l'ARN d'un nucléoside modifié (5-éthynyl-uridine) dans les cellules en culture suivie d'une réaction dite « click », c'est-à-dire la ligation d'un azide fluorescent avec le nucléoside modifié et 2) L'analyse des *R-loops* dans les cellules de SMD *SF3B1*^{MUT}. Nous avons engagé une collaboration avec une équipe américaine spécialiste des *R-loops* et de la technique de DRIP-qPCR (pour *DNA/RNA Immunoprecipitation-qPCR*). La DRIP-qPCR permet la détection des *R-loops*, par exemple dans des régions promotrices de gènes cibles. Elle combine deux technologies : l'immunoprécipitation des hybrides ARN/ADN dans les extraits d'ADN génomique (ADNg) grâce à l'anticorps S9.6, puis la PCR quantitative des ADN immunoprécipités. L'anticorps S9.6 reconnaît, en plus des hybrides ARN/ADN, les ARN double-brin. Toutefois, le traitement des extraits d'ADNg avec des nucléases permet d'éliminer les ARN contaminants, ce qui rend la technique de DRIP-qPCR plus robuste que l'IF. L'identification des gènes d'intérêt sera réalisée à partir des données de RNA-Seq d'érythroblastes primaires de patients porteurs d'un SMD avec ou sans mutation du gène

SF3B1, par la mise en évidence des transcrits avec rétention d'intron dans les cellules sauvages ayant perdu l'intron détenu dans les cellules mutées.

2. La vitesse accrue de progression des fourches est-elle liée à un nombre augmenté d'origines de réplication activées ?

L'augmentation de l'incorporation du BrdU visualisée dans les expériences de cycle cellulaire peut s'expliquer par une augmentation de la vitesse de progression des fourches ou du nombre d'origines de réplication activées. Pour trancher entre ces deux hypothèses, la longueur des fibres d'ADN sera mesurée dans un réplicat biologique dans les clones G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E} et *Sf3b1*^{WT}, pour acquérir un plus grand nombre de données. Le taux d'initiation de la réplication de l'ADN (nombre d'origines activées/longueur des fibres), sera calculé à partir des mesures des distances inter-origines (DIO) : plus la DIO est longue, moins il y a d'activation d'origines de réplication (Maya-Mendoza et al., 2018). Le taux d'initiation est normalement inversement corrélé à la vitesse de progression des fourches de réplication. En cas de ralentissement de la fourche, les origines de réplication dites « dormantes » sont allumées. *A contrario*, l'augmentation du nombre d'origines de réplication activées entraîne un ralentissement des fourches puisque les réplisomes entrent en compétition pour l'utilisation d'une quantité limitante de dNTPs. Si le nombre d'origines de réplication activées diminue, la vitesse de progression des fourches augmente en raison de l'augmentation de la disponibilité en dNTPs. Par exemple, dans la Levure, la perte d'activité de la kinase cycline-dépendante Cdc7 – enzyme indispensable à l'allumage des origines de réplication – induit une augmentation de la vitesse des fourches de réplication (Zhong et al., 2013).

3. Quel est le lien entre stress répliatif et oncogénicité des mutations récurrentes des gènes d'épissage ?

Le stress répliatif se définit habituellement comme un ralentissement voire un arrêt des fourches de réplication, l'exposition anormalement prolongée d'ADNsb pouvant induire la formation de CDB secondaires sources d'instabilité génomique.

Les travaux de Chen et al. (2018) attribuent l'oncogénicité des mutations des gènes *SRSF2* ou *U2AF1* à l'apparition de CDB secondaires au ralentissement des fourches. Ces CDB pourraient favoriser l'acquisition de nouvelles mutations et l'évolution leucémique.

Cependant, les auteurs soulignent le fait que les cellules leucémiques myéloïdes accumulent très peu de mutations dans leur génome comparativement aux tumeurs solides. De plus, dans les modèles cellulaires utilisés HEK 293T *SRSF2*^{MUT} ou *U2AF1*^{MUT}, la voie ATM/CHK2 n'est pas activée suggérant que les CDB marquées par l'anticorps dirigé contre la protéine peu spécifique γ -H2AX ne seraient qu'un phénomène marginal. Dès lors il est difficile d'expliquer le caractère oncogénique des mutations *SRSF2* ou *U2AF1* par ce mécanisme.

En 2018, Maya-Mendoza *et al.* ont démontré que l'accélération aberrante des fourches de réplication dans les cellules U2OS traitées avec un inhibiteur des poly(ADP-ribose) polymérasés (PARP) conduit aussi à l'accumulation de foyers p-RPA32 et à l'activation de la voie de réparation des dommages de l'ADN ATR/CHK1 (Maya-Mendoza *et al.*, 2018). Pour information, les PARP-1 et -2 sont activées lors de cassures simple-brins et double-brins et catalysent l'addition d'une chaîne de poly(ADP-ribose) (PAR) sur des protéines ciblées, tels que les histones H1 et H2B. La PARylation des histones permet la réorganisation de la chromatine et le recrutement des protéines de réparation de l'ADN au niveau des sites de cassures (pour revue Ciccia and Elledge, 2010). Nos résultats se rapprochent de ces travaux puisque nous observons une augmentation des foyers p-RPA32 et une accélération des fourches de réplication dans les précurseurs érythroïdes immatures G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E}.

La réplication nécessitant l'ouverture de l'ADN double-brin parental, on peut supposer que la perte de contrôle de la vitesse de progression des fourches induit une exposition d'ADNsb. Seule la voie ATR/CHK1 devrait être activée dans ces conditions – et pas la voie ATM/CHK2 qui est déclenchée par la présence de CDB –, ce que nous vérifierons par de nouvelles expériences d'immuno-empreinte.

RPA est le senseur et le protecteur de l'ADNsb (Nguyen *et al.*, 2017; Zeman and Cimprich, 2014; Zou and Elledge, 2003). Il se fixe sur l'ADNsb et active les *checkpoints* de la phase S. De plus, RPA interagit avec la RNase H1 nucléaire – l'enzyme qui hydrolyse les ARN des hybrides ARN/ADN – et stimule son activation (cf. Chapitre II-Introduction 2-I-2-a) (Nguyen *et al.*, 2017). Ces données soulignent le rôle central de RPA dans le processus de réplication de l'ADN, sa disponibilité étant le facteur limitant. Nos données de transcriptomique établissent que l'expression des protéines RPA ainsi que celle des hélicases MCM2-7 est augmentée dans les CMN avec mutation *SF3B1* suggérant une adaptation au stress répliatif (cf. Annexe 4). Il résulterait de ce mécanisme adaptatif une protection à

l'égard des dommages à l'ADN. Ceci pourrait expliquer la rareté des événements oncogéniques secondaires et le risque modéré de transformation leucémique comparativement aux SMD avec mutation *SRSF2* ou *U2AF1*. Enfin, le stress répliatif des érythroblastes avec mutation *SF3B1* se distinguent de celui décrit dans les progéniteurs hématopoïétiques de la souris âgée, dans lesquels un ralentissement des fourches par déficit en protéines impliquées dans l'activation des origines de réplication – telles que les hélicases MCM – a été décrit (Flach et al., 2014).

En conclusion, la présence d'une mutation *SF3B1* contribuerait au phénotype de l'anémie sidéroblastique en induisant une réplication accrue de l'ADN et donc en soutenant l'amplification du compartiment érythroblastique.

Conclusion générale

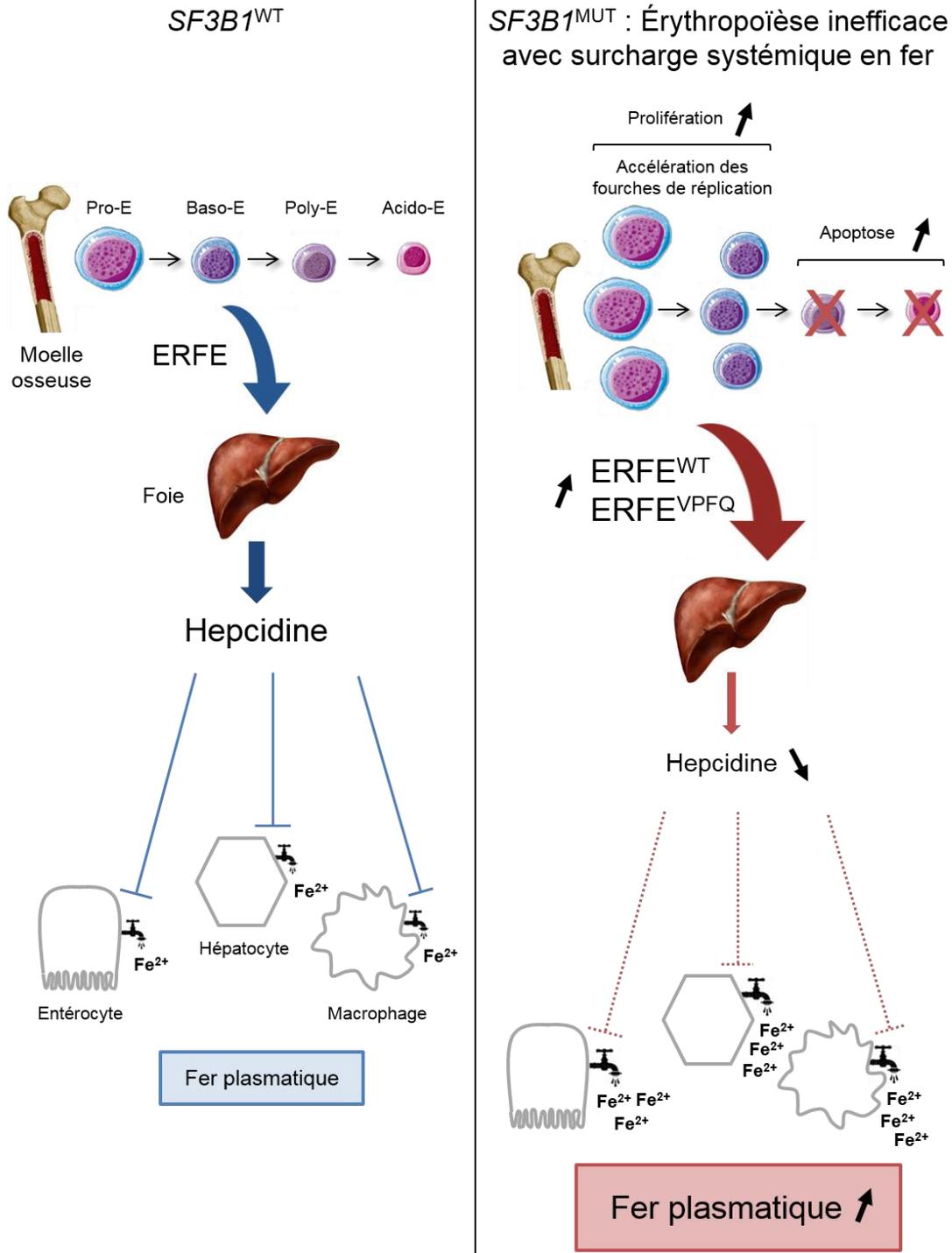


Figure 37 : Conséquences fonctionnelles des mutations du gène $SF3B1$ dans les SMD-SC

L'augmentation de la prolifération des précurseurs érythroïdes immatures (Pro-E et Baso-E) SMD-SC $SF3B1^{MUT}$ s'accompagne d'une accélération de la progression des fourches de réplication avec exposition d'ADNs b mais sans formation de CDB de l'ADN. Les précurseurs érythroïdes $SF3B1^{MUT}$ surexpriment les transcrits ERFE canonique et alternatif, respectivement traduits en protéines ERFE^{WT} et ERFE^{VPFQ}. Ces deux protéines contribuant à diminuer l'expression de l'hépcidine au niveau hépatique, les patients SMD $SF3B1^{MUT}$ développent *in fine* une surcharge systémique en fer. Les érythroblastes matures (Poly-E) $SF3B1^{MUT}$ meurent massivement par apoptose – notamment à cause de l'accumulation de fer dans les mitochondries –, d'où finalement une érythropoïèse inefficace.

Discussion et perspectives

L'objectif principal de ma thèse était de mieux comprendre les mécanismes oncogénétiques des mutations du gène *SF3B1* dans les SMD-SC, puisqu'actuellement aucun lien n'est clairement établi entre ces dernières et l'initiation d'un processus pré-leucémique. Dans un premier temps, nous avons montré que les mutations *hotspot* du gène *SF3B1* dans les SMD sont associées, comme dans les cellules de MU ou de LLC, à l'utilisation préférentielle de sites 3' accepteurs alternatifs. D'après les études précédemment publiées, les transcrits cryptiques ainsi générés sont ensuite soit dégradés dans le cytoplasme par NMD, soit traduits – avec plus ou moins d'efficacité – en des variants protéiques. Selon le domaine protéique où se situe l'ajout d'acides aminés, les variants peuvent changer ou perdre leur fonction et/ou leur stabilité. Finalement, les anomalies d'épissage participent à la modification de l'expression du gène canonique. Nous nous sommes alors demandé si les mutations de *SF3B1* pouvaient, par l'intermédiaire des épissages aberrants, participer au phénotype particulier des SMD-SC, à savoir une surcharge systémique en fer. Nous avons identifié une jonction alternative dans le pré-ARNm du gène *FAM132B/ERFE* codant pour un régulateur négatif de l'expression de l'hepcidine. Le transcrypt aberrant $ERFE^{+12}$ est traduit en un variant protéique $ERFE^{VPFQ}$ qui, avec la protéine canonique, contribue à la surexpression d'ERFE chez les patients porteurs d'un SMD $SF3B1^{MUT}$. Tout comme la protéine ERFE canonique, $ERFE^{VPFQ}$ diminue l'expression de l'hepcidine et est donc, *in fine*, responsable de l'augmentation de la biodisponibilité du fer.

1. L'érythroferrone en tant que cible thérapeutique

Le transcrypt ERFE alternatif – marqueur de l'érythropoïèse clonale dans les SMD-SC – n'a été identifié que dans nos données de RNA-Seq. La faible expression, voire quasi-absence d'expression, du gène *ERFE* dans les tissus non érythroïdes explique l'absence d'expression de l'AG' d'ERFE dans les cellules de LLC ou de MU $SF3B1^{MUT}$. En revanche, pourquoi l'AG' d'ERFE n'a-t-il pas été identifié dans les autres études de RNA-Seq portant sur les SMD ? Plusieurs hypothèses peuvent être envisagées : 1) dans les RNA-seq de progéniteurs $CD34^+$, l'absence d'expression de l'AG' d'ERFE peut être attribuée à la très faible expression du gène *ERFE* dans ces cellules. En effet, dans les érythroblastes murins et humains, l'expression du transcrypt *ERFE* n'apparaît que progressivement au cours de l'érythropoïèse, pour être maximale aux stades Baso-E et Poly-E ; 2) dans les RNA-seq d'échantillons de CMN, la probabilité de détecter l'AG' d'ERFE est en théorie plus grande étant donné la richesse érythroblastique des moelles osseuses de patients porteurs d'un SMD-SC.

Néanmoins, le gène *ERFE* et donc les jonctions alternatives restent faiblement exprimés par rapport à de nombreux autres gènes. Notre pipeline d'analyse a permis de l'identifier grâce aux filtres que nous avons appliqués.

La régulation de l'expression de l'hepcidine fait intervenir divers facteurs (cf. Chapitre I-I-E-4-b-ii). Dans le cas des anémies profondes – comme c'est souvent le cas chez les patients atteints d'un SMD-SC – il devient difficile de dissocier les effets des régulateurs érythroïdes produits au sein de la moelle osseuse des effets des facteurs produits par les tissus hypoxiques. À titre d'exemple, en cas d'hypoxie, le facteur de transcription HIF-1 α stabilisé induit l'expression de la furine, la protéase responsable du clivage de l'HJV membranaire en HJV soluble (Silvestri et al., 2008b). L'HJV membranaire est le co-récepteur hépatique du récepteur aux BMPs (cf. Chapitre I-I-E-4-b-ii). La liaison des BMPs 2 et 6 au complexe récepteur aux BMPs active les voies de signalisation SMAD1/5/8 et, donc, la transcription du gène *HAMP* dans les hépatocytes. D'après les travaux de l'équipe de Tomas Ganz, l'HJV soluble inhibe l'expression de l'hepcidine *in vitro* (Lin et al., 2005). Ces observations ont par la suite été confirmées dans des modèles *in vivo* (Babitt et al., 2007). Pour justifier l'effet opposé de l'HJV soluble par rapport à l'HJV membranaire sur l'expression de l'hepcidine, Lin *et al.* proposent un mécanisme de compétition entre ces deux protéines vis-à-vis d'un ligand encore non identifié. Ainsi, l'HJV soluble pourrait, elle aussi, participer à la mise en place d'une surcharge systémique en fer chez les patients anémiés. Toutefois, l'association entre hypoxie tissulaire et génération d'une HJV soluble par induction de l'expression de la furine n'a pour le moment été démontrée que dans la lignée cellulaire HeLa transfectée avec un ADNc de l'HJV. De plus, Silvestri *et al.* n'ont pas mesuré la diminution exacte de la pression partielle en O₂ dans les cellules après traitement au CoCl₂, un inhibiteur des PHD. Peut-on alors comparer les conditions de l'hypoxie artificiellement induite dans des cellules cultivées *in vitro* avec celles de l'hypoxie des patients anémiés ? Enfin, pour rappel également, la surcharge systémique en fer chez les patients atteints d'un SMD-SC se développe de façon précoce, c'est-à-dire avant la mise en place d'une anémie trop profonde (hémoglobine < 8 g/dl), ce qui modère le potentiel rôle de l'hypoxie tissulaire. En revanche, l'hyperplasie du compartiment érythroïde est un événement précocoe responsable de la surproduction d'ERFE ; donc indépendamment de l'hypoxie, cette surproduction d'ERFE est probablement le déterminant de l'hyperferritinémie au diagnostic comme nous le suggérons par nos travaux.

La diminution de l'expression d'ERFE⁺¹² permettrait-elle de rétablir des taux physiologiques d'hepcidine plasmatique ? Pour répondre à cette question, nous avons décidé d'explorer les effets de trois oligonucléotides anti-sens (OAS) dirigés spécifiquement contre le transcrite ERFE⁺¹². Les OAS sont des petits fragments d'ADNs (\approx 15-25 pb) complémentaire de l'ARNm et/ou du pré-ARNm ciblé. Ils entrent dans la cellule par endocytose, puis forment un hétéroduplex ARNm/OAS ou pré-ARNm/OAS dans le cytosol ou le noyau respectivement. Les OAS diminuent l'expression du gène ciblé selon divers mécanismes : 1) dégradation de l'ARNm par les RNases H qui reconnaissent les hybrides ARN/ADN ; 2) encombrement stérique empêchant l'assemblage des sous-unités du ribosome au niveau de l'ARNm ; 3) inhibition de la formation de la coiffe en 5' du pré-ARNm et 4) inhibition de l'épissage du pré-ARNm (pour revue Chan et al., 2006). Nous disposons d'un LNA-OAS (pour *locked nucleic acid*, acide nucléique bloqué) et de deux morpholino-OAS (Annexe 6). Les modifications structurelles de type LNA ou morpholino augmentent la stabilité des OAS et leur affinité vis-à-vis de l'ARN cible. Elles favorisent également leur adsorption au niveau de la membrane plasmique, ce qui permet, en théorie, leur internalisation dans la cellule sans transfection préalable. Des érythroblastes primaires de patients porteurs d'un SMD *SF3BI*^{MUT} ou *SF3BI*^{WT} sont actuellement cultivés. Les OAS ERFE⁺¹² ou contrôles sont ajoutés au milieu de culture à partir du 10^{ème} jour post-tri CD34⁺, c'est-à-dire au moment de la différenciation érythroïde terminale. Les jours suivants, les ARNm des érythroblastes primaires sont extraits et le niveau d'expression des transcrits ERFE^{WT} et ERFE⁺¹² est évalué par RT-qPCR. L'idée serait ensuite d'ajouter le surnageant de culture des érythroblastes traités avec les OAS aux hépatocytes (lignée Hep3B ou HepG2), de façon à vérifier si la modification de l'expression d'ERFE⁺¹² entraîne une modification de l'expression de l'hepcidine. Une dizaine d'OAS font à ce jour l'objet d'essais cliniques, notamment pour le traitement des cancers, des maladies génétiques ou des infections virales.

2. Quels transcrits aberrants pour quel phénotype ?

Beaucoup de transcrits aberrants sont communément retrouvés dans différents types de tumeurs *SF3BI*^{MUT}. C'est le cas, par exemple, du transcrite ABCB7 alternatif non spécifique du tissu érythroïde, puisqu'il a été identifié dans les SMD, les MU et les LLC *SF3BI*^{MUT} (cf. Chapitre II-Partie 1-II-C-3). C'est aussi le cas, comme l'a relevé l'équipe de Marc-Henri Stern et mon équipe d'accueil, des transcrits GUSBP11, ANKHD1, CRNDE, UQCC et ABCC5 alternatifs, identifiés, là encore, dans les SMD, les MU et les LLC *SF3BI*^{MUT} (Gentien et al.,

2014). Le fait que les mêmes anomalies d'épissage soient observées dans des cellules tumorales de lignages et d'origine embryonnaire différents suggère qu'elles sont probablement les conséquences directes des mutations de *SF3B1*.

L'épissage alternatif pourrait contribuer au processus de transformation en favorisant la sélection clonale (avantage de croissance et survie) et le blocage de la différenciation. Par exemple, un transcrit aberrant du facteur de transcription FOXP1 a été identifié dans les LLC avec mutation *SF3B1* (Quesada et al., 2012). FOXP1 favorise l'expansion des lymphocytes B normaux par répression de l'expression de gènes pro-apoptotiques (Keimpema et al., 2014). Dans les cellules LLC *SF3B1*^{MUT}, le transcrit FOXP1 aberrant serait, selon les prédictions bio-informatiques, traduit en un variant tronqué au niveau de la région C-terminale qui contient les séquences régulant la dégradation de la protéine. Le transcrit FOXP1 aberrant est surexprimé dans les cellules LLC *SF3B1*^{MUT} et contribue ainsi à la survie des cellules mutées.

Dans les SMD-SC, les mutations de *SF3B1* semblent conférer un avantage de prolifération aux cellules mutées. Ceci a été confirmé par l'observation d'une augmentation des cellules souches hématopoïétiques chez la souris KI *Sf3b1*^{K700E} (Obeng et al., 2016). Nous avons identifié une diminution de l'expression du facteur de transcription RUNX1 (pour *Runt-related transcription factor 1*) dans les données de RNA-seq issues des CMN de patients. Il serait intéressant de vérifier si un épissage aberrant est à l'origine de ce changement d'expression et si ce dernier contribue à la sélection clonale de la CSH mutée, tout comme il a été montré que l'épissage aberrant d'EZH2 (pour *Enhancer of zeste homolog 2*) est responsable d'une diminution de son expression et du blocage partiel de différenciation observé dans les SMD avec mutation *SRSF2* (Kim et al., 2015).

Au-delà de l'analyse des sites 3' accepteurs alternatifs, il conviendra d'analyser l'ensemble des événements d'épissage alternatifs pouvant induire une modification de l'expression et/ou de la fonction de la protéine canonique. Physiologiquement, la fréquence des rétentions d'intron augmente au cours de la différenciation érythroïde ce qui se traduit par une diminution d'expression des transcrits correspondants (Pimentel et al., 2016). Or, Shiozawa *et al.* ont par exemple montré une réduction de la rétention d'intron dans les CMN et CD34⁺ de SMD *SF3B1*^{MUT} par rapport aux cellules *SF3B1*^{WT} (Shiozawa et al., 2018). Il nous apparaît donc important d'étudier le devenir des transcrits retenant des introns dans le contexte des mutations *SF3B1* à partir des données de RNA-Seq d'érythroblastes primaires *SF3B1*^{MUT} et *SF3B1*^{WT} (Baso-E et Poly-E).

3. Du transcrit aberrant à la protéine aberrante

Nous menons une analyse protéogénomique combinant RNA-Seq et protéomique des érythroblastes pour identifier de nouveaux variants protéiques. Brièvement, les protéines sont digérées par la trypsine, fractionnées en peptides qui sont injectés sur un système de nano-chromatographie en phase liquide couplé à un SM (Orbitrap Fusion, Thermofisher). L'analyse est réalisée dans MaxQuant et selon la méthode décrite par mon équipe d'accueil (Gautier et al., 2016). L'interrogation de la base de données Swiss-Prot (UniProt) a permis d'identifier plus de 6 000 protéines par échantillon. En parallèle, l'analyse des RNA-Seq a permis d'identifier tous les transcrits alternatifs des les échantillons *SF3BI*^{MUT}. Les séquences peptidiques sont déduites pour constituer une base de données de référence. Les néopeptides sont recherchés dans les données de protéomique (Annexe 7). Actuellement, nous avons identifié deux potentiels variants protéiques dus à un saut d'exon dans les érythroblastes *SF3BI*^{MUT}. Cette étude est en cours.

L'un des intérêts majeurs de la protéogénomique dans les cancers – outre l'identification de nouveaux biomarqueurs – est l'identification de néo-antigènes. En effet, certains variants protéiques sont susceptibles d'être présentés à la surface des cellules cancéreuses par des molécules du complexe majeur d'histo-compatibilité de classe I, et sont alors des potentielles cibles d'anticorps thérapeutiques. En 2018, Kahles *et al.* ont ré-analysé les données de génomique (RNA-Seq et séquençage de l'exome entier) et de protéomique répertoriées dans le TCGA (pour *The Cancer Genome Atlas*) chez 8 705 patients atteints de différents types de cancers et 670 contrôles sains (Kahles et al., 2018). Trois observations majeures ressortent de l'analyse des données de génomique : 1) les échantillons tumoraux – y compris ceux sans mutation des gènes d'épissage – ont jusqu'à 30% de plus d'évènements d'épissage alternatif que les tissus sains correspondants ; 2) les transcrits alternatifs non répertoriés dans les bases de données actuelles – et étant donc la conséquence de nouveaux évènements d'épissage – sont omniprésents dans les échantillons tumoraux et 3) comme dit précédemment, de nombreux mêmes transcrits alternatifs sont retrouvés dans différents types de tumeurs. Les techniques actuelles de protéomique identifient au moins un potentiel néo-épitope issu d'un épissage alternatif dans 68% des échantillons de cancer du sein ou de l'ovaire analysés. Toutefois, la validation fonctionnelle de l'immunogénicité des néo-antigènes n'a pas encore été évaluée. L'amélioration future des techniques de protéomique – notamment la sensibilité de détection des peptides – permettra probablement d'identifier d'autres néo-épitopes.

4. Comment les mutations de *SF3B1* altèrent-elles l'épissage des pré-ARNm ?

Les études récentes de cristallographie n'ont pas totalement élucidé le mécanisme de dérégulation de l'activité de la protéine SF3B1 mutée, puisque son affinité pour le pré-ARNm semble identique à celle de la protéine SF3B1 sauvage (Cretu et al., 2016).

Quelle est la structure exacte de la protéine SF3B1 mutée ? Un changement du repliement tridimensionnel de la région protéique liée au pré-ARNm peut-il expliquer le choix d'un BS aberrant ? La modification du site de liaison de la protéine SF3B1 mutée au pré-ARNm gêne-t-elle la fonction des autres facteurs et/ou cofacteurs d'épissage, comme l'hélicase Prp5 qui aide au remodelage de l'ARN au cours de l'épissage (pour revue Jenkins and Kielkopf, 2017) ? Le CLIP-Seq (pour *UV-cross-linking and RNA Immunoprecipitation-RNA sequencing*) pourrait aider à mieux comprendre les mécanismes de liaison de la protéine SF3B1 mutée à l'ARN (Rejeski et al., 2015). En effet, cette technique explore les associations protéines/ARN par réticulation UV-dépendante, puis immunoprécipitation de l'ARN et séquençage à haut débit. *In fine*, le CLIP-Seq identifie les motifs de l'ARNm spécifiquement liés à la protéine étudiée. Cette méthode pourrait permettre de comprendre pourquoi seule une petite partie (1%) des pré-ARNm normalement ciblés par la protéine SF3B1 ubiquitaire subit un épissage alternatif lorsque celle-ci est mutée. Sachant que les protéines se liant à l'ARN ont une expression spécifique à chaque tissu, il nous semble important de poser ces questions dans le tissu hématopoïétique – et plus particulièrement au cours de l'érythropoïèse –, pour identifier les mécanismes favorisant l'émergence de la maladie. Enfin, étant donné que les mutations des gènes d'épissage se produisent préférentiellement chez les personnes âgées de plus de 70 ans, une attention particulière devra prochainement être accordée aux changements de l'épissage liés au vieillissement dans les CSH et progéniteurs hématopoïétiques (pour revue Joshi et al., 2017).

5. SF3B1, un régulateur de la réplication de l'ADN ?

Nos travaux suggèrent que les mutations du gène *SF3B1* sont associées à une augmentation de la dynamique de réplication de l'ADN des précurseurs érythroïdes immatures humains (érythroblastes primaires) et murins (lignée G1E-ER4 CRISPR-Cas9). Notre hypothèse mécanistique est que les fourches progressent sans rencontrer d'obstacles, tels que les *R-loops* décrits dans le contexte des mutations des gènes *SRSF2* ou *U2AF1*.

Les anomalies de la réplication de l'ADN observées dans les cellules *SF3B1*^{MUT} peuvent-elles être dues à un épissage aberrant des pré-ARNm de la réplication de l'ADN ? Nos analyses de RNA-Seq générées à partir des échantillons de CMN de patients ne révèlent pas de jonction alternative sur un gène connu pour contrôler la réparation ou la réplication de l'ADN. Les rétentions d'intron, sauts d'exon, exons mutuellement exclusifs et promoteurs ou sites de polyadénylation alternatifs, ainsi que leur association avec une éventuelle modification de l'expression génique, sont en cours d'étude, dans les échantillons de CMN mais aussi dans les érythroblastes des patients.

Plusieurs travaux ont démontré que les protéines de la voie AF participaient au maintien de la stabilité génomique en contrôlant les processus simultanés de réplication et de transcription de l'ADN (Lossaint et al., 2013; Schwab et al., 2015). En effet, les expériences de PLA (pour *Proximity ligation assay*) révèlent que FANCD2 colocalise avec l'ARN polymérase II totale et, surtout, phosphorylée. La diminution de l'expression de FANCD2 dans les cellules U2OS traitées avec un siFANCD2 induit une accumulation de *R-loops* et un arrêt des fourches de réplication avec formation de CDB. La diminution de l'expression de FANCA – enzyme nécessaire à la mono-ubiquitinylation de FANCD2 – entraînant également un excès de *R-loops*, l'activation physiologique de la voie AF durant la transcription de l'ADN permet donc de limiter la formation des hybrides ARN/ADN bloquant la progression des fourches de réplication. En outre, la protéine FANCM est impliquée dans la résolution des *R-loops* via son activité de translocation de l'ADN et de l'ARN. Au vu de ces résultats, et du fait de l'interaction de la protéine SF3B1 sauvage avec FANCD2 et FANCI (Moriel-Carretero et al., 2017), il serait intéressant de déterminer le potentiel rôle de SF3B1 dans le maintien du réplisome et de préciser l'impact de la mutation sur la formation du complexe SF3B1-FANCD2/I.

Enfin, bien que les mutations de *SF3B1* soient transmises par la CSH tout au long de la hiérarchie hématopoïétique (Chesnais et al., 2017), seuls les érythroblastes mutés présentent un stress réplicatif. En effet, nos expériences préliminaires d'IF ne montrent pas de foyers intranucléaires de p-RPA32 dans les granulocytes des patients porteurs d'un SMD-SC *SF3B1*^{MUT}. Pourquoi l'érythropoïèse est-elle plus sensible aux anomalies de l'épissage et/ou de la réplication et réparation de l'ADN que la granulopoïèse ? L'une des hypothèses vient probablement du fait que notre organisme renferme 2000 à 6000 fois plus de globules rouges que de globules blancs. Le tissu érythroïde est en perpétuel renouvellement, et la progression de chaque cellule dans le cycle cellulaire nécessite une régulation soignée de son expression

génique et une duplication de son matériel génétique. En outre, le tissu érythroïde est plus sensible au stress réplicatif induit par déplétion en nucléotides (cf. Chapitre II-Partie 2-I-A-3) (Austin et al., 2012).

6. Quelles options thérapeutiques raisonnables ?

D'un point de vue thérapeutique, l'une des stratégies en cours de développement clinique pour la prise en charge de l'anémie des patients atteints d'un SMD est le luspatercept (cf. Chapitre I-II-D-2). Cette molécule est aussi testée dans les β -thalassémies, pathologies associant, comme les SMD-SC, un arrêt de la différenciation érythroïde et une surcharge systémique en fer. Pour rappel, le luspatercept piège les ligands de la superfamille du TGF β , dont le GDF11 produit par les progéniteurs et précurseurs érythroïdes. Grâce à des expériences de RT-qPCR, l'équipe d'Olivier Hermine a, dès 2014, mis en évidence une augmentation de l'expression de GDF11 dans les érythroblastes de la rate des souris thalassémiques (*Hbb*^{th1/th1}) (Dussiot et al., 2014). D'après les dosages ELISA, ces souris ont une concentration plus élevée de GDF11 sérique. Par conséquent, la voie de signalisation SMAD2 est activée, comme le confirment les analyses immunohistochimiques. Dussiot *et al.* ont ensuite étudié les effets d'une autre molécule issue de la même classe thérapeutique que le luspatercept : le sotatercept (ou ACE-011), une protéine recombinante de fusion formée du domaine extracellulaire modifié du récepteur à l'activine de type IIA lié au Fc de l'immunoglobuline G1. Le traitement des souris thalassémiques pendant 60 jours avec l'équivalent murin du sotatercept (RAP-011) montre, entre autres, une correction de l'anémie avec augmentation des taux d'hémoglobine et amélioration de la différenciation érythrocytaire terminale. Les concentrations sériques en EPO augmentent de façon drastique. Concernant l'aspect mécanistique, la phosphorylation de SMAD2 est bien réduite chez les souris thalassémiques traitées par rapport aux souris non traitées. Enfin, la maturation terminale des érythroblastes murins thalassémiques cultivés *in vitro* est rétablie en présence d'un anticorps anti-GDF11, mais pas en présence des anticorps anti-GDF8, anti-activine A et anti-activine B, les potentielles autres cibles du RAP-011. Le RAP-011 diminue la quantité de ROS produits par les érythroblastes murins thalassémiques en culture. L'érythropoïèse inefficace est souvent associée à une surcharge systémique en fer. Les souris thalassémiques sous RAP-011 voient leur surcharge systémique en fer diminuer au fur et à mesure de la restauration de l'érythropoïèse. En effet, le taux de saturation de la transferrine sérique baisse significativement, tandis que l'expression du gène *Hamp* dans les hépatocytes augmente. Il

conviendra de préciser l'impact du sotatercept et du luspatercept sur le métabolisme du fer chez l'homme. Les dosages ELISA révélant une augmentation des concentrations sériques de GDF11 chez les patients porteurs d'un SMD ou d'une β -thalassémie, peut-on supposer qu'il existe un mécanisme commun de mise en place de l'érythropoïèse inefficace dans ces deux pathologies ? Des travaux complémentaires sont nécessaires, notamment pour mieux explorer les effets du GDF11 sur le stress oxydatif, mais également pour identifier les gènes ciblés par les facteurs de transcription SMAD2/3 et leur rôle dans l'érythropoïèse.

Parmi les anti-cancéreux en cours d'évaluation clinique, figurent des molécules ciblant le complexe SF3b du spliceosome, telles que la spliceostatine A (SSA), la meayamycine ou encore le composé E7107, un dérivé du pladienolide B (Bonnal et al., 2012). L'ensemble de ces molécules a été obtenu par hémisynthèse à partir de composés naturels produits par les bactéries *Pseudomonas* spp. ou *Streptomyces* spp. La cible protéique des modulateurs de l'épissage n'est pas clairement identifiée : bien que SF3b130 soit, d'après les expériences de co-immunoprécipitation avec une SSA biotinylée, la sous-unité du complexe SF3b la plus enrichie dans la fraction associée à la SSA biotinylée, la sous-unité SF3B1 affiche quant à elle la liaison la plus forte (Kaida et al., 2007). Plus récemment, Teng *et al.* ont déterminé la structure de la sous-unité SF3b14b par cristallographie (Teng et al., 2017). D'après leurs travaux, SF3B1 et SF3b14b forment une petite poche dans laquelle vient se fixer la molécule inhibitrice de l'épissage. Cette poche se situe au niveau du BS du pré-ARNm. L'association des modulateurs de l'épissage au complexe SF3b induit un changement conformationnel puis une destabilisation de la particule nucléaire ribonucléoprotéique U2. Cette dernière n'est alors plus capable d'interagir avec le pré-ARNm, d'où un arrêt de l'épissage et de la prolifération cellulaire. Il est important de noter que ces molécules ne ciblent pas spécifiquement la protéine SF3B1 – ou une quelconque autre protéine du spliceosome – mutée. Leur non spécificité d'action vis-à-vis de l'épissage alternatif représente un obstacle majeur à leur application clinique. En effet, étant donné la place centrale du processus d'épissage dans la biologie des cellules eucaryotes, l'utilisation de molécules inhibitrices de l'épissage peut être source d'effets secondaires très importants. D'après les premières études *in vitro*, il semblerait toutefois que les modulateurs de l'épissage aient plutôt un effet cytostatique (et non cytotoxique) à concentration thérapeutique. Ils n'inhibent pas entièrement l'épissage mais diminuent l'expression – par NMD, rétention d'intron ou formation de petits ARNm traduits en des variants protéiques non fonctionnels – de certains gènes surexprimés dans les cellules cancéreuses et/ou impliqués dans la progression tumorale, comme le facteur de croissance de

l'endothélium vasculaire (Furumai et al., 2010) ou MDM2, un régulateur négatif du gène suppresseur de tumeur p53 (Fan et al., 2011). Pourquoi les pré-ARNm ne présentent-ils pas tous la même sensibilité à ces composés ? Les modulateurs de l'épissage ciblent-ils une catégorie particulière de pré-ARNm ? Et si oui, quelles sont les particularités de ces pré-ARNm ? Diverses zones d'ombre persistent.

Dans le cas des SMD ou LAM avec mutation sur l'un des gènes du spliceosome, l'intérêt thérapeutique des modulateurs de l'épissage s'appuie sur les travaux démontrant que l'expression de l'allèle sauvage en regard de l'allèle mutant est indispensable à la survie de la cellule (Lee et al., 2016; Zhou et al., 2015). Les cellules hétérozygotes $SF3BI^{WT/MUT}$ sont donc, en théorie, plus sensibles à ces molécules que les cellules homozygotes $SF3BI^{WT/WT}$. Récemment, l'équipe de Silvia Buonamici a par exemple étudié les effets de la molécule H3B-8800, un dérivé du pladienolide B synthétisé par le laboratoire H3 Biomedicine, dans des lignées cellulaires $SF3BI^{WT}$ et $SF3BI^{MUT}$ (Seiler et al., 2018b). Comme attendu, la molécule H3B-8800 inhibe l'épissage canonique et aberrant des cellules $SF3BI^{MUT}$, mais a un effet pro-apoptotique plus conséquent sur les cellules $SF3BI^{MUT}$ par rapport aux cellules $SF3BI^{WT}$. Selon cette même logique, la molécule H3B-8800 a un effet anti-tumoral chez les souris immunodéficientes greffées avec des cellules CD34⁺ leucémiques humaines $SF3BI^{MUT}$ ou $SRSF2^{MUT}$, mais pas chez les souris greffées avec des cellules CD34⁺ $SF3BI^{WT}$ ou $SRSF2^{WT}$. De plus, d'après les analyses de RNA-Seq, les introns majoritairement retenus dans les cellules traitées avec la molécule H3B-8800 sont petits et riches en GC. Or, de façon intéressante, les gènes impliqués dans l'épissage des pré-ARNm sont significativement enrichis en ce type d'introns. Pour résumer, Seiler *et al.* suggèrent que la molécule H3B-8800 induit surtout la rétention des introns contenus dans les gènes du spliceosome, ce qui potentialise l'apoptose préférentielle des cellules mutées sur $SF3BI$ ou $SRSF2$. Actuellement, cette molécule est testée dans le cadre d'un essai clinique de phase I chez les patients porteurs d'un SMD, d'une LAM ou d'une leucémie myélomonocytaire chronique et résistants ou intolérants aux thérapies classiques (Steensma et al., 2017). Seuls les patients porteurs d'un SMD de haut risque (ou d'un SMD de bas risque avec dépendance aux transfusions sanguines et non réponse aux ASE) sont éligibles à cet essai clinique, puisque, comme dit dans le chapitre I, la stratégie thérapeutique chez les patients porteurs d'un SMD de bas risque vise surtout à corriger l'anémie et non à supprimer l'hématopoïèse clonale de sujet âgé dont l'hématopoïèse normale résiduelle est de mauvaise qualité (Flach et al., 2014; Walter et al., 2015).

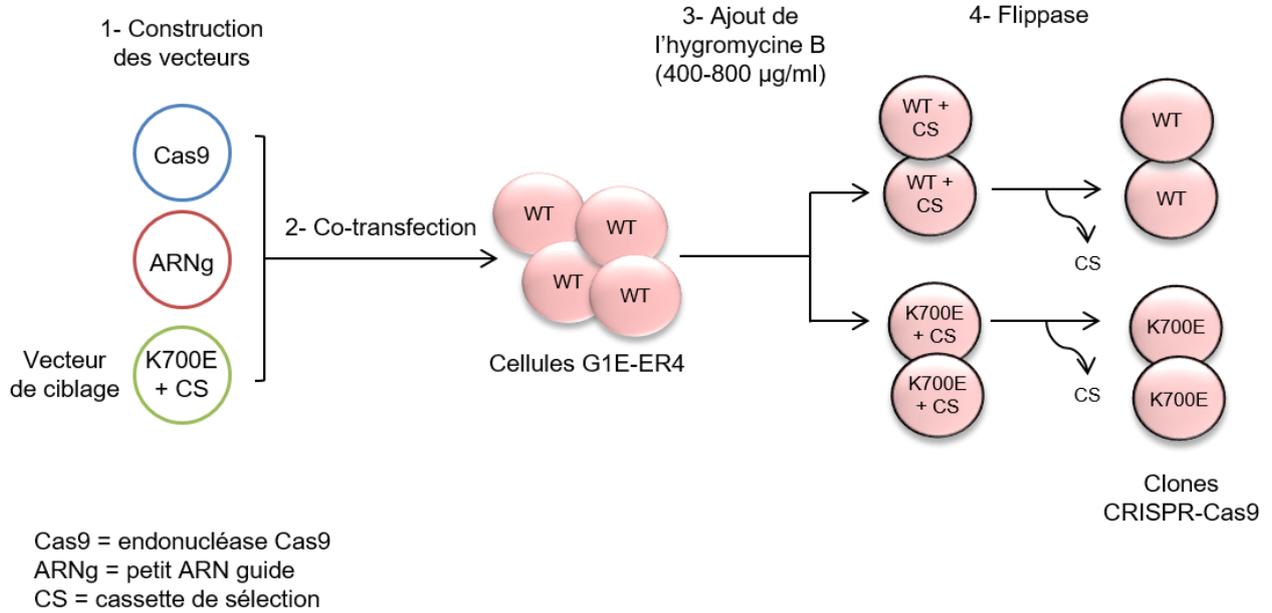
D'après les travaux *in vitro* de Nguyen *et al.*, les modulateurs de l'épissage – ici le composé E7107 – conduisent à l'accumulation de *R-loops* dans les cellules HeLa et K562 (Nguyen *et al.*, 2018). En effet, l'arrêt de l'épissage augmente la formation des hybrides ARN/ADN, et active donc la voie ATR par blocage des fourches de réplication. L'accumulation de *R-loops* est d'autant plus importante lorsque les cellules expriment de façon stable une mutation sur le gène *U2AF1* (cf. Chapitre II-Partie 2-II). Par conséquent, l'apoptose des cellules K562 *U2AF1*^{MUT} est supérieure à l'apoptose des cellules *U2AF1*^{WT} après 24h de traitement au composé E7107 (10% de cellules *U2AF1*^{WT} annexine V positives *versus* 15% de cellules *U2AF1*^{MUT} annexine V positives). L'association d'un inhibiteur d'ATR avec le composé E7107 renforce considérablement l'apoptose des cellules K562 *U2AF1*^{MUT} (25% de cellules *U2AF1*^{MUT} annexine V positives), puisque, la fourche n'étant plus arrêtée le temps de résoudre les *R-loops*, des dommages à l'ADN se forment et entraînent une mort cellulaire. Les inhibiteurs d'ATR validés *in vitro* dans les lignées cancéreuses sont en cours d'études cliniques dans les tumeurs solides. Nguyen *et al.* proposent de les associer avec des modulateurs de l'épissage de façon à « sensibiliser » les cellules aux dommages à l'ADN, mais également de les tester chez les patients porteurs d'un SMD à haut risque de transformation lorsque le gène *SRSF2* ou *U2AF1* est muté. Les mutations de *SF3B1* étant de bon pronostic, il ne devrait pas y avoir d'intérêt à une telle association dans les SMD-SC.

Pour conclure, les fonctions et localisations subnucléaires de la protéine SF3B1 sont multiples. L'exploration des conséquences fonctionnelles des mutations du gène *SF3B1* nécessite donc de s'intéresser à divers processus et structures cellulaires, tels que l'épissage co-transcriptionnel au niveau des brins d'ARN naissants et la réplication de l'ADN.

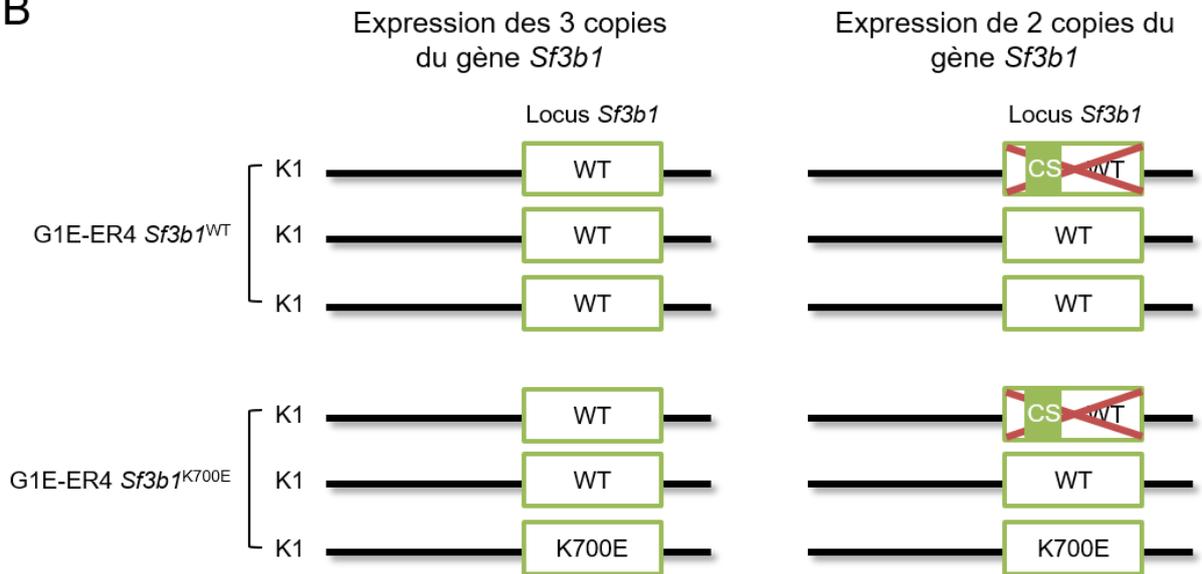
Annexes

Annexe 1 : Clones G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E}

A

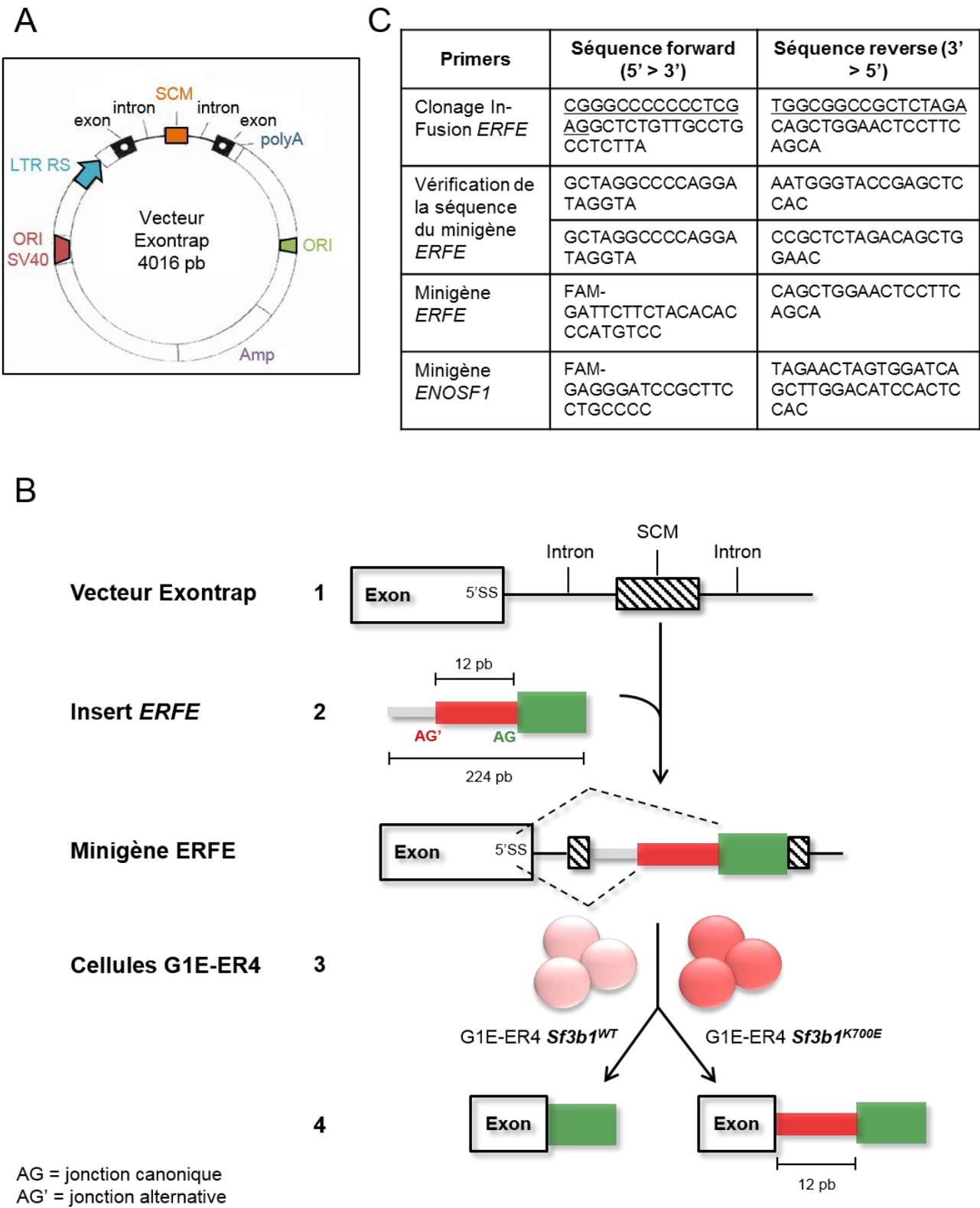


B



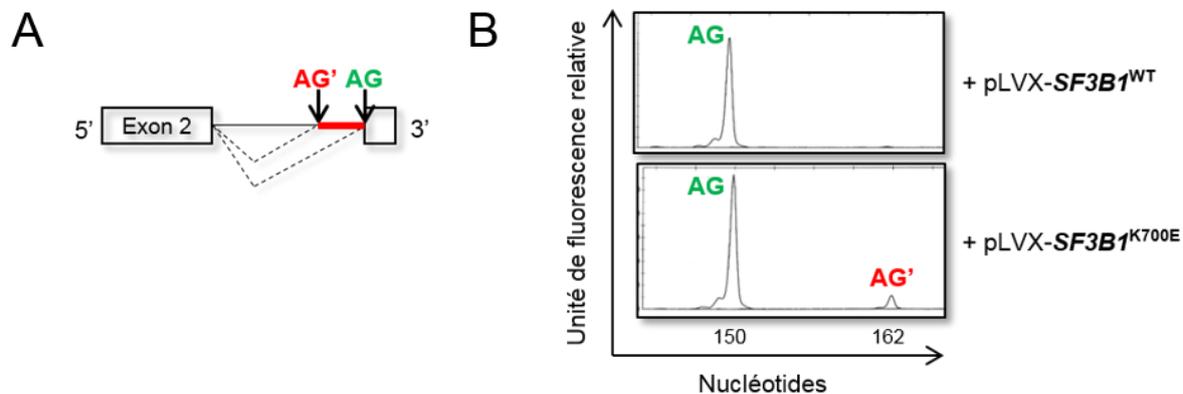
Annexe 1 : Clones G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E}. (A) **Construction des clones G1E-ER4 CRISPR-Cas9** (pour *Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated protein 9*) *Sf3b1*^{K700E}. **1-** Construction des trois vecteurs plasmidiques codant pour : **l'endonuclease Cas9**, **le petit ARN guide** et **le vecteur de ciblage** porteur de la séquence *Sf3b1*^{K700E} et d'une cassette de sélection. La cassette de sélection correspond à un gène de résistance à l'hygromycine B entouré des séquences FRT (pour *Flippase recognition target*). **2-** Co-transfection des trois vecteurs dans les cellules G1E-ER4. La lignée murine érythroïde G1E-ER4 dérive de cellules souches embryonnaires bloquées au stade Pro-E par invalidation de *GATA1* (Weiss et al., 1997). Les cellules G1E-ER4 sont c-KIT⁺ et Ter119⁻. Leur temps de doublement est d'environ 12h. Elles expriment de manière stable une protéine de fusion GATA1/ER (récepteur aux estrogènes), dans laquelle la région codante du gène GATA1 est fusionnée avec le domaine de liaison du ligand du récepteur humain aux estrogènes (Welch et al., 2004). Ainsi, l'ajout de β -estradiol 10^{-7} M dans le milieu de culture restaure l'activité de GATA1 et les cellules entrent dans un processus de différenciation érythrocytaire. Après 72h de traitement aux estrogènes, les cellules G1E-ER4 atteignent un stade de développement équivalent à l'érythroblaste acidophile. **3-** Sélection des cellules ayant intégré le vecteur de ciblage par l'hygromycine B. **4-** Vérification de la présence de la mutation *Sf3b1*^{K700E} par PCR et séquençage Sanger, puis retrait de la cassette de sélection sous l'action de la flippase qui reconnaît les séquences FRT. (B) **Différents clones G1E-ER4 obtenus.** Chez la souris, le gène *Sf3b1* se situe sur le chromosome 1. Les cellules G1E-ER4 ont trois chromosomes 1, donc trois copies du gène *Sf3b1*. La flippase ne retire pas systématiquement la cassette de sélection. Cette dernière se situant en plein milieu du gène *Sf3b1*, elle empêche son expression. Certains clones n'expriment finalement que deux copies du gène *Sf3b1*, comme dans les cellules primaires humaines. Les clones G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E} ont été établis par le Dr Carine Lefèvre, chercheur post-doctorante au sein de mon équipe d'accueil.

Annexe 2 : Minigène *ERFE*



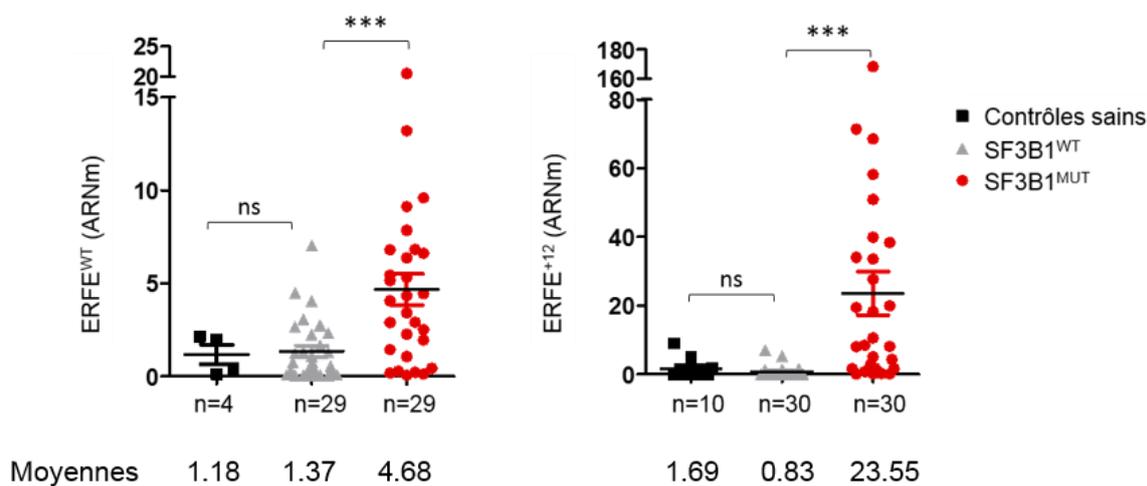
Annexe 2 : Minigène *ERFE*. (A) Vecteur Exontrap (MoBiTec, Goettingen, Allemagne). Le vecteur Exontrap est un petit fragment de gène formé de deux séquences exoniques séparées par une séquence intronique, à l'intérieur de laquelle se trouve un site de clonage multiple. L'intron contient toutes les séquences consensus nécessaires à l'expression des exons. **SCM** = site de clonage multiple, **polyA** = site de polyadénylation en 3', **ORI** = origine de réplication bactérienne (pBR322), **amp** = gène de résistance à l'ampicilline, **ORI SV40** = origine de réplication SV40 (virus simien 40), **LTR RS** = *long terminal repeat* = promoteur RS (virus du sarcome de Rous). **(B) Construction du minigène *ERFE*. 1-** Le vecteur Exontrap est dans un premier temps linéarisé par digestion avec les enzymes de restriction XhoI et XbaI. **2-** L'ADN génomique humain (cellules MOLM-14, \approx 200 pb) autour de la jonction alternative (AG') identifiée sur *ERFE* est amplifié par PCR avec l'ADN polymérase Phusion (Thermo Fisher Scientific). Puis, l'insert *ERFE* est cloné dans le vecteur Exontrap par clonage In-Fusion (Clontech Takara Bio, Mountain View, CA). La technique de clonage In-Fusion repose, comme son nom l'indique, sur la fusion de deux fragments d'ADN par homologie d'une quinzaine de pb situées à l'extrémité de ces deux fragments. En pratique, les primers entourant la séquence d'intérêt, la séquence du vecteur Exontrap ainsi que les enzymes de restriction choisies sont tous trois renseignés sur le site internet <http://bioinfo.clontech.com/infusion/convertPcrPrimers.do> développé par Clontech. Le logiciel détermine alors les primers In-Fusion, avec introduction d'une quinzaine de pb (annexe 2c, pb soulignées) de part et d'autre des primers de la séquence d'intérêt, ces pb étant homologues aux extrémités du vecteur linéarisé. Les primers In-Fusion sont utilisés pour amplifier la séquence d'intérêt. Le minigène *ERFE* est produit dans les bactéries *Escherichia coli*, puis la séquence est vérifiée par séquençage Sanger (primers en annexe 2c). **3-** Le minigène *ERFE* est nucléofecté dans les cellules murines G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E} ou *Sf3b1*^{WT} (Lonza, kit L, programme A-024). **4-** Vingt-quatre heures post transfection, l'ARNm des cellules G1E-ER4 est extrait puis converti en ADN complémentaire (ADNc) par RT-PCR. L'ADNc est amplifié par PCR avec des primers spécifiques fluorescents (annexe 2c). Enfin, l'électrophorèse capillaire sépare les fragments canoniques des fragments alternatifs (+ 12 pb). Le minigène *ENOSF1* sert de contrôle (construit par le Dr Samar Alsafadi, Institut Curie) (primers en annexe 2c). **(C) Séquences des primers utilisés pour les expériences de minigènes.**

Annexe 3 : Transfection des HUDEP-2 avec les vecteurs d'expression *SF3B1*^{K700E} ou *SF3B1*^{WT}



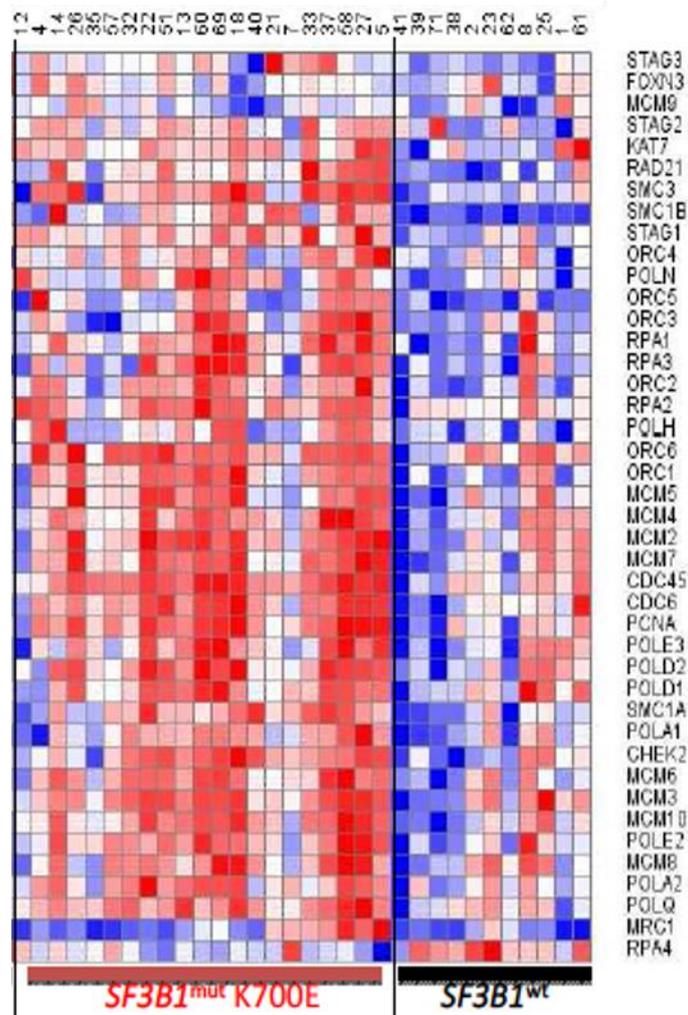
(A) Le site 3' accepteur alternatif (AG') d'ERFE se situe entre les exons 2 et 3, 12 pb en amont du site 3' accepteur canonique (AG). (B) Les cellules érythroïdes HUDEP-2 ont été transfectées avec un plasmide codant pour une forme WT ou mutée de *SF3B1*. Après extraction de l'ARN, RT-PCR autour de la jonction d'intérêt, et analyse de fragment, nous détectons le transcrit aberrant ERFE⁺¹² uniquement dans les cellules transfectées avec le plasmide p-LVX *SF3B1*^{K700E}.

Annexe 4 : Quantification des transcrits ERFE^{WT} et ERFE⁺¹² par RT-qPCR



La quantification des transcrits ERFE^{WT} et ERFE⁺¹² dans les CMN des patients ou des contrôles sains a été faite par RT-qPCR. Les valeurs d'expression d'ERFE^{WT} et d'ERFE⁺¹² ont été normalisées par rapport à celles des gènes de ménage (*B2M* codant pour la β 2-microglobuline et *ACTB* codant pour la β -actine). Les séquences des primers figurent dans la sup. table S5 de l'article soumis à STM. Le test de Student compare les moyennes des expressions normalisées d'ERFE^{WT} ou ERFE⁺¹² entre les différents groupes (***) ($P < 0.001$).

Annexe 5 : Analyse transcriptomique de l'expression de 48 gènes de la réplication et de la réparation de l'ADN dans les CMN de patients porteurs d'un SMD



Rouge : gènes surexprimés au niveau ARNm

Bleu : gènes sous-exprimés au niveau ARNm

Puce GeneChip® Human Genome U133 Plus 2.0 (Affymetrix, Santa Clara, CA) sur CMN. 50 patients SMD au screening, dont 39 *SF3B1*^{MUT} (22 *SF3B1*^{K700E}) et 11 *SF3B1*^{WT}. Les analyses ont été réalisées par le Dr Virginie Chesnais (résultats non publiés). De façon générale, les 48 gènes étudiés, dont les hélicases MCM, les protéines de réplication A (RPA) et les ADN polymérase répliquatives, sont surexprimées dans les CMN des patients porteurs d'une mutation *SF3B1*^{K700E} par rapport aux patients non mutés.

Annexe 6 : Oligonucléotides anti-sens dirigés spécifiquement contre le transcrit aberrant ERFE⁺¹²

A

ADNg : 5' cctgctcctcacatgctcctcactgccaa**ag**gttcctttc**ag**TTCGGCTTGCCAGGG 3'

ag : site 3' accepteur cryptique

ag : site 3' accepteur canonique

B

1. OAS n°1 : Qiagen

i : Séquence de l'OAS : 5' AACTGAAAGGGAAC 3'

ii : Séquence ERFE complémentaire : 3' UUGACUUUCCCUUG 5'

iii : 5' gttcctttc**ag**TT 3'

2. OAS n°2 : GeneTools

i : Séquence de l'OAS : 5' AAAGGGAACCTTGGCAGTGAGGACA 3'

ii : Séquence ERFE complémentaire : 3' UUUCCCUUGGAACCGUCACUCCUGU 5'

iii : 5' gtcctcactgccaa**ag**gttccttt 3'

3. OAS n°3 : GeneTools

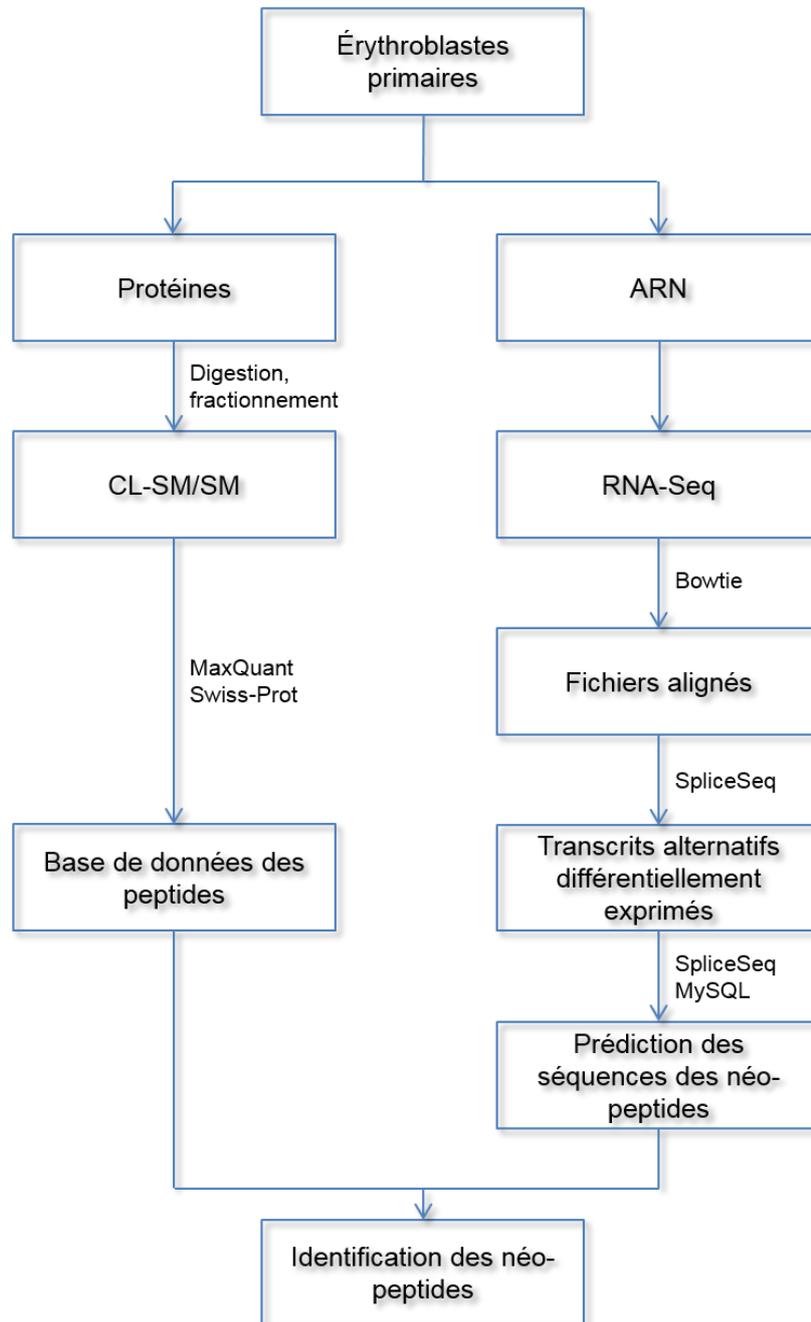
i : Séquence de l'OAS : 5' ACCTTGGCAGTGAGGACATGT 3'

ii : Séquence ERFE complémentaire : 3' UGGAACCGUCACUCCUGUACA 5'

iii : 5' acatgctcactgccaa**aggt** 3'

(A) Une partie de la séquence ADNg du gène ERFE autour de la jonction alternative, entre les exons 2 et 3. Minuscules = fin de l'intron 2-3, Majuscules = début de l'exon 3. **(B) Séquences des OAS.** i : séquences des OAS (ADN), dans le sens 5'>3' ; ii : séquences complémentaires de l'OAS sur le pré-ARNm ERFE dans le sens 3'>5' ; iii : séquences correspondantes sur ERFE dans le sens 5'>3'. Les OAS ont été conçus et commandés chez Qiagen et GeneTools (Philomath, OR).

Annexe 7 : Flowchart de la stratégie de protéogénomique sur les échantillons d'érythroblastes primaires



La protéomique permet de générer une base de données pour la recherche des néo-peptides prédits après analyse des données de RNA-Seq. **CL** = chromatographie liquide, **SM** = spectrométrie de masse, Bowtie = mapper (logiciel d'alignement).

Annexe 8 : Liste des publications

- ❖ Ladli M, Richard C, Cantero Aguilar L, Ducamp S, Bondu S, *et al.* Finely-tuned regulation of AMP-activated protein kinase (AMPK) is crucial for human adult erythropoiesis. *Haematologica*. 2018; 191403.
- ❖ Lefèvre C, Bondu S, Le Goff S, *et al.* Dyserythropoiesis of myelodysplastic syndromes. *Curr Opin Hematol*. 2017; 24: 191-197.
- ❖ Chesnais V, Arcangeli ML, Delette C, Rousseau A, Guermouche H, Lefevre C, Bondu S, *et al.* Architectural and functional heterogeneity of hematopoietic stem/progenitor cells in non-del(5q) myelodysplastic syndromes. *Blood*. 2017; 129: 484-496.

Des résumés scientifiques concernant les travaux en cours de révisions à *Science Translational Medicine* ont été soumis pour une communication orale aux congrès de l'European School of Haematology et de la Société Française d'Hématologie (mars 2019).

Bibliographie

Abboud, S., and Haile, D.J. (2000). A novel mammalian iron-regulated protein involved in intracellular iron metabolism. *J. Biol. Chem.* 275, 19906–19912.

Adolfsson, J., Månsson, R., Buza-Vidas, N., Hultquist, A., Liuba, K., Jensen, C.T., Bryder, D., Yang, L., Borge, O.-J., Thoren, L.A.M., et al. (2005). Identification of Flt3+ lympho-myeloid stem cells lacking erythro-megakaryocytic potential a revised road map for adult blood lineage commitment. *Cell* 121, 295–306.

Aguilera, A., and García-Muse, T. (2012). R Loops: From Transcription Byproducts to Threats to Genome Stability. *Molecular Cell* 46, 115–124.

Allikmets, R., Raskind, W.H., Hutchinson, A., Schueck, N.D., Dean, M., and Koeller, D.M. (1999). Mutation of a Putative Mitochondrial Iron Transporter Gene (ABC7) in X-Linked Sideroblastic Anemia and Ataxia (XLSA/A). *Hum Mol Genet* 8, 743–749.

Alsafadi, S., Houy, A., Battistella, A., Popova, T., Wassef, M., Henry, E., Tirode, F., Constantinou, A., Piperno-Neumann, S., Roman-Roman, S., et al. (2016). Cancer-associated *SF3B1* mutations affect alternative splicing by promoting alternative branchpoint usage. *Nature Communications* 7, 10615.

Ambaglio, I., Malcovati, L., Papaemmanuil, E., Laarakkers, C.M., Porta, M.G.D., Gallì, A., Vià, M.C.D., Bono, E., Ubezio, M., Travaglino, E., et al. (2013). Inappropriately low hepcidin levels in patients with myelodysplastic syndrome carrying a somatic mutation of SF3B1. *Haematologica* 98, 420–423.

Amigo, J.D., Yu, M., Troadec, M.-B., Gwynn, B., Cooney, J.D., Lambert, A.J., Chi, N.C., Weiss, M.J., Peters, L.L., Kaplan, J., et al. (2011). Identification of distal cis-regulatory elements at mouse mitoferrin loci using zebrafish transgenesis. *Mol. Cell. Biol.* 31, 1344–1356.

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

Anguita, E., Hughes, J., Heyworth, C., Blobel, G.A., Wood, W.G., and Higgs, D.R. (2004). Globin gene activation during haemopoiesis is driven by protein complexes nucleated by GATA-1 and GATA-2. *EMBO J.* 23, 2841–2852.

Arber, D.A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M.J., Beau, M.M.L., Bloomfield, C.D., Cazzola, M., and Vardiman, J.W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood* 127, 2391–2405.

Arezes, J., Foy, N., McHugh, K., Sawant, A., Quinkert, D., Terraube, V., Brinth, A., Tam, M., Lavallie, E., Taylor, S., et al. (2018). Erythroferrone inhibits the induction of hepcidin by BMP6. *Blood*.

Aschemeyer, S., Qiao, B., Stefanova, D., Valore, E.V., Sek, A.C., Ruwe, T.A., Vieth, K.R., Jung, G., Casu, C., Rivella, S., et al. (2018). Structure-function analysis of ferroportin defines the binding site and an alternative mechanism of action of hepcidin. *Blood* 131, 899–910.

Ashby, D.R., Gale, D.P., Busbridge, M., Murphy, K.G., Duncan, N.D., Cairns, T.D., Taube, D.H., Bloom, S.R., Tam, F.W.K., Chapman, R., et al. (2010). Erythropoietin administration in humans causes a marked and prolonged reduction in circulating hepcidin. *Haematologica* 95, 505–508.

Bibliographie

- Aul, C., Arning, M., Runde, V., and Schneider, W. (1991). Serum erythropoietin concentrations in patients with myelodysplastic syndromes. *Leukemia Research* 15, 571–575.
- Austin, W.R., Armijo, A.L., Campbell, D.O., Singh, A.S., Hsieh, T., Nathanson, D., Herschman, H.R., Phelps, M.E., Witte, O.N., Czernin, J., et al. (2012). Nucleoside salvage pathway kinases regulate hematopoiesis by linking nucleotide metabolism with replication stress. *J. Exp. Med.* 209, 2215–2228.
- Babitt, J.L., Huang, F.W., Wrighting, D.M., Xia, Y., Sidis, Y., Samad, T.A., Campagna, J.A., Chung, R.T., Schneyer, A.L., Woolf, C.J., et al. (2006). Bone morphogenetic protein signaling by hemojuvelin regulates hepcidin expression. *Nature Genetics* 38, 531–539.
- Babitt, J.L., Huang, F.W., Xia, Y., Sidis, Y., Andrews, N.C., and Lin, H.Y. (2007). Modulation of bone morphogenetic protein signaling in vivo regulates systemic iron balance. *J Clin Invest* 117, 1933–1939.
- de Back, D.Z., Kostova, E.B., van Kraaij, M., van den Berg, T.K., and van Bruggen, R. (2014). Of macrophages and red blood cells; a complex love story. *Front Physiol* 5.
- Bandyopadhyay, S., Gama, F., Molina-Navarro, M.M., Gualberto, J.M., Claxton, R., Naik, S.G., Huynh, B.H., Herrero, E., Jacquot, J.P., Johnson, M.K., et al. (2008). Chloroplast monothiol glutaredoxins as scaffold proteins for the assembly and delivery of [2Fe–2S] clusters. *EMBO J* 27, 1122–1133.
- Baron, M.H., Isern, J., and Fraser, S.T. (2012). The embryonic origins of erythropoiesis in mammals. *Blood* 119, 4828–4837.
- Berget, S.M., Moore, C., and Sharp, P.A. (1977). Spliced segments at the 5' terminus of adenovirus 2 late mRNA. *Proc. Natl. Acad. Sci. U.S.A.* 74, 3171–3175.
- Bessis, M.C., and Breton-Gorius, J. (1962). Iron Metabolism in the Bone Marrow as Seen by Electron Microscopy: A Critical Review. *Blood* 19, 635–663.
- Besson-Fournier, C., Latour, C., Kautz, L., Bertrand, J., Ganz, T., Roth, M.-P., and Coppin, H. (2012). Induction of activin B by inflammatory stimuli up-regulates expression of the iron-regulatory peptide hepcidin through Smad1/5/8 signaling. *Blood* 120, 431–439.
- Bessonov, S., Anokhina, M., Krasauskas, A., Golas, M.M., Sander, B., Will, C.L., Urlaub, H., Stark, H., and Lührmann, R. (2010). Characterization of purified human Bact spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA* 16, 2384–2403.
- Beyer, A.L., and Osheim, Y.N. (1988). Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev.* 2, 754–765.
- Biankin, A.V., Waddell, N., Kassahn, K.S., Gingras, M.-C., Muthuswamy, L.B., Johns, A.L., Miller, D.K., Wilson, P.J., Patch, A.-M., Wu, J., et al. (2012). Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature* 491, 399–405.
- Bonnal, S., Vigevani, L., and Valcárcel, J. (2012). The spliceosome as a target of novel antitumour drugs. *Nature Reviews Drug Discovery* 11, 847–859.
- Boudrez, A., Beullens, M., Waelkens, E., Stalmans, W., and Bollen, M. (2002). Phosphorylation-dependent Interaction between the Splicing Factors SAP155 and NIPP1. *J. Biol. Chem.* 277, 31834–31841.

Bibliographie

- Boultonwood, J., Pellagatti, A., Nikpour, M., Pushkaran, B., Fidler, C., Cattani, H., Littlewood, T.J., Malcovati, L., Della Porta, M.G., Jädersten, M., et al. (2008). The role of the iron transporter ABCB7 in refractory anemia with ring sideroblasts. *PLoS ONE* 3, e1970.
- Brada, S.J.L., Wolf, J.T. de, Hendriks, D., Louwes, H., van den Berg, E., and Vellenga, E. (1998). Characterization of the erythropoiesis in myelodysplasia by means of ferrokinetic studies, *in vitro* erythroid colony formation and soluble transferrin receptor. *Leukemia* 12, 340–345.
- Brissot, P., Ropert, M., Le Lan, C., and Loréal, O. (2012). Non-transferrin bound iron: a key role in iron overload and iron toxicity. *Biochim. Biophys. Acta* 1820, 403–410.
- Broudy, V.C. (1997). Stem Cell Factor and Hematopoiesis. *Blood* 90, 1345–1364.
- Broudy, V.C., Lin, N., Brice, M., Nakamoto, B., and Papayannopoulou, T. (1991). Erythropoietin receptor characteristics on primary human erythroid cells. *Blood* 77, 2583–2590.
- Burge, C.B., Padgett, R.A., and Sharp, P.A. (1998). Evolutionary fates and origins of U12-type introns. *Mol. Cell* 2, 773–785.
- Burma, S., Chen, B.P., Murphy, M., Kurimasa, A., and Chen, D.J. (2001). ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *J. Biol. Chem.* 276, 42462–42467.
- Camaschella, C., Campanella, A., De Falco, L., Boschetto, L., Merlini, R., Silvestri, L., Levi, S., and Iolascon, A. (2007). The human counterpart of zebrafish shiraz shows sideroblastic-like microcytic anemia and iron overload. *Blood* 110, 1353–1358.
- Canali, S., Zumbrennen-Bullough, K.B., Core, A.B., Wang, C.-Y., Nairz, M., Bouley, R., Swirski, F.K., and Babitt, J.L. (2017). Endothelial cells produce bone morphogenetic protein 6 required for iron homeostasis in mice. *Blood* 129, 405–414.
- Canonne-Hergaux, F., Gruenheid, S., Ponka, P., and Gros, P. (1999). Cellular and subcellular localization of the Nramp2 iron transporter in the intestinal brush border and regulation by dietary iron. *Blood* 93, 4406–4417.
- Casanovas, G., Vujić Spasic, M., Casu, C., Rivella, S., Strelau, J., Unsicker, K., and Muckenthaler, M.U. (2013). The murine growth differentiation factor 15 is not essential for systemic iron homeostasis in phlebotomized mice. *Haematologica* 98, 444–447.
- Casey, J.L., Koeller, D.M., Ramin, V.C., Klausner, R.D., and Harford, J.B. (1989). Iron regulation of transferrin receptor mRNA levels requires iron-responsive elements and a rapid turnover determinant in the 3' untranslated region of the mRNA. *EMBO J.* 8, 3693–3699.
- Castillo Bosch, P., Segura-Bayona, S., Koole, W., van Heteren, J.T., Dewar, J.M., Tijsterman, M., and Knipscheer, P. (2014). FANCI promotes DNA synthesis through G-quadruplex structures. *EMBO J.* 33, 2521–2533.
- Caudill, J.S., Imran, H., Porcher, J.C., and Steensma, D.P. (2008). Congenital sideroblastic anemia associated with germline polymorphisms reducing expression of FECH. *Haematologica* 93, 1582–1584.
- Cazzola, M., Barosi, G., Gobbi, P.G., Invernizzi, R., Riccardi, A., and Ascarì, E. (1988). Natural history of idiopathic refractory sideroblastic anemia. *Blood* 71, 305–312.
- Cerritelli, S.M., and Crouch, R.J. (2009). Ribonuclease H: the enzymes in eukaryotes. *FEBS J.* 276, 1494–1505.

Bibliographie

- Chan, J.H.P., Lim, S., and Wong, W.S.F. (2006). Antisense oligonucleotides: from design to therapeutic application. *Clin. Exp. Pharmacol. Physiol.* *33*, 533–540.
- de la Chapelle, A., Träskelin, A.L., and Juvonen, E. (1993). Truncated erythropoietin receptor causes dominantly inherited benign human erythrocytosis. *Proc. Natl. Acad. Sci. U.S.A.* *90*, 4495–4499.
- Chen, L., Chen, J.-Y., Zhang, X., Gu, Y., Xiao, R., Shao, C., Tang, P., Qian, H., Luo, D., Li, H., et al. (2017). R-ChIP Using Inactive RNase H Reveals Dynamic Coupling of R-loops with Transcriptional Pausing at Gene Promoters. *Mol. Cell* *68*, 745-757.e5.
- Chen, L., Chen, J.-Y., Huang, Y.-J., Gu, Y., Qiu, J., Qian, H., Shao, C., Zhang, X., Hu, J., Li, H., et al. (2018). The Augmented R-Loop Is a Unifying Mechanism for Myelodysplastic Syndromes Induced by High-Risk Splicing Factor Mutations. *Mol. Cell* *69*, 412-425.e6.
- Chen, T.H.-P., Kambal, A., Krysiak, K., Walshauer, M.A., Raju, G., Tibbitts, J.F., and Walter, M.J. (2011). Knockdown of Hspa9, a del(5q31.2) gene, results in a decrease in hematopoietic progenitors in mice. *Blood* *117*, 1530–1539.
- Chesnais, V., Renneville, A., Toma, A., Lambert, J., Passet, M., Dumont, F., Chevret, S., Lejeune, J., Raimbault, A., Stamatoullas, A., et al. (2016). Effect of lenalidomide treatment on clonal architecture of myelodysplastic syndromes without 5q deletion. *Blood* *127*, 749–760.
- Chesnais, V., Arcangeli, M.-L., Delette, C., Rousseau, A., Guermouche, H., Lefevre, C., Bondu, S., Diop, M., Cheok, M., Chapuis, N., et al. (2017). Architectural and functional heterogeneity of hematopoietic stem/progenitor cells in non-del(5q) myelodysplastic syndromes. *Blood* *129*, 484–496.
- Chou, S.T., Khandros, E., Bailey, L.C., Nichols, K.E., Vakoc, C.R., Yao, Y., Huang, Z., Crispino, J.D., Hardison, R.C., Blobel, G.A., et al. (2009). Graded repression of PU.1/Sfp1 gene transcription by GATA factors regulates hematopoietic cell fate. *Blood* *114*, 983–994.
- Chow, L.T., Gelinas, R.E., Broker, T.R., and Roberts, R.J. (1977). An amazing sequence arrangement at the 5' ends of adenovirus 2 messenger RNA. *Cell* *12*, 1–8.
- Claessens, Y.-E., Bouscary, D., Dupont, J.-M., Picard, F., Melle, J., Gisselbrecht, S., Lacombe, C., Dreyfus, F., Mayeux, P., and Fontenay-Roupie, M. (2002). In vitro proliferation and differentiation of erythroid progenitors from patients with myelodysplastic syndromes: evidence for Fas-dependent apoptosis. *Blood* *99*, 1594–1601.
- Collins, A.R. (1999). Oxidative DNA damage, antioxidants, and cancer. *BioEssays* *21*, 238–246.
- Conboy, J.G. (2017). RNA splicing during terminal erythropoiesis. *Curr. Opin. Hematol.* *24*, 215–221.
- Cortelezzi, A., Cattaneo, C., Cristiani, S., Duca, L., Sarina, B., Deliliers, G.L., Fiorelli, G., and Cappellini, M.D. (2000). Non-transferrin-bound iron in myelodysplastic syndromes: a marker of ineffective erythropoiesis? *Hematol. J.* *1*, 153–158.
- Cortez, D., Glick, G., and Elledge, S.J. (2004). Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *PNAS* *101*, 10078–10083.
- Corvelo, A., Hallegger, M., Smith, C.W.J., and Eyras, E. (2010). Genome-Wide Association between Branch Point Properties and Alternative Splicing. *PLOS Computational Biology* *6*, e1001016.
- Cretu, C., Schmitzová, J., Ponce-Salvatierra, A., Dybkov, O., De Laurentiis, E.I., Sharma, K., Will, C.L., Urlaub, H., Lührmann, R., and Pena, V. (2016). Molecular Architecture of SF3b and Structural Consequences of Its Cancer-Related Mutations. *Molecular Cell* *64*, 307–319.

Bibliographie

- Csere, P., Lill, R., and Kisfal, G. (1998). Identification of a human mitochondrial ABC transporter, the functional orthologue of yeast Atm1p. *FEBS Lett.* *441*, 266–270.
- Cui, R., Gale, R.P., Zhu, G., Xu, Z., Qin, T., Zhang, Y., Huang, G., Li, B., Fang, L., Zhang, H., et al. (2014). Serum iron metabolism and erythropoiesis in patients with myelodysplastic syndrome not receiving RBC transfusions. *Leukemia Research* *38*, 545–550.
- Damm, F., Kosmider, O., Gelsi-Boyer, V., Renneville, A., Carbuccia, N., Hidalgo-Curtis, C., Valle, V.D., Couronné, L., Scourzic, L., Chesnais, V., et al. (2012). Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. *Blood* *119*, 3211–3218.
- Darman, R.B., Seiler, M., Agrawal, A.A., Lim, K.H., Peng, S., Aird, D., Bailey, S.L., Bhavsar, E.B., Chan, B., Colla, S., et al. (2015). Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Reports* *13*, 1033–1045.
- Davies, A.A., Huttner, D., Daigaku, Y., Chen, S., and Ulrich, H.D. (2008). Activation of ubiquitin-dependent DNA damage bypass is mediated by replication protein a. *Mol. Cell* *29*, 625–636.
- De Maria, R., Zeuner, A., Eramo, A., Domenichelli, C., Bonci, D., Grignani, F., Srinivasula, S.M., Alnemri, E.S., Testa, U., and Peschle, C. (1999). Negative regulation of erythropoiesis by caspase-mediated cleavage of GATA-1. *Nature* *401*, 489–493.
- DeBoever, C., Ghia, E.M., Shepard, P.J., Rassenti, L., Barrett, C.L., Jepsen, K., Jamieson, C.H.M., Carson, D., Kipps, T.J., and Frazer, K.A. (2015). Transcriptome Sequencing Reveals Potential Mechanism of Cryptic 3' Splice Site Selection in SF3B1-mutated Cancers. *PLoS Comput Biol* *11*.
- Desmet, F.-O., Hamroun, D., Lalande, M., Collod-Bérout, G., Claustres, M., and Bérout, C. (2009). Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* *37*, e67–e67.
- Dolatshad, H., Pellagatti, A., Fernandez-Mercado, M., Yip, B.H., Malcovati, L., Attwood, M., Przychodzen, B., Sahgal, N., Kanapin, A.A., Lockstone, H., 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*, 1798.
- Dolatshad, H., Pellagatti, A., Liberante, F.G., Llorian, M., Repapi, E., Steeples, V., Roy, S., Scifo, L., Armstrong, R.N., Shaw, J., et al. (2016). Cryptic splicing events in the iron transporter *ABC7* and other key target genes in *SF3B1*-mutant myelodysplastic syndromes. *Leukemia* *30*, 2322–2331.
- Donovan, A., Brownlie, A., Zhou, Y., Shepard, J., Pratt, S.J., Moynihan, J., Paw, B.H., Drejer, A., Barut, B., Zapata, A., et al. (2000). Positional cloning of zebrafish ferroportin1 identifies a conserved vertebrate iron exporter. *Nature* *403*, 776–781.
- Donzelli, M., and Draetta, G.F. (2003). Regulating mammalian checkpoints through Cdc25 inactivation. *EMBO Rep* *4*, 671–677.
- Dulić, V., Kaufmann, W.K., Wilson, S.J., Tlsty, T.D., Lees, E., Harper, J.W., Elledge, S.J., and Reed, S.I. (1994). p53-dependent inhibition of cyclin-dependent kinase activities in human fibroblasts during radiation-induced G1 arrest. *Cell* *76*, 1013–1023.
- Dussiot, M., Maciel, T.T., Fricot, A., Chartier, C., Negre, O., Veiga, J., Grapton, D., Paubelle, E., Payen, E., Beuzard, Y., et al. (2014). An activin receptor IIA ligand trap corrects ineffective erythropoiesis in β -thalassemia. *Nature Medicine* *20*, 398–407.
- Ekert, P.G., Silke, J., and Vaux, D.L. (1999). Caspase inhibitors. *Cell Death Differ.* *6*, 1081–1086.

Bibliographie

- Fakan, S. (1994). Perichromatin fibrils are in situ forms of nascent transcripts. *Trends Cell Biol.* *4*, 86–90.
- Fan, L., Lagiseti, C., Edwards, C.C., Webb, T.R., and Potter, P.M. (2011). Sudemycins, Novel Small Molecule Analogues of FR901464, Induce Alternative Gene Splicing. *ACS Chemical Biology* *6*, 582–589.
- Felli, N., Pedini, F., Zeuner, A., Petrucci, E., Testa, U., Conticello, C., Biffoni, M., Di Cataldo, A., Winkles, J.A., Peschle, C., et al. (2005). Multiple members of the TNF superfamily contribute to IFN-gamma-mediated inhibition of erythropoiesis. *J. Immunol.* *175*, 1464–1472.
- Fenaux, P., and Ades, L. (2013). How we treat lower-risk myelodysplastic syndromes. *Blood* *121*, 4280–4286.
- Finger, C., Escher, C., and Schneider, D. (2009). The single transmembrane domains of human receptor tyrosine kinases encode self-interactions. *Sci Signal* *2*, ra56.
- Flach, J., Bakker, S.T., Mohrin, M., Conroy, P.C., Pietras, E.M., Reynaud, D., Alvarez, S., Diolaiti, M.E., Ugarte, F., Forsberg, E.C., et al. (2014). Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* *512*, 198–202.
- Flanagan, J.G., Chan, D.C., and Leder, P. Transmembrane Form of the kit Ligand Growth Factor Is Determined by Alternative Splicing and Is Missing in the Nd Mutant. *11*.
- Fleming, M.D., Romano, M.A., Su, M.A., Garrick, L.M., Garrick, M.D., and Andrews, N.C. (1998). Nramp2 is mutated in the anemic Belgrade (b) rat: evidence of a role for Nramp2 in endosomal iron transport. *Proc. Natl. Acad. Sci. U.S.A.* *95*, 1148–1153.
- Fontenay-Roupie, M., Bouscary, D., Guesnu, M., Picard, F., Melle, J., Lacombe, C., Gisselbrecht, S., Mayeux, P., and Dreyfus, F. (1999). Ineffective erythropoiesis in myelodysplastic syndromes: correlation with Fas expression but not with lack of erythropoietin receptor signal transduction. *Br. J. Haematol.* *106*, 464–473.
- Freire, R., van Vugt, M.A.T.M., Mamely, I., and Medema, R. (2006). Claspin: Timing the Cell Cycle Arrest When the Genome is Damaged. *Cell Cycle* *5*, 2831–2834.
- Frisan, E., Vandekerckhove, J., de Thonel, A., Pierre-Eugène, C., Sternberg, A., Arlet, J.-B., Floquet, C., Gyan, E., Kosmider, O., Dreyfus, F., et al. (2012). Defective nuclear localization of Hsp70 is associated with dyserythropoiesis and GATA-1 cleavage in myelodysplastic syndromes. *Blood* *119*, 1532–1542.
- Fu, X.-D. (2017). Exploiting the Hidden Treasure of Detained Introns. *Cancer Cell* *32*, 393–395.
- Fujiwara, Y., Browne, C.P., Cunniff, K., Goff, S.C., and Orkin, S.H. (1996). Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc. Natl. Acad. Sci. U.S.A.* *93*, 12355–12358.
- Furumai, R., Uchida, K., Komi, Y., Yoneyama, M., Ishigami, K., Watanabe, H., Kojima, S., and Yoshida, M. (2010). Spliceostatin A blocks angiogenesis by inhibiting global gene expression including VEGF. *Cancer Sci.* *101*, 2483–2489.
- Gaillard, H., García-Muse, T., and Aguilera, A. (2015). Replication stress and cancer. *Nat. Rev. Cancer* *15*, 276–289.
- Ganz, T., Jung, G., Naeim, A., Ginzburg, Y., Pakbaz, Z., Walter, P.B., Kautz, L., and Nemeth, E. (2017). Immunoassay for human serum erythroferrone. *Blood* *130*, 1243–1246.

Bibliographie

- García-Gómez, S., Reyes, A., Martínez-Jiménez, M.I., Chocrón, E.S., Mourón, S., Terrados, G., Powell, C., Salido, E., Méndez, J., Holt, I.J., et al. (2013). PrimPol, an archaic primase/polymerase operating in human cells. *Mol. Cell* 52, 541–553.
- Gardenghi, S., Ramos, P., Marongiu, M.F., Melchiorri, L., Breda, L., Guy, E., Muirhead, K., Rao, N., Roy, C.N., Andrews, N.C., et al. (2010). Heparin as a therapeutic tool to limit iron overload and improve anemia in β -thalassemic mice. *J Clin Invest* 120, 4466–4477.
- Gattermann, N., Finelli, C., Porta, M.D., Fenaux, P., Stadler, M., Guerci-Bresler, A., Schmid, M., Taylor, K., Vassilief, D., Habr, D., et al. (2012). Hematologic responses to deferasirox therapy in transfusion-dependent patients with myelodysplastic syndromes. *Haematologica* 97, 1364–1371.
- Gautier, E.-F., Ducamp, S., Leduc, M., Salnot, V., Guillonneau, F., Dussiot, M., Hale, J., Giarratana, M.-C., Raimbault, A., Douay, L., et al. (2016). Comprehensive Proteomic Analysis of Human Erythropoiesis. *Cell Rep* 16, 1470–1484.
- Ge, X.Q., and Blow, J.J. (2010). Chk1 inhibits replication factory activation but allows dormant origin firing in existing factories. *The Journal of Cell Biology* 191, 1285–1297.
- Genovese, G., Kähler, A.K., Handsaker, R.E., Lindberg, J., Rose, S.A., Bakhoum, S.F., Chambert, K., Mick, E., Neale, B.M., Fromer, M., et al. (2014). Clonal Hematopoiesis and Blood-Cancer Risk Inferred from Blood DNA Sequence. *New England Journal of Medicine* 371, 2477–2487.
- Gentien, D., Kosmider, O., Nguyen-Khac, F., Albaud, B., Rapinat, A., Dumont, A.G., Damm, F., Popova, T., Marais, R., Fontenay, M., et al. (2014). A common alternative splicing signature is associated with *SF3B1* mutations in malignancies from different cell lineages. *Leukemia* 28, 1355–1357.
- Gerstung, M., Pellagatti, A., Malcovati, L., Giagounidis, A., Porta, M.G.D., Jädersten, M., Dolatshad, H., Verma, A., Cross, N.C.P., Vyas, P., et al. (2015). Combining gene mutation with gene expression data improves outcome prediction in myelodysplastic syndromes. *Nat Commun* 6, 5901.
- Girard, C., Will, C.L., Peng, J., Makarov, E.M., Kastner, B., Lemm, I., Urlaub, H., Hartmuth, K., and Lührmann, R. (2012). Post-transcriptional spliceosomes are retained in nuclear speckles until splicing completion. *Nature Communications* 3, 994.
- Gobert, S., Duprez, V., Lacombe, C., Gisselbrecht, S., and Mayeux, P. (1995). The signal transduction pathway of erythropoietin involves three forms of mitogen-activated protein (MAP) kinase in UT7 erythroleukemia cells. *Eur. J. Biochem.* 234, 75–83.
- Gorgoulis, V.G., Vassiliou, L.-V.F., Karakaidos, P., Zacharatos, P., Kotsinas, A., Liloglou, T., Venere, M., Dittullo, R.A., Kastrinakis, N.G., Levy, B., et al. (2005). Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* 434, 907–913.
- Goswami, T., and Andrews, N.C. (2006). Hereditary hemochromatosis protein, HFE, interaction with transferrin receptor 2 suggests a molecular mechanism for mammalian iron sensing. *J. Biol. Chem.* 281, 28494–28498.
- Gozani, O., Potashkin, J., and Reed, R. (1998). A potential role for U2AF-SAP 155 interactions in recruiting U2 snRNP to the branch site. *Mol. Cell. Biol.* 18, 4752–4760.
- de Graaf, K., Czajkowska, H., Rottmann, S., Packman, L.C., Lilischkis, R., Lüscher, B., and Becker, W. (2006). The protein kinase DYRK1A phosphorylates the splicing factor SF3b1/SAP155 at Thr434, a novel in vivo phosphorylation site. *BMC Biochem* 7, 7.

Bibliographie

- Graubert, T.A., Shen, D., Ding, L., Okeyo-Owuor, T., Lunn, C.L., Shao, J., Krysiak, K., Harris, C.C., Koboldt, D.C., Larson, D.E., et al. (2011). Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat. Genet.* *44*, 53–57.
- Greenberg, P., Cox, C., LeBeau, M.M., Fenaux, P., Morel, P., Sanz, G., Sanz, M., Vallespi, T., Hamblin, T., Oscier, D., et al. (1997). International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood* *89*, 2079–2088.
- Greenberg, P.L., Tuechler, H., Schanz, J., Sanz, G., Garcia-Manero, G., Solé, F., Bennett, J.M., Bowen, D., Fenaux, P., Dreyfus, F., et al. (2012). Revised International Prognostic Scoring System for Myelodysplastic Syndromes. *Blood* *120*, 2454–2465.
- Gregory, T., Yu, C., Ma, A., Orkin, S.H., Blobel, G.A., and Weiss, M.J. (1999). GATA-1 and erythropoietin cooperate to promote erythroid cell survival by regulating bcl-xL expression. *Blood* *94*, 87–96.
- Groves, M.R., Hanlon, N., Turowski, P., Hemmings, B.A., and Barford, D. (1999). The Structure of the Protein Phosphatase 2A PR65/A Subunit Reveals the Conformation of Its 15 Tandemly Repeated HEAT Motifs. *Cell* *96*, 99–110.
- Guernsey, D.L., Jiang, H., Campagna, D.R., Evans, S.C., Ferguson, M., Kellogg, M.D., Lachance, M., Matsuoka, M., Nightingale, M., Rideout, A., et al. (2009). Mutations in mitochondrial carrier family gene SLC25A38 cause nonsyndromic autosomal recessive congenital sideroblastic anemia. *Nat. Genet.* *41*, 651–653.
- Guilliam, T.A., Brissett, N.C., Ehlinger, A., Keen, B.A., Kolesar, P., Taylor, E.M., Bailey, L.J., Lindsay, H.D., Chazin, W.J., and Doherty, A.J. (2017). Molecular basis for PrimPol recruitment to replication forks by RPA. *Nature Communications* *8*, 15222.
- Gunshin, H., Mackenzie, B., Berger, U.V., Gunshin, Y., Romero, M.F., Boron, W.F., Nussberger, S., Gollan, J.L., and Hediger, M.A. (1997). Cloning and characterization of a mammalian proton-coupled metal-ion transporter. *Nature* *388*, 482–488.
- Gyan, E., Frisan, E., Beyne-Rauzy, O., Deschemin, J.-C., Pierre-Eugene, C., Randriamampita, C., Dubart-Kupperschmitt, A., Garrido, C., Dreyfus, F., Mayeux, P., et al. (2008). Spontaneous and Fas-induced apoptosis of low-grade MDS erythroid precursors involves the endoplasmic reticulum. *Leukemia* *22*, 1864–1873.
- Haile, D.J., Rouault, T.A., Tang, C.K., Chin, J., Harford, J.B., and Klausner, R.D. (1992). Reciprocal control of RNA-binding and aconitase activity in the regulation of the iron-responsive element binding protein: role of the iron-sulfur cluster. *Proc. Natl. Acad. Sci. U.S.A.* *89*, 7536–7540.
- Hanada, K., Budzowska, M., Davies, S.L., van Drunen, E., Onizawa, H., Beverloo, H.B., Maas, A., Essers, J., Hickson, I.D., and Kanaar, R. (2007). The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nature Structural & Molecular Biology* *14*, 1096–1104.
- Hanspal, M., Smockova, Y., and Uong, Q. (1998). Molecular identification and functional characterization of a novel protein that mediates the attachment of erythroblasts to macrophages. *Blood* *92*, 2940–2950.
- Harrison, P.M., Fischbach, F.A., Hoy, T.G., and Haggis, G.H. (1967). Ferric oxyhydroxide core of ferritin. *Nature* *216*, 1188–1190.

Bibliographie

- Hartono, S.R., Malapert, A., Legros, P., Bernard, P., Chédin, F., and Vanoosthuyse, V. (2018). The Affinity of the S9.6 Antibody for Double-Stranded RNAs Impacts the Accurate Mapping of R-Loops in Fission Yeast. *J. Mol. Biol.* *430*, 272–284.
- Heinrich, M.C., Dooley, D.C., Freed, A.C., Band, L., Hoatlin, M.E., Keeble, W.W., Peters, S.T., Silvey, K.V., Ey, F.S., and Kabat, D. (1993). Constitutive expression of steel factor gene by human stromal cells. *Blood* *82*, 771–783.
- Horrigan, S.K., Arbieva, Z.H., Xie, H.Y., Kravarusic, J., Fulton, N.C., Naik, H., Le, T.T., and Westbrook, C.A. (2000). Delineation of a minimal interval and identification of 9 candidates for a tumor suppressor gene in malignant myeloid disorders on 5q31. *Blood* *95*, 2372–2377.
- Hu, J., Liu, J., Xue, F., Halverson, G., Reid, M., Guo, A., Chen, L., Raza, A., Galili, N., Jaffray, J., et al. (2013). Isolation and functional characterization of human erythroblasts at distinct stages: implications for understanding of normal and disordered erythropoiesis in vivo. *Blood* *121*, 3246–3253.
- Huang, Y., Hale, J., Wang, Y., Li, W., Zhang, S., Zhang, J., Zhao, H., Guo, X., Liu, J., Yan, H., et al. (2018). SF3B1 deficiency impairs human erythropoiesis via activation of p53 pathway: implications for understanding of ineffective erythropoiesis in MDS. *J Hematol Oncol* *11*.
- Hubert, N., and Hentze, M.W. (2002). Previously uncharacterized isoforms of divalent metal transporter (DMT)-1: implications for regulation and cellular function. *Proc. Natl. Acad. Sci. U.S.A.* *99*, 12345–12350.
- Hustedt, N., and Durocher, D. (2016). The control of DNA repair by the cell cycle. *Nat. Cell Biol.* *19*, 1–9.
- Ilgan, J.O., Ramakrishnan, A., Hayes, B., Murphy, M.E., Zebari, A.S., Bradley, P., and Bradley, R.K. (2015). *U2AF1* mutations alter splice site recognition in hematological malignancies. *Genome Research* *25*, 14–26.
- Ishiai, M., Kitao, H., Smogorzewska, A., Tomida, J., Kinomura, A., Uchida, E., Saberi, A., Kinoshita, E., Kinoshita-Kikuta, E., Koike, T., et al. (2008). FANCI phosphorylation functions as a molecular switch to turn on the Fanconi anemia pathway. *Nat. Struct. Mol. Biol.* *15*, 1138–1146.
- Isono, K. (2005). Mammalian Polycomb-mediated repression of Hox genes requires the essential spliceosomal protein Sf3b1. *Genes & Development* *19*, 536–541.
- Jabbour, E., Kantarjian, H.M., Koller, C., and Taher, A. (2008). Red blood cell transfusions and iron overload in the treatment of patients with myelodysplastic syndromes. *Cancer* *112*, 1089–1095.
- Jaiswal, S., Fontanillas, P., Flannick, J., Manning, A., Grauman, P.V., Mar, B.G., Lindsley, R.C., Mermel, C.H., Burt, N., Chavez, A., et al. (2014). Age-Related Clonal Hematopoiesis Associated with Adverse Outcomes. *New England Journal of Medicine* *371*, 2488–2498.
- Jenkins, J.L., and Kielkopf, C.L. (2017). Splicing Factor Mutations in Myelodysplasias: Insights from Spliceosome Structures. *Trends in Genetics* *33*, 336–348.
- Ji, P., Yeh, V., Ramirez, T., Murata-Hori, M., and Lodish, H.F. (2010). Histone deacetylase 2 is required for chromatin condensation and subsequent enucleation of cultured mouse fetal erythroblasts. *Haematologica* *95*, 2013–2021.
- Ji, X., Zhou, Y., Pandit, S., Huang, J., Li, H., Lin, C.Y., Xiao, R., Burge, C.B., and Fu, X.-D. (2013). SR Proteins Collaborate with 7SK and Promoter-Associated Nascent RNA to Release Paused Polymerase. *Cell* *153*, 855–868.

Bibliographie

- Joshi, P., Halene, S., and Abdel-Wahab, O. (2017). How do messenger RNA splicing alterations drive myelodysplasia? *Blood* *129*, 2465–2470.
- Kahles, A., Lehmann, K.-V., Toussaint, N.C., Hüser, M., Stark, S.G., Sachsenberg, T., Stegle, O., Kohlbacher, O., Sander, C., Caesar-Johnson, S.J., et al. (2018). Comprehensive Analysis of Alternative Splicing Across Tumors from 8,705 Patients. *Cancer Cell* *34*, 211-224.e6.
- Kaida, D., Motoyoshi, H., Tashiro, E., Nojima, T., Hagiwara, M., Ishigami, K., Watanabe, H., Kitahara, T., Yoshida, T., Nakajima, H., et al. (2007). Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nature Chemical Biology* *3*, 576–583.
- Kautz, L. (2014). [Erythroferrone, an erythroid regulator of iron metabolism]. *Med Sci (Paris)* *30*, 834–836.
- Kautz, L., Jung, G., Valore, E.V., Rivella, S., Nemeth, E., and Ganz, T. (2014). Identification of erythroferrone as an erythroid regulator of iron metabolism. *Nature Genetics* *46*, 678–684.
- Kavanaugh, G., Ye, F., Mohni, K.N., Luzwick, J.W., Glick, G., and Cortez, D. (2015). A whole genome RNAi screen identifies replication stress response genes. *DNA Repair (Amst.)* *35*, 55–62.
- Kawabata, H., Yang, R., Hiramata, T., Vuong, P.T., Kawano, S., Gombart, A.F., and Koeffler, H.P. (1999). Molecular cloning of transferrin receptor 2. A new member of the transferrin receptor-like family. *J. Biol. Chem.* *274*, 20826–20832.
- Kawane, K., Fukuyama, H., Kondoh, G., Takeda, J., Ohsawa, Y., Uchiyama, Y., and Nagata, S. (2001). Requirement of DNase II for definitive erythropoiesis in the mouse fetal liver. *Science* *292*, 1546–1549.
- Keimpema, M. van, Grüneberg, L.J., Mokry, M., Boxtel, R. van, Koster, J., Coffey, P.J., Pals, S.T., and Spaargaren, M. (2014). FOXP1 directly represses transcription of proapoptotic genes and cooperates with NF- κ B to promote survival of human B cells. *Blood* *124*, 3431–3440.
- Kemp, M.G., Akan, Z., Yilmaz, S., Grillo, M., Smith-Roe, S.L., Kang, T.-H., Cordeiro-Stone, M., Kaufmann, W.K., Abraham, R.T., Sancar, A., et al. (2010). Tipin-Replication Protein A Interaction Mediates Chk1 Phosphorylation by ATR in Response to Genotoxic Stress. *J Biol Chem* *285*, 16562–16571.
- Kennedy, S.A., Frazier, M.L., Steiniger, M., Mast, A.M., Marzluff, W.F., and Redinbo, M.R. (2009). Crystal Structure of the HEAT Domain from the Pre-mRNA Processing Factor Symplekin. *J Mol Biol* *392*, 115–128.
- Kerdivel, G., Chesnais, V., Becht, E., Toma, A., Cagnard, N., Dumont, F., Rousseau, A., Fenaux, P., Chevret, S., Chapuis, N., et al. (2018). Lenalidomide-mediated erythroid improvement in non-del(5q) myelodysplastic syndromes is associated with bone marrow immuno-remodeling. *Leukemia* *32*, 558–562.
- Kfir, N., Lev-Maor, G., Gleich, O., Alajem, A., Datta, A., Sze, S.K., Meshorer, E., and Ast, G. (2015). SF3B1 Association with Chromatin Determines Splicing Outcomes. *Cell Reports* *11*, 618–629.
- Kim, A.R., Ulirsch, J.C., Wilmes, S., Unal, E., Moraga, I., Karakukcu, M., Yuan, D., Kazerounian, S., Abdulhay, N.J., King, D.S., et al. (2017). Functional Selectivity in Cytokine Signaling Revealed Through a Pathogenic EPO Mutation. *Cell* *168*, 1053-1064.e15.
- Kim, E., Ilagan, J.O., Liang, Y., Daubner, G.M., Lee, S.C.-W., Ramakrishnan, A., Li, Y., Chung, Y.R., Micol, J.-B., Murphy, M.E., et al. (2015). SRSF2 Mutations Contribute to Myelodysplasia by Mutant-Specific Effects on Exon Recognition. *Cancer Cell* *27*, 617–630.

Bibliographie

- Kingsley, P.D., Malik, J., Fantauzzo, K.A., and Palis, J. (2004). Yolk sac-derived primitive erythroblasts enucleate during mammalian embryogenesis. *Blood* 104, 19–25.
- Koch, P.-S., Olsavszky, V., Ulbrich, F., Sticht, C., Demory, A., Leibing, T., Henzler, T., Meyer, M., Zierow, J., Schneider, S., et al. (2017). Angiocrine Bmp2 signaling in murine liver controls normal iron homeostasis. *Blood* 129, 415–419.
- Kohtz, J.D., Jamison, S.F., Will, C.L., Zuo, P., Lührmann, R., Garcia-Blanco, M.A., and Manley, J.L. (1994). Protein-protein interactions and 5'-splice-site recognition in mammalian mRNA precursors. *Nature* 368, 119–124.
- Kolinjivadi, A.M., Sannino, V., De Antoni, A., Zadorozhny, K., Kilkenny, M., Técher, H., Baldi, G., Shen, R., Ciccia, A., Pellegrini, L., et al. (2017). Smarcal1-Mediated Fork Reversal Triggers Mre11-Dependent Degradation of Nascent DNA in the Absence of Brca2 and Stable Rad51 Nucleofilaments. *Mol. Cell* 67, 867–881.e7.
- Krishna, S., Wagener, B.M., Liu, H.P., Lo, Y.-C., Sterk, R., Petrini, J.H.J., and Nickoloff, J.A. (2007). Mre11 and Ku Regulation of Double-Strand Break Repair by Gene Conversion and Break-Induced Replication. *DNA Repair (Amst)* 6, 797–808.
- Krönke, J., Fink, E.C., Hollenbach, P.W., MacBeth, K.J., Hurst, S.N., Udeshi, N.D., Chamberlain, P.P., Mani, D.R., Man, H.W., Gandhi, A.K., et al. (2015). Lenalidomide induces ubiquitination and degradation of CK1 α in del(5q) MDS. *Nature* 523, 183–188.
- Krystal, G. (1994). Transforming growth factor beta 1 is an inducer of erythroid differentiation. *Journal of Experimental Medicine* 180, 851–860.
- Kumagai, A., and Dunphy, W.G. (2000). Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in *Xenopus* egg extracts. *Mol. Cell* 6, 839–849.
- Kumagai, A., Lee, J., Yoo, H.Y., and Dunphy, W.G. (2006). TopBP1 activates the ATR-ATRIP complex. *Cell* 124, 943–955.
- Lacombe, C., Da Silva, J.L., Bruneval, P., Fournier, J.G., Wendling, F., Casadevall, N., Camilleri, J.P., Bariety, J., Varet, B., and Tambourin, P. (1988). Peritubular cells are the site of erythropoietin synthesis in the murine hypoxic kidney. *J. Clin. Invest.* 81, 620–623.
- Lamarche, B.J., Orazio, N.I., and Weitzman, M.D. (2010). The MRN complex in Double-Strand Break Repair and Telomere Maintenance. *FEBS Lett* 584, 3682–3695.
- Lambert, S., and Lopez, B.S. (2002). Inactivation of the RAD51 recombination pathway stimulates UV-induced mutagenesis in mammalian cells. *Oncogene* 21, 4065–4069.
- Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. *Nature Reviews Molecular Cell Biology* 4, 605–612.
- Lavin, M.F., Kozlov, S., Gatei, M., and Kijas, A.W. (2015). ATM-Dependent Phosphorylation of All Three Members of the MRN Complex: From Sensor to Adaptor. *Biomolecules* 5, 2877–2902.
- Layer, G., Reichelt, J., Jahn, D., and Heinz, D.W. (2010). Structure and function of enzymes in heme biosynthesis. *Protein Sci* 19, 1137–1161.
- Le Tallec, B., Koundrioukoff, S., Wilhelm, T., Letessier, A., Brison, O., and Debatisse, M. (2014). Updating the mechanisms of common fragile site instability: how to reconcile the different views? *Cell Mol Life Sci* 71, 4489–4494.

Bibliographie

- Lee, S.C.-W., Dvinge, H., Kim, E., Cho, H., Micol, J.-B., Chung, Y.R., Durham, B.H., Yoshimi, A., Kim, Y.J., Thomas, M., et al. (2016). Modulation of splicing catalysis for therapeutic targeting of leukemia with mutations in genes encoding spliceosomal proteins. *Nat. Med.* *22*, 672–678.
- Lefebvre, T., Dessendier, N., Houamel, D., Ialy-Radio, N., Kannengiesser, C., Manceau, H., Beaumont, C., Nicolas, G., Gouya, L., Puy, H., et al. (2015). LC-MS/MS method for hepcidin-25 measurement in human and mouse serum: clinical and research implications in iron disorders. *Clinical Chemistry and Laboratory Medicine (CCLM)* *53*, 1557–1567.
- Leimberg, M.J., Prus, E., Konijn, A.M., and Fibach, E. (2008). Macrophages function as a ferritin iron source for cultured human erythroid precursors. *J. Cell. Biochem.* *103*, 1211–1218.
- Leipuviene, R., and Theil, E.C. (2007). The family of iron responsive RNA structures regulated by changes in cellular iron and oxygen. *Cell. Mol. Life Sci.* *64*, 2945–2955.
- Leitch, H.A. (2007). Improving clinical outcome in patients with myelodysplastic syndrome and iron overload using iron chelation therapy. *Leukemia Research* *31*, S7–S9.
- Leong, W.Y., Nguyen, H.D., Li, W., Walter, M.J., Zou, L., and Graubert, T.A. (2017). Spliceosomal Mutations Induce R Loop-Associated ATR Signaling. *Blood* *130*, 116–116.
- Li, X., and Manley, J.L. (2005). Inactivation of the SR protein splicing factor ASF/SF2 results in genomic instability. *Cell* *122*, 365–378.
- Li, J., Hale, J., Bhagia, P., Xue, F., Chen, L., Jaffray, J., Yan, H., Lane, J., Gallagher, P.G., Mohandas, N., et al. (2014). Isolation and transcriptome analyses of human erythroid progenitors: BFU-E and CFU-E. *Blood* *124*, 3636–3645.
- Lin, L., Goldberg, Y.P., and Ganz, T. (2005). Competitive regulation of hepcidin mRNA by soluble and cell-associated hemojuvelin. *Blood* *106*, 2884–2889.
- Lin, L., Valore, E.V., Nemeth, E., Goodnough, J.B., Gabayan, V., and Ganz, T. (2007). Iron transferrin regulates hepcidin synthesis in primary hepatocyte culture through hemojuvelin and BMP2/4. *Blood* *110*, 2182–2189.
- List, A. (2018). The Medalist Trial: Results of a Phase 3, Randomized, Double-Blind, Placebo-Controlled Study of Luspatercept to Treat Anemia in Patients with Very Low-, Low-, or Intermediate-Risk Myelodysplastic Syndromes (MDS) with Ring Sideroblasts (RS) Who Require Red Blood Cell (RBC) Transfusions. (ASH), p.
- List, A.F., Baer, M.R., Steensma, D.P., Raza, A., Esposito, J., Martinez-Lopez, N., Paley, C., Feigert, J., and Besa, E. (2012). Deferasirox Reduces Serum Ferritin and Labile Plasma Iron in RBC Transfusion-Dependent Patients With Myelodysplastic Syndrome. *JCO* *30*, 2134–2139.
- Liu, H.X., Chew, S.L., Cartegni, L., Zhang, M.Q., and Krainer, A.R. (2000). Exonic splicing enhancer motif recognized by human SC35 under splicing conditions. *Mol. Cell. Biol.* *20*, 1063–1071.
- Liu, Q., Davidoff, O., Niss, K., and Haase, V.H. (2012). Hypoxia-inducible factor regulates hepcidin via erythropoietin-induced erythropoiesis. *J. Clin. Invest.* *122*, 4635–4644.
- Lopes, M., Foiani, M., and Sogo, J.M. (2006). Multiple mechanisms control chromosome integrity after replication fork uncoupling and restart at irreparable UV lesions. *Mol. Cell* *21*, 15–27.
- Lossaint, G., Larroque, M., Ribeyre, C., Bec, N., Larroque, C., Décaillet, C., Gari, K., and Constantinou, A. (2013). FANCD2 binds MCM proteins and controls replisome function upon activation of s phase checkpoint signaling. *Mol. Cell* *51*, 678–690.

Bibliographie

- Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15, 550.
- Macheret, M., and Halazonetis, T.D. (2015). DNA replication stress as a hallmark of cancer. *Annu Rev Pathol* 10, 425–448.
- Madan, V., Kanojia, D., Li, J., Okamoto, R., Sato-Otsubo, A., Kohlmann, A., Sanada, M., Grossmann, V., Sundaresan, J., Shiraishi, Y., et al. (2015). Aberrant splicing of U12-type introns is the hallmark of ZRSR2 mutant myelodysplastic syndrome. *Nature Communications* 6, 6042.
- Mailand, N., Bekker-Jensen, S., Bartek, J., and Lukas, J. (2006). Destruction of Claspin by SCFbetaTrCP restrains Chk1 activation and facilitates recovery from genotoxic stress. *Mol. Cell* 23, 307–318.
- Malcovati, L., and Cazzola, M. (2016). Recent advances in the understanding of myelodysplastic syndromes with ring sideroblasts. *Br. J. Haematol.* 174, 847–858.
- Malcovati, L., Porta, M.G.D., Pascutto, C., Invernizzi, R., Boni, M., Travaglino, E., Passamonti, F., Arcaini, L., Maffioli, M., Bernasconi, P., et al. (2005). Prognostic Factors and Life Expectancy in Myelodysplastic Syndromes Classified According to WHO Criteria: A Basis for Clinical Decision Making. *JCO* 23, 7594–7603.
- Malcovati, L., Papaemmanuil, E., Bowen, D.T., Boulton, J., Porta, M.G.D., Pascutto, C., Travaglino, E., Groves, M.J., Godfrey, A.L., Ambaglio, I., et al. (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., Jädersten, M., Jansson, M., Elena, C., Galli, A., Walldin, G., Porta, M.G.D., et al. (2015). SF3B1 mutation identifies a distinct subset of myelodysplastic syndrome with ring sideroblasts. *Blood* 126, 233–241.
- Manwani, D., and Bieker, J.J. (2008). The Erythroblastic Island. *Curr Top Dev Biol* 82, 23–53.
- Martin, M., Maßhöfer, L., Temming, P., Rahmann, S., Metz, C., Bornfeld, N., van de Nes, J., Klein-Hitpass, L., Hinnebusch, A.G., Horsthemke, B., et al. (2013). Exome sequencing identifies recurrent somatic mutations in EIF1AX and SF3B1 in uveal melanoma with disomy 3. *Nat. Genet.* 45, 933–936.
- Mathis, S., Chapuis, N., Debord, C., Rouquette, A., Radford-Weiss, I., Park, S., Dreyfus, F., Lacombe, C., Béné, M.C., Kosmider, O., et al. (2013). Flow cytometric detection of dyserythropoiesis: a sensitive and powerful diagnostic tool for myelodysplastic syndromes. *Leukemia* 27, 1981–1987.
- Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y., et al. (2007). ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 316, 1160–1166.
- Matthes, T.W., Meyer, G., Samii, K., and Beris, P. (2000). Increased apoptosis in acquired sideroblastic anaemia. *Br. J. Haematol.* 111, 843–852.
- Maxwell, A.P., Lappin, T.R., Johnston, C.F., Bridges, J.M., and McGeown, M.G. (1990). Erythropoietin production in kidney tubular cells. *Br. J. Haematol.* 74, 535–539.
- McKie, A.T., Marciani, P., Rolfs, A., Brennan, K., Wehr, K., Barrow, D., Miret, S., Bomford, A., Peters, T.J., Farzaneh, F., et al. (2000). A novel duodenal iron-regulated transporter, IREG1, implicated in the basolateral transfer of iron to the circulation. *Mol. Cell* 5, 299–309.

Bibliographie

- McKie, A.T., Barrow, D., Latunde-Dada, G.O., Rolfs, A., Sager, G., Mudaly, E., Mudaly, M., Richardson, C., Barlow, D., Bomford, A., et al. (2001). An iron-regulated ferric reductase associated with the absorption of dietary iron. *Science* 291, 1755–1759.
- Metzgeroth Georgia, Rosée Paul La, Kuhn Christian, Schultheis Beate, Dorn-Beineke Alexandra, Hehlmann Rüdiger, and Hastka Jan (2007). The soluble transferrin receptor in dysplastic erythropoiesis in myelodysplastic syndrome. *European Journal of Haematology* 79, 8–16.
- Meunier, M., Ancelet, S., Lefebvre, C., Arnaud, J., Garrel, C., Pezet, M., Wang, Y., Faure, P., Szymanski, G., Duployez, N., et al. (2017). Reactive oxygen species levels control NF- κ B activation by low dose deferasirox in erythroid progenitors of low risk myelodysplastic syndromes. *Oncotarget* 8, 105510–105524.
- Mian, S.A., Rouault-Pierre, K., Smith, A.E., Seidl, T., Pizzitola, I., Kizilors, A., Kulasekararaj, A.G., Bonnet, D., and Mufti, G.J. (2015). SF3B1 mutant MDS-initiating cells may arise from the haematopoietic stem cell compartment. *Nature Communications* 6, 10004.
- Misteli, T., and Spector, D.L. (1999). RNA polymerase II targets pre-mRNA splicing factors to transcription sites in vivo. *Mol. Cell* 3, 697–705.
- Moraga, I., Wernig, G., Wilmes, S., Gryshkova, V., Richter, C.P., Hong, W.-J., Sinha, R., Guo, F., Fabionar, H., Wehrman, T.S., et al. (2015). Tuning Cytokine Receptor Signaling by Re-orienting Dimer Geometry with Surrogate Ligands. *Cell* 160, 1196–1208.
- Moriel-Carretero, M., Ovejero, S., Gêrus-Durand, M., Vryzas, D., and Constantinou, A. (2017). Fanconi anemia FANCD2 and FANCI proteins regulate the nuclear dynamics of splicing factors. *J. Cell Biol.* 216, 4007–4026.
- Mortera-Blanco, T., Dimitriou, M., Woll, P.S., Karimi, M., Elvarsdottir, E., Conte, S., Tobiasson, M., Jansson, M., Douagi, I., Moarii, M., et al. (2017). SF3B1-initiating mutations in MDS-RSs target lymphomyeloid hematopoietic stem cells. *Blood* 130, 881–890.
- Muckenthaler, M., Gray, N.K., and Hentze, M.W. (1998). IRP-1 binding to ferritin mRNA prevents the recruitment of the small ribosomal subunit by the cap-binding complex eIF4F. *Mol. Cell* 2, 383–388.
- Mufti, G.J., Bennett, J.M., Goasguen, J., Bain, B.J., Baumann, I., Brunning, R., Cazzola, M., Fenaux, P., Germing, U., Hellström-Lindberg, E., 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, 1712–1717.
- Munugalavadla, V., Dore, L.C., Tan, B.L., Hong, L., Vishnu, M., Weiss, M.J., and Kapur, R. (2005). Repression of c-Kit and Its Downstream Substrates by GATA-1 Inhibits Cell Proliferation during Erythroid Maturation. *Mol Cell Biol* 25, 6747–6759.
- Murthy, T., Bluemn, T., Gupta, A.K., Reimer, M., Rao, S., Pillai, M.M., and Minella, A.C. (2018). Cyclin-dependent kinase 1 (CDK1) and CDK2 have opposing roles in regulating interactions of splicing factor 3B1 with chromatin. *J. Biol. Chem.* 293, 10220–10234.
- Nandakumar, S.K., Ulirsch, J.C., and Sankaran, V.G. (2016). Advances in understanding erythropoiesis: evolving perspectives. *British Journal of Haematology* 173, 206–218.
- Nemeth, E., and Ganz, T. (2006). Regulation of Iron Metabolism by Hcpidin. *Annual Review of Nutrition* 26, 323–342.

Bibliographie

- Nemeth, E., Valore, E.V., Territo, M., Schiller, G., Lichtenstein, A., and Ganz, T. (2003). Heparin, a putative mediator of anemia of inflammation, is a type II acute-phase protein. *Blood* 101, 2461–2463.
- Nemeth, E., Tuttle, M.S., Powelson, J., Vaughn, M.B., Donovan, A., Ward, D.M., Ganz, T., and Kaplan, J. (2004a). Heparin Regulates Cellular Iron Efflux by Binding to Ferroportin and Inducing Its Internalization. *Science* 306, 2090–2093.
- Nemeth, E., Rivera, S., Gabayan, V., Keller, C., Taudorf, S., Pedersen, B.K., and Ganz, T. (2004b). IL-6 mediates hypoferrremia of inflammation by inducing the synthesis of the iron regulatory hormone heparin. *J Clin Invest* 113, 1271–1276.
- Neubauer, H., Cumano, A., Müller, M., Wu, H., Huffstadt, U., and Pfeffer, K. (1998). Jak2 deficiency defines an essential developmental checkpoint in definitive hematopoiesis. *Cell* 93, 397–409.
- Neukirchen, J., Schoonen, W.M., Strupp, C., Gattermann, N., Aul, C., Haas, R., and Germing, U. (2011). Incidence and prevalence of myelodysplastic syndromes: data from the Düsseldorf MDS-registry. *Leuk. Res.* 35, 1591–1596.
- Nguyen, H.D., Yadav, T., Giri, S., Saez, B., Graubert, T.A., and Zou, L. (2017). Functions of Replication Protein A as a Sensor of R Loops and a Regulator of RNaseH1. *Mol. Cell* 65, 832–847.e4.
- Nguyen, H.D., Leong, W.Y., Li, W., Reddy, P.N.G., Sullivan, J.D., Walter, M.J., Zou, L., and Graubert, T.A. (2018). Spliceosome Mutations Induce R Loop-Associated Sensitivity to ATR Inhibition in Myelodysplastic Syndromes. *Cancer Res.* 78, 5363–5374.
- Nicolas, G., Bennoun, M., Devaux, I., Beaumont, C., Grandchamp, B., Kahn, A., and Vaulont, S. (2001). Lack of heparin gene expression and severe tissue iron overload in upstream stimulatory factor 2 (USF2) knockout mice. *Proc. Natl. Acad. Sci. U.S.A.* 98, 8780–8785.
- Nicolas, G., Bennoun, M., Porteu, A., Mativet, S., Beaumont, C., Grandchamp, B., Sirito, M., Sawadogo, M., Kahn, A., and Vaulont, S. (2002). Severe iron deficiency anemia in transgenic mice expressing liver heparin. *Proc. Natl. Acad. Sci. U.S.A.* 99, 4596–4601.
- Nicolas, G., Viatte, L., Lou, D.-Q., Bennoun, M., Beaumont, C., Kahn, A., Andrews, N.C., and Vaulont, S. (2003). Constitutive heparin expression prevents iron overload in a mouse model of hemochromatosis. *Nature Genetics* 34, 97–101.
- Nikpour, M., Scharenberg, C., Liu, A., Conte, S., Karimi, M., Mortera-Blanco, T., Gai, V., Fernandez-Mercado, M., Papaemmanuil, E., Höglstrand, K., et al. (2013). The transporter *ABCB7* is a mediator of the phenotype of acquired refractory anemia with ring sideroblasts. *Leukemia* 27, 889–896.
- Notta, F., Zandi, S., Takayama, N., Dobson, S., Gan, O.I., Wilson, G., Kaufmann, K.B., McLeod, J., Laurenti, E., Dunant, C.F., et al. (2016). Distinct routes of lineage development reshape the human blood hierarchy across ontogeny. *Science* 351, aab2116–aab2116.
- Obeng, E.A., Chappell, R.J., Seiler, M., Chen, M.C., Campagna, D.R., Schmidt, P.J., Schneider, R.K., Lord, A.M., Wang, L., Gambe, R.G., et al. (2016). Physiologic Expression of Sf3b1K700E Causes Impaired Erythropoiesis, Aberrant Splicing, and Sensitivity to Therapeutic Spliceosome Modulation. *Cancer Cell* 30, 404–417.
- Ogawa, M., Matsuzaki, Y., Nishikawa, S., Hayashi, S., Kunisada, T., Sudo, T., Kina, T., Nakauchi, H., and Nishikawa, S. (1991). Expression and function of c-kit in hemopoietic progenitor cells. *J. Exp. Med.* 174, 63–71.

Bibliographie

- Olson, E., Nievera, C.J., Klimovich, V., Fanning, E., and Wu, X. (2006). RPA2 is a direct downstream target for ATR to regulate the S-phase checkpoint. *J. Biol. Chem.* *281*, 39517–39533.
- Omar, S., Feki, M., and Kaabachi, N. (2006). Le métabolisme du fer : revue générale et récents développements. *Annales de Biologie Clinique* *64*, 523–534.
- Origa, R., Galanello, R., Ganz, T., Giagu, N., Maccioni, L., Faa, G., and Nemeth, E. (2007). Liver iron concentrations and urinary hepcidin in β -thalassemia. *Haematologica* *92*, 583–588.
- Orkin, S.H. (1992). GATA-binding transcription factors in hematopoietic cells. *Blood* *80*, 575–581.
- Osaki, S., Johnson, D.A., and Frieden, E. (1966). The possible significance of the ferrous oxidase activity of ceruloplasmin in normal human serum. *J. Biol. Chem.* *241*, 2746–2751.
- Pak, M., Lopez, M.A., Gabayan, V., Ganz, T., and Rivera, S. (2006a). Suppression of hepcidin during anemia requires erythropoietic activity. *Blood* *108*, 3730–3735.
- Pak, M., Lopez, M.A., Gabayan, V., Ganz, T., and Rivera, S. (2006b). Suppression of hepcidin during anemia requires erythropoietic activity. *Blood* *108*, 3730–3735.
- Palis, J. (2014). Primitive and definitive erythropoiesis in mammals. *Front Physiol* *5*.
- Palis, J., Chan, R.J., Koniski, A., Patel, R., Starr, M., and Yoder, M.C. (2001). Spatial and temporal emergence of high proliferative potential hematopoietic precursors during murine embryogenesis. *Proc Natl Acad Sci U S A* *98*, 4528–4533.
- Pan, Q., Shai, O., Lee, L.J., Frey, B.J., and Blencowe, B.J. (2008). Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nature Genetics* *40*, 1413–1415.
- Papaemmanuil, E., Cazzola, M., Boulton, J., Malcovati, L., Vyas, P., Bowen, D., Pellagatti, A., Wainscoat, J.S., Hellstrom-Lindberg, E., Gambacorti-Passerini, C., et al. (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., Loo, P.V., Yoon, C.J., Ellis, P., Wedge, D.C., Pellagatti, A., et al. (2013). Clinical and biological implications of driver mutations in myelodysplastic syndromes. *Blood* *122*, 3616–3627.
- Parajuli, S., Teasley, D.C., Murali, B., Jackson, J., Vindigni, A., and Stewart, S.A. (2017). Human ribonuclease H1 resolves R-loops and thereby enables progression of the DNA replication fork. *J. Biol. Chem.* *292*, 15216–15224.
- Park, C.H., Valore, E.V., Waring, A.J., and Ganz, T. (2001). Hepcidin, a urinary antimicrobial peptide synthesized in the liver. *J. Biol. Chem.* *276*, 7806–7810.
- Park, S., Grabar, S., Kelaidi, C., Beyne-Rauzy, O., Picard, F., Bardet, V., Coiteux, V., Leroux, G., Lepelley, P., Daniel, M.-T., et al. (2008). Predictive factors of response and survival in myelodysplastic syndrome treated with erythropoietin and G-CSF: the GFM experience. *Blood* *111*, 574–582.
- Pasricha, S.-R., Frazer, D.M., Bowden, D.K., and Anderson, G.J. (2013). Transfusion suppresses erythropoiesis and increases hepcidin in adult patients with β -thalassemia major: a longitudinal study. *Blood* *122*, 124–133.

Bibliographie

- Patil, M., Pabla, N., and Dong, Z. (2013). Checkpoint kinase 1 in DNA damage response and cell cycle regulation. *Cell Mol Life Sci* 70, 4009–4021.
- Paulsen, R.D., Soni, D.V., Wollman, R., Hahn, A.T., Yee, M.-C., Guan, A., Hesley, J.A., Miller, S.C., Cromwell, E.F., Solow-Cordero, D.E., et al. (2009). A Genome-wide siRNA Screen Reveals Diverse Cellular Processes and Pathways that Mediate Genome Stability. *Mol Cell* 35, 228–239.
- Pellagatti, A., Cazzola, M., Giagounidis, A.A.N., Malcovati, L., Porta, M.G.D., Killick, S., Campbell, L.J., Wang, L., Langford, C.F., Fidler, C., 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, 337–345.
- Perkins, A.C., Sharpe, A.H., and Orkin, S.H. (1995). Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* 375, 318–322.
- Perutz, M.F. (1963). X-ray analysis of hemoglobin. *Science* 140, 863–869.
- Petermann, E., Orta, M.L., Issaeva, N., Schultz, N., and Helleday, T. (2010). Hydroxyurea-stalled replication forks become progressively inactivated and require two different RAD51-mediated pathways for restart and repair. *Mol. Cell* 37, 492–502.
- Pigeon, C., Ilyin, G., Courselaud, B., Leroyer, P., Turlin, B., Brissot, P., and Loréal, O. (2001). A new mouse liver-specific gene, encoding a protein homologous to human antimicrobial peptide hepcidin, is overexpressed during iron overload. *J. Biol. Chem.* 276, 7811–7819.
- Pimentel, H., Parra, M., Gee, S., Ghanem, D., An, X., Li, J., Mohandas, N., Pachter, L., and Conboy, J.G. (2014). A dynamic alternative splicing program regulates gene expression during terminal erythropoiesis. *Nucleic Acids Res.* 42, 4031–4042.
- Pimentel, H., Parra, M., Gee, S.L., Mohandas, N., Pachter, L., and Conboy, J.G. (2016). A dynamic intron retention program enriched in RNA processing genes regulates gene expression during terminal erythropoiesis. *Nucleic Acids Res.* 44, 838–851.
- Platzbecker, U., Germing, U., Götze, K.S., Kiewe, P., Mayer, K., Chromik, J., Radsak, M., Wolff, T., Zhang, X., Laadem, A., et al. (2017). Luspatercept for the treatment of anaemia in patients with lower-risk myelodysplastic syndromes (PACE-MDS): a multicentre, open-label phase 2 dose-finding study with long-term extension study. *The Lancet Oncology* 18, 1338–1347.
- Pondarré, C., Antiochos, B.B., Campagna, D.R., Clarke, S.L., Greer, E.L., Deck, K.M., McDonald, A., Han, A.-P., Medlock, A., Kutok, J.L., 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, 953–964.
- Potts, K.S., Sargeant, T.J., Markham, J.F., Shi, W., Biben, C., Josefsson, E.C., Whitehead, L.W., Rogers, K.L., Liakhovitskaia, A., Smyth, G.K., et al. (2014). A lineage of diploid platelet-forming cells precedes polyploid megakaryocyte formation in the mouse embryo. *Blood* 124, 2725–2729.
- Prades, E., Chambon, C., Dailey, T.A., Dailey, H.A., Brière, J., and Grandchamp, B. (1995). A new mutation of the *ALAS2* gene in a large family with X-linked sideroblastic anemia. *Hum. Genet.* 95, 424–428.
- Pugh, C.W., Tan, C.C., Jones, R.W., and Ratcliffe, P.J. (1991). Functional analysis of an oxygen-regulated transcriptional enhancer lying 3' to the mouse erythropoietin gene. *Proc Natl Acad Sci U S A* 88, 10553–10557.

Bibliographie

- Qiao, B., Sugianto, P., Fung, E., Del-Castillo-Rueda, A., Moran-Jimenez, M.-J., Ganz, T., and Nemeth, E. (2012). Hepcidin-induced endocytosis of ferroportin is dependent on ferroportin ubiquitination. *Cell Metab.* *15*, 918–924.
- Quesada, V., Conde, L., Villamor, N., Ordóñez, G.R., Jares, P., Bassaganyas, L., Ramsay, A.J., Beà, S., Pinyol, M., Martínez-Trillos, A., et al. (2012). Exome sequencing identifies recurrent mutations of the splicing factor *SF3B1* gene in chronic lymphocytic leukemia. *Nature Genetics* *44*, 47–52.
- Ramey, G., Deschemin, J.-C., and Vaulont, S. (2009). Cross-talk between the mitogen activated protein kinase and bone morphogenetic protein/hemojuvelin pathways is required for the induction of hepcidin by holotransferrin in primary mouse hepatocytes. *Haematologica* *94*, 765–772.
- Reijns, M.A.M., Rabe, B., Rigby, R.E., Mill, P., Astell, K.R., Lettice, L.A., Boyle, S., Leitch, A., Keighren, M., Kilanowski, F., et al. (2012). Enzymatic Removal of Ribonucleotides from DNA Is Essential for Mammalian Genome Integrity and Development. *Cell* *149*, 1008–1022.
- Rejeski, K., Liang, Y., Tebaldi, T., Stefani, G., Taylor, A., Maziarz, J., Song, Y., Balasubramanian, K., Vasic, R., Kapetanovic, E., et al. (2015). Integrative Genome-Wide Analysis of RNA Binding and Splicing Reveals Complex Loss and Gain of Function Alterations By SRSF2 P95 Mutations in Myelodysplasia. *Blood* *126*, 141–141.
- Remy, I., Wilson, I.A., and Michnick, S.W. (1999). Erythropoietin receptor activation by a ligand-induced conformation change. *Science* *283*, 990–993.
- Ribeil, J.-A., Zermati, Y., Vandekerckhove, J., Cathelin, S., Kersual, J., Dussiot, M., Coulon, S., Moura, I.C., Zeuner, A., Kirkegaard-Sørensen, T., et al. (2007). Hsp70 regulates erythropoiesis by preventing caspase-3-mediated cleavage of GATA-1. *Nature* *445*, 102–105.
- Richardson, D.R., and Ponka, P. (1997). The molecular mechanisms of the metabolism and transport of iron in normal and neoplastic cells. *Biochim. Biophys. Acta* *1331*, 1–40.
- Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K.E., Guyot, B., Heck, A.J.R., Vyas, P., de Boer, E., Grosveld, F., and Strouboulis, J. (2005). GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J.* *24*, 2354–2366.
- Roepstorff, P., and Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom.* *11*, 601.
- Rouault, T.A., Hentze, M.W., Caughman, S.W., Harford, J.B., and Klausner, R.D. (1988). Binding of a cytosolic protein to the iron-responsive element of human ferritin messenger RNA. *Science* *241*, 1207–1210.
- Roy, C.N., and Enns, C.A. (2000). Iron homeostasis: new tales from the crypt. *Blood* *96*, 4020–4027.
- Rübe, C.E., Fricke, A., Widmann, T.A., Fürst, T., Madry, H., Pfreundschuh, M., and Rübe, C. (2011). Accumulation of DNA damage in hematopoietic stem and progenitor cells during human aging. *PLoS ONE* *6*, e17487.
- Saleque, S., Cameron, S., and Orkin, S.H. (2002). The zinc-finger proto-oncogene Gfi-1b is essential for development of the erythroid and megakaryocytic lineages. *Genes Dev.* *16*, 301–306.
- Santini, V., Girelli, D., Sanna, A., Martinelli, N., Duca, L., Campostrini, N., Cortelezzi, A., Corbella, M., Bosi, A., Reda, G., et al. (2011). Hepcidin Levels and Their Determinants in Different Types of Myelodysplastic Syndromes. *PLOS ONE* *6*, e23109.

Bibliographie

- Santini, V., Almeida, A., Giagounidis, A., Gröpper, S., Jonasova, A., Vey, N., Mufti, G.J., Buckstein, R., Mittelman, M., Platzbecker, U., et al. (2016). Randomized Phase III Study of Lenalidomide Versus Placebo in RBC Transfusion-Dependent Patients With Lower-Risk Non-del(5q) Myelodysplastic Syndromes and Ineligible for or Refractory to Erythropoiesis-Stimulating Agents. *J. Clin. Oncol.* *34*, 2988–2996.
- Sartori, A.A., Lukas, C., Coates, J., Mistrik, M., Fu, S., Bartek, J., Baer, R., Lukas, J., and Jackson, S.P. (2007). Human CtIP promotes DNA end resection. *Nature* *450*, 509–514.
- Savage, K.I., Gorski, J.J., Barros, E.M., Irwin, G.W., Manti, L., Powell, A.J., Pellagatti, A., Lukashchuk, N., McCance, D.J., McCluggage, W.G., et al. (2014). Identification of a BRCA1-mRNA splicing complex required for efficient DNA repair and maintenance of genomic stability. *Mol. Cell* *54*, 445–459.
- Schade, A.L., and Caroline, L. (1946). An Iron-binding Component in Human Blood Plasma. *Science* *104*, 340–341.
- Schanz, J., Tüchler, H., Solé, F., Mallo, M., Luño, E., Cervera, J., Granada, I., Hildebrandt, B., Slovak, M.L., Ohyashiki, K., 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*, 820–829.
- Schellenberg, M.J., Edwards, R.A., Ritchie, D.B., Kent, O.A., Golas, M.M., Stark, H., Lührmann, R., Glover, J.N.M., and MacMillan, A.M. (2006). Crystal structure of a core spliceosomal protein interface. *Proc Natl Acad Sci U S A* *103*, 1266–1271.
- Schmitz-Abe, K., Ciesielski, S.J., Schmidt, P.J., Campagna, D.R., Rahimov, F., Schilke, B.A., Cuijpers, M., Rieneck, K., Lausen, B., Linenberger, M.L., et al. (2015). Congenital sideroblastic anemia due to mutations in the mitochondrial HSP70 homologue HSPA9. *Blood* *126*, 2734–2738.
- Schwab, R.A., Nieminuszczy, J., Shah, F., Langton, J., Lopez Martinez, D., Liang, C.-C., Cohn, M.A., Gibbons, R.J., Deans, A.J., and Niedzwiedz, W. (2015). The Fanconi Anemia Pathway Maintains Genome Stability by Coordinating Replication and Transcription. *Mol. Cell* *60*, 351–361.
- Seghezzi, W., Chua, K., Shanahan, F., Gozani, O., Reed, R., and Lees, E. (1998). Cyclin E Associates with Components of the Pre-mRNA Splicing Machinery in Mammalian Cells. *Molecular and Cellular Biology* *18*, 4526–4536.
- Seiler, M., Peng, S., Agrawal, A.A., Palacino, J., Teng, T., Zhu, P., Smith, P.G., Caesar-Johnson, S.J., Demchok, J.A., Felau, I., et al. (2018a). Somatic Mutational Landscape of Splicing Factor Genes and Their Functional Consequences across 33 Cancer Types. *Cell Reports* *23*, 282-296.e4.
- Seiler, M., Yoshimi, A., Darman, R., Chan, B., Keaney, G., Thomas, M., Agrawal, A.A., Caleb, B., Csibi, A., Sean, E., et al. (2018b). H3B-8800, an orally available small-molecule splicing modulator, induces lethality in spliceosome-mutant cancers. *Nat. Med.* *24*, 497–504.
- Seldin, M.M., Tan, S.Y., and Wong, G.W. (2014). Metabolic function of the CTRP family of hormones. *Rev Endocr Metab Disord* *15*, 111–123.
- Sender, R., Fuchs, S., and Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLOS Biology* *14*, e1002533.
- Shaw, G.C., Cope, J.J., Li, L., Corson, K., Hersey, C., Ackermann, G.E., Gwynn, B., Lambert, A.J., Wingert, R.A., Traver, D., et al. (2006). Mitoferrin is essential for erythroid iron assimilation. *Nature* *440*, 96–100.

Bibliographie

- Shi, Y., and Massagué, J. (2003). Mechanisms of TGF- β Signaling from Cell Membrane to the Nucleus. *Cell* *113*, 685–700.
- Shiozawa, Y., Sato-Otsubo, S., Gallì, A., Yoshida, K., Yoshizato, T., Sato, Y., Kataoka, K., Sanada, M., Shiraishi, Y., Chiba, K., et al. (2014). Comprehensive Analysis of Aberrant RNA Splicing in Myelodysplastic Syndromes. *Blood* *124*, 826–826.
- Shiozawa, Y., Malcovati, L., Gallì, A., Sato-Otsubo, A., Kataoka, K., Sato, Y., Watatani, Y., Suzuki, H., Yoshizato, T., Yoshida, K., et al. (2018). Aberrant splicing and defective mRNA production induced by somatic spliceosome mutations in myelodysplasia. *Nat Commun* *9*.
- Shivdasani, R.A., Mayer, E.L., and Orkin, S.H. (1995). Absence of blood formation in mice lacking the T-cell leukaemia oncoprotein tal-1/SCL. *Nature* *373*, 432–434.
- Silvestri, L., Pagani, A., Nai, A., De Domenico, I., Kaplan, J., and Camaschella, C. (2008a). The serine protease matriptase-2 (TMPRSS6) inhibits hepcidin activation by cleaving membrane hemojuvelin. *Cell Metab.* *8*, 502–511.
- Silvestri, L., Pagani, A., and Camaschella, C. (2008b). Furin-mediated release of soluble hemojuvelin: a new link between hypoxia and iron homeostasis. *Blood* *111*, 924–931.
- Sinha, M., Jang, Y.C., Oh, J., Khong, D., Wu, E.Y., Manohar, R., Miller, C., Regalado, S.G., Loffredo, F.S., Pancoast, J.R., et al. (2014). Restoring systemic GDF11 levels reverses age-related dysfunction in mouse skeletal muscle. *Science* *344*, 649–652.
- So, S., Davis, A.J., and Chen, D.J. (2009). Autophosphorylation at serine 1981 stabilizes ATM at DNA damage sites. *The Journal of Cell Biology* *187*, 977–990.
- Socolovsky, M., Fallon, A.E., Wang, S., Brugnara, C., and Lodish, H.F. (1999). Fetal anemia and apoptosis of red cell progenitors in Stat5a-/-5b-/- mice: a direct role for Stat5 in Bcl-X(L) induction. *Cell* *98*, 181–191.
- Sollier, J., Stork, C.T., García-Rubio, M.L., Paulsen, R.D., Aguilera, A., and Cimprich, K.A. (2014). Transcription-coupled nucleotide excision repair factors promote R-loop-induced genome instability. *Mol Cell* *56*, 777–785.
- Soni, S., Bala, S., Gwynn, B., Sahr, K.E., Peters, L.L., and Hanspal, M. (2006). Absence of erythroblast macrophage protein (Emp) leads to failure of erythroblast nuclear extrusion. *J. Biol. Chem.* *281*, 20181–20189.
- Sørensen, C.S., Hansen, L.T., Dziegielewska, J., Syljuåsen, R.G., Lundin, C., Bartek, J., and Helleday, T. (2005). The cell-cycle checkpoint kinase Chk1 is required for mammalian homologous recombination repair. *Nat. Cell Biol.* *7*, 195–201.
- Steensma, D.P., Bejar, R., Jaiswal, S., Lindsley, R.C., Sekeres, M.A., Hasserjian, R.P., and Ebert, B.L. (2015). Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood* *126*, 9–16.
- Steensma, D.P., Maris, M.B., Yang, J., Donnellan, W.B., Brunner, A.M., McMasters, M., Greenberg, P., Komrokji, R.S., Klimek, V.M., Goldberg, J.M., et al. (2017). H3B-8800-G0001-101: A first in human phase I study of a splicing modulator in patients with advanced myeloid malignancies. *Journal of Clinical Oncology*.
- Stehling, O., Wilbrecht, C., and Lill, R. (2014). Mitochondrial iron-sulfur protein biogenesis and human disease. *Biochimie* *100*, 61–77.

Bibliographie

- Stephens, P.J., Tarpey, P.S., Davies, H., Van Loo, P., Greenman, C., Wedge, D.C., Nik-Zainal, S., Martin, S., Varela, I., Bignell, G.R., et al. (2012). The landscape of cancer genes and mutational processes in breast cancer. *Nature* *486*, 400–404.
- Stucki, M., Clapperton, J.A., Mohammad, D., Yaffe, M.B., Smerdon, S.J., and Jackson, S.P. (2005). MDC1 Directly Binds Phosphorylated Histone H2AX to Regulate Cellular Responses to DNA Double-Strand Breaks. *Cell* *123*, 1213–1226.
- Suragani, R.N.V.S., Cadena, S.M., Cawley, S.M., Sako, D., Mitchell, D., Li, R., Davies, M.V., Alexander, M.J., Devine, M., Loveday, K.S., et al. (2014). Transforming growth factor- β superfamily ligand trap ACE-536 corrects anemia by promoting late-stage erythropoiesis. *Nature Medicine* *20*, 408–414.
- Suzuki, M., Kobayashi-Osaki, M., Tsutsumi, S., Pan, X., Ohmori, S., Takai, J., Moriguchi, T., Ohneda, O., Ohneda, K., Shimizu, R., et al. (2013). GATA factor switching from GATA2 to GATA1 contributes to erythroid differentiation. *Genes Cells* *18*, 921–933.
- Šviković, S., Crisp, A., Tan-Wong, S.M., Guillian, T.A., Doherty, A.J., Proudfoot, N.J., Guilbaud, G., and Sale, J.E. (2018). R-loop formation during S phase is restricted by PrimPol-mediated repriming. *The EMBO Journal* e99793.
- Swart, L. de, Reiniers, C., Bagguley, T., Marrewijk, C. van, Bowen, D., Hellström-Lindberg, E., Tatic, A., Symeonidis, A., Huls, G., Cermak, J., et al. (2018). Labile plasma iron levels predict survival in patients with lower-risk myelodysplastic syndromes. *Haematologica* *103*, 69–79.
- Syed, R.S., Reid, S.W., Li, C., Cheetham, J.C., Aoki, K.H., Liu, B., Zhan, H., Osslund, T.D., Chirino, A.J., Zhang, J., et al. (1998). Efficiency of signalling through cytokine receptors depends critically on receptor orientation. *Nature* *395*, 511–516.
- Taketani, S., Kakimoto, K., Ueta, H., Masaki, R., and Furukawa, T. (2003). Involvement of ABC7 in the biosynthesis of heme in erythroid cells: interaction of ABC7 with ferrochelatase. *Blood* *101*, 3274–3280.
- Tamary, H., Shalev, H., Perez-Avraham, G., Zoldan, M., Levi, I., Swinkels, D.W., Tanno, T., and Miller, J.L. (2008). Elevated growth differentiation factor 15 expression in patients with congenital dyserythropoietic anemia type I. *Blood* *112*, 5241–5244.
- Tanno, T., Bhanu, N.V., Oneal, P.A., Goh, S.-H., Staker, P., Lee, Y.T., Moroney, J.W., Reed, C.H., Luban, N.L., Wang, R.-H., et al. (2007). High levels of GDF15 in thalassemia suppress expression of the iron regulatory protein hepcidin. *Nature Medicine* *13*, 1096–1101.
- Tanno, T., Porayette, P., Sripichai, O., Noh, S.-J., Byrnes, C., Bhupatiraju, A., Lee, Y.T., Goodnough, J.B., Harandi, O., Ganz, T., et al. (2009). Identification of TWSG1 as a second novel erythroid regulator of hepcidin expression in murine and human cells. *Blood* *114*, 181–186.
- Te Raa, G.D., Derks, I. a. M., Navrkalova, V., Skowronska, A., Moerland, P.D., van Laar, J., Oldreive, C., Monsuur, H., Trbusek, M., Malcikova, J., et al. (2015). The impact of SF3B1 mutations in CLL on the DNA-damage response. *Leukemia* *29*, 1133–1142.
- Tehranchi, R., Fadeel, B., Forsblom, A.-M., Christensson, B., Samuelsson, J., Zhivotovsky, B., and 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.

Bibliographie

- Tehranchi, R., Invernizzi, R., Grandien, A., Zhivotovsky, B., Fadeel, B., Forsblom, A.-M., Travaglino, E., Samuelsson, J., Hast, R., Nilsson, L., et al. (2005). Aberrant mitochondrial iron distribution and maturation arrest characterize early erythroid precursors in low-risk myelodysplastic syndromes. *Blood* *106*, 247–253.
- Teng, T., Tsai, J.H., Puyang, X., Seiler, M., Peng, S., Prajapati, S., Aird, D., Buonamici, S., Caleb, B., Chan, B., et al. (2017). Splicing modulators act at the branch point adenosine binding pocket defined by the PHF5A–SF3b complex. *Nature Communications* *8*, 15522.
- Testa, U., Pelosi, E., and Peschle, C. (1993). The transferrin receptor. *Crit Rev Oncog* *4*, 241–276.
- Thépot, S., Ben Abdelali, R., Chevret, S., Renneville, A., Beyne-Rauzy, O., Prébet, T., Park, S., Stamatoullas, A., Guerci-Bresler, A., Cheze, S., et al. (2016). A randomized phase II trial of azacitidine +/- epoetin- β in lower-risk myelodysplastic syndromes resistant to erythropoietic stimulating agents. *Haematologica* *101*, 918–925.
- Thol, F., Kade, S., Schlarmann, C., Löffeld, P., Morgan, M., Krauter, J., Wlodarski, M.W., Kölking, B., Wichmann, M., Görlich, K., et al. (2012). Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood* *119*, 3578–3584.
- Toma, A., Kosmider, O., Chevret, S., Delaunay, J., Stamatoullas, A., Rose, C., Beyne-Rauzy, O., Banos, A., Guerci-Bresler, A., Wickenhauser, S., et al. (2016). Lenalidomide with or without erythropoietin in transfusion-dependent erythropoiesis-stimulating agent-refractory lower-risk MDS without 5q deletion. *Leukemia* *30*, 897–905.
- Ünsal-Kaçmaz, K., Chastain, P.D., Qu, P.-P., Mino, P., Cordeiro-Stone, M., Sancar, A., and Kaufmann, W.K. (2007). The Human Tim/Tipin Complex Coordinates an Intra-S Checkpoint Response to UV That Slows Replication Fork Displacement. *Molecular and Cellular Biology* *27*, 3131–3142.
- Uto, K., Inoue, D., Shimuta, K., Nakajo, N., and Sagata, N. (2004). Chk1, but not Chk2, inhibits Cdc25 phosphatases by a novel common mechanism. *EMBO J* *23*, 3386–3396.
- Van Handel, B., Prashad, S.L., Hassanzadeh-Kiabi, N., Huang, A., Magnusson, M., Atanassova, B., Chen, A., Hamalainen, E.I., and Mikkola, H.K.A. (2010). The first trimester human placenta is a site for terminal maturation of primitive erythroid cells. *Blood* *116*, 3321–3330.
- Vandekerckhove, J., Ribeil, J.-A., Zermati, Y., Garrido, C., Courtois, G., Solary, E., and Hermine, O. (2008). Hsp70, l'ange gardien de GATA-1 lors de la différenciation des globules rouges. *Med Sci (Paris)* *24*, 37–40.
- Vogelstein, B., Papadopoulos, N., Velculescu, V.E., Zhou, S., Diaz, L.A., and Kinzler, K.W. (2013). Cancer genome landscapes. *Science* *339*, 1546–1558.
- Vokurka, M., Krijt, J., Sulc, K., and Necas, E. (2006). Hcpidin mRNA levels in mouse liver respond to inhibition of erythropoiesis. *Physiol Res* *55*, 667–674.
- Vujanovic, M., Krietsch, J., Raso, M.C., Terraneo, N., Zellweger, R., Schmid, J.A., Tagliatela, A., Huang, J.-W., Holland, C.L., Zwicky, K., et al. (2017). Replication Fork Slowing and Reversal upon DNA Damage Require PCNA Polyubiquitination and ZRANB3 DNA Translocase Activity. *Molecular Cell* *67*, 882–890.e5.
- Vulpe, C.D., Kuo, Y.M., Murphy, T.L., Cowley, L., Askwith, C., Libina, N., Gitschier, J., and Anderson, G.J. (1999). Hephaestin, a ceruloplasmin homologue implicated in intestinal iron transport, is defective in the sla mouse. *Nat. Genet.* *21*, 195–199.

Bibliographie

- Wahl, M.C., Will, C.L., and Lührmann, R. (2009). The Spliceosome: Design Principles of a Dynamic RNP Machine. *Cell* 136, 701–718.
- Walter, D., Lier, A., Geiselhart, A., Thalheimer, F.B., Huntscha, S., Sobotta, M.C., Moehrle, B., Brocks, D., Bayindir, I., Kaschutnig, P., et al. (2015). Exit from dormancy provokes DNA-damage-induced attrition in haematopoietic stem cells. *Nature* 520, 549–552.
- Wang, C., Chua, K., Seghezzi, W., Lees, E., Gozani, O., and Reed, R. (1998a). Phosphorylation of spliceosomal protein SAP 155 coupled with splicing catalysis. *Genes Dev* 12, 1409–1414.
- Wang, C., Sashida, G., Saraya, A., Ishiga, R., Koide, S., Oshima, M., Isono, K., Koseki, H., and Iwama, A. (2014). Depletion of Sf3b1 impairs proliferative capacity of hematopoietic stem cells but is not sufficient to induce myelodysplasia. *Blood* 123, 3336–3343.
- Wang, E.T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S.F., Schroth, G.P., and Burge, C.B. (2008). Alternative isoform regulation in human tissue transcriptomes. *Nature* 456, 470–476.
- Wang, G.L., Jiang, B.H., Rue, E.A., and Semenza, G.L. (1995). Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O₂ tension. *Proc Natl Acad Sci U S A* 92, 5510–5514.
- Wang, J., Takagaki, Y., and Manley, J.L. (1996). Targeted disruption of an essential vertebrate gene: ASF/SF2 is required for cell viability. *Genes Dev.* 10, 2588–2599.
- Wang, J., Xiao, S.H., and Manley, J.L. (1998b). Genetic analysis of the SR protein ASF/SF2: interchangeability of RS domains and negative control of splicing. *Genes Dev.* 12, 2222–2233.
- Wang, L., Brooks, A.N., Fan, J., Wan, Y., Gambe, R., Li, S., Hergert, S., Yin, S., Freeman, S.S., Levin, J.Z., et al. (2016). Transcriptomic Characterization of SF3B1 Mutation Reveals Its Pleiotropic Effects in Chronic Lymphocytic Leukemia. *Cancer Cell* 30, 750–763.
- Ward, I.M., and Chen, J. (2001). Histone H2AX Is Phosphorylated in an ATR-dependent Manner in Response to Replicational Stress. *J. Biol. Chem.* 276, 47759–47762.
- Weinstein, D.A., Roy, C.N., Fleming, M.D., Loda, M.F., Wolfsdorf, J.I., and Andrews, N.C. (2002). Inappropriate expression of hepcidin is associated with iron refractory anemia: implications for the anemia of chronic disease. *Blood* 100, 3776–3781.
- Weiss, M.J., Keller, G., and Orkin, S.H. (1994). Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes Dev.* 8, 1184–1197.
- Weiss, M.J., Yu, C., and Orkin, S.H. (1997). Erythroid-cell-specific properties of transcription factor GATA-1 revealed by phenotypic rescue of a gene-targeted cell line. *Mol. Cell. Biol.* 17, 1642–1651.
- Welch, J.J., Watts, J.A., Vakoc, C.R., Yao, Y., Wang, H., Hardison, R.C., Blobel, G.A., Chodosh, L.A., and Weiss, M.J. (2004). Global regulation of erythroid gene expression by transcription factor GATA-1. *Blood* 104, 3136–3147.
- Wetterberg, I., Baurén, G., and Wieslander, L. (1996). The intranuclear site of excision of each intron in Balbiani ring 3 pre-mRNA is influenced by the time remaining to transcription termination and different excision efficiencies for the various introns. *RNA* 2, 641–651.
- William Roy, S., and Gilbert, W. (2006). The evolution of spliceosomal introns: patterns, puzzles and progress. *Nature Reviews Genetics* 7, 211–221.

Bibliographie

- Wrighting, D.M., and Andrews, N.C. (2006). Interleukin-6 induces hepcidin expression through STAT3. *Blood* 108, 3204–3209.
- Wu, H., Liu, X., Jaenisch, R., and Lodish, H.F. (1995). Generation of committed erythroid BFU-E and CFU-E progenitors does not require erythropoietin or the erythropoietin receptor. *Cell* 83, 59–67.
- Yamamoto, R., Morita, Y., Ooehara, J., Hamanaka, S., Onodera, M., Rudolph, K.L., Ema, H., and Nakauchi, H. (2013). Clonal analysis unveils self-renewing lineage-restricted progenitors generated directly from hematopoietic stem cells. *Cell* 154, 1112–1126.
- Yan, C., Wan, R., Bai, R., Huang, G., and Shi, Y. (2016). Structure of a yeast activated spliceosome at 3.5 Å resolution. *Science* 353, 904–911.
- Yang, X.H., Shiotani, B., Classon, M., and Zou, L. (2008). Chk1 and Claspin potentiate PCNA ubiquitination. *Genes Dev.* 22, 1147–1152.
- Yoder, M.C. (2014). Inducing definitive hematopoiesis in a dish. *Nature Biotechnology* 32, 539–541.
- Yoshida, H., Kawane, K., Koike, M., Mori, Y., Uchiyama, Y., and Nagata, S. (2005). Phosphatidylserine-dependent engulfment by macrophages of nuclei from erythroid precursor cells. *Nature* 437, 754–758.
- Yoshida, K., Sanada, M., Shiraishi, Y., Nowak, D., Nagata, Y., Yamamoto, R., Sato, Y., Sato-Otsubo, A., Kon, A., Nagasaki, M., et al. (2011). Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 478, 64–69.
- Yoshizawa-Sugata, N., and Masai, H. (2007). Human Tim/Timeless-interacting protein, Tipin, is required for efficient progression of S phase and DNA replication checkpoint. *J. Biol. Chem.* 282, 2729–2740.
- Yu, J.S.L., and Cui, W. (2016). Proliferation, survival and metabolism: the role of PI3K/AKT/mTOR signalling in pluripotency and cell fate determination. *Development* 143, 3050–3060.
- Yu, K., Chedin, F., Hsieh, C.-L., Wilson, T.E., and Lieber, M.R. (2003). R-loops at immunoglobulin class switch regions in the chromosomes of stimulated B cells. *Nat. Immunol.* 4, 442–451.
- Zamai, L., Secchiero, P., Pierpaoli, S., Bassini, A., Papa, S., Alnemri, E.S., Guidotti, L., Vitale, M., and Zauli, G. (2000). TNF-related apoptosis-inducing ligand (TRAIL) as a negative regulator of normal human erythropoiesis. *Blood* 95, 3716–3724.
- Zellweger, R., Dalcher, D., Mutreja, K., Berti, M., Schmid, J.A., Herrador, R., Vindigni, A., and Lopes, M. (2015). Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. *J. Cell Biol.* 208, 563–579.
- Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. *Nature Cell Biology* 16, 2–9.
- Zermati, Y., Fichelson, S., Valensi, F., Freyssinier, J.M., Rouyer-Fessard, P., Cramer, E., Guichard, J., Varet, B., and Hermine, O. (2000). Transforming growth factor inhibits erythropoiesis by blocking proliferation and accelerating differentiation of erythroid progenitors. *Exp. Hematol.* 28, 885–894.
- Zermati, Y., Garrido, C., Amsellem, S., Fishelson, S., Bouscary, D., Valensi, F., Varet, B., Solary, E., and Hermine, O. (2001). Caspase activation is required for terminal erythroid differentiation. *J. Exp. Med.* 193, 247–254.

Bibliographie

- Zhang, X., Yan, C., Zhan, X., Li, L., Lei, J., and Shi, Y. (2018). Structure of the human activated spliceosome in three conformational states. *Cell Research* 28, 307–322.
- Zhao, B., Mei, Y., Schipma, M.J., Roth, E.W., Bleher, R., Rappoport, J.Z., Wickrema, A., Yang, J., and Ji, P. (2016). Nuclear Condensation during Mouse Erythropoiesis Requires Caspase-3-Mediated Nuclear Opening. *Dev. Cell* 36, 498–510.
- Zhou, L., Nguyen, A.N., Sohal, D., Ying Ma, J., Pahanish, P., Gundabolu, K., Hayman, J., Chubak, A., Mo, Y., Bhagat, T.D., et al. (2008). Inhibition of the TGF- receptor I kinase promotes hematopoiesis in MDS. *Blood* 112, 3434–3443.
- Zhou, Q., Derti, A., Ruddy, D., Rakiec, D., Kao, I., Lira, M., Gibaja, V., Chan, H., Yang, Y., Min, J., et al. (2015). A Chemical Genetics Approach for the Functional Assessment of Novel Cancer Genes. *Cancer Res* 75, 1949–1958.
- Zhu, Y., Li, X., Chang, C., Xu, F., He, Q., Guo, J., Tao, Y., Liu, Y., Liu, L., and Shi, W. (2016). SF3B1-mutated myelodysplastic syndrome with ring sideroblasts harbors more severe iron overload and corresponding over-erythropoiesis. *Leukemia Research* 44, 8–16.
- Zou, L., and Elledge, S.J. (2003). Sensing DNA damage through ATRIP recognition of RPA-ssDNA complexes. *Science* 300, 1542–1548.
- Zou, L., Liu, D., and Elledge, S.J. (2003). Replication protein A-mediated recruitment and activation of Rad17 complexes. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13827–13832.

BONDU Sabrina
Université Paris Descartes, année 2019

Conséquences fonctionnelles des mutations du gène *SF3B1* dans les syndromes myélodysplasiques avec sidéroblastes en couronne

Résumé de la thèse : Les syndromes myélodysplasiques (SMD) avec sidéroblastes en couronne (SMD-SC) sont des pathologies clonales de la cellule souche hématopoïétique caractérisées par une érythropoïèse inefficace et une surcharge mitochondriale et systémique en fer. Le gène *SF3B1* est muté chez 90% des patients porteurs d'un SMD-SC. Il code pour un facteur d'épissage appartenant au complexe U2 du spliceosome. L'interaction de SF3B1 avec l'intron au niveau du site de branchement permet le recrutement d'U2. SF3B1 s'associe aussi avec la chromatine et sa dynamique de distribution dans le noyau est sous le contrôle de diverses protéines de réparation de l'ADN, telles FANCD2, FANCI, et BRCA1. Les conséquences fonctionnelles des mutations du gène *SF3B1* (*SF3B1*^{MUT}) dans les SMD-SC sont mal connues. Lors de ma thèse, je me suis intéressée aux anomalies d'épissage générées en aval des mutations de *SF3B1* dans les cellules des patients porteurs d'un SMD. En accord avec les précédentes analyses de séquençage de l'ARN dans les pathologies *SF3B1*^{MUT}, les mutations de *SF3B1* sont associées à une utilisation préférentielle de sites accepteurs alternatifs dans les cellules mononucléées (CMN) de SMD. Nous avons identifié la surexpression d'un transcrite alternatif de l'érythroferrone (ERFE) dans les CMN et érythroblastes en culture *SF3B1*^{MUT}. ERFE est une hormone sécrétée par les érythroblastes qui régule l'expression hépatique de l'hepcidine, cette dernière limitant la biodisponibilité du fer dans le plasma. Le transcrite alternatif d'ERFE (ERFE⁺¹²) n'est retrouvé que dans les cellules érythroïdes *SF3B1*^{MUT} et son expression reflète l'érythropoïèse clonale induite par les mutations de *SF3B1*. ERFE⁺¹² est traduit en un variant protéique fonctionnel. L'augmentation des concentrations des protéines ERFE canonique et alternative dans le plasma des patients SMD *SF3B1*^{MUT} est corrélée à la diminution des taux d'hepcidine, et explique le développement précoce d'une surcharge systémique en fer. J'ai ensuite caractérisé les processus de réplication et de réparation de l'ADN dans les érythroblastes primaires humains *SF3B1*^{MUT} et dans la lignée érythroïde murine G1E-ER4 CRISPR-Cas9 *Sf3b1*^{K700E}. Nos expériences démontrent que les mutations de *SF3B1* confèrent un avantage de prolifération aux précurseurs érythroïdes immatures en augmentant leur dynamique de réplication de l'ADN, *via* l'accélération de la progression des fourches. L'augmentation de la vitesse des fourches de réplication s'accompagne d'une exposition d'ADN simple-brin recouvert par la protéine p-RPA32, mais ne conduit pas à la formation de cassures double-brin de l'ADN. Les mutations de *SF3B1* contribuent donc à accroître de façon inadéquate le compartiment des cellules érythroïdes immatures dans la moelle osseuse des patients SMD-SC. Nos prochains travaux viseront à déterminer par quels mécanismes les fourches de réplication sont accélérées, notamment avec l'exploration des intermédiaires de transcription *R-loops*.

MOTS CLÉS : SMD-SC, SF3B1, épissage, érythroferrone, réplication et réparation de l'ADN

Directrice de thèse : Pr Michaela Fontenay

Rapporteurs : Pr Marie-Bérengère Troadec, Dr Sarah Lambert

Examineurs : Dr Sophie Vaultont, Dr Marc-Henri Stern, Dr Jean-Baptiste Micol