



**UNIVERSITÉ
DE LORRAINE**

**BIBLIOTHÈQUES
UNIVERSITAIRES**

AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

Ecole Doctorale BioSE (Biologie-Santé-Environnement)

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITE DE LORRAINE

Mention : « Sciences de la Vie et de la Santé »

par **Marion DIVOUX**

APPROCHE GÉNOMIQUE DES NEOPLASIES MYÉLODYSPLASIQUES/MYÉLOPROLIFÉRATIVES AVEC MUTATION DE *SF3B1* ET THROMBOCYTOSE TRIPLE NEGATIVES

25/11/2024

Membres du jury :

Rapporteurs :	Madame Laurence LEGROS	Maître de conférences des universités, Unité Inserm U1197, Villejuif
	Monsieur Éric SOLARY	Professeur des universités, Institut Gustave Roussy, Villejuif
Examineurs :	Madame Maud D'AVENI	Professeur des universités, Faculté de médecine, Nancy
	Monsieur Éric LIPPERT	Professeur des universités, Université de Brest
	Monsieur Julien BROSEUS	Professeur des universités, Unité Inserm U1256, Faculté de médecine, Nancy, co-directeur de thèse
	Monsieur Pierre FEUGIER	Professeur des universités, Unité Inserm U1256, Faculté de médecine, Nancy, directeur de thèse

200919208U : Unité Inserm U1256 - NGERE, Université de Lorraine - Faculté de
Médecine - 9 avenue de la Forêt de Haye - CS 50184 54500 VANDOEUVRE LES NANCY
CEDEX

TABLE DES MATIÈRES

TABLE DES MATIÈRES	3
PUBLICATIONS AU COURS DU DOCTORAT	7
PUBLICATIONS RÉFÉRENCÉES PUBMED.....	7
PUBLICATION SOUMISE EN COURS DE REVIEWING	7
PUBLICATIONS PRÉ-SOUMISSION	8
PUBLICATIONS EN COURS DE RÉDACTION	8
ABRÉVIATIONS	9
LISTE DES TABLEAUX, FIGURES ET ANNEXES	13
LISTE DES TABLEAUX.....	13
LISTE DES FIGURES.....	15
LISTE DES ANNEXES.....	19
RÉSUMÉS	21
RÉSUMÉ DE LA THÈSE EN FRANÇAIS	21
RÉSUMÉ DE LA THÈSE EN ANGLAIS.....	23
RÉSUMÉ VULGARISÉ POUR LE GRAND PUBLIC EN FRANÇAIS	25
RÉSUMÉ VULGARISÉ POUR LE GRAND PUBLIC EN ANGLAIS	25
INTRODUCTION	27
LES SYNDROMES MYÉLODYSPLASIQUES	27
1) Définitions.....	27
2) Classification OMS 2022 des MDS	30
1) Pronostic.....	31
2) Traitements recommandés des MDS.....	33
MDS ET ÉPIGÉNÉTIQUE	37
1) Expression génique et épigénétique - définitions.....	37
2) Méthylation de l'ADN	39
3) Modifications post-traductionnelles des histones.....	43
MDS ET ALTÉRATIONS DE L'ÉPISSAGE	49
1) L'épissage	49
2) Mutations récurrentes impactant l'épissage dans les MDS.....	52
3) Les MDS- <i>SF3B1</i>	55

LES NÉOPLASIES MYÉLOPROLIFÉRATIVES.....	57
1) Les types principaux de MPN	57
2) La voie JAK-STAT	58
3) La thrombopoïèse	70
4) Thrombocytémie Essentielle	73
5) Méthylome des TE	75
LES MDS/MPN- <i>SF3B1</i> -T.....	78
1) Découverte de l'entité frontière MDS/MPN- <i>SF3B1</i> -T et évolution de la nomenclature :	78
2) <i>Signature moléculaire des MDS/MPN-<i>SF3B1</i>-T</i>	81
3) Pronostic	83
4) Traitement.....	85
WHOLE EXOME SEQUENCING (WES).....	87
ANALYSES BIOINFORMATIQUES DU NGS	88
1) Données brutes et contrôles qualité	88
2) Alignement	88
3) Détermination des variants	88
4) Annotation des variants :	89
SÉQUENCAGE DU MÉTHYLOME PAR RRBS	89
OBJECTIF DE L'ÉTUDE	91
PARTIE 1 : séquençage d'exome et validations fonctionnelles.....	91
PARTIE 2 : méthylome et expression.	91
MATÉRIEL ET MÉTHODES	93
PARTIE 1 – PREMIÈRE COHORTE - PANEL IGR.....	94
PARTIE 1 - DEUXIÈME COHORTE - EXOME de 6 patients	95
1) Collection biologique de MDS/MPN- <i>SF3B1</i> -T, fractions CD3+ et CD3-.....	95
2) <i>Whole exome sequencing</i> (WES).....	96
3) Analyses bioinformatiques	98
PARTIE 1 - TROISIÈME COHORTE – COHORTE DE CONFIRMATION	101
PARTIE 1 - ANALYSES FONCTIONNELLES	101
1) Variant c.2117 A>T de <i>STAT5B</i>	101
2) <i>PTPN12</i>	104
PARTIE 2 - QUATRIÈME COHORTE – COHORTE MYST.....	107
1) Collection biologique et données cliniques	107
2) Extraction et préparation des librairies.....	109
3) Caractérisation moléculaire	110
4) Disponibilités des données de séquençage et type d'analyses.....	110

5) Méthodologie Méthylkit.....	111
6) Méthodologie « <i>in-house</i> ».....	113
7) Analyse ciblée.....	118
8) Gestion des réplicats.....	118
9) qPCR des candidats issus du RRBS.....	118
RÉSULTATS	121
PARTIE 1 – PREMIÈRE COHORTE D'ÉCHANTILLONS : NGS IGR.....	121
PARTIE 1 - DEUXIÈME ET TROISIÈME COHORTES : COHORTE DE SCREENING ET COHORTE DE VALIDATION.....	121
1) Analyses primaires WES de la deuxième cohorte (cohorte de screening).....	121
2) Analyses secondaires WES de la deuxième cohorte.....	122
3) Analyses tertiaires WES de la deuxième cohorte (cohorte de screening).....	124
4) Troisième cohorte (cohorte de validation).....	128
PARTIE 1 – ANALYSES FONCTIONNELLES	131
1) La mutation c.2117 A>T entraîne un saut d'exon de <i>STAT5B</i>	131
2) Surexpression de <i>PTPN12</i> dans un modèle cellulaire HEK293.....	132
PARTIE 2 - QUATRIÈME COHORTE – COHORTE MYST	133
1) Collection biologique et données cliniques.....	133
2) Caractérisation moléculaire.....	134
3) Contrôles qualité	134
4) Prise en compte des réplicats	134
5) Méthylation globale des 48 échantillons.....	135
6) Analyse de la méthylation des 48 échantillons selon la méthode <i>MethylKit</i>	135
7) Analyse de la méthylation selon la méthode <i>in-house</i>	137
8) qPCR des candidats.....	144
9) Analyses ciblées de la méthylation	146
DISCUSSION	147
PARTIE 1 - Analyse d'exome et validations fonctionnelles	147
PARTIE 2 - Approche épigénétique par l'étude du méthylome	155
CONCLUSION.....	159
RÉFÉRENCES.....	161
ANNEXES	171
RÉSUMÉS	215

PUBLICATIONS AU COURS DU DOCTORAT

PUBLICATIONS RÉFÉRENCÉES PUBMED

1. D'Aveni-Piney M, **Divoux M**, Busby-Venner H, Muller M, Broséus J, Feugier P. Idelalisib in a patient with refractory Waldenström's macroglobulinemia complicated by anuric renal failure: a case report. *J Med Case Reports*. 12 juin 2018;12(1):164.
2. **Divoux M**, Plocque A, Sevin M, Voillat L, Feugier P, Guerci-Bresler A, et al. Efficacy of lenalidomide in myelodysplastic/myeloproliferative neoplasms with ring sideroblasts and an extreme platelet count. *Clin Case Rep*. sept 2020;8(9):1774-80.
3. Cazaubiel T, Leleu X, Perrot A, Manier S, Buisson L, ..., **Divoux M**, et al. Primary plasma cell leukemias displaying t(11;14) have specific genomic, transcriptional, and clinical features. *Blood*. 28 avr 2022;139(17):2666-72.
4. Hergalant S, Saurel C, **Divoux M**, Rech F, Pouget C, Godfraind C, et al. Correlation between DNA Methylation and Cell Proliferation Identifies New Candidate Predictive Markers in Meningioma. *Cancers*. 17 déc 2022;14(24):6227.
5. Lannes R, Samur M, Perrot A, Mazzotti C, **Divoux M**, Cazaubiel T, et al. In Multiple Myeloma, High-Risk Secondary Genetic Events Observed at Relapse Are Present From Diagnosis in Tiny, Undetectable Subclonal Populations. *J Clin Oncol Off J Am Soc Clin Oncol*. 20 mars 2023;41(9):1695-702.
6. Notarantonio AB, Roth-Guépin G, Bonmati C, **Divoux M**, Kicki C, Pagliuca S, et al. A new sequential conditioning regimen based on CPX- 351/Vyxeos (« Vyx-Seq ») in patients with higher risk myelodysplastic syndromes. *Leuk Res*. déc 2023;135:107405.
7. Schavgoulidze A, Talbot A, Perrot A, Cazaubiel T, ..., **Divoux M**, et al. Biallelic deletion of 1p32 defines ultra-high-risk myeloma, but monoallelic del(1p32) remains a strong prognostic factor. *Blood*. 16 mars 2023;141(11):1308-15.
8. Schavgoulidze A, Lauwers-Cances V, Perrot A, Cazaubiel T, Chretien ML... **Divoux M** et al. Heterogeneity in long-term outcomes for patients with Revised International Staging System stage II, newly diagnosed multiple myeloma. *Haematologica*. 1 mai 2023;108(5):1374-84.
9. Thomas H, Alix T, Renard É, Renaud M, Wourms J, ..., **Divoux M**, et al. Expanding the genetic and clinical spectrum of Tatton-Brown-Rahman syndrome in a series of 24 French patients. *J Med Genet*. 29 août 2024;61(9):878-85.
10. Aubin L, Vilas Boas R, Daltro De Oliveira R, Le Brun V, **Divoux M**, Rey J, et al. CALR-mutated patients with low allele burden represent a specific subtype of essential thrombocythemia: A study on behalf of FIM and GBMHM. *Am J Hematol*. mai 2024;99(5):1001-4.

PUBLICATION SOUMISE EN COURS DE REVIEWING

1. **Divoux M**, Resche-Rigon M, Michonneau D, Sutra Del Galy A, Dhedin N, Xhaard A, Sicre de Fontbrune F, Robin M, Socié G, Peffault de Latour R. Anti-human T lymphocyte globulin Fresenius in graft-versus-host disease prophylaxis for unrelated hematopoietic stem cell transplantation after myeloablative conditioning: a long-term real-life retrospective study. *Am J Hematol*. per reviewing.

PUBLICATIONS PRÉ-SOUMISSION

1. **Divoux M et al.** DNA methylation comparative analysis in triple negative or JAK2V617F MDS/MPN with SF3B1 mutation and thrombocytosis, JAK2V617F essential thrombocytemia and MDS with SF3B1 mutation (**cf ANNEXE 5**).

PUBLICATIONS EN COURS DE RÉDACTION

1. **Divoux M** et al. Mutational landscape of Triple Negative Myelodysplastic/Myeloproliferative Neoplasms with *SF3B1* mutation and Thrombocytosis (MDS/MPN-*SF3B1*-T). Analyses fonctionnelles enc ours.
2. **Divoux M** et al. Methylation analysis in high-risk myelodysplastic syndromes with excess of blasts (SMD-EB). Analyses en attente.
3. **Divoux M** et al. Treatment of overlap Hemolytic Paroxysmal Nocturnal Hemoglobinuria and severe aplastic anemia: a review. En cours de rédaction. Cible : Medicina (IF : 2,9)

ABRÉVIATIONS

5^mC : 5-methyl-cytosine	EXAC : EXome aggregation consortium
5q- : Délétion du bras long du chromosome 5	EZH2 Enhancer of zeste 2 polycomb repressive complex 2 subunit
AATK : Apoptosis associated tyrosine Kinase	FAB : Franco-américano-britannique
ABCB7 : ATP binding cassettesSubfamily B member 7	FACS : Fluorescence activated cell sorting
ARNm : Acide ribonucléique messenger	FDR : False discovery rate
ADN : Acide désoxyribonucléique	FL : FLT3 ligand
AGK : Acylglycerol kinase	FLT3 : Fms related receptor tyrosine kinase 3
ASE : Agents stimulants de l'érythropoïèse	FLI1 : Friend leukemia virus integration 1
ASXL1 : Additional sex combs-like 1	FOG-1 : Friend of GATA-1
BAP1 : BRCA1-associated protein 1	FT : Facteurs de transcription
BCLAP : Bladder cancer associated protein	GATA-1 : GATA-binding protein 1
BCOR : BCL6 corepressor	GFP : Green fluorescent protein
BER : Base excision repair	GM-CSF : Granulocyte macrophage colony-stimulating factor
BOM : Biopsie ostéo-médullaire	H2AK11 : Lysine 11 de l'histone H2A
CALR : Calreticulin	H3K27 : Lysine 27 de l'histone H3
CBL : Casitas B-lineage lymphoma proto-oncogene	H3K9me : Lysine 9 de l'histone H3 méthylée
CCL5 : Chemokine (C-C motif) ligand 5	HAT : Histone acteyl transférases
CHIP : Hématopoïèse de potentiel indéterminé	HDAC : Histone desacetylases
CHU : Centre hospitalo-universitaire	HEK293 : Human embryonic kidney 293 cells
CHRU : Centre hospitalier régional universitaire	HeLa : Henrietta Lacks cells
CNV : Copy number variation	HMT : Histone methyl transferase
CRB : Centre de ressources biologiques	hnRNP : Heterogeneous nuclear ribonucleoprotein particle
CSH : cellule souche hématopoïétique	ICC : International consensus classification
c-src : Cellular-src	IDH1 : Isocitrate dehydrogenase (NADP(+)) 1
DBD : DNA binding domain	IGF-1 : Insulin like growth factor
DEG : Régions différenciellement exprimées	IGR : Institut Gustave Roussy
DMEM : Dulbecco modified eagle medium	IL-11 : Interleukine-11
DMR : Differentiated methylated region	IL-12 : Interleukine-12
DNMT : DNA methyl transférases	IL-13 : Interleukine-13
DNMT3a : DNA methyltransferase 3a	IL-1α : Interleukine-1 α
DUSP2 : Dual specificity phosphatase 2	IL-3 : Interleukin-3
ELN : European leukemia net	IL5 : Interleukin-5
EPO : Erythropoïétine	IL-6 : Interleukine-6
ESE : Exonic splicing enhancer	IL-9 : Interleukine-9
ESS : Exonic splicing silencer	IPSS : International prognosis score system
ET-RS : Thrombocytémie essentielle avec sidéroblastes en couronne	IPSS-M : IPSS moléculaire
ETV6 : ETS variant transcription factor 6	ISE : Intronic splicing enhancer

ISS : Intronic splicing silencer	MKP : Mégacaryocytes précurseurs
ITGA2B : Integrin subunit alpha 2b	MKL1 : Megakaryoblastic leukemia 1
ITGB3 : Integrin subunit beta 3	MN : Mononucléées
ISS : Site d'initiation de la transcription	MPL : Myeloproliferative Leukemia
IWG : International working group	MPN : Néoplasie myéloproliférative
JAK : Just another kinase	MUTYH : MutY DNAGlycosylase
JAK2 : Janus tyrosine kinase 2	MYD88 : Myeloid differentiation primary response 88
JH : JAK homology	NaCl : Chlorure de sodium
laboratoire MLL : Münchner leukemia labor	NF-E2 : Nuclear factor erythroid 2
LAM : Leucémie aiguë myéloïde	NGS : Séquençage haut débit
LMC : Leucémie myéloïde chronique	NMD : Nonsense-mediated decay
LMMC : Leucémie myélomonocytaire chronique	NRAS : Neuroblastoma RAS viral oncogene homolog
LGL : Large granular lymphocytes	OMS : Organisation mondiale de la santé
Leucémie LGL : Large granular lymphocytic leukemia	PCR : Polymerase chain reaction
LHX6 : LIM homeobox 6	PHD : Plant homeodomain
LMX1B : LIM Homeobox Transcription Factor 1 Beta	Phi : Chromosome « Philadelphie »
LNK : Lymphocyte specific adaptor protein	PI3K : Phosphatidylinositol 3-kinase
K : Lysine	PIAS : Protein inhibitor of activated STAT
MAF : Minor Allele Frequency	PN : Polynucléées
MAPK : Mitogen-activated protein kinase	PNN : Polynucléaires neutrophiles
MDS : Myéodysplasie	PRC : Polycomb repressor complexe
MDS/MPN-SF3B1-T : Néoplasie myéodysplasique, myéloproliférative avec mutation de SF3B1 et thrombocytose	PRC1 : Polycomb repressor complex 1
MPN-SF3B1-T-JAK2^{V617F} : MDS/MPN-SF3B1-T avec mutation JAK2 ^{V617F}	PRC2 : Polycomb repressor complex 2
MDS/MPN-SF3B1-T TN : MDS/MPN-SF3B1-T triple négatifs	PTK : Protéines kinases
MDS-5q : Myéodysplasie 5q-	PTP : Protein tyrosine phosphatase
MDS-biTP53 : Myéodysplasie avec atteinte multiple de TP53	PTPN11 : protein tyrosine phosphatase non receptor type 11
MDS-f : Myéodysplasie avec fibrose	PTPN12 : Protein tyrosine phosphatase non-receptor Type 12
MDS-h : Myéodysplasie hypoplasique	PTPN6 : Protein tyrosine phosphatase non-receptor type 6
MDS-IB : Myéodysplasie avec augmentation des blastes	PV : Polyglobulie de Vaquez
MDS-LB : MDS sans augmentation des blastes	qPCR : Quantitative PCR
MDS-SF3B1 : Myéodysplasie, avec mutation de SF3B1	R : Arginine
MEP : Progéniteurs multipotents engagés dans les lignées érythrocytaire et plaquettaire	RARS : Refractory anaemia with ring sideroblasts
MF : Myélofibrose	RARS-MT : RARS with marked thrombocytosis
MIGB : Metabolism, integrative genomics & bioinformatics	RARS-nMT : RARS not marked thrombocytosis
miRNA : ARN endogènes interférents	RARS-T : RARS with thrombocytosis
MK : Mégacaryocyte	R-IPSS : Revised international prognosis scoring system
	RNA : ARN
	RNA-seq : Séquençage complet de l'ARN
	RPMI : Roswell park memorial institute
	RPPH1 : Ribonuclease P RNA component H1
	RRBS : Reduced representation bisulfite sequencing

RUNX1 : <i>Runt-related transcription factor 1</i>	TAL-1 : <i>T-cell acute lymphocytic leukemia 1</i>
RSV : <i>Rous sarcoma virus</i>	TBP : <i>TATA Box Binding Protein</i>
S : <i>Sérine</i>	TE : <i>Thrombocytémie essentielle</i>
SAM : <i>S-adenosine-méthionine</i>	TET2 : <i>Ten eleven translocation 2</i>
SC : <i>Sidéroblastes en couronne</i>	THP-1 : <i>Tohoku hospital pediatrics-1 cells</i>
SCF : <i>Stem cell factor</i>	TN : <i>Triple négatives</i>
SF3B1 : <i>Splicing factor 3b subunit 1</i>	TPO : <i>Thrombopoïétine</i>
SH : <i>Src homology domain</i>	TLR : <i>Toll like receptor</i>
SHP-1 : <i>SH2 containing protein tyrosine phosphatase 1</i>	TMEM14C : <i>Transmembrane protein 14C</i>
SMAD : <i>Mothers against decapentaplegic homolog</i>	TP53 : <i>tumor protein p53</i>
SnRNA : <i>Petits ARN nucléaires (Small nuclear RNA)</i>	U2AF1 : <i>U2 small nuclear auxiliary factor 1</i>
SnRNPs : <i>Small nuclear ribonucleoproteins</i>	UMS : <i>Unité mixte de service</i>
SNV : <i>Single nucleotide variant</i>	UTR : <i>Untranslated region</i>
SOCS : <i>Suppressor of cytokine signalling</i>	VEP : <i>Variant effect predictor</i>
SOCS3 : <i>Suppressor of cytokine signaling</i>	V-Src : <i>Tyrosine-protein kinase transforming protein Src</i>
SPRY2 : <i>Sprouty RTK signaling Antagonist 2</i>	WES : <i>Whole exome sequencing</i>
SR : <i>Serin-arginin repeat</i>	WGBS : <i>Whole genome bisulfite sequencing</i>
SRSF2 : <i>Serine and arginine rich splicing factor 2</i>	Wt : <i>wild type</i>
STATs : <i>Signal transducer and activator of transcription</i>	Y : <i>Tyrosine</i>
SVF : <i>Sérum de veau fœtal</i>	YRS : <i>Tyrosyl-tRNA synthetase</i>
TAD : <i>Domaine transactivateur</i>	ZRSF2 : <i>Zinc finger CCCH-type RNA binding motif and serine/arginine rich 2</i>
	ZRSR2 : <i>Zinc finger CCCH-Type, RNA binding motif and serine/arginine rich 2</i>

LISTE DES TABLEAUX, FIGURES ET ANNEXES

LISTE DES TABLEAUX

Tableau 1. Classification 2022 des MDS selon l'OMS.....	31
Tableau 2. <i>International Prognosis Scoring System</i>	32
Tableau 3. <i>Revised International Prognosis Scoring System</i>	33
Tableau 4. Score IPSET révisé pour la TE.....	75
Tableau 5. Critères diagnostiques des MDS/MPN- <i>SF3BI-T</i> selon l'OMS 2022	80
Tableau 6. Score de stratification de risque pour les MDS/MPN- <i>SF3BI-T</i>	84
Tableau 7. Séquence des oligonucléotides utilisés pour la production de plasmide pXJ41- <i>STAT5b^{exon17WT}</i> et pXJ41- <i>STAT5b^{exon17MUT}</i>	102
Tableau 8. Protocole de PCR pour la mutagenèse dirigée de <i>STAT5b</i> dans le plasmide pXJ41- <i>STAT5b^{exon17WT}</i>	103
Tableau 9. Séquence des oligonucléotides utilisés pour le clonage de <i>PTPN12</i> dans le plasmide peGFP-C2.....	105
Tableau 10. Présentation morphologique et mutationnelle des 5 groupes faisant l'objet d'une analyse comparative du méthylome.	107
Tableau 11. Plan de l'utilisation des échantillons dans la cohorte MYST.....	108
Tableau 12. Score de pertinence appliqué aux DMRs issues de la comparaison MDS/MPN- <i>SF3BI-T-JAK2^{V617F}</i> vs MDS/MPN- <i>SF3BI-T-TN</i>	116
Tableau 13. Couples d'amorces utilisés pour la qPCR des candidats de la cohorte MYST. .	119
Tableau 14. Données cliniques des 48 échantillons (39 patients) de la cohorte MYST.	133
Tableau 15. Candidats retenus après analyse finale <i>in-house</i> des données du RRBS des MDS/MPN- <i>SF3BI-T</i>	143
Tableau 16. Mesure du niveau d'expression par qPCR des gènes candidats issus du RRBS.	144

LISTE DES FIGURES

Figure 1. Prévalence des mutations somatiques dans les CSH en fonction de l'âge.....	27
Figure 2. Mutations récurrentes des MDS classées selon leur fonction biologique.....	29
Figure 3. Les différentes fonctions cellulaires impactées dans les MDS.....	29
Figure 4. Expansion clonale dans les MDS.....	30
Figure 5. Courbes de survie et d'évolution en LAM secondaire selon le score IPSS.....	32
Figure 6. Régulation de la transcription par les <i>enhancers</i>	38
Figure 7. Méthylation des cytosines de l'ADN.....	39
Figure 8. Méthylation et déméthylation des cytosines.	40
Figure 9. Principales mutations de <i>DNMT3a</i> dans les MDS.....	41
Figure 10. Compaction de l'ADN et régulation de l'expression.....	43
Figure 11. Principales modifications possibles des histones du nucléosome.....	45
Figure 12. Etat de la chromatine en fonction des modifications des histones.....	45
Figure 13. Effet des mutations d' <i>ASXL1</i> sur les PRCs, H2A et H3.	47
Figure 14. Réaction d'épissage d'un intron.....	49
Figure 15. Eléments nucléotidiques du pré-ARN messager nécessaires à l'épissage.....	50
Figure 16. Etapes de l'épissage catalysé par le spliceosome.....	51
Figure 17. Positionnement de SF3B1 au sein du spliceosome.....	51
Figure 18. Régulation positive et négative de l'épissage.	52
Figure 19. Mutations du spliceosome dans les MDS.....	52
Figure 20. Modèle d'altération de l'épissage par les mutations de <i>SF3B1</i>	53
Figure 21. Synthèse de transcrits aberrants induits par les mutations de <i>SF3B1</i>	54
Figure 22. Sidéroblaste en couronne en coloration de Perls, en microscopie optique (x100).....	55
Figure 23. Distribution des mutations de <i>JAK2</i> , <i>CALR</i> et <i>MPL</i> dans la PV, la TE et la MF.	58
Figure 24. Mécanisme moléculaire de la fonction du domaine SH2 dans la protéine src.....	59
Figure 25. Voie de signalisation JAK-STAT.	60
Figure 26. Structure des protéines JAKs et leurs sites de phosphorylation conservés.	61
Figure 27. Liaison sélective des différentes protéines JAK à leurs récepteurs membranaires.....	62
Figure 28. Structure et sites de phosphorylation des protéines de la famille STAT.	63
Figure 29. Exemples d'interactions de STAT3 avec les acteurs de la régulation épigénétique.....	64
Figure 30. Régulation négative de la voie JAK-STAT.	65
Figure 31. Interactions de la voie JAK-STAT avec les voies RAS/MAPK et PI3K/Akt.....	67
Figure 32. Interconnexions de la voie JAK-STAT et de divers mécanismes cellulaires.	67
Figure 33. Mutations de <i>STAT5b</i> dans les hémopathies malignes.	69
Figure 34. Schéma de la thrombopoïèse.	70
Figure 35. Rôle de l'apoptose dans la biologie des mégacaryocytes et des plaquettes.....	73
Figure 36. Incidence cumulée de transformation blastique pour les MPN non Phi.	74
Figure 37. Différences significatives d'expression entre les volontaires sains et le groupe des TE.	76

Figure 38. Classification hiérarchique des différents sous-types moléculaires de TE et des échantillons contrôles selon les profils d'expression génique.	77
Figure 39. Répartition des principales mutations caractéristiques des MPN non-Phi dans les MDS/MPN- <i>SF3B1</i> -T.	81
Figure 40. Survies globales et survie sans progression en LAM des RARS-T, MDS avec SC et TE... 83	
Figure 41. Survie globale selon la stratification de risque de Patnaik et al.	84
Figure 42. Représentation du logiciel <i>IGV</i> pour l'exon 14 du gène <i>SF3B1</i>	87
Figure 43. Exemple de fichier « .vcf ».	88
Figure 44. Cohortes d'échantillons utilisés pour l'approche génomique de la partie 1.....	94
Figure 45. <i>Workflow</i> du kit <i>Agilent SureSelect^{QXT} target enrichment</i>	97
Figure 46. Génération des clusters par réplication en pont selon le système <i>cBot</i>	97
Figure 47. Étapes clés de l'analyse d'exome.....	100
Figure 48. Représentation de la construction pXJ41- <i>STAT5b^{exon17}</i>	102
Figure 49. Comparaison des groupes par les CpGs selon <i>MethylKit</i>	111
Figure 50. Création de DMR selon la méthode <i>Methylkit</i>	112
Figure 51. Représentation de la méthodologie de comparaison <i>in-house</i> pour les CpGs.....	113
Figure 52. Représentation du calcul du ratio de méthylation d'une DMR pour chaque échantillon et selon la méthodologie <i>in-house</i>	114
Figure 53. Justification du score de pertinence des DMRs issues de la comparaison MDS/MPN- <i>SF3B1</i> -T- <i>JAK2</i> vs MDS/MPN- <i>SF3B1</i> -T-TN.	116
Figure 54. Représentation schématique de l'application du score de pertinence à 5/5.	117
Figure 55. Séquençage ciblé de 15 échantillons de patients avec MDS/MPN- <i>SF3B1</i> -T TN.....	121
Figure 56. Contrôles qualité du séquençage de l'exome.	122
Figure 57. Profondeur moyenne du séquençage par chromosome et par kilobase.....	123
Figure 58. Nombre d'exons par catégorie de couverture et par échantillon.	123
Figure 59. Types de variants identifiés dans l'analyse d'exome de six patients porteurs d'un MDS/MPN- <i>SF3B1</i> -T triple négatif.	124
Figure 60. Système de réparation BER selon la base de données KEGG (Kyoto Encyclopedia of Genes and Genomes).	125
Figure 61. Représentation des 49 gènes candidats retenus après l'application du pipeline bioinformatique.	126
Figure 62. Exome ciblé de la deuxième cohorte (cohorte de screening, n=6 patients, panel similaire à celui utilisé pour la première cohorte).	127
Figure 63. qPCR de <i>PTPN12</i> pour trois échantillons de TE TN (TE7, TE8 et TE9).	128
Figure 64. Séquençage ciblé de la cohorte de validation (troisième cohorte, n=42 patients, panel de 46 gènes candidats).....	129
Figure 65. Analyse prédictive <i>in silico</i> du variant c.2117 A>Tselon <i>Ex-skip</i>	131
Figure 66. Minigène <i>STAT5B</i>	131

Figure 67. Expression de ERK, Phospho-ERK, AKT, Phospho-AKT selon la surexpression de <i>PTPN12</i> dans le modèle HEK293.....	132
Figure 68. a. Nombre de <i>reads</i> issus du RRBS selon leur longueur. b. Nombre de <i>reads</i> issus du RRBS selon le score <i>Phred</i> de leurs bases.	134
Figure 69. Méthylation moyenne de l'ADN par échantillon et par groupe de pathologie (analyses globales).	135
Figure 70. Manhattan plot des CpGs significativement différentiellement méthylés entre les MDS/MPN- <i>SF3BI</i> -T TN et les MDS/MPN- <i>SF3BI</i> -T- <i>JAK2</i> ^{V617F} sur l'analyse selon <i>MethylKit</i>	136
Figure 71. Manhattan plot des DMRs significativement différentiellement méthylées entre les MDS/MPN- <i>SF3BI</i> -T TN et les MDS/MPN- <i>SF3BI</i> -T- <i>JAK2</i> ^{V617F} selon <i>MethylKit</i>	136
Figure 72. Analyse non supervisée par PCA des CpGs significatifs du RRBS avec un seuil de différence de méthylation à 5%.....	137
Figure 73. Analyse non supervisée par PCA des 15 000 CpGs les plus variants après analyse du RRBS.	138
Figure 74. Heatmap des 15 000 CpGs les plus variants après analyse du RRBS.	138
Figure 75. Manhattan plot des CpGs significativement différentiellement méthylées entre les MDS/MPN- <i>SF3BI</i> -T TN et les MDS/MPN- <i>SF3BI</i> -T- <i>JAK2</i> ^{V617F} sur l'analyse selon la méthode <i>in-house</i>	139
Figure 76. Manhattan plot des DMRs significativement différentiellement méthylées entre les MDS/MPN- <i>SF3BI</i> -T TN et les MDS/MPN- <i>SF3BI</i> -T- <i>JAK2</i> ^{V617F} sur l'analyse selon la méthode <i>in-house</i>	140
Figure 77. Analyse non supervisée par PCA des 959 CpGs différentiels significatifs pour un seuil de différence de méthylation de 5% et avec score de pertinence à 5/5.	141
Figure 78. Analyse non supervisée par PCA des 113 CpGs avec score de pertinence à 5/5 parmi les 15 000 CpGs les plus variants.	141
Figure 79. Heatmap des 113 CpGs avec score de pertinence à 5/5 parmi les 15 000 CpGs les plus variants.	142
Figure 80. PCR quantitative des gènes candidats issus du RRBS.....	145
Figure 81. Hypothèse de l'impact du variant c.2117 A>T de <i>STAT5B</i> dans les MDS/MPN- <i>SF3BI</i> -T TN.	151

LISTE DES ANNEXES

Annexe 1. Epigenomic landscape selon Waddington	171
Annexe 2. Liste des 53 variants dans 46 gènes candidats identifiés dans la cohorte de screening et constituent le panel de la cohorte de validation (GRCh37).....	172
Annexe 3. Séquençage ciblé de la cohorte de validation de 42 MDS/MPN- <i>SF3B1</i> -T	174
Annexe 4. Variants candidats de la cohorte de validation retenus pour annotations fonctionnelles... 179	
Annexe 5. Manuscrit pré-soumission. Divoux et al. DNA methylation comparative analysis in triple negative or <i>JAK2</i> ^{V617F} MDS/MPN with <i>SF3B1</i> mutation and thrombocytosis, <i>JAK2</i> ^{V617F} essential thrombocytemia and MDS with <i>SF3B1</i> mutation.	180
Annexe 6. Publication – Divoux et al. Efficacy of lenalidomide in MDS/MPN-SF3B1-T and an extreme platelet count.	180
Annexe 7. Publication en cours de reviewing <i>American Journal of Haematology</i> . Divoux et al. Anti-human T lymphocyte globulin Fresenius in graft-versus-host disease prophylaxis for unrelated hematopoietic stem cell transplantation after myeloablative conditioning: a long-term real-life retrospective study	208
Annexe 8. Publication – Hergalant, Saurel, Divoux et al. Correlation between DNA Methylation and Cell Proliferation Identifies New Candidate Predictive Markers in Meningioma.	209
Annexe 9. Publication. Aubin, Vilas Boas, Daltro De Oliveira, Le Brun, Divoux et al. <i>CALR</i> -mutated patients with low allele burden represent a specific subtype of essential thrombocytemia: A study on behalf of FIM and GBMHM.	210

RÉSUMÉS

RÉSUMÉ DE LA THÈSE EN FRANÇAIS

Les néoplasies myélodysplasiques, myéloprolifératives avec mutation de *SF3B1* et thrombocytose sont des pathologies clonales de la cellule souche hématopoïétique regroupant des caractéristiques communes aux néoplasies myélodysplasiques (MDS) et aux néoplasies myéloprolifératives (MPN). La mutation *JAK2^{V617F}* est retrouvée dans seulement 45 à 50% des cas. Les mutations de *CALR* ou *MPL* ne sont retrouvées que dans moins de 5% des cas. La prolifération plaquettaire est donc inexplicée pour les 50% restant, conduisant à des difficultés diagnostiques pour ces MDS/MPN-*SF3B1*-T triple négatifs (MDS/MPN-*SF3B1*-T TN). Dans le but de comprendre le mécanisme moléculaire expliquant la thrombocytose, nous avons réalisé un séquençage complet d'exome et exploré la méthylation des MDS/MPN-*SF3B1*-T TN.

Pour l'analyse d'exome, la cohorte de screening était constituée de 12 échantillons de 6 patients atteints de MDS/MPN-*SF3B1*-T TN : un échantillon cellulaire CD3+ (lymphocytes T, considérés comme tissu sain) et un échantillon CD3- (enrichi en granulocytes, considéré tumoral). L'analyse comparative a permis d'identifier 49 gènes candidats. Il a également mis en évidence un biais de transversion pouvant refléter une atteinte de la réparation BER et la duplication d'une région incluant le gène *PTPN12*. Le NGS ciblé des gènes candidats sur une cohorte de confirmation a identifié 5 gènes particulièrement pertinents : *BRAT1*, *LMTK3*, *LTK*, *POU4F1*, *SGK223* et *STAT5b*. L'un des variants de *STAT5B* induit un épissage alternatif avec saut d'un exon aboutissant à une potentielle protéine tronquée non phosphorylable et pouvant être impliqué dans la prolifération hématopoïétique.

Nous avons également développé ici une nouvelle approche d'analyse du méthylome à partir de données obtenues par RRBS comparatif sur une cohorte composée d'échantillons issus de 7 MDS/MPN-*SF3B1*-T *JAK2^{V617F}*, 9 MDS/MPN-*SF3B1*-T TN, 8 MDS-*SF3B1*, 10 thrombocytémies essentielles *JAK2^{V617F}* (TE *JAK2^{V617F}*) et 4 moelles normales (MO).

L'analyse du RRBS a été réalisée avec une nouvelle approche intégrant la position de chaque CpG dans les DMRs (*Differentially methylated regions*). Cette méthode a permis d'identifier 131 DMRs significativement hypométhylées et 103 DMRs significativement hyperméthylées dans les MDS/MPN-*SF3B1*-T TN par rapport aux MDS/MPN-*SF3B1*-T *JAK2*^{V617F}. Nous avons établi un score clinique de pertinence pour sélectionner les DMRs en lien avec la thrombocytose non reliée à *JAK2*^{V617F}, identifiant ainsi 15 DMRs annotées à 11 gènes.

Neuf gènes candidats étaient reliés à la voie MAPK : 7 gènes hypométhylés dans les MDS/MPN-*SF3B1*-T TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} (*AATK*, *BLCAP*, *DUSP2*, *FLT4*, *GPM6A*, *LHX6* et *SPRY2*) et 2 gènes hyperméthylés dans les MDS/MPN-*SF3B1*-T TN (*LMX1B* et *IMPA2*). Les transcrits de 8 de ces gènes ont été testés par qPCR dans des échantillons issus du RRBS. La qPCR a confirmé la surexpression de *AATK*, *BLCAP*, *DUSP2*, *FLT4*, *LMX1B* et *SPRY2* dans les MDS/MPN-*SF3B1*-T TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et par rapport aux MO. *IMPA2* était sous-exprimé dans les MDS/MPN-*SF3B1*-T TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et par rapport aux MO. *GPM6A* était surexprimé dans les MDS/MPN-*SF3B1*-T TN et les MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} par rapport aux MO.

En conclusion, nous avons identifié des candidats génomiques pertinents tels que *STAT5B* dont les analyses fonctionnelles doivent se poursuivre. Nous avons construit une méthode innovante d'analyse des données de RRBS pour l'identification et la sélection de DMRs potentiellement reliées à la thrombocytose indépendante de *JAK2*^{V617F} dans les MDS/MPN-*SF3B1*-T-TN. Nous avons ainsi mis en évidence l'implication de la voie MAPK dans les MDS/MPN-*SF3B1*-T, en particulier TN. La régulation positive et négative des MAPKs dans les MDS/MPN-*SF3B1*-T ouvre la voie à de nouvelles thérapies spécifiques et approches diagnostiques dans les MDS/MPN-*SF3B1*-T et plus généralement dans les MPN non mutés *JAK2*/*CALR*/*MPL*.

RÉSUMÉ DE LA THÈSE EN ANGLAIS

Myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis (MDS/MPN-*SF3B1*-T) is a clonal hematopoietic stem cell disorder with overlapping features of myelodysplastic neoplasms (MDS) and myeloproliferative neoplasms (MPN). *JAK2*^{V617F} mutations are found in only 45-50% of cases. *CALR* or *MPL* mutations are found in less than 5% of cases. Platelet's proliferation remains then unexplained in the residual 50%, leading to diagnosis difficulties for these triple negative MDS/MPN-*SF3B1*-T (TN MDS/MPN-*SF3B1*-T). To understand the molecular mechanism underlying thrombocytosis, we explored whole exome sequencing and DNA methylation in TN MDS/MPN-*SF3B1*-T.

Here, we developed a novel approach of DNA methylation profiling from the data obtained with comparative reduced representation bisulfite sequencing (RRBS) on a cohort composed of 7 *JAK2*^{V617F} MDS/MPN-*SF3B1*-T, 9 TN MDS/MPN-*SF3B1*-T, 8 myelodysplastic neoplasms with *SF3B1* mutation (MDS-*SF3B1*), 10 *JAK2*^{V617F} essential thrombocytemia (*JAK2*^{V617F} TE) and 4 normal bone marrow (BM) samples. Comparative methylation analysis was initially performed by using R software package *MethylKit* and thereafter with a new designed approach integrating CpG position in differentially methylated regions (DMRs).

This method identified 103 significant hypomethylated DMRs and 131 significant hypermethylated DMRs between *JAK2*^{V617F} MDS/MPN-*SF3B1*-T and TN MDS/MPN-*SF3B1*-T. The filtering of significant DMRs relied on a clinical pertinence score, devised for selecting DMRs associated with *JAK2*^{V617F} unrelated thrombocytosis. This approach enabled the identification of 15 significant DMRs associated with 11 genes overrepresenting the mitogen-activated protein kinases (MAPK) and related pathways.

Nine candidate genes were particularly relevant and were related to MAPK pathway: 7 genes (*AATK*, *BLCAP*, *DUSP2*, *FLT4*, *GPM6A*, *LHX6* and *SPRY2*) hypomethylated in MDS/MPN-*SF3B1*-T TN as compared to MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} and two genes (*LMX1B* and *IMPA2*) hypermethylated. The hypermethylation of *LMX1B* was reported to lead to upregulation of *LMX1B* messenger RNA in previous works. Corresponding transcripts were analyzed by quantitative polymerase chain reaction (qPCR) in RRBS samples.

Consistently, qPCR confirmed overexpression of *AATK*, *BLCAP*, *DUSP2*, *FLT4*, *LMX1B* and *SPRY2* transcripts in TN MDS/MPN-*SF3B1*-T as compared to *JAK2*^{V617F} MDS/MPN-*SF3B1*-T, MDS-*SF3B1* and normal BM. *IMPA2* transcript was underexpressed in TN MDS/MPN-*SF3B1*-T as compared to *JAK2*^{V617F} MDS/MPN-*SF3B1*-T, MDS-*SF3B1* and normal BM. *GPM6A* transcript was overexpressed in TN MDS/MPN-*SF3B1*-T and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T compared to normal BM.

In conclusion, we devised an innovative profiling method from RRBS measurements for selecting DMRs potentially associated with thrombocytosis but *JAK2*^{V617F}-independent in TN MDS/MPN-*SF3B1*-T. We highlighted the implication of MAPK pathway in MDS/MPN-*SF3B1*-T and especially in TN MDS/MPN-*SF3B1*-T. The upregulation of MAPKs in MDS/MPN-*SF3B1*-T paves the way to specific therapeutic and diagnostic approaches in MDS/MPN-*SF3B1*-T and more generally in *JAK2*/*CALR*/*MPL* unmutated MPNs.

RÉSUMÉ VULGARISÉ POUR LE GRAND PUBLIC EN FRANÇAIS

La moelle osseuse produit toutes les cellules sanguines (globules rouges, blancs et plaquettes). Au cours du vieillissement, il peut y survenir des mutations génétiques à l'origine de maladies, dont les néoplasies myélodysplasiques/myéloprolifératives avec mutation de *SF3B1* et thrombocytose (MDS/MPN-*SF3B1*-T). Dans cette maladie, il existe une production excessive de plaquettes, inexpliquée dans 50% des cas (MDS/MPN-*SF3B1*-T TN).

Le but de ces travaux est de comprendre son mécanisme sous-jacent, pour mieux diagnostiquer et traiter. Nous avons analysé une partie du génome d'échantillons de patients avec MDS/MPN-*SF3B1*-T TN, à la recherche de mutations inconnues à ce jour et avons également développé une nouvelle méthode d'analyse de la méthylation du génome.

Nos résultats montrent une atteinte du gène *STAT5B* et de la méthylation en lien avec la voie de signalisation MAPK.

Cela ouvre la voie à de nouvelles thérapeutiques spécifiques et approches diagnostiques dans les MDS/MPN-*SF3B1*-T.

RÉSUMÉ VULGARISÉ POUR LE GRAND PUBLIC EN ANGLAIS

The bone marrow produces all blood cells (red blood cells, white blood cells and platelets). During oldering, genetic mutations may occur and lead to bone marrow diseases of which myelodysplastic/myeloproliferative neoplasms with *SF3B1* mutation and thrombocytosis (MDS/MPN-*SF3B1*-T). Excessive platelets production in MDS/MPN-*SF3B1*-T is unexplained in 50% of cases (MDS/MPN-*SF3B1*-T TN).

To improve diagnosis and care, the aim of our work is to understand the underlying mechanisms of this platelet's overproduction. To identify unknown mutations, we analyzed a genome subset from MDS/MPN-*SF3B1*-T TN patients' samples. We also developed a new genome methylation analysis method.

Our results show that *STAT5B* and methylation related to signaling pathway MAPK are altered.

It paves the way to specific therapeutic and diagnostic approaches in MDS/MPN-*SF3B1*-T.

INTRODUCTION

LES SYNDROMES MYÉLODYSPLASIQUES

1) Définitions

Les syndromes myélodysplasiques ou néoplasies myélodysplasiques (MDS) sont les hémopathies malignes myéloïdes les plus fréquentes des pays industrialisés, avec une incidence de 5 à 13 cas pour 100 000 habitants et par an, et un âge médian au diagnostic de 71 à 76 ans (1). Ces pathologies de la cellule souche hématopoïétique (CSH) se manifestent par une hématopoïèse inefficace (entraînant des cytopénies périphériques), associées à des anomalies morphologiques des cellules myéloïdes (dysplasie > 10% des cellules d'une lignée) et un risque d'évolution agressive vers une leucémie aiguë myéloïde (LAM) secondaire (2). La physiopathologie est complexe, favorisée par un contexte inflammatoire sur un terrain de vieillissement de la CSH. Ces conditions aboutissent à des dommages de l'ADN (acide désoxyribonucléique) et à l'émergence d'une hématopoïèse clonale (un ou deux clones majoritaires). Au cours du vieillissement, certaines CSH peuvent en effet acquérir des mutations somatiques (**Figure 1**) et occuper une proportion non négligeable du compartiment hématopoïétique (3,4). Quand elles existent, ces mutations confèrent un avantage de prolifération par rapport aux autres CSH. Lorsque la fréquence allélique est supérieure à 2%, en l'absence de pathologie hématologique (cytopénie, dysplasie), on parle d'hématopoïèse clonale de potentiel indéterminé (CHIP) (5).

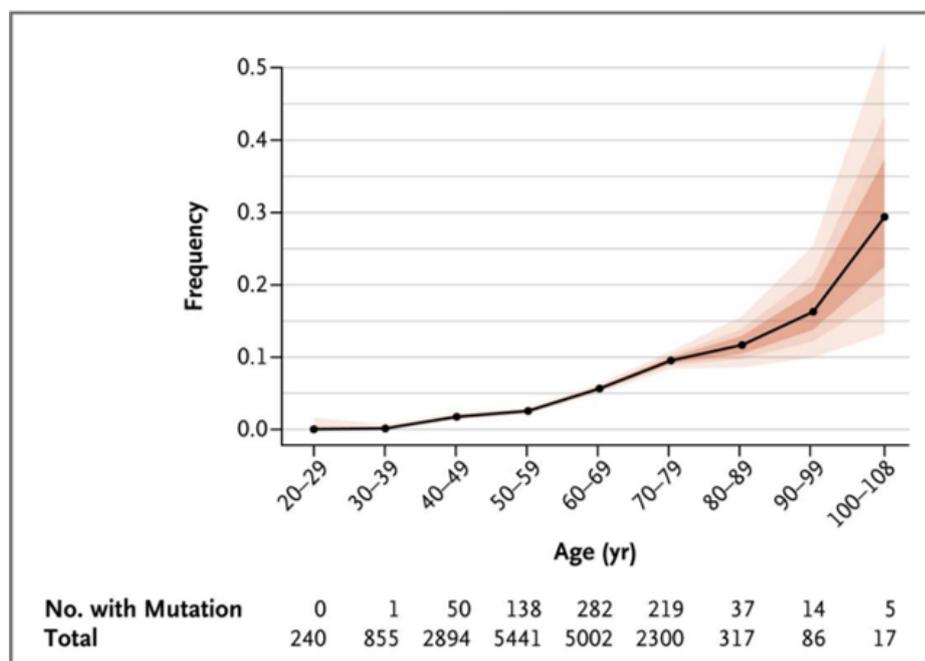


Figure 1. Prévalence des mutations somatiques dans les CSH en fonction de l'âge.

Les CHIP constituent un facteur de risque de développement d'hémopathie myéloïde (notamment les MDS), de pathologie cardio-vasculaire et de surmortalité (3,6). Plus de 10% des patients de plus de 60 ans et plus de 20% des patients de plus de 90 ans présentent ce type d'altérations somatiques (7). Elles concernent principalement des mutations de gènes impliqués dans la méthylation de l'ADN tels que *DNMT3a* (*DNA methyltransferase 3a*) et *TET2* (*ten eleven translocation 2*). Le panorama mutationnel des MDS comprend surtout des altérations de gènes impliqués dans l'épissage et la méthylation de l'ADN (**Figure 2, Figure 3**) (8–11).

Dans un contexte de dysfonction du microenvironnement médullaire et de l'épigénétique, l'expansion de clones porteurs de mutations initiatrices, puis l'occurrence d'évènements secondaires, participent à l'apparition d'une MDS et/ou d'une LAM (**Figure 4**) (1). Les mutations initiatrices de la pathologie concernent des gènes impliqués dans l'épissage et le modelage de la chromatine tels que *DNMT3a*, *TET2*, *ASXL1* (*additional sex combs-like 1*), *SF3B1* (*splicing factor 3b subunit 1*), *ZRSF2* (*zinc finger CCCH-type RNA binding motif and serine/arginine rich 2*) et *U2AF1* (*U2 small nuclear auxiliary factor 1*). Les évènements secondaires concernent le plus souvent des gènes codant pour des facteurs de transcription : *RUNX1* (*Runt-Related Transcription Factor 1*), *ETV6* (*ETS Variant Transcription Factor 6*), la voie TP53 (*tumor protein p53*), ou des acteurs de signalisation impliqués dans la différenciation (12,13).

Sur le plan cytogénétique, des réarrangements chromosomiques détectables au caryotype sont rapportés dans la moitié des cas de MDS (1). Parmi les anomalies récurrentes, les délétions isolées du bras long du chromosome 5 (5q-) et les délétions (partielles ou totales) du chromosome 7 sont particulièrement corrélées au pronostic.

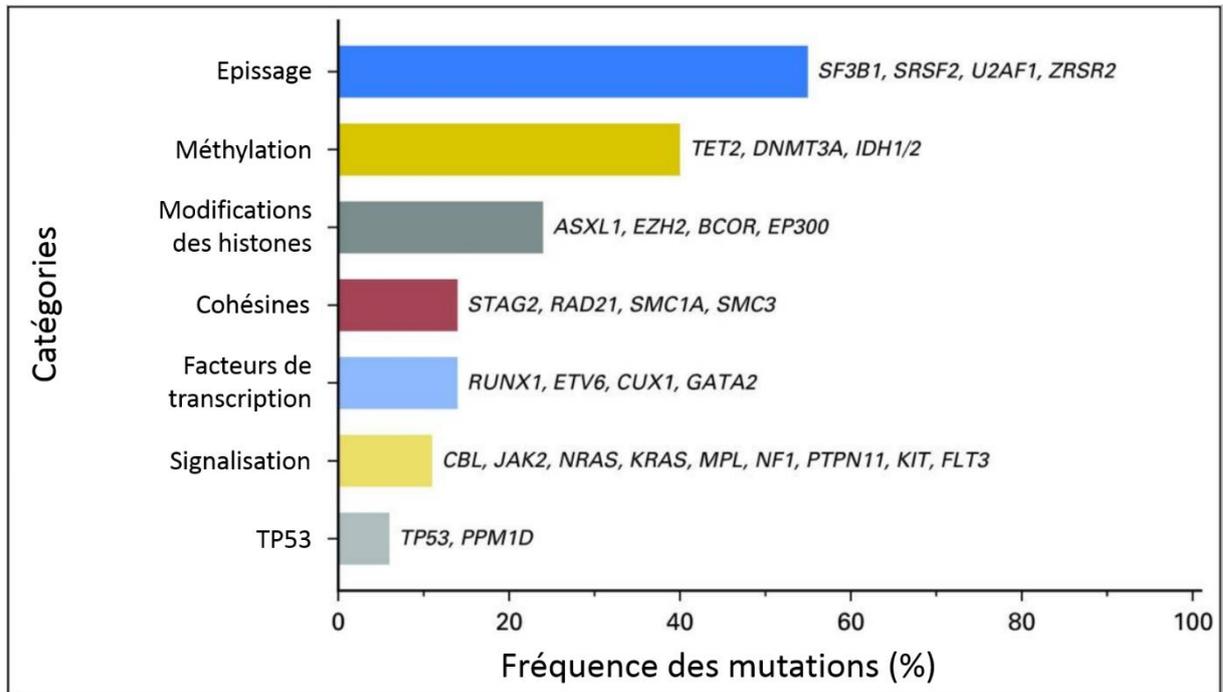


Figure 2. Mutations récurrentes des MDS classées selon leur fonction biologique.

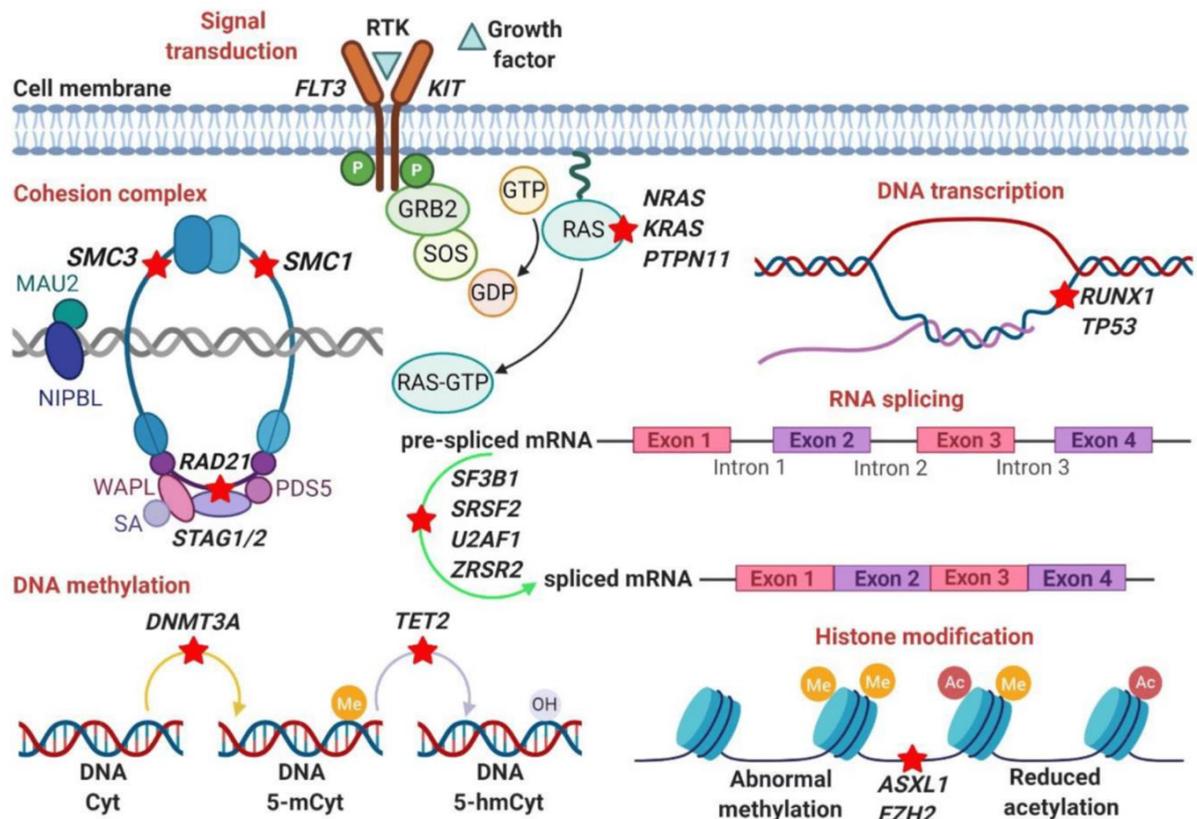


Figure 3. Les différentes fonctions cellulaires impactées dans les MDS.

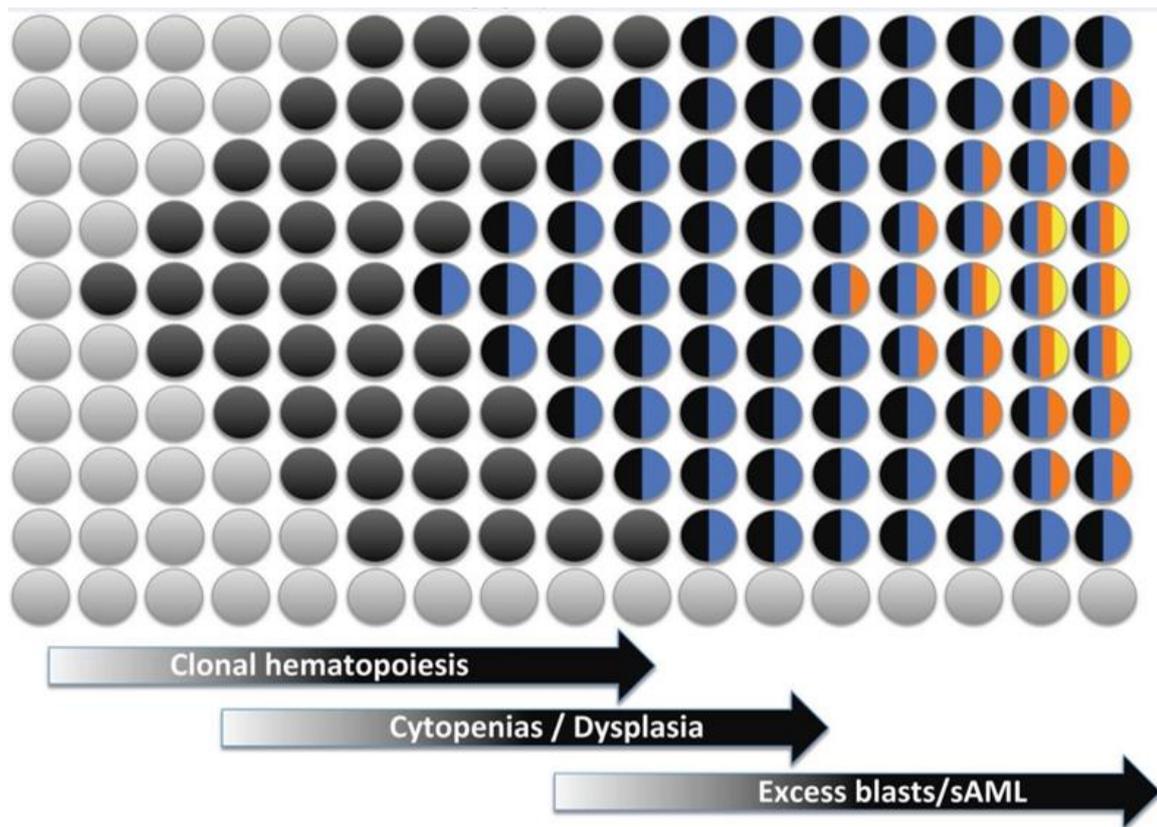


Figure 4. Expansion clonale dans les MDS.

2) Classification OMS 2022 des MDS

La 5^{ème} édition de la classification de l'organisation mondiale de la santé (OMS) 2022 des hémopathies (2) distingue deux catégories de MDS, comportant elles-mêmes plusieurs sous-catégories (**Tableau 1**). Dans cet ensemble très hétérogène, les MDS peuvent en effet être regroupées selon leur morphologie et leurs altérations génétiques. Cette classification impacte les traitements proposés (altération du pronostic en cas d'excès de blastes ou possibilités de thérapies ciblées selon l'anomalie génétique par exemple). Il existe par ailleurs une classification ICC (*International Consensus Classification*) 2022 présentant quelques différences avec la classification OMS 2022 (14).

Tableau 1. Classification 2022 des MDS selon l'OMS.

	Blastes	Cytogénétique	Mutations
MDS définis par leur anomalie cytogénétique			
<u>MDS-5q</u> : MDS sans excès de blastes et avec délétion 5q isolée	< 5% (moelle) < 2% (sang)	Délétion 5q isolée ou associée à une autre anomalie autre que la monosomie 7 ou la délétion 7q	
<u>MDS-SF3B1</u> : MDS sans excès de blastes et avec mutation de SF3B1*		Absence de délétion 5q, de monosomie 7 ou de caryotype complexe	SF3B1
<u>MDS-biTP53</u> : MDS avec mutation bi-allélique de TP53	< 20% (moelle et sang)	Habituellement complexe	Deux (ou plus) mutations de TP53 ou une mutation avec preuve de délétion de l'autre allèle ou perte d'hétérozygotie
MDS définis par leur morphologie			
<u>MDS-LB</u> : MDS faible blastes	< 5% (moelle) < 2% (sang)		
<u>MDS-h</u> : MDS hypoplasique ¹			
<u>MDS-IB</u> : MDS avec augmentation des blastes			
MDS-IB 1	5-9% (moelle) 2-4% (sang)		
MDS-IB 2	10-19 % (moelle ou sang) ou présence de corps d'Auer		
MDS-f : MDS avec fibrose	10-19 % (moelle ou sang)		

*la détection de plus de 15% de sidéroblastes en couronne peut se substituer à la présence d'une mutation de SF3B1. MDS-LB: myelodysplastic syndrome with low blast count; MDS-IB: myelodysplastic syndrome with increased blast count.

¹par définition : richesse médullaire inférieure ou égale à 25% (ajustée à l'âge).

1) Pronostic

Sur le plan pronostic, l'ensemble des MDS est encore une fois très hétérogène. Les scores IPSS (*International Prognosis Scoring System*) (**Tableau 2, Figure 5**) et R-IPSS (*Revised International Prognosis Scoring System*) (**Tableau 3**) s'appuient sur le pourcentage de blastes médullaires, le nombre de cytopénies et les altérations cytogénétiques (15,16). Ils permettent de dichotomiser les MDS en deux grandes catégories de risque (de survie et d'évolution vers la LAM) : élevé et faible. Plus récemment, un score IPSS moléculaire (IPSS-M) propose d'intégrer les données génomiques dans le calcul du pronostic (17). En effet, le profil moléculaire influence le pronostic. Par exemple, les patients présentant deux altérations de TP53 ou plus, et ceux avec un faible nombre de cooccurrences de mutations ont un risque élevé d'évolution vers la LAM indépendamment de leur groupe IPSS ou R-IPSS (18).

Tableau 2. International Prognosis Scoring System.

Variable	Nombre de points				
	0	0,5	1,0	1,5	2,0
% de blastes médullaires	<5	5-10	—	11-20	21-30
Caryotype*	Bon	Intermédiaire	Défavorable		
Cytopenies	0/1	2/3			

*Bon : normal, -Y, del(5q), del(20q); Défavorable : complexe (≥ 3 anomalies), anomalie du chromosome 7; Intermédiaire : autres anomalies.

Groupes de risque : Faible : 0 ; INT-1 : 0.5-1.0 ; INT-2 : 1.5-2.0 ; Elevé : ≥ 2.5 .

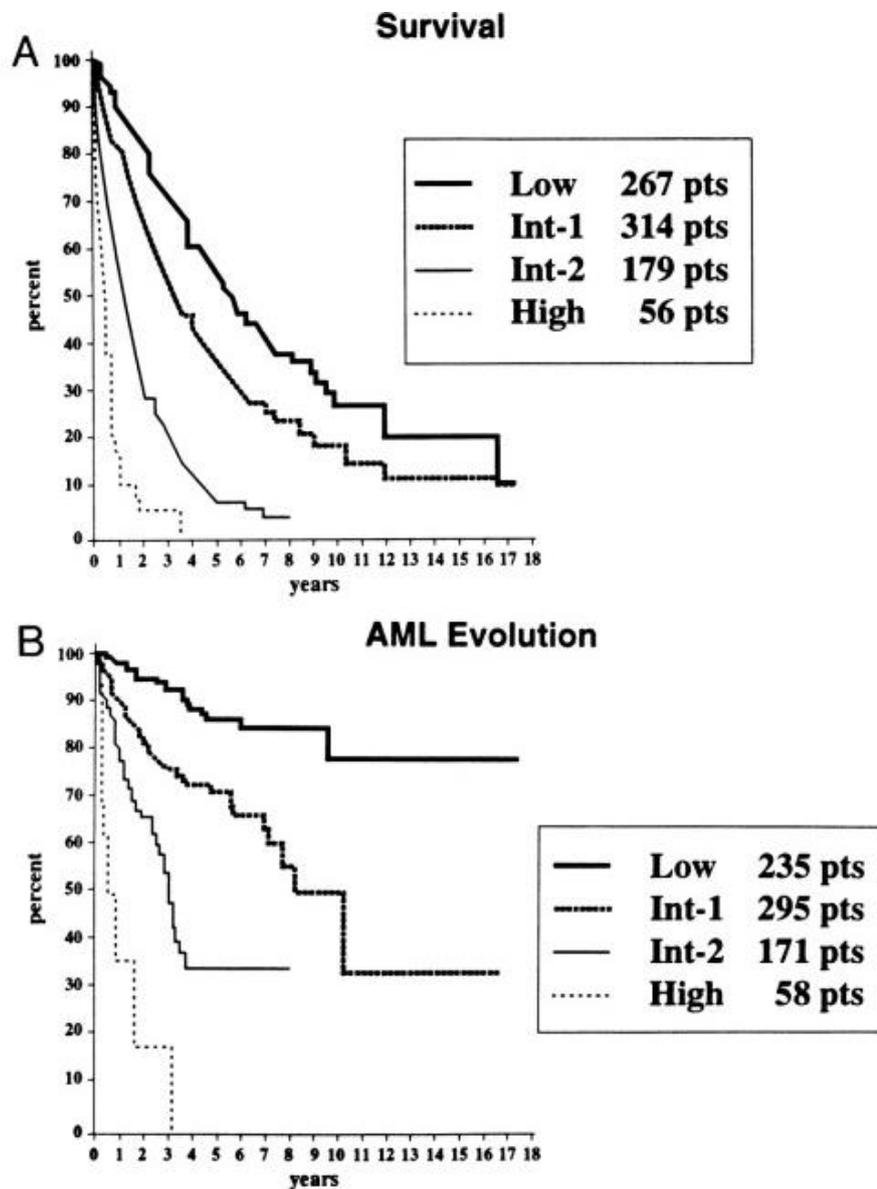


Figure 5. Courbes de survie et d'évolution en LAM secondaire selon le score IPSS.

Tableau 3. Revised International Prognosis Scoring System

Variable	Points						
	0	0,5	1	1,5	2	3	4
Cytogénétique*	Très bon	—	Bon	—	Intermédiaire	Défavorable	Très défavorable
% blastes médullaires	≤ 2	—	>2% - < 5%	—	5%-10%	> 10%	—
Hémoglobine (g/dL)	≥ 10	—	8- < 10	< 8	—	—	—
Plaquettes	≥ 100	50 - <100	< 50	—	—	—	—
Compte absolu de polynucléaires (G/L)	≥ 0,8	< 0,8	—	—	—	—	—

*Très bon : -Y, del(11q) ; Bon : normal, del(5q), del(12p), del(20q), del(5q) + une autre anomalie ; Intermédiaire : del(7q), +8, +19, i(17q), autre anomalie non concernée par un autre groupe ; Défavorable : -7, inv(3)/t(3q)/del(3q), -7/del(7q) + une autre anomalie, complexe à 3 anomalies ; très défavorable : complexe à plus de 3 anomalies.

Groupes de risque : Très faible : ≤ 1,5 ; Faible : 1.5-3 ; Intermédiaire : 3-4,5 ; Elevé : 4,5-6 ; Très élevé : > 6

2) Traitements recommandés des MDS

La prise en charge des MDS repose sur la classification OMS et l'évaluation des scores pronostiques. L'allogreffe de CSH constitue le seul traitement curatif des MDS. Compte-tenu de l'âge des patients et de leurs comorbidités, elle n'est cependant réalisable que dans certains cas.

a) MDS de faible risque (R-IPSS < 3,5)

L'enjeu de prise en charge des MDS dits « de faible risque » est surtout lié à la dépendance transfusionnelle et à la qualité de vie des patients. La cytopénie la plus fréquente est l'anémie et les risques liés aux transfusions itératives ne sont pas anodins. Malgré une réglementation stricte et une sécurité transfusionnelle de plus en plus importante, les patients souvent âgés restent exposés à des complications immédiates, telles que les œdèmes pulmonaires de surcharge, et à des complications retardées telles que l'hémochromatose secondaire (surcharge martiale) (19). Les ASE (agents stimulants de l'érythropoïèse) comme l'érythropoïétine (EPO) ou le *luspatercept*, peuvent permettre d'obtenir une indépendance transfusionnelle. En cas de surcharge martiale post-transfusionnelle, il est par ailleurs recommandé d'introduire des chélateurs du fer.

- *La supplémentation en EPO*

La supplémentation en EPO chez les patients atteints de MDS de faible risque avec un dosage sérique d'EPO inférieur à 500 UI/L constitue le standard de traitement en première ligne. En cas d'efficacité, bien que suspensive, cette supplémentation retarde le recours aux transfusions (20,21). Il est à noter que seule l'*epoietin α* est recommandée, même si la *darbepoietin α* est largement utilisée.

- *Le luspatercept*

Le luspatercept est une protéine recombinante inhibitrice du TGF β , lui-même impliqué dans la voie de signalisation SMAD (*Mothers against decapentaplegic homolog*). Il est recommandé pour les MDS-*SF3B1* en échec thérapeutique de la supplémentation en EPO (22). Son utilisation n'est pas encore recommandée pour tous les types de MDS de faible risque mais certains essais montrent des résultats intermédiaires encourageants (23).

- *Le lenalidomide*

Le lenalidomide est un immuno-modulateur dont l'action permet la restauration de l'érythropoïèse pour les MDS de faible risque avec 5q- isolée (24). Son utilisation n'est pas encore recommandée pour tous les types de MDS.

b) *MDS de haut risque (R-IPSS > 3,5)*

En dehors de l'allogreffe de CSH, les stratégies thérapeutiques actuelles des MDS présentant un score IPSS ou R-IPSS élevé reposent sur l'utilisation d'agents « hypo-méthylants » comme la 5'-azacytidine (azacytidine), pour le moment en monothérapie (25).

Développés en 1964, les agents hypo-méthylants (azacytidine, decitabine) sont des analogues de la cytosine (26) qui s'insèrent dans l'ADN (mais aussi l'ARN pour l'azacytidine) durant la phase S du cycle cellulaire. Ils ont la propriété d'inhiber les DNA méthyl-transférases (DNMT) par liaison covalente et de favoriser leur destruction par le protéasome. L'azacytidine permettrait de corriger l'hyper-méthylation des gènes suppresseurs de tumeur impliqués dans la pathogénèse des MDS (27). Les changements du profil de méthylation semblent persister après la fin de l'exposition aux hypo-méthylants, suggérant une reprogrammation épigénétique durable (28).

Le traitement par azacytidine est fondé sur la répétition de cycles de 4 semaines (injections journalières durant une semaine toutes les quatre semaines), jusqu'à progression ou mise en évidence d'un échec de traitement (absence de réponse). Les profils cytogénétiques et mutationnels au diagnostic ne sont que très peu corrélés à la survie et à la réponse au traitement par azacytidine (10).

L'évaluation de l'efficacité de la 5'-azacytidine est réalisée selon les critères cliniques et biologiques simples de l'*International Working Group* (IWG) (29). Dans plus de 85% des cas, ces critères ne sont cependant pas atteints avant au moins 6 mois. Environ 50% des patients sont finalement considérés comme répondeurs. En cas d'échec du traitement, les options thérapeutiques sont limitées et l'inclusion dans des essais cliniques est recommandée (30).

MDS ET ÉPIGÉNÉTIQUE

1) Expression génique et épigénétique - définitions

L'expression des gènes codants de l'ADN commence par leur transcription en acide ribonucléique messager (ARNm). Dans la cellule eucaryote, la transcription est régulée par un ensemble d'éléments nucléotidiques et protéiques. Les « éléments Cis » sont des séquences nucléotidiques régulatrices intrinsèques à l'ADN transcrit. Les « facteurs Trans » (par opposition à « Cis ») sont des éléments régulateurs externes à la séquence d'ADN. Ils sont également appelés « facteurs de transcription » (FT) et interagissent avec les éléments Cis. Les éléments Cis situés, pour la plupart, en amont du site d'initiation de la transcription (nucléotide +1) constituent le promoteur du gène. Celui-ci est variable d'un gène à l'autre mais certaines séquences promotrices sont conservées à travers le génome.

- Promoteur basal et séquences de type I

Les séquences promotrices de type I permettent la fidélité de la transcription (pour un gène donné, la transcription débute toujours au même endroit) et définissent le site d'initiation de la transcription (ISS). Elles sont proches de celui-ci (à moins de 40 bases) et constituent le promoteur basal. La « boîte TATA » est un exemple classique de séquence de type I. Sa séquence est très conservée dans le génome et d'un individu à l'autre. Elle interagit avec le facteur Trans TBP (*TATA Box Binding Protein*) pour recruter le complexe ARN polymérase II.

- Promoteur proximal et séquences de type II

Les séquences promotrices de type II permettent de définir la fréquence à laquelle la transcription du gène doit être réalisée. Elles sont situées en amont des séquences de type I mais restent proches de l'ISS (à moins de 200 bases) et constituent le « promoteur proximal ». La « boîte CAAT » est un exemple de séquence de type II.

- Promoteur distal et séquences de type III

Les séquences promotrices de type III permettent de définir la vitesse à laquelle la transcription du gène doit être réalisée. Elles peuvent être situées très en amont ou très en aval du gène mais interagissent avec le complexe ARN polymérase grâce à la structure en trois dimensions de l'ADN (maintenue par les cohésines) (**Figure 6**) (31). Leur fonctionnement est donc intimement lié à l'état de condensation de la chromatine et aux modifications des histones. Elles sont appelées « *enhancers* » lorsqu'elles sont activatrices et « *repressors* » lorsqu'elles sont inhibitrices.

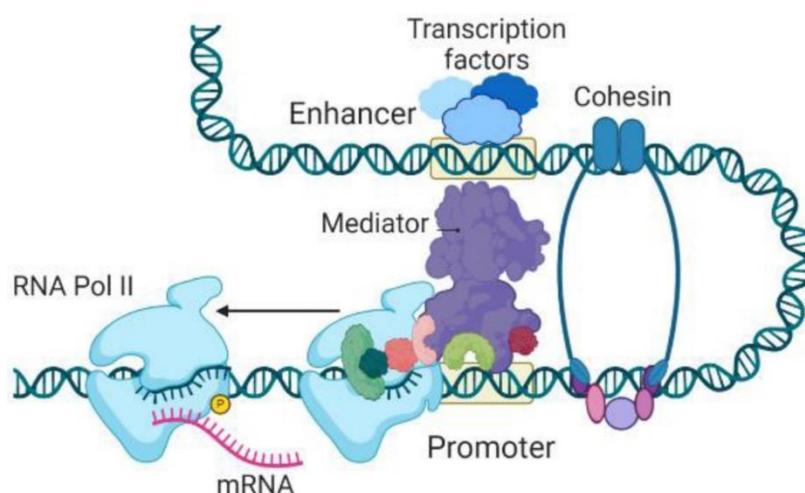


Figure 6. Régulation de la transcription par les *enhancers*.

La notion « d'épigénotype » apparaît en 1942, lorsque Conrad Hal Waddington s'intéresse à l'embryologie et aux interactions entre le génotype et le phénotype. Il considère qu'il existe un ensemble complexe de processus développementaux qui relie le génotype au phénotype (32). Le terme « épigénotype » dérive de « l'épigénèse » d'Aristote, théorie selon laquelle le développement de la forme provient de l'interaction entre ses parties (rien n'est préformé). L'épigénotype serait comparable à un ensemble de cordes qui façonnent le paysage selon leurs points de traction (« *epigenomic landscape* ») (**Annexe 1**) (32,33).

En 2024, l'épigénétique désigne l'ensemble des mécanismes de modulation de l'expression des gènes, indépendamment de leur séquence d'ADN. Elle permet l'obtention de phénotypes différents sans modification du génotype. A l'échelle de l'individu, deux cellules possédant le même patrimoine génique peuvent donc l'exprimer différemment et se différencier. Cette modulation s'effectue notamment par méthylation de l'ADN, modifications post-traductionnelles des histones et régulation des ARN messagers.

L'épigénétique est particulièrement importante en physiologie (inactivation du chromosome X, différenciation cellulaire) mais également en physiopathologie. Il est important de noter que les différentes modifications épigénétiques peuvent être héréditaires. Certains gènes sont par exemple soumis au phénomène d'empreinte, selon lequel la méthylation d'un allèle est dépendante de son origine paternelle ou maternelle. Les « marques » épigénétiques peuvent être encodées par la séquence d'ADN elle-même et dépendent également de l'environnement (nutrition de l'individu par exemple).

2) Méthylation de l'ADN

a) Méthylation physiologique de l'ADN

Chez les vertébrés, la méthylation de l'ADN concerne principalement les cytosines, formant des ^{5m}C (5-methyl-cytosines) (**Figure 7A**) (27). Les ^{5m}C peuvent être isolées ou regroupées en îlots dits « îlots CpG ».

Environ 60% des promoteurs du génome sont associés à des îlots CpG (27). Ce sont des régions riches en dinucléotides CpG (C et G contigus sur le même brin et liés par une liaison phosphodiester). Le plus souvent, leur méthylation réprime les interactions entre l'ADN et les facteurs trans-activateurs d'une part et recrute des éléments inhibiteurs d'autre part (**Figure 7B**) (27). La méthylation d'un promoteur est donc fréquemment associée à la répression de l'expression du gène régulé par ce dernier.

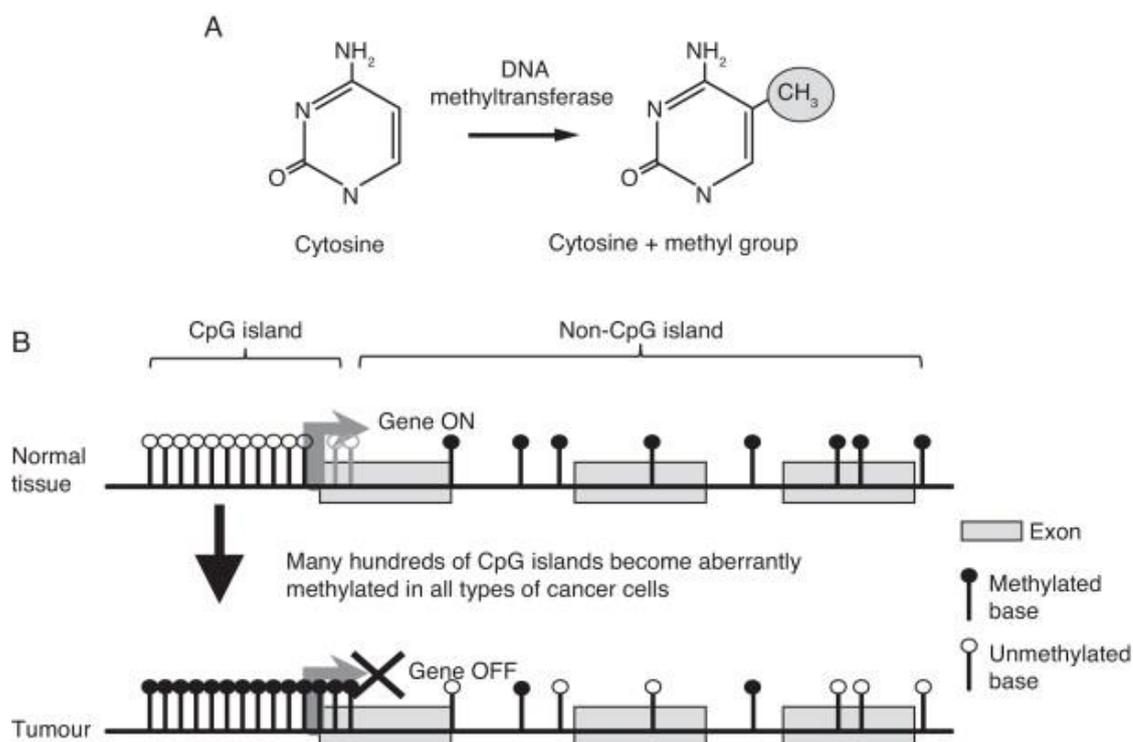


Figure 7. Méthylation des cytosines de l'ADN.

La méthylation des cytosines implique de nombreux acteurs de la machinerie cellulaire. Les groupements méthyles sont issus des cycles des folates et de la vitamine B12 via la S-adénosine-méthionine (SAM). La réaction qui transfère le groupement méthyle de la SAM sur la cytosine est catalysée par la famille des DNMT (**Figure 8**) (34). Parmi elles, **DNMT3a** est l'enzyme la plus impliquée dans les méthylations dites *de novo* post-réplcatives.

La déméthylation de l'ADN est cruciale afin de maintenir un équilibre dans l'expression génique. La conversion du groupement 5-méthyle en 5-hydroxyméthyle par les protéines **TET** initie la déméthylation active (**Figure 8**). Les bases modifiées produites sont ensuite excisées par le système de réparation BER (*base excision repair*) (35).

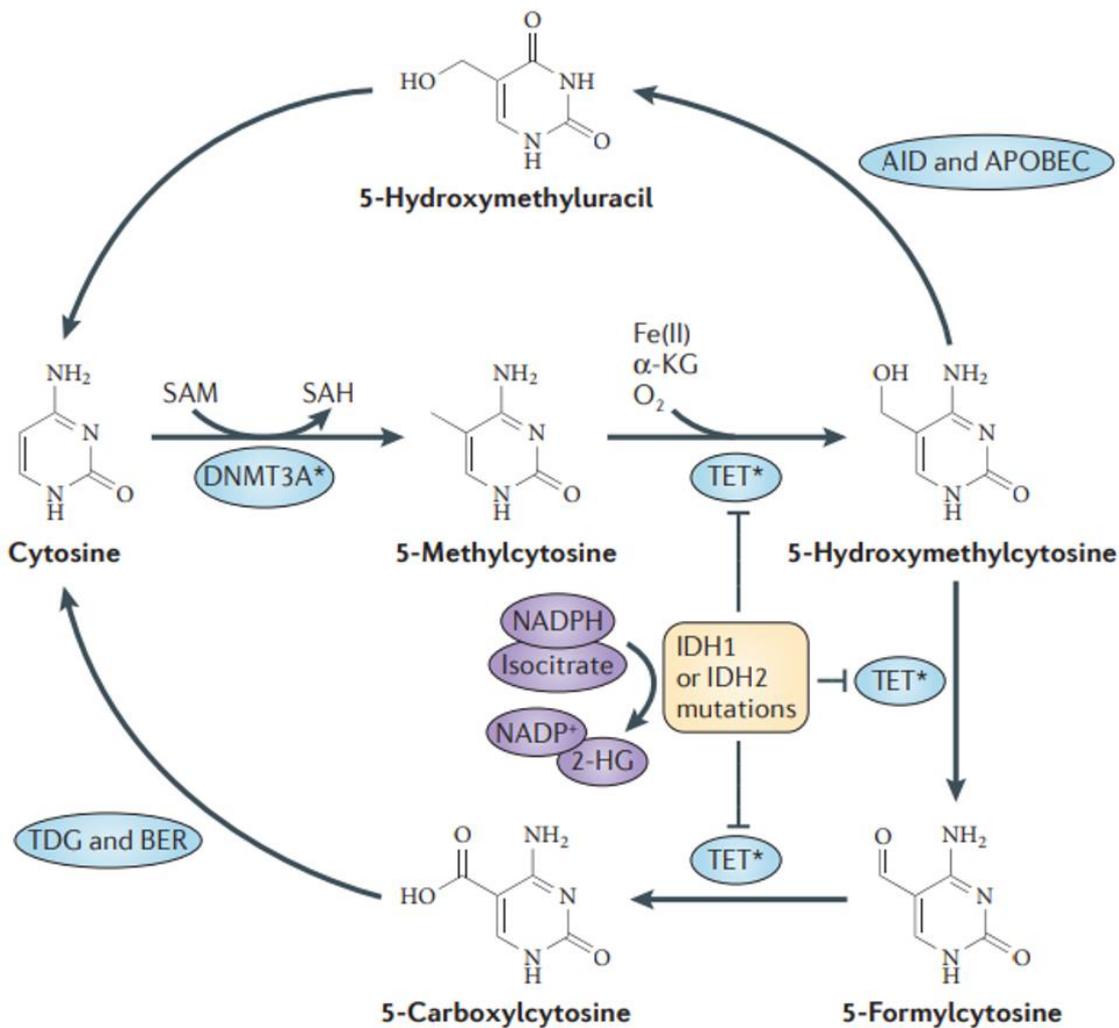


Figure 8. Méthylation et déméthylation des cytosines.
 α-KG = α-ketoglutarate

b) *Altération de la méthylation de l'ADN dans les MDS*

Les désordres de la méthylation ont été identifiés comme des acteurs majeurs de la pathogénèse des MDS (10,36,37). L'analyse du méthylome de cellules hématopoïétiques CD34+ de patients avec MDS en comparaison de cellules issues de moelles normales a montré une hyper-méthylation aberrante dispersée sur le génome et affectant des centaines de gènes (38).

- *Les mutations de DNMT3a*

Elles sont présentes dans environ 8% des MDS (27). Il peut s'agir de décalages du cadre de lecture, de mutations faux-sens ou non-sens (**Figure 9**) (39), aboutissant à une perte de fonction (soit par perte de l'activité catalytique du domaine méthyltransferase, soit par défaut de fixation à l'ADN) et entraînant donc un défaut de méthylation de certaines cytosines.

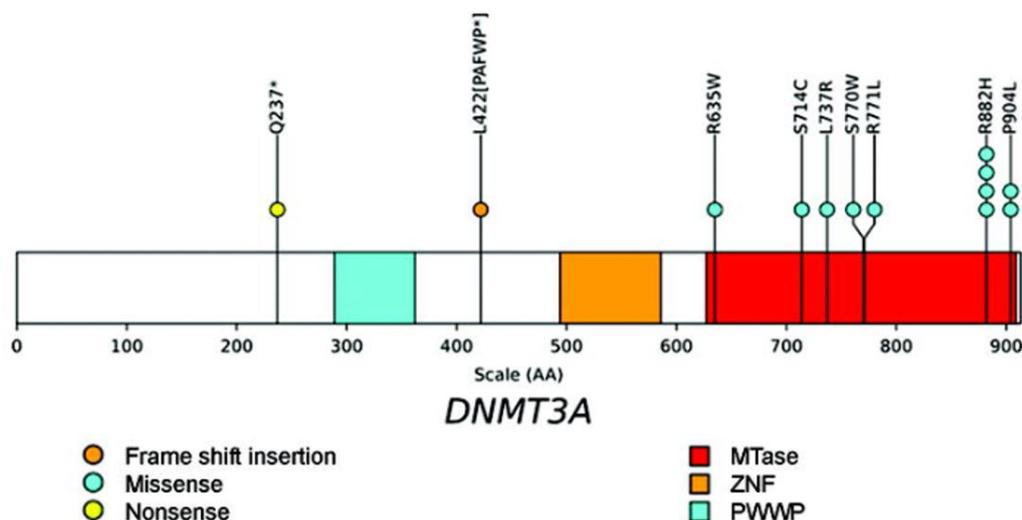


Figure 9. Principales mutations de DNMT3a dans les MDS.

Domaine PWWP = Proline-tryptophan-tryptophan-proline, domaine ZNF= zinc finger et domaine MTase = méthyltransferase.

- *Mutations de TET2*

Même si leurs conséquences fonctionnelles sur l'hyper-méthylation varient, les mutations de *TET2* sont particulièrement fréquentes (40–42). *TET2* est le seul gène de la famille TET fréquemment muté dans une grande variété de pathologies hématologiques, dont les MDS (43). Les mutations inactivatrices de *TET2* sont associées à une hyperméthylation des *enhancers* et donc à la répression de nombreux gènes impliqués dans la différenciation myéloïde (1). Les mutations de *TET2* du clone majoritaire ont été reconnues comme facteur prédictif de bonne réponse aux agents hypo-méthylants pour les MDS et les LAM pauciblastiques (44). Cet effet pronostique est perdu en cas de mutations additionnelles d'*ASXL1* et n'est pas retrouvé pour les mutations de *TET2* sous-clonales.

- Mutations d'*IDH1* et *IDH2*

Les mutations affectant *IDH1* ou *IDH2* sont retrouvées dans environ 5% des MDS. *IDH1* et *IDH2* mutés catalysent le α -ketoglutarate (nécessaire à l'activité de TET2) en 2-hydroxyglutarate. Le gain de fonction d'*IDH* (1 ou 2) a plusieurs conséquences, dont l'inhibition de TET2, de la méthylation de certaines histones et de la différenciation cellulaire, ainsi que l'altération de l'homéostasie oxydative de la cellule (45).

La cooccurrence de mutations de *TET2* et de *DNMT3a* est très fréquente dans les MDS. Leurs effets, bien qu'à première vue opposés, sont synergiques dans la pathogénèse des MDS (46). L'hyperméthylation de gènes suppresseurs de tumeurs liées aux mutants de *TET2* et les vastes régions hypométhylées liées aux mutations de *DNMT3a* s'intégreraient dans une même machinerie complexe de régulation de l'expression des gènes.

Sous traitement hypo-méthylant, l'analyse ciblée du méthylome de certains gènes peut parfois être corrélée la réponse au traitement. Le gène *p15* présente par exemple une hyperméthylation aberrante de son promoteur dans environ 50% des MDS (47). Certaines études corrélient le retour à un niveau de méthylation normal de *p15* et la réponse au traitement (48).

Certaines signatures de méthylation ont également été corrélées à la réponse à la décitabine dans la leucémie myélo-monocytaire chronique (LMMC) (26). Considérée jusqu'en 2001 comme une MDS, cette pathologie est maintenant classée comme syndrome frontière entre MDS et néoplasie myéloproliférative (MPN). Sur le plan épigénétique, elle présente des similitudes moléculaires avec les MDS (*DNMT3a*, *TET2*) et les agents hypo-méthylants restent une pierre angulaire du traitement. L'étude du méthylome et du transcriptome de patients porteurs de LMMC et traités par décitabine a révélé au diagnostic des signatures de méthylation pouvant prédire la réponse au traitement (26). De manière intéressante, cette signature était enrichie en régions comportant des *enhancers* (promoteurs distaux), souvent non intégrés dans les études précédentes qui ciblaient les promoteurs proximaux.

Il est à noter que l'azacytidine et la décitabine ont des effets différents sur le profil de méthylation et d'expression protéique (27).

3) Modifications post-traductionnelles des histones

a) *Modifications post-traductionnelles physiologiques des histones*

- *La compaction de l'ADN*

Le patrimoine génétique nucléaire de la cellule humaine contient deux fois trois milliards de paires de bases, réparties en 23 paires de chromosomes. Deux nucléotides contigus étant séparés par 0,34 nm, la longueur de ce matériel est de deux mètres, maintenue dans un noyau de quelques micromètres de diamètre. Pour être contenu dans le noyau, l'ADN nucléaire est donc compacté, enroulé autour de « nucléosomes » (**Figure 10**) (49).

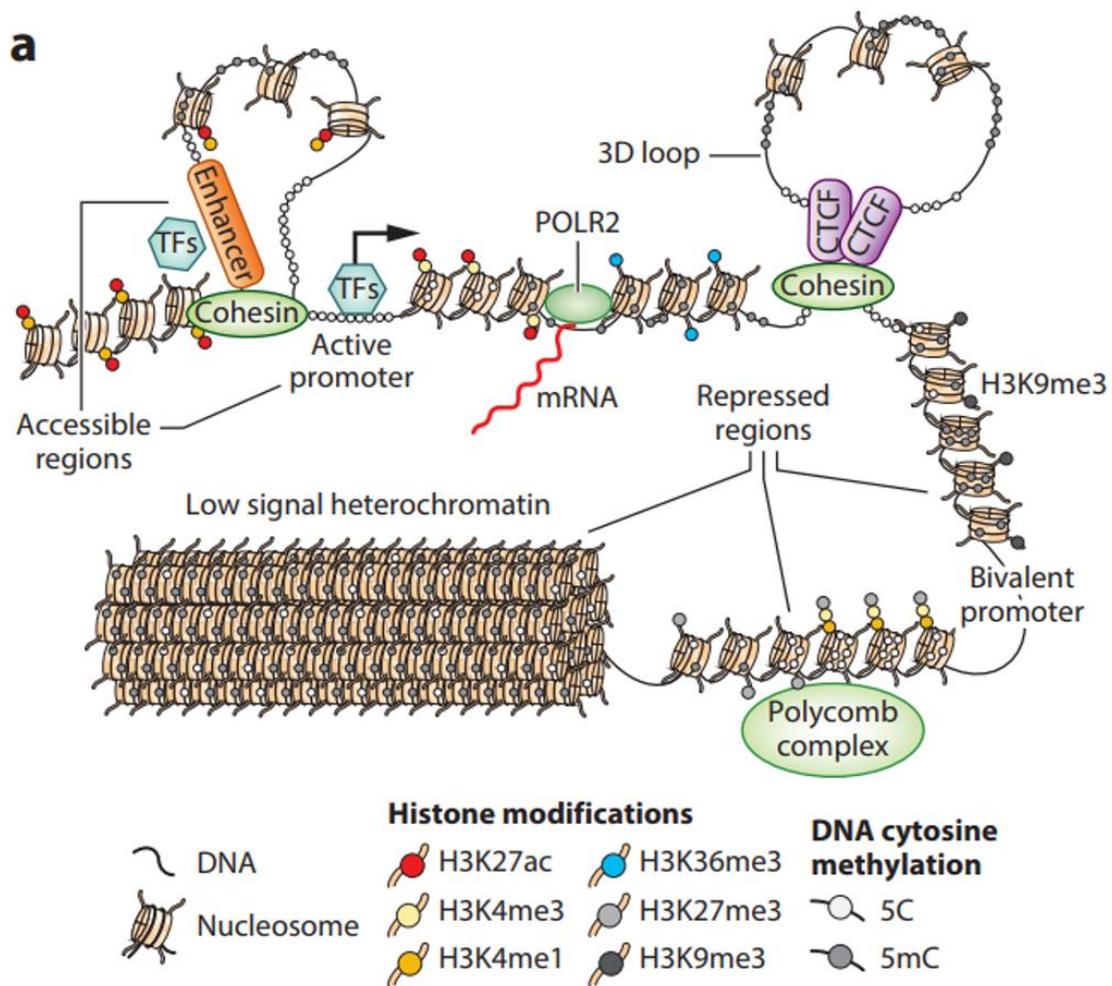


Figure 10. Compaction de l'ADN et régulation de l'expression.

Chaque nucléosome est un octamère de protéines de type histone (deux H2A, deux H2B, deux H3 et deux H4) et d'une fraction d'ADN génomique (ADN nucléosomique) qui s'y enroule. L'ADN parsemé de nucléosomes s'enroule à nouveau sur lui-même pour former un « solénoïde » (**Figure 10**). Cette structure est maintenue par des histones de type H1. Puis le solénoïde s'enroule à nouveau. Après plusieurs degrés de compaction, l'ADN prend la forme du chromosome, visible en microscopie optique. L'accès de la machinerie de transcription à l'ADN nécessite sa décompaction.

- *Modifications post-traductionnelles des histones*

Le domaine amino-terminal des histones est riche en lysines (K) et arginines (R) qui sont des acides aminés porteurs d'un groupement $-NH_3^+$ libre. La charge positive de $-NH_3^+$ interagit avec les charges négatives de l'ADN, ce qui permet le maintien de l'association ADN/protéine. L'ajout de groupements sur ces extrémités (un, deux ou trois méthyle(s), un acétyle, un phosphate ou une ubiquitine) permet de moduler ces charges et donc de réguler le degré de compaction de l'ADN autour des histones (**Figure 11**) (50).

L'**acétylation** des histones est réalisée par des enzymes appelée histone acétyl-transférases (HAT) et la désacétylation par les « histone désacétylases » (HDAC). L'acétylation neutralise la charge positive de l'histone et aboutit le plus souvent à une diminution des contacts avec l'ADN, le rendant plus accessible. Toutes les histones peuvent être ciblées par cette modification.

La **méthylation** des histones est réalisée par les « histone méthyl-transférases » (HMT). Elle concerne principalement les histones H3 et H4. Les HMT peuvent ajouter un nombre variable de groupements méthyles sur un même acide aminé.

La **phosphorylation** des histones est réalisée par des protéines kinases au niveau de résidus de type sérine (S). Son impact peut être très variable.

Selon l'acide aminé impliqué et le type/nombre de groupement(s) ajouté(s), chaque modification d'histones joue un rôle précis dans l'expression de l'ADN nucléosomique. Par exemple, la méthylation de la lysine 9 de l'histone H3 (H3K9me) est une « marque » corrélée à la compaction de la chromatine et donc à la répression de l'expression génique. Les différentes marques peuvent s'influencer les unes les autres. Le profil global des modifications post-traductionnelles des histones d'une région génique influence son expression (**Figure 11, Figure 12**) (49,50).

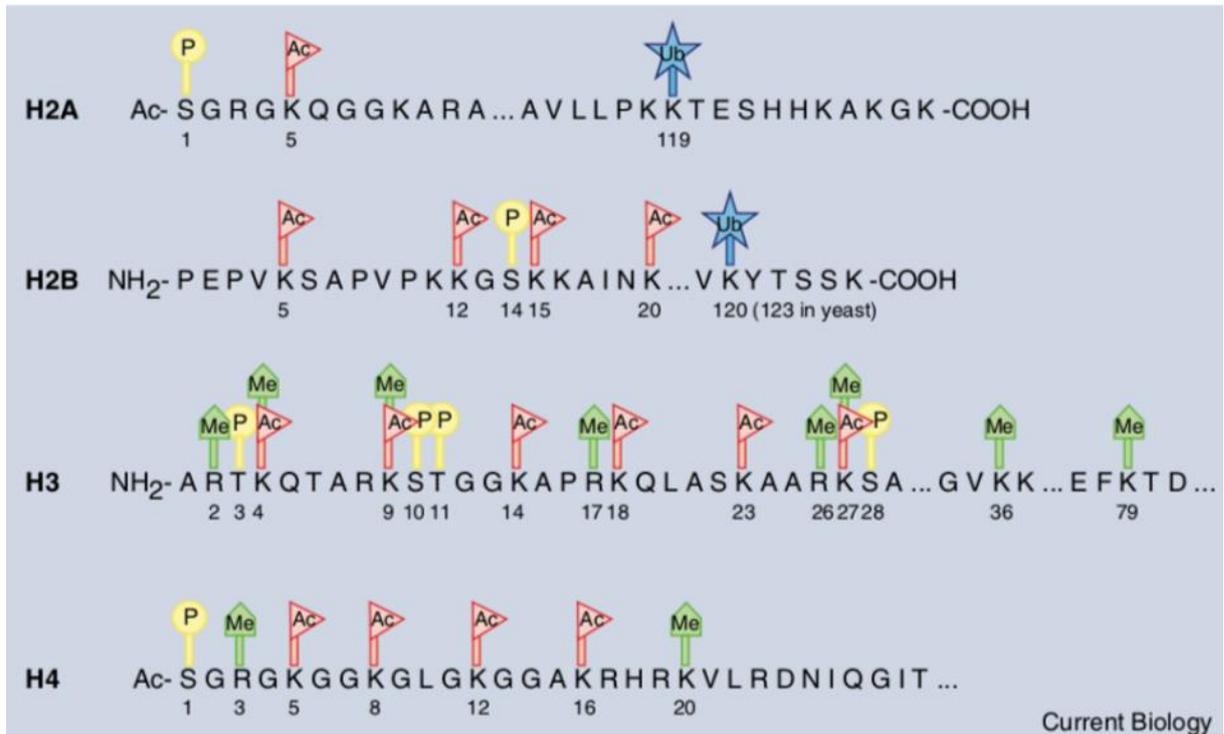


Figure 11. Principales modifications possibles des histones du nucléosome. P = phosphate ; Ac = acétyle ; Ub = ubiquitine ; Me = méthyle

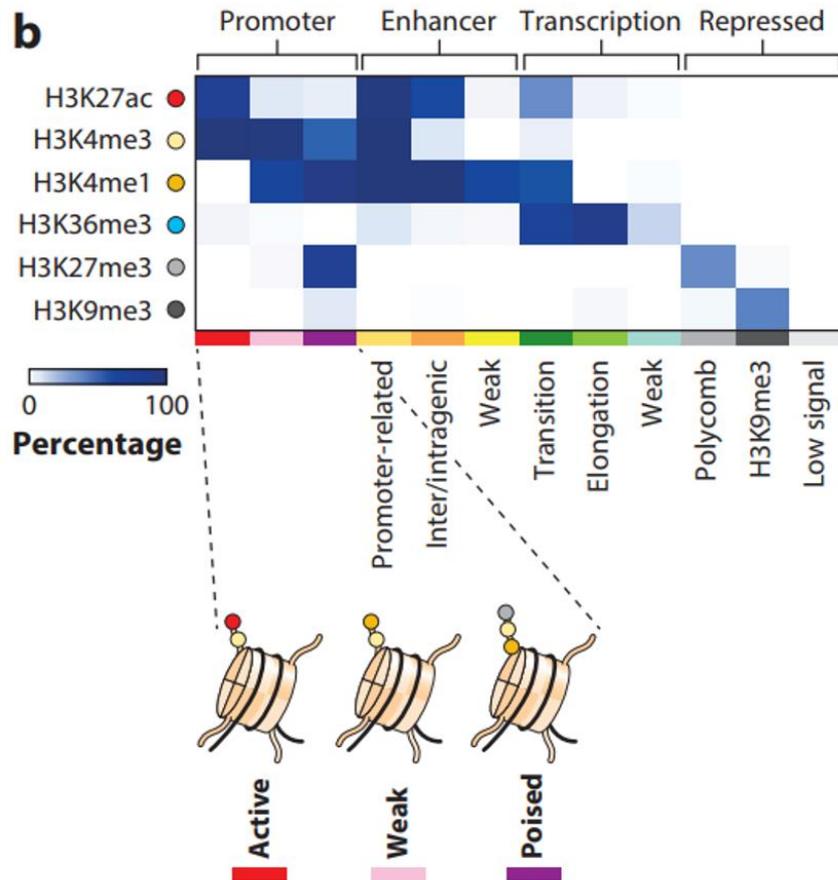


Figure 12. Etat de la chromatine en fonction des modifications des histones.

- *Le complexe Polycomb*

Les *Polycomb repressor complexes* (PRCs) correspondent à des ensembles protéiques responsables des modifications post-traductionnelles des histones (51). Il comporte notamment PRC1 (*polycomb repressor complex 1*) et PRC2 (*polycomb repressor complex 2*). PRC1 est impliqué dans l'ubiquitinylation de la lysine 11 de l'histone H2A (H2AK11) et PRC2 est impliqué dans la mono-, di- ou tri-méthylation de la lysine 27 de l'histone H3 (H3K27). H2AK11ub1 et H3K27me3 sont fortement associées à la compaction de la chromatine. Les PRCs se fixent préférentiellement sur des régions riches en promoteurs et régions régulatrices, suggérant des activités plus larges que la modification directe des histones, dont l'interaction avec des îlots CpG. Il existe plusieurs conformations et compositions de PRC1 et PRC2. PRC2.1 est un complexe PRC2 composé de 4 sous-unités dont EZH2 (*Enhancer Of Zeste 2 Polycomb Repressive Complex 2 Subunit*) (ou son paralogue EZH1).

L'action des PRCs est également régulée par différents acteurs. ASXL1 est une protéine qui s'associe à BAP1 (*BRCA1-associated protein 1*) et PRC2 pour désubiquitinyler H2A. Elle régule également l'activité de PRC1. Le domaine PHD (*plant homeodomain*) d'ASXL1 permet de détecter les modifications des histones.

b) *Modifications des histones dans les MDS*

- *Les mutations d'EZH2*

Elles sont décrites dans 2 à 6% des MDS et aboutissent à une perte de fonction. Il en résulte une diminution de la méthylation des histones et du recrutement des DNMT (1,27).

- *Les mutations de BCOR*

Elles sont décrites dans 5% des MDS. BCOR (*BCL6 Corepressor*) correspond à une sous-unité de PRC1.1. Elles sont corrélées à un pronostic défavorable (1).

- *Les mutations d'ASXL1*

Ce gène est muté dans 20% des MDS (27). Les mutations aboutissent à une protéine tronquée qui perd son domaine PHD et donc sa capacité d'interaction avec les PRCs, contribuant aux désordres d'expression des MDS (**Figure 13**) (52).

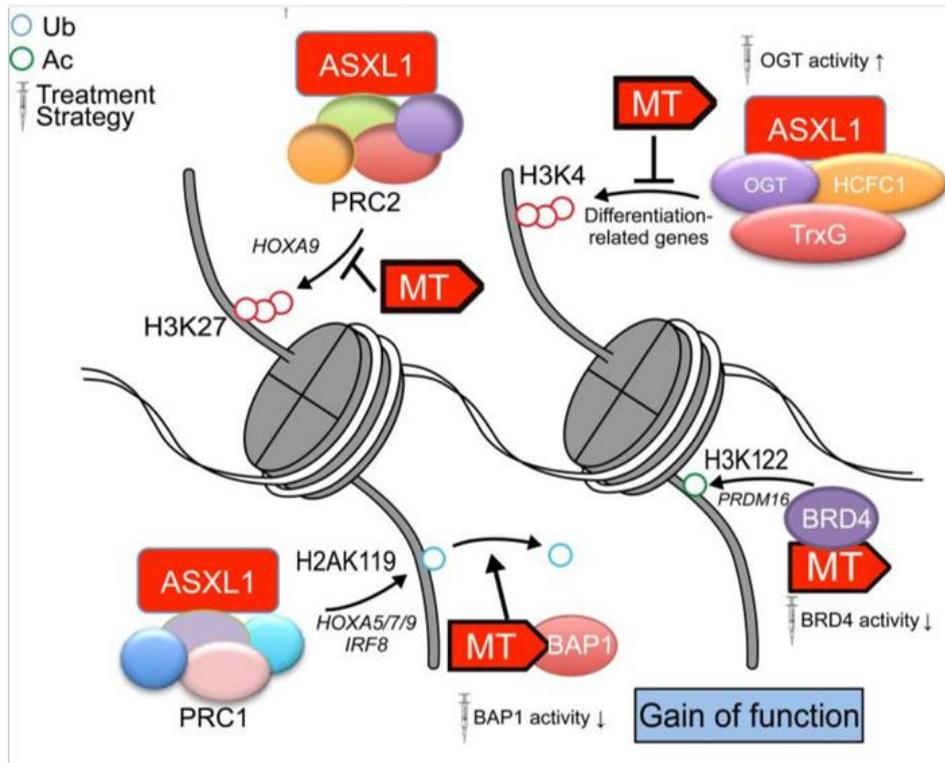


Figure 13. Effet des mutations d'*ASXL1* sur les PRCs, H2A et H3.
MT = *ASXL1* muté.

MDS ET ALTÉRATIONS DE L'ÉPISSAGE

1) L'épissage

La maturation du pré-ARN messager comprend trois étapes : (i) la coiffe 5', (ii) l'épissage et (iii) la coiffe 3' (ou poly-adénylation). L'épissage correspond à l'excision des introns et à la réassociation des exons. Il met en jeu des réactions de trans-estérification par attaque nucléophile de groupements –OH (**Figure 14**).

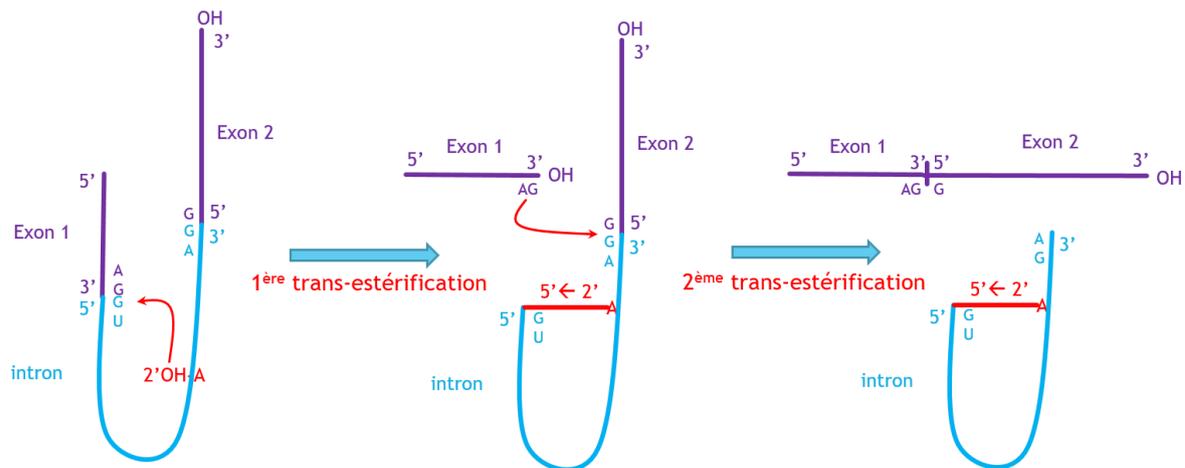


Figure 14. Réaction d'épissage d'un intron.

La réaction implique trois éléments majeurs de la séquence du pré-ARN messager, (**Figure 15**) :

- **Un site de branchement** situé dans l'intron à épisser : **adénosine** dans la majorité des cas, au sein d'une séquence très conservée. Son groupement 2'-OH est responsable de la première attaque nucléophile.
- **Des dinucléotides** (très conservés) situés aux extrémités de l'intron à épisser : **GU** et **AG** (configuration la plus fréquente, appelée « spliceosome majeur »). Le groupement 3'-OH du dinucléotide AG, libéré après la première trans-estérification, est responsable de la deuxième attaque nucléophile.
- Des **séquences flanquantes** aux dinucléotides, plus variables et riches en pyrimidines.

Les dinucléotides et leurs séquences flanquantes constituent les sites **donneur** et **accepteur** de l'épissage : le site 5' (extrémité 5' de l'intron à épisser) correspond au site « donneur » et le site 3' (extrémité 3' de l'intron à épisser) correspond au site « accepteur ».

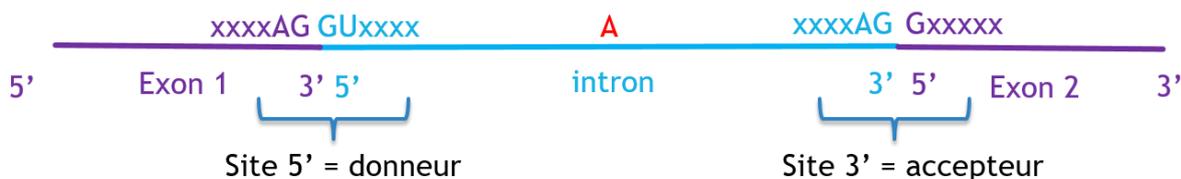


Figure 15. Eléments nucléotidiques du pré-ARN messager nécessaires à l'épissage.

Afin d'être compatible avec le rythme cellulaire, l'épissage nécessite une catalyse. Celle-ci est réalisée par un complexe nucléoprotéique comportant des petits ARN nucléaires (SnRNA) associés à des protéines et formant alors des SnRNPs (*small nuclear ribo-nucleoproteins*). Les SnRNPs impliquées dans l'épissage sont U1, U2, U4, U5 et U6. L'ensemble des acteurs nécessaires à la catalyse de la réaction d'épissage (séquences nucléotidiques et SnRNPs) est appelé « spliceosome ».

La catalyse de l'épissage au sein du spliceosome implique des rapprochements stériques entre les groupements –OH responsables des attaques nucléophiles et les sites d'épissage (**Figure 16**).

La réaction d'épissage est donc fortement dépendante du positionnement des différents acteurs du spliceosome sur la séquence nucléotidique à épisser. La ribo-nucléoprotéine U2 (U2 SnRNP) permet notamment d'associer le site de branchement et le site d'épissage 3'.

La protéine SF3B1 est une sous-unité de U2 (**Figure 17**) qui interagit avec la protéine U2AF1 pour faire correspondre la « bonne » adénosine au « bon » site 3'(53). En effet, pour un pré-ARNm donné, bien qu'il existe un site d'épissage « canonique » majoritairement recruté, d'autres sites alternatifs peuvent être utilisés. Cet « épissage alternatif » permet d'augmenter le répertoire protéomique et explique certaines isoformes d'une même protéine. Chez l'Homme, plus de 95% des gènes sont concernés par l'épissage alternatif (54).

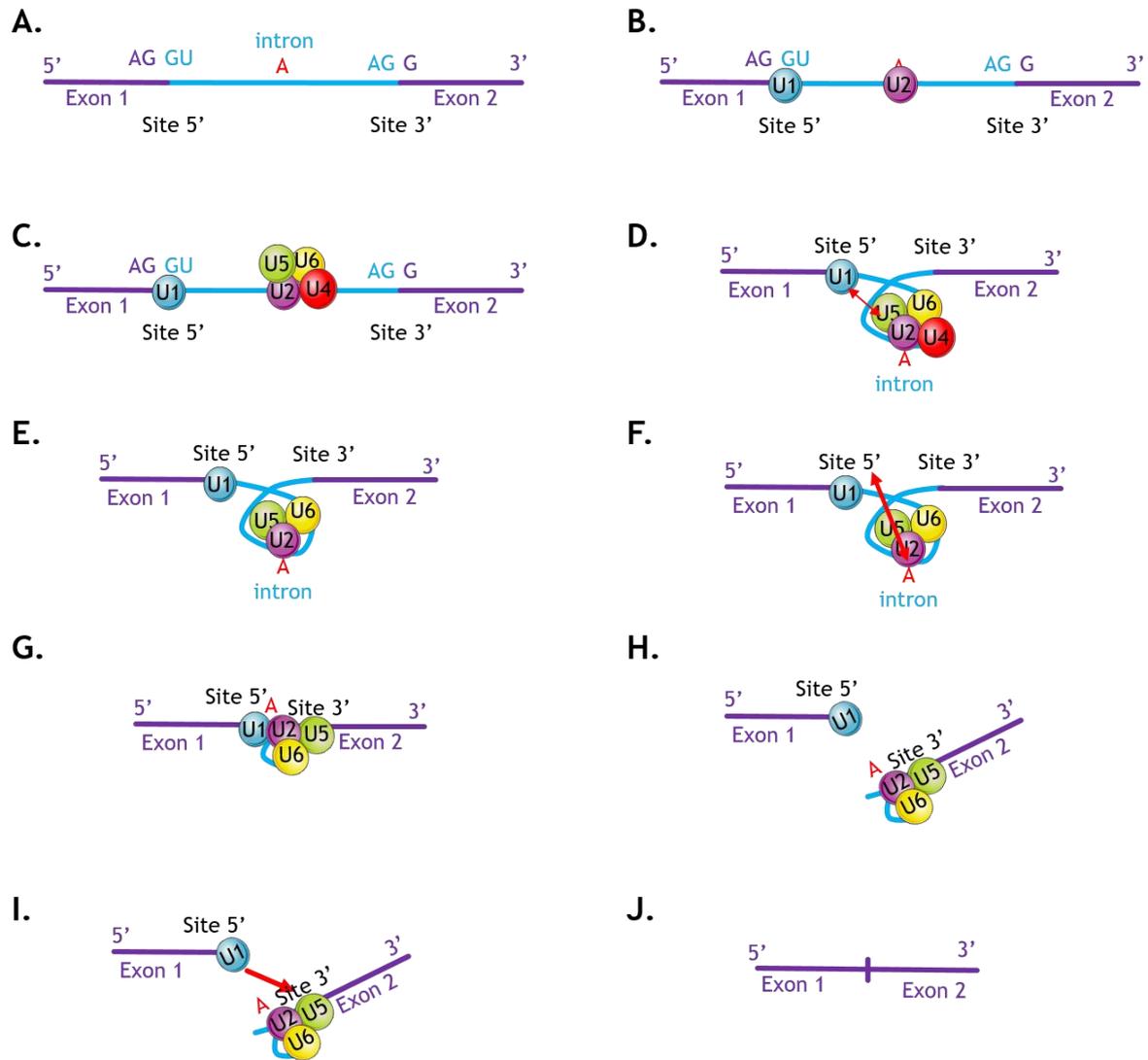


Figure 16

. Etapes de l'épissage catalysé par le spliceosome.

A. pré-ARN messager. **B.** Fixation de U1 sur le site 5' et de U2 sur le site de branchement. **C.** Fixation de U4, U5 et U6 sur le site de branchement et interactions de U4 avec U2. **D.** Interactions de U5 avec le site 5'. **E.** Libération du site catalytique de U2 par éloignement de U4. **F.** Rapprochement du site de branchement et du site 5' grâce à U2. **G et H.** Attaque nucléophile du groupement 2' –OH de l'adénosine de branchement sur le site 5' : **première trans-estérification**. **I.** Attaque nucléophile du groupement 3' –OH libre du site 5' sur le site 3' : **deuxième trans-estérification**. **J.** ARN messager épissé.

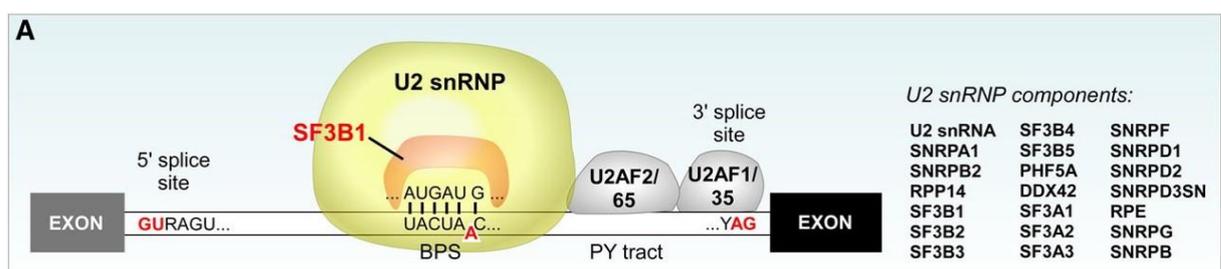


Figure 17. Positionnement de SF3B1 au sein du spliceosome.

La régulation de l'épissage met en jeux des éléments Cis (de la séquence nucléotidique) qui interagissent avec des facteurs protéiques (54). Les séquences nucléotidiques favorisant l'épissage sont appelées ESE (*exonic splicing enhancer*) et ISE (*intronic splicing enhancer*) selon leur localisation. Celles qui le répriment sont appelées ESS (*exonic splicing silencer*) ou ISS (*intronic splicing silencer*) (**Figure 18**).

Les protéines SR (*Serine-arginin repeat*) telles que SRSF2 (*Serine And Arginine Rich Splicing Factor 2*) interagissent généralement avec les ESE et les ISE pour promouvoir l'épissage. Les protéines hnRNP (*heterogeneous nuclear ribonucleoprotein particle*) quant à elles sont plutôt associées à la répression via leurs interactions avec les ESS et ISS (**Figure 18**).

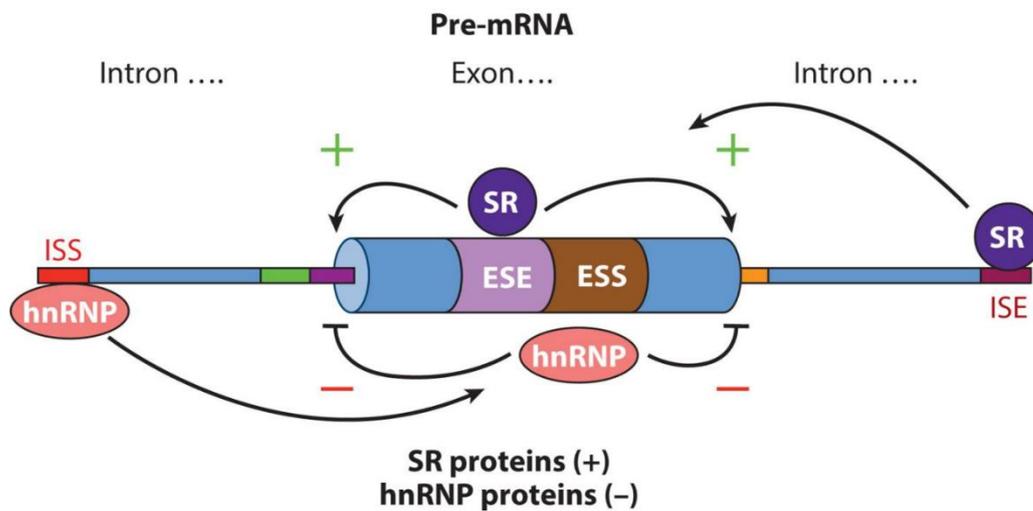


Figure 18. Régulation positive et négative de l'épissage.

ISS = *intronic splicing silencer*; ESS = *exonic splicing silencer*; ISE = *intronic splicing enhancer*; ESE = *exonic splicing enhancer*; SR = *serine-arginine repeat protein*; hnRNP = *heterogeneous nuclear ribonucleoprotein particle*.

2) Mutations récurrentes impactant l'épissage dans les MDS

Environ 60% des MDS présentent des mutations impactant le spliceosome, notamment au niveau des gènes *SF3B1*, *SRSF2*, *U2AF1* et *ZRSR2* (*Zinc Finger CCCH-Type, RNA Binding Motif And Serine/Arginine Rich 2*) (**Figure 19**) (8,55,56).

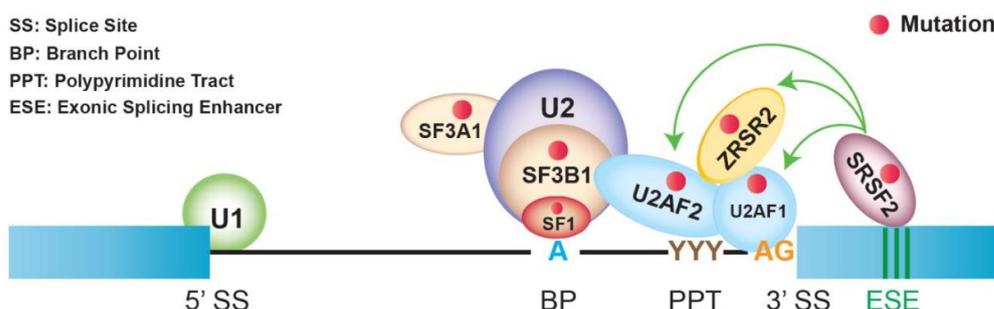


Figure 19. Mutations du spliceosome dans les MDS.

a) *Impact des mutations de SF3B1*

Les mutations de *SF3B1* aboutissent à un « changement de fonction » avec reconnaissance aberrante, via les protéines U2AF, d'un site 3' alternatif et de son association à un site de branchement alternatif (**Figure 20**) (57,58).

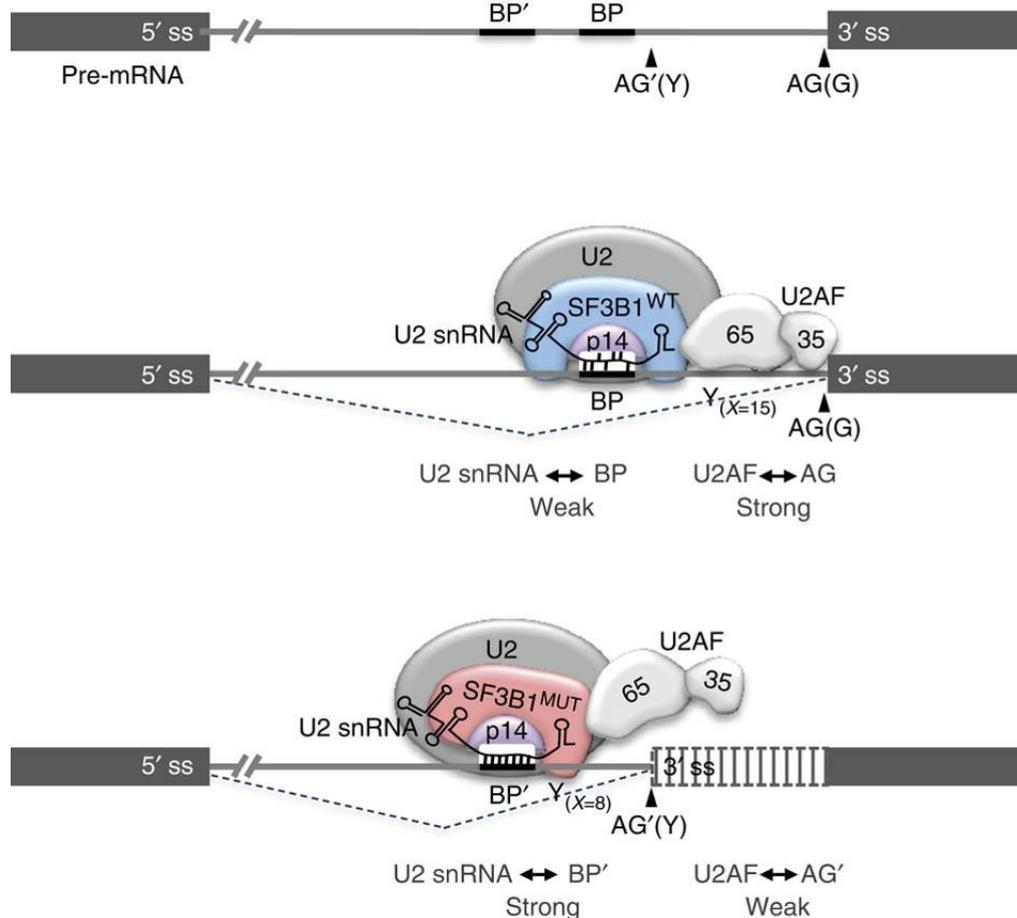


Figure 20. Modèle d'altération de l'épissage par les mutations de *SF3B1*.

BP = site de branchement ; BP' = site de branchement alternatif ; SS = site de *splicing*.

Il en résulte donc des transcrits aberrants qui, s'ils ne sont pas détruits par le NMD (*nonsense-mediated decay*) aboutissent à des protéines aberrantes (**Figure 21**) (57). La mutation de *SF3B1* peut donc aboutir à l'haplo-insuffisance ou au gain de fonction d'autres gènes.

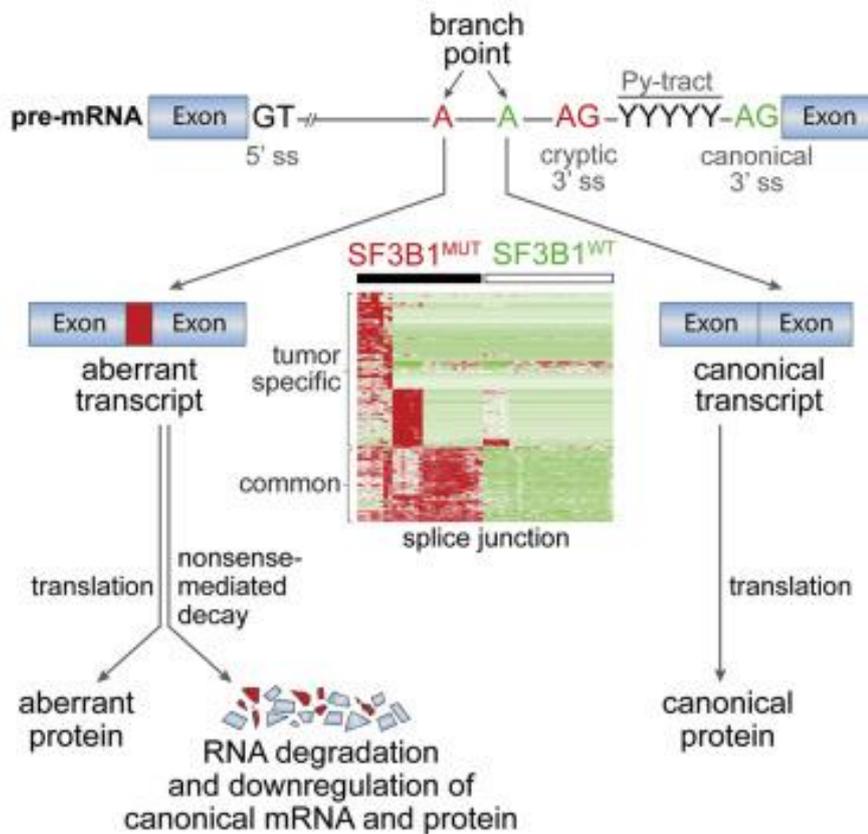


Figure 21. Synthèse de transcrits aberrants induits par les mutations de *SF3B1*.

Cela concerne notamment une centaine de gènes impliqués dans le déroulement de l'érythropoïèse et dans l'homéostasie du fer, ce qui explique la forte corrélation (même si elle n'est pas absolue) entre la présence d'une mutation de *SF3B1* et celle de sidéroblastes en couronne (SC) (59). Plus précisément, il a été démontré que les altérations d'épissage conduisant à la sous-expression des gènes *TMEM14C* (*Transmembrane Protein 14C*) et *ABCB7* (*ATP Binding Cassette Subfamily B Member 7*) étaient suffisantes pour induire la formation des SC.

Les mutations de *SF3B1* sont les événements les plus courants d'altération du spliceosome dans les MDS et lorsqu'elles sont isolées, sont plutôt corrélées à un pronostic favorable (1).

b) *Impact des mutations de SRSF2*

Les mutations de *SRSF2* sont également très fréquentes dans les MDS. L'altération de leurs interactions avec les ESE aboutit à une dérégulation de l'épissage de certains transcrits dont les transcrits *EZH2* (1). Les mutations affectant le spliceosome peuvent donc indirectement impacter la méthylation des histones, rappelant la complexité de la physiopathologie des MDS.

c) *Impact des mutations de U2AF1*

L'altération des interactions entre U2AF1 (et donc le reste du spliceosome) et le site 3' peut aboutir à un « saut d'exon ». En effet, si la reconnaissance se fait avec le site 3' de l'intron suivant, l'exon positionné entre les deux introns sera épissé dans le même temps.

3) *Les MDS-SF3B1*

Selon la classification OMS 2022 (2), l'ensemble « MDS-*SF3B1* » regroupe les MDS avec une faible proportion de blastes (moins de 5% de blastes médullaires et moins de 2% de blastes sanguins) associés à la présence d'une mutation de *SF3B1* (exclusion de cet ensemble en cas de délétion 5q, de monosomie 7 ou de complexité au caryotype) **ou** d'au moins 15% de SC (« MDS sans excès de blastes avec SC »).

Les SC correspondent à des érythroblastes dont les mitochondries sont surchargées en fer et disposées le long de la circonférence nucléaire (60). Ils sont mis en évidence par la coloration de Perls (bleu de Prusse) (**Figure 22**).

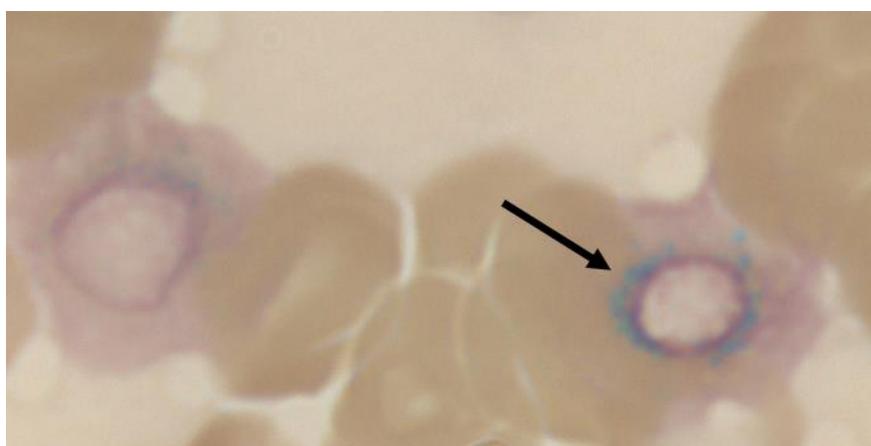


Figure 22. Sidéroblaste en couronne en coloration de Perls, en microscopie optique (x100).

Sur le plan moléculaire, les MDS présentant des SC sont surtout reliés aux mutations de *SF3B1*, *SRSF2* et *TP53* (*multi hit*). *SF3B1* est muté dans 80% des MDS présentant des SC. La cooccurrence de mutations de *SF3B1* et de *SRSF2* est rare, de même que *SF3B1* avec *TP53*. La présence de SC peut toutefois en être indépendante et mise en évidence dans des MDS-LB, des syndromes frontière entre MDS et MPN ou même des LAM (61). Certains toxiques peuvent également être responsables de cet aspect (notamment l'alcool, l'isoniazide ou le linezolid). Il existe également des formes congénitales (62).

Le pronostic des MDS-*SF3B1* est favorable en l'absence de facteur cytogénétique ou mutationnel supplémentaire de haut risque. Quand *SF3B1* est muté, il n'y a pas de différence de survie entre les MDS avec plus de 15% de SC et ceux qui en présentent 5 à 14%.

Concernant les MDS avec SC non mutés pour *SF3B1*, en cas de mutation de *TP53* ou *SRSF2*, les données de survie ne montrent pas de différence entre les MDS-LB avec et les MDS-LB sans SC. Ces données expliquent que la classification ICC 2022 ne conserve, dans l'entité des MDS-*SF3B1*, que les MDS mutés *SF3B1* (avec une fréquence allélique supérieure à 10%), indépendamment du nombre de SC (et à l'exclusion des mutés *TP53* multi-hit) (14).

La prise en charge des MDS-*SF3B1* repose sur la supplémentation en EPO en première ligne et sur l'utilisation du luspaterecept en cas d'échec ou de perte de réponse (22). Des études sur l'utilisation du luspaterecept en première ligne sont en cours.

LES NÉOPLASIES MYÉLOPROLIFÉRATIVES

1) Les types principaux de MPN

Les MPN constituent un ensemble d'hémopathies malignes chroniques acquises et clonales, caractérisés par l'hyperplasie médullaire d'une ou plusieurs lignées myéloïdes, sans blocage de maturation, conduisant à une augmentation du nombre de cellules circulantes (granulocytes, érythrocytes et/ou plaquettes). Ces pathologies partagent des caractéristiques cliniques telles que la splénomégalie (fréquente dans tous les MPN) ou le risque de thrombose (MPN non liés au chromosome Philadelphie).

Comme les MDS, les MPN sont associés à des anomalies génétiques récurrentes et sont classés en deux grandes catégories selon la présence ou non du chromosome « Philadelphie » (Phi), chromosome 22 amoindri issu de la translocation $t(9;22)(q34;q11)$. Cette translocation aboutit au transcrite de fusion *BCR-ABL*, caractéristique de la leucémie myéloïde chronique (LMC). Les MPN ne présentant pas de chromosome Phi sont surtout représentées par la *polycythaemia vera* ou polyglobulie de Vaquez (PV), la thrombocytémie essentielle (TE) et la myélofibrose (MF). Leur physiopathologie implique notamment la mutation gain de fonction *JAK2^{V617F}* située dans l'exon 14 de *JAK2* (*Janus Tyrosine Kinase 2*) (63), les mutations faux sens situées dans l'exon 10 de *MPL* (*Myeloproliferative Leukemia*) (64) et les insertions/délétions avec décalage du cadre de lecture situées dans l'exon 9 de *CALR* (*Calreticulin*) (65). La mutation *JAK2^{V617F}* est rapportée dans plus de 95% des cas de PV, mais également dans 50 à 60% des cas de TE et de MF (66).

En l'absence de mutation de *JAK2*, on retrouve des mutations au sein des gènes *CALR* et *MPL* : environ 30% des TE et des MF pour *CALR* et 1% et 5% pour *MPL* dans les TE et MF respectivement (**Figure 23**) (65). Les MF et les TE ne présentant aucune mutation de *JAK2*, *CALR* ou *MPL* sont dites « triple négatives » (TN).

Bien que non spécifiques des MPN, des mutations affectant les régulateurs épigénétiques tels que *TET2*, *DNMT3a*, *EZH2*, *IDH1* et *IDH2* ont également été décrites dans les MPN (67,68).

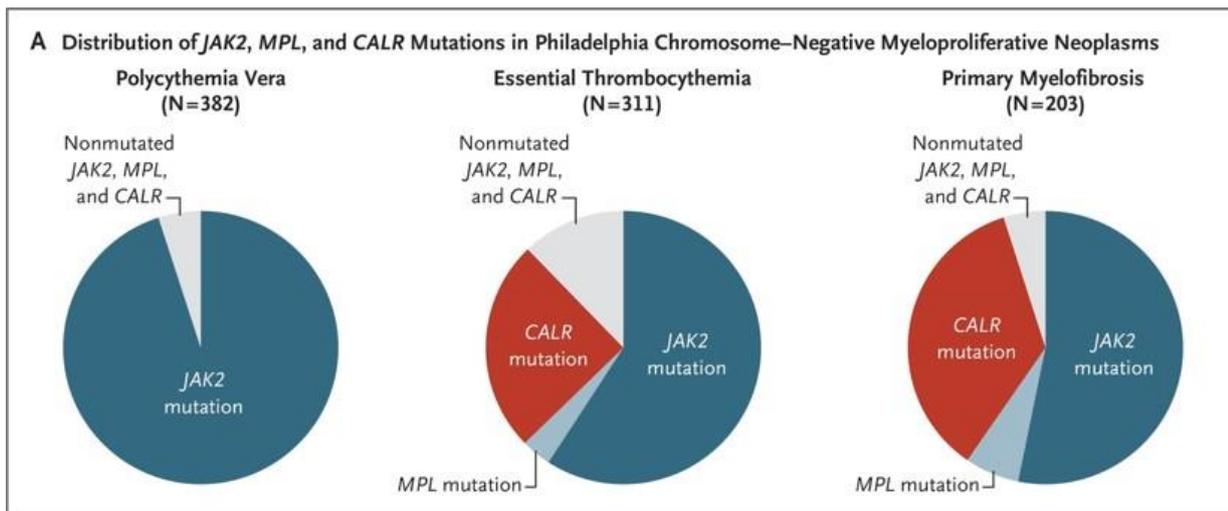


Figure 23. Distribution des mutations de JAK2, CALR et MPL dans la PV, la TE et la MF.

2) La voie JAK-STAT

a) *Les Tyrosines kinases*

Les protéines de type tyrosine kinase correspondent à une large famille d'enzymes, dont l'activité permet la phosphorylation d'autres protéines impliquées dans la transduction du signal menant à une réponse cellulaire après un stimulus extracellulaire.

V-Src (Tyrosine-protein kinase transforming protein Sarcoma) est le premier gène codant pour une tyrosine kinase qui a pu être identifié, au sein du *Rous Sarcoma Virus (RSV)* en 1970 (69). Il a rapidement été relié à la cancérogénèse de cellules infectées par le RSV. L'activité de phosphorylation de la protéine correspondante et de son homologue eucaryote *c-src (cellular-src)* a été mise en évidence dans les années 80.

C-src a rapidement été suspectée d'interagir avec de nombreux partenaires, conduisant à la découverte de domaines d'homologie SH (*Src homology domain*) avec d'autres protéines kinases : (i) SH1 portant l'activité catalytique, (ii) SH2 correspondant à un domaine permettant la reconnaissance de motifs phosphotyrosine et étant fortement conservé entre différentes tyrosines kinases (**Figure 24**) (70), et (iii) SH3.

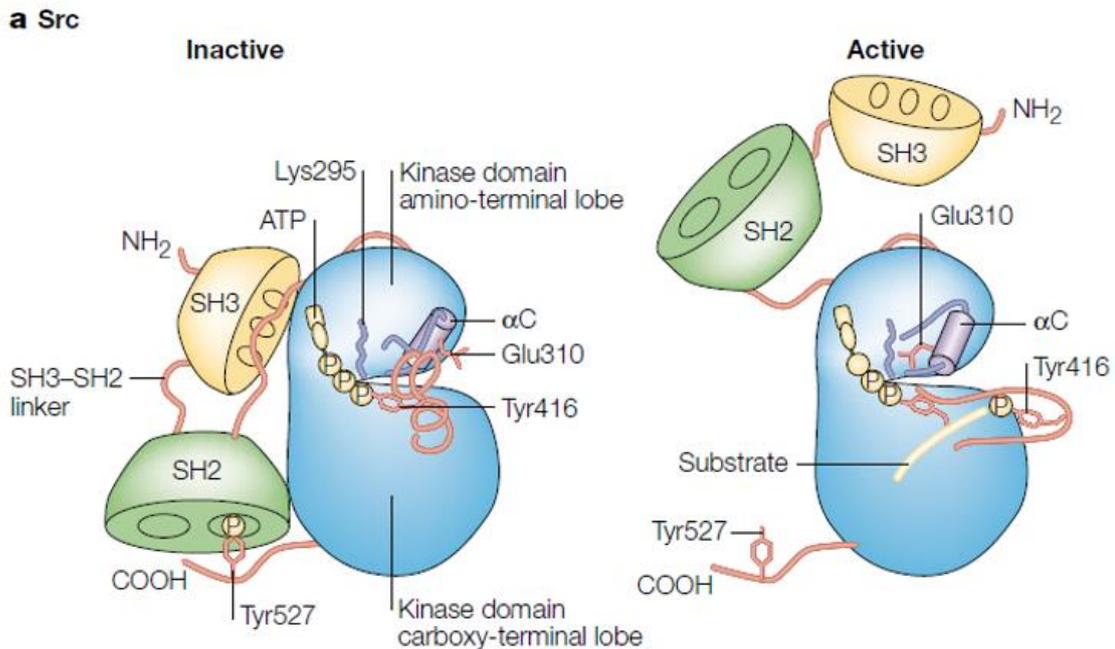


Figure 24. Mécanisme moléculaire de la fonction du domaine SH2 dans la protéine src.

Devant la forte conservation de la séquence du domaine SH2 dans les tyrosine kinases, l'utilisation d'amorces qui lui sont complémentaires ont permis d'identifier de nouvelles kinases. En 1988, Hanks *et al.* découvrent 65 nouvelles protéines kinases (PTK) (71). Les PTK sont classifiées selon leur domaine d'homologie (famille Src, Fes ou CSF par exemple).

b) *JAK et STAT*

- *JAK*

En 1989, de nombreuses tyrosine kinases ont déjà été identifiées comme étant des récepteurs de certains facteurs de croissance. Wilks *et al.* utilisent des amorces recombinées (PTK I et PTK II) pour isoler dans le tissu hématopoïétique une nouvelle classe de protéines kinases : les protéines JAK (72,73). Initialement appelées *Just Another Kinase* (JAK), elles ont ensuite été renommées *Janus tyrosine Kinase*, en référence au dieu romain Janus, qui avait la particularité de posséder deux visages (évoquant les domaines kinase et pseudokinase de JAK).

De manière générale, les protéines JAK et leurs partenaires STAT (*Signal transducer and activator of transcription*) sont associées à la superfamille des récepteurs cytokiniques de type I. L'activation du récepteur membranaire par sa cytokine induit une dimérisation de celui-ci, ce qui aboutit au rapprochement de deux protéines JAK inactives qui vont s'activer par phosphorylation croisée mutuelle.

Cette activation conduit classiquement au recrutement et à la phosphorylation de protéines STAT, constituant la voie d'activation JAK-STAT. Deux protéines STAT activées peuvent se dimériser à leur tour (par interaction entre leurs domaines SH2) avant d'être importées dans le noyau, où elles jouent le rôle de facteurs Trans et promeuvent la transcription de gènes spécifiques (**Figure 25**). La voie JAK-STAT est donc une voie clé de l'expression génique intégrant des signaux cytokiniques, hormonaux et des facteurs de croissances. Les protéines STAT non phosphorylées agissent également sur de multiples cibles par d'autres mécanismes moins classiques.

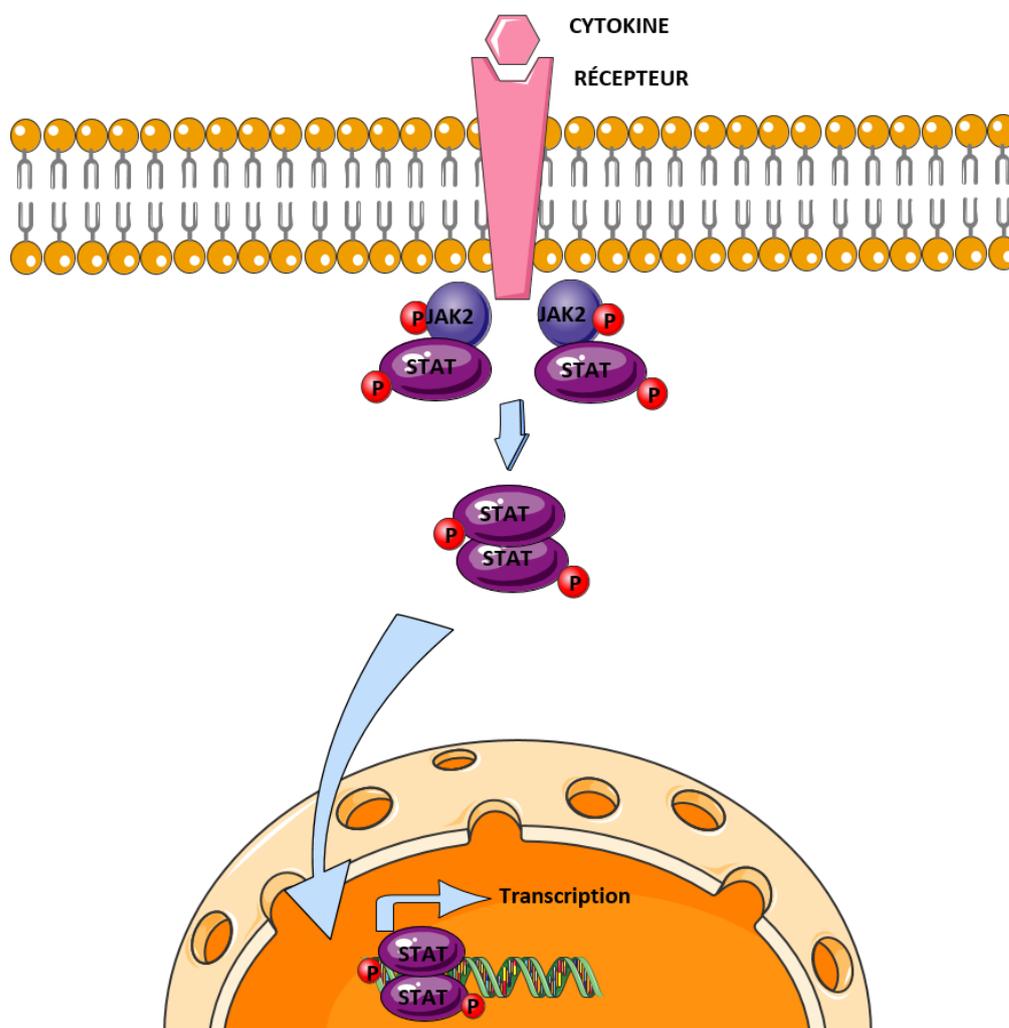


Figure 25. Voie de signalisation JAK-STAT.

La famille JAK comprend quatre protéines : JAK1, JAK2, JAK3 et TYK2. Ces protéines homologues ont un poids moléculaire de 120 à 130 kDa et sont retrouvées de manière ubiquitaire dans les cellules des mammifères, à l'exception de JAK3 qui est surtout présente dans le tissu hématopoïétique. Une fonction primordiale des protéines JAK est la transduction du signal cytokinique (74,75).

Chez l'Homme, le gène *JAK2* est situé sur le chromosome 9 (9p24.1). Comme toutes les protéines de la famille JAK, JAK2 comporte sept domaines d'homologie (*JAK homology ou JH*) (**Figure 26**) (76) parmi lesquels JH1 et JH2 (74). JH2 (domaine pseudokinase), dépourvu d'activité catalytique (73,77), est responsable de l'inhibition du domaine portant l'activité kinase JH1 (78). JH3 et une partie de JH4 forment le domaine SH2. JH5, JH6, JH7 et la partie résiduelle de JH4 forment le domaine FERM, impliqué dans les interactions avec le récepteur membranaire. JAK2 peut être issu de deux transcrits différents et comporte de multiples sites d'autophosphorylation (74).

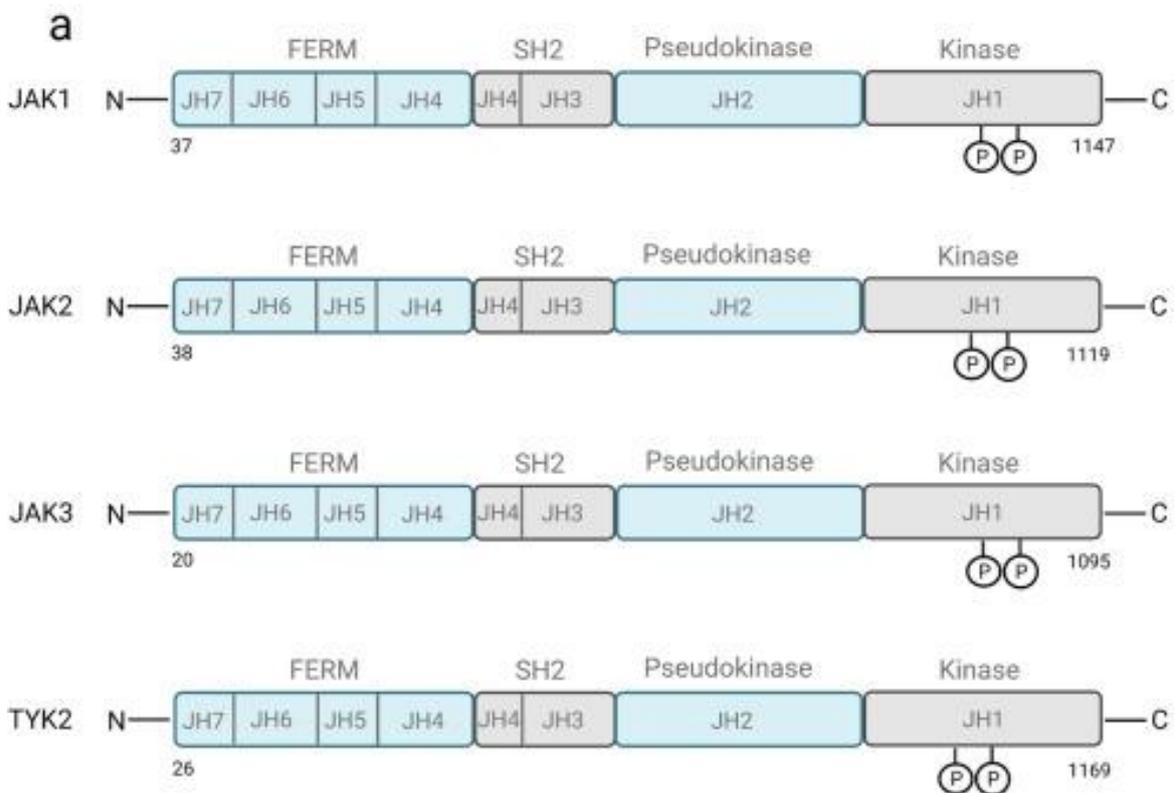


Figure 26. Structure des protéines JAKs et leurs sites de phosphorylation conservés.

Les protéines JAKs sont couplées à de nombreux récepteurs (**Figure 27**) (79). JAK2 est ubiquitaire mais particulièrement associé à l'hématopoïèse. Cette protéine est notamment couplée aux récepteurs de la thrombopoïétine (TPO) et de l'EPO (80,81), mais peut être activée par les récepteurs de cytokines de type IL3 (*interleukin-3*), IL5 (*interleukin-5*) ou encore GM-CSF (*Granulocyte Macrophage Colony-Stimulating Factor*) (74).

Son rôle est essentiel dans la transduction du signal de survie cellulaire et de prolifération et sa délétion germinale entraîne des troubles létaux de l'hématopoïèse embryonnaire chez la Souris (82). La délétion conditionnelle somatique de *JAK2* dans les CSH entraîne également une aplasie médullaire sévère.

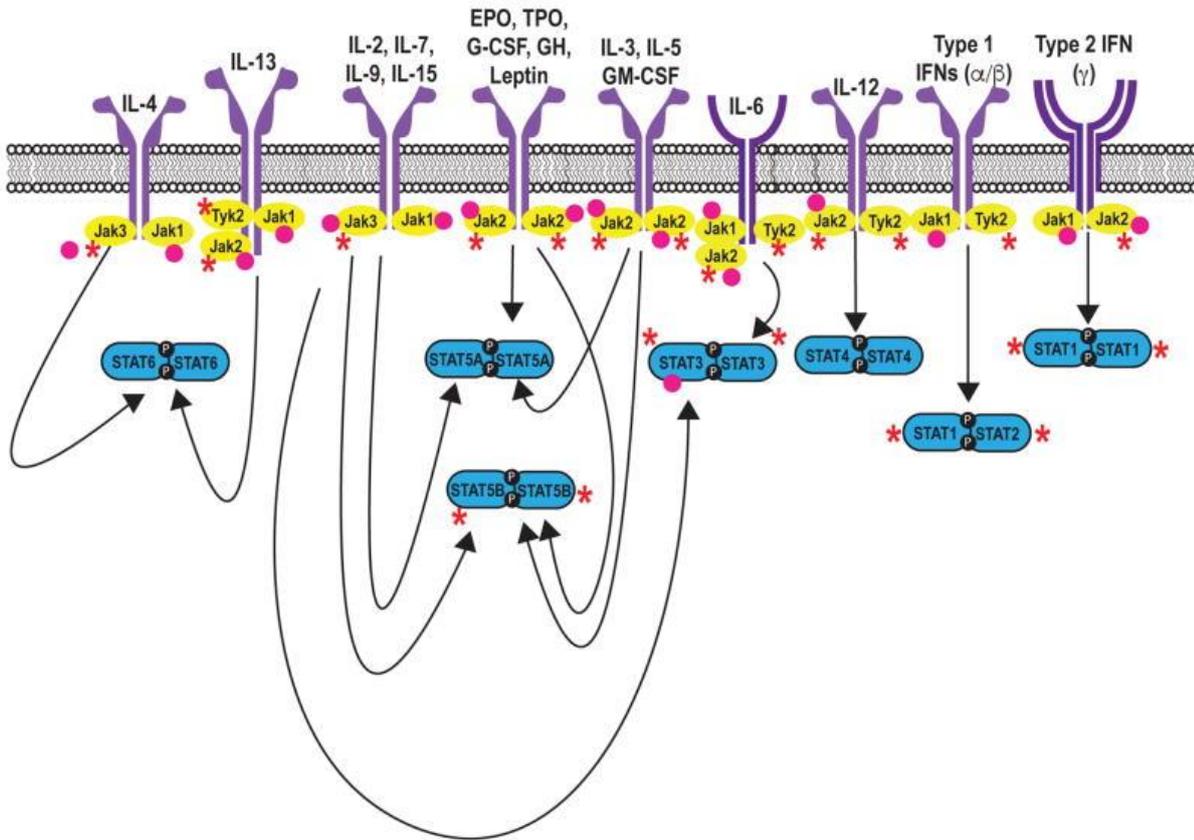


Figure 27. Liaison sélective des différentes protéines JAK à leurs récepteurs membranaires.

- *STAT*

Il existe sept facteurs de transcription STAT : STAT1 (avec deux isoformes : STAT1 α et STAT1 β), STAT2, STAT3 (avec deux isoformes : STAT3 α , STAT3 β), STAT4, STAT5a, STAT5b et STAT6 (dont l'influence est variable) (83). Les protéines STATs sont constituées de 750 à 900 acides aminés et comportent 6 domaines (**Figure 28**) (76) dont :

- un domaine superhélice (*coiled-coil domain*) participant à l'interaction de STAT avec ses protéines régulatrices ainsi que son export/import cytoplasme/noyau.
- un domaine de liaison à l'ADN (DBD (*DNA binding domain*)).
- un domaine SH2.
- un domaine transactivateur (TAD) (activation de la transcription). Il contient des sites de phosphorylation de type Tyrosine (Y) et Sérine (S), essentiels pour l'activation/dimérisation de STAT et le recrutement de coactivateurs de la transcription. Il est aussi impliqué dans la déphosphorylation des STATs.

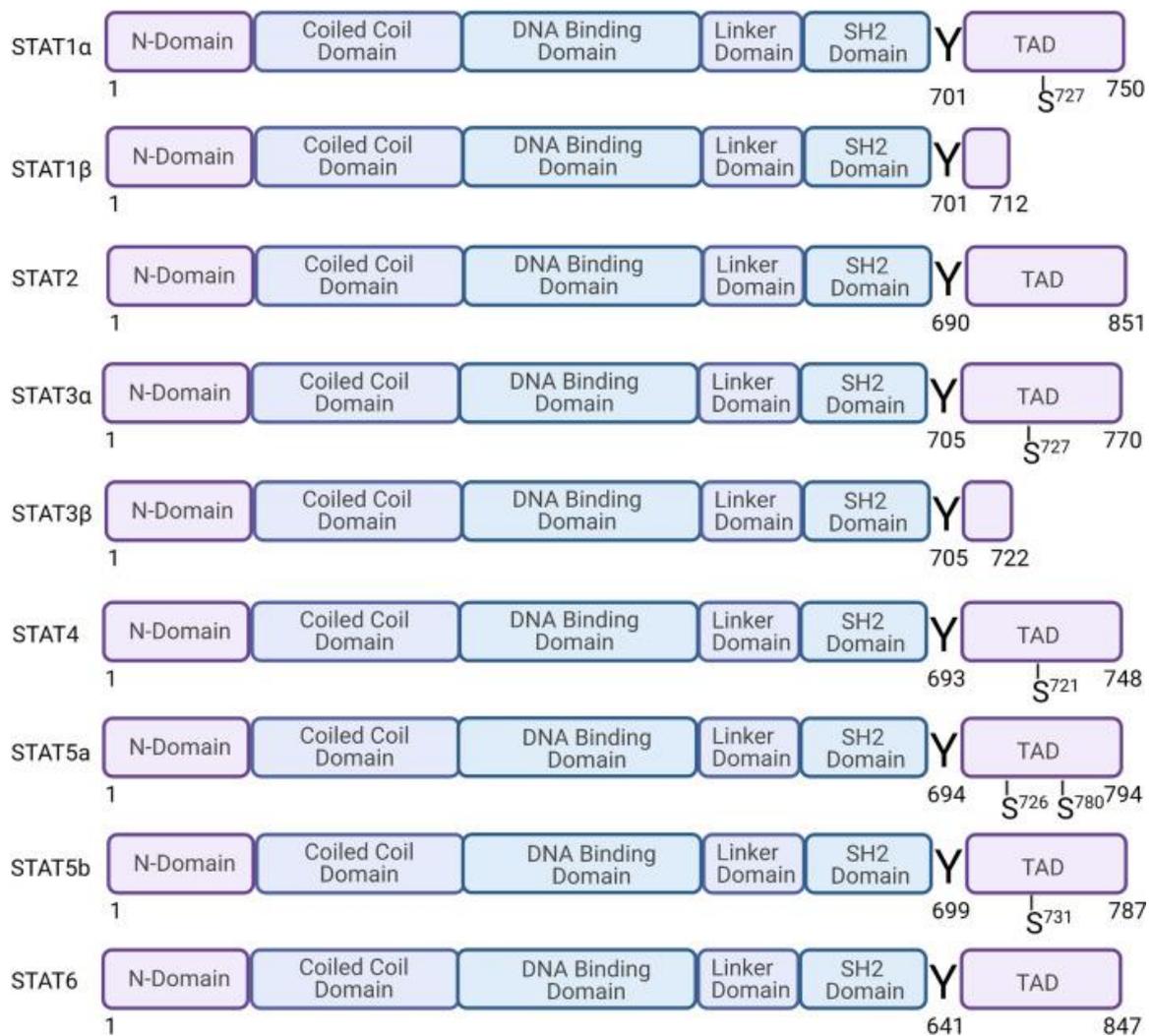


Figure 28. Structure et sites de phosphorylation des protéines de la famille STAT.

La rapidité d'action du signal médié par ces protéines leur a valu la dénomination par l'acronyme « STAT » (*Signal Transducer and Activator of Transcription*) (74). En effet, dérivé du terme latin *statim* signifiant « immédiatement », l'expression anglophone « STAT ORDER » désigne une situation clinique mettant en jeu le pronostic vital et nécessitant une réaction très rapide.

L'implication de *STAT3* et *STAT5* est particulièrement décrite dans l'oncogénèse de certains cancers. Leur hyperactivation est associée à des processus variés tels que l'inflammation, l'auto-immunité et les infections. Elle influence la prolifération et la survie cellulaire, ce qui explique l'implication des protéines STAT dans des mécanismes oncogéniques. *STAT5* (chromosome 17) peut être activée via de nombreux récepteurs cytokiniques et est notamment impliquée dans les phénomènes de renouvellement des CSH et de différenciation mégacaryocytaire (84).

Il a été démontré que STAT3 et STAT5 sont impliquées dans la régulation de l'expression génique via le modelage de la chromatine et sa méthylation (85). Lorsqu'elles sont phosphorylées et dimérisées, elles interagissent avec des enzymes telles que les DNMTs, les HATs, les HDACs ou même TET2 et EZH2 (**Figure 29**) (83).

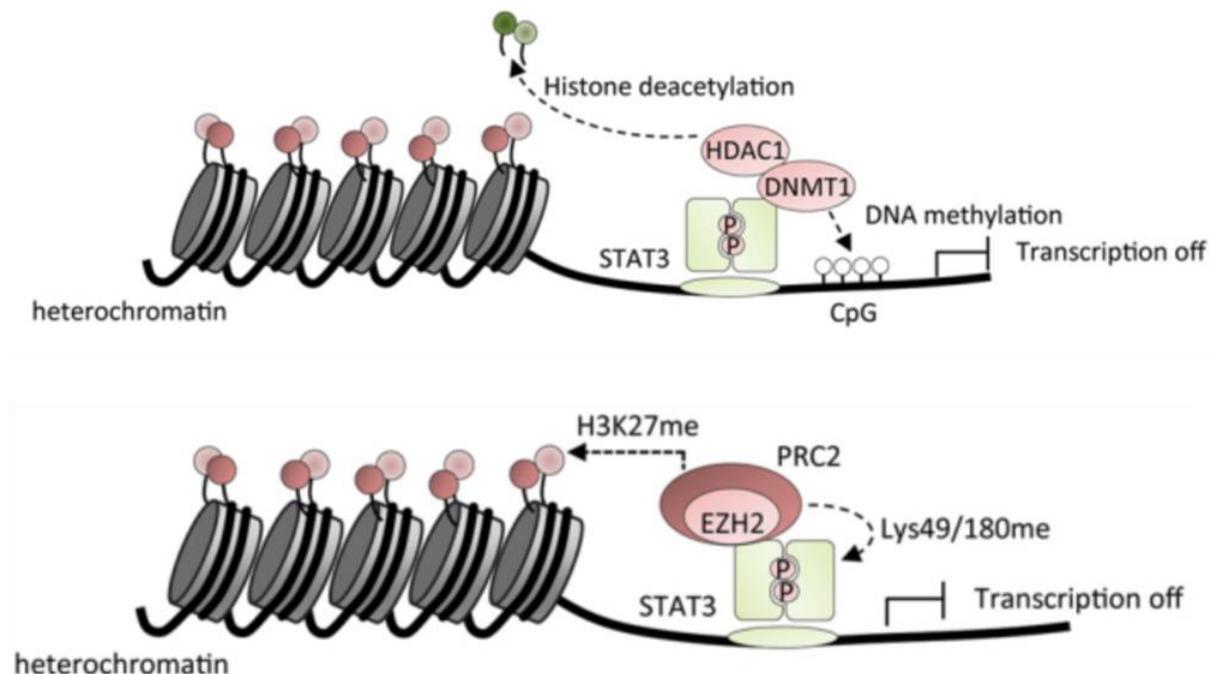


Figure 29. Exemples d'interactions de STAT3 avec les acteurs de la régulation épigénétique.

Il a également été démontré que les protéines STATs non phosphorylées n'étaient pas inactives. En effet, STAT5b non phosphorylée et STAT5b phosphorylée commandent des programmes transcriptionnels différents (86) impactant la mégacaryopoïèse et l'apoptose.

STAT3 et STAT5 peuvent chacune se lier aux mêmes régions oncogéniques régulatrices avec une action de redondance ou de signal antagoniste (83). La redondance est décrite dans l'érythropoïèse où la perte de STAT5 peut être compensée par STAT3.

c) *Régulation de la voie JAK-STAT*

La voie JAK-STAT est soumise à des phénomènes de régulation positive et négative (76). Concernant la régulation négative, trois mécanismes sont identifiés : (i) les protéines PIAS (*protein inhibitor of activated STAT*), (ii) les protéines CIS (*cytokine induced SH2 containing protein*) et SOCS (*suppressor of cytokine signaling*) et (iii) les protéines PTP (*protein tyrosine phosphatase*) (**Figure 30**) (76).

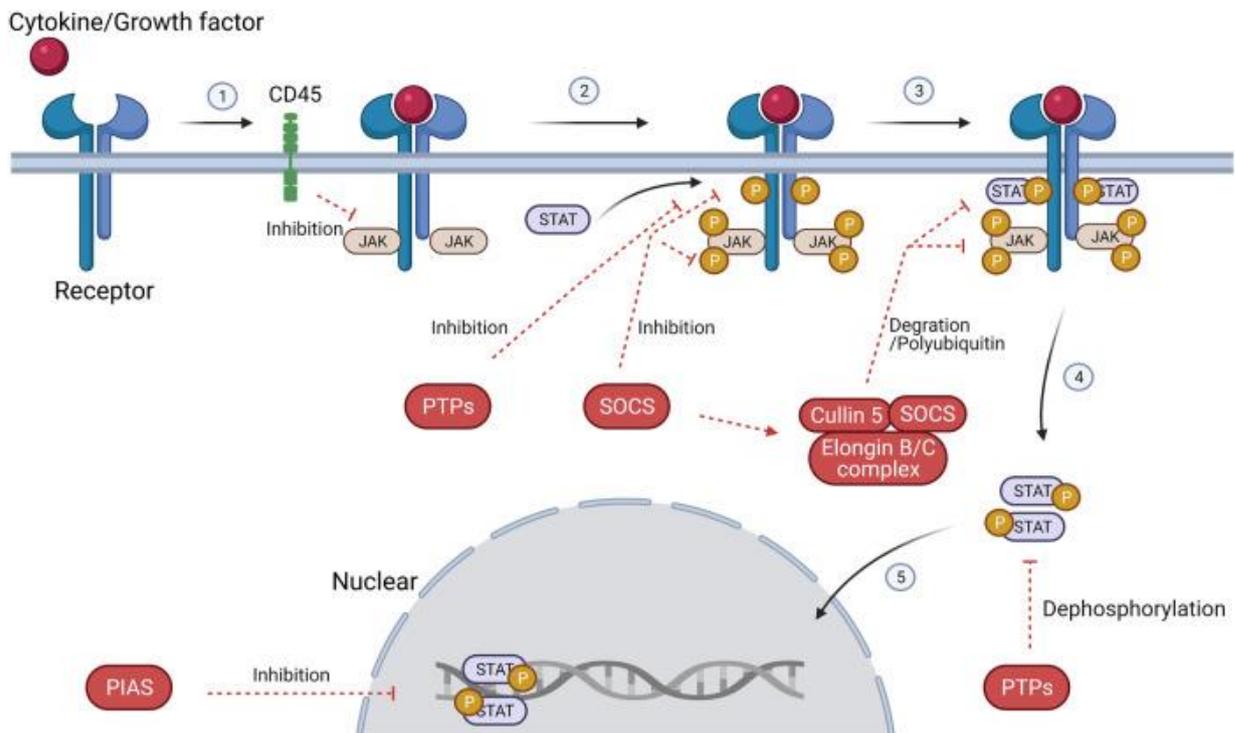


Figure 30. Régulation négative de la voie JAK-STAT.

- Protéines PIAS

Les protéines PIAS se lient aux protéines STATs activées (phosphorylées et dimérisées). Elles inhibent le DBD de STAT et donc les interactions avec l'ADN et l'activité transcriptionnelle. Elles facilitent également la formation de complexes répresseurs de la transcription.

- Protéines SOCS

Les protéines CIS/SOCS se fixent sur les protéines JAK2 phosphorylées et le récepteur cytokinique qui y est associé pour inhiber leur action. Le dimère phosphorylé de STAT5 favorise la transcription des gènes codant pour les protéines SOCS afin d'exercer un rétrocontrôle.

- Protéines PTP

Les PTP sont des protéines capables de retirer un groupement phosphate de leur substrat. Il en existe plus d'une centaine. Leur domaine SH cible des molécules de signalisation, leur récepteur ou directement JAK ou STAT. Certaines sont cytoplasmiques et d'autres membranaires, comme le CD45 qui inhibe la phosphorylation de JAK induite par l'interleukine 3.

Parmi les PTP, deux possèdent des domaines SH2 : PTPN6 (*protein tyrosine phosphatase non-receptor type 6*), aussi appelée Shp-1 (*SH2 containing protein tyrosine phosphatase 1*) et PTPN11 (*protein tyrosine phosphatase non receptor type 11*), aussi appelée Shp-2 (*SH2 containing protein tyrosine phosphatase 2*).

Shp-1 est particulièrement exprimée dans le tissu hématopoïétique, alors que Shp-2 est ubiquitaire (87). Shp-1 cible notamment STAT5b lorsqu'elle est activée par des hormones de croissance et transloquée dans le noyau.

Malgré leur homologie Shp-1 et Shp-2 ont des activités très différentes, voire opposées. La fonction principale de Shp-2 serait en effet plutôt activatrice des voies de signalisation par interaction avec de multiples cibles via son domaine SH2, dont JAK2 (88).

La régulation des STATs met également en jeu des phénomènes d'épissage alternatif qui aboutissent à des protéines tronquées, amputées de leur domaine transactivateur (89). La forme tronquée naturelle de STAT5B (STAT5B β) (718 acides aminés) est phosphorylable et possède un effet dominant négatif. En effet, elle garde la capacité de constituer des hétérodimères avec la forme entière STAT5B phosphorylée. Elle peut donc interagir avec les promoteurs de gènes cibles de STAT5B mais, étant dépourvue du domaine TAD, bloquer leur transcription (90).

Concernant la régulation positive, de nombreux co-activateurs peuvent renforcer le signal de la voie JAK-STAT, tels que le récepteur aux glucocorticoïdes, LNK (*Lymphocyte specific adaptor protein*), CBL (*Casitas B-lineage lymphoma proto-oncogene*) et SOCS3 (*Suppressor of cytokine signaling 3*).

d) *Interactions avec d'autres voies de signalisation*

La voie JAK-STAT interagit avec de nombreuses autres voies de signalisation, contribuant à l'équilibre de prolifération cellulaire (76). Certaines de ces voies sont particulièrement impliquées dans la prolifération cellulaire (**Figure 31**) (91), dont par exemple :

- La voie *RAS/mitogen-activated protein kinase* (MAPK), interagissant avec la voie JAK-STAT par le biais de Shp-2 (PTPN11) et GRB2.
- La voie NOTCH.
- La voie phosphatidylinositol 3-kinase (PI3K)/Akt via Shp-2.

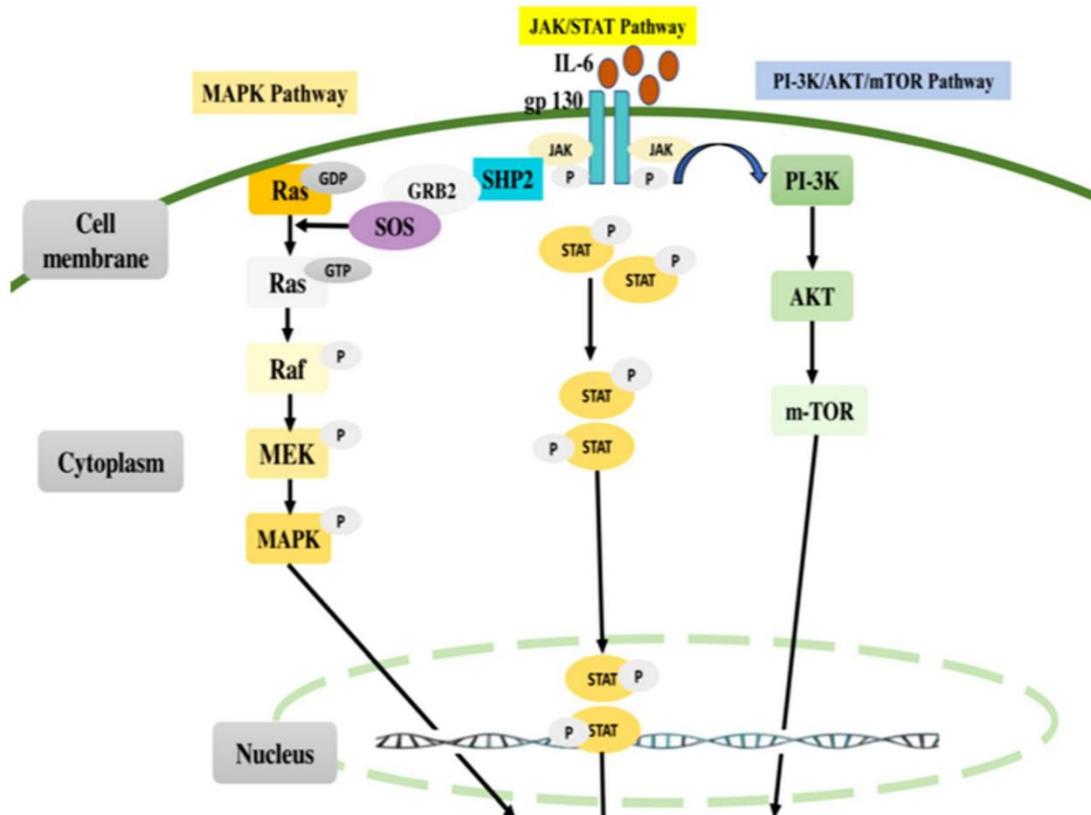


Figure 31. Interactions de la voie JAK-STAT avec les voies RAS/MAPK et PI3K/Akt.

e) *La voie JAK-STAT en pathologie cancéreuse*

La voie JAK-STAT constitue un carrefour pour de nombreuses voies de signalisation et divers mécanismes cellulaires (Figure 32) (83). Les atteintes de la voie JAK-STAT sont donc largement décrites dans de multiples pathologies cancéreuses et inflammatoires, qui varient selon la protéine JAK ou STAT affectée.

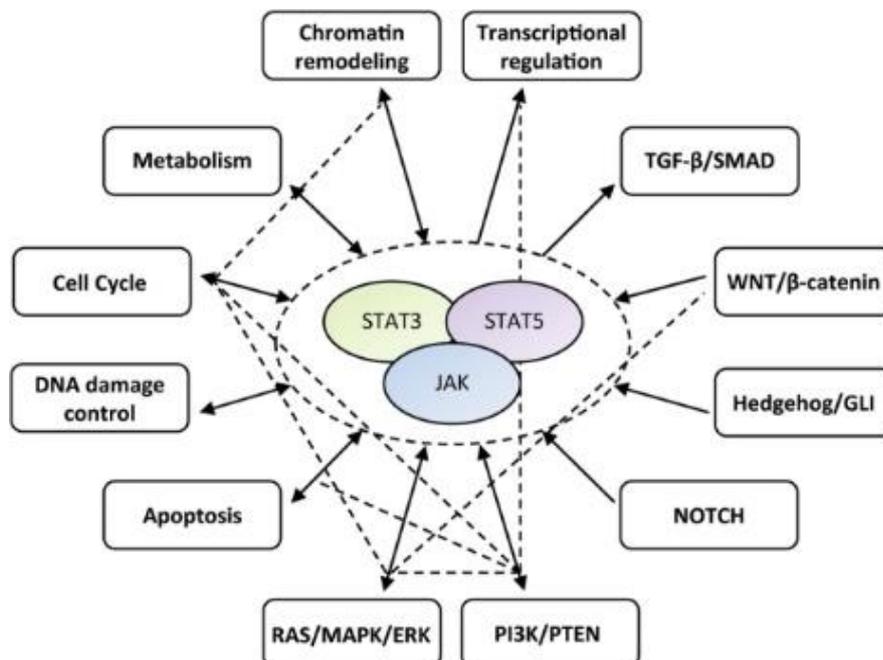


Figure 32. Interconnexions de la voie JAK-STAT et de divers mécanismes cellulaires.

En cancérologie, la voie de signalisation JAK-STAT peut être activée par mutations de *JAK* ou *STAT*, ou encore indirectement : (i) par des mutations affectant ses régulateurs comme LNK, CBL et SOCS3, ou (ii) par des mutations gain de fonction touchant des gènes de la voie RAS comme *NRAS* (*Neuroblastoma RAS Viral Oncogene Homolog*), *NF1* (*Neurofibromin 1*) et *Shp-2*.

- *Mutations de JAK2*

James *et al* ont décrit en 2005 la mutation récurrente *JAK2*^{V617F} associée aux MPN (63). Cette mutation conduit à la perte de fonction du domaine pseudokinase JH2 de JAK2. Il en résulte une perte d'inhibition et donc une activation constitutive de JAK2, même en l'absence de signal cytokinique. L'association de JAK2 aux récepteurs de l'EPO, la TPO, l'IL-3 et du GM-CSF explique la prolifération des différentes lignées myéloïdes et l'indépendance à ces cytokines dans les MPN porteurs d'une mutation *JAK2*^{V617F}.

La mutation *JAK2*^{V617F} a été initialement recherchée dans des échantillons de cellules myéloïdes et dans des échantillons de lymphocytes T issus des mêmes patients (63), orientant vers un caractère somatique acquis qui a par la suite été confirmé (92).

Pouvant s'intégrer dans le contexte d'une hématopoïèse clonale liée à l'âge, la détection de la mutation *JAK2*^{V617F} à elle seule, surtout à faible charge allélique, ne confirme pas le diagnostic de MPN. Elle est un argument fort mais non suffisant, comme le stipule les critères diagnostiques des MPN établis par l'OMS (2). Concernant la TE, la présence de la mutation *JAK2*^{V617F} impacte le pronostic (risque de thrombose) (93).

En 2007, de nouvelles mutations « gain de fonction » situées dans l'exon 12 de *JAK2* ont été décrites chez des patients atteints de PV en l'absence de la mutation *JAK2*^{V617F} (94). Les atteintes de l'exon 12 de *JAK2* concernent 3% des patients avec PV (95).

- *Mutations de STAT3 et STAT5*

Si la suractivation de STAT3 et STAT5 est le plus souvent secondaire à l'activation constitutive de tyrosines kinases telles que JAK2 ou FLT3 (*Fms Related Receptor Tyrosine Kinase 3*), elle peut également être directe, par mutation de l'un des deux gènes correspondants (85).

De manière générale, les mutations de *STAT3* et *STAT5* sont mutuellement exclusives dans les hémopathies malignes, mais fréquemment associées à des mutations d'acteurs de la régulation épigénétique comme les DNMT (83). En hématologie, les mutations de *STAT3* et de *STAT5b* ont majoritairement été rapportées dans des hémopathies lymphoïdes.

Les mutations de *STAT3* sont surtout décrites comme impactant l'inflammation et l'immunité, particulièrement au sein de l'environnement tumoral (96). Elles participent à la physiopathologie des lymphomes B diffus à grandes cellules et des proliférations de type LGL (*large granular lymphocytes*) (97). Par ailleurs, des populations clonales de lymphocytes T porteurs de mutations de *STAT3* sont retrouvées dans des hémopathies myéloïdes telles que l'aplasie médullaire idiopathique ou les MDS.

Les mutations de *STAT5b* impliquées dans la tumorigénèse sont pour la plupart des mutations « gain de fonction » affectant le domaine SH2, la tyrosine phosphorylable ou le domaine TAD (**Figure 33**) (98). La plus fréquente est la mutation N642H du domaine SH2 qui est décrite dans diverses hémopathies malignes de la lignée T.

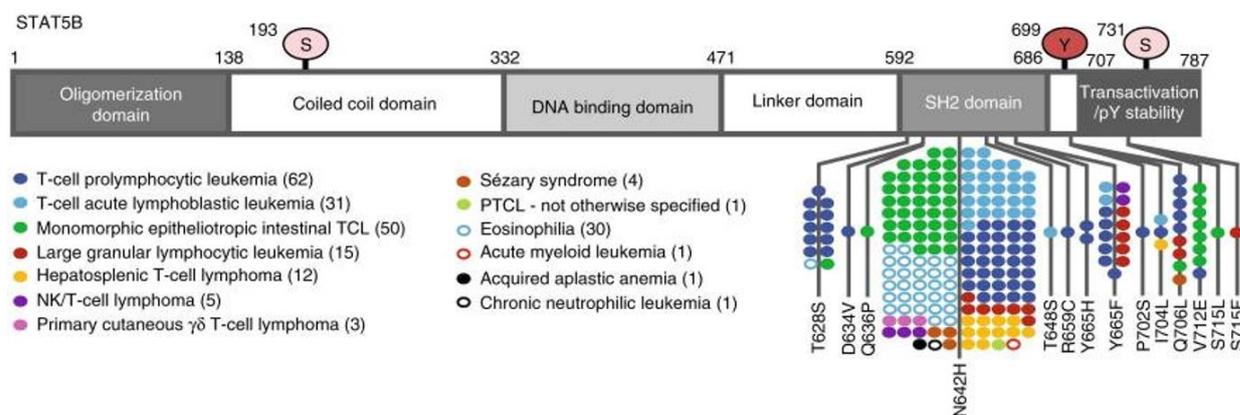


Figure 33. Mutations de STAT5b dans les hémopathies malignes.

3) La thrombopoïèse

a) Généralités

La thrombopoïèse correspond aux mécanismes aboutissant à la genèse et au relargage quotidien d'environ 100 milliards de plaquettes fonctionnelles dans la circulation sanguine (**Figure 34**) (99). Le précurseur médullaire engagé dans la différenciation plaquettaire est le mégacaryocyte (MK). Les MK matures sont générés lors du phénomène de mégacaryopoïèse, à partir de la CSH. Il existe deux types de mégacaryopoïèse (**Figure 34**) : (i) la différenciation de la CSH en progéniteurs multipotents engagés dans les lignées érythrocytaire et plaquettaire (MEP), eux-mêmes se différenciant en mégacaryocytes précurseurs (MKP) unipotents ; (ii) la différenciation biaisée et directe de la CSH en MKP (99). Ce type de mégacaryopoïèse permet de répondre rapidement à un besoin aigu en plaquettes.

Le MKP entre ensuite dans un processus d'endomitoses et de polyploïdisation pour former le MK mature (**Figure 34**). Celui-ci génère, à travers l'endothélium vasculaire, des extensions cytoplasmiques appelées pro-plaquettes, desquelles le flux sanguin arrache des « débris » qui constituent les plaquettes. Chez l'Homme, ces dernières étapes sont réalisées en environ 5 jours et les plaquettes formées ont une durée de vie de 7 à 10 jours. Toutes ces étapes sont sous la dépendance de multiples cytokines et FT et mettent en jeu des phénomènes de régulation tels que les ARN endogènes interférents (miRNA (miR)) et l'apoptose.

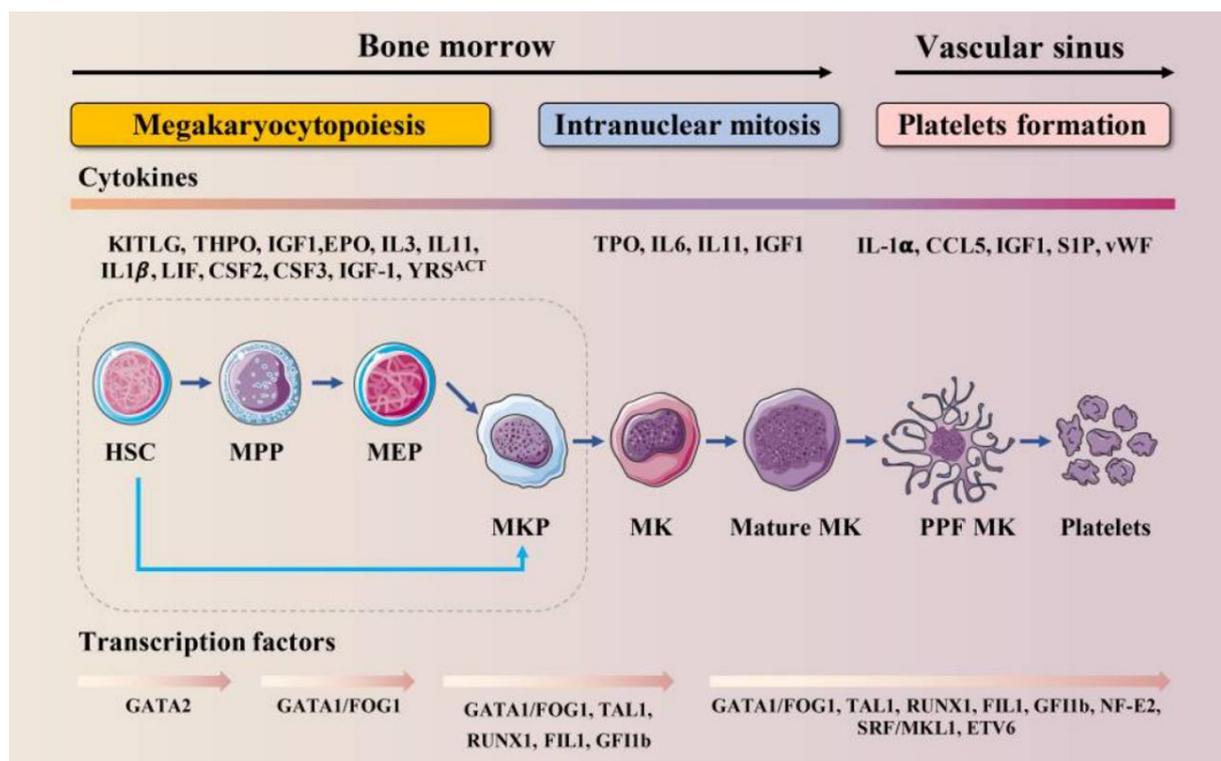


Figure 34. Schéma de la thrombopoïèse.

b) *Cytokines et signalisation*

De nombreuses cytokines et voies de signalisation sont impliquées dans le phénomène de thrombopoïèse (99,100). En premier lieu, la signalisation médiée par la TPO intervient *via* sa reconnaissance par le récepteur membranaire MPL, lui-même couplé à la protéine JAK2 (**Figure 25**).

L'autophosphorylation de JAK2 qui en résulte active notamment la voie JAK-STAT, mais aussi la voie PI3K (via Shp-2 et la voie NOTCH) et la voie des MAPK (**Figure 31**). L'équilibre d'activation des voies PI3K et MAPK est essentiel. En effet, la suractivation de la voie PI3K sans activation des MAPK conduit à la prolifération de MK immatures incapables de produire des pro-plaquettes. De même, l'hyperactivation pathologique des MAPK sans activation de la voie PI3K inhibe la formation des pro-plaquettes.

L'activation *via* la TPO permettrait par ailleurs de stimuler la fonction mitochondriale des CSH myéloïdes pour rompre leur quiescence et induire le biais de différenciation mégacaryocytaire (100). Certaines protéines mitochondriales ainsi stimulées, telles qu'AGK (*Acylglycerol kinase*) pourraient également amplifier la réponse à la TPO par liaison cytosolique avec JAK2.

La mégacaryopoïèse est également influencée par d'autres cytokines activatrices directes de JAK2 (**Figure 27**) telles que SCF (*Stem Cell factor*), IL3, IL5, IL6 (*interleukine-6*) (qui par ailleurs stimule la synthèse hépatique de la TPO), IL11 (*interleukine-11*) par voie autocrine, IL12 (*interleukine-12*), IL13 (*interleukine-13*), CCL5 (*chemokine (C-C motif) ligand 5*) et IGF-1 (*insulin like growth factor*) (100).

D'autres cytokines co-stimulent la voie TPO-JAK2 telles que l'IL1 α (*interleukine-1 α*), l'IL9 (*interleukine-9*) et FL (*FLT3 ligand*). Enfin, il existe une activation de la mégacaryopoïèse indépendante de JAK2 et de TPO, médiée par YRS (*tyrosyl-tNA synthetase*) via la voie du TLR (*Toll like receptor*) et MYD88 (*myeloid differentiation primary response 88*) (99).

Ce profil d'activation témoigne de l'influence de l'inflammation et du microenvironnement sur la mégacaryopoïèse.

c) *Facteurs de transcription*

Un réseau complexe de FT influence le devenir des CSH. La thrombopoïèse répond à un programme transcriptionnel spécifique, contrôlé par plusieurs de ces FT, dont notamment GATA-1 (*GATA-binding protein 1*), FOG-1 (*Friend of GATA-1*), FLI1 (*Friend leukemia virus integration 1*), RUNX1 (*Runt-related transcription factor 1*), MKL1 (*Megakaryoblastic leukemia 1*), TAL-1 (*T-cell acute lymphocytic leukemia 1*) et NF-E2 (*Nuclear factor erythroid 2*). L'expression génique dépendante de ces FT évolue tout au long de la différenciation mégacaryocytaire (**Figure 34**).

d) *ARN interférents (mi-RNA)*

Les ARN interférents endogènes (mi-RNA) sont décrits comme influençant de nombreux mécanismes de différenciation dont notamment l'hématopoïèse myéloïde (101). Cette influence existe également pour la thrombopoïèse. L'expression de *RUNX1* est par exemple négativement corrélée à celle de miR-9 dont il est une cible (100).

e) *Influence de l'apoptose*

La cinétique du relargage de plaquettes par les mégacaryocytes est influencée par l'apoptose de ces derniers. En effet, l'apoptose et la genèse des pro-plaquettes partagent des mécanismes morphologiques et architecturaux communs. Par des effets positifs ou négatifs, les phénomènes apoptotiques joueraient un rôle crucial dans l'équilibre de la numération plaquettaire (99) (**Figure 35**) (100).

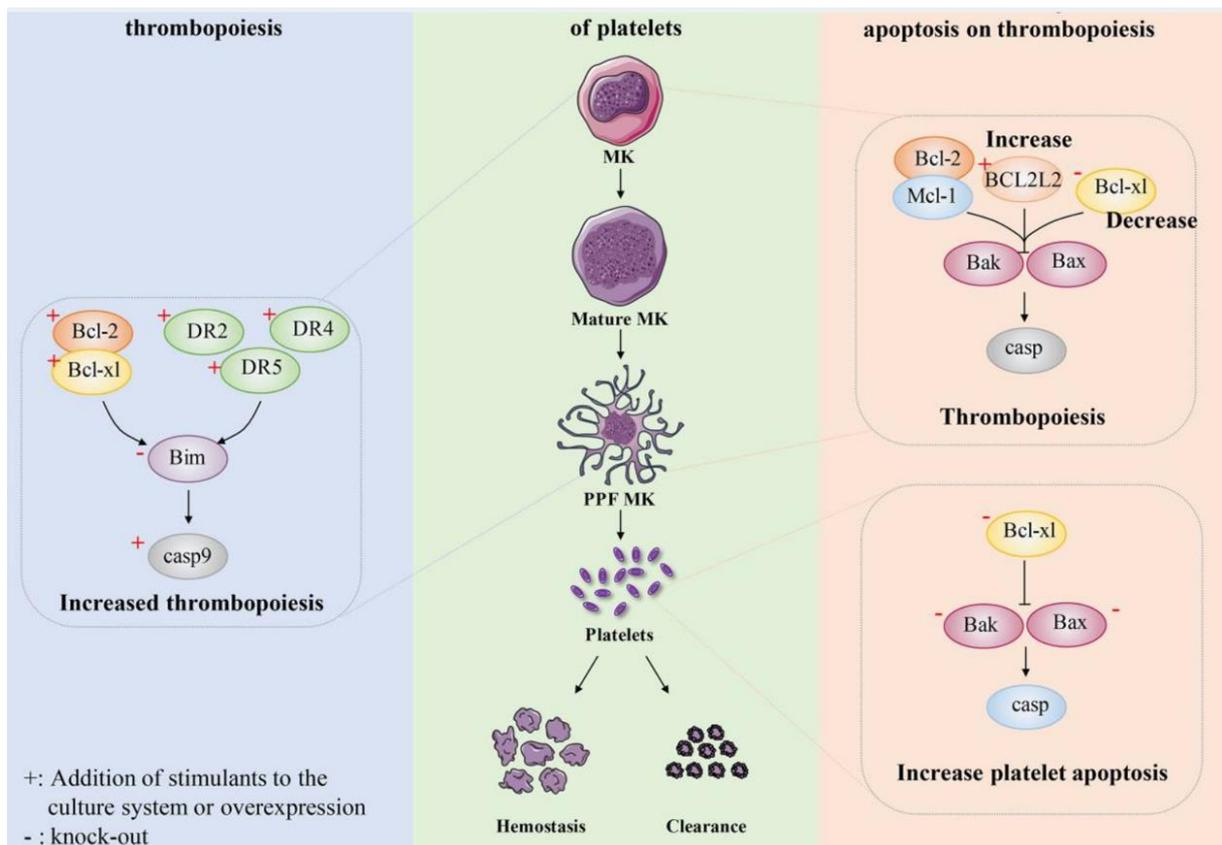


Figure 35. Rôle de l'apoptose dans la biologie des mégacaryocytes et des plaquettes.

4) Thrombocyémie Essentielle

La TE est une MPN caractérisée par un excès primitif chronique de production de plaquettes, lié à une expansion clonale mégacaryocytaire (2).

- Diagnostic

Le diagnostic repose sur des critères établis par l'OMS selon la classification de 2016 (non modifiés en 2022). Les critères majeurs sont : (i) une numération plaquettaire supérieure à 450 G/L, (ii) une biopsie ostéo-médullaire (BOM) montrant une prolifération principalement de la lignée mégacaryocytaire avec nombre augmenté de grands mégacaryocytes matures au noyau hyperlobé, l'absence d'augmentation significative ou de hiatus dans la lignée granulocytaire ou érythroblastique, et absence d'augmentation (ou très mineure = grade 1) des fibres de réticuline, (iii) l'absence de critère OMS évoquant une LMC *BCR-ABL*, une PV, une mastocytose systémique, un MDS ou une autre MPN et (iv) la présence d'une mutation de *JAK2*, *CALR* ou *MPL*. Le critère mineur correspond à la présence d'un autre marqueur clonal ou à l'absence d'argument en faveur d'une thrombocytose réactionnelle (inflammation, carence martiale, asplénie). Le diagnostic est confirmé en cas de présence des quatre critères majeurs ou de trois critères majeurs et du critère mineur.

- *Pronostic*

La survie médiane pour les patients atteints de TE est de 20 ans (tous âges confondus) et de 33 ans pour les patients de moins de 60 ans (102). La transformation en LAM est peu fréquente (**Figure 36**). Le pronostic est essentiellement lié au risque de thromboses veineuses et artérielles (103). Le risque de thrombose est par ailleurs moins important pour les TE avec *CALR* muté par rapport aux TE *JAK2* ou *MPL* mutés.

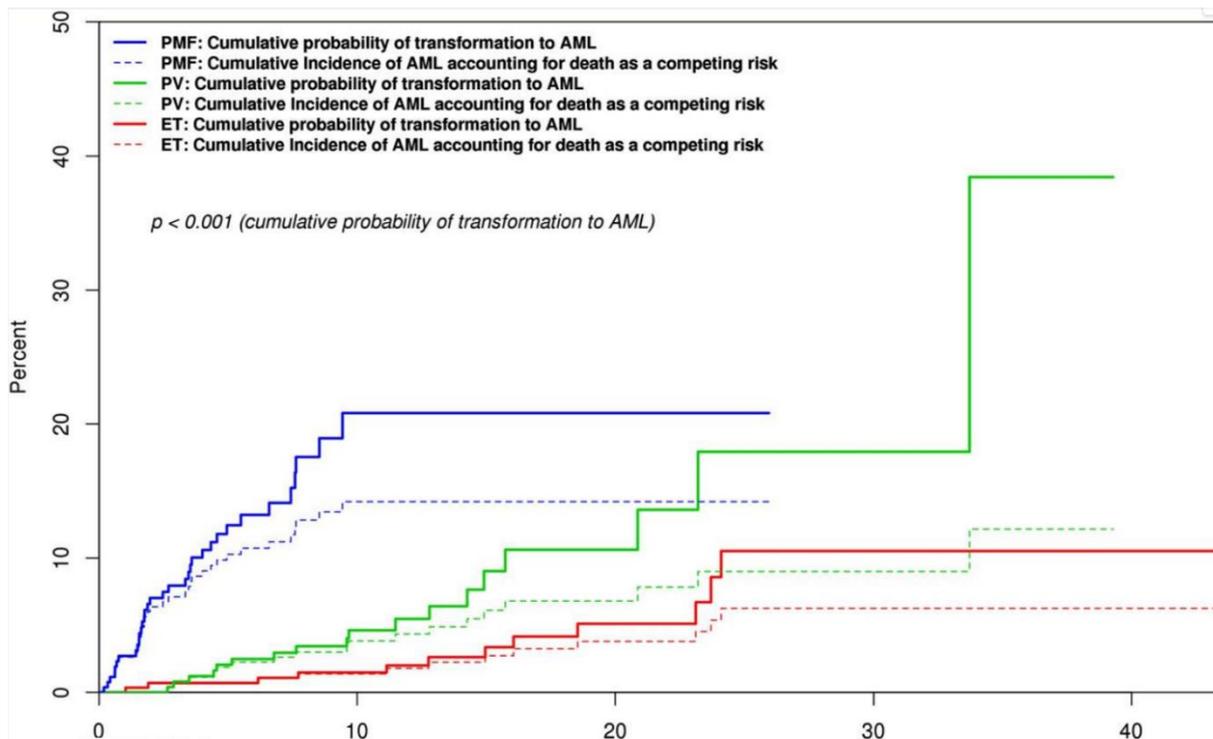


Figure 36. Incidence cumulée de transformation blastique pour les MPN non Phi.
PMF = myélofibrose primitive ; PV = *polycythaemia vera* ; ET = thrombocytémie essentielle ;
AML = leucémie aiguë myéloïde.

- *Traitement*

L'indication du traitement des TE repose sur l'établissement d'un score pronostique évaluant le risque de thrombose. Le risque est estimé élevé en cas d'âge supérieur à 60 ans et/ou d'antécédent de thrombose. La *MAYO clinic* (Rochester, Minnesota, USA) a proposé le score « IPSET révisé » (**Tableau 4**) qui prend également en compte la présence d'une mutation de *JAK2* et les facteurs de risque cardiovasculaires (93).

L'introduction d'un traitement cytoréducteur, visant à diminuer le chiffre plaquettaire en dessous de 450 G/L, suit les recommandations de l'ELN (*European Leukemia Net*). Ce traitement n'est pas recommandé pour les patients de risque faible ou intermédiaire si les facteurs de risque cardiovasculaires sont bien contrôlés. Les cytoréducteurs recommandés en première ligne sont l'hydroxyurée ou l'interféron α -pegylé (104).

Tableau 4. Score IPSET révisé pour la TE.

Facteur de risque	Point
Age supérieur à 60 ans	1
Facteurs de risque cardiovasculaires	1
Antécédent de thrombose	2
Présence de la mutation <i>JAK2</i> ^{V617F}	2

Facteurs de risque cardiovasculaires : tabac, diabète, hypertension artérielle, dyslipidémie.
IPSET < 2 : faible risque ; IPSET = 2 : risque intermédiaire ; IPSET ≥ 3 : risque élevé.

Tous les patients de risque élevé doivent également recevoir un traitement antiagrégant plaquettaire par acide acétylsalicylique, de même que les patients de risque faible ou intermédiaire en cas d'âge supérieur à 60 ans, de facteur de risque cardiovasculaire non contrôlé ou de présence de la mutation *JAK2*^{V617F} (104).

5) Méthylome des TE

Bien que des mutations affectant des régulateurs épigénétiques soient bien décrites, peu de données sont disponibles concernant la méthylation de l'ADN dans les MPNs, particulièrement pour les TE « triple négatives » (68,105).

L'analyse du méthylome de MPNs et de prélèvements issus de donneurs sains a montré une signature spécifique aux MPNs, affectant des régions régulatrices en lien avec l'inflammation et la voie NF-kB. mais non spécifiques d'un type de MPN (MF ou PV/TE) (106).

Une analyse multi-omique de patients atteints de TE, MF secondaire à une TE et de volontaires sains a inclus, à partir d'échantillons de cellules périphériques mononuclées, un méthylome sur puce et un séquençage complet de l'ARN (RNA-seq) (105).

L'analyse du méthylome n'a pas montré de différence significative entre les sous-types moléculaires de TE (mutées *JAK2*^{V617F}, *CALR* mutées ou TN) (105). Cependant, la comparaison de l'ensemble des TE avec les volontaires sains a permis de mettre en évidence des régions différenciellement méthylées (DMRs) et des gènes différenciellement exprimés (DEGs) (**Figure 37**) (105).

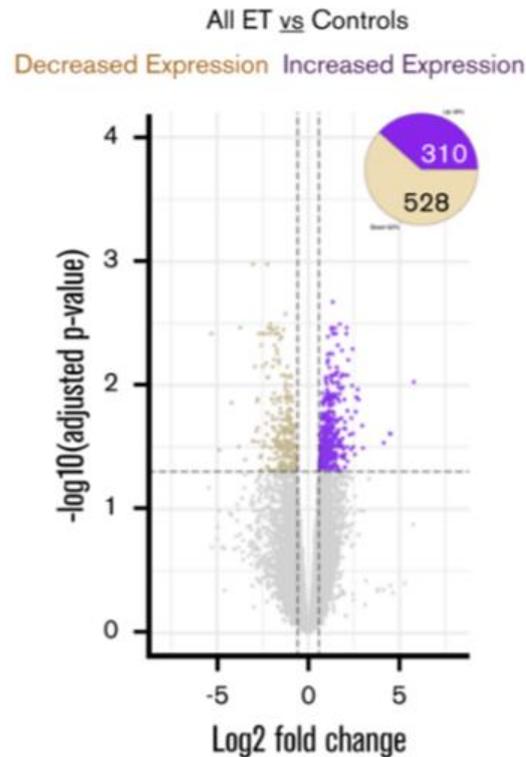


Figure 37. Différences significatives d'expression entre les volontaires sains et le groupe des TE.

Cette signature d'expression (**Figure 38**) (106), commune aux différents sous-types de TE, comprenait un panel de gènes surexprimés (comparativement aux donneurs sains) impliqués dans la différenciation myéloïde et la prolifération, l'oncogénèse et l'inflammation, notamment la voie des MAPK. La liste des gènes sous-exprimés était enrichie en gènes associés à des signatures transcriptionnelles associées à la thrombopénie.

Pour valider une vingtaine de gènes identifiés comme surexprimés, une cohorte de confirmation avec de nouveaux échantillons de patients TE TN et de volontaires sains a été constituée pour réaliser des qPCR (*Quantitative Polymerase Chain Reaction*). La surexpression relative de ces 20 gènes dans les TE a pu être confirmée.

Toutefois, les DMRs et les DEGs se recouvraient peu. Les auteurs ont alors émis l'hypothèse de DMRs reliées à des gènes codant des FT impliqués dans la régulation des gènes correspondant aux DEGs.

En effet, la liste des FT associés aux promoteurs des gènes surexprimés était enrichie en FT impliqués dans le développement des cellules myéloïdes, notamment via la voie des MAPK. Vingt-cinq de ces FT étaient par ailleurs surexprimés dans la cohorte des TE.

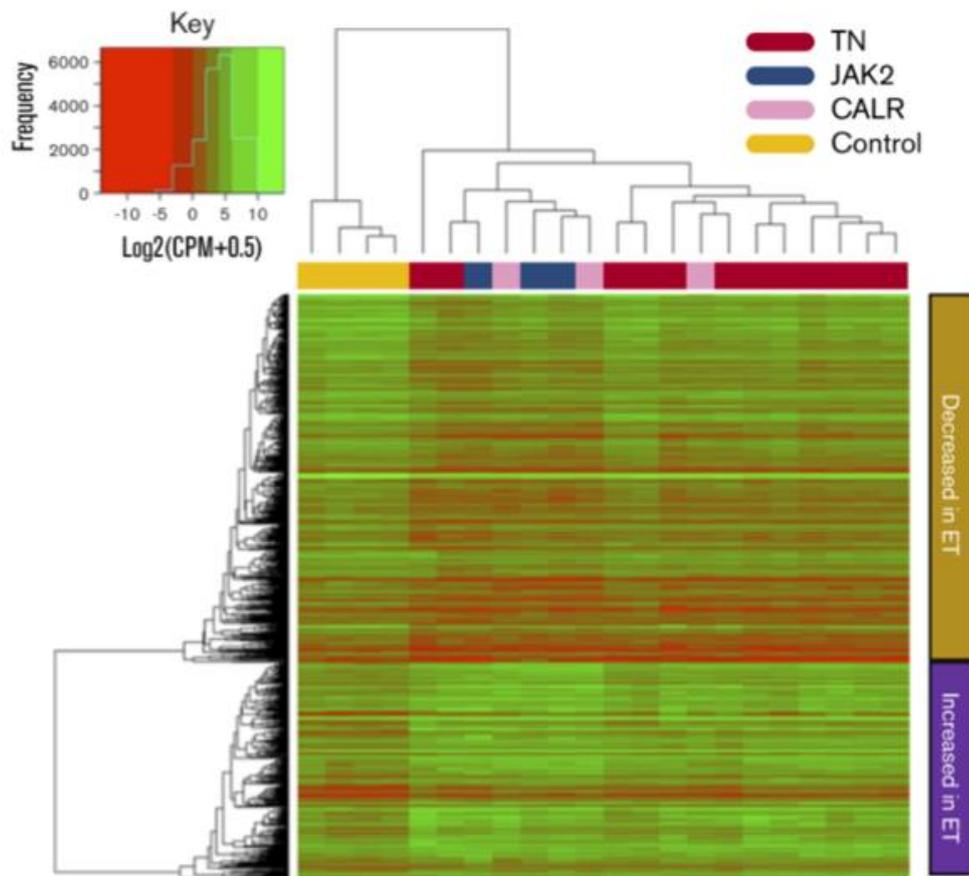


Figure 38. Classification hiérarchique des différents sous-types moléculaires de TE et des échantillons contrôles selon les profils d'expression génique.

Cette hypothèse a été testée *in vitro* sur la lignée cellulaire de TE « SET-2 ». Le *RNA-seq* réalisé sur ces cellules confirmait la surexpression de la plupart des DEGs, dont celles des FT associés aux voies NF- κ B et MAPK (en particulier *ITGA2B* (*Integrin Subunit Alpha 2b*) et *ITGB3* (*Integrin Subunit Beta 3*) de la famille JUN, par ailleurs décrits comme surexprimés dans certaines séries de MPNs (107)).

Les auteurs concluaient que différentes mutations *driver* pouvaient activer des voies moléculaires similaires qui convergeaient vers le même profil d'expression génique, aboutissant au phénotype de TE.

LES MDS/MPN-SF3B1-T

1) Découverte de l'entité frontière MDS/MPN-SF3B1-T et évolution de la nomenclature :

L'association entre MDS avec SC et thrombocytose est décrite depuis les années 1970 (108–112). En 1977, Streeter *et al* pointaient l'occurrence de thrombocytose pour quelques patients atteints de MDS avec SC (113), associée à une moindre fréquence d'évolution vers la LAM secondaire. L'observation d'une série de 39 cas de MDS avec SC (114) montrait 17 malades initialement porteurs d'une thrombocytose supérieure à 400 G/L, dont 5 supérieures à 800 G/L. Lors du suivi, 4 cas ont dépassé les 1 000 G/L, posant la question d'une transition vers une MPN avec un tableau proche de la TE.

Deux cas de diagnostics concomitant de TE et de MDS avec SC ont ensuite été publiés (115) avec au myélogramme la présence de plus de 15% de SC, mais également d'une hypercellularité avec des clusters de mégacaryocytes (MK). Les auteurs évoquaient pour la première fois le terme de « *Refractory Anaemia with Ring Sideroblasts (RARS) with Thrombocytosis* » ou RARS-T, devenu « Anémie réfractaire avec Sidéroblastes en couronne et Thrombocytose » (ARS-T) en français. Il lui a cependant été initialement préféré le terme de thrombocytémie essentielle avec sidéroblastes en couronne (ET-RS).

L'avancée des connaissances a mis en avant la nécessité d'une classification plus précise que la classification FAB (Franco-Américano-Britannique) de 1982. Dans ce but, les classifications successives de l'OMS définissent des entités pathologiques cliniquement significatives en prenant en compte les caractéristiques cliniques, morphologiques, cytochimiques, immunophénotypiques et génétiques. Elles se basent sur un consensus entre les différentes sociétés savantes dédiées à chaque hémopathie maligne. Plusieurs versions ont été proposées avec des actualisations en 2001 (2^{ème}), 2008 (3^{ème}), 2016 (4^{ème}) et 2022 (5^{ème}). La classification de 2001 (116) mentionne les ET-RS comme sous-entité provisoire des TE, à considérer comme un syndrome myélodysplasique/myéloprolifératif inclassable et avec pour caractéristiques : (i) moins de 5% de blastes, (ii) plus de 15% de SC et (iii) une thrombocytose persistante supérieure à 600 G/l. À ce stade, 3 hypothèses nosologiques sont envisagées pour les RARS-T : une entité distincte, un sous-groupe des RARS, ou l'occurrence simultanée de deux pathologies bien distinctes (117).

Schmitt-Graeff et al. (118) ont comparé les caractéristiques cliniques et morphologiques (BOM) de 39 patients ET-RS et ont déterminé 3 entités distinctes au sein du spectre des ET-RS : superposable à une TE, superposable à une myélofibrose primitive et superposable à une myélodysplasie. Ils ont conclu davantage à un spectre de pathologies présentant des SC et une thrombocytose qu'à une réelle entité mixte MDS/MPN.

Il apparaissait donc déjà que la difficulté de classification et de diagnostic des RARS-T réside dans la distinction entre thrombocytose primitive ou réactionnelle, associée aux SC (119). En effet, la plupart des syndromes mixtes MDS/MPN présentent une prolifération des lignées neutrophiles et/ou monocytaires et bien plus rarement une prolifération des lignées érythrocytaire ou mégacaryocytaire. De plus, la morphologie mégacaryocytaire apparaît plutôt normale dans les MDS avec SC comparativement aux RARS-T (pour lesquelles la morphologie variait sur un spectre allant de la MDS à la TE) (120).

La survenue concomitante de deux pathologies de type TE et MDS avec SC semble peu probable étant donné la faible prévalence de chacune dans la population générale. Il apparaît plus probable qu'un mécanisme indépendant prédispose l'une des pathologies à l'autre (121).

La classification révisée de 2008 (122) conserve l'entité provisoire des ET-RS sous le terme RARS-T en abaissant cependant le seuil de thrombocytose persistante à 450 G/L - incluant ainsi dans cette entité les RARS-MT (« *marked thrombocytosis* ») avec numération plaquettaire supérieure à 600G/L) et les RARS-nMT (« *not marked thrombocytosis* » avec plaquettes comprises entre 400 et 600G/L).

Le seuil plaquettaire utilisé pour définir la RARS-T a évolué au fil des classifications. Initialement placé à 600 G/L, il a ensuite été abaissé à 450 G/L dans la classification OMS de 2008 par similarité avec le seuil utilisé pour le diagnostic de TE. Cependant, il existe dans cette classification un chevauchement d'entité car une thrombocytose supérieure à 500 G/L est fréquente dans les MDS avec SC.

La classification révisée de 2016 (122) a entériné les RARS-T comme entité distincte parmi les syndromes mixtes MDS/MPN sous le terme de « MDS/MPN-SC-T » (MDS/MPN avec SC et thrombocytose). Les critères diagnostiques conservent le seuil de 450 G/L pour la numération plaquettaire. La présence de 15% de SC est toujours requise, même en présence d'une mutation de *SF3B1*. Cette classification a mis l'accent sur le fait qu'une cooccurrence de mutations de *JAK2* et *SF3B1* était un argument fort en faveur du diagnostic. Par ailleurs, dans cette classification, en présence d'une mutation de *SF3B1*, l'observation de 5% (et non 15%) de SC devient un critère diagnostique des MDS-SC.

Enfin, la classification actualisée de 2022 (2) insiste sur l'importance de la mutation de *SF3B1*. Elle modifie l'entité des MDS-SC en MDS-*SF3B1*, regroupant ces pathologies sous un ensemble caractérisé par la génétique. En effet, les MDS présentant une mutation de *SF3B1* regroupent plus de 90% des MDS-SC avec plus de 5% de SC. La présence de 15% de SC peut toutefois encore se substituer à la présence d'une mutation de *SF3B1*. En conséquence, les MDS/MPN-SC-T sont renommés « MDN/MPN-*SF3B1*-T », sans autre retentissement sur la description de la pathologie et ses critères diagnostiques (**Tableau 5**). L'âge médian pour cette présentation varie de 71 à 75 ans (62).

Tableau 5. Critères diagnostiques des MDS/MPN-*SF3B1*-T selon l'OMS 2022

Critères diagnostiques des MDS/MPN associé aux mutations de SF3B1 avec thrombocytose
Anémie associée à une dysplasie de la lignée érythrocytaire , avec ou sans dysplasie multilignée, <1% de blastes dans le sang et < 5% de blastes dans la moelle osseuse.
Mutation de <i>SF3B1</i>
Thrombocytose persistante ≥ 450 G/L
Absence de fusion <i>BCR-ABL1</i> , de réarrangement de <i>PDGFRA</i> , <i>PDGFRB</i> , <i>FGFR1</i> , <i>PCMI-JAK2</i> , t(3 ;3)(q21,q26), inv(3)(q21q26) ou del(5q)
Absence d'antécédent de MPN, MDS (sauf MDS- <i>SF3B1</i>) ou d'un autre type de MDS/MPN
- La présence de 15% de SC peut se substituer à la mutation de <i>SF3B1</i>.
- Dans un cas qui remplit par ailleurs les critères diagnostiques de MDS avec del(5q) : absence de basophilie absolue isolée ou minime ; basophiles généralement < 2% des leucocytes.

2) Signature moléculaire des MDS/MPN-SF3B1-T

Sur le plan caryotypique, la plupart des cas de MDS/MPN-SF3B1-T ne présentent pas d'anomalie (123–127). Cependant, même si leur prévalence diffère, ils partagent des traits moléculaires avec les MPN « classiques » comme les mutations de *JAK2*, *CALR* et *MPL* (**Figure 39**).

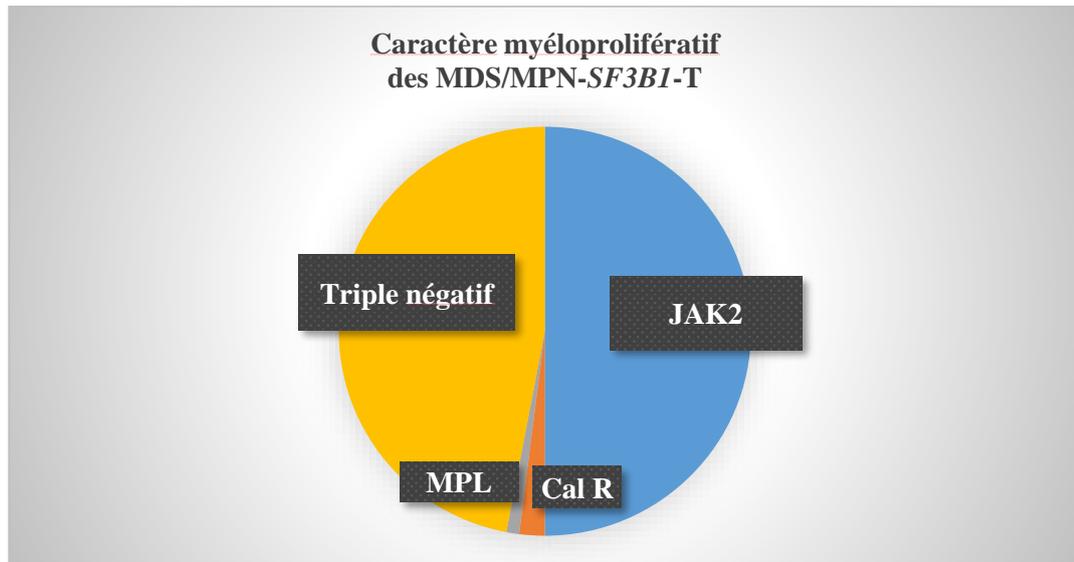


Figure 39. Répartition des principales mutations caractéristiques des MPN non-Phi dans les MDS/MPN-SF3B1-T.

a) *JAK2* dans les MDS/MPN-SF3B1-T

Suite à la découverte de la mutation *JAK2*^{V617F}, une série de 16 patients porteurs de MDS/MPN-SF3B1-T ont été séquencés pour *JAK2* (124). 50% de ces patients présentaient des critères OMS de TE et 31% étaient porteurs de la mutation *JAK2*^{V617F}. À la différence des TE isolées, il n'y avait cependant pas de pousse spontanée des progéniteurs mégacaryocytaires. Les séries de MDS/MPN-SF3B1-T identifiant la mutation *JAK2*^{V617F} se multiplient (117,120,125,126,128,129) avec jusqu'à 71%, voire 90% des cas de MDS/MPN-SF3B1-T *JAK2* mutés (126,128).

Bien que très commune dans les MPN classiques non Phi, la mutation de *JAK2* n'est que peu fréquemment retrouvée dans les MDS, notamment les MDS avec SC (121,125) et les autres MPN non Phi (121,130). En 2006, sur une série de MPN atypiques (notamment JMML, LMMC, LMC atypiques, RARS-T), 67% des patients RARS-T présentaient une mutation de *JAK2*. De manière intéressante, même en l'absence de mutation de *JAK2*, on retrouvait un profil aberrant de phosphorylation de *STAT5* (non retrouvé pour les autres MPN atypiques *JAK2* non mutés) (123).

Dans certaines séries, la mutation de *JAK2* est plus fréquemment retrouvée pour les MDS/MPN-*SF3B1*-T avec des plaquettes supérieures à 600 G/L (61%) que dans les cas avec plaquettes comprises entre 450 et 600 G/L (12,5%). Il n'y a cependant pas de différence en termes de survie ou de progression vers une myélofibrose (121).

En 2006, Wang et al. (120) ont comparé des MDS/MPN-*SF3B1*-T (avec plaquettes supérieures à 600 G/L) avec des MDS avec SC associés à une thrombocytose inférieure à 600 G/L, des MDS sans SC avec une thrombocytose supérieure à 400 G/L (deux groupes distincts si plaquettes supérieures ou inférieures à 600 G/L) et des MDS de type syndrome 5q-. Leur analyse a confirmé l'occurrence de la mutation *JAK2*^{V617F} dans 50% des cas de MDS/MPN-*SF3B1*-T (et son absence dans les autres groupes), mais elle a également montré une différence de survie globale avec les autres groupes, sauf pour les MDS avec SC. Au total, les MDS/MPN-*SF3B1*-T présentent des caractéristiques clinico-pathologiques des MDS et des MPN mais différent des MDS sans SC sur le plan clinique et moléculaire (vis-à-vis de la mutation de *JAK2*).

La recherche d'une mutation *JAK2*^{V617F} est un outil diagnostique majeur des MPN non-Phi et il apparait également très utile dans le diagnostic des MDS/MPN-*SF3B1*-T. Au travers des différentes publications, le seuil plaquettaire à retenir comme critère diagnostique de MDS/MPN-*SF3B1*-T est largement débattu (500 ou 600 G/L). Selon les séries, 10 à 50% des cas de MDS/MPN-*SF3B1*-T ne présentent donc pas la mutation *JAK2*^{V617F}.

b) *CALR* et *MPL* dans les MDS/MPN-*SF3B1*-T

Il est décrit une occurrence de mutation de *CALR* dans les MDS/MPN-*SF3B1*-T dans 0 à 3% des cas (65,131–134). Les mutations de *MPL* quant à elles sont détectées pour 1 à 2% des cas (65,131–135).

c) *Autres signatures moléculaires*

D'autres gènes sont mutés de manière récurrente dans les MDS/MPN-*SF3B1*-T : *TET2* dans 20% des cas, *DNMT3A* dans 15% des cas, *ASXL1* dans 14% des cas, *SETBP1* (*SET Binding Protein 1*) dans 10% des cas et *SH2B3* (*SH2B Adaptor Protein 3*) dans 5% des cas (132).

3) Pronostic

Dans la série de Wang et al de 2006 (120), la survie médiane des RARS-T était plus courte que celle des TE et des MDS 5q-, mais meilleure que celle des MDS avec SC et thrombocytose inférieure à 600 G/L (à l'époque, toutes étaient non considérées comme des MDS avec SC). Il n'y avait pas de différence avec les MDS avec SC. Broséus et al ont confirmé certains de ces résultats en 2012 sur une large cohorte. Ils ont cependant démontré que la médiane de survie des RARS-T était meilleure que celle des MDS avec SC et que cette survie n'était pas impactée par la présence de la mutation *JAK2*^{V617F}, ni par le seuil plaquettaire inférieur ou supérieur à 600 G/L. Le risque de progression vers une LAM secondaire était plus important pour les RARS-T que pour les TE (**Figure 40**) (136).

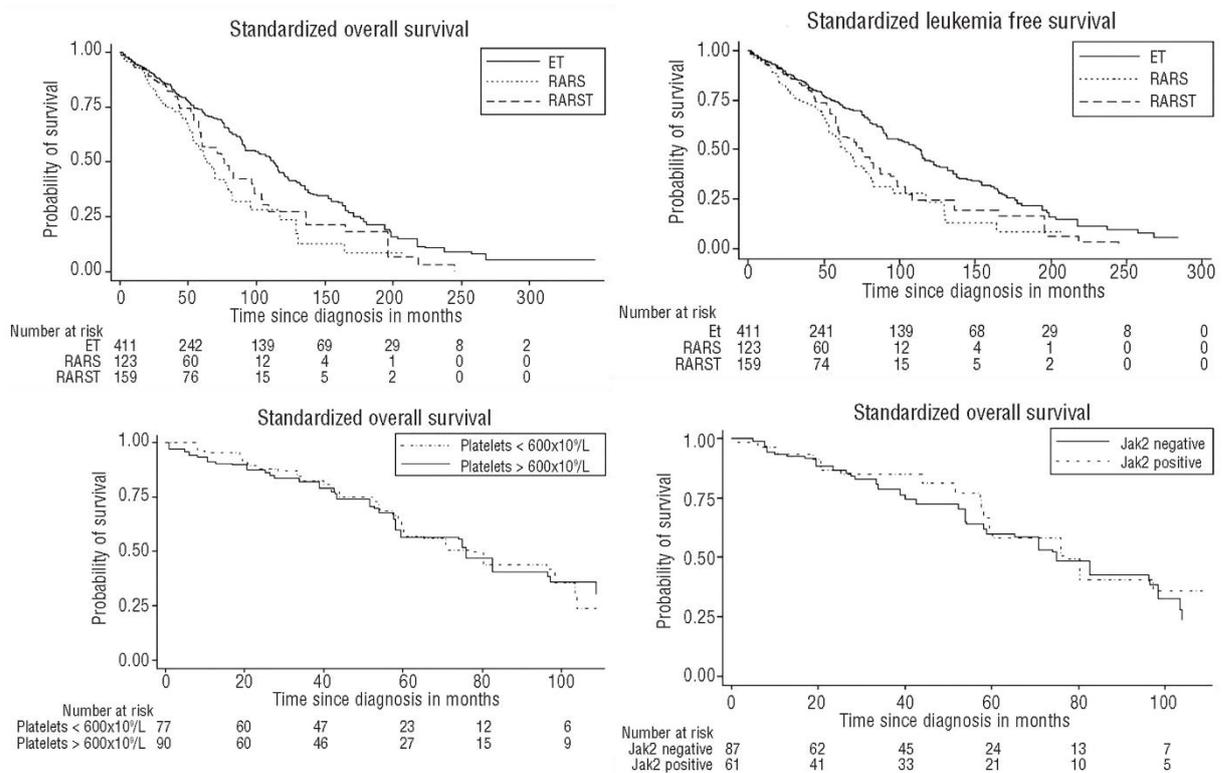


Figure 40. Survies globales et survie sans progression en LAM des RARS-T, MDS avec SC et TE.

La survie des MDS/MPN-*SF3B1*-T est impactée de manière indépendante par la présence d'une anémie ou d'un caryotype anormal. Par ailleurs, les mutations de *SETBP1* et *ASXL1* ont également un rôle pronostic négatif, sans impact supplémentaire du nombre de mutations retrouvées sur un panel myéloïde en séquençage haut débit (NGS) (137). Patnaik et al ont établi un score prenant en compte ces quatre éléments (anémie inférieure à 10 g/dl, caryotype anormal, statut mutationnel *SETBP1* et *ASXL1*) (**Tableau 6**) et qui permet de stratifier les patients en trois catégories de risque (**Figure 41**) (137).

Tableau 6. Score de stratification de risque pour les MDS/MPN-SF3B1-T.

Facteur de risque	Point
Caryotype anormal	2
Mutation de <i>SETBP1</i> et/ou <i>ASXL1</i>	1
Hémoglobine inférieure à 10 g/dl	1

Score = 0 : faible risque ; score = 1 : risque intermédiaire ; score ≥ 2 : risque élevé.

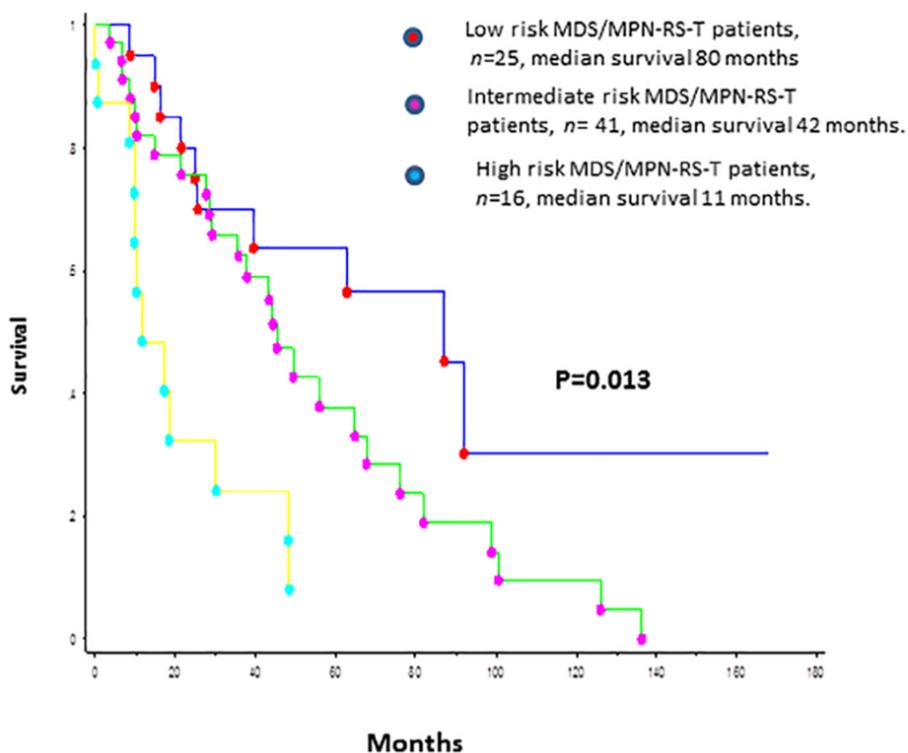


Figure 41. Survie globale selon la stratification de risque de Patnaik et al.

Le risque thrombotique dans les MDS/MPN-SF3B1-T est quant à lui similaire à celui des TE, soit d'environ 4% par an (62). Cependant, contrairement aux TE, l'antécédent de thrombose précédant le diagnostic n'impactait pas la survie globale. La présence de mutation de *SF3B1* semblait avoir un impact péjoratif sur la survie sans thrombose.

Il est à noter que les patients atteints de MDS/MPN-SF3B1-T présentent un risque accru de développer une maladie de Willebrand acquise (à caractère hémorragique), surtout en cas de thrombocytose supérieure à 1 000 G/l (62).

4) Traitement

La complication la plus fréquente des MDS/MPN-SF3B1-T est l'anémie, dont la prise en charge se superpose avec celle des MDS-SF3B1 avec utilisation des ASE et/ou du luspatercept (62,138). L'utilisation du lenalidomide a démontré une efficacité plus importante que celle des agents hypo-méthylants, sans différence significative par rapport à celle des ASE (139).

L'introduction d'une cytoréduction peut diminuer le nombre de plaquettes, mais elle peut aggraver l'anémie. Patnaik et al (62) recommandent une stratification du traitement en fonction des facteurs de risque du score IPSET : (i) les patients de plus de 60 ans ou avec antécédent de thrombose doivent bénéficier d'une cytoréduction associée à une antiagrégation plaquettaire, (ii) les autres patients doivent bénéficier d'une antiagrégation plaquettaire seule (cas particulier des patients de ce groupe ne présentant ni de mutation de *JAK2*, ni facteurs de risque cardiovasculaire qui peuvent bénéficier d'une simple surveillance).

La cytoréduction est classiquement effectuée par hydroxyurée. En cas d'intolérance ou de non-efficacité, il peut être discuté une deuxième ligne selon la présence d'une anémie associée. Les molécules recommandées en deuxième ligne sont le lenalidomide (138), l'interferon α , l'anagrelide et le busulfan (62).

En cas de saignement et ou de numération plaquettaire supérieure à 1 000 G/l, l'activité du facteur Willebrand doit être mesurée et une cytoréduction introduite si celle-ci est inférieure à 30% (62).

WHOLE EXOME SEQUENCING (WES)

Le séquençage complet de l'exome (*Whole Exome Sequencing - WES*) est une technique de NGS qui cible les séquences codantes (chez l'Homme, les exons représentent environ 2% du génome) (140). Il permet de détecter des anomalies de type (i) SNV (*Single Nucleotide Variant*) qui correspondent à des variations d'une seule base de la séquence et (ii) CNV (*Copy Number Variation*) qui correspondent à des duplications ou délétions de séquences. La préparation des bibliothèques repose sur une technique de capture des séquences exoniques (exons).

Le WES génère un nombre élevé de fragments de petite taille appelés « *reads* » (plusieurs millions par échantillon), qui sont séquencés en parallèles. Le chevauchement et l'alignement de ces *reads* sur un génome de référence permet de reconstruire l'exome initial et de l'y comparer. Il permet également de définir la couverture et la profondeur du séquençage (**Figure 42**) :

- l'étalement des *reads* sur le génome de référence (et les régions qui ont été séquencées) constitue la couverture.
- le nombre de *reads* couvrant une région donnée correspond à la profondeur.

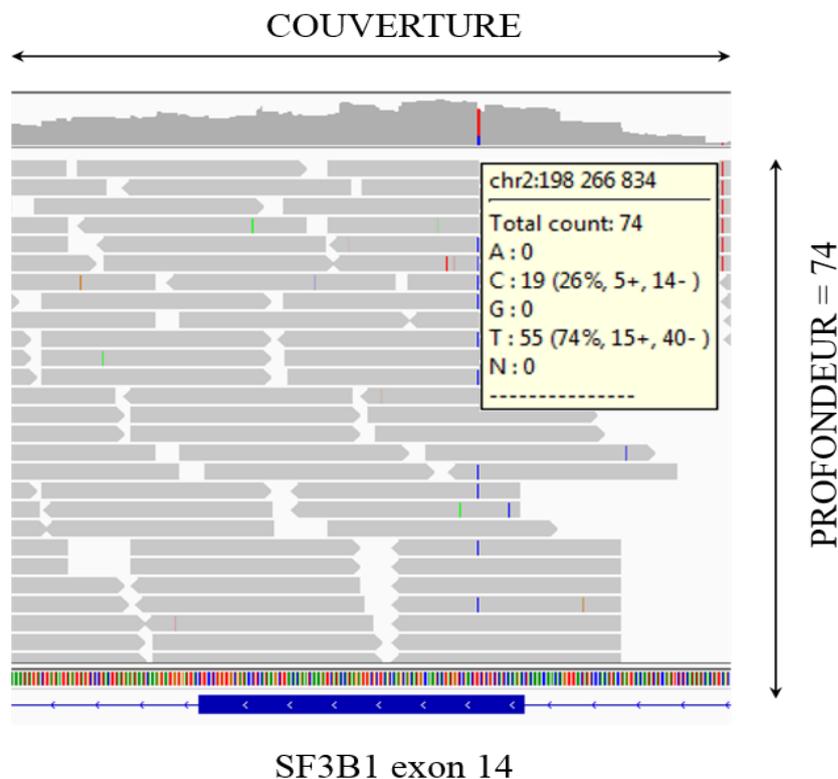


Figure 42. Représentation du logiciel IGV pour l'exon 14 du gène SF3B1
Mise en évidence d'un variant C>T. Pour cette position, la profondeur était de 74 *reads* et la profondeur allélique du variant C>T était de 55 *reads*.

ANALYSES BIOINFORMATIQUES DU NGS

1) Données brutes et contrôles qualité

Les données brutes du séquenceur sont classées par échantillon grâce à l'insertion préalable de code-barres spécifiques de chaque échantillon (démultiplexage) puis converties au format FASTQ. Le format FASTQ permet d'associer un score de qualité Q appelé *Phred* (échelle logarithmique) aux différents nucléotides séquencés. Il reflète la probabilité P d'erreur de séquençage, avec $Q = -10 \log P$. Plus le score est élevé, meilleure est la qualité.

2) Alignement

Les *reads* de bonne qualité sont alignés sur un génome de référence. Cette étape se base sur la concordance des *reads* avec le génome de référence. Un score d'alignement est ainsi déterminé pour chaque *read*. Toutes ces informations sont compilées dans des fichiers «.sam» et décrivent exhaustivement les "*hits*" des *reads* sur le génome de référence pour chaque échantillon. L'alignement est réalisé par des outils comme *Bowtie2* (141) ou *BWA* (142).

3) Détermination des variants

Les variants sont appelés à partir des empilements des *reads* alignés à l'aide des outils SAMTOOLS et BCFTOOLS. On génère ainsi des fichiers «.vcf» contenant des informations comme le génotype identifié, un score de qualité, différents biais de détection, ou encore la profondeur à la position donnée et la fréquence allélique (**Figure 43**).

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	Échantillon 03
chr2	198257299	.	G	GGC	999	PASS	MQ0F=0;MQ=44;MQSB=1;DP=339;DFGT:DP:SP:AC0/0:2:0:.,.,.,.,.,.		
chr2	198257795	.	C	T	238	PASS	Gene_refGene=SF3B1; MQSB=1;MQCGT:PL:DP:SP 1/1:172,18,0:6		

Figure 43. Exemple de fichier «.vcf».

Mise en évidence d'un variant C>T pour le gène *SF3B1*.

La qualité d'appel du variant dépend directement de la profondeur du séquençage. Dans les régions mal couvertes, il est impossible de différencier un variant somatique potentiel d'un génotype sauvage avec une bonne probabilité. Ceux-ci sont filtrés par la suite.

4) Annotation des variants :

Cette étape consiste à associer chaque variant à des données biologiques et cliniques. Le but est d'obtenir la description la plus exhaustive possible en fonction des informations disponibles (dans des bases de données comme par exemple *Ensembl*, *Refseq*, *ClinVar*, *dbSNP*). Cette description permet ensuite de répondre à des questions fonctionnelles (quel gène ? Quelle fonction biologique ?), des questions d'organisation du génome (localisation précise, déséquilibre de liaison ?), d'impact potentiel de la (des) mutation(s) sur les mécanismes moléculaires (perte/gain de fonction des protéines, site de régulation touché), ou encore cliniques (le variant a-t-il déjà été décrit ? Dans quelles maladies ?). L'annotation permet de mieux évaluer l'effet potentiel des variants identifiés.

SÉQUENCAGE DU MÉTHYLOME PAR RRBS

Le RRBS (*Reduced representation bisulfite sequencing*) est une technique de NGS permettant une analyse de l'état de méthylation des cytosines d'une sous-partie de l'ADN (143).

Le traitement de l'ADN par bisulfite permet, par sulfonation, de convertir les cytosines (C) de l'ADN en uraciles (U, qui deviendront des thymines (T) lors du séquençage). La présence d'un groupement méthyle sur le carbone 5 de la cytosine empêche cette conversion et les ^{5m}C restent donc des ^{5m}C.

La comparaison du séquençage d'un ADN bisulfité à un ADN de référence bisulfité d'une part et non bisulfité d'autre part permet donc d'identifier ses positions de méthylation. Il est ensuite possible de déterminer le pourcentage de reads méthylés pour une position donnée, en évaluant pour cette position le nombre de reads présentant une base C et ceux présentant une base T. La répartition de la méthylation sur le génome est appelée « méthylome ». L'analyse du méthylome peut être réalisée sur génome entier (WGBS, *Whole genome bisulfite sequencing*) ou sur une fraction du génome.

Le RRBS s'appuie sur l'utilisation de l'enzyme de conversion *MspI* qui reconnaît le motif palindromique CCGG et clive le brin d'ADN en amont du dinucléotide CG. Cette enzyme permet donc d'isoler des fragments d'ADN enrichis en CpG et donc en région promotrices méthylables. Cette technique réduit considérablement l'étendue à couvrir lors du séquençage pour se focaliser sur des régions méthylables. La quantité de données générées et le coût en sont donc réduits. Cette technique est cependant moins exhaustive qu'un WGBS (*Whole genome bisulfite sequencing*).

OBJECTIF DE L'ÉTUDE

L'objectif de mon travail est d'identifier les mécanismes moléculaires responsables de la thrombocytose dans les MDS/MPN-*SF3B1*-T triples négatifs.

PARTIE 1 : séquençage d'exome et validations fonctionnelles.

HYPOTHESE : Il existe une ou plusieurs mutations géniques inconnue(s) à ce jour, responsable(s) de la thrombocytose dans les MDS/MPN-*SF3B1*-T TN.

PARTIE 2 : méthylome et expression.

HYPOTHESE : Les désordres de la méthylation étant fréquents dans les MDS, il nous paraissait essentiel de ne pas occulter la possibilité d'un mécanisme épigénétique responsable de la prolifération dans les MDS/MPN-*SF3B1*-T TN. Il pourrait également en résulter un intérêt thérapeutique (utilisation potentielle des agents hypométhylants ou inhibiteurs ciblés d'une voie de signalisation).

MATÉRIEL ET MÉTHODES

Afin d'explorer la première partie de mon travail, nous avons réuni trois cohortes d'échantillons indépendantes de MDS/MPN-*SF3B1*-T TN (**Figure 44**) :

- Cohorte 1 : constituée de 15 échantillons de sang issus de 15 patients avec MDS/MPN-*SF3B1*-T TN. Cette première cohorte a permis la réalisation, à l'institut Gustave Roussy (Villejuif), d'un NGS ciblé sur un panel de 23 gènes fréquemment atteints dans les hémopathies myéloïdes.
- Cohorte 2 : correspond à une cohorte de screening de 6 échantillons de sang issus de 6 patients avec MDS/MPN-*SF3B1*-T TN (cohorte de screening) sur lesquels un tri cellulaire a permis de générer 12 échantillons en séparant une fraction enrichie en lymphocytes T et une fraction enrichie en polynucléaires neutrophiles (PNN) pour chacun des 6 patients. Cette seconde cohorte a permis la réalisation d'un séquençage complet d'exome.
- Cohorte 3 : correspond à une cohorte de validation de 42 échantillons d'ADN issus de 42 patients avec MDS/MPN-*SF3B1*-T TN. Cette troisième cohorte a permis la réalisation d'un NGS ciblé (*targeted genome sequencing*) sur les gènes candidats identifiés à la suite de l'analyse d'exome de la cohorte de screening.

Afin d'explorer la 2^{ème} partie de mon travail, nous avons travaillé sur une 4^{ème} cohorte d'échantillons issus de patients avec MDS/MPN-*SF3B1*-T TN (dont deux sont communs avec la deuxième cohorte), MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}, TE ou MDS-*SF3B1* ainsi que des échantillons de moelles normales.

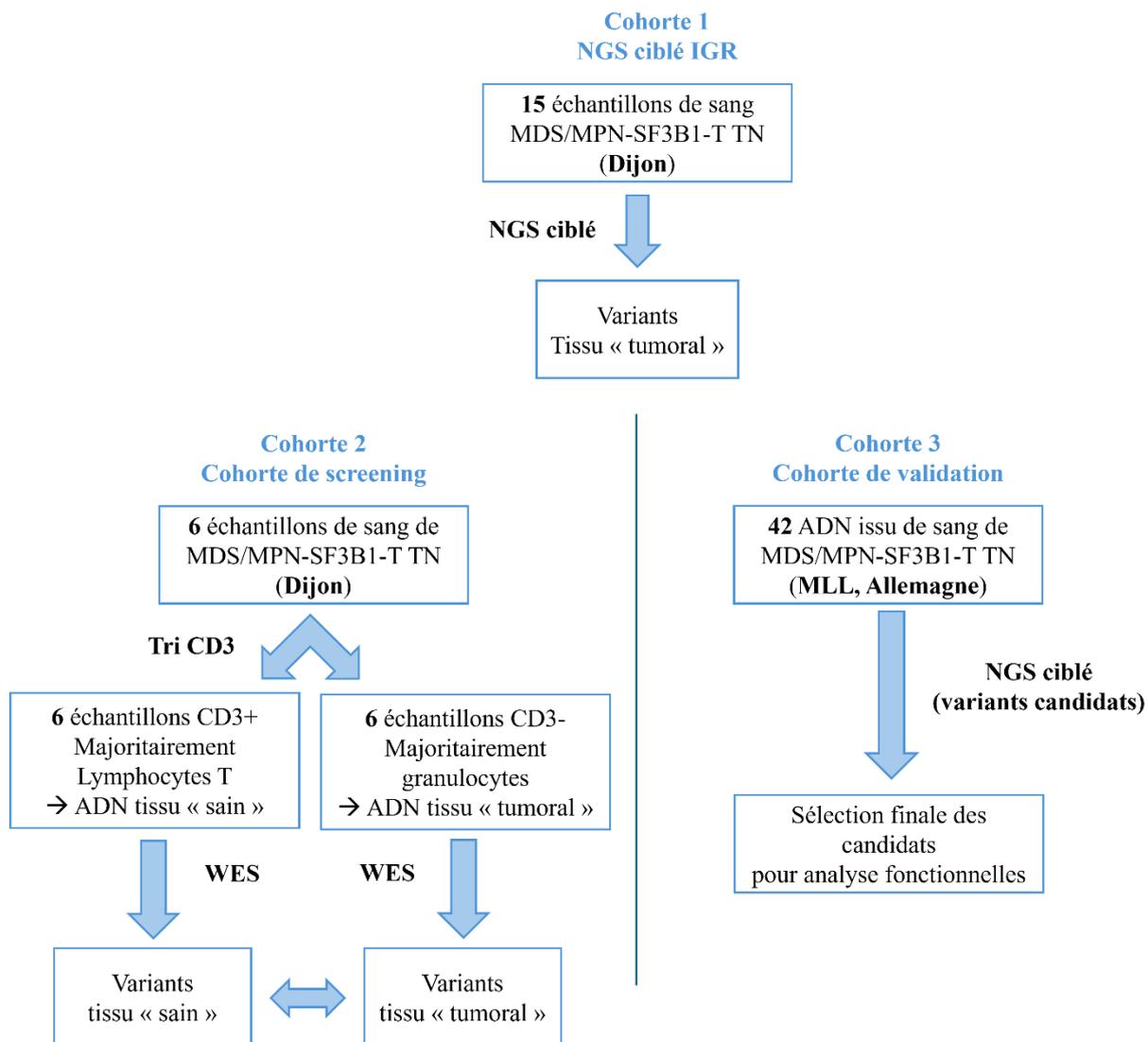


Figure 44. Cohortes d'échantillons utilisés pour l'approche génomique de la partie 1.

PARTIE 1 – PREMIÈRE COHORTE - PANEL IGR

Le panel ciblé réalisé à l'institut Gustave Roussy couvrait les gènes suivants, classiquement impliqués dans les hémopathies myéloïdes : *AKT1*, *ASXL1*, *CALR*, *CBL*, *CHEK2*, *DNMT3a*, *EZH2*, *IDH1*, *IDH2*, *JAK2*, *MPL*, *NFE2*, *PIK3CD*, *PIK3R2*, *PRSS3*, *SF3B1*, *SH2B3*, *SRSF2*, *TET2*, *TNK2*, *TNS1*, *TP53* et *U2AF1*.

PARTIE 1 - DEUXIÈME COHORTE - EXOME de 6 patients

1) Collection biologique de MDS/MPN-SF3B1-T, fractions CD3+ et CD3-.

Une cohorte de **screening** a été constituée à partir de **6 échantillons** de sang périphérique issus de patients suivis au centre hospitalo-universitaire (CHU) de Dijon. Les patients sélectionnés présentaient un MDS/MPN-SF3B1-T TN, selon les critères OMS 2022 (**Tableau 5. Critères diagnostiques des MDS/MPN-SF3B1-T selon l'OMS 2022**) (2). Les hémopathies ne devaient pas exprimer de mutation somatique de *JAK2*, *CALR* ou *MPL* pour être classées dans les « MDS/MPN-SF3B1-T TN ».

Dans le but de mieux cibler notre recherche de variants d'intérêt, nous avons considéré que :

- les variations d'intérêt sont **d'origine somatique**.
- les variations d'intérêt sont survenues **après différenciation myéloïde** lors de l'hématopoïèse. Elles sont donc absentes du tissu lymphoïde.
- tout comme les mutations connues de *JAK2*, *CALR* et *MPL*, les variations d'intérêt sont survenues dans des cellules progénitrices **en amont de la différenciation vers la lignée plaquettaire** et sont donc présentes dans les granulocytes.

Chaque échantillon de sang de la cohorte de *screening* a donc été divisé par FACS (*fluorescence activated cell sorting*) en deux fractions leucocytaires : une fraction exprimant le CD3 (marqueur lymphocytaire T) et une fraction soustraite n'exprimant pas le CD3, constituée principalement de granulocytes. La fraction CD3+ a été considérée comme « contrôle négatif » (génomique « germinale ») à comparer aux modifications moléculaires identifiées dans les granulocytes (génomique « tumorale »).

Une cohorte indépendante de **validation** a été constituée à partir de **42 échantillons** de MDS/MPN-SF3B1-T TN au laboratoire MLL (*Münchner Leukemie Labor*, Munich, Allemagne).

Dans la mesure où des variants somatiques qui expliqueraient la thrombocytose dans les MDS/MPN-SF3B1-T TN pourraient également être impliqués dans les MPN classiques triples négatifs, nous avons testé l'un de nos candidats sur une **cohorte additionnelle** de **12 TE TN et de 6 MF TN**.

2) *Whole exome sequencing* (WES)

Nous avons choisi cette méthode d'exploration afin de réduire l'étendue du séquençage et donc d'obtenir une profondeur de lecture suffisante pour détecter les variants génotypiques. En comparaison d'un séquençage du génome complet, le WES permet un enrichissement des impacts potentiels, tout en étant à moindre coût.

La préparation des bibliothèques a été réalisée en vue d'un séquençage selon la méthode *Illumina*[®] avec le kit *Agilent SureSelect*^{QXT} *target enrichment* (*Agilent, technologies*, Santa Clara, CA, USA), fondé sur le principe de la capture des séquences cibles à l'aide d'un pool d'ARN complémentaires de ces séquences (sondes ARN biotinylées).

La préparation des échantillons selon *Illumina*[®] a compris plusieurs étapes (**Figure 45**) : (i) **tagmentation** (fragmentation de l'ADN et ligation d'adaptateurs en 3' et 5'), (ii) **purification**, (iii) **amplification**, (iv) **hybridation** (sondes de capture), et enfin (v) **capture** de ces fragments par billes porteuses de Streptavidine. Les adaptateurs et les indexes incorporés dans les bibliothèques ont permis respectivement l'amarrage à la surface de la *Flowcell* et l'identification spécifique de chaque échantillon avant et après multiplexage.

L'amplification en clusters a été réalisée sur support solide *Flowcell* par le système *cbot* (*Illumina*[®]). La réplication en ponts (**Figure 46**) génère des clusters comprenant environ 1 000 fragments identiques, avec une densité cible de 10^6 clusters/mm²(144).

La réaction de séquence a été réalisée par détection de fluorescence avec la technologie *Reverse Dye Terminator*, sur un séquenceur HiSeq1000 (*Illumina*[®], San Diego, CA, USA) situé dans l'Unité Mixte de Service (UMS) Ingénierie, Biologie et Santé en Lorraine (IBSLor, UMS2008, CNRS-UL-INSERM). Les *reads* obtenus ont été alignés sur le génome de référence *Hg19* (GRCH37 dans sa version *Ensembl* (145)). Nous avons choisi pour cette expérience un séquençage *paired-end* (*reads* appariés R1 et R2 sur brins d'ADN complémentaires et séparés par un insert d'environ 100 nucléotides).

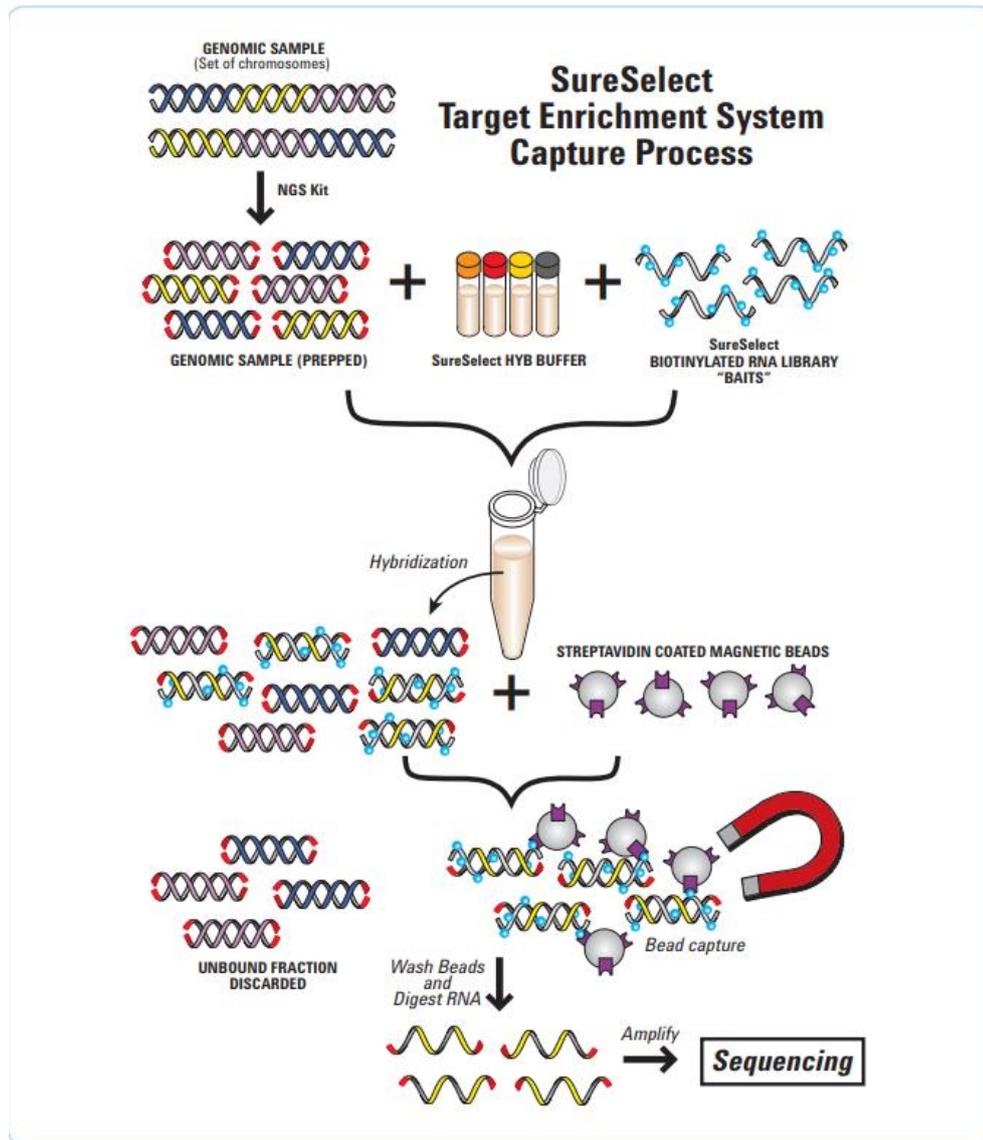


Figure 45. Workflow du kit Agilent SureSelect^{QXT} target enrichment.

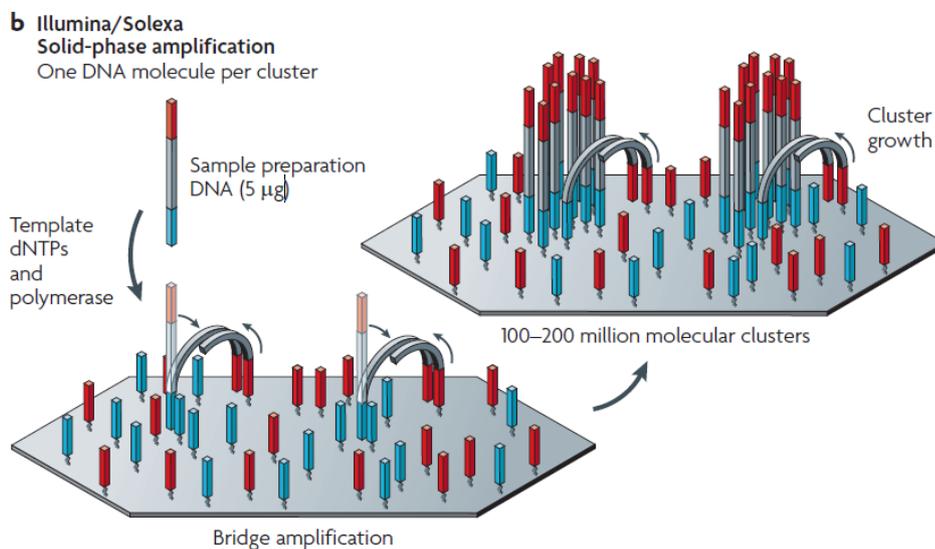


Figure 46. Génération des clusters par réplication en pont selon le système *cBot*.

3) Analyses bioinformatiques

Les analyses bioinformatiques ont été effectuées au sein de l'équipe MIGB (*Metabolism, Integrative Genomics & Bioinformatics*) de l'Unité INSERM N-GERE - UMR_S U1256.

Le pipeline d'analyse utilisé et présenté ci-après (**Figure 47**) constitue un enchaînement d'étapes clés, faisant fonctionner différents logiciels, afin d'obtenir une série de résultats exploitables à partir des données brutes du séquençage.

Nous avons individualisé trois types d'analyse dans notre pipeline :

- les « **analyses primaires** » : correspondant aux éléments de contrôle qualité et d'alignement.
- les « **analyses secondaires** » : correspondant aux éléments d'appel, d'identification et d'annotation des variants.
- les « **analyses tertiaires** » : comprenant notamment l'application de filtres de sélection ou de suppression des variants trouvés.

Les annotations ont été réalisées à partir des bases de données *Ensembl*, *Refseq*, *ClinVar* et *dbSNP*, principalement grâce à l'outil VEP (*Variant Effect Predictor*) (146).

La sélection des variants somatiques candidats a été permise par comparaison des fractions à prédominance granulocytaire (CD3-) et lymphocytaire (CD3+).

Parmi les variants identifiés lors des premières analyses (primaires et secondaires), la sélection de gènes candidats significatifs et pertinents a nécessité l'application de deux protocoles distincts (analyses tertiaires).

Tout d'abord, la sélection de variants somatiques d'intérêt s'est faite par l'application de filtres reposant sur des paramètres intrinsèques aux analyses primaires et secondaires. Nous les avons appelés filtres « paramétriques » :

- Suppression des variants avec un ou plusieurs génotype(s) manquant(s) (information manquante pour un ou plusieurs échantillons).
- Sélection des variants avec un score de **qualité de l'appel supérieur à 30** (type *Phred*).
- Sélection des variants correspondants à un « **gain somatique** » pour au moins 1 des 6 patients (mutation identifiée sur les CD3- mais n'existant pas sur les CD3+).
- Exclusion des variants se situant hors des **régions exoniques, zones d'épissage ou UTRs (*untranslated regions*)**.
- **Exclusion des mutations synonymes** (en première instance).

Un second protocole de sélection des variants a ensuite été appliqué, celui-ci s'appuyant sur leurs annotations et sur l'application de tests statistiques (basés notamment sur leur distribution), que nous avons appelés filtres « sémantiques » :

- Exclusion des variants rapportés comme « bénins » ou « tolérés » dans les bases de données d'impacts *SIFT* (*Craig Venter Institute*) et *Polyphen* (*Harvard University*).
- Exclusion des variants avec une **MAF** (*Minor Allele Frequency*) > 1% et conservation des variants avec une fréquence allélique < 1% dans *EXAC* (*EXome Aggregation Consortium*).
- Recherche des mutations statistiquement significatives par rapport au bruit de fond observé : utilisation de l'outil *MuSIC* (147) et correction des probabilités par FDR (*False Discovery Rate*) suivant la procédure de Benjamini-Hochberg.
- Calcul de la fréquence de gain somatique (nombre d'allèles mutés dans les CD3- par rapport aux CD3+) ramenée au nombre de patients touchés : obtention d'un **taux de mutation corrigé**, que l'on sélectionne > 20%.

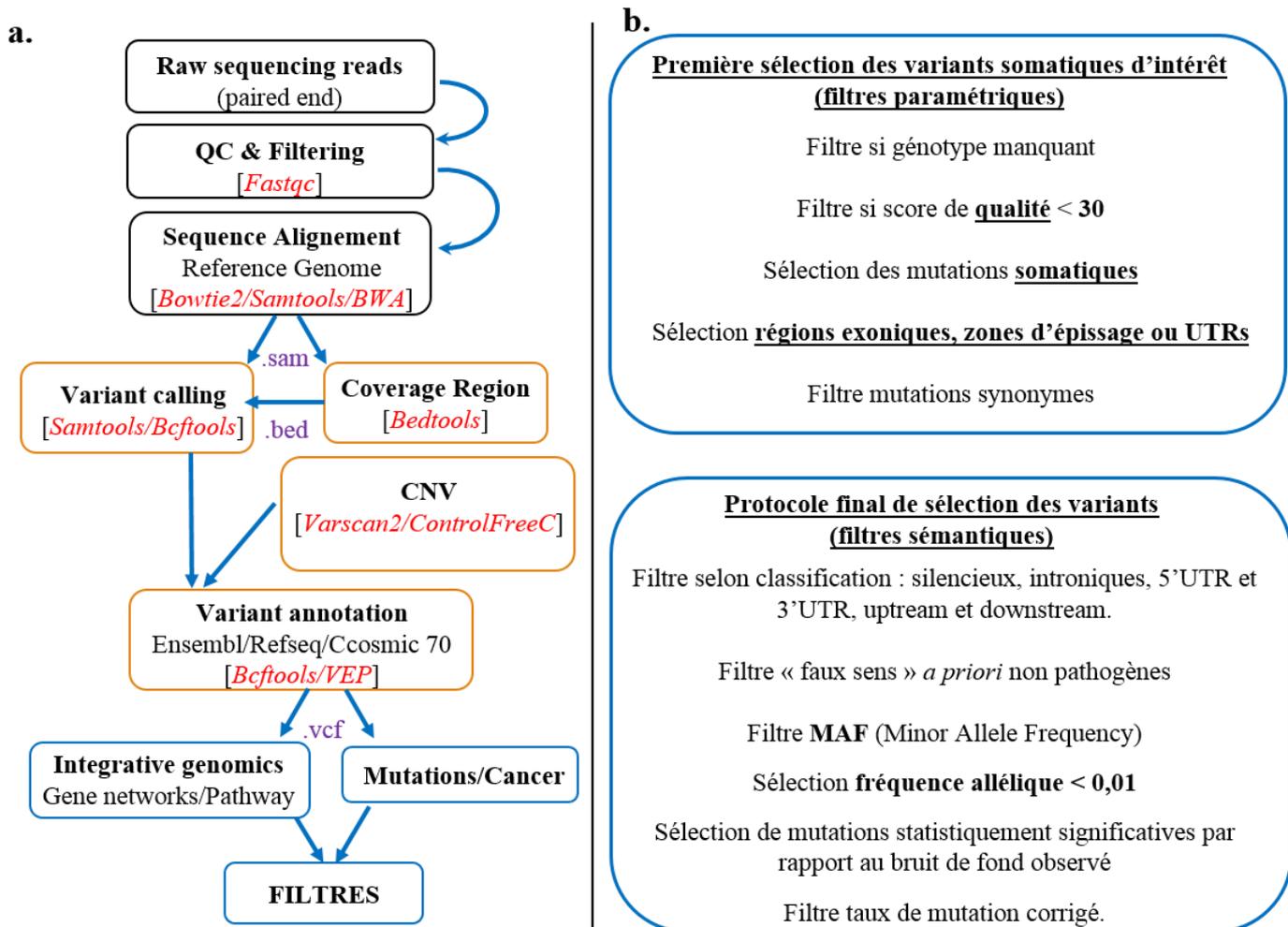


Figure 47. Étapes clés de l'analyse d'exome.

a. Pipeline d'analyse de l'exome utilisé pour notre étude. Trois grandes étapes se distinguent : l'analyse primaire (cadres noirs), l'analyse secondaire (cadres orange) et l'analyse tertiaire (cadres bleus). L'analyse primaire est automatique et comprend les contrôles qualité et l'alignement ; l'analyse secondaire est fortement paramétrable et comprend l'appel/l'identification et l'annotation des variants ; l'analyse tertiaire est plus approfondie et comprend notamment l'application de filtres de sélection ou de suppression des variants trouvés. **b. Filtres** appliqués lors de l'analyse tertiaire.

Chaque gène de la liste initiale de candidats a été discuté sur le plan de sa pertinence scientifique théorique. Cette discussion s'est appuyée sur les données de la littérature, celle des bases de données disponibles (*GeneCards*, *Ensembl*) et en comité d'experts (qui associait un bioinformaticien, un hématocytologiste et un généticien, tous expérimentés dans le domaine de l'hématopoïèse). La liste a ensuite été enrichie par la recherche de variants impliqués dans la mégacaryopoïèse et qui auraient été filtrés au cours du dernier protocole d'analyse bioinformatique.

L'analyse des CNV (*Copy Number Variation*) a été réalisée avec l'outil *Control FreeC*.

PARTIE 1 - TROISIÈME COHORTE – COHORTE DE CONFIRMATION

Afin de confirmer et de tester la liste finale de gènes candidats issues de l'analyse de l'exome de la deuxième cohorte, nous avons travaillé en partenariat avec le laboratoire MLL à Munich. Un génome ciblé de nos gènes candidats a été réalisé sur 42 échantillons d'ADN de MDS/MPN-*SF3B1*-T TN et pour lesquels les fichiers « .vcf » nous ont été fournis.

Les résultats de ce séquençage ont permis la sélection finale de gènes candidats dont les variants ont été ou seront testés sur le plan fonctionnel.

PARTIE 1 - ANALYSES FONCTIONNELLES

1) Variant c.2117 A>T de *STAT5B*

a) *Analyses in silico*

Pour prédire l'impact potentiel de la mutation c.2117 A>T (Q706L), nous avons utilisé les outils bioinformatiques prédictifs *exskip* (<https://ex-skip.img.cas.cz/>) et *ESE finder* (148) version 3.0 (<https://esefinder.ahc.umn.edu/cgi-bin/tools/ESE3/esefinder.cgi>).

b) *Production de plasmide STAT5b*

Le gène humain sauvage *STAT5b*^{WT} a été amplifié par PCR à partir d'ADN génomique extrait de cellules HEK293 (*Human Embryonic Kidney 293 cells*) avec les oligonucléotides PM-21.0126 et PM-21.0127 (**Tableau 7**) et le kit CloneAMP TM HIFI PCR Premix® (Takara Bio, Kyoto, Japon) en suivant les recommandations du fournisseur. Cette amplification a permis la synthèse d'un insert de 400 bases comportant l'exon 17 de *STAT5b*^{exon17WT} ainsi que les régions introniques adjacentes et des sites de restriction d'*Eco-RI*. Les produits de PCR obtenus ont ensuite été digérés par l'enzyme de restriction *EcoRI* (ThermoFisher, Waltham Massachusetts, USA). La digestion a été contrôlée sur gel d'agarose à 0,8 % et l'insert obtenu a été purifié sur colonne (Qiagen, Hilden, Allemagne).

Tableau 7. Séquence des oligonucléotides utilisés pour la production de plasmide pXJ41-STAT5b^{exon17^{WT}} et pXJ41-STAT5b^{exon17^{MUT}}.

Nom de l'oligonucléotide	Orientation	Séquence
PM-21.0126	Forward	5' AAAAAGAATTCGACTGTTTCAAAAAACAAA 3'
PM-21.0127	Reverse	5' TTTTGAATTCGGGCCTCAGGTACTGGCAT 3'
PM-21.0150	Forward	5' ACGGTGCATTGGAACGGACC 3'
PM-21.0010	Reverse	5' ACCACTAGCTTGATCTGTGGCTTCACG 3'
PM-21.0123	Forward	5' GATCAAGCTAGTGGTCCCTGAGTAAGTGTC 3'
PM-21.0010	Reverse	5' ACCACTAGCTTGATCTGTGGCTTCACG 3'
PM-21.0123	Forward	5' ACGGTGCATTGGAACGGACC 3'
PM-21.0124	Reverse	5' CAAGTTCTGCTTTAATAAGATC 3'

Quatre µg de plasmide pXJ41 contenant les exons 2 et 3 de la bêta-globine de lapin ont été digérés par l'enzyme de restriction *MfeI-Hf* (New england biolab, Ipswich, Massachussets, USA). Cent dix ng d'insert *STAT5b^{WT}* ont été insérés dans 280 ng de plasmide pXJ41 linéarisé avec le kit in-Fusion HD EcoDryTM cloning ® (Takara Bio, Kyoto, Japon) (**Figure 48**).

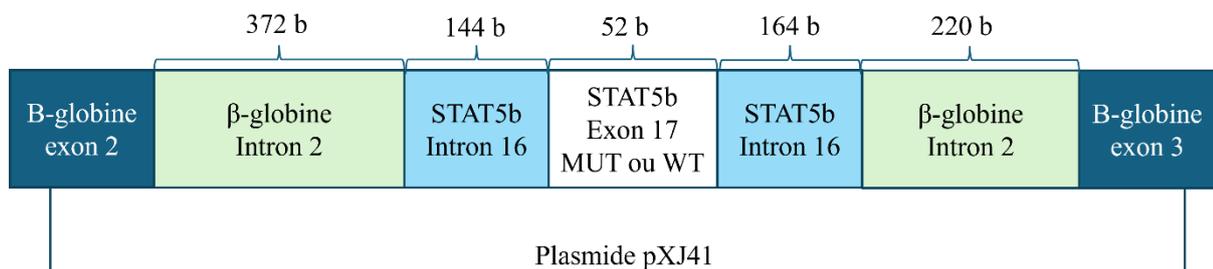


Figure 48. Représentation de la construction pXJ41-STAT5b^{exon17}.

La purification sur gel a été réalisée avec le kit *QIAquick* (Qiagen, Hilden, Allemagne). Les plasmides pXJ41-STAT5b^{exon17^{WT}} et pXJ41 « vides » (contrôle) ont été utilisés pour la transformation de bactéries compétentes OneShotTM TOP10 (F- *mcrA* Δ (*mrr-hsdRMS-mcrBC*) Φ80*lacZ*Δ*M15* Δ *lacX74* *recA1* *araD139* Δ (*araleu*) 7697 *galU* *galK* *rpsL* (StrR) *endA1* *nupG* ; ThermoFisher, Waltham Massachussets, USA). L'ADN plasmidique des clones positifs a été purifié avec le kit *QIAGEN Plasmid Mini Kit* (Qiagen, Hilden, Allemagne) et vérifiés par séquençage Sanger (Eurofins Genomics Europe, Ebersberg, Allemagne). L'ADN plasmidique utilisé pour les transfections a été préparé à l'aide du kit *QIAGEN Plasmid Midi Kit* (Qiagen, Hilden, Allemagne).

c) *Mutagenèse dirigée*

La mutagenèse a été réalisée par PCR (**Tableau 8**) avec le kit CloneAMP™ HIFI PCR Premix® (Takara Bio, Kyoto, Japon) à partir de 5 ng de la construction pXJ41-*STAT5b*^{exon17WT} et les oligonucléotides PM-21.0150 (*forward*) et PM-21.0010 (*reverse*) permettant d'introduire la mutation c.2117 A>T située dans l'exon 17 de *STAT5b* (**Tableau 7**).

Tableau 8. Protocole de PCR pour la mutagenèse dirigée de *STAT5b* dans le plasmide pXJ41-*STAT5b*^{exon17WT}.

Etape	Température	Durée (secondes)	Nombre de cycles
Fusion	98°C	10	35
Hybridation	55°C	5	
Elongation	72°C	30	

L'ADN plasmidique des clones obtenus a été séquencé auprès de la société Eurofins® avec les amorces PM-21.0123 (*forward*) et PM-21.0010 (*reverse*) (**Tableau 7**). L'ADN utilisé pour la transfection a été préparé à l'aide du kit *QIAGEN Plasmid Midi Kit* (Qiagen, Hilden, Allemagne).

d) *Minigene assay*

Les plasmides pXJ41-*STAT5b*^{exon17WT} et pXJ41-*STAT5b*^{exon17MUT} ont été transfectés dans des cellules HEK293, cultivées dans du milieu DMEM (Sigma Aldrich, Saint-Louis, Missouri, USA) supplémenté avec 5% de Sérum de Veau Fœtal (SVF) et 1% d'antibiotiques (Streptomycine et Pénicilline). Les cellules ont étéensemencées dans des plaques de 6 puits, à raison de 2×10^5 cellules par puit, 24 heures avant la transfection.

Une quantité de 2,5 µg de plasmides (1,5 µg de plasmide pBS-SK et 1 µg de plasmide pXJ41-*STAT5b*^{exon17WT} ou 1 µg de plasmide pXJ41-*STAT5b*^{exon17MUT}) a été mélangée avec 75 µL de chlorure de sodium (NaCl) 150 mM puis ajouté à un mélange contenant 6 µL de JetPEI (Polyplus, Illkirch-Graffenstaden, Alsace, France) et 94 µL de NaCl 150 mM. Après une incubation de 30 minutes, la solution a été déposée goutte à goutte dans le milieu de culture avec les cellules.

L'ARN des cellules HEK293 transfectées a été extrait avec le kit *Nucleospin RNA plus* (Macherey Nagel, Düren, Allemagne). Il a été rétro-transcrit en ADN complémentaire (ADNc) puis amplifié par PCR avec les amorces PM-21.0123 (*forward*) et PM-21.0124 (*reverse*) (**Tableau 7**). Les produits de PCR ont été déposés pour migration sur gel d'agarose à 2 %. La taille attendue des produits de PCR était de 269 bases ou 217 bases selon l'inclusion ou non de l'exon 17 de *STAT5b*.

Une cotransfection avec des plasmides permettant de surexprimer hRNPA1, SC35 ou SRp40 a également été réalisée.

2) PTPN12

a) *qPCR de PTPN12*

Les échantillons analysés pour la qPCR (*quantitative PCR*) de *PTPN12* (*Protein Tyrosine Phosphatase Non-Receptor Type 12*) étaient des échantillons d'ADN de patients atteints de TE TN ou de MF TN. L'ADN était extrait sur sang total (populations non triées). L'amplification a été réalisée avec un master mix *SYBR Green* (Thermofisher, Foster City, Californie, USA), avec trois couples d'amorce spécifique de *PTPN12* (PTPN12-1, PTPN12-10 et PTPN12-18) et un couple d'amorces spécifique du gène de référence, en triplicats et contre un témoin pour lequel aucun polymorphisme de *PTPN12* n'a été décrit. Le gène de référence était *RPPH1* (*Ribonuclease P RNA Component H1*), situé sur le chromosome 14. La réaction de qPCR a été effectuée sur un appareil *7500 Fast Real System* (Applied Biosystem®, Thermofisher, Foster City, Californie, USA) et l'analyse a été réalisée sur le logiciel *7500 Fast System Software*. Les niveaux d'expression ont été calculés avec la méthode $2^{-\Delta\Delta-Ct}$.

b) *Production des plasmides PTPN12*

L'ADN complémentaire codant pour le gène humain sauvage *PTPN12^{WT}* (**Tableau 9**) a été inséré avec le kit de clonage *In-fusion* (Takara bio Inc, Kusatsu, préfecture de Shiga, Japon) dans un plasmide peGFP-C2 contenant un promoteur viral issu du CMV contrôlant l'expression d'une protéine fluorescente (EGFP) ainsi que le promoteur SV40 contrôlant l'expression d'un gène de résistance à la kanamycine. Les plasmides ont été préparés en utilisant l'enzyme de restriction *EcoRI* et les oligonucléotides de *PTPN12* PM-19.0684 (*Forward*) et PM-19.0685 (*Reverse*). La purification sur gel a été réalisée avec le kit d'extraction sur gel *QIAquick* (Qiagen, Hilden, Allemagne).

Les plasmides ont été amplifiés dans des bactéries StellarTM (*F*⁻, *ara*, $\Delta(lac-proAB)$ [$\Phi80d lacZAM15$], *rpsL(str)*, *thi*, $\Delta(mrr-hsdRMS-mcrBC)$, $\Delta mcrA$, *dam*, *dcm* ; Takara bio Inc) et purifiés avec le kit *QIAGEN Plasmid Mini Kit* (Qiagen, Hilden, Allemagne). Les plasmides ont été séquencés avant leur amplification dans les bactéries compétentes avec le kit *QIAGEN Plasmid Midi Kit* (Qiagen, Hilden, Allemagne).

Tableau 9. Séquence des oligonucléotides utilisés pour le clonage de *PTPN12* dans le plasmide peGFP-C2.

Orientation	Sequence
Forward	5' TCGAGCTCAAGCTTCGAATTATGGAGCAAGTGGAGATCC 3'
Reverse	5' GTACCGTCGACTGCAGAATTCATGTCCATTCTGAAGGTGG 3'

c) *Cultures cellulaires et transfection*

Nous avons travaillé avec deux lignées cellulaires : THP-1 (*Tohoku Hospital Pediatrics-1 cells*) (lignée de LAM5) et HEK293 (*Human Embryonic Kidney 293 cells*).

Les cellules THP-1 sont cultivées dans du milieu RPMI (*Roswell Park Memorial Institute*) supplémenté avec 10% de SVF (Sigma Aldrich), 1% de glutamine et 1% d'antibiotiques (Streptomycine et Pénicilline).

Les cellules HEK293 sont quant à elles cultivées dans du milieu DMEM (*Dulbecco Modified Eagle Medium* ; Sigma Aldrich, Saint-Louis, Missouri, USA) supplémenté avec 5% de SVF et 1% d'antibiotiques (Streptomycine et Pénicilline). Elles ont étéensemencées dans des plaques de 6 puits, à raison de 2×10^5 cellules par puit.

Une quantité de 2,5 μ g de plasmides (1,5 μ g de plasmide pBS-SK et 1 μ g de plasmide peGFP-C2-PTPN12 ou 1 μ g de plasmide peGFP-C2 sans *PTPN12*) a été mélangée avec 75 μ L de chlorure de sodium (NaCl) 150 mM puis ajouté à un mélange contenant 6 μ L de JetPEI (Polyplus, Illkirch-Graffenstaden, Alsace, France) et 94 μ L de NaCl 150 mM. Après une incubation de 30 minutes, la solution a été transférée dans le milieu de culture avec les cellules. L'efficacité de la transfection a été évaluée grâce à la GFP fluorescente co-transfectée avec un microscope confocal laser.

d) *Immunoblotting*

Les protéines ont été extraites dans du tampon RIPA (0,5M Tris-HCl, pH 7.4, 1,5M NaCl, 2,5% acide déoxycholique, 10% NP-40, 10mM EDTA). Les protéines (30 µg) ont ensuite été séparées par électrophorèse sur un gel de SDS-polyacrylamide et transférées sur une membrane de nitrocellulose par transfert semi-sec.

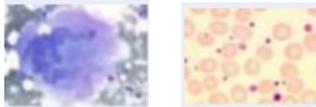
L'immunoblot a été réalisé avec un anticorps anti-PTP-PEST (Ozyme, Saint-Cyr-l'École, France), un anticorps anti-GFP (Abcam, Cambridge, Royaume-Uni), un anticorps anti-ERK (Cell Signaling, Danvers, USA) et anticorps anti-AKT (Cell Signaling, Danvers, USA). Les bandes ont été révélées par chimioluminescence (ECL, Pierce Perbio, Brebières, France).

PARTIE 2 - QUATRIÈME COHORTE – COHORTE MYST

1) Collection biologique et données cliniques

L'analyse du méthylome des MDS/MPN-*SF3B1*-T TN a été réalisée par méthode comparative avec 4 groupes « contrôle » : MDS-*SF3B1*, TE *JAK2*^{V617F}, MDS/MPN-*SF3B1*-T *JAK2*^{V617F} et moelles osseuses normales (**Tableau 10**). La collection biologique formée pour l'analyse du RRBS est constituée d'échantillons issus du Centre de ressources biologiques (CRB) du CHU de Dijon (MDS-*SF3B1*, MDS/MPN-*SF3B1*-T), du CRB du CHRU (centre hospitalier régional universitaire) de Nancy (moelles osseuses normales) et du laboratoire de génétique du CHRU de Nancy (TE *JAK2*^{V617F}, MDS/MPN-*SF3B1*-T) entre 2007 et 2021. Les échantillons issus des CRB ont été conservés avec le consentement des patients. Les échantillons du laboratoire de génétique du CHRU de Nancy correspondaient à des reliquats d'ADN extraits dans le cadre du soin. Les patients ont reçu une information et n'ont pas émis d'opposition à leur utilisation. Tous les échantillons ont été anonymisés.

Tableau 10. Présentation morphologique et mutationnelle des 5 groupes faisant l'objet d'une analyse comparative du méthylome.

	Groupe 1 MDS- <i>SF3B1</i>	Groupe 2 MDS/MPN- <i>SF3B1</i> -T triple négatifs	Groupe 3 MDS/MPN- <i>SF3B1</i> -T <i>JAK2</i> ^{V617F}	Groupe 4 TE <i>JAK2</i> ^{V617F}	Groupe 5 Moelles normales
	Dysplasie <i>SF3B1</i> induite	Thrombocytose d'origine indéterminée et dysplasie <i>SF3B1</i> induite	Thrombocytose <i>JAK2</i> ^{V617F} induite et dysplasie <i>SF3B1</i> induite	Thrombocytose <i>JAK2</i> ^{V617F} induite	
Composante myéloproliférative Mégacaryocytes géants – thrombocytose					
	Non	Oui	Oui	Oui	Non
<i>JAK2</i> ^{V617F}	Non	Non	Oui	Oui	Non
<i>MPL</i> ^{W515K/L}	Non	Non	Non	Non	Non
<i>CALR</i> exon 9	Non	Non	Non	Non	Non
Composante myélodysplasique Dysérythropoïèse – SC					
	Oui	Oui	Oui	Non	Non
<i>SF3B1</i> muté	Oui	Oui	Oui	Non	Non
Profil de méthylation					
	(1)	?	(1) + (2)	(2)	(3)

(1) Représentatif du versant myélodysplasique ; (2) représentatif du versant myéloprolifératif *JAK2*-induit ; (3) méthylome médullaire de référence.

Pour chaque patient, les données collectées ont compris : la source de l'échantillon, le type de cellules dont l'échantillon a été issu, le sexe, l'âge, les antécédents médicaux (néoplasie et risque cardiovasculaire notamment), le diagnostic, la date du diagnostic, les caractéristiques cliniques telles que la présence d'une splénomégalie, l'hémogramme au diagnostic, l'analyse cytologique médullaire au diagnostic (dont la coloration de Perls), le caryotype médullaire, l'analyse moléculaire ciblée de *SF3B1*, *JAK2*, *CALR* et *MPL*, l'analyse histologique d'une biopsie ostéo-médullaire (si disponible), l'évolution de la pathologie, la dépendance transfusionnelle, l'occurrence de thromboses, l'occurrence d'hémorragies, l'évolution vers une LAM secondaire, le statut vital à la dernière visite.

Compte-tenu du caractère rétrospectif de notre étude, de la variabilité des pathologies et de leurs prises en charge diagnostiques, les échantillons étaient issus de différentes origines (sang, moelle). Lorsque plusieurs sources (cellules polynucléées, cellules mononucléées, fraction totale) ou origines différentes étaient disponibles pour un même patient, elles ont toutes été prises en compte dans le plan de plaque (**Tableau 11**).

Tableau 11. Plan de l'utilisation des échantillons dans la cohorte MYST.

<u>MDS-SF3B1</u>	<u>MDS/MPN-SF3B1-T</u> <u>triple négatives</u>	<u>MDS/MPN-SF3B1-T</u> <u>JAK2^{V617F}</u>	<u>TE</u> <u>JAK2^{V617F}</u>	<u>Moelles</u> <u>normales</u>
MYST 1 moelle MN	MYST 14 sang PN	MYST 9 moelle MN	MYST 24 moelle totale	MYST 20 moelle totale
MYST 1 moelle PN	MYST 14 moelle MN	MYST 9 moelle PN	MYST 26 sang total	MYST 21 moelle totale
MYST 2 moelle MN	MYST 14 moelle PN	MYST 9 moelle MN	MYST 27 sang total	MYST 22 moelle totale
MYST 2 moelle PN	MYST 15 sang PN	MYST 9 sang PN	MYST 28 sang total	MYST 23 moelle totale
MYST 3 moelle PN	MYST 16 sang PN	MYST 10 sang MN	MYST 29 sang total	
MYST 4 moelle PN	MYST 17 moelle totale	MYST 10 moelle PN	MYST 30 sang total	
MYST 5 moelle MN	MYST 18 moelle MN	MYST 11 moelle totale	MYST 31 sang total	
MYST 6 moelle MN	MYST 19 sang total	MYST 12 moelle MN	MYST 32 sang total	
MYST 7 moelle MN	MYST 19 moelle MN	MYST 13 sang MN	MYST 33 sang total	
MYST 8 moelle MN	MYST 37 sang total	MYST 38 moelle totale	MYST 34 sang total	
	MYST 42 sang PN	MYST 38 sang total		
	MYST 43 sang PN	MYST 41 sang total		

MN = mononucléées ; PN = polynucléées

2) Extraction et préparation des librairies

Pour chaque échantillon des CRB, l'ADN a été extrait manuellement avec un kit *NucleoSpin Tissue* (Macherey-Nagel, Düren, Allemagne).

Pour les échantillons du laboratoire de génétique du CHRU de Nancy, l'ADN a été extrait avec un kit *QIAamp DNA Blood Mini QIAcube Kit* (Qiagen, Hilden, Allemagne) sur un automate Qiacube (Qiagen, Hilden, Germany).

Après un contrôle qualité sur *Tape Station* (Agilent, technologies, Santa Clara, CA, USA), 100 ng d'ADN extrait ont été utilisés pour réaliser un RRBS avec le kit *Diagenode Premium RRBS* (Denville, New Jersey, USA), suivant les instructions du constructeur, tel que précédemment décrit (149). Ce kit est basé sur une digestion enzymatique de motifs riches en GC puis d'une capture des fragments selon leur taille par billes magnétiques.

Cette méthode permet d'obtenir des librairies enrichies en sites de méthylation (îlots CpG). La technique a compris les étapes suivantes :

- (i) digestion de l'ADN avec l'enzyme de restriction *MspI*
- (ii) réparation des extrémités et ligation des adaptateurs Illumina
- (iii) sélection des fragments d'ADN avec un insert d'une taille comprise entre 40 et 220 pb
- (iv) bisulfitation avec le kit *EZ DNA Methylation TM Kit* (Zymo research, Irvine, CA, USA)
- (v) amplification par PCR et séquençage sur l'automate Illumina Nextseq 500/550 (Illumina®, San Diego, CA, USA) avec la technologie *Reverse Dye Terminator*.

L'appel des bases a été réalisé selon la méthode décrite avec le module *bclConvert* du pipeline Illumina pendant 71 cycles de lecture en *single end*. Le processus de *trimming* a été effectué avec l'outil *TrimGalore* (Babraham Institute, University of Cambridge). Les séquences ont ensuite été alignées avec la méthode Bismark (150) sur le génome humain de référence Hg38 et le génome bisulfité. Afin de conserver une information suffisante pour l'alignement, seules les séquences de plus de 30 bases ont été sélectionnées pour les analyses suivantes (séquences de 30 à 71 bases).

3) Caractérisation moléculaire

Dans le but de mieux caractériser et de confirmer le statut mutationnel de chaque échantillon, nous avons réalisé un séquençage haut débit de routine sur un panel myéloïde comportant les gènes suivants : *ABL1* exons 4-9, *ASXL1* exons 10,12,13, *BRAF* exon 15, *CALR* exon 9, *CBL* exons 8,9, *CEBPA*, *CSF3R*, *DNMT3A*, *ETV6*, *EZH2*, *FLT3* exons 13-15,20, *HRAS* exons 2-3, *IDH1* exon 4, *IDH2* exon 4, *JAK2*, *KIT* exons 2,8-11,13,17,18, *KRAS* exons 2,3, *MPL* exon 10, *NPM1* exons 10,11, *NRAS* exons 2-3, *PTPN11* exons 3,7-13, *RUNX1*, *SETBP1* exon 4, *SF3B1* exons 10-16, *SRSF2* exon 1, *TET2*, *TP53*, *U2AF1* exons 2-6, *WT1* exons 6-10, *ZRSR2*.

4) Disponibilités des données de séquençage et type d'analyses

L'analyse finale des fichiers «.bam» correspondants aux 48 échantillons a été réalisée en 2023 selon deux méthodologies différentes :

- avec le package *MethylKit* du logiciel R v.4.1 (151)
- avec une méthode « *in-house* » dont le design a été créé pour cette étude.

Dans les deux cas, le taux de méthylation a été calculé pour chaque CpG avec une profondeur d'au moins 20 *reads*. Les CpGs avec une profondeur située au-delà du 99,5^{ème} percentile ont été éliminés afin de limiter le biais de PCR. Les positions situées sur le chromosome X ont également été filtrées. Les fichiers ont ensuite été fusionnés dans chaque groupe.

5) Méthodologie Méthylkit

a) Méthodologie *MethylKit* pour les CpGs

Avec la méthodologie *MethylKit*, la comparaison des groupes par les CpGs est faite selon une analyse de la variation du pourcentage de méthylation sur une position d'un groupe à l'autre (**Figure 49**).

Les positions étaient rapportées seulement si l'information (« méthylié » ou « non méthylié ») était connue pour au moins 3 individus du groupe (2 pour le groupe des moelles normales).

La différence de méthylation était considérée significative si elle était d'au moins 20% entre les deux groupes (soit une différence de 0,2 entre les deux groupes sur une échelle de 0 à 1).

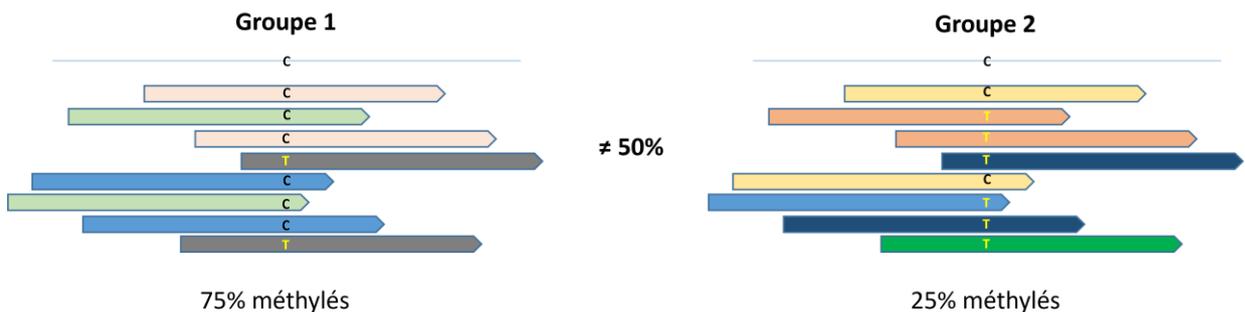


Figure 49. Comparaison des groupes par les CpGs selon *MethylKit*.
Chaque couleur de *read* représente un échantillon différent.

b) Méthodologie *MethylKit* pour les DMRs

Cette méthodologie s'appuie sur la création de régions de 1000 bases dont le niveau de méthylation est comparé d'un groupe à l'autre. Les régions différentiellement méthylées entre les groupes sont appelées « *differentiated methylated regions* » (DMRs).

Selon la méthode *Methylkit*, ces régions sont définies par une fenêtre de 1000 bases glissant de 500 bases en 500 bases le long du génome (**Figure 50**).

Le niveau de méthylation est défini par la moyenne de méthylation de toutes les positions de cytosines méthylables au sein de cette même région, en cumulant tous les échantillons du groupe (nombre de C sur nombre total de *reads* du groupe pour une même position).

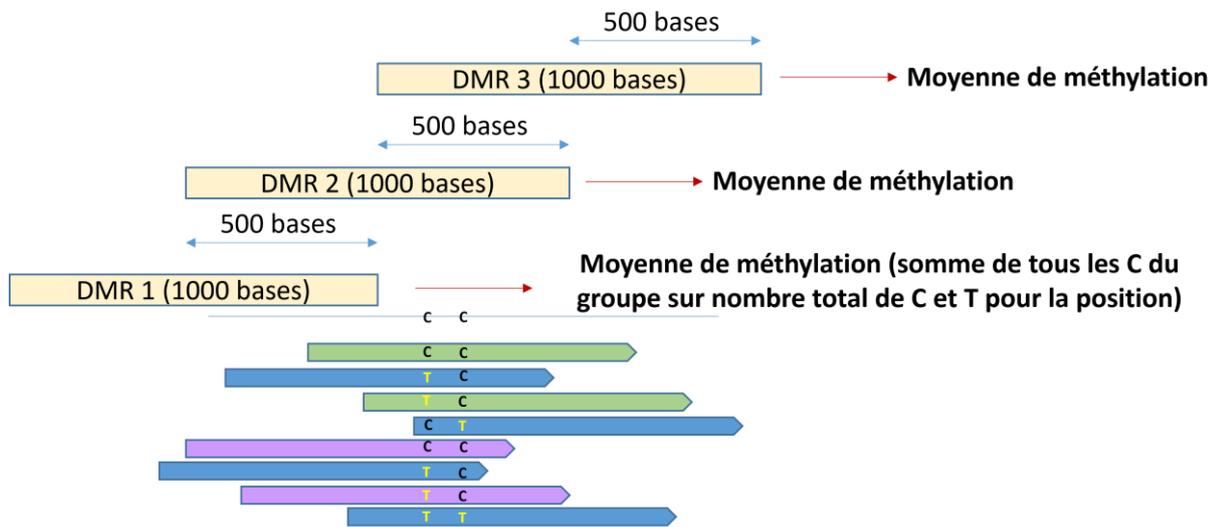


Figure 50. Création de DMR selon la méthode *Methylkit*.

Ces DMRs ont été reliées à des gènes après annotations selon les bases de données :

- *UCSC RefGene Name* (152) lorsqu'elles étaient situées à moins de 2 000 bases d'un gène
- *NCBI Hgnc* (153) lorsqu'elles étaient situées à plus de 2 000 bases d'un gène et non disponibles via Illumina .

6) Méthodologie « *in-house* »

a) *Méthodologie in-house pour les CpG*

Selon la méthodologie que nous avons développée pour cette étude, à partir des données normalisées issues de *MethylKit*, la comparaison des groupes par les CpG est faite en comparant des vecteurs par un test de Student (**Figure 51**).

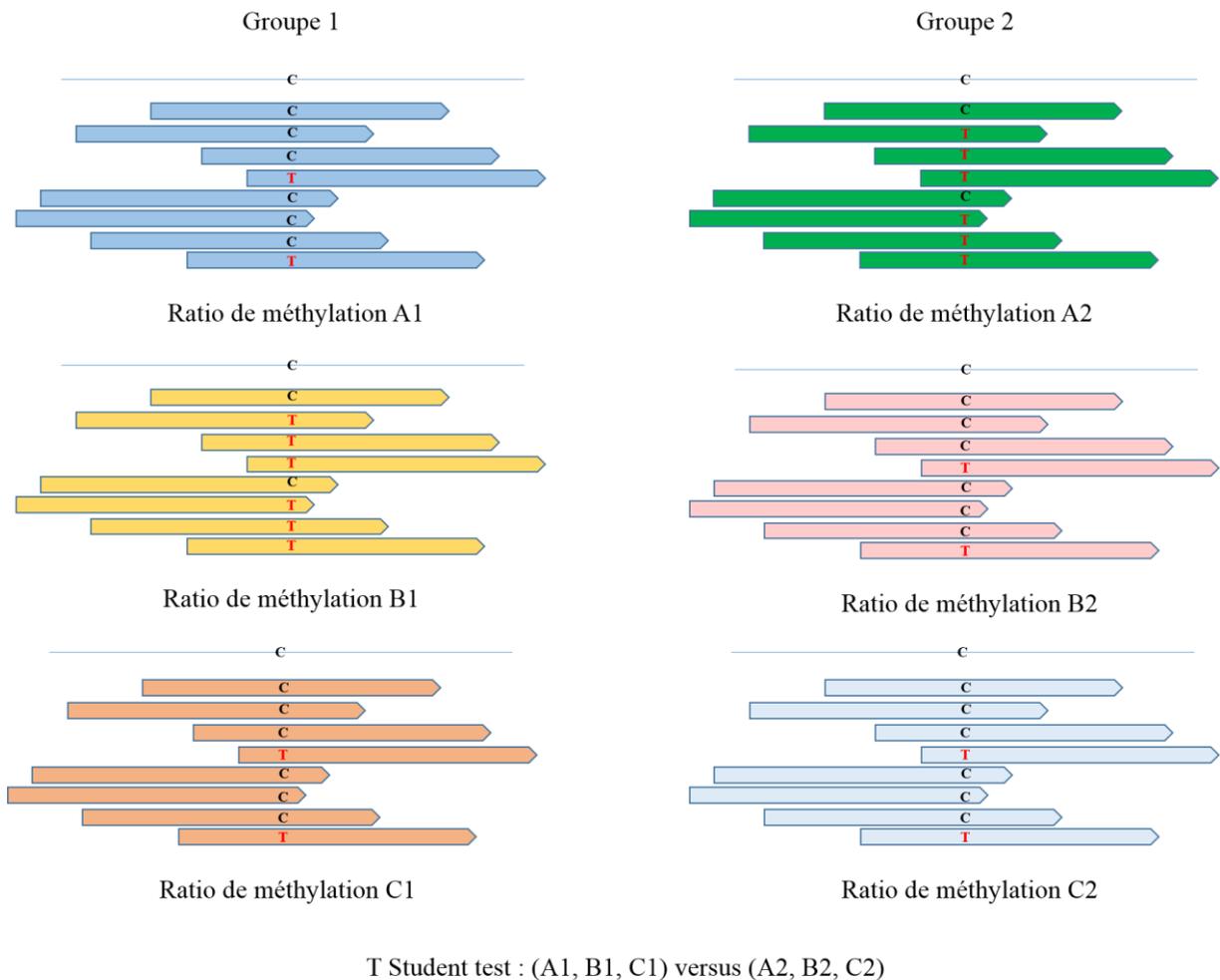


Figure 51. Représentation de la méthodologie de comparaison *in-house* pour les CpGs.

Pour chaque échantillon au sein d'un groupe et pour une position donnée, nous avons établi un ratio de méthylation correspondant au nombre de bases C (méthylées) divisé par le nombre total de bases C et T (bases méthylées et non méthylées pour une position).

Pour chaque position et pour chaque groupe, nous avons donc constitué un vecteur de ratios de méthylation. Cette méthode permet de conserver la notion de variabilité interindividuelle au sein d'un même groupe. Les positions étaient rapportées seulement si l'information (« méthylé » ou « non méthylé ») était connue pour au moins 2 individus du groupe.

Pour une analyse non supervisée globale de la méthylation, nous avons étudié d'une part les 15 000 CpGs les plus variants et d'autre part les CpGs présentant une différence de méthylation d'au moins 5% (soit une différence de méthylation de 0,05 sur une échelle de 0 à 1). Afin d'augmenter la stringence de notre pipeline, nous avons fait varier ce seuil lors de nos analyses définitives : à chaque comparaison, nous avons retenu comme significatifs les CpGs qui présentaient une différence de méthylation de plus de 20% (soit une différence de 0,2 sur une échelle de 0 à 1) entre les deux groupes.

b) *Méthodologie in-house pour les DMRs*

Lors de l'analyse incluant l'ensemble des données, nous avons réorienté notre stratégie méthodologique pour la sélection des DMRs. La méthodologie appliquée ici s'inspire de *Methylkit*. Les régions à comparer sont également issues d'une fenêtre glissante de 1000 bases toutes les 500 bases mais la méthode de comparaison diffère.

Pour un même groupe, un ratio de méthylation est calculé pour chaque position et pour chaque échantillon (comme lors de l'analyse des CpGs). Il est ensuite établi, pour le groupe et pour chaque position, une moyenne de ces ratios (**Figure 52**

). La comparaison de deux DMRs s'effectue par t-test en comparant, non plus une somme globale ne différenciant pas les positions, mais les vecteurs composés des moyennes obtenues pour chaque position au sein des groupes :

- premier vecteur constitué de l'ensemble des moyennes de ratios par position dans le premier groupe
- second vecteur constitué de l'ensemble des moyennes de ratios par position dans le second groupe

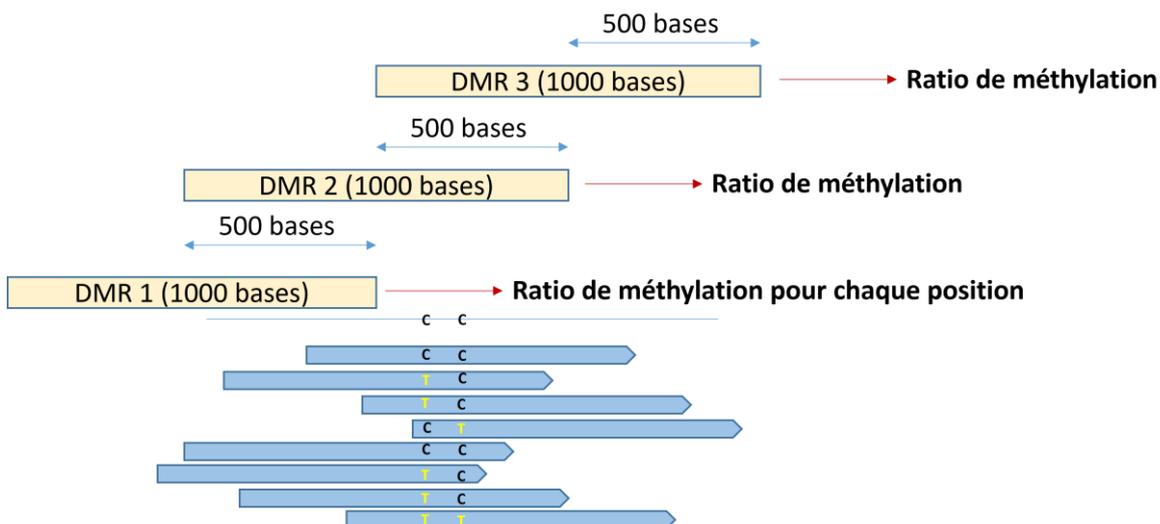


Figure 52. Représentation du calcul du ratio de méthylation d'une DMR pour chaque échantillon et selon la méthodologie *in-house*.

À chaque comparaison, nous avons retenu comme significatives les régions qui présentaient une différence de méthylation de plus de 20% (soit une différence de 0,2 sur une échelle de 0 à 1) entre les deux groupes et qui contenaient au moins 10 CpGs considérés significatifs avec une profondeur d'au moins 10 *reads* et couverts pour au moins 2 échantillons de chaque groupe. Nous avons exclu le chromosome X de nos analyses afin de s'affranchir du biais de méthylation lié au sexe. Les valeurs p ont été ajustées selon la méthode de Benjamini Hochberg (154). Ces DMRs ont été reliées à des gènes après annotations selon les bases de données :

- *UCSC RefGene Name* (152) lorsque qu'elles étaient disponibles
- *NCBI Hgnc* (153) lorsqu'elles n'étaient pas disponibles via Illumina .

c) *Élimination du bruit de fond dans la méthode in-house*

Parmi les DMRs obtenues par la comparaison entre MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et MDS/MPN-*SF3B1*-T-TN, nous avons éliminé de notre analyse celles qui étaient :

- **Communes aux comparaisons TE-*JAK2*^{V617F} versus MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et TE-*JAK2*^{V617F} versus MDS/MPN-*SF3B1*-T-TN** : correspondant à des régions différentiellement méthylées en lien avec le caractère dysplasique non spécifique des MDS/MPN-*SF3B1*-T.
- **Communes aux comparaisons MDS-*SF3B1* versus MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et MDS-*SF3B1* versus MDS/MPN-*SF3B1*-T-TN** : correspondant à des régions différentiellement méthylées en lien avec le caractère prolifératif non spécifique des MDS/MPN-*SF3B1*-T.

d) *Etablissement d'un score de pertinence selon les comparaisons*

Parmi les DMRs obtenues par la comparaison entre MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et MDS/MPN-*SF3B1*-T-TN selon la méthodologie spécifiquement conçue dans notre laboratoire, nous avons appliqué un score de pertinence sur 5 points.

Ce score a été établi afin d'identifier les DMRs les plus pertinentes pour caractériser la prolifération des MDS/MPN-*SF3B1*-T **triple négatifs** (DMRs non liées à la dysplasie, non liées à la prolifération induite par *JAK2*^{V617F}, spécifique du caractère MDS/MPN-*SF3B1*-T et du caractère TN) (**Tableau 12, Figure 53, Figure 54**). En effet, nous cherchons à filtrer les DMRs non pertinentes car une DMR peut refléter une différence entre deux groupes, parfois indépendamment de la problématique posée.

La liste finale de DMRs candidates est constituée des DMRs issues de la comparaison MDS/MPN-SF3B1-T-JAK2^{V617F} versus MDS/MPN-SF3B1-T-TN et pour lesquelles ce score était de 5/5 (ensemble des critères réunis).

Tableau 12. Score de pertinence appliqué aux DMRs issues de la comparaison MDS/MPN-SF3B1-T-JAK2^{V617F} vs MDS/MPN-SF3B1-T-TN.

Comparaisons	Présence ou absence de la DMR	Justification	Point
TE- JAK2 ^{V617F} vs MDS/MPN-SF3B1-T-TN	Présence (DMR pertinente)	DMR reliée à la dysplasie ou à la prolifération non liée à JAK2 ^{V617F}	1
MDS/MPN-SF3B1-T- JAK2 ^{V617F} + TE- JAK2 ^{V617F} vs MDS/MPN-SF3B1-T-TN	Présence (DMR pertinente)	DMR reliée à la prolifération induite par le « caractère TN »	1
MDS-SF3B1 vs TE- JAK2 ^{V617F}	Absence (DMR non spécifique)	DMR reliée à la dysplasie ou à la prolifération induite par JAK2 ^{V617F}	1
TE- JAK2 ^{V617F} vs moelles normales	Absence (DMR non spécifique)	DMR reliée à la prolifération induite par JAK2 ^{V617F}	1
MDS-SF3B1 vs moelles normales	Absence (DMR non spécifique)	DMR reliée à la dysplasie	1

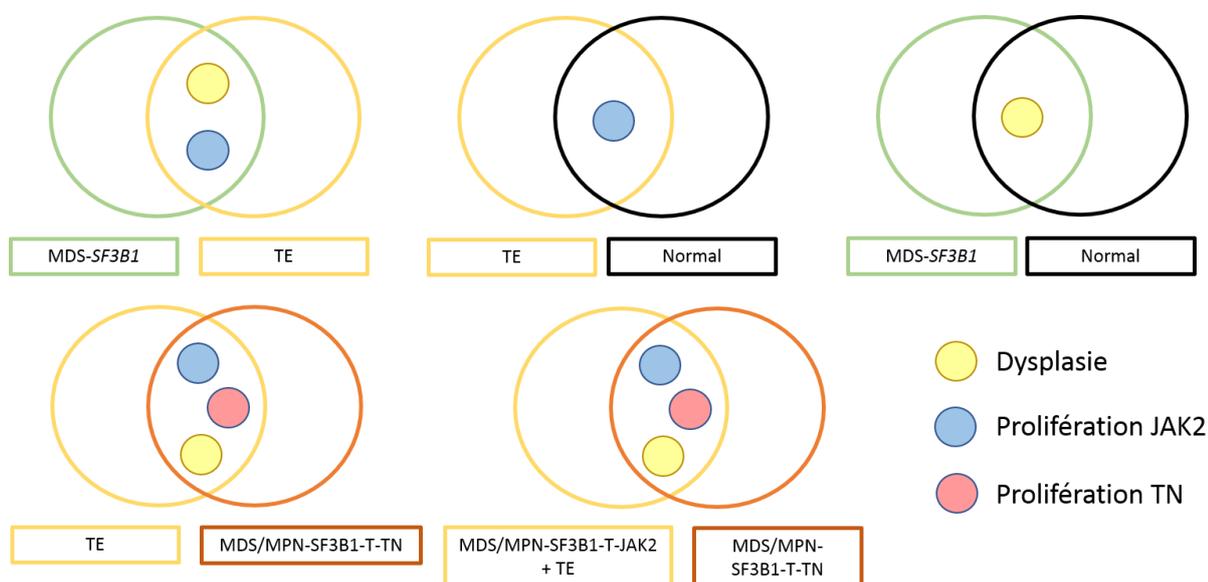


Figure 53. Justification du score de pertinence des DMRs issues de la comparaison MDS/MPN-SF3B1-T-JAK2 vs MDS/MPN-SF3B1-T-TN.

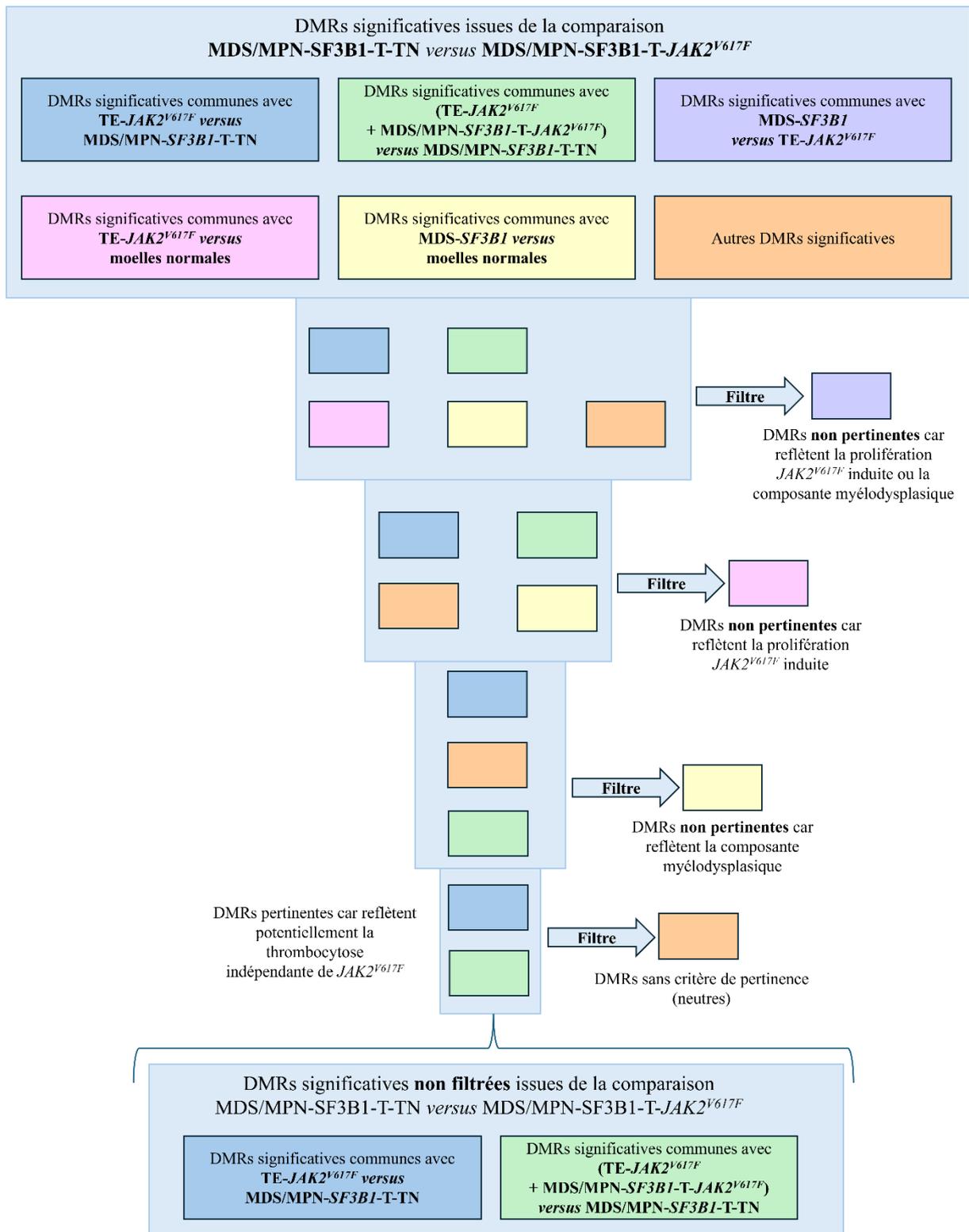


Figure 54. Représentation schématique de l'application du score de pertinence à 5/5.
 A noter que cette représentation ne tient pas compte du chevauchement des différents ensembles entre eux et qu'elle ne s'applique pas à un score cible inférieur à 5/5.

7) Analyse ciblée

Nous avons également réalisé une analyse ciblée de la méthylation (avec la méthode *in-house* pour la comparaison des DMRs) : d'une part sur 32 gènes d'intérêt impliqués dans les pathologies myéloïdes (*WT1, SETBP1, ABL1, HRAS, DNMT3A, RUNX1, CEBPA, KIT, EZH2, NPM1, STAT5B, PTPN11, PTPN12, FLT3, IDH1, IDH2, KRAS, CBL, BRAF, ASXL1, ETV6, JAK2, TET2, U2AF1, CALR, CSF3R, MPL, NRAS, SF3B1, SRSF2, TP53 et ZRSR2*) et d'autre part sur les gènes codant pour différents acteurs des voies de signalisation (voies JAK/STAT, PI3k/Akt et Ras/MAPK).

8) Gestion des réplicats

Pour évaluer l'impact des différents types cellulaires dont nous disposons pour un même patient (mononucléés (MN), polynucléés (PN)), nous les avons comparés avec *Methylkit*. Nous avons analysé 17 000 gènes pour lesquels nous disposons d'informations de méthylation pour les CpGs situés à proximité des ISS. Nous avons ensuite recherché les gènes pour lesquels il existait des DMRs avec une différence de méthylation de plus de 20%, dans le but de mettre en évidence une signature commune aux PN ou aux MN.

Même en l'absence de signature spécifique des MN ou PN, lorsque nous disposons (pour un même échantillon) de données issues de MN, PN ou matériel total, nous avons sélectionné pour chaque comparaison le type cellulaire le plus proche de celui des autres échantillons du groupe comparé.

9) qPCR des candidats issus du RRBS

La disponibilité d'ARN pour 4 échantillons (MYS19, MYS30, MYS36 et MYS53) issus de MDS-*SF3B1*, MDS/MPN-*SF3B1*-T-TN et MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et pour les 4 moelles normales nous a permis d'effectuer des qPCR afin de corrélérer les niveaux d'expression de 5 gènes candidats à la méthylation visualisée au RRBS.

L'amplification a été réalisée avec un *master mix SYBR Green* (Thermofisher, Foster City, Californie, USA) et un couple d'amorces spécifiques des candidats listés (**Tableau 13**), en triplicats. La réaction de qPCR s'est déroulée sur un appareil *7500 Fast Real System* (Applied Biosystem®, Thermofisher, Foster City, Californie, USA) et l'analyse sur le logiciel *Step One Software*. Les niveaux d'expression ont été calculés avec la méthode $2^{-\Delta\Delta Ct}$.

Tableau 13. Couples d'amorces utilisés pour la qPCR des candidats de la cohorte MYST.

Gène	sens	Séquence
<i>DUSP2</i>	<i>Forward</i>	5' CTGAGGCTGGCCCTCATTCG 3'
	<i>Reverse</i>	5' TGGTGGTGGACCAGGGAGAG 3'
<i>LMX1B</i>	<i>Forward</i>	5' CCAGCCAGAGCCTCATCTGC 3'
	<i>Reverse</i>	5' TGTGCTGGCACCTTGGTCTG 3'
<i>AATK</i>	<i>Forward</i>	5' ACATCCCGTCCTCAGCCAGT 3'
	<i>Reverse</i>	5' CCTTCCTGCGCCTCCTTGAG 3'
<i>BCLAP</i>	<i>Forward</i>	5' ACCTGGGTGTCTGGGCTCAA 3'
	<i>Reverse</i>	5' GAGCATGCACGGGAGACACA 3'
<i>SPRY2</i>	<i>Forward</i>	5' CCAGCACTCGCAGGTCCATT 3'
	<i>Reverse</i>	5' GGCCCGGAGGAGAAGGATGA 3'
<i>FLT4</i>	<i>Forward</i>	5' CGGCGTCATCGCTGTCTTCT 3'
	<i>Reverse</i>	5' CCACTGGCTGGCATCGTAGG 3'
<i>GPM6A</i>	<i>Forward</i>	5' CGGGTGTTGCCCTGTTCTGT 3'
	<i>Reverse</i>	5' GCAGCTGCGATGCCGTAGAT 3'
<i>IMPA2</i>	<i>Forward</i>	5' CCGACTGTGGCGGTTAGCAT 3'
	<i>Reverse</i>	5' CGTGTACAGCCGCTCCTCTG 3'
	<i>Reverse</i>	5' TGGGTGCTCAGACAGAAGGC 3'

RÉSULTATS

PARTIE 1 – PREMIÈRE COHORTE D'ÉCHANTILLONS : NGS IGR

L'analyse en NGS ciblé des patients de la première cohorte d'échantillons a montré la présence de variant de *SF3B1* dans 93% des cas (**Figure 55**). Conformément aux données de la littérature (130,131), d'autres anomalies affectant les acteurs du spliceosome ou la méthylation étaient aussi retrouvées : 21% des patients présentaient des variants de *DNMT3a*, 21% des variants de *TET2*, 21% des variants d'*ASXL1* et 21% des variants d'*U2AF1*, soit au moins une anomalie par cas, plusieurs dans certains cas.

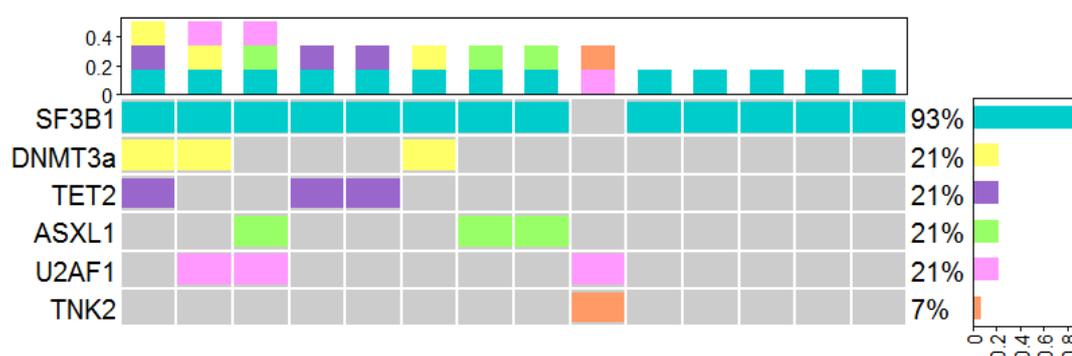


Figure 55. Séquençage ciblé de 15 échantillons de patients avec MDS/MPN-*SF3B1*-T TN.

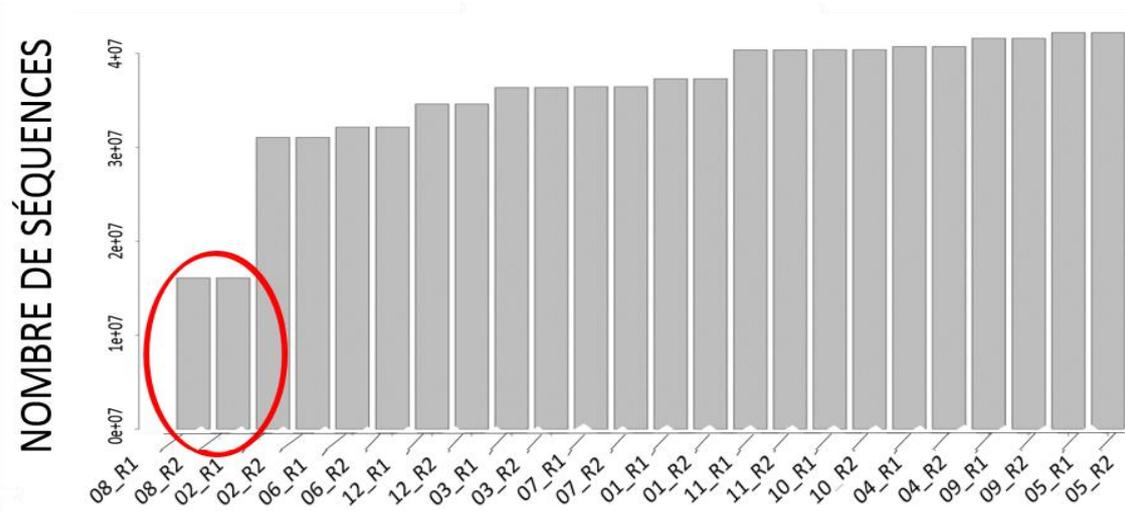
PARTIE 1 - DEUXIÈME ET TROISIÈME COHORTES : COHORTE DE SCREENING ET COHORTE DE VALIDATION

1) Analyses primaires WES de la deuxième cohorte (cohorte de screening)

Nous disposons de 12 échantillons numérotés de 01 à 12, provenant de six patients (deux échantillons par patient : un échantillon de granulocytes (numéros impairs) et un échantillon de lymphocytes T circulants (numéros pairs)).

En moyenne, **72 millions** de *reads* R1+R2 (*reads* appairés après séquençage en *paired end*) ont été produits pour chaque échantillon. Le nombre de *reads* était nettement plus faible pour l'échantillon de lymphocytes T 08 (**Figure 56a**). Une quantité non négligeable de *reads* (7,8%) ont été séquencés avec une qualité moyenne < 33 (**Figure 56b**) et ont donc été marqués et filtrés lors des étapes suivantes.

a.



b.

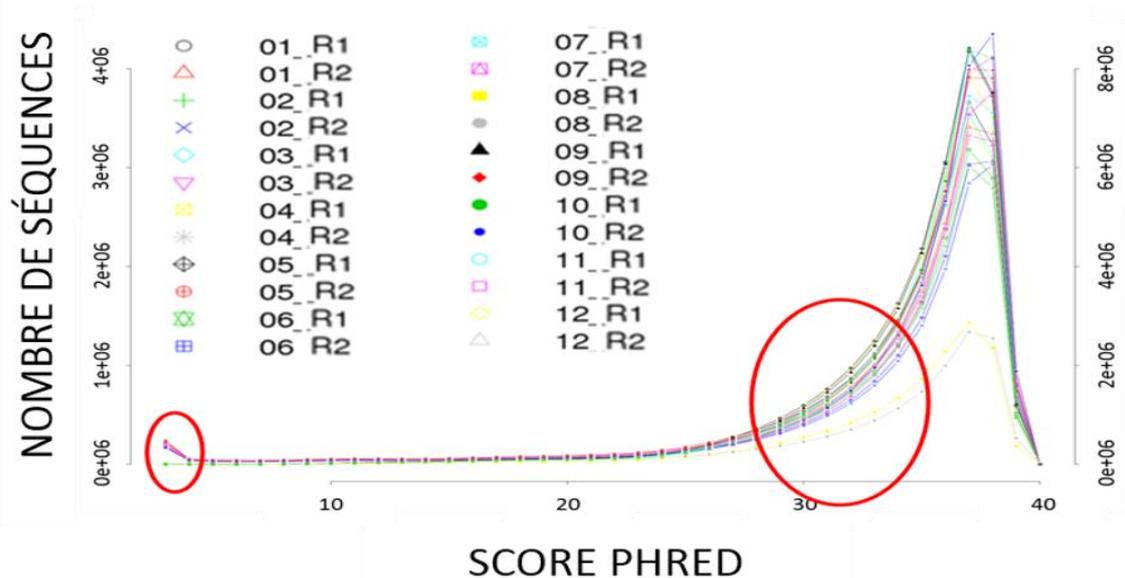


Figure 56. Contrôles qualité du séquençage de l'exome.

a. Nombre de *reads* R1 et R2 par échantillon. **b.** Nombre de *reads* en fonction de la qualité moyenne des bases incorporées. Cerclés en rouge à gauche, quelques *reads* avec des scores de qualité très faibles. Le cercle de droite montre que la plupart des *reads* avec des scores de qualité inférieurs à 33 restent cependant avec des scores aux alentours de 30, ce qui correspond à un risque d'erreur de 0,1%.

2) Analyses secondaires WES de la deuxième cohorte

a) *Alignement des reads, les fichiers «.sam » :*

Après alignement sur le génome de référence Hg19, 2,78% des paires R1-R2 n'étaient pas alignés, **79,57% des *reads* étaient alignés une fois exactement** et 17,81% étaient alignés plus d'une fois (correspondance possible avec plusieurs positions sur Hg19). La plupart de ces *reads* qui présentent plusieurs "*hits*" correspondent à des régions répétées ou à des pseudogènes. Les *reads* non alignés peuvent provenir de diverses contaminations.

b) *Profondeur et couverture*

La **profondeur moyenne** du séquençage rapportée en kilobase d'exons était de 283. Une différence de couverture a été observée entre les fractions CD3- (granulocytes) et les CD3+ (lymphocytes T) (**Figure 57**), imposant une normalisation lors de l'analyse tertiaire.

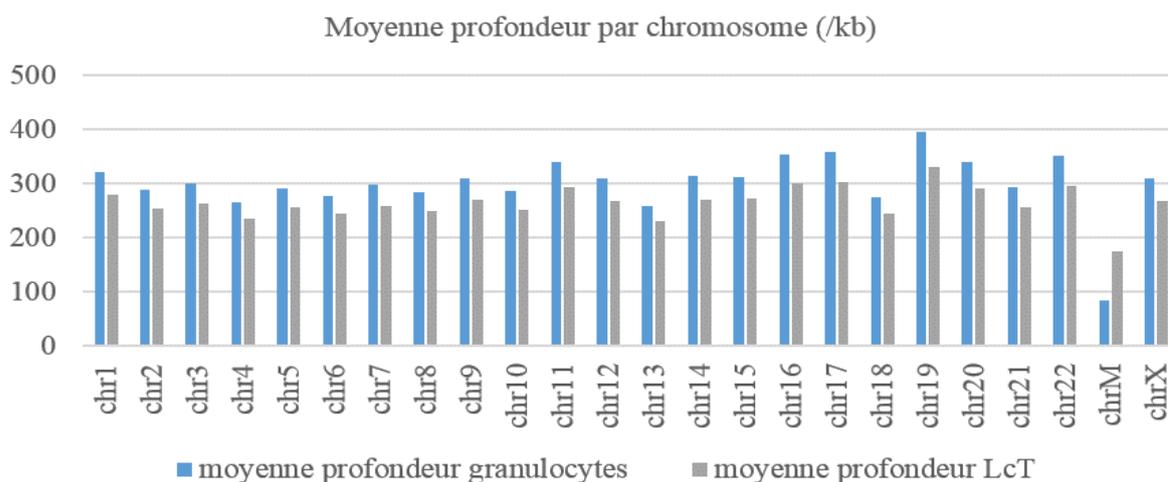


Figure 57. Profondeur moyenne du séquençage par chromosome et par kilobase. LcT = lymphocytes T

La profondeur était nettement plus faible pour l'ADN mitochondrial des CD3-. Le pourcentage moyen de couverture des exons était de **74%** avec une forte hétérogénéité de couverture (**Figure 58**), surtout pour les gènes mono-exoniques à petits exons. La région centrale des très grands exons était généralement mal couverte. **194 457** positions variantes ont été appelées sur l'ensemble des échantillons.

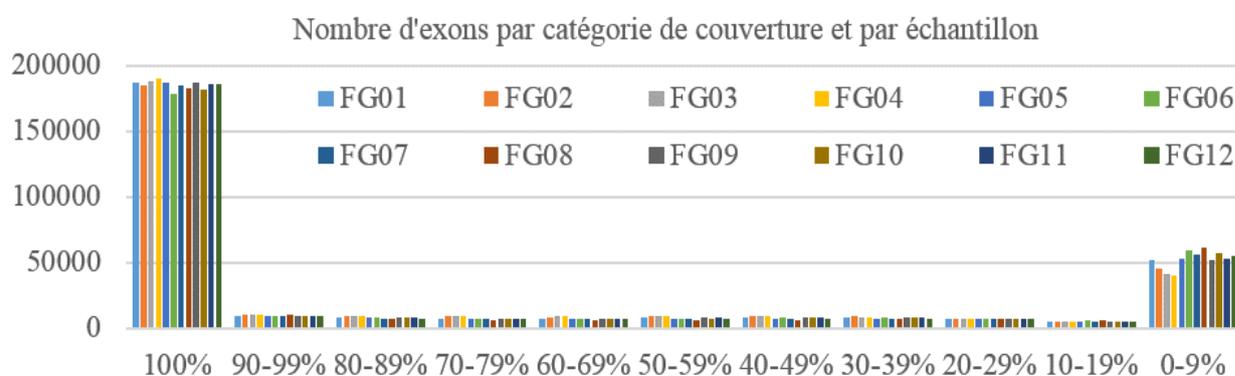


Figure 58. Nombre d'exons par catégorie de couverture et par échantillon. La grande majorité des exons sont couverts à 100% mais une catégorie d'exon reste mal couverte à moins de 10%. Il s'agit des exons de taille extrême (très grands exons ou très petits exons).

3) Analyses tertiaires WES de la deuxième cohorte (cohorte de screening)

a) Identification d'un biais de transversion

Plus de 80% des variants identifiés par l'exome des 6 patients MDS/MPN-SF3B1-T TN correspondent à des SNV, dont la majorité étaient des transversions C vers A (**Figure 59**). Ces résultats sont retrouvés dans l'analyse globale incluant tous les échantillons mais également dans l'analyse individuelle. Ils ne sont cependant pas retrouvés dans la cohorte de confirmation.

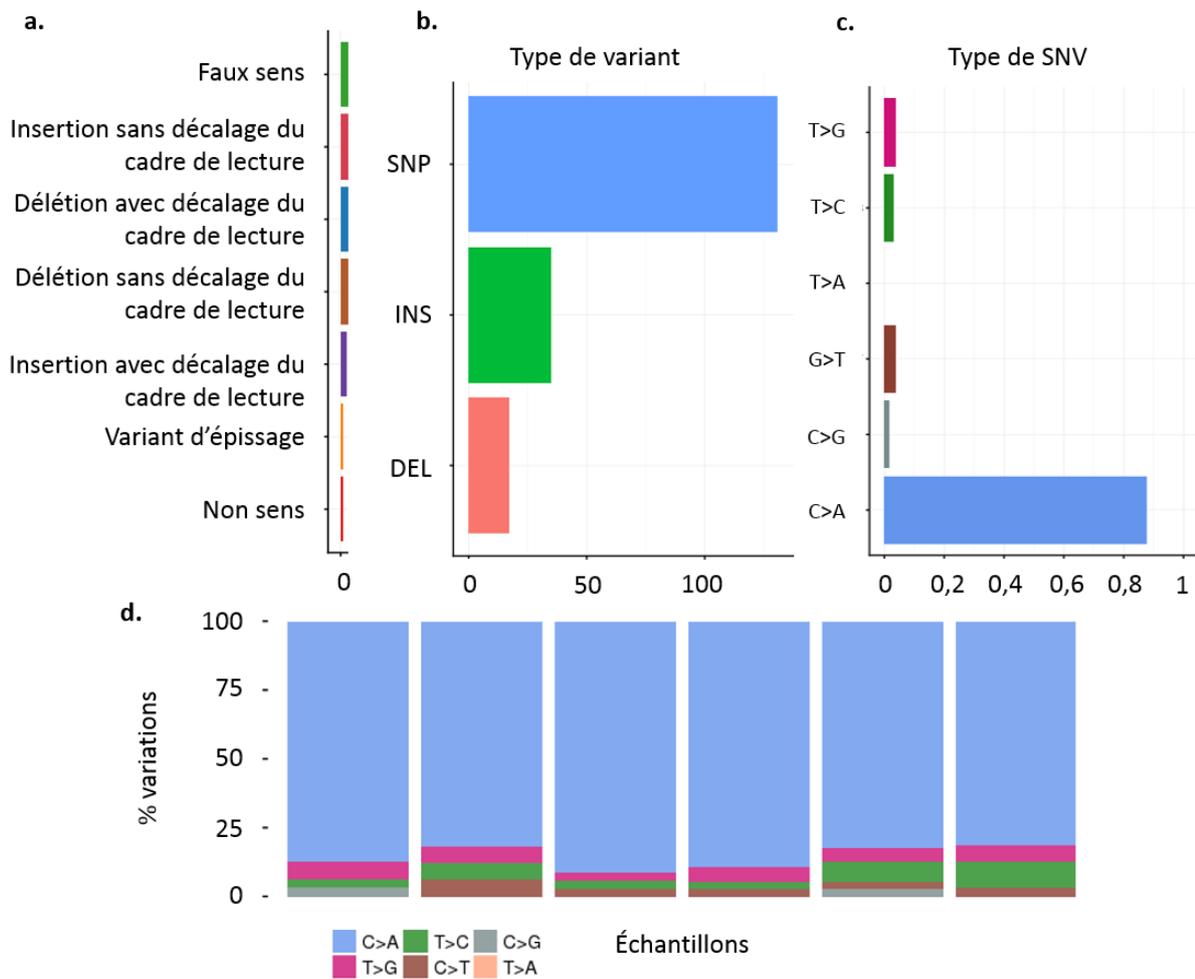


Figure 59. Types de variants identifiés dans l'analyse d'exome de six patients porteurs d'un MDS/MPN-SF3B1-T triple négatif.

a. conséquences des variant. **b.** Type de variants sur l'analyse globale de tous les patients. SNP = *single nucleotide polymorphism*, INS = insertion, DEL = délétion. **c.** Catégorie de SNV pour tous les variants. **d.** Catégorie de SNV par échantillon.

Devant la présence d'un biais de transversion C vers A, les données d'exome concernant les gènes codant pour les protéines du système BER (*base excision repair*) (réparation simple brin de l'ADN), impliquant la protéine MUTYH (*mutY DNAGlycosylase*), ont été spécifiquement analysées.

La protéine MUTYH code pour une ADN glycosylase impliquée dans la réparation des dommages oxydatifs de l'ADN, via le système BER (**Figure 60**). Elle joue également un rôle dans la signalisation de l'apoptose par l'introduction de coupures simple brin. Les mutations germinales de ce gène sont décrites comme prédisposant au cancer colorectal (polypose associée à *MUTYH*) (155,156).

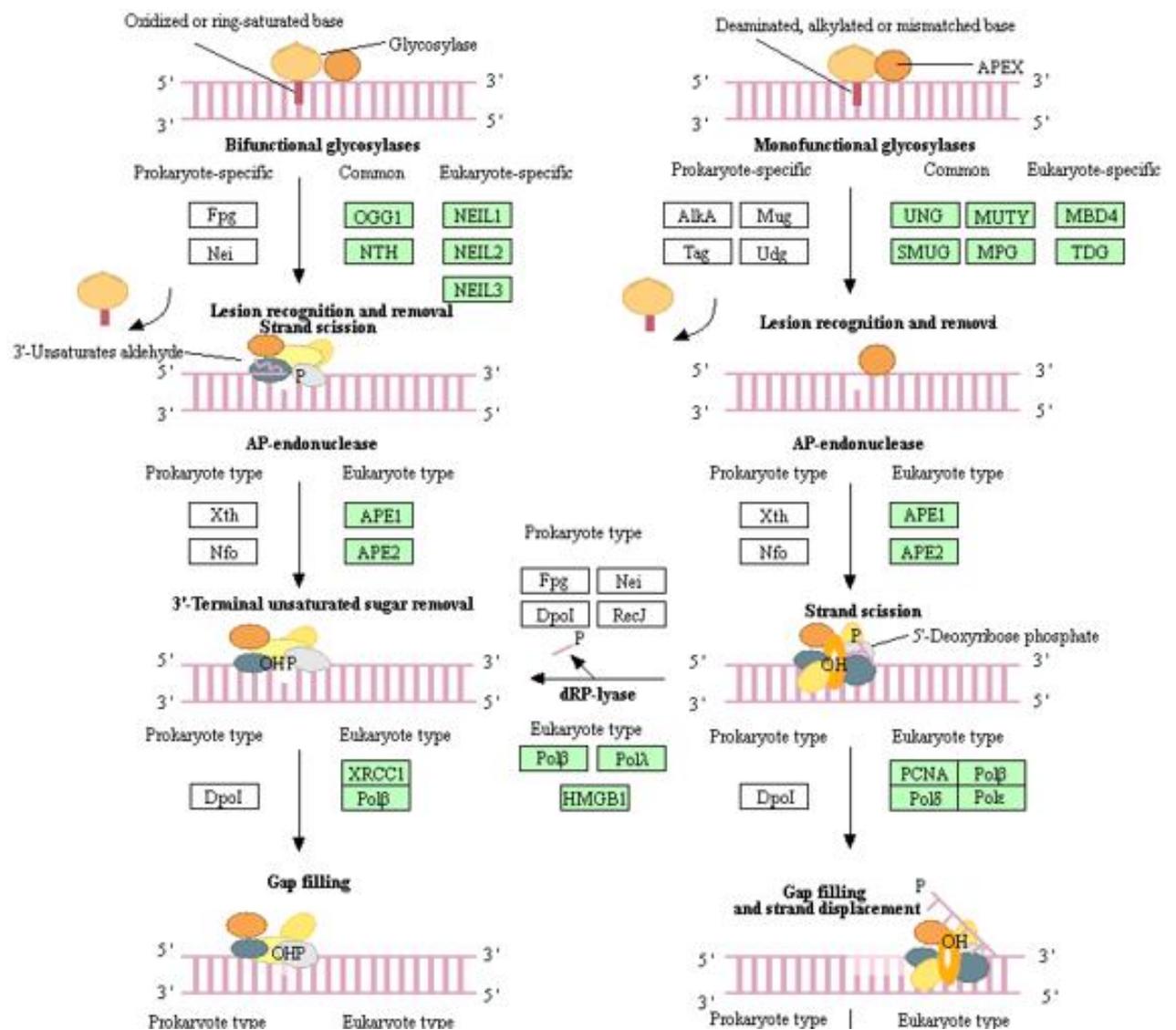


Figure 60. Système de réparation BER selon la base de données KEGG (Kyoto Encyclopedia of Genes and Genomes).

L'analyse d'exome n'a retrouvé aucun variant de type somatique (présent dans la fraction CD3- et absent dans la fraction CD3+) ou d'allure constitutionnelle au sein du système MUTYH pouvant expliquer ce résultat.

b) *Identification de gènes « candidats » touchés par les variants somatiques du WES*

À l'issue de l'application des filtres « paramétriques », 1210 positions multi-alléliques ont été retenues, ce qui correspond à **1311 variants bi-alléliques** après décomposition/normalisation.

L'application des filtres sémantiques a sélectionné **53 variants significatifs**, soit **49 gènes touchés par un ou plusieurs variants (Figure 61, Annexe 2)**. Parmi ces 49 gènes, *SF3B1* était atteint par un ou des variants pour l'ensemble des patients. Aucun patient ne présentait de variant significatif de *JAK2*, *CALR* ou *MPL* (Figure 62). Les fréquences alléliques déterminées en routine au diagnostic et lors de notre analyse bioinformatique étaient similaires pour *SF3B1*. Les analyses ont également mis en évidence des variants au niveau du gène *TET2*.



Figure 61. Représentation des 49 gènes candidats retenus après l'application du pipeline bioinformatique.

L'ensemble des 49 gènes significatifs était enrichi en gènes impliqués dans la signalisation, notamment des MAPK (*FGF18*, *MLK4*, *TRPV3*, *MAPK7*), des protéines G (*ADRA2C*, *PLEKHG3*, *HTR6*, *LPHN1*, *DEF6*, *UBXN11*, *CNGB1*) et de l'apoptose (*POU4F1*, *CRIPAK*). Cet ensemble comporte également des gènes codant pour les facteurs Trans CEBPZ (déjà décrit dans les LAM) et *ETV2* (impliqué dans la régulation de PDGFRa).

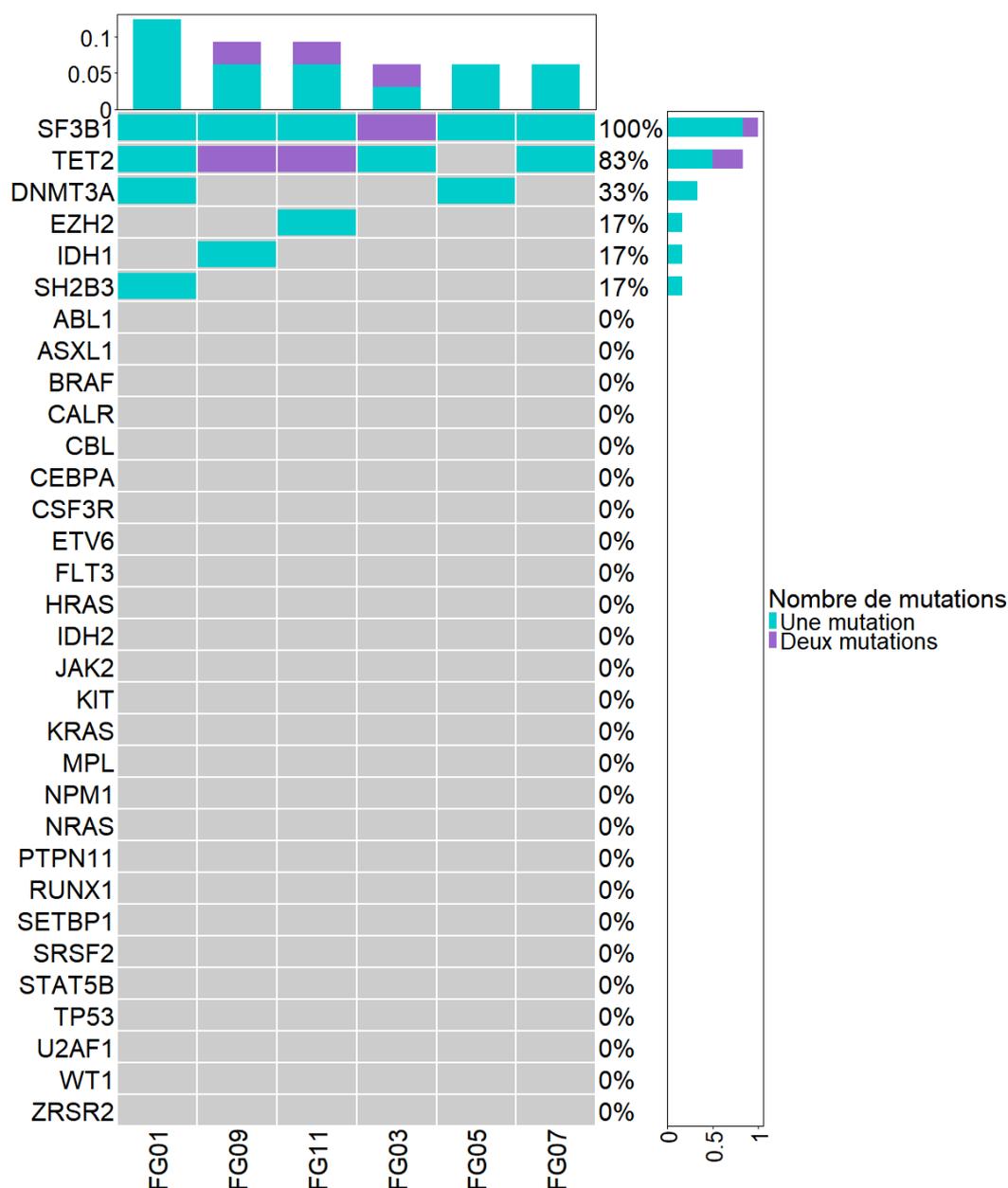


Figure 62. Exome ciblé de la deuxième cohorte (cohorte de screening, n=6 patients, panel similaire à celui utilisé pour la première cohorte).

Certains gènes impliqués dans la signature moléculaire des MDS/MPN-*SF3B1*-T n'ont pas été retenus après application des filtres (variants non exoniques, variants retrouvés dans les lignées CD3+ et CD3- et non considéré comme « somatique », variants somatiques retrouvés uniquement pour un patient). Le profil mutationnel des patients inclus, sur un panel similaire à celui effectué sur la première cohorte d'échantillons montrait des variants de *DNMT3a* dans 33% des cas, de *TET2* dans 83% des cas, d'*IDH1* dans 17% des cas, d'*EZH2* dans 17% des cas et *SH2B3* dans 17% des cas (**Figure 62**).

c) *L'analyse des CNV révèle une duplication de PTPN12*

L'analyse des CNV sur la deuxième cohorte (exome complet de 6 patients) a révélé un gain de la région 77 167 346 – 77 230 282, contenant le gène *PTPN12* pour 3 patients, soit 50% de l'effectif.

L'analyse des données de qPCR de 11 ADN de TE TN et 4 ADN de MF TN, montre qu'un échantillon de TE TN présente une duplication de *PTPN12* (**Figure 63**).

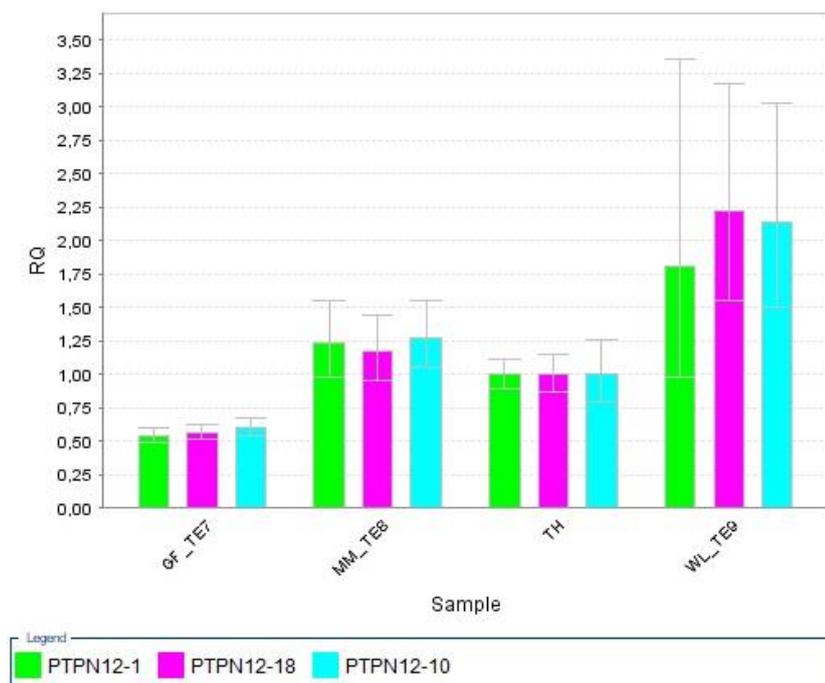


Figure 63. qPCR de *PTPN12* pour trois échantillons de TE TN (TE7, TE8 et TE9). Amplification détectée pour le patient TE9 sur les trois amorces de *PTPN12* par rapport au témoin négatif et au gène de référence (*RPPH1*). TH = Témoin Homme.

4) Troisième cohorte (cohorte de validation)

Un panel de 46 gènes candidats a été conçu pour l'analyse génomique ciblée de la troisième cohorte, comportant (i) des gènes retenus à l'issue du pipeline d'analyse de la cohorte de screening (2^{ème} cohorte) et (ii) des gènes particulièrement pertinents sur le plan fonctionnel ayant été filtrés à la dernière étape du pipeline (**Annexe 2**).

Le panel comportait les gènes suivants : *ADRA2C, ANO7, ARHGEF39, ATXN1, BMP4, BRAT1, CBL, CDK6, CEBPZ, COPG1, DCP1B, DCP1B, DOT1L, DTX1, EPHA10, ETV2, FGF18, HCN3, HTR6, LMTK3, LPHN1, LTK, LURAP1L, MAG11, MAML2, MAPK7, MLK4, NFKBIE, NUTF2, OGF, PHLPP1, PLEKHG3, POU4F1, POU4F2, SATB1, SGK223, SH2B3, SIRT7, STAT5B, TET2, TNFAIP6, TNFRSF19, TRIO, TRPV3, UBXN11, WDR37.*

Les fréquences alléliques des variants des autres gènes variaient de 2 à 10 %. Après bibliographie et recherche des données disponibles sur ces 31 gènes, nous avons sélectionné un panel final de 6 candidats nous paraissant particulièrement pertinents (Erreur ! Source du renvoi introuvable.) :

- **BRATI** : gène codant pour une protéine impliquée dans la prolifération, la progression cellulaire, la réparation des anomalies double-brin et en lien avec mTOR. Présence de quatre variants faux-sens (soit 9,52% d'occurrence) dans la cohorte de validation.

- **LMTK3** : gène codant pour une protéine tyrosine kinase impliquée dans les voies de signalisation des MAPK et NANOG. Activateur de FOXO3 et elle a des interactions avec MYC via un intermédiaire. Présence de trois variants faux-sens et une délétion (soit 9,52% d'occurrence) dans la cohorte de validation.

- **LTK** : gène codant pour une protéine de type récepteur couplé à une tyrosine kinase, dont la fonction est mal connue mais en lien avec la prolifération. Ses variants sont associés à des phénotypes d'augmentation du nombre de granulocytes. Elle interagit directement avec CBL, PI3K et PTPN1. Présence de trois variants faux-sens (soit 7,14 % d'occurrence) dans la cohorte de validation.

- **POU4F1** : gène codant pour une protéine impliquée dans les phénomènes d'apoptose. Elle est associée à la régulation de TP53 et interagit avec BCL2 (*B-cell lymphoma 2*) et BAX (*BCL2 associated X*). Présence de deux variants faux-sens et une insertion (soit 9,52% d'occurrence) dans la cohorte de validation.

- **SGK223** : gène codant pour une protéine co-activatrice de la voie NOTCH. Présence de trois variants faux-sens (soit 9,52% d'occurrence) dans la cohorte de validation.

- **STAT5B** : gène codant pour une protéine impliquée dans la voie JAK/STAT. Elle interagit directement avec l'érythropoïétine et la thrombopoïétine. Présence de deux variants faux-sens (soit 4,76 % d'occurrence) dans la cohorte de validation.

PARTIE 1 – ANALYSES FONCTIONNELLES

1) La mutation c.2117 A>T entraîne un saut d'exon de *STAT5B*.

La cohorte de validation retrouve deux variants de *STAT5B* (4,76% des cas) :

- Le variant Y665F dans le domaine SH2 de *STAT5B*.
- Le variant c.2117 A>T (décrit comme Q706L dans la littérature) du domaine transactivateur de *STAT5B*.

Les analyses *in silico* selon *Ex-skip* (<http://www.ex-skip.img.cas.cz>) pour le variant c.2117 A>T ont montré la création potentielle de site alternatifs pour la fixation des facteurs d'épissage SC35 (SRSF2), hnRNPA1 and SRp40 (**Figure 65**).

<i>STAT5B</i> : p.Gln706Leu : Altération significative des motifs ESE/ESS		
Modifications d'épissage	Position	Séquence
EIE (Cassure site ESE)	chr17: 42202769	AAGTGG
IIE (Nouveau site ESS)	chr17: 42202769	TAGTGG
Fas ESS (Nouveau site ESS)	chr17: 42202769	TAGTGG
ESE-SRp40 (Nouveau site ESE)	chr17: 42202770	CTAGTGG
EIE (Cassure site ESE)	chr17: 42202771	GCAAGT
ESE-SC35 (Nouveau site ESE)	chr17: 42202772	AGCTAGTG
PESE (Cassure site ESE)	chr17: 42202772	AGCAAGTG

Figure 65. Analyse prédictive *in silico* du variant c.2117 A>T selon *Ex-skip*.

Lors du test du minigène (**Figure 66**), le variant c.2117 A>T induit un épissage aberrant de l'exon 17. La surexpression de SC35 et SRp40 entraîne l'élimination de l'exon 17 pour la construction sauvage et majore l'effet du variant. La surexpression de hnRNPA1 n'affecte pas le profil d'épissage.

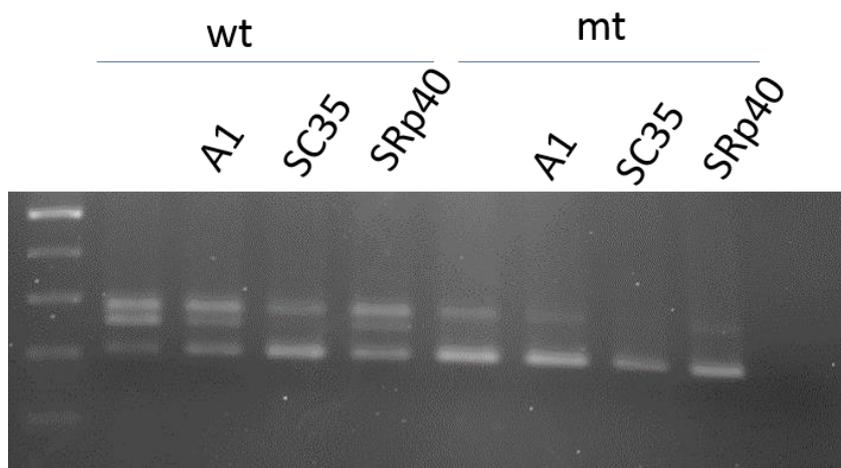


Figure 66. Minigène *STAT5B*.

Wt = Wild Type; mt = muté. A1 = surexpression de hnRNPA1; SC35 = surexpression de SC35; SRp40 = surexpression de SRp40.

L'épissage de l'exon 17 de l'ARNm de *STAT5B* entraîne un décalage du cadre de lecture, qui peut conduire à la synthèse d'une protéine tronquée au niveau du codon 695.

La surexpression de la forme tronquée ou sauvage de *STAT5B* dans des cellules HEK293 a montré une localisation identique des deux protéines (nucléaire et cytoplasmique).

2) Surexpression de *PTPN12* dans un modèle cellulaire HEK293.

La surexpression de *PTPN12* n'augmente pas la phosphorylation des protéines ERK et AKT dans le modèle cellulaire HEK293 (**Figure 67**).

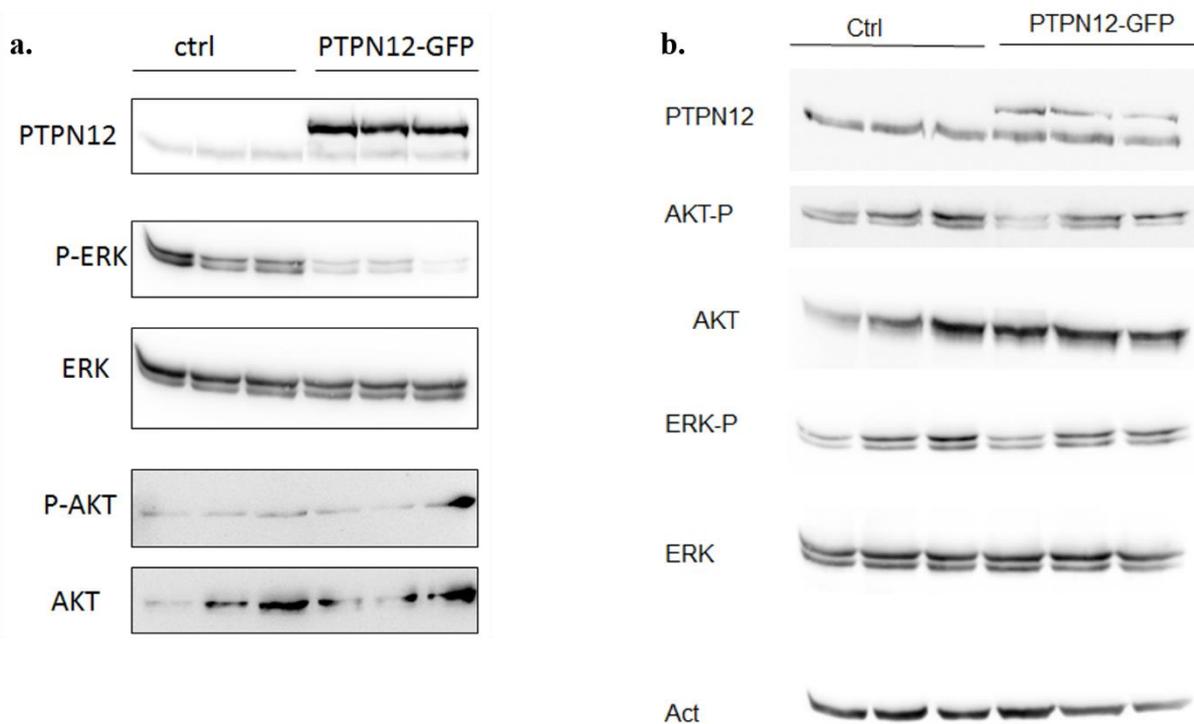


Figure 67. Expression de ERK, Phospho-ERK, AKT, Phospho-AKT selon la surexpression de *PTPN12* dans le modèle HEK293.

(contrôle (Ctrl, pEGFP) *versus* cellules transfectées PTPN12-GFP).

a. Aspect de réduction de la phosphorylation lorsque la protéine PTPN12 est surexprimée.
b. Absence de différence significative d'expression avec utilisation de l'actine comme protéine de référence.

PARTIE 2 - QUATRIÈME COHORTE – COHORTE MYST

1) Collection biologique et données cliniques

Nous avons réuni un total de 48 échantillons issus de 39 patients (**Tableau 14**) : 10 TE-*JAK2^{V617F}*, 12 MDS/MPN-*SF3B1*-T TN, 12 MDS/MPN-*SF3B1*-T-*JAK2^{V617F}*, 10 MDS-*SF3B1* et 4 moelles normales.

Tableau 14. Données cliniques des 48 échantillons (39 patients) de la cohorte MYST.

	MDS- <i>SF3B1</i>	MDS/MPN- <i>SF3B1</i> -T TN	MDS/MPN- <i>SF3B1</i> -T- <i>JAK2^{V617F}</i>	TE	Moelles Normales
N (échantillons/patients)	10/8	12/10	12/7	10/10	4/4
Age au diagnostic (années)					
<i>Médiane</i>	70,5	76,5	71	53,5	73
<i>Min-Max</i>	62-88	58-92	62-80	23-79	58-81
Genre, n (%)					
<i>Masculin</i>	5 (63%)	4 (40%)	6 (86%)	3 (30%)	1 (25%)
<i>Féminin</i>	3 (37%)	6 (60%)	1 (14%)	7 (70%)	3 (75%)
Origine des échantillons, n (%)					
<i>Moelle osseuse</i>	10(100%)	5 (42%)	7 (38%)	1 (10%)	4 (100%)
<i>Sang</i>	0 (0%)	7 (38%)	5 (42%)	9 (90%)	0 (0%)
Type d'échantillon					
<i>MN</i>	6 (60%)	3 (25%)	5 (42%)	0 (0%)	0 (0%)
<i>PN</i>	4 (40%)	6 (50%)	3 (25%)	0 (0%)	0 (0%)
<i>Total</i>	0 (0%)	3 (25%)	4 (33%)	10 (100%)	4 (100%)
Evènement thrombotique, n (%)					
<i>Oui</i>	2 (25%)	0	1 (14%)	4 (40%)	0
<i>Non</i>	6 (75%)	7 (70%)	5 (72%)	6 (60%)	4 (100%)
<i>Donnée manquante</i>	0	3 (30%)	1 (14%)	0	0
LAM secondaire, n (%)					
<i>Oui</i>	1 (12%)	1 (10%)	1 (14%)	0	0
<i>Non</i>	7 (88%)	6 (60%)	5 (72%)	10 (100%)	4 (100%)
<i>Donnée manquante</i>	0	3 (30%)	1 (14%)	0	0
myélofibrose secondaire, n (%)					
<i>Oui</i>	0	0	3 (43%)	0	0
<i>Non</i>	8 (100%)	7 (70%)	4 (57%)	10 (100%)	4 (100%)
<i>Donnée manquante</i>	0	3 (30%)	0	0	0
Traitements reçus, n (%)					
<i>EPO</i>	6 (75%)	6 (60%)	4 (57%)	0	0
<i>Transfusions</i>	8 (100%)	7 (70%)	4 (57%)	0	0
<i>Lenalidomide</i>	2 (25%)	1 (10%)	1 (14%)	0	0
<i>Luspatercept</i>	0	0	0	0	0
<i>Hydroxyurée</i>	0	4 (40%)	3 (43%)	4 (40%)	0
<i>Azacytidine</i>	3	1 (10%)	0	0	0
<i>Interferon</i>	0	0	0	4 (40%)	0
<i>Antiagrégants</i>	4 (50%)	5 (50%)	2 (29%)	10 (100%)	0
<i>Anticoagulants</i>	2 (25%)	1 (10%)	2 (29%)	0	0
<i>Donnée manquante</i>	0	3 (30%)	2 (29%)	0	0

2) Caractérisation moléculaire

La caractérisation moléculaire par NGS avec un panel myéloïde de routine est en cours.

3) Contrôles qualité

La bisulfitation moyenne était supérieure à 99% pour tous les échantillons. Le pourcentage de méthylation des C dans un contexte de CpG était comparable d'un échantillon à l'autre.

Après *trimming* et filtrage, il n'y avait pas de *reads* d'une longueur inférieure à 30 bases (**Figure 68a**). Les *reads* dont les bases avaient un score *Phred* inférieur à 30 ont été filtrés (**Figure 68b**).

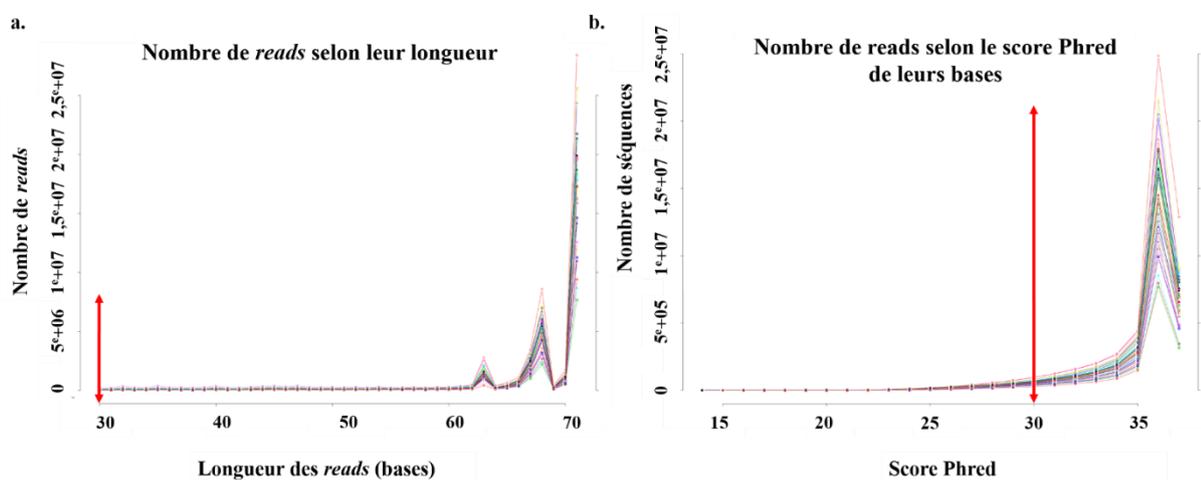


Figure 68. a. Nombre de *reads* issus du RRBS selon leur longueur. b. Nombre de *reads* issus du RRBS selon le score *Phred* de leurs bases.

Après alignement selon *Bismark*, en moyenne 64,9% des *reads* de chaque échantillon étaient alignés 1 fois, 12,3% des *reads* ne s'alignaient pas du tout et 22,8% des *reads* s'alignaient plus d'une fois. Cet alignement multiple peut s'expliquer par le caractère homologue des séquences Cis riches en CpGs régulatrices de la transcription.

4) Prise en compte des réplicats

Parmi les 17 000 gènes étudiés afin de rechercher une signature de méthylation spécifique aux cellules mononucléées ou polynucléées, 32 seulement permettaient de différencier PN et MN pour chacun des patients. Les analyses par simulation numérique nous ont permis de conclure qu'il n'existait pas de signature significative différenciant MN et PN. Nous avons donc poursuivi les analyses en estimant qu'il n'y avait pas de biais entre MN et PN.

5) Méthylation globale des 48 échantillons

L'analyse des 48 échantillons n'a pas montré de différence significative de méthylation globale entre les groupes (**Figure 69**).

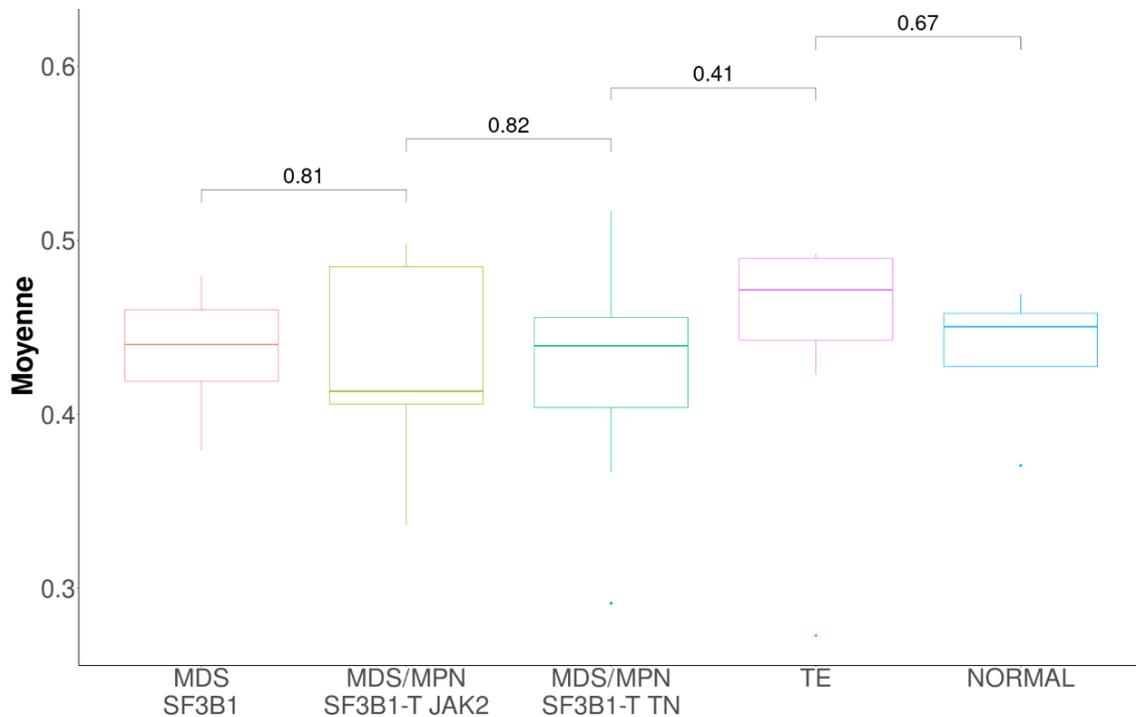


Figure 69. Méthylation moyenne de l'ADN par échantillon et par groupe de pathologie (analyses globales).

6) Analyse de la méthylation des 48 échantillons selon la méthode *MethylKit*

a) *Analyse globale des CpGs selon MethylKit*

L'analyse globale des CpGs selon *Methylkit* a montré 63 962 CpGs différenciellement méthylés entre les MDS/MPN-SF3B1-T TN et les MDS/MPN-SF3B1-T-JAK2^{V617F} (**Figure 70**).

Les annotations de gènes associées aux CpGs les plus significatifs étaient : *RAPH1*, *SATB2*, *MIR4656*, *HOXA3*, *DYNCH11*, *GRIN3A*, *C10orf105*, *C10orf71* et *RBL1*.

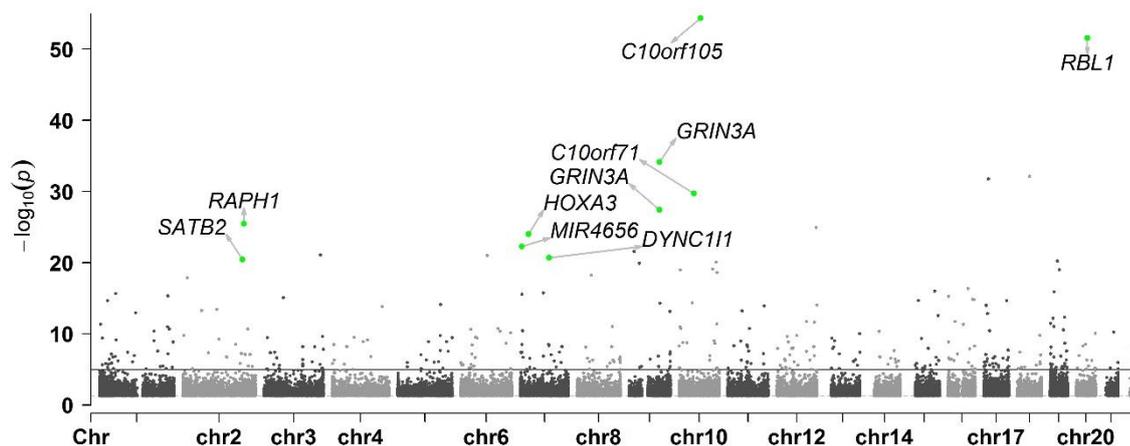


Figure 70. Manhattan plot des CpGs significativement différentiellement méthylés entre les MDS/MPN-*SF3B1*-T TN et les MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} sur l'analyse selon *MethylKit*.

b) *Analyse globale des DMRs selon MethylKit*

L'analyse globale des DMRs selon *Methylkit* a montré 2945 DMRs significativement différentiellement méthylées sur 22 599 DMRs analysées entre les MDS/MPN-*SF3B1*-T TN et les MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} (**Figure 71**). Les annotations de gènes associées aux DMRs les plus significatives étaient : *SPTBN1*, *BHLHE40*, *E2F3*, *CLCN2*, *USP54*, *LTA4H*, *MPRIP* et *CAMTA2*.

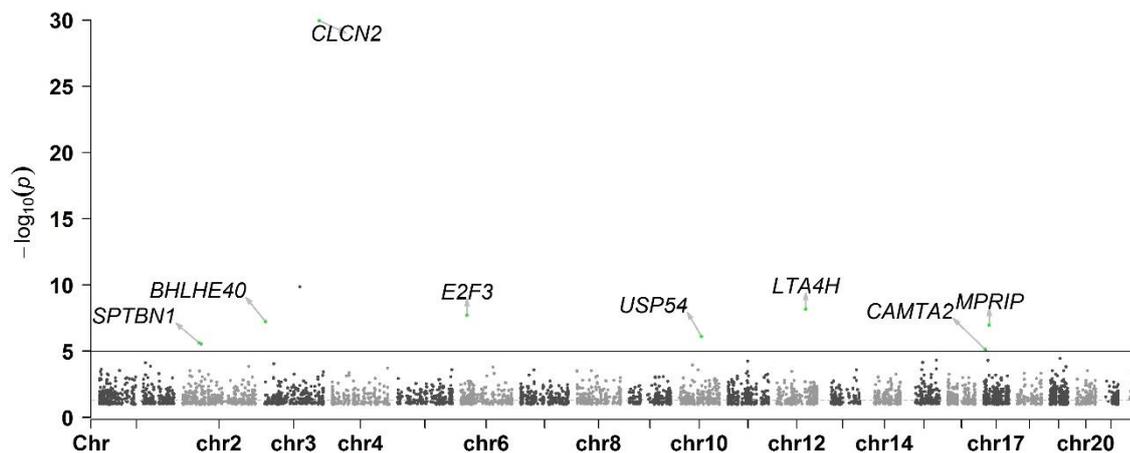


Figure 71. Manhattan plot des DMRs significativement différentiellement méthylées entre les MDS/MPN-*SF3B1*-T TN et les MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} selon *MethylKit*.

7) Analyse de la méthylation selon la méthode *in-house*

a) *Analyse globale des CpGs selon la méthode in-house*

La comparaison MDS/MPN-*SF3B1*-T TN *versus* MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} selon la méthode *in-house* retenaient 81 510 CpGs présentant une différence de méthylation significative d'au moins 5%. L'analyse non supervisée de type PCA (*Principal Component Analysis*) de ces CpGs ne permettait pas de regrouper les échantillons selon le groupe dont ils étaient issus (**Figure 72**).

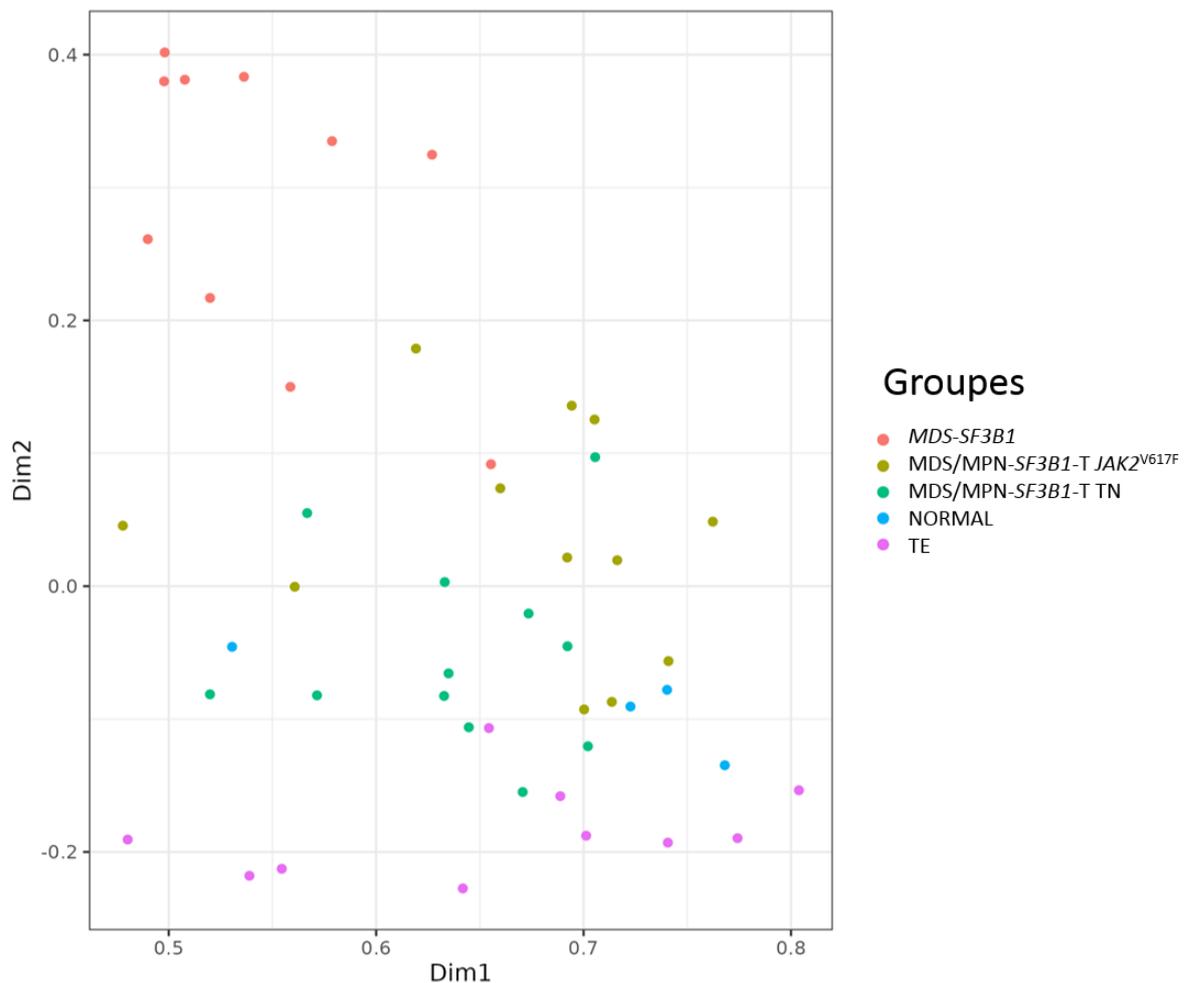


Figure 72. Analyse non supervisée par PCA des CpGs significatifs du RRBS avec un seuil de différence de méthylation à 5%.

L'analyse non supervisée de type PCA et heatmap des 15 000 CpGs les plus variants dans la comparaison MDS/MPN-*SF3B1*-*JAK2* *versus* MDS/MPN-*SF3B1*-TN ne permettait pas de regrouper clairement les échantillons MDS/MPN-*SF3B1*-T-TN et de les distinguer des autres groupes (**Figure 73**, **Figure 74**).

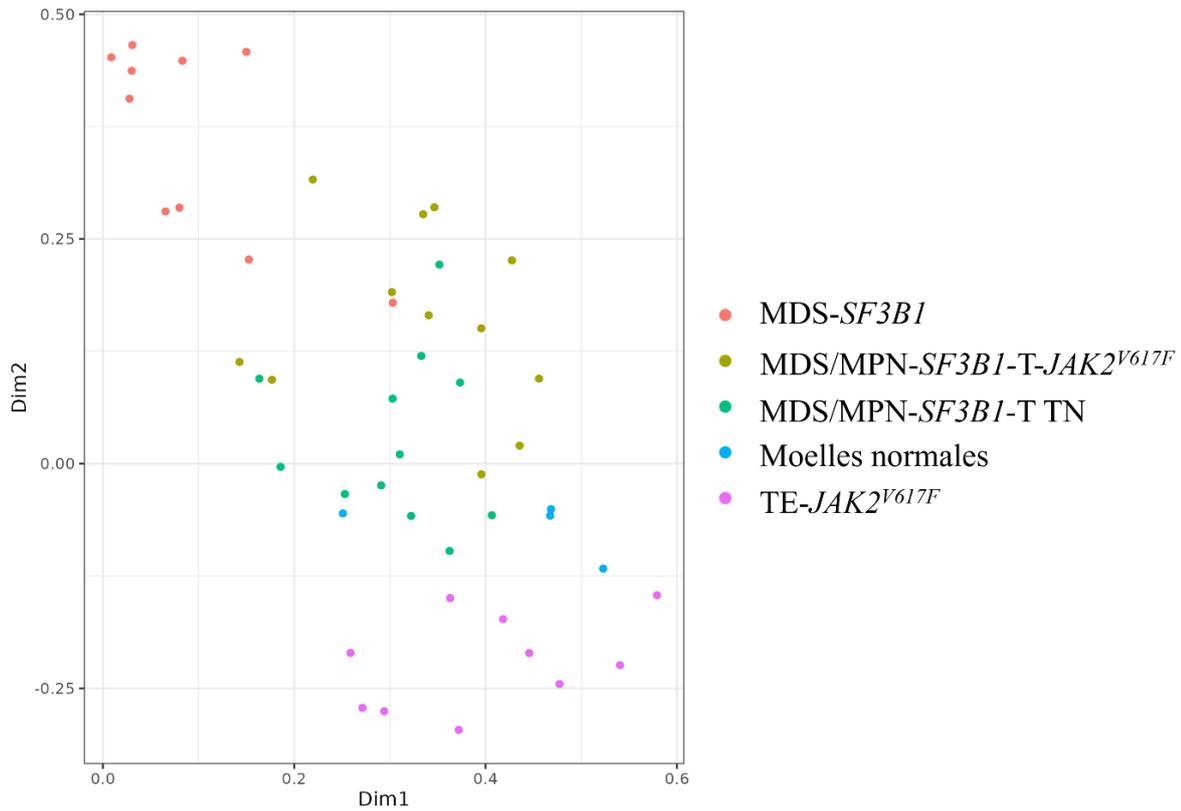


Figure 73. Analyse non supervisée par PCA des 15 000 CpGs les plus variants après analyse du RRBS.

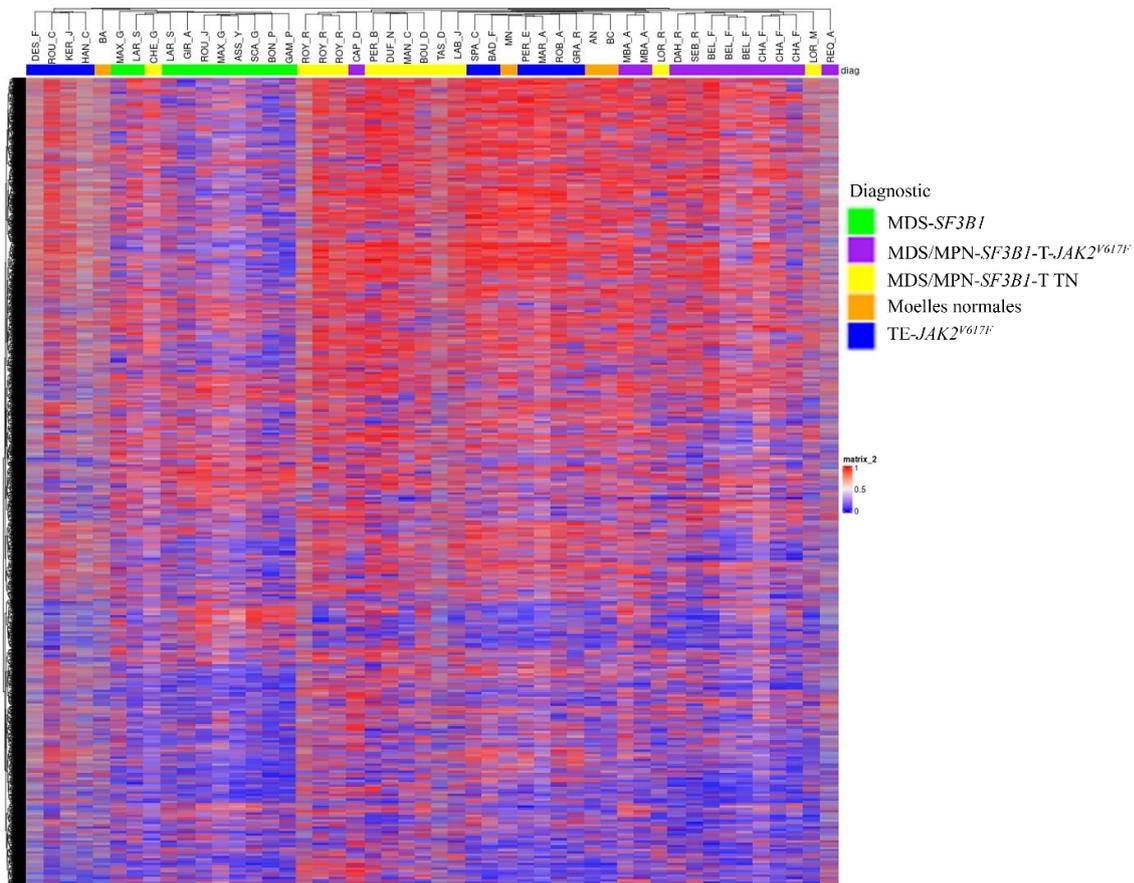


Figure 74. Heatmap des 15 000 CpGs les plus variants après analyse du RRBS.

L'utilisation d'un seuil de différence de méthylation à 20% dans la méthode *in-house* a permis d'identifier **6 246 CpGs** significativement différenciellement méthylés entre les groupes MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} et MDS/MPN-*SF3B1*-T-TN (**Figure 75**) :

- **3 849 hyperméthylées** dans le groupe MDS/MPN-*SF3B1*-T-TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}.
- **2 297 hypométhylées** dans le groupe MDS/MPN-*SF3B1*-T-TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}.

Les annotations de gènes associées aux CpGs les plus significatifs étaient : *G3BP1*, *JAKMIP3*, *SLC6A15*, *CHGA*, *PPIL2*, *RAB5IF* et *CLMN* (**Figure 75**).

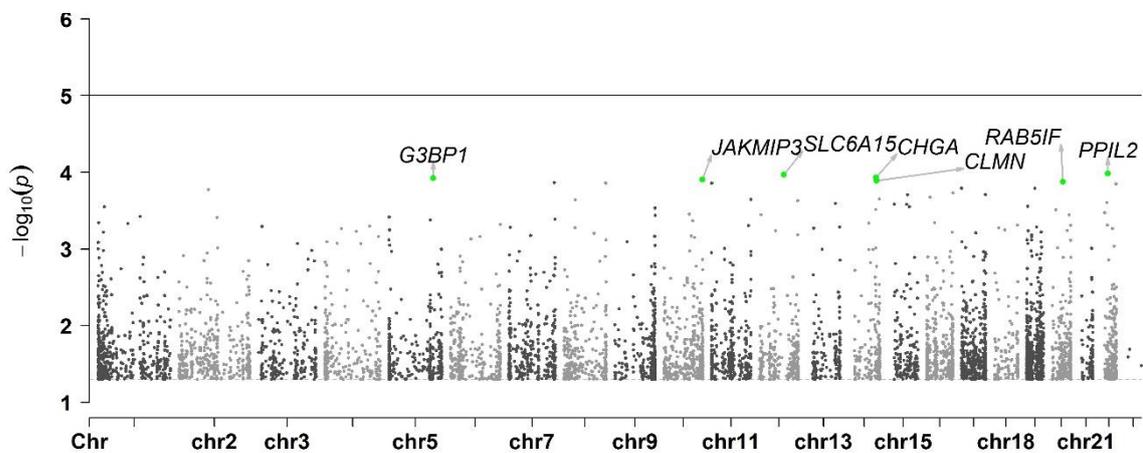


Figure 75. Manhattan plot des CpGs significativement différenciellement méthylés entre les MDS/MPN-*SF3B1*-T TN et les MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} sur l'analyse selon la méthode *in-house*.

Parmi ces CpGs, il y avait **944 CpGs** avec une valeur p inférieure à 0,01, dont :

- **618 hyperméthylées** dans le groupe MDS/MPN-*SF3B1*-T-TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}.
- **326 hypométhylées** dans le groupe MDS/MPN-*SF3B1*-T-TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}.

b) *Analyse globale des DMRs selon la méthode in-house*

L'utilisation d'un seuil à 20% de différence de méthylation dans la méthode *in-house* a permis d'identifier **234 DMRs** significatives entre les groupes MDS/MPN-*SF3B1*-TN et MDS/MPN-*SF3B1*-*JAK2*^{V617F} (**Figure 76**) :

- **131 hyperméthylées** dans le groupe MDS/MPN-*SF3B1*-T-TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}.
- **103 hypométhylées** dans le groupe MDS/MPN-*SF3B1*-T-TN par rapport aux MDS/MPN-*SF3B1*-T-*JAK2*^{V617F}.

Les annotations de gènes associées aux DMRs les plus significatives étaient : *CITED4*, *FLT4*, *ADAMTS2*, *MGC21881*, *CACNA1A* et *CDH20*.

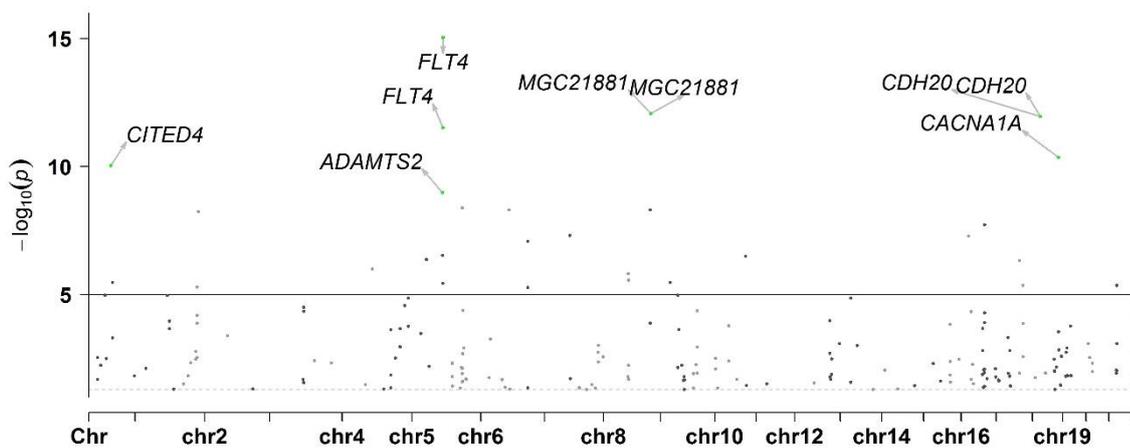


Figure 76. Manhattan plot des DMRs significativement différentiellement méthylées entre les MDS/MPN-*SF3B1*-T TN et les MDS/MPN-*SF3B1*-T *JAK2*^{V617F} sur l'analyse selon la méthode *in-house*.

c) *Élimination du bruit de fond et application du score de pertinence*

Parmi les 81 510 CpGs significatifs avec un seuil de différence de méthylation à 5%, dans la comparaison MDS/MPN-*SF3B1*-*JAK2*^{V617F} *versus* MDS/MPN-*SF3B1*-TN, 959 CpGs présentaient un score de pertinence à 5/5. L'analyse non supervisée de type PCA de ces 959 CpGs permettait de séparer les MDS/MPN-*SF3B1*-T TN des autres pathologies (**Figure 77**).

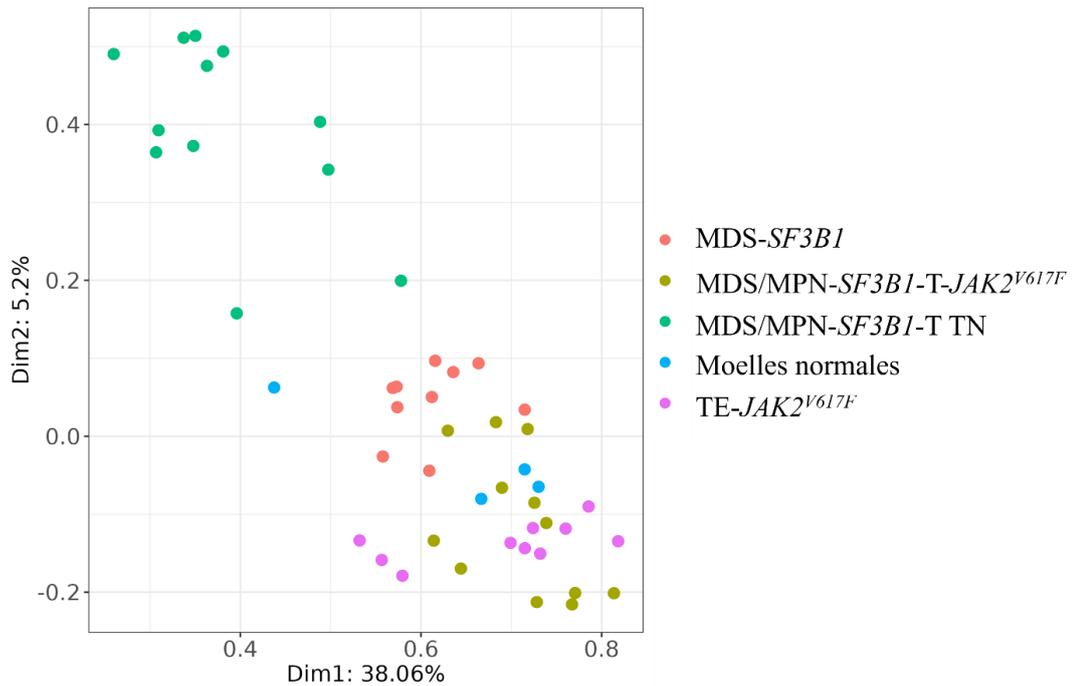


Figure 77. Analyse non supervisée par PCA des 959 CpGs différentiels significatifs pour un seuil de différence de méthylation de 5% et avec score de pertinence à 5/5.

Parmi les 15 000 CpGs les plus variants dans la comparaison MDS/MPN-*SF3B1*-*JAK2*^{V617F} versus MDS/MPN-*SF3B1*-TN, 113 CpGs présentaient un score de pertinence à 5/5. L'analyse non supervisée par PCA et heatmap de ces 113 CpGs permettait de séparer les MDS/MPN-*SF3B1*-T TN (**Figure 78**, **Figure 79**).

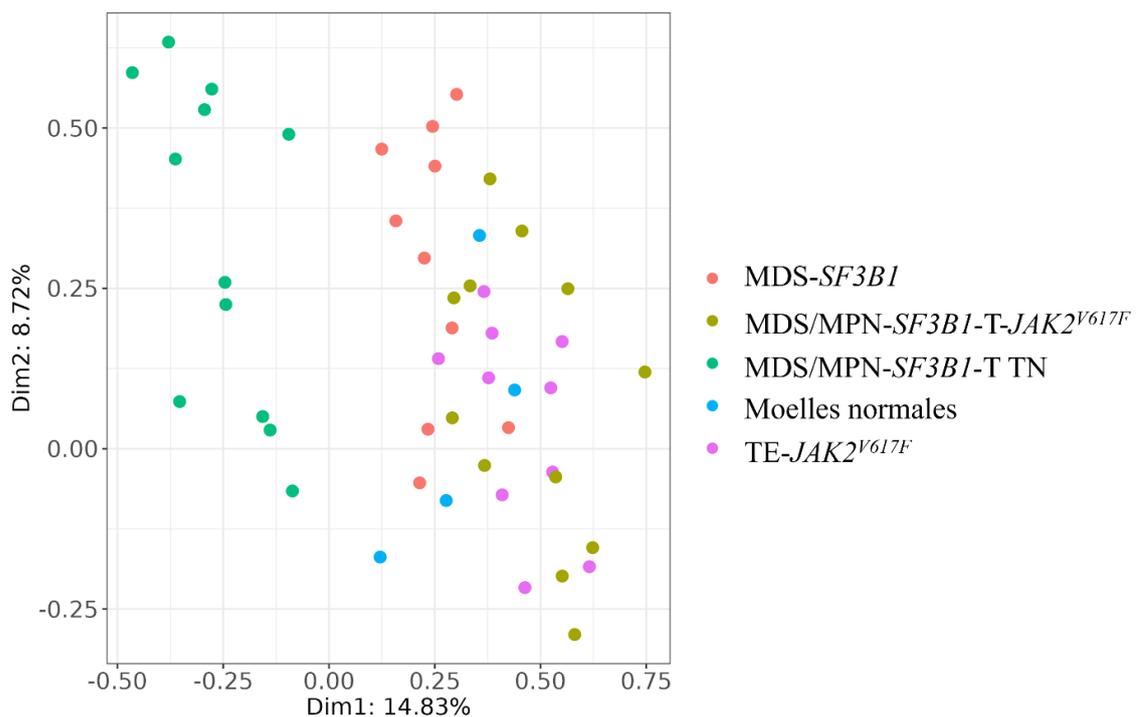


Figure 78. Analyse non supervisée par PCA des 113 CpGs avec score de pertinence à 5/5 parmi les 15 000 CpGs les plus variants.

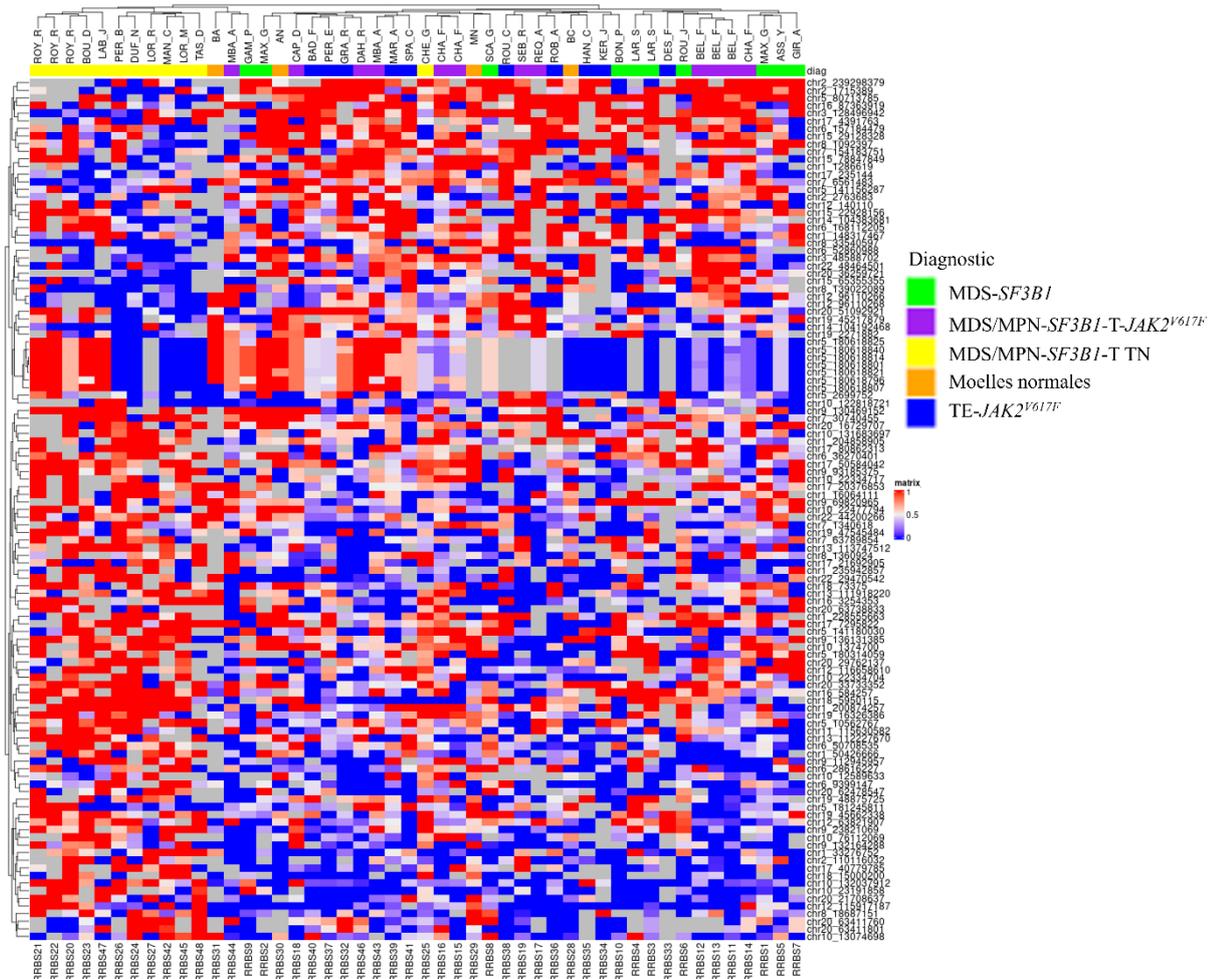


Figure 79. Heatmap des 113 CpGs avec score de pertinence à 5/5 parmi les 15 000 CpGs les plus variants.

L'utilisation d'un seuil à 20%, l'élimination du bruit de fond et la sélection des DMRs présentant un score de pertinence maximal (5/5) dans la méthode *in-house* ont permis de retenir, avant correction des *p values*, 46 DMRs annotées à 33 gènes.

Après correction des *p values*, parmi les 234 DMRs identifiées, l'élimination du bruit de fond et la sélection des DMRs présentant un score de pertinence maximal (5/5) ont permis de retenir 15 DMRs annotées à 11 gènes.

La sélection des DMRs présentant un score de pertinence supérieur ou égal à 4/5 permettait d'obtenir, parmi les 234 DMRs identifiées, 38 DMRs annotées à 28 gènes.

Les annotations fonctionnelles des DMRs montraient un enrichissement en gènes en lien la voie des MAPK :

- 6/33 pour le score maximal de 5/5 et p non corrigée : *AATK*, *LHX6*, *LMX1B*, *BCLAP*, *DUSP2* et *SPRY2*

- 4/11 pour le score maximal de 5/5 et p corrigée : *FLT4*, *GPM6A*, *LHX6*, *SPRY2*.

- 5/28 pour un score d'au moins 4/5 et p corrigée : *FLT4*, *GPM6A*, *LHX6*, *SPRY2*, *IMPA2*.

Les annotations fonctionnelles des DMRs avec un score d'au moins 4/5 contenaient également les gènes *RBI* et *ZSWIM4*.

Parmi ces DMRs, nous avons sélectionné **9 candidats** dont l'annotation fonctionnelle était particulièrement pertinente (**Tableau 15**). **Huit** d'entre eux ont pu être testés en qPCR.

Tableau 15. Candidats retenus après analyse finale *in-house* des données du RRBS des MDS/MPN-SF3B1-T.

Gène	Fonction	Voie de signalisation	Atteinte décrite dans les cancers
<i>DUSP2</i>	Phosphatase. Exprimé dans les plaquettes	Ras/MAPK	Oui
<i>LHX6</i>	LIM homeobox/facteur Trans. Famille LIM	MAPK	Oui
<i>LMX1B</i>	LIM homeobox/facteur Trans. Famille LIM	MAPK	oui
<i>AATK</i>	- tyrosine kinase en lien avec P53 - Contient miR-1250 (répresseur MAPK et dépendant du promoteur de <i>AATK</i>)	TP53/Ras/MAPK	Oui
<i>BCLAP</i>	Apoptose	STAT3/MAPK	Oui
<i>SPRY2</i>	Kinase, inhibiteur de tyrosine kinase (dont PDGFR)	Ras/MAPK, PTPN11	Oui
<i>FLT4</i>	Tyrosine kinase Active la voie MAPK/JUN et AKT	MAPK/AKT	Oui
<i>GPM6A</i>	Inhibe indirectement les MAPK et la prolifération	MAPK/src	Oui
<i>IMPA2</i>	Inhibe la phosphorylation de ERK	MAPK	Oui

8) qPCR des candidats

La qPCR a été réalisée sur 4 échantillons de moelles normales, 1 échantillon de MDS/MPN-*SF3B1*-T *JAK2*^{V617F}, 1 échantillon de MDS/MPN-*SF3B1*-T TN et 1 échantillon de MDS-*SF3B1*. Elle a montré dans l'échantillon de MDS/MPN-*SF3B1*-T TN la **surexpression relative de 5 gènes dont les séquences régulatrices étaient hypométhylées** par rapport à l'échantillon de MDS/MPN-*SF3B1*-T *JAK2*^{V617F} et aux contrôles : ***AATK*, *BLCAP*, *DUSP2*, *FLT4*, et *SPRY2*** (Tableau 16, Figure 80). Ces gènes étaient également surexprimés, dans une moindre mesure, dans l'échantillon de MDS/MPN-*SF3B1*-T-*JAK2*^{V617F} par rapport au MDS-*SF3B1* et aux moelles normales. ***LMX1B* était hyperméthylé et surexprimé** dans l'échantillon de MDS/MPN-*SF3B1*-T TN, en cohérence avec les données de la littérature (157). Malgré une hypométhylation dans l'échantillon de MDS/MPN-*SF3B1*-T TN, ***GPM6A* était sous-exprimé par rapport à l'échantillon de MDS/MPN-*SF3B1*-T *JAK2*^{V617F}, mais bien surexprimé par rapport aux moelles normales. *IMPA2* était hyperméthylé et sous-exprimé** dans l'échantillon de MDS/MPN-*SF3B1*-T TN par rapport à l'échantillon de MDS/MPN-*SF3B1*-T *JAK2*^{V617F}. *IMPA2* était sous-exprimé dans les deux échantillons de MDS/MPN-*SF3B1*-T (TN et *JAK2*^{V617F}) par rapport aux moelles normales.

Tableau 16. Mesure du niveau d'expression par qPCR des gènes candidats issus du RRBS.
MO : moelles osseuses normales

Gène	Méthylation		Niveau d'expression	
	TN <i>versus</i> <i>JAK2</i> ^{V617F}	MDS/MPN- <i>SF3B1</i> -T TN <i>versus</i> MO	MDS/MPN- <i>SF3B1</i> -T <i>JAK2</i> ^{V617F} <i>versus</i> MO	MDS- <i>SF3B1</i> <i>versus</i> MO
<i>DUSP2</i>	-24%	18	4	1
<i>SPRY2</i>	-29%	2052	5	0,3
<i>AATK</i>	-31%	25	10	0,03
<i>LMX1B</i>	+23%	272	77	0,18
<i>BLCAP</i>	-31%	32	13	0,2
<i>FLT4</i>	-56%	16	14	<0,01
<i>GPM6A</i>	-32%	7	16	7
<i>IMPA2</i>	+34%	0,06	0,14	< 0,01

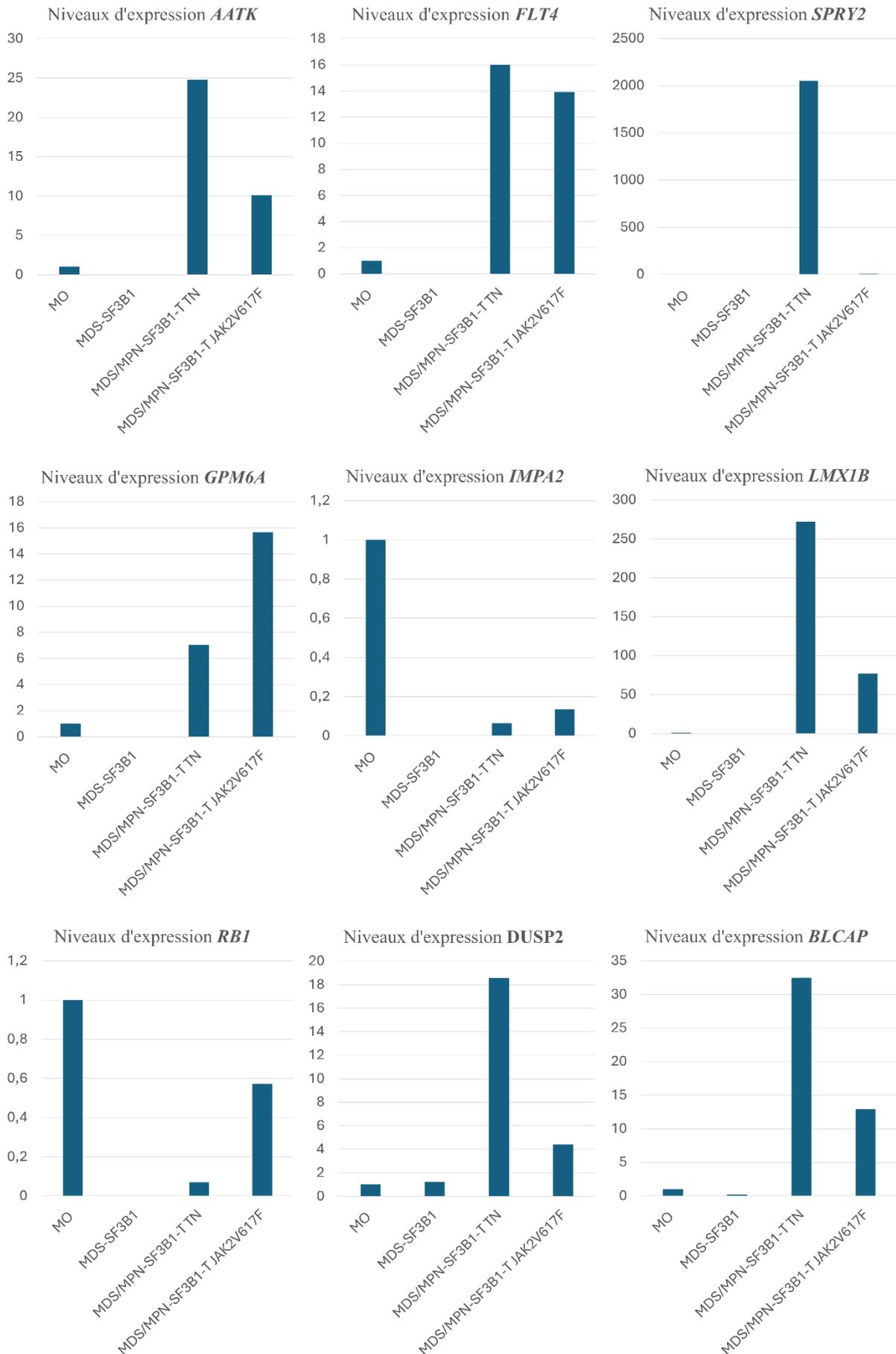


Figure 80. PCR quantitative des gènes candidats issus du RRBS.
MO = moelle osseuse.

9) Analyses ciblées de la méthylation

d) *Analyse ciblée du panel myéloïde*

L'analyse ciblée du panel myéloïde n'a pas mis en évidence de candidat dont le niveau de méthylation était significativement variable d'un groupe de pathologie à l'autre.

e) *Analyse ciblée des voies JAK/STAT, PI3k/Akt et Ras/MAPK*

L'analyse des trois grandes voies de signalisation n'a pas mis en évidence de candidat dont le niveau de méthylation était significativement variable d'un groupe de pathologie à l'autre.

DISCUSSION

PARTIE 1 - Analyse d'exome et validations fonctionnelles

L'identification de nouveaux déterminants moléculaires et des mécanismes physiopathologiques associés, expliquant la prolifération cellulaire dans les MPN permettrait d'améliorer les stratégies diagnostiques et la prise en charge thérapeutique. Cela apporterait également de nouveaux arguments quant au fondement des MDS/MPN-*SF3B1*-T en tant qu'entité distincte. Dans le cas d'une MDS avec plus de 15% de SC, une thrombocytose et présence d'une mutation *SF3B1* sans mutation de *JAK2*, *MPL* ou *CALR*, le diagnostic de MDS/MPN-*SF3B1*-T demeure incertain.

Notre première approche a été de rechercher, sur une cohorte de 16 échantillons de MDS/MPN-*SF3B1*-T, des mutations classiquement décrites dans les MPN (hors *JAK2*, *MPL* et *CALR*) en réalisant un séquençage ciblé sur un panel de 23 gènes connus. Sur le versant myélodysplasique, les résultats obtenus sont cohérents avec les données de la littérature concernant la signature moléculaire des MDS/MPN-*SF3B1*-T avec des variants de *SF3B1* dans presque tous les cas, mais également de *DNMT3A*, *TET2* et *ASXL1*. En revanche, cette approche ne nous a pas permis d'identifier au sein de ce panel de 23 gènes une ou plusieurs mutations somatiques pouvant expliquer le versant myéloprolifératif. L'effectif d'échantillons n'a cependant pas eu la puissance nécessaire pour identifier des mutations de *SETBP* et *SH2B3* décrite dans la littérature (131,132).

Nous avons donc posé l'hypothèse de l'existence de mutations somatiques encore non décrites, ou pour lesquelles le lien avec la prolifération cellulaire n'a pas encore été établi. Ceci avec une stratégie fondée sur l'analyse successive de deux cohortes : une cohorte de screening et une cohorte de validation.

En routine, le diagnostic des TE repose sur la recherche de la mutation *JAK2*^{V617F} sur sang total car si elle est présente, la mutation est retrouvée dans les granulocytes (elle survient à un stade myéloïde antérieur à la différenciation en mégacaryocyte ou granulocyte). Lors de la mise en place de l'analyse d'exome, nous avons choisi de nous focaliser sur des variations somatiques spécifiques du tissu myéloïde, en considérant que la fraction leucocytaire CD3- issue de sang total, enrichie en PNN, était le reflet du tissu myéloïde (à l'image de la recherche de mutation *JAK2*^{V617F} en routine lors du diagnostic de TE). A l'inverse, la fraction CD3+ était quant à elle presque exclusivement constituée de lymphocytes T et donc non porteuse de variations somatiques survenant à un stade hématopoïétique ultérieur à la différenciation myéloïde/lymphoïde.

Ce choix sélectif peut se discuter. En effet, une analyse ciblée avant application de ce filtre somatique (sur un panel identique à celui appliqué sur la cohorte de screening) montre que certaines mutations somatiques décrites pathogènes sont présentes dans les deux fractions CD3+ et CD3- (par exemple les mutations de *TET2* et *DNMT3A*), alors que d'autres ne sont présentes que dans la fraction CD3- (par exemple les mutations de *SF3B1*). Ainsi, les mutations de *DNMT3A* et de *TET2* apparaîtraient antérieurement à celles de *SF3B1* et sont filtrées par notre pipeline.

Ce filtre a cependant permis de réduire le nombre de candidats pertinents et nos résultats restent cohérents avec le statut mutationnel au diagnostic de nos 6 patients MDS/MPN-*SF3B1*-T TN (*SF3B1* muté avec des VAF similaires), et de confirmer l'absence de mutation de *JAK2*, *CALR* ou *MPL*.

L'analyse de la cohorte de screening (n=6 patients) a permis d'identifier 49 gènes candidats, se classant dans différentes catégories fonctionnelles pouvant impacter la prolifération. En effet, les protéines concernées sont impliquées dans la signalisation (notamment liée aux protéines G, à la survie et à l'apoptose), dans la réparation ou encore dans des mécanismes immuno-hématologiques. D'autres sont des facteurs de transcription. Presque tous sont exprimés dans la moelle osseuse selon *GeneCards* (<https://www.genecards.org>). Ces résultats confortent le choix des filtres appliqués pour une sélectivité efficace mais non excessive. Le pipeline a également mis en évidence des variants au niveau du gène *TET2*, en cohérence avec le panorama mutationnel des myélodysplasies.

La cohorte de confirmation a permis d'évaluer nos gènes candidats sur un plus grand nombre d'échantillons (n=42) d'ADN issus de patients MDS/MPN-*SF3B1*-T TN. Les données obtenues (fichiers vcf et résultats finaux) résultent toutefois d'un pipeline d'analyses bioinformatiques sur lequel nous ne sommes pas intervenus. La configuration d'un tel pipeline implique parfois des choix objectifs. En conséquence, nous n'avons pas toujours accès à la raison pour laquelle certains variants ont été filtrés. Cependant, cette cohorte a pu confirmer la forte proportion de mutation du gène *TET2* (33,3% des patients de la cohorte de validation) et la prévalence de variants faux-sens pour des candidats pertinents que nous avons déjà identifiés dans notre cohorte de screening : ***BRAT1*, *LMTK3*, *LTK*, *POU4F1*, *SGK223* et *STAT5b***.

Dans cette cohorte de validation (n=42), deux patients (4,76%) sont porteurs de variants de *STAT5b*: le variant Y665F et le variant c.2117 A>T (décrit comme Q706L dans la littérature) (158). Le variant Y665F du domaine SH2 est déjà décrit chez des patients porteurs de leucémie LGL (*Large Granular Lymphocytic Leukemia*) ou de leucémie pro-lymphocytaire T comme majorant l'activité transcriptionnelle et la phosphorylation de STAT5B (158,159). Il est donc un candidat pertinent pour participer au phénotype prolifératif des MDS/MPN-*SF3B1*-T TN. Le variant c.2117 A>T est situé dans le domaine transactivateur de STAT5B et a été également décrit chez des patients porteurs de leucémie LGL, sans effet démontré en mutagenèse dirigée sur un modèle cellulaire HeLa (*Henrietta Lacks cells*) (158). Il n'avait cependant pas été identifié comme un variant d'épissage. Les analyses *in silico* montrent que ce variant induit la création potentielle de sites de liaison à certains facteurs d'épissage tels que SC35 (SRSF2), hnRNPA1 et SRp40.

Le minigène a confirmé le saut de l'exon 17, qui aboutit à une protéine tronquée au niveau de l'acide aminé 695 de STAT5B et donc dépourvu de la tyrosine 699. Il n'est pas démontré que cet ARN mutant persiste dans la cellule et ne soit pas dégradé par le NMD. Pour parvenir à le démontrer, compte-tenu de la rareté des MDS/MPN-*SF3B1*-TN, nous projetons d'obtenir de l'ARN issu de sang de patient atteint de leucémie LGL avec le variant c.2117 A>T. Cependant, les examens de routine pouvant mettre en évidence ce variant (NGS ciblé sur un panel spécifique aux syndromes lymphoprolifératifs T) sont effectués sur ADN, ce qui limite la probabilité d'un échantillon d'ARN disponible dans un autre centre.

Nous avons donc déterminé que le variant c.2117 A>T de STAT5B induit un épissage alternatif avec saut de l'exon 17 de l'ARNm de STAT5B et aboutissant potentiellement à une protéine tronquée dépourvue de sa tyrosine phosphorylable 699. La persistance de l'ARN issu du variant c.2117 A>T aboutirait donc à la production et la surexpression de protéines STAT5B tronquées et non phosphorylées (car non phosphorylables, uSTAT5B-t). Il a été démontré que des isoformes tronquées de STAT3 et STAT5 existaient à l'état naturel (89). Une expression significative de formes tronquées participant à la leucémogénèse a également été observée dans des échantillons de LAM (160). Les formes non phosphorylées des protéines STATs (uSTATs) ont longtemps été considérées comme des précurseurs inactifs des formes phosphorylées (pSTATs) mais de plus en plus de données confirment leur présence nucléaire et leur impact sur différentes fonctions cellulaires (84,86).

Trois clusters de régions nucléiques cibles ont été déterminées au sein de la CSH et sa différenciation mégacaryocytaire (**Figure 81a**) : (i) un cluster de cibles spécifiques de uSTAT5 presque exclusivement adjacentes de sites de fixation de CTCF (CCCTC-binding factor qui est une protéine activatrice ou répressive de la transcription) et réprimées au niveau de la CSH, en l'absence de TPO (Cluster I), (ii) un cluster de cibles spécifiques de pSTAT5 (séquences GAS, *gamma interferon-activated site*) exprimées sous l'influence de la TPO (Cluster II) et (iii) un cluster de cibles à la fois de uSTAT5 et de pSTAT5 (non influencés par la TPO) (Cluster III) (86). Les liaisons de uSTAT5B à ses cibles du cluster I sont par ailleurs presque entièrement perdues dans des CSH exposées à la TPO et le *knock down* de *STAT5B* dans des cellules non exposées à la TPO induisent la différenciation mégacaryocytaire (84). uSTAT5B inhiberait donc un programme transcriptionnel nécessaire à la différenciation mégacaryocytaire et pSTAT5b favoriserait sous l'influence de la TPO, *via* ses cibles GAS, un programme transcriptionnel lié à la survie et la différenciation cellulaire. Plus précisément, la fixation d'uSTAT5B empêcherait la fixation d'ERG (*ETS transcription factor*) sur ses séquences cibles (de type ETS). L'inactivation de ERG est quant à elle décrite comme impactant fortement la mégacaryopoïèse ainsi que la destruction plaquettaire (161). Ces données dans leur ensemble suggèrent d'une part qu'un ARNm tronqué de STAT5B puisse persister et ne pas être dégradé par le NMD et d'autre part qu'une forme uSTAT5B-t puisse impacter le programme transcriptionnel de uSTAT5B (à l'instar de l'impact des formes pSTAT5B-t sur le programme de pSTAT5B).

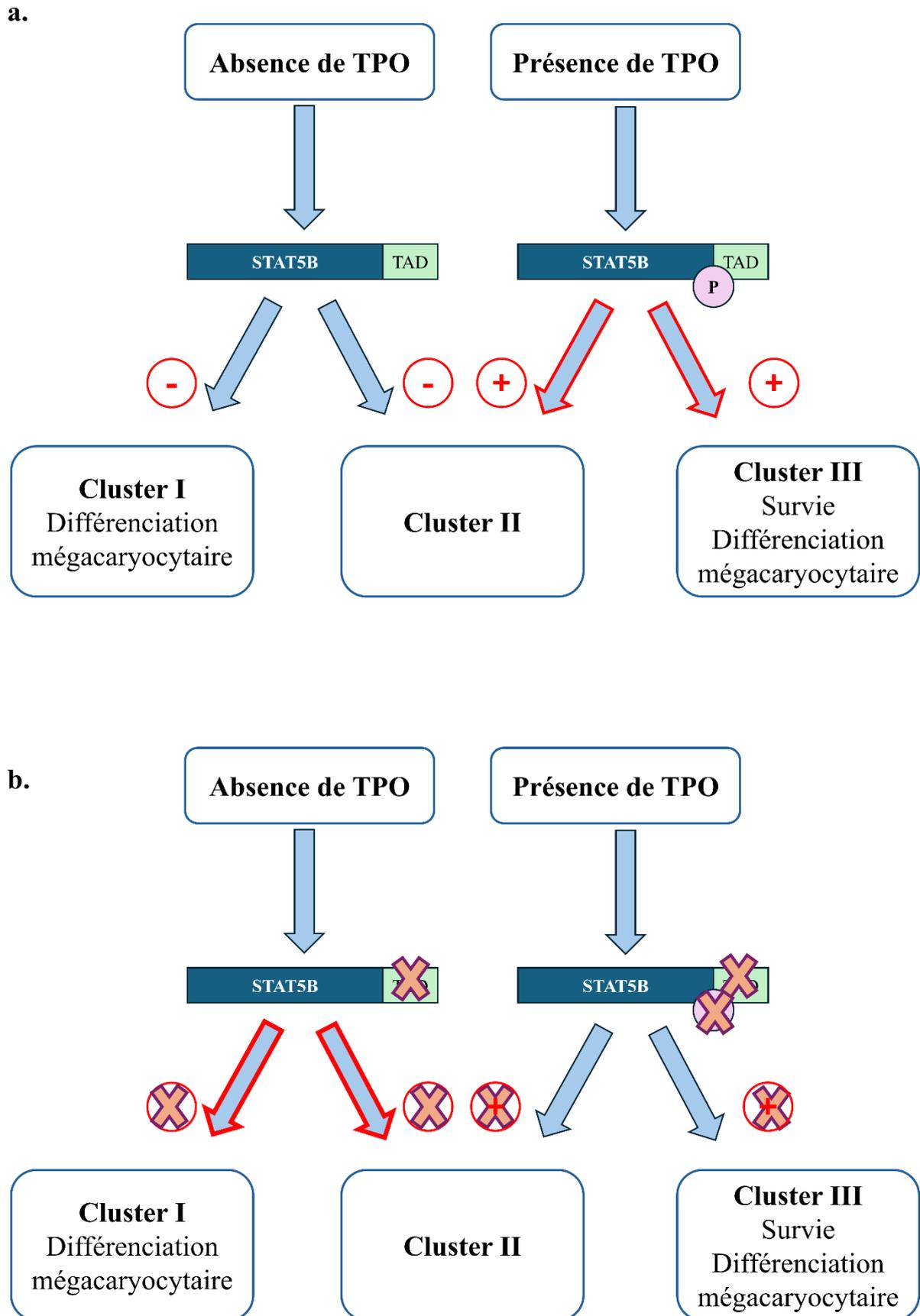


Figure 81. Hypothèse de l'impact du variant c.2117 A>T de STAT5B dans les MDS/MPN-SF3B1-T TN.

a. Effet de STAT5B phosphorylé et non phosphorylé en présence ou absence de TPO. **b.** Effet supposé de STAT5B tronquée non phosphorylée en présence ou absence de TPO.

Notre hypothèse est ici que l'expression de uSTAT5B-t causée par le variant d'épissage que nous avons identifié induit une thrombocytose en modifiant l'expression des gènes cibles de uSTAT5B (en inhibant son effet inhibiteur sur les cluster II et III par perte de son domaine transactivateur) et pSTAT5B (**Figure 81b**). Ce déséquilibre serait à l'origine d'une modification de l'activité des voies de signalisation JAK-STAT, Ras-MAPK et/ou Pi3K-Akt-mTor. L'utilisation d'un modèle cellulaire surexprimant la forme tronquée de STAT5B nous permettra d'identifier les gènes cibles impactés et les partenaires protéiques impliqués. Nous explorerons également les voies de signalisation JAK-STAT, Ras-MAPK et Akt-mTOR.

Nous disposons déjà de plasmides permettant d'exprimer soit la forme sauvage de STAT5B, soit la forme tronquée uSTAT5B-t, fusionnées à la GFP. L'étiquette GFP permet de détecter la protéine uSTAT5B-t. En effet, les anticorps commerciaux reconnaissent l'extrémité C-terminale de la protéine, absente dans uSTAT5B-t. Des premiers résultats de transfection ont montré une localisation identique (nucléaire et cytoplasmique) pour les deux protéines dans des cellules HEK-293, suggérant que la forme uSTA5B-t puisse agir également comme facteur de transcription.

Nous utiliserons dans un premier temps des cellules HEK293 : ces cellules sont faciles à transfecter et les voies de signalisation JAK-STAT, Ras-MAPK et/ou Pi3K-Akt-mTor sont présentes. En fonction des résultats obtenus, nous confirmerons dans un deuxième temps dans un modèle d'UT-7 (CVCL_2233 : lignée mégacaryocytaire), ces cellules étant beaucoup plus difficiles à cultiver et à transfecter.

La perte du domaine de transactivation dans l'isoforme uSTAT5B-t laisse supposer une perte d'interaction avec certains partenaires protéiques, notamment avec un rôle inhibiteur sur la fonction de STAT5B. Nous souhaitons donc identifier les partenaires communs et spécifiques à chacune des isoformes par co-immunoprécipitation sur les cellules triées GFP+.

Nous souhaitons également examiner et comparer, par RT-qPCR, l'expression d'un panel de gènes cibles de STAT5B dans les cellules mutées et sauvages. Nous pourrions enfin analyser l'activation des différentes voies de signalisation (JAK-STAT, Ras-MAPK et/ou Pi3K-Akt-mTor) par *western blot* sur les cellules transfectées exprimant soit la protéine sauvage, soit la protéine uSTAT5B-t. La contribution des partenaires protéiques identifiés précédemment sera vérifiée en inhibant leur expression par siARN.

Des biais de transversion C vers A sont décrits dans la littérature comme « signature » de certains cancers (162,163), notamment le cancer du côlon, où une atteinte du système de réparation simple brin est classiquement retrouvée au niveau du gène *MUTYH*. L'analyse *a posteriori* des variants issus des derniers protocoles d'analyse bioinformatique de l'exome des 6 patients de la cohorte de screening n'a pas révélé d'atteinte des gènes impliqués dans le système *MUTYH*. L'atteinte du système de réparation BER peut cependant expliquer un biais de transversion tel que nous le retrouvons. Il serait donc intéressant d'analyser, de manière plus large, les résultats de l'exome pour l'ensemble des acteurs du BER. Ces éléments pourraient également se compléter avec les résultats du méthylome, par une analyse ciblée sur le BER.

PTPN12 est un homologue de *PTPN11*, un gène qui code pour une phosphatase impliquée dans de multiples processus de signalisation cellulaire (164), notamment la régulation des voies Ras/Erk, Pi3k/Akt, JAK/STAT ou encore NF-Kb. Plusieurs mutations récurrentes (E76K, D61Y) de *PTPN11* ont été décrites dans la leucémie myélo-monocytaire chronique juvénile (LMMJ), une MPN de l'enfance non associée aux mutations de *JAK2*^{V617F}. Le gain de fonction de *PTPN11* joue un rôle positif dans la transduction des signaux de prolifération. La mise en évidence d'une duplication de *PTPN12* n'a pas pu être confirmée sur la cohorte de validation, en l'absence de comparaison possible entre tissu tumoral et tissu sain. L'implication de son homologue *PTPN11* dans la LMMJ nous a poussé à rechercher la duplication de *PTPN12* dans d'autres MPN triple négatifs, notamment la TE. Nous avons disposé d'échantillons ADN issus de sang total et avons choisi de tester également de l'ADN issu de MF, dont la charge allélique tumorale (souvent plus importante) augmenterait nos chances de mettre en évidence un tel variant s'il existe. Les premières plaques de qPCR retrouvent une duplication de *PTPN12* au niveau d'un échantillon de TE TN (sur 11 testés). Aucune duplication n'a été mise en évidence pour les 4 échantillons de MF testés.

Le modèle cellulaire constitué sur lignée HEK293 n'a pas permis de mettre en évidence une majoration de la phosphorylation des voies AKT et ERK suite à la surexpression de *PTPN12*.

La possible duplication de *PTPN12* reste une piste sérieuse. Elle conforterait la « parenté » des MDS/MPN-*SF3B1*-T avec d'autres MPN telles que la LMMJ. Elle constituerait également une piste intéressante pour explorer les TE et MF TN qui, comme pour les MDS/MPN-*SF3B1*-T peuvent être difficiles à diagnostiquer en l'absence de marqueur moléculaire spécifique. Son implication moléculaire pourrait conduire à l'utilisation de nouvelles thérapeutiques.

PARTIE 2 - Approche épigénétique par l'étude du méthylome

Le caractère triple négatif de certains MDS/MPN-*SF3B1*-T est à l'origine de difficultés diagnostiques. Il en résulte également une mauvaise compréhension de la physiopathologie de la prolifération plaquettaire. Le traitement des MDS/MPN-*SF3B1*-T repose sur la prise en charge de l'anémie et la prévention de la thrombose. Il nécessite donc parfois la mise en place d'une cytoréduction médullaire, qui doit impérativement tenir compte de l'accentuation de l'anémie. Une meilleure compréhension des mécanismes prolifératifs pourrait permettre la mise en place de nouvelles stratégies thérapeutiques cytoréductrices ciblées et moins toxiques.

Il a été démontré que la prolifération cellulaire des MPN étaient en lien avec une activation constitutive de la voie JAK-STAT, mais également avec un environnement inflammatoire riche en cytokines telles que l'IL3, la TPO ou le GM-CSF (165). La voie JAK-STAT communique avec la voie des MAPK par le biais de PTPN11 (dont les mutations sont fréquentes dans la LMMJ) et GRB2 (91). Paradoxalement, dans les cellules portant la mutation *JAK2*^{V617F}, il existe un rétrocontrôle de la voie des MAPK avec, en particulier, la surexpression de *DUSP1*, régulateur négatif de la voie des MAPK et de TP53. Cette surexpression de *DUSP1* n'est pas retrouvée dans les cellules *MPL*^{W515L} mais donne une explication plausible à l'échappement thérapeutique aux inhibiteurs de JAK2. En effet, l'élévation du seuil apoptotique des cellules souches tumorales entraîne leur survie sous traitement et explique l'absence d'éradication complète du clone. Il existe donc un lien étroit entre MPN et voie MAPK, dont la régulation positive comme négative est impactée par la pathologie.

La méthodologie d'analyse du RRBS et l'application du score de pertinence a permis de mettre en évidence, dans les MDS/MPN-*SF3B1*-T TN, un enrichissement des DMRs significatives en séquences régulatrices de gènes reliés à la voie MAPK. Nous avons notamment mis en évidence l'hypométhylation et la surexpression, par rapport aux contrôles et MDS/MPN-*SF3B1*-T *JAK2*^{V617F}, de quatre gènes régulateurs négatifs de la voie des MAPK : ***LMTK1***, ***BLCAP***, ***DUSP2*** (paralogue de *DUSP1*) et ***SPRY2***. *LMTK1* code pour une tyrosine kinase impliquée dans la croissance et la prolifération des précurseurs myéloïdes. Des études ont démontré que les protéines LMTK inhibaient la phosphorylation d'Akt et ERK et promouvait l'activation de la MAPK p38 et des kinases JUN (166). La surexpression des protéines JUN a déjà été documentée dans les TE (107).

BLCAP et *SPRY2* sont des gènes suppresseurs de tumeur (167). *SPRY2* est un régulateur négatif direct de la voie MAPK et décrit comme régulé négativement par hyperméthylation dans des pathologies myéloïdes telles que la LMMC (168). Cette protéine inhibe également les protéines GRB2 et SOS dans les interconnexions entre les voies JAK-STAT et MAPK.

Cet ensemble cohérent de surexpressions franches (surexpression d'un facteur 18 à 2052 par rapport aux contrôles) pourrait être le reflet d'une réaction de la cellule à l'activation constitutive des voies de prolifération. Cela renforce l'hypothèse selon laquelle les MDS/MPN-*SF3B1*-T, et les MPN au sens large, sont des maladies affectant la voie des MAPK et la prolifération au-delà de la voie JAK-STAT. L'absence de mutation affectant des acteurs de la voie JAK-STAT dans les MPN triple négatifs n'impacterait donc pas l'apparition d'un phénotype similaire et pourrait orienter vers des possibilités thérapeutiques communes extérieures à JAK-STAT.

Nous avons aussi mis en évidence une surexpression de *FLT4* (une tyrosine kinase activatrice de la voie MAPK) (169), ainsi qu'une surexpression par hyperméthylation de *LMX1B*, en cohérence avec les données de la littérature (157). *LMX1B* est un facteur transcriptionnel oncogène de la famille des LIM-homeodomains. Sa surexpression est associée à la prolifération, à la migration et la radiorésistance dans divers cancers (170).

IMPA2, codant pour une protéine inhibitrice de ERK et donc de la voie MAPK était sous-exprimé (sous-expression d'un facteur 0,06 par rapport aux contrôles et 0,45 par rapport à l'échantillon les MDS/MPN-*SF3B1*-T *JAK2*^{V617F}), ce qui a déjà été observé dans certains cancers (171).

De récents travaux ont démontré le rôle inhibiteur de *GPM6A* sur la voie MAPK et sa sous-expression dans le cancer du poumon, proposant cette protéine comme nouvelle cible thérapeutique (172). Cette protéine est ici sous-exprimée par hyperméthylation dans les MDS/MPN-*SF3B1*-T par rapport aux contrôles, et particulièrement pour l'échantillon de MDS/MPN-*SF3B1*-T TN.

Les différences d'expression de *FLT4*, *GPM6A* et *IMPA2* pourraient être directement causées dans le phénotype des MDS/MPN-*SF3B1*-T TN. L'altération de leur régulation épigénétique pourrait par ailleurs être une conséquence des anomalies de méthylation imputables au caractère dysplasique de la pathologie.

Il est intéressant de noter que, à l'exception de *GPM6A*, les surexpressions de l'ensemble des gènes est également observée dans l'échantillon de MDS/MPN-*SF3B1*-T *JAK2*^{V617F} par rapport aux contrôles, mais dans une moindre mesure (4 à 77 fois par rapport aux contrôles). Ces mêmes gènes étaient plutôt sous-exprimés dans l'échantillon de MDS-*SF3B1* (de 0,03 à 1 fois par rapport aux contrôles). Cette hiérarchisation se retrouvait aussi pour la sous-expression de *IMPA2*, d'avantage sous-exprimé dans les MDS/MPN-*SF3B1*-T TN que les MDS/MPN-*SF3B1*-T *JAK2*^{V617F} par rapport aux contrôles. Cela renforce l'idée selon laquelle les altérations observées sont réactionnelles pour les quatre gènes régulateurs négatifs mais soulève également la probabilité que *LMX1B*, *FLT4* et *IMPA2* ne soient impliqués que partiellement dans le phénotype des MDS/MPN-*SF3B1*-T TN.

L'analyse de données de méthylation issues de méthylome de type RRBS nécessite l'utilisation d'outils d'identification de DMR tels que *MethylKit*, *Metilene*, *BiSeq* ou *DMRfinder* (151,173–176). La méthodologie *MethylKit* se base sur la création arbitraire de régions par fenêtre glissante le long du génome et par comparaison de ces régions afin de mettre en évidence les DMRs. Cette stratégie est fréquemment utilisée pour ce type d'analyse mais il n'existe pas de méthode de référence pour définir les seuils de sélection. Le calcul du niveau de méthylation de chaque DMR selon *MethylKit* se base sur le pourcentage de cytosines non converties par le traitement bisulfite sur l'ensemble de la région. La notion de position de chaque CpG est donc perdue, ainsi que le niveau de méthylation différentiel entre les CpGs d'une même région.

Nous avons donc proposé ici une nouvelle méthodologie inspirée par *MethylKit*, tout en conservant ces informations positionnelles dans notre comparaison. Ce ne sont plus les moyennes globales de méthylation que nous comparons, mais des vecteurs composés des ratios de méthylation de chaque CpG de la région. L'une des caractéristiques des données issues de séquençage haut débit est leur important volume. L'analyse bio-informatique nécessite l'application d'un pipeline comportant des étapes de filtre afin d'extraire les données les plus pertinentes. Les annotations fonctionnelles sont d'une grande importance pour sélectionner les candidats les plus pertinents mais sont limitées par les connaissances établies et la capacité d'explorer en profondeur un grand nombre de ces candidats.

Dans le but de ne relever que les candidats pouvant répondre à la problématique étudiée, nous avons établi un score de pertinence clinique. Ce score permet d'une part de traduire la cohérence phénotypique d'une DMR candidate en une échelle chiffrée, mais également d'exploiter au mieux les comparaisons multiples réalisées à partir d'échantillons de pathologies exprimant diversement les versants myélodysplasiques et myéloprolifératifs, et de contrôles. Cette stratégie aide à l'élimination de DMRs significatives reflétant une différence entre deux groupes mais ne répondant pas à la problématique proliférative *JAK2*-indépendante étudiée.

Nous sommes toutefois limités dans notre étude. En effet, notre méthode peut potentiellement éliminer des DMRs pertinentes mais non reconnues comme telles. De plus, la rareté de la pathologie et le coût des analyses nous conduisent à de faibles effectifs et à une analyse non exhaustive. Par ailleurs, nous ne pouvons maîtriser toutes les variables impactant le méthylome. Nous avons notamment dû adapter la sélection des échantillons selon leur origine tissulaire et cellulaire, mais ces différences n'ont pu être totalement effacées. Le RRBS est de moins en moins utilisé pour l'observation du méthylome. En effet, à ce jour, l'analyse complète du méthylome ou l'utilisation de puces de méthylation disposant à présent d'une meilleure couverture tendent à réduire la pertinence de notre choix initial.

De récents travaux en bioinformatique ont démontré une plus grande pertinence biologique de « e-values » dans la détection des DMRs en RRBS par les outils classiques tels que *MethylKit* (176). En effet, les p-values et p-values corrigées, bien que largement utilisées pour estimer la significativité statistique de très nombreux types d'analyses, sont de plus en plus discutées et controversées. Les *p-values* permettent d'évaluer la probabilité de pouvoir rejeter une hypothèse nulle après un test statistique. Plus elle est faible et plus l'hypothèse peut être rejetée. Yang et al ont proposé, en 2023, l'utilisation de e-value qui serait un ratio de probabilité d'accepter l'hypothèse nulle en prenant en compte l'ensemble des paramètres de position et de dispersion (176), dans la détection de DMRs en RRBS. Il pourrait donc être intéressant de reprendre nos données et d'associer l'utilisation de e-values, la conservation des informations de position et l'application du score de pertinence.

Malgré tout, les résultats obtenus par notre nouvelle stratégie d'interprétation du RRBS sont en cohérence avec l'état actuel des connaissances et avec le phénotype de notre pathologie. Si la problématique d'un mécanisme épigénétique précis à l'origine des MDS/MPN-*SF3B1*-T TN n'est pas résolue, ces résultats nous permettent de penser que notre stratégie pourrait être applicable à d'autres analyses de type RRBS.

CONCLUSION

L'analyse génétique et épigénétique d'échantillons de patients atteints de MDS/MPN-*SF3B1*-T TN a révélé plusieurs pistes pouvant expliquer la prolifération des plaquettes. Nous avons retenu plusieurs gènes candidats dont les variants pourraient être à l'origine du phénotype, notamment *STAT5B*, dont les analyses fonctionnelles sont en cours. L'existence d'un biais de transversion pour tous nos patients nous amène également à suspecter une atteinte de la réparation qui pourrait être d'origine germinale. La mise au point d'une nouvelle stratégie d'analyse du méthylome par RRBS a conforté l'hypothèse d'une maladie affectant la voie des MAPK.

En conclusion, les MDS/MPN-*SF3B1*-T TN correspondent à une maladie complexe, dans laquelle le phénotype est probablement d'origine multifactorielle, avec implication de mutations somatiques et peut être germinales. La coexistence de mutations du spliceosome et de mutations affectant l'épigénétique pourrait être à l'origine des dérégulations des voies de prolifération par hypométhylation aberrante de promoteurs proto-oncogènes et hyperméthylation de promoteurs suppresseurs de tumeur.

RÉFÉRENCES

1. Sperling AS, Gibson CJ, Ebert BL. The genetics of myelodysplastic syndrome: from clonal haematopoiesis to secondary leukaemia. *Nat Rev Cancer*. janv 2017;17(1):5-19.
2. Khoury JD, Solary E, Abla O, Akkari Y, Alaggio R, Apperley JF, et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Myeloid and Histiocytic/Dendritic Neoplasms. *Leukemia*. juill 2022;36(7):1703-19.
3. Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al. Age-related clonal hematopoiesis associated with adverse outcomes. *N Engl J Med*. 25 déc 2014;371(26):2488-98.
4. Demajo S, Ramis-Zaldivar JE, Muiños F, Grau ML, Andrianova M, López-Bigas N, et al. Identification of Clonal Hematopoiesis Driver Mutations through In Silico Saturation Mutagenesis. *MedRxiv Prepr Serv Health Sci*. 14 déc 2023;2023.12.13.23299893.
5. Steensma DP, Bejar R, Jaiswal S, Lindsley RC, Sekeres MA, Hasserjian RP, et al. Clonal hematopoiesis of indeterminate potential and its distinction from myelodysplastic syndromes. *Blood*. 2 juill 2015;126(1):9-16.
6. Calvillo-Argüelles O, Jaiswal S, Shlush LI, Moslehi JJ, Schimmer A, Barac A, et al. Connections Between Clonal Hematopoiesis, Cardiovascular Disease, and Cancer: A Review. *JAMA Cardiol*. 1 avr 2019;4(4):380-7.
7. Genovese G, Kähler AK, Handsaker RE, Lindberg J, Rose SA, Bakhoum SF, et al. Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N Engl J Med*. 25 déc 2014;371(26):2477-87.
8. Kennedy JA, Ebert BL. Clinical Implications of Genetic Mutations in Myelodysplastic Syndrome. *J Clin Oncol Off J Am Soc Clin Oncol*. 20 mars 2017;35(9):968-74.
9. Haferlach T, Nagata Y, Grossmann V, Okuno Y, Bacher U, Nagae G, et al. Landscape of genetic lesions in 944 patients with myelodysplastic syndromes. *Leukemia*. févr 2014;28(2):241-7.
10. Bejar R, Stevenson K, Abdel-Wahab O, Galili N, Nilsson B, Garcia-Manero G, et al. Clinical effect of point mutations in myelodysplastic syndromes. *N Engl J Med*. 30 juin 2011;364(26):2496-506.
11. Awada H, Thapa B, Visconte V. The Genomics of Myelodysplastic Syndromes: Origins of Disease Evolution, Biological Pathways, and Prognostic Implications. *Cells*. 20 nov 2020;9(11):2512.
12. Will B, Zhou L, Vogler TO, Ben-Neriah S, Schinke C, Tamari R, et al. Stem and progenitor cells in myelodysplastic syndromes show aberrant stage-specific expansion and harbor genetic and epigenetic alterations. *Blood*. 6 sept 2012;120(10):2076-86.
13. Chesnais V, Arcangeli ML, Delette C, Rousseau A, Guermouche H, Lefevre C, et al. Architectural and functional heterogeneity of hematopoietic stem/progenitor cells in non-del(5q) myelodysplastic syndromes. *Blood*. 26 janv 2017;129(4):484-96.
14. Arber DA, Orazi A, Hasserjian RP, Borowitz MJ, Calvo KR, Kvasnicka HM, et al. International Consensus Classification of Myeloid Neoplasms and Acute Leukemias: integrating morphologic, clinical, and genomic data. *Blood*. 15 sept 2022;140(11):1200-28.
15. Greenberg P, Cox C, LeBeau MM, Fenau P, Morel P, Sanz G, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. *Blood*. 15 mars 1997;89(6):2079-88.
16. Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Solé F, et al. Revised international prognostic scoring system for myelodysplastic syndromes. *Blood*. 20 sept 2012;120(12):2454-65.
17. Molecular International Prognostic Scoring System for Myelodysplastic Syndromes | NEJM Evidence [Internet]. [cité 9 janv 2024]. Disponible sur: <https://evidence.nejm.org/doi/full/10.1056/EVIDoa2200008>
18. Bernard E, Nannya Y, Hasserjian RP, Devlin SM, Tuechler H, Medina-Martinez JS, et al. Implications of TP53 allelic state for genome stability, clinical presentation and outcomes in myelodysplastic syndromes. *Nat Med*. oct 2020;26(10):1549-56.
19. Steensma DP, Gattermann N. When is iron overload deleterious, and when and how should iron chelation therapy be administered in myelodysplastic syndromes? *Best Pract Res Clin Haematol*. déc 2013;26(4):431-44.
20. Malcovati L, Hellström-Lindberg E, Bowen D, Adès L, Cermak J, Del Cañizo C, et al. Diagnosis and treatment of primary myelodysplastic syndromes in adults: recommendations from the European LeukemiaNet. *Blood*. 24 oct 2013;122(17):2943-64.

21. Park S, Kelaidi C, Sapena R, Vassilieff D, Beyne-Rauzy O, Coiteux V, et al. Early introduction of ESA in low risk MDS patients may delay the need for RBC transfusion: a retrospective analysis on 112 patients. *Leuk Res.* nov 2010;34(11):1430-6.
22. Fenaux P, Platzbecker U, Mufti GJ, Garcia-Manero G, Buckstein R, Santini V, et al. Luspatercept in Patients with Lower-Risk Myelodysplastic Syndromes. *N Engl J Med.* 9 janv 2020;382(2):140-51.
23. Platzbecker U, Della Porta MG, Santini V, Zeidan AM, Komrokji RS, Shortt J, et al. Efficacy and safety of luspatercept versus epoetin alfa in erythropoiesis-stimulating agent-naïve, transfusion-dependent, lower-risk myelodysplastic syndromes (COMMANDS): interim analysis of a phase 3, open-label, randomised controlled trial. *Lancet Lond Engl.* 29 juill 2023;402(10399):373-85.
24. Fenaux P, Giagounidis A, Selleslag D, Beyne-Rauzy O, Mufti G, Mittelman M, et al. A randomized phase 3 study of lenalidomide versus placebo in RBC transfusion-dependent patients with Low-/Intermediate-1-risk myelodysplastic syndromes with del5q. *Blood.* 6 oct 2011;118(14):3765-76.
25. Fenaux P, Mufti GJ, Hellstrom-Lindberg E, Santini V, Finelli C, Giagounidis A, et al. Efficacy of azacitidine compared with that of conventional care regimens in the treatment of higher-risk myelodysplastic syndromes: a randomised, open-label, phase III study. *Lancet Oncol.* mars 2009;10(3):223-32.
26. Meldi K, Qin T, Buchi F, Droin N, Sotzen J, Micol JB, et al. Specific molecular signatures predict decitabine response in chronic myelomonocytic leukemia. *J Clin Invest.* mai 2015;125(5):1857-72.
27. Santini V, Melnick A, Maciejewski JP, Duprez E, Nervi C, Cocco L, et al. Epigenetics in focus: pathogenesis of myelodysplastic syndromes and the role of hypomethylating agents. *Crit Rev Oncol Hematol.* nov 2013;88(2):231-45.
28. Tsai HC, Li H, Van Neste L, Cai Y, Robert C, Rassool FV, et al. Transient low doses of DNA-demethylating agents exert durable antitumor effects on hematological and epithelial tumor cells. *Cancer Cell.* 20 mars 2012;21(3):430-46.
29. Cheson BD, Greenberg PL, Bennett JM, Lowenberg B, Wijermans PW, Nimer SD, et al. Clinical application and proposal for modification of the International Working Group (IWG) response criteria in myelodysplasia. *Blood.* 15 juill 2006;108(2):419-25.
30. Santini V, Prebet T, Fenaux P, Gattermann N, Nilsson L, Pfeilstöcker M, et al. Minimizing risk of hypomethylating agent failure in patients with higher-risk MDS and practical management recommendations. *Leuk Res.* déc 2014;38(12):1381-91.
31. Belver L, Albero R, Ferrando AA. Deregulation of enhancer structure, function, and dynamics in acute lymphoblastic leukemia. *Trends Immunol.* mai 2021;42(5):418-31.
32. Ingram N. Waddington, Holmyard and Alchemy: Perspectives on the Epigenetic Landscape. *Endeavour.* sept 2019;43(3):100690.
33. Baedke J. The epigenetic landscape in the course of time: Conrad Hal Waddington's methodological impact on the life sciences. *Stud Hist Philos Biol Biomed Sci.* déc 2013;44(4 Pt B):756-73.
34. Shih AH, Abdel-Wahab O, Patel JP, Levine RL. The role of mutations in epigenetic regulators in myeloid malignancies. *Nat Rev Cancer.* sept 2012;12(9):599-612.
35. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. *Cell.* 16 sept 2011;146(6):866-72.
36. Esteller M. Epigenetics in cancer. *N Engl J Med.* 13 mars 2008;358(11):1148-59.
37. Yamazaki J, Issa JPJ. Epigenetic aspects of MDS and its molecular targeted therapy. *Int J Hematol.* févr 2013;97(2):175-82.
38. Figueroa ME, Skrabanek L, Li Y, Jiemjit A, Fandy TE, Paietta E, et al. MDS and secondary AML display unique patterns and abundance of aberrant DNA methylation. *Blood.* 15 oct 2009;114(16):3448-58.
39. Walter MJ, Ding L, Shen D, Shao J, Grillot M, McLellan M, et al. Recurrent DNMT3A mutations in patients with myelodysplastic syndromes. *Leukemia.* juill 2011;25(7):1153-8.
40. Langemeijer SMC, Kuiper RP, Berends M, Knops R, Aslanyan MG, Massop M, et al. Acquired mutations in TET2 are common in myelodysplastic syndromes. *Nat Genet.* juill 2009;41(7):838-42.
41. Jankowska AM, Szpurka H, Tiu RV, Makishima H, Afable M, Huh J, et al. Loss of heterozygosity 4q24 and TET2 mutations associated with myelodysplastic/myeloproliferative neoplasms. *Blood.* 18 juin 2009;113(25):6403-10.
42. Ko M, Huang Y, Jankowska AM, Pape UJ, Tahiliani M, Bandukwala HS, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. *Nature.* 9 déc 2010;468(7325):839-43.

43. Rasmussen KD, Jia G, Johansen JV, Pedersen MT, Rapin N, Bagger FO, et al. Loss of TET2 in hematopoietic cells leads to DNA hypermethylation of active enhancers and induction of leukemogenesis. *Genes Dev.* 1 mai 2015;29(9):910-22.
44. Itzykson R, Kosmider O, Cluzeau T, Mansat-De Mas V, Dreyfus F, Beyne-Rauzy O, et al. Impact of TET2 mutations on response rate to azacitidine in myelodysplastic syndromes and low blast count acute myeloid leukemias. *Leukemia.* juill 2011;25(7):1147-52.
45. Sengoku T, Yokoyama S. Structural basis for histone H3 Lys 27 demethylation by UTX/KDM6A. *Genes Dev.* 1 nov 2011;25(21):2266-77.
46. Jeong M, Sun D, Luo M, Huang Y, Challen GA, Rodriguez B, et al. Large conserved domains of low DNA methylation maintained by Dnmt3a. *Nat Genet.* janv 2014;46(1):17-23.
47. Chen H, Wu S. Hypermethylation of the p15(INK4B) gene in acute leukemia and myelodysplastic syndromes. *Chin Med J (Engl).* juill 2002;115(7):987-90.
48. Daskalakis M, Nguyen TT, Nguyen C, Guldborg P, Köhler G, Wijermans P, et al. Demethylation of a hypermethylated P15/INK4B gene in patients with myelodysplastic syndrome by 5-Aza-2'-deoxycytidine (decitabine) treatment. *Blood.* 15 oct 2002;100(8):2957-64.
49. Duran-Ferrer M, Martín-Subero JI. Epigenomic Characterization of Lymphoid Neoplasms. *Annu Rev Pathol.* 24 janv 2024;19:371-96.
50. Peterson CL, Laniel MA. Histones and histone modifications. *Curr Biol CB.* 27 juill 2004;14(14):R546-551.
51. Blackledge NP, Klose RJ. The molecular principles of gene regulation by Polycomb repressive complexes. *Nat Rev Mol Cell Biol.* déc 2021;22(12):815-33.
52. Inoue D, Fujino T, Kitamura T. ASXL1 as a critical regulator of epigenetic marks and therapeutic potential of mutated cells. *Oncotarget.* 16 oct 2018;9(81):35203-4.
53. Wan Y, Wu CJ. SF3B1 mutations in chronic lymphocytic leukemia. *Blood.* 6 juin 2013;121(23):4627-34.
54. Lee Y, Rio DC. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev Biochem.* 2015;84:291-323.
55. Graubert TA, Shen D, Ding L, Okeyo-Owuor T, Lunn CL, Shao J, et al. Recurrent mutations in the U2AF1 splicing factor in myelodysplastic syndromes. *Nat Genet.* 11 déc 2011;44(1):53-7.
56. Rahman M. Oncogenic Landscapes of Splicing-Factor Mutant MDS. In 2019 [cité 10 janv 2024]. Disponible sur: <https://www.semanticscholar.org/paper/Oncogenic-Landscapes-of-Splicing-Factor-Mutant-MDS-Rahman/b04a217360339bf2f106117c09812dfbc239a6a6#citing-papers>
57. Darman RB, Seiler M, Agrawal AA, Lim KH, Peng S, Aird D, et al. Cancer-Associated SF3B1 Hotspot Mutations Induce Cryptic 3' Splice Site Selection through Use of a Different Branch Point. *Cell Rep.* 3 nov 2015;13(5):1033-45.
58. Alsafadi S, Houy A, Battistella A, Popova T, Wassef M, Henry E, et al. Cancer-associated SF3B1 mutations affect alternative splicing by promoting alternative branchpoint usage. *Nat Commun.* 4 févr 2016;7:10615.
59. Clough CA, Pangallo J, Sarchi M, Ilagan JO, North K, Bergantinos R, et al. Coordinated missplicing of TMEM14C and ABCB7 causes ring sideroblast formation in SF3B1-mutant myelodysplastic syndrome. *Blood.* 31 mars 2022;139(13):2038-49.
60. Mufti GJ, Bennett JM, Goasguen J, Bain BJ, Baumann I, Brunning R, et al. 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.* nov 2008;93(11):1712-7.
61. Todisco G, Creignou M, Bernard E, Björklund AC, Moura PL, Tesi B, et al. Integrated Genomic and Transcriptomic Analysis Improves Disease Classification and Risk Stratification of MDS with Ring Sideroblasts. *Clin Cancer Res Off J Am Assoc Cancer Res.* 13 oct 2023;29(20):4256-67.
62. Patnaik MM, Tefferi A. Myelodysplastic syndromes with ring sideroblasts (MDS-RS) and MDS/myeloproliferative neoplasm with RS and thrombocytosis (MDS/MPN-RS-T) - « 2021 update on diagnosis, risk-stratification, and management ». *Am J Hematol.* 1 mars 2021;96(3):379-94.
63. James C, Ugo V, Le Couédic JP, Staerk J, Delhommeau F, Lacout C, et al. A unique clonal JAK2 mutation leading to constitutive signalling causes polycythaemia vera. *Nature.* 28 avr 2005;434(7037):1144-8.
64. Pardanani AD, Levine RL, Lasho T, Pikman Y, Mesa RA, Wadleigh M, et al. MPL515 mutations in myeloproliferative and other myeloid disorders: a study of 1182 patients. *Blood.* 15 nov 2006;108(10):3472-6.

65. Klampfl T, Gisslinger H, Harutyunyan AS, Nivarthi H, Rumi E, Milosevic JD, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med.* 19 déc 2013;369(25):2379-90.
66. Campbell PJ, Green AR. The myeloproliferative disorders. *N Engl J Med.* 7 déc 2006;355(23):2452-66.
67. Grinfeld J, Nangalia J, Green AR. Molecular determinants of pathogenesis and clinical phenotype in myeloproliferative neoplasms. *Haematologica.* janv 2017;102(1):7-17.
68. Greenfield G, McMullin MF. Epigenetics in myeloproliferative neoplasms. *Front Oncol.* 2023;13:1206965.
69. Martin GS. The hunting of the Src. *Nat Rev Mol Cell Biol.* juin 2001;2(6):467-75.
70. Yaffe MB. Phosphotyrosine-binding domains in signal transduction. *Nat Rev Mol Cell Biol.* mars 2002;3(3):177-86.
71. Hanks SK, Quinn AM, Hunter T. The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science.* 1 juill 1988;241(4861):42-52.
72. Wilks AF. Two putative protein-tyrosine kinases identified by application of the polymerase chain reaction. *Proc Natl Acad Sci U S A.* mars 1989;86(5):1603-7.
73. Wilks AF, Harpur AG, Kurban RR, Ralph SJ, Zürcher G, Ziemiecki A. Two novel protein-tyrosine kinases, each with a second phosphotransferase-related catalytic domain, define a new class of protein kinase. *Mol Cell Biol.* avr 1991;11(4):2057-65.
74. Leonard WJ, O'Shea JJ. Jaks and STATs: biological implications. *Annu Rev Immunol.* 1998;16:293-322.
75. Velazquez L, Fellous M, Stark GR, Pellegrini S. A protein tyrosine kinase in the interferon alpha/beta signaling pathway. *Cell.* 24 juill 1992;70(2):313-22.
76. Hu X, Li J, Fu M, Zhao X, Wang W. The JAK/STAT signaling pathway: from bench to clinic. *Signal Transduct Target Ther.* 26 nov 2021;6(1):402.
77. Velazquez L, Mogensen KE, Barbieri G, Fellous M, Uzé G, Pellegrini S. Distinct domains of the protein tyrosine kinase tyk2 required for binding of interferon-alpha/beta and for signal transduction. *J Biol Chem.* 17 févr 1995;270(7):3327-34.
78. Frank SJ, Gilliland G, Kraft AS, Arnold CS. Interaction of the growth hormone receptor cytoplasmic domain with the JAK2 tyrosine kinase. *Endocrinology.* nov 1994;135(5):2228-39.
79. O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB, Laurence A. The JAK-STAT pathway: impact on human disease and therapeutic intervention. *Annu Rev Med.* 2015;66:311-28.
80. Argetsinger LS, Campbell GS, Yang X, Witthuhn BA, Silvennoinen O, Ihle JN, et al. Identification of JAK2 as a growth hormone receptor-associated tyrosine kinase. *Cell.* 30 juill 1993;74(2):237-44.
81. Witthuhn BA, Quelle FW, Silvennoinen O, Yi T, Tang B, Miura O, et al. JAK2 associates with the erythropoietin receptor and is tyrosine phosphorylated and activated following stimulation with erythropoietin. *Cell.* 30 juill 1993;74(2):227-36.
82. Akada H, Akada S, Hutchison RE, Sakamoto K, Wagner KU, Mohi G. Critical role of Jak2 in the maintenance and function of adult hematopoietic stem cells. *Stem Cells Dayt Ohio.* juill 2014;32(7):1878-89.
83. Wingelhofer B, Neubauer HA, Valent P, Han X, Constantinescu SN, Gunning PT, et al. Implications of STAT3 and STAT5 signaling on gene regulation and chromatin remodeling in hematopoietic cancer. *Leukemia.* 2018;32(8):1713-26.
84. Decker T. Emancipation from transcriptional latency: unphosphorylated STAT5 as guardian of hematopoietic differentiation. *EMBO J.* 15 mars 2016;35(6):555-7.
85. Orlova A, Wagner C, de Araujo ED, Bajusz D, Neubauer HA, Herling M, et al. Direct Targeting Options for STAT3 and STAT5 in Cancer. *Cancers.* 3 déc 2019;11(12).
86. Park HJ, Li J, Hannah R, Biddie S, Leal-Cervantes AI, Kirschner K, et al. Cytokine-induced megakaryocytic differentiation is regulated by genome-wide loss of a uSTAT transcriptional program. *EMBO J.* 15 mars 2016;35(6):580-94.
87. Lorenz U. SHP-1 and SHP-2 in T cells: two phosphatases functioning at many levels. *Immunol Rev.* mars 2009;228(1):342-59.
88. Qu CK. Role of the SHP-2 tyrosine phosphatase in cytokine-induced signaling and cellular response. *Biochim Biophys Acta.* 11 nov 2002;1592(3):297-301.
89. Wang D, Stravopodis D, Teglund S, Kitazawa J, Ihle JN. Naturally occurring dominant negative variants of Stat5. *Mol Cell Biol.* nov 1996;16(11):6141-8.

90. Shchelkunova A, Ermolinsky B, Boyle M, Mendez I, Lehker M, Martirosyan KS, et al. Tuning of alternative splicing--switch from proto-oncogene to tumor suppressor. *Int J Biol Sci.* 2013;9(1):45-54.
91. Ibrahim SSA, Huttunen KM. Orchestrated modulation of rheumatoid arthritis via crosstalking intracellular signaling pathways. *Inflammopharmacology.* août 2021;29(4):965-74.
92. Kralovics R, Passamonti F, Buser AS, Teo SS, Tiedt R, Passweg JR, et al. A gain-of-function mutation of JAK2 in myeloproliferative disorders. *N Engl J Med.* 28 avr 2005;352(17):1779-90.
93. Haider M, Gangat N, Lasho T, Abou Hussein AK, Elala YC, Hanson C, et al. Validation of the revised International Prognostic Score of Thrombosis for Essential Thrombocythemia (IPSET-thrombosis) in 585 Mayo Clinic patients. *Am J Hematol.* juin 2016;91(4):390-4.
94. Scott LM, Tong W, Levine RL, Scott MA, Beer PA, Stratton MR, et al. JAK2 exon 12 mutations in polycythemia vera and idiopathic erythrocytosis. *N Engl J Med.* 1 févr 2007;356(5):459-68.
95. Vainchenker W, Kralovics R. Genetic basis and molecular pathophysiology of classical myeloproliferative neoplasms. *Blood.* 9 févr 2017;129(6):667-79.
96. Yu H, Pardoll D, Jove R. STATs in cancer inflammation and immunity: a leading role for STAT3. *Nat Rev Cancer.* nov 2009;9(11):798-809.
97. O'Shea JJ, Schwartz DM, Villarino AV, Gadina M, McInnes IB, Laurence A. The JAK-STAT Pathway: Impact on Human Disease and Therapeutic Intervention. *Annu Rev Med.* 2015;66:311-28.
98. de Araujo ED, Erdogan F, Neubauer HA, Meneksedag-Erol D, Manaswiyoungkul P, Eram MS, et al. Structural and functional consequences of the STAT5BN642H driver mutation. *Nat Commun.* 7 juin 2019;10(1):2517.
99. Yang S, Wang L, Wu Y, Wu A, Huang F, Tang X, et al. Apoptosis in megakaryocytes: Safeguard and threat for thrombopoiesis. *Front Immunol.* 2022;13:1025945.
100. Li L, Ni R, Li Z, Ming Y, Liu L, Peng D, et al. Insights into Regulatory Factors in Megakaryocyte Development and Function: Basic Mechanisms and Potential Targets. *Front Biosci Landmark Ed.* 25 nov 2022;27(11):313.
101. El Gazzar M. microRNAs as potential regulators of myeloid-derived suppressor cell expansion. *Innate Immun.* avr 2014;20(3):227-38.
102. Tefferi A, Guglielmelli P, Larson DR, Finke C, Wassie EA, Pieri L, et al. Long-term survival and blast transformation in molecularly annotated essential thrombocythemia, polycythemia vera, and myelofibrosis. *Blood.* 16 oct 2014;124(16):2507-13; quiz 2615.
103. Hulcrantz M, Björkholm M, Dickman PW, Landgren O, Derolf ÅR, Kristinsson SY, et al. Risk for Arterial and Venous Thrombosis in Patients With Myeloproliferative Neoplasms: A Population-Based Cohort Study. *Ann Intern Med.* 6 mars 2018;168(5):317-25.
104. Barbui T, Tefferi A, Vannucchi AM, Passamonti F, Silver RT, Hoffman R, et al. Philadelphia chromosome-negative classical myeloproliferative neoplasms: revised management recommendations from European LeukemiaNet. *Leukemia.* mai 2018;32(5):1057-69.
105. Alimam S, Villiers W, Dillon R, Simpson M, Runglall M, Smith A, et al. Patients with triple-negative, JAK2V617F- and CALR-mutated essential thrombocythemia share a unique gene expression signature. *Blood Adv.* 23 févr 2021;5(4):1059-68.
106. Pérez C, Pascual M, Martín-Subero JI, Bellosillo B, Segura V, Delabesse E, et al. Aberrant DNA methylation profile of chronic and transformed classic Philadelphia-negative myeloproliferative neoplasms. *Haematologica.* sept 2013;98(9):1414-20.
107. Müller CI, Luong QT, Shih LY, Jones LC, Desmond JC, Kawamata N, et al. Identification of marker genes including RUNX3 (AML2) that discriminate between different myeloproliferative neoplasms and normal individuals. *Leukemia.* sept 2008;22(9):1773-8.
108. Kushner JP, Lee GR, Wintrobe MM, Cartwright GE. Idiopathic refractory sideroblastic anemia: clinical and laboratory investigation of 17 patients and review of the literature. *Medicine (Baltimore).* mai 1971;50(3):139-59.
109. Juneja SK, Imbert M, Jouault H, Scoazec JY, Sigaux F, Sultan C. Haematological features of primary myelodysplastic syndromes (PMDS) at initial presentation: a study of 118 cases. *J Clin Pathol.* oct 1983;36(10):1129-35.
110. Cheng DS, Kushner JP, Wintrobe MM. Idiopathic refractory sideroblastic anemia: incidence and risk factors for leukemic transformation. *Cancer.* août 1979;44(2):724-31.
111. Cazzola M, Barosi G, Gobbi PG, Invernizzi R, Riccardi A, Ascari E. Natural history of idiopathic refractory sideroblastic anemia. *Blood.* févr 1988;71(2):305-12.

112. Garand R, Gardais J, Bizet M, Bremond JL, Accard F, Callat MP, et al. Heterogeneity of acquired idiopathic sideroblastic anaemia (AISA). *Leuk Res.* 1992;16(5):463-8.
113. Streeter RR, Presant CA, Reinhard E. Prognostic significance of thrombocytosis in idiopathic sideroblastic anemia. *Blood.* sept 1977;50(3):427-32.
114. Fenaux P, Estienne MH, Lai JL, Beuscart R, Dupriez B, Jouet JP, et al. [Acquired idiopathic sideroblastic anemia. Apropos of 39 cases]. *Rev Med Interne.* août 1989;10(4):323-9.
115. Gupta R, Abdalla SH, Bain BJ. Thrombocytosis with sideroblastic erythropoiesis: a mixed myeloproliferative myelodysplastic syndrome. *Leuk Lymphoma.* août 1999;34(5-6):615-9.
116. Hossfeld DK. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Haematopoietic and Lymphoid Tissues. *Ann Oncol.* 1 mars 2002;13(3):490.
117. Raya JM, Arenillas L, Domingo A, Bellosillo B, Gutiérrez G, Luño E, et al. Refractory anemia with ringed sideroblasts associated with thrombocytosis: comparative analysis of marked with non-marked thrombocytosis, and relationship with JAK2 V617F mutational status. *Int J Hematol.* nov 2008;88(4):387-95.
118. Schmitt-Graeff A, Thiele J, Zuk I, Kvasnicka HM. Essential thrombocythemia with ringed sideroblasts: a heterogeneous spectrum of diseases, but not a distinct entity. *Haematologica.* avr 2002;87(4):392-9.
119. Pérez Sánchez I, Pérez Corrala A, Menarguez Palanca J, Mayayo Crespo M, Escudero Soto A, Pintado Cros T. Sideroblastic anaemia with reactive thrombocytosis versus myelodysplastic/myeloproliferative disease. *Leuk Lymphoma.* mars 2003;44(3):557-9.
120. Wang SA, Hasserjian RP, Loew JM, Sechman EV, Jones D, Hao S, et al. Refractory anemia with ringed sideroblasts associated with marked thrombocytosis harbors JAK2 mutation and shows overlapping myeloproliferative and myelodysplastic features. *Leukemia.* sept 2006;20(9):1641-4.
121. Hellström-Lindberg E, Cazzola M. The role of JAK2 mutations in RARS and other MDS. *Hematol Am Soc Hematol Educ Program.* 2008;52-9.
122. Vardiman JW, Thiele J, Arber DA, Brunning RD, Borowitz MJ, Porwit A, et al. The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes. *Blood.* 30 juill 2009;114(5):937-51.
123. Szpurka H, Tiu R, Murugesan G, Aboudola S, Hsi ED, Theil KS, et al. Refractory anemia with ringed sideroblasts associated with marked thrombocytosis (RARS-T), another myeloproliferative condition characterized by JAK2 V617F mutation. *Blood.* 1 oct 2006;108(7):2173-81.
124. Boissinot M, Garand R, Hamidou M, Hermouet S. The JAK2-V617F mutation and essential thrombocythemia features in a subset of patients with refractory anemia with ring sideroblasts (RARS). *Blood.* 1 sept 2006;108(5):1781-2.
125. Ceesay MM, Lea NC, Ingram W, Westwood NB, Gäken J, Mohamedali A, et al. The JAK2 V617F mutation is rare in RARS but common in RARS-T. *Leukemia.* nov 2006;20(11):2060-1.
126. Gattermann N, Billiet J, Kronenwett R, Zipperer E, Germing U, Nollet F, et al. High frequency of the JAK2 V617F mutation in patients with thrombocytosis (platelet count >600x10⁹/L) and ringed sideroblasts more than 15% considered as MDS/MPD, unclassifiable. *Blood.* 1 févr 2007;109(3):1334-5.
127. Shaw GR. Ringed sideroblasts with thrombocytosis: an uncommon mixed myelodysplastic/myeloproliferative disease of older adults. *Br J Haematol.* oct 2005;131(2):180-4.
128. Renneville A, Quesnel B, Charpentier A, Terriou L, Crinquette A, Lai JL, et al. High occurrence of JAK2 V617 mutation in refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Leukemia.* nov 2006;20(11):2067-70.
129. Remacha AF, Nomdedéu JF, Puget G, Estivill C, Sarda MP, Canals C, et al. Occurrence of the JAK2 V617F mutation in the WHO provisional entity: myelodysplastic/myeloproliferative disease, unclassifiable-refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Haematologica.* mai 2006;91(5):719-20.
130. Steensma DP, Dewald GW, Lasho TL, Powell HL, McClure RF, Levine RL, et al. The JAK2 V617F activating tyrosine kinase mutation is an infrequent event in both « atypical » myeloproliferative disorders and myelodysplastic syndromes. *Blood.* 15 août 2005;106(4):1207-9.
131. Broséus J, Lippert E, Harutyunyan AS, Jeromin S, Zipperer E, Florensa L, et al. Low rate of calreticulin mutations in refractory anaemia with ring sideroblasts and marked thrombocytosis. *Leukemia.* juin 2014;28(6):1374-6.
132. Jeromin S, Haferlach T, Weissmann S, Meggendorfer M, Eder C, Nadarajah N, et al. Refractory anemia with ring sideroblasts and marked thrombocytosis cases harbor mutations in SF3B1 or other

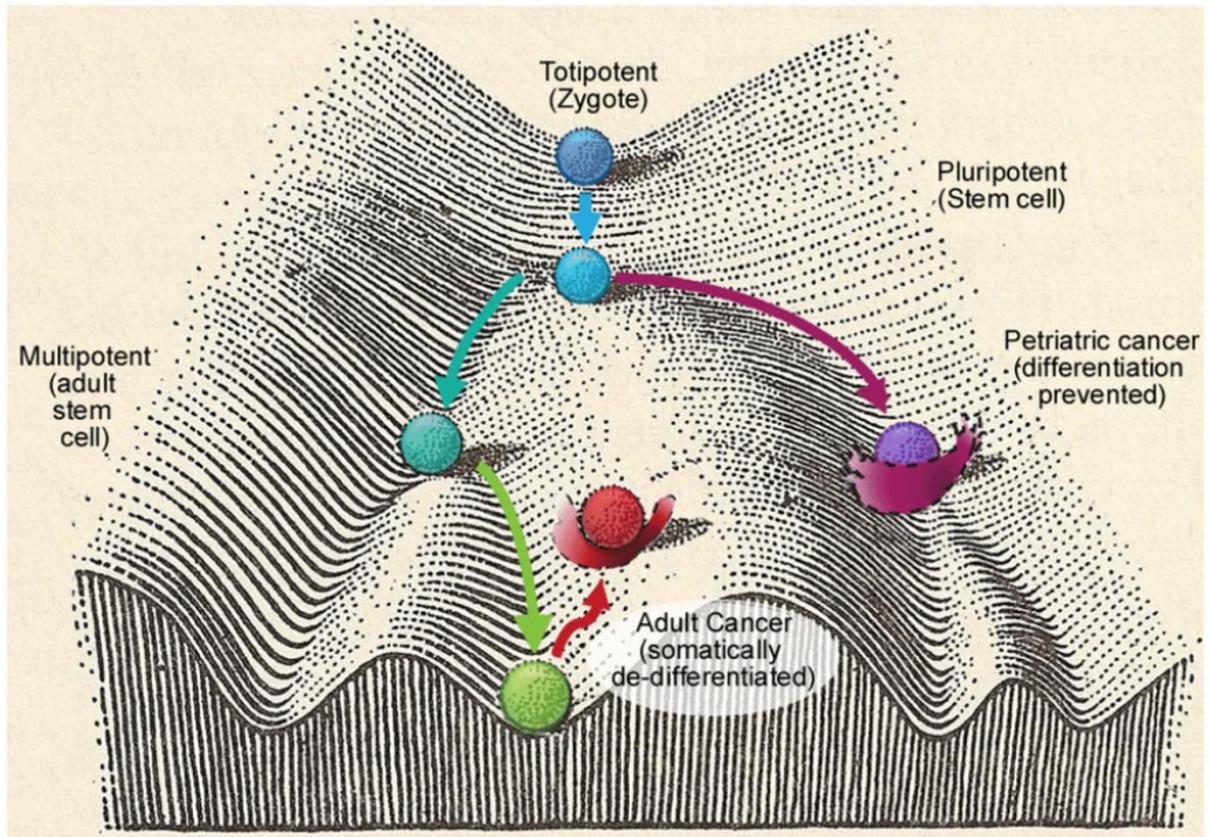
- spliceosome genes accompanied by JAK2V617F and ASXL1 mutations. *Haematologica*. avr 2015;100(4):e125-127.
133. Nangalia J, Massie CE, Baxter EJ, Nice FL, Gundem G, Wedge DC, et al. Somatic CALR mutations in myeloproliferative neoplasms with nonmutated JAK2. *N Engl J Med*. 19 déc 2013;369(25):2391-405.
 134. Patnaik MM, Belachew A, Finke C, Lasho TL, Hanson CA, Tefferi A. CALR mutations are infrequent in WHO-defined refractory anemia with ring sideroblasts. *Leukemia*. juin 2014;28(6):1370-1.
 135. Schnittger S, Bacher U, Haferlach C, Dengler R, Kröber A, Kern W, et al. Detection of an MPLW515 mutation in a case with features of both essential thrombocythemia and refractory anemia with ringed sideroblasts and thrombocytosis. *Leukemia*. févr 2008;22(2):453-5.
 136. Broseus J, Florensa L, Zipperer E, Schnittger S, Malcovati L, Richebourg S, et al. Clinical features and course of refractory anemia with ring sideroblasts associated with marked thrombocytosis. *Haematologica*. juill 2012;97(7):1036-41.
 137. Patnaik MM, Lasho TL, Finke CM, Hanson CA, King RL, Ketterling RP, et al. Predictors of survival in refractory anemia with ring sideroblasts and thrombocytosis (RARS-T) and the role of next-generation sequencing. *Am J Hematol*. mai 2016;91(5):492-8.
 138. Divoux M, Plocque A, Sevin M, Voillat L, Feugier P, Guerci-Bresler A, et al. Efficacy of lenalidomide in myelodysplastic/myeloproliferative neoplasms with ring sideroblasts and an extreme platelet count. *Clin Case Rep*. sept 2020;8(9):1774-80.
 139. Komrokji R, Melody M, Al Ali N, Chan O, Klimek V, Ball BJ, et al. Treatment outcomes for patients with myelodysplastic syndrome/myeloproliferative neoplasms with ring sideroblasts and thrombocytosis. *Leuk Lymphoma*. janv 2022;63(1):199-204.
 140. Majewski J, Schwartzenuber J, Lalonde E, Montpetit A, Jabado N. What can exome sequencing do for you? *J Med Genet*. sept 2011;48(9):580-9.
 141. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 4 mars 2012;9(4):357-9.
 142. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinforma Oxf Engl*. 15 juill 2009;25(14):1754-60.
 143. Meissner A, Gnirke A, Bell GW, Ramsahoye B, Lander ES, Jaenisch R. Reduced representation bisulfite sequencing for comparative high-resolution DNA methylation analysis. *Nucleic Acids Res*. 2005;33(18):5868-77.
 144. Metzker ML. Sequencing technologies - the next generation. *Nat Rev Genet*. janv 2010;11(1):31-46.
 145. Zerbino DR, Achuthan P, Akanni W, Amode MR, Barrell D, Bhai J, et al. Ensembl 2018. *Nucleic Acids Res*. 4 janv 2018;46(D1):D754-61.
 146. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. *Genome Biol*. 6 juin 2016;17(1):122.
 147. Tsai YT, Huang YP, Yu CT, Lu CL. MuSiC: a tool for multiple sequence alignment with constraints. *Bioinforma Oxf Engl*. 22 sept 2004;20(14):2309-11.
 148. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res*. 1 juill 2003;31(13):3568-71.
 149. Gu H, Bock C, Mikkelsen TS, Jäger N, Smith ZD, Tomazou E, et al. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat Methods*. févr 2010;7(2):133-6.
 150. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinforma Oxf Engl*. 1 juin 2011;27(11):1571-2.
 151. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol*. 3 oct 2012;13(10):R87.
 152. Kuhn RM, Karolchik D, Zweig AS, Wang T, Smith KE, Rosenbloom KR, et al. The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res*. janv 2009;37(Database issue):D755-761.
 153. Braschi B, Denny P, Gray K, Jones T, Seal R, Tweedie S, et al. Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Res*. 8 janv 2019;47(D1):D786-92.
 154. Ghosh D. Incorporating the empirical null hypothesis into the Benjamini-Hochberg procedure. *Stat Appl Genet Mol Biol*. 26 juill 2012;11(4):j/sagmb.2012.11.issue-4/1544-6115.1735/1544-6115.1735.xml.

155. Plotz G, Casper M, Raedle J, Hinrichsen I, Heckel V, Brieger A, et al. MUTYH gene expression and alternative splicing in controls and polyposis patients. *Hum Mutat.* juill 2012;33(7):1067-74.
156. Kamiza AB, Hsieh LL, Tang R, Chien HT, Lai CH, Chiu LL, et al. Polymorphisms of DNA repair genes are associated with colorectal cancer in patients with Lynch syndrome. *Mol Genet Genomic Med.* 17 avr 2018;6(4):533-40.
157. Fan L, Zhang A, Deng P. LMX1B mRNA expression and its gene body CpG methylation are valuable prognostic biomarkers for laryngeal squamous cell carcinoma. *Biomed Pharmacother Biomedecine Pharmacother.* sept 2019;117:109174.
158. Andersson EI, Tanahashi T, Sekiguchi N, Gasparini VR, Bortoluzzi S, Kawakami T, et al. High incidence of activating STAT5B mutations in CD4-positive T-cell large granular lymphocyte leukemia. *Blood.* 17 nov 2016;128(20):2465-8.
159. Teramo A, Barilà G, Calabretto G, Vicenzetto C, Gasparini VR, Semenzato G, et al. Insights Into Genetic Landscape of Large Granular Lymphocyte Leukemia. *Front Oncol.* 2020;10:152.
160. Xia Z, Sait SN, Baer MR, Barcos M, Donohue KA, Lawrence D, et al. Truncated STAT proteins are prevalent at relapse of acute myeloid leukemia. *Leuk Res.* juin 2001;25(6):473-82.
161. Kruse EA, Loughran SJ, Baldwin TM, Josefsson EC, Ellis S, Watson DK, et al. Dual requirement for the ETS transcription factors Fli-1 and Erg in hematopoietic stem cells and the megakaryocyte lineage. *Proc Natl Acad Sci U S A.* 18 août 2009;106(33):13814-9.
162. Pilati C, Shinde J, Alexandrov LB, Assié G, André T, Hélias-Rodzewicz Z, et al. Mutational signature analysis identifies MUTYH deficiency in colorectal cancers and adrenocortical carcinomas. *J Pathol.* mai 2017;242(1):10-5.
163. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SAJR, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. *Nature.* 22 août 2013;500(7463):415-21.
164. Liu W, Yu WM, Zhang J, Chan RJ, Loh ML, Zhang Z, et al. Inhibition of the Gab2/PI3K/mTOR signaling ameliorates myeloid malignancy caused by Ptpn11 (Shp2) gain-of-function mutations. *Leukemia.* juin 2017;31(6):1415-22.
165. Kesarwani M, Kincaid Z, Azhar M, Menke J, Schwieterman J, Ansari S, et al. MAPK-negative feedback regulation confers dependence to JAK2V617F signaling. *Leukemia.* août 2023;37(8):1686-97.
166. Sun P, Sun X, Zhao W, Ren M, Zhang C, Wang Z, et al. Lemur Tyrosine Kinase-3 Suppresses Growth of Prostate Cancer Via the AKT and MAPK Signaling Pathways. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol.* 2017;42(6):2582-92.
167. Han F, Hu M, Zhang L, Fan X, Wang J, Lou Z, et al. A-to-I RNA editing of BLCAP promotes cell proliferation by losing the inhibitory of Rb1 in colorectal cancer. *Exp Cell Res.* 15 août 2022;417(2):113209.
168. Kunimoto H, Meydan C, Nazir A, Whitfield J, Shank K, Rapaport F, et al. Cooperative Epigenetic Remodeling by TET2 Loss and NRAS Mutation Drives Myeloid Transformation and MEK Inhibitor Sensitivity. *Cancer Cell.* 8 janv 2018;33(1):44-59.e8.
169. Liang J, Lei P, Su X, Gao J, Ren B, Zhang Y, et al. Dauricine Inhibits Non-small Cell Lung Cancer Development by Regulating PTEN/AKT/mTOR and Ras/MEK1/2/ERK1/2 Pathways in a FLT4-dependent Manner. *Curr Cancer Drug Targets.* 2024;24(11):1157-68.
170. Zang C, Zhao F, Pu Y. LMX1B involved in the radioresistance, proliferation and migration of esophageal cancer cells. *Biomed Pharmacother Biomedecine Pharmacother.* oct 2019;118:109358.
171. Zhang K, Liu L, Wang M, Yang M, Li X, Xia X, et al. A novel function of IMPA2, plays a tumor-promoting role in cervical cancer. *Cell Death Dis.* 14 mai 2020;11(5):371.
172. Zhang Q, Deng S, Li Q, Wang G, Guo Z, Zhu D. Glycoprotein M6A Suppresses Lung Adenocarcinoma Progression via Inhibition of the PI3K/AKT Pathway. *J Oncol.* 2022;2022:4601501.
173. Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.* févr 2016;26(2):256-62.
174. Hebestreit K, Dugas M, Klein HU. Detection of significantly differentially methylated regions in targeted bisulfite sequencing data. *Bioinforma Oxf Engl.* 1 juill 2013;29(13):1647-53.
175. Liu Y, Han Y, Zhou L, Pan X, Sun X, Liu Y, et al. A comprehensive evaluation of computational tools to identify differential methylation regions using RRBS data. *Genomics.* nov 2020;112(6):4567-76.
176. Yang Y, Liu H, Liu Y, Zhou L, Zheng X, Yue R, et al. E-value: a superior alternative to P-value and its adjustments in DNA methylation studies. *Brief Bioinform.* 20 juill 2023;24(4):bbad241.

177. Rossi C, Madl P, Foletti A, Mocenni C. Equilibrium and far-from equilibrium states. In 2015. p. 71-94.
178. Asaulenko ZP, Spiridonov IN, Baram DV, Krivolapov YA. [WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues, 2022 (5th edition): Myeloid and Histiocytic Tumors]. *Arkh Patol.* 2023;85(5):36-44.

ANNEXES

Annexe 1. Epigenomic landscape selon Waddington.



Annexe 2. Liste des 53 variants dans 46 gènes candidats identifiés dans la cohorte de screening et constituent le panel de la cohorte de validation (GRCh37).

Chr : chromosome ; SNP : single nucleotide polymorphism ; Mis : Missense variant (variant faux-sens) ; If_Ins : In frame insertion (insertion sans décalage du cadre de lecture) ; If_Del : In frame deletion (délétion sans décalage du cadre de lecture) ; Fs_Var: Frameshift variant (variant avec décalage du cadre de lecture) ; Spl: Splice site (site d'épissage) ; Stop : variant non-sens.

Gène	Chr	Position	Type de variant	Allèle de référence	Allèle fraction CD3+	Allèle fraction CD3-	dbSNP
ADRA2C	4	3769414	Mis/SNP	C	C	A	rs765307281
ANO7	2	242157293	Mis/SNP	C	C	A	novel
ANO7	2	242142799	Mis/SNP	C	C	A	rs368075617
ARHGEF39	9	35665121	Mis/SNP	G	G	T	novel
ATXN1	6	16327867	If Ins	-	TGC	TGC	rs765686161
ATXN1	6	16327913	If Del	TGA	-	-	rs751377396
BMP4	14	54418601	Mis/SNP	C	C	A	rs747238009
BRAT1	7	2579180	Mis/SNP	C	C	A	novel
CBL	11	119077233	If Del	CAC	CAC	-	rs778672208
CDK6	7	92462634	Ns/SNP	C	C	A	novel
CEBPZ	2	37454859	Mis/SNP	C	C	A	
COPG1	3	128974928	Mis/SNP	C	C	A	novel
DCP1B	12	2062336	If Ins	-	-	GCG	novel
DCP1B	12	2062340	If Ins	-	-	CTT	novel
DOT1L	19	2227046	Mis/SNP	A	A	G	novel
DTX1	12	113531017	Mis/SNP	C	C	A	novel
EPHA10	1	38184063	Mis/SNP	C	C	A	rs747802363
ETV2	19	36135558	Mis/SNP	C	C	A	rs747982924
FGF18	5	170883731	Mis/SNP	G	G	T	rs755637038
HCN3	1	155257913	Mis/SNP	G	G	T	rs767366757
HCN3	1	155258117	Mis/SNP	G	G	T	novel
HTR6	1	20005583	Mis/SNP	G	G	T	rs767832258
LMTK3	19	49002852	Mis/SNP	C	C	A	rs755811752
LMTK3	19	49003067	Mis/SNP	G	G	T	novel
LPHN1	19	14266331	Mis/SNP	G	G	T	rs755607962

LTK	15	41804124	Mis/SNP	G	G	T	novel
LURAP1L	9	12775855	If Ins	-	-	GGTGGG	
MAG1	3	65425563	If Ins	-	CTG	CTG	rs773463513
MAML2	11	95825372	If Del	TGT	TGT	-	rs766133743
MAPK7	17	19285118	Mis/SNP	C	C	A	rs748090158
MLK4	1	233497836	Mis/SNP	C	C	A	novel
NFKBIE	6	44233377	Mis/SNP	C	C	A	novel
NUTF2	16	67899046	Mis/SNP	C	C	A	rs751898570
OGFR	20	61444540	Mis/SNP	A	A	C	rs61729550
PHLPP1	18	60383814	Mis/SNP	G	G	T	novel
PLEKHG3	14	65208963	Mis/SNP	C	C	A	
POU4F1	13	79176484	If Del	TGG	TGG	-	rs762014663
POU4F2	4	147560469	If Ins	-	GGC	GGC	
SATB1	3	18458570	Mis/SNP	C	C	A	novel
SGK223	8	8234872	If Ins	-	-	CTGCCT	novel
SH2B3	12	111885323	Mis/SNP	C	C	T	novel
SIRT7	17	79875732	Ns/SNP	C	C	A	novel
STAT5B	17	40369418	Mis/SNP	G	G	T	novel
TET2	4	106157244	Fs Del	CT	CT	-	
TET2	4	106164794	Mis/SNP	G	G	T	novel
TET2	4	106190782	Fs Del	AG	AG	-	
TET2	4	106158509	Spl	G	G	C	rs772634266
TNFAIP6	2	152236047	Mis/SNP	AA	-	-	rs753982588
TNFRSF19	13	24233255	Mis/SNP	C	C	A	novel
TRIO	5	14488224	Mis/SNP	G	G	T	rs375694147
TRPV3	17	3438998	Mis/SNP	G	G	T	rs776957787
UBXN11	1	26608852	Mis/SNP	G	G	A	rs61775089
WDR37	10	1170957	Mis/SNP	G	G	T	

Annexe 3. Séquençage ciblé de la cohorte de validation de 42 MDS/MPN-SF3B1-T

Pat : patient ; Chr : chromosome ; VAF : Variant allele frequency ; Mis : Missense variant (variant faux-sens) ; If_Ins : In frame insertion (insertion sans décalage du cadre de lecture) ; If_Del : In frame deletion (délétion sans décalage du cadre de lecture) ; Fs_Var: Frameshift variant (variant avec décalage du cadre de lecture) ; Spl: Splice site (site d'épissage) ; Stop : variant non-sens.

Pat	Gène	Chr	Position	Allèle de référence	Allèle alternatif	VAF (%)	Type de variant	HGVSc
1	SF3B1	2	198266834	T	C	39	Mis	ENST00000335508.6:c.2098A>G
2	LMTK3	19	49002561	A	C	18	Mis	ENST00000270238.3:c.1852T>G
	NFKBIE	6	44233152	T	G	13	Mis	ENST00000275015.5:c.349A>C
3	PHLPP1	18	60383003	G	GGCAGCA	51	If_Ins	ENST00000262719.5:c.99_104dup
	SF3B1	2	198267484	G	A	43	Mis	ENST00000335508.6:c.1873C>T
	SGK223	8	8235398	T	G	11	Mis	ENST00000520004.1:c.521A>C
	TET2	4	106190797	C	T	13	Mis	ENST00000540549.1:c.4075C>T
	WDR37	10	1149762	T	C	48	Mis	ENST00000358220.1:c.947T>C
	BRAT1	7	2582895	C	G	47	Mis	ENST00000340611.4:c.866G>C
	SF3B1	2	198266834	T	C	41	Mis	ENST00000335508.6:c.2098A>G
4	SH2B3	12	111885508	C	G	49	Mis	ENST00000341259.2:c.1285C>G
	ATXN1	6	16327916	T	TGCTGCTGCTGCTGCT	49	If_Ins	ENST00000244769.4:c.612_626dup
5	DOT1L	19	2226731	C	A	47	Mis	ENST00000398665.3:c.4211C>A
	LMTK3	19	49001295	A	C	8	Mis	ENST00000270238.3:c.3118T>G
	SF3B1	2	198267491	C	A	42	Mis	ENST00000335508.6:c.1866G>T
	TET2	4	106196834	C	T	58	Mis	ENST00000540549.1:c.5167C>T
	TET2	4	106157698	T	C	36	Mis	ENST00000540549.1:c.2599T>C
	UBXN11	1	26629304	C	T	54	Mis,Spl	ENST00000374222.1:c.71G>A
	SF3B1	2	198266834	T	C	43	Mis	ENST00000335508.6:c.2098A>G
6	STAT5B	17	40354787	T	A	23	Mis	ENST00000293328.3:c.2117A>T
	BRAT1	7	2584629	G	A	50	Mis	ENST00000340611.4:c.344C>T
	MAGI1	3	65456135	T	C	46	Mis	ENST00000402939.2:c.782A>G
	PLEKHG3	14	65208879	C	G	53	Mis	ENST00000247226.7:c.2476C>G
	SF3B1	2	198266834	T	C	34	Mis	ENST00000335508.6:c.2098A>G

7	SF3B1	2	198267371	G	C	36	Mis	ENST00000335508.6:c.1986C>G
8	ATXN1	6	16327770	C	A	45	Mis	ENST00000244769.4:c.772G>T
	ATXN1	6	16327916	T	TGCTGCTGCT	42	If_Ins	ENST00000244769.4:c.618_626dup
	SF3B1	2	198267491	C	G	41	Mis	ENST00000335508.6:c.1866G>C
	TET2	4	106196770	G	A	47	Mis	ENST00000540549.1:c.5103G>A
9	COPG1	3	128994049	A	G	54	Mis	ENST00000314797.6:c.2408A>G
	SF3B1	2	198267491	C	G	45	Mis	ENST00000335508.6:c.1866G>C
10	ADRA2C	4	3768389	A	C	10	Mis	ENST00000330055.5:c.56A>C
	LTK	15	41805009	T	G	10	Mis	ENST00000263800.6:c.255A>C
	OGFR	20	61443643	G	A	48	Mis	ENST00000290291.6:c.676G>A
	PLEKHG3	14	65205465	G	C	44	Mis	ENST00000247226.7:c.1242G>C
	SF3B1	2	198266834	T	C	42	Mis	ENST00000335508.6:c.2098A>G
	TET2	4	106190797	C	T	41	Mis	ENST00000540549.1:c.4075C>T
	TET2	4	106156330	CCTCTT	C	9	Fs_Var	ENST00000540549.1:c.1233_1237del
11	BRAT1	7	2577977	G	C	44	Mis	ENST00000340611.4:c.2192C>G
	EPHA10	1	38227127	C	T	49	Mis	ENST00000373048.4:c.800G>A
	SF3B1	2	198266834	T	C	27	Mis	ENST00000335508.6:c.2098A>G
12	PHLPP1	18	60383872	C	T	0,25	Mis	ENST00000262719.5:c.956C>T
	SF3B1	2	198267491	C	A	28	Mis	ENST00000335508.6:c.1866G>T
	TET2	4	106196834	C	T	35	Mis	ENST00000540549.1:c.5167C>T
	TET2	4	106157698	T	C	42	Mis	ENST00000540549.1:c.2599T>C
13	SF3B1	2	198267371	G	C	39	Mis	ENST00000335508.6:c.1986C>G
14	SF3B1	2	198266494	T	C	45	Mis	ENST00000335508.6:c.2342A>G
	TET2	4	106156747	C	T	31	Stop	ENST00000540549.1:c.1648C>T
15	DOT1L	19	2211776	T	G	18	Mis	ENST00000398665.3:c.1492T>G
	SF3B1	2	198266834	T	C	41	Mis	ENST00000335508.6:c.2098A>G
	SGK223	8	8185869	T	G	12	Mis	ENST00000520004.1:c.2423A>C
16	SF3B1	2	198266834	T	C	44	Mis	ENST00000335508.6:c.2098A>G
	TRIO	5	14487860	T	G	19	Mis	ENST00000344204.4:c.7123T>G
17	ATXN1	6	16327916	T	TGCTGCTGCTGCTGCT	40	If_Ins	ENST00000244769.4:c.612_626dup

	SF3B1	2	198267491	C	G	38	Mis	ENST00000335508.6:c.1866G>C
18	SF3B1	2	198267371	G	T	50	Mis	ENST00000335508.6:c.1986C>A
19	LMTK3	19	49002587	A	C	14	Mis	ENST00000270238.3:c.1826T>G
	SF3B1	2	198267484	G	A	47	Mis	ENST00000335508.6:c.1873C>T
20	ATXN1	6	16327827	G	T	43	Mis	ENST00000244769.4:c.715C>A
	CBL	11	119148478	C	T	34	Mis	ENST00000264033.4:c.1019C>T
	PLEKHG3	14	65205465	G	C	50	Mis	ENST00000247226.7:c.1242G>C
	SF3B1	2	198266834	T	C	37	Mis	ENST00000335508.6:c.2098A>G
21	LTK	15	41797702	C	T	49	Mis	ENST00000263800.6:c.1724G>A
	SF3B1	2	198266834	T	C	34	Mis	ENST00000335508.6:c.2098A>G
	TRIO	5	14143875	C	T	45	Mis	ENST00000344204.4:c.41C>T
22	ATXN1	6	16327915	A	ATGC	41	If Ins	ENST00000244769.4:c.624_626dup
	MAPK7	17	19284104	T	G	7	Mis	ENST00000308406.5:c.582T>G
	PHLPP1	18	60646054	G	A	39	Mis	ENST00000262719.5:c.4544G>A
	SF3B1	2	198267371	G	T	38	Mis	ENST00000335508.6:c.1986C>A
	SH2B3	12	111885228	G	GA	27	Fs_Var	ENST00000341259.2:c.1119dup
	UBXN11	1	26620806	T	C	56	Mis	ENST00000374222.1:c.449A>G
23	SF3B1	2	198266834	T	C	25	Mis	ENST00000335508.6:c.2098A>G
24	CBL	11	119169085	G	A	48	Mis	ENST00000264033.4:c.2269G>A
	PLEKHG3	14	65210369	G	A	4	Mis	ENST00000247226.7:c.3440G>A
	POU4F1	13	79176483	C	CTGG	48	If Ins	ENST00000377208.5:c.324_326dup
	SF3B1	2	198266834	T	C	43	Mis	ENST00000335508.6:c.2098A>G
25	SF3B1	2	198267360	T	C	34	Mis	ENST00000335508.6:c.1997A>G
26	BRAT1	7	2582895	C	G	56	Mis	ENST00000340611.4:c.866G>C
27	PLEKHG3	14	65209909	G	A	45	Mis	ENST00000247226.7:c.2980G>A
	SF3B1	2	198267491	C	G	29	Mis	ENST00000335508.6:c.1866G>C

28	POU4F1	13	79175619	G	C	54	Mis	ENST00000377208.5:c.1191C>G
	SF3B1	2	198266494	T	C	45	Mis	ENST00000335508.6:c.2342A>G
	TET2	4	106182974	AGAAACTT	A	21	Fs Var	ENST00000540549.1:c.4015_4021del
	TNFRSF19	13	24243162	G	A	42	Mis	ENST00000382258.4:c.1171G>A
29	SF3B1	2	198266834	T	C	48	Mis	ENST00000335508.6:c.2098A>G
	SH2B3	12	111856273	CG	C	7	Fs Var	ENST00000341259.2:c.326del
	TRPV3	17	3432281	A	C	13	Mis	ENST00000301365.4:c.1251T>G
30	COPG1	3	128971732	G	A	48	Mis	ENST00000314797.6:c.257G>A
	LTK	15	41797721	G	T	45	Mis	ENST00000263800.6:c.1705C>A
	SF3B1	2	198266834	T	C	35	Mis	ENST00000335508.6:c.2098A>G
31	SF3B1	2	198267360	T	C	45	Mis	ENST00000335508.6:c.1997A>G
32	LURAP1L	9	12821695	T	C	45	Mis	ENST00000319264.3:c.623T>C
	SF3B1	2	198266834	T	C	46	Mis	ENST00000335508.6:c.2098A>G
	TRPV3	17	3436058	T	C	48	Mis	ENST00000301365.4:c.958A>G
33	LMTK3	19	49000877	AGCGGTGGCG	A	47	If Del	ENST00000270238.3:c.3527_3535del
	PHLPP1	18	60383003	G	GGCAGCA	37	If Ins	ENST00000262719.5:c.99_104dup
	POU4F2	4	147561147	C	A	38	Mis	ENST00000281321.3:c.417C>A
	SF3B1	2	198267491	C	G	43	Mis	ENST00000335508.6:c.1866G>C
	TRIO	5	14498281	C	T	42	Mis	ENST00000344204.4:c.8131C>T
34	ANO7	2	242142761	C	T	49	Mis	ENST00000274979.8:c.899C>T
	POU4F2	4	147561803	G	A	50	Mis	ENST00000281321.3:c.1073G>A
	SF3B1	2	198266834	T	C	35	Mis	ENST00000335508.6:c.2098A>G
35	DCP1B	12	2058442	A	G	45	Mis	ENST00000280665.6:c.1583T>C
	SF3B1	2	198266834	T	C	29	Mis	ENST00000335508.6:c.2098A>G
	TET2	4	106193748	C	T	54	Stop	ENST00000540549.1:c.4210C>T
36	POU4F1	13	79176269	C	T	44	Mis	ENST00000377208.5:c.541G>A
	SF3B1	2	198266834	T	C	45	Mis	ENST00000335508.6:c.2098A>G
37	ATXN1	6	16327915	A	ATGC	49	If Ins	ENST00000244769.4:c.624_626dup
	SF3B1	2	198267491	C	G	45	Mis	ENST00000335508.6:c.1866G>C
38	SF3B1	2	198266834	T	C	32	Mis	ENST00000335508.6:c.2098A>G

	UBXN11	1	26620806	T	C	47	Mis	ENST00000374222.1:c.449A>G
39	ATXN1	6	16327918	C	CTGA	48	If_Ins	ENST00000244769.4:c.623_624insTC A
	SF3B1	2	198267360	T	C	41	Mis	ENST00000335508.6:c.1997A>G
	STAT5B	17	40359659	T	A	41	Mis	ENST00000293328.3:c.1994A>T
	TRPV3	17	3432281	A	C	10	Mis	ENST00000301365.4:c.1251T>G
40	SF3B1	2	198267491	C	A	32	Mis	ENST00000335508.6:c.1866G>T
	TET2	4	106182919	G	T	33	Stop	ENST00000540549.1:c.3958G>T
41	BMP4	14	54418669	G	C	52	Mis	ENST00000245451.4:c.272C>G
	LURAP1L	9	12821695	T	C	61	Mis	ENST00000319264.3:c.623T>C
	MAGI1	3	65425582	CTGCTGCTG T	C	34	If_Del	ENST00000402939.2:c.1233_1241del
	PHLPP1	18	60612478	G	A	52	Mis	ENST00000262719.5:c.3298G>A
	SF3B1	2	198266834	T	C	41	Mis	ENST00000335508.6:c.2098A>G
42	SF3B1	2	198266834	T	C	48	Mis	ENST00000335508.6:c.2098A>G
	TET2	4	106156097	C	T	52	Mis	ENST00000540549.1:c.998C>T
	TET2	4	106197285	T	C	44	Mis	ENST00000540549.1:c.5618T>C

Annexe 4. Variants candidats de la cohorte de validation retenus pour annotations fonctionnelles.

Chr : chromosome

	Chr	position hg19	Allèle de référence	Allèle alternatif	conséquence	occurrence dans la cohorte 3 (%)
<u>BRATI</u>	7	2582895	C	G	missense_variant	9.52
		2584629	G	A	missense_variant	
		2577977	G	C	missense_variant	
		2582895	C	G	missense_variant	
<u>LMTK3</u>	19	49002561	A	C	missense_variant	9.52
		49001295	A	C	missense_variant	
		49002587	A	C	missense_variant	
		49000877	AGCGGTGGCG	A	inframe_deletion	
<u>LTK</u>	15	41805009	T	G	missense_variant	7.14
		41797702	C	T	missense_variant	
		41797721	G	T	missense_variant	
<u>POU4F1</u>	13	79176483	C	CTGG	inframe_insertion	9.52
		79175619	G	C	missense_variant	
		79176269	C	T	missense_variant	
<u>SGK223</u>	8	8235398	T	G	missense_variant	9.52
		8185869	T	G	missense_variant	
		8234904	C	T	missense_variant	
<u>STAT5B</u>	17	40354787	T	A	missense_variant	4.76
		40359659	T	A	missense_variant	

Annexe 5. Manuscrit pré-soumission. Divoux et al. DNA methylation comparative analysis in triple negative or *JAK2*^{V617F} MDS/MPN with *SF3B1* mutation and thrombocytosis, *JAK2*^{V617F} essential thrombocytemia and MDS with *SF3B1* mutation.

DNA methylation comparative analysis in triple negative or *JAK2*^{V617F} MDS/MPN with *SF3B1* mutation and thrombocytosis, *JAK2*^{V617F} essential thrombocytemia and MDS with *SF3B1* mutation

Marion Divoux¹, Ghislain Fievet¹, Natacha Dreumont¹, Laurence Schenone², Maud D'Aveni², François Girodon³, Pierre Feugier¹, Sébastien Hergalant¹, Julien Broséus¹

1. UMR Inserm 1256 NGERE (Nutrition, Génétique et Exposition aux Risques Environnementaux), Université de Lorraine, 54000 Nancy, France
2. Department of Haematology, CHRU Nancy, 54500 Nancy, France
3. Haematology Laboratory, CHRU Dijon, France

INTRODUCTION

Myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis (MDS/MPN-*SF3B1*-T) is a clonal hematopoietic stem cell disorder with overlapping features of myelodysplastic neoplasms with *SF3B1* mutation (MDS-*SF3B1*) and myeloproliferative neoplasms (MPN) like essential thrombocytemia (ET).

According to 2022 WHO (World Health Organization) classification (1), MDS-*SF3B1* or MDS with ringed sideroblasts (MDS-RS) corresponds to myelodysplastic neoplasm (MDS) with low bone marrow blasts associated with *SF3B1* mutation or at least 15% of ringed sideroblasts with exclusion of MDS with 5q deletion, 7 monosomy or complex karyotype. The management of MDS-*SF3B1* frequently consists with stimulating erythropoiesis agents or blood transfusion.

ET is a MPN characterized by a chronic primitive excess in platelets production due to megakaryocytic clonal expansion (1). Diagnosis of ET relies on major criterion including thrombocytosis over 450 G/l, histological bone marrow assessment showing megakaryocytic proliferation without fibrosis, differential diagnosis exclusion and the presence of mutation in *JAK2* (*JAK2*^{V617F}), *CALR* or *MPL*. Because of the thrombosis risk, cytoreduction may be indicated with targeted platelet count under 450 G/l. MDS/MPN-*SF3B1*-T is now considered as an individual entity and no more as the concomitant occurrence of ET and MDS-*SF3B1* but share strong pattern with them (**Table 1**).

Table 1. Comparative morphological and molecular description of TN MDS/M TN MDS/MPN-*SF3B1*-T, *JAK2* MDS/MPN-*SF3B1*-T, MDS-*SF3B1*, *JAK2* ET and BM.

JAK2 MDS/MPN-*SF3B1*-T = *JAK2*^{V617F} mutated myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis with mutated *JAK2*^{V617F}, TN MDS/MPN-*SF3B1*-T = triple negative myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, *JAK2* ET = *JAK2*^{V617F} mutated essential thrombocytemia, MDS-*SF3B1* = myelodysplastic neoplasm with mutated *SF3B1*, BM = healthy bone marrows.

	MDS- <i>SF3B1</i>	TN MDS/MPN- <i>SF3B1</i> -T	<i>JAK2</i> ^{V617F} MDS/MPN- <i>SF3B1</i> -T	<i>JAK2</i> ^{V617F} ET	BM
	<i>SF3B1</i> induced dysplasia	Thrombocytosis of undetermined origin <i>SF3B1</i> induced dysplasia	<i>JAK2</i> ^{V617F} induced thrombocytosis <i>SF3B1</i> induced dysplasia	<i>JAK2</i> ^{V617F} induced thrombocytosis	
<u>Myeloproliferative component</u> Giant megakaryocytes thrombocytosis 	no	yes	yes	yes	no
<i>JAK2</i>^{V617F} mutation	no	no	yes	yes	no
<i>MPL</i>^{W515K/L} mutation	no	no	no	no	no
<i>CALR</i> exon 9 mutation	no	no	no	no	no
<u>Myelodysplastic component</u> Dyserythropoiesis ringed sideroblasts 	yes	yes	yes	no	no
<i>SF3B1</i> mutation	yes	yes	yes	no	no
Comparative methylation profile 	?	?	?	?	?

The diagnosis relies on MDS-*SF3B1* and ET criterion, with platelet count over 450 G/l and the absence of MPN, MDS or other MDS/MPN history. The detection of *JAK2*^{V617F} mutation facilitates the diagnosis but is only detected in 10 to 50% of cases (2,3), even if platelet count over 600 G/l increases its prevalence (4). *CALR* and *MPL* are only mutated in 0 to 3% of cases (2,3,5–7). As in both MPN and MDS, mutations in *TET2*, *ASXL1*, *SETBP1* and *SH2B3* are also described (3).

Despite mutations in epigenetic regulatory elements are widely reported, few data about DNA methylation in MPN are available, especially for ET (8,9). DNA methylation analysis in MPN and healthy samples revealed a specific MPN signature involving inflammation and NF- κ B pathway regulatory regions. However, this signature was not specific of any MPN in particular (primary myelofibrosis or polycythemia vera or ET) (10). A multi-omic analysis was performed on peripheral blood mononuclear cells samples from ET, secondary myelofibrosis from ET and healthy volunteers. It included methylation chip analysis and RNA sequencing (9). Methylation analysis did not highlight significant differences between ET molecular subtypes ($JAK2^{V617F}$, mutated *CALR* or TN) but underlined differentially methylated regions (DMRs) and differentially expressed genes (DEGs) between all ET and healthy volunteers. This common expression signature for all ET molecular subtypes included overexpression of genes involved in myeloid differentiation, proliferation, oncogenesis and inflammation, especially in MAPK pathway. Nonetheless, DMRs and DEGs were poorly overlapped, and the authors hypothesized that these DMRs were related to regulatory element for DEGs. RNA sequencing on TE SET-2 cell line confirmed that hypothesis with overexpression of associated NF- κ B and MAPK transcription factors.

Platelet's proliferation remains unexplained in 50% of cases of MDS/MPN-*SF3B1*-T, leading to diagnosis difficulties. As some genomic molecular mechanisms of thrombocytosis are shared by ET and MDS/MPN-*SF3B1*-T, epigenetic patterns may also be common. We hypothesized that an epigenetic mechanism is responsible of thrombocytosis in TN MDS/MPN-*SF3B1*-T. To understand this mechanism, we explored comparative DNA methylation by reduced representation bisulfite sequencing (RRBS) in TN MDS/MPN-*SF3B1*-T, $JAK2^{V617F}$ mutated MDS/MPN-*SF3B1*-T ($JAK2^{V617F}$ MDS/MPN-*SF3B1*-T), MDS-*SF3B1* and ET.

MATERIALS AND METHODS

Biological collection and Clinical Data

Comparative analysis was performed with five groups including samples from patients with $JAK2^{V617F}$ MDS/MPN-*SF3B1*-T, TN MDS/MPN-*SF3B1*-T, $JAK2^{V617F}$ mutated ET ($JAK2^{V617F}$ ET), MDS-*SF3B1* and healthy bone marrow (BM) as control group.

Blood or bone marrow derived samples from Biological Resources Centers in Nancy and Dijon (France) and from the Nancy CHRU Genetics Department between 2007 and 2021 were included. Samples from biological resources centers were preserved after patient's consents. Samples from the Nancy CHRU Genetics Department were residual extracted DNA from classical care and patients received information and did not provide opposition to their utilization. All samples were anonymized. The study was conducted in accordance with local ethical guidelines.

For each patient, we collected data including: source type of sample, cellular type of sample, gender, age, medical background (neoplasia and cardiovascular risk factors especially), diagnosis, date of diagnosis, clinical features at diagnosis as splenomegaly, biological features at diagnosis as hemogram, bone marrow aspiration cytological analysis (including Perls coloration), molecular analysis (*JAK2*^{V617F}, *CALR*, *MPL* and *SF3B1* mutations), bone marrow karyotype, bone marrow biopsy (if available), evolution features as transfusions dependence, treatment, thrombosis occurrence, hemorrhage occurrence, secondary acute myeloid leukemia occurrence, vital status at last news.

Replicates from different sources

Due to the retrospective feature and the heterogeneity of the pathologies, methylation analysis included samples from different sources: total blood cells, total bone marrow cells, mononuclear blood cells, polynuclear blood cells, mononuclear bone marrow cells and polynuclear bone marrow cells. For some patients, different sources were available and enclosed in sequencing. To identify specific methylation signature for each cell's source, we analyzed with the R package *MethylKit* (11) 17 000 genes for which methylation data for CpG near initiation site of transcription were available.

Molecular characterization

Samples from Biological Resources Centers, DNA was manually extracted with *NucleoSpin Tissue kit* (Macherey-Nagel, Düren, Allemagne). Samples from the Nancy CHRU Genetics Department, DNA was extracted with *QIAamp DNA Blood Mini QIAcube Kit* (Qiagen, Hilden, Germany) on *Qiacube Automate* (Qiagen, Hilden, Germany). To better characterize and confirm mutational status of each sample, local routine "myeloid" targeted next generation sequencing (NGS) was performed. Targeted panel included the following genes: *ABL1* exons 4-9, *ASXL1* exons 10,12,13, *BRAF* exon 15, *CALR* exon 9, *CBL* exons 8,9, *CEBPA*, *CSF3R*, *DNMT3A*, *ETV6*, *EZH2*, *FLT3* exons 13-15,20, *HRAS* exons 2-3, *IDH1* exon 4, *IDH2* exon 4, *JAK2*, *KIT* exons 2,8-11,13,17,18, *KRAS* exons 2,3, *MPL* exon 10, *NPM1* exons 10,11, *NRAS* exons 2-3, *PTPN11* exons 3,7-13, *RUNX1*, *SETBP1* exon 4, *SF3B1* exons 10-16, *SRSF2* exon 1, *TET2*, *TP53*, *U2AF1* exons 2-6, *WT1* exons 6-10, *ZRSR2*.

DNA Methylation Analysis

We devised specific RRBS analysis pipeline for our cohort (**Figure 1**). After quality control on *Agilent Technologies® 2200 Tape Station* (Agilent, technologies, Santa Clara, CA, USA), 100 ng of extracted DNA were used to perform Reduced-representation bisulfite sequencing (RRBS) with *Diagenode's Premium RRBS technology kit* (Denville, New Jersey, USA) following the manufacturer's instructions, as previously described (12), with the following steps: DNA digestion with *MspI* restriction enzyme, end repair and Illumina adapters DNA ligation, selection of DNA fragments with insert sizes ranging between 40 and 220 bp, bisulfitation with *EZ DNA Methylation™ Kit* (Zymo research, Irvine, CA, USA), PCR amplification and sequencing using an *Illumina Nextseq 500/550 automate* (Illumina®, San Diego, CA, USA).

Base calling was performed as described with the Illumina pipeline module *bclConverter* process during 71 single end read cycles. The trim process was fulfilled with the *TrimGalore Tool* (Babraham Institute, University of Cambridge). The trimmed sequences were then mapped using the *Bismark* method (13) with the human genome (hg19) and bisulfite human genome. To keep sufficient information for alignment, only sequences over 30 pb length were selected for downstream analysis (length ranging 30 to 71 pb).

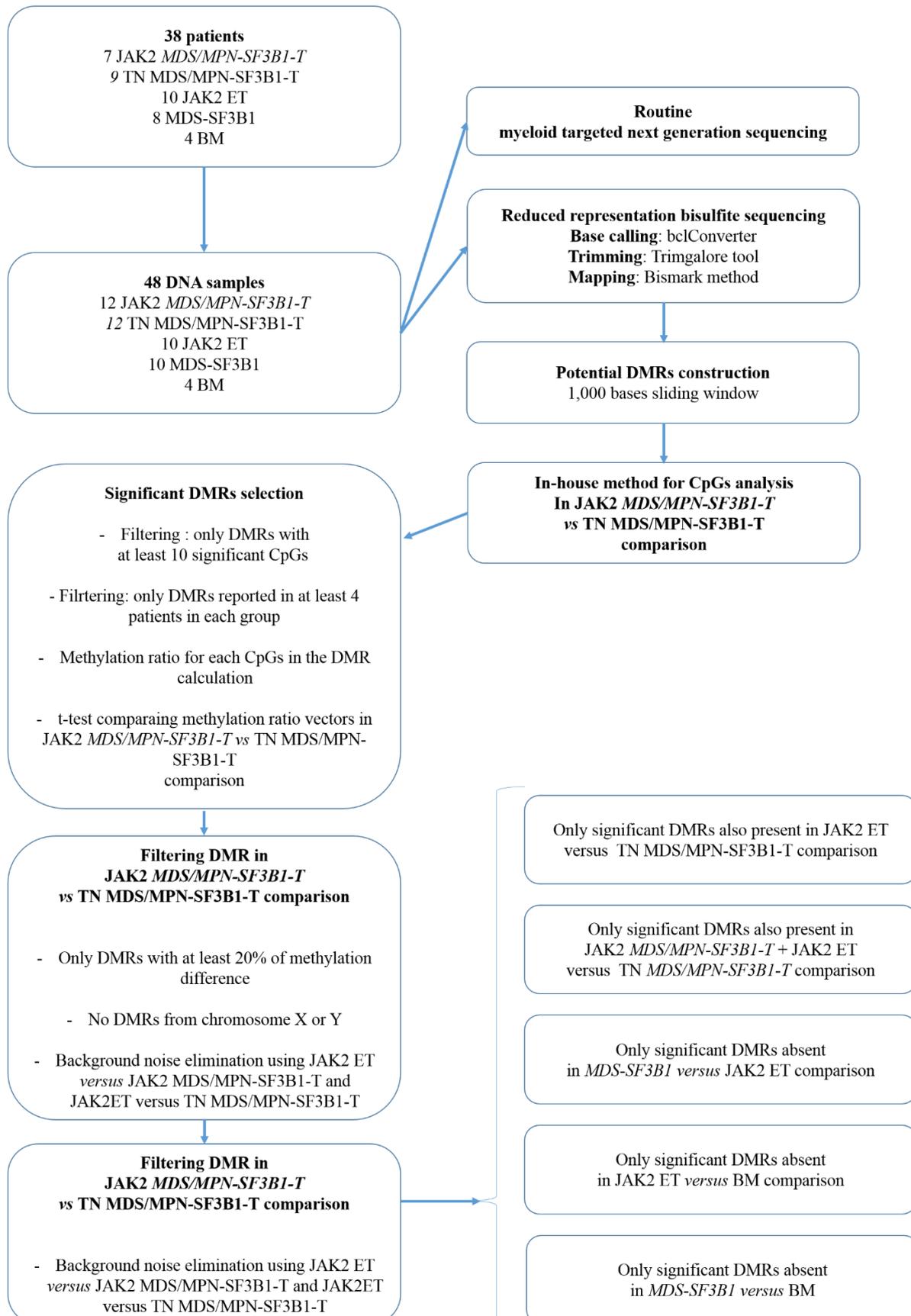


Figure 1. Reduced representation bisulfite sequencing analysis pipeline.

The .bam files were treated with *R v.4.1*. In CpGs analysis, the methylation rate was calculated for each CpG with at least 10 reads depth. CpGs with a depth over the 99.5th percentile were removed to avoid PCR bias. All CpGs on X chromosome were removed. Only CpGs with methylation data available for 2 samples per group were selected for methylation comparison. Comparison was performed using t student test (vector 1 corresponding to methylation ratios at the position for each sample of group 1 and vector 2 corresponding to methylation ratios at the position for each sample of group 2) and only significant CpGs with methylation difference over 20% were considered.

The DMRs analysis was inspired by the *Methylkit* method (11). Potential DMRs were defined with a 1,000 bases “sliding window” every 500 bases on the genome. Only sliding windows containing at least 10 significantly differentially methylated CpGs with minimal depth of 10 reads for at least 2 patients were selected for DMRs analysis. In every group, for each sample and each position, we calculated a methylation ratio. Potential DMRs were compared using t Student test (vector 1 corresponding to averages of methylation ratios for each position in the potential DMR for group 1 and vector 2 corresponding to averages of methylation ratios for each position in the potential DMR for group 2). P values were adjusted with Benjamini Hochberg method. For each comparison, DMRs were considered as significant if the methylation difference was at least over 20%. For patients with multiple samples, we selected the most relevant sample according to each comparison.

Some significant DMRs from the comparison MDS/MPN-*SF3B1*-T- *JAK2*^{V617F} versus TN MDS/MPN-*SF3B1*-T were considered as background noise and filtered: (i) common significant DMRs with the *JAK2*^{V617F} ET versus *JAK2*^{V617F} MDS/MPN-*SF3B1*-T comparison and the *JAK2*^{V617F} ET versus TN MDS/MPN-*SF3B1*-T. Those DMRs were considered as related to the non-specific dysplastic feature of the MDS/MPN-*SF3B1*-T; (ii) common significant DMRs with the MDS-*SF3B1* versus *JAK2*^{V617F} MDS/MPN-*SF3B1*-T and MDS-*SF3B1* versus TN MDS/MPN-*SF3B1*-T. Those DMRs were considered as related to the non-specific proliferative feature of the MDS/MPN-*SF3B1*-T. To identify the most relevant CpGs and DMRs, we applied for each CpG or DMR a five points clinical pertinence score, based on the presence or absence of the same significant CpGs or DMRs in other comparison (**Table 2, Figure 1**).

CpG and DMR annotations

Annotations were performed with both *HGNC* (14) and *UCSC RefGene Name* (15). If no annotation was available in *UCSC* database, the DMR was annotated with *HGNC*.

qPCR of the most relevant DMR associated genes

Available RNA for 4 samples (included in RRBS analysis) from MDS-*SF3B1*, TN MDS/MPN-*SF3B1*-T, *JAK2*^{V617F} MDS/MPN-*SF3B1*-T and BM allowed us to assess the relative transcript expression by quantitative polymerase chain reaction (qPCR) of 8 candidates genes. Amplification was performed in triplicates with a *master mix SYBR Green* (ThermoFisher, Foster City, California, USA) and specific primers. The qPCR reaction was performed on *7500 Fast Real System* device (Applied Biosystem®, ThermoFisher, Foster City, California, USA) and the datas were analyzed with *Step One Software*.

Table 2. Applied five points clinical pertinence score to significant DMRs from JAK2 MDS/MPN-SF3B1-T versus TN MDS/MPN-SF3B1-T comparison.

$JAK2^{V617F}$ MDS/MPN-SF3B1-T = $JAK2^{V617F}$ mutated myelodysplastic/myeloproliferative neoplasm with SF3B1 mutation and thrombocytosis, TN MDS/MPN-SF3B1-T = triple negative myelodysplastic/myeloproliferative neoplasm with SF3B1 mutation and thrombocytosis, $JAK2^{V617F}$ ET = $JAK2^{V617F}$ mutated essential thrombocytemia, MDS-SF3B1 = myelodysplastic neoplasm with mutated SF3B1, BM = healthy bone marrows.

Comparison	targeted DMR	Justification	Point
$JAK2^{V617F}$ ET versus TN MDS/MPN-SF3B1-T	not missing	DMRs related to dysplasia or prolifération induced by "TN" feature	1
$JAK2^{V617F}$ MDS/MPN-SF3B1-T + $JAK2^{V617F}$ ET versus TN MDS/MPN-SF3B1-T	not missing	DMRs related to prolifération induced by "TN" feature	1
MDS-SF3B1 versus $JAK2^{V617F}$ ET	missing	DMRs related to dysplasia or prolifération induced by " $JAK2^{V617F}$ " feature	1
$JAK2^{V617F}$ ET versus BM	missing	DMRs related to prolifération induced by " $JAK2^{V617F}$ " feature	1
MDS-SF3B1 versus BM	missing	DMRs related to dysplasia	1

Targeted methylation analysis

We performed targeted analysis on 32 classical genes involved in myeloid neoplasms, including: *WT1*, *SETBP1*, *ABL1*, *HRAS*, *DNMT3A*, *RUNX1*, *CEBPA*, *KIT*, *EZH2*, *NPM1*, *STAT5B*, *PTPN11*, *PTPN12*, *FLT3*, *IDH1*, *IDH2*, *KRAS*, *CBL*, *BRAF*, *ASXL1*, *ETV6*, *JAK2*, *TET2*, *U2AF1*, *CALR*, *CSF3R*, *MPL*, *NRAS*, *SF3B1*, *SRSF2*, *TP53* and *ZRSR2* (3,5,16).

RESULTS

Forty-eight samples from 38 patients with $JAK2^{V617F}$ MDS/MPN-SF3B1-T (12 samples), TN MDS/MPN-SF3B1-T (12 samples), $JAK2^{V617F}$ ET (10 samples), MDS-SF3B1 (10 samples) BM (4 samples) were included for RRBS analysis (**Figure 1, Table 3**). Features are presented in **Table 4**. On average 64.9% of reads were uniquely mapped, 22.8% of reads were mapped more than on time and 12.3% of reads were not mapped at all. The bisulfite rate conversion was 99%. Cytosine methylation rate and average cytosine methylation according to the position in the read was similar in every groups.

Molecular characterization by NGS

The molecular analysis by routine NGS is in progress.

Table 3. Samples repartition according to diagnosis and cell's source.

BM = bone marrow; MN = mononuclear; PN = polynuclear. $JAK2^{V617F}$ MDS/MPN- $SF3B1$ -T = $JAK2^{V617F}$ mutated myelodysplastic/myeloproliferative neoplasm with $SF3B1$ mutation and thrombocytosis, TN MDS/MPN- $SF3B1$ -T = triple negative myelodysplastic/myeloproliferative neoplasm with $SF3B1$ mutation and thrombocytosis, $JAK2^{V617F}$ ET = $JAK2^{V617F}$ mutated essential thrombocytemia, MDS- $SF3B1$ = myelodysplastic neoplasm with mutated $SF3B1$, BM = healthy bone marrows.

MDS- $SF3B1$	TN MDS/MPN- $SF3B1$ -T	$JAK2^{V617F}$ MDS/MPN- $SF3B1$ -T	$JAK2^{V617F}$ ET	BM
Patient 1 BM MN	Patient 9 blood PN	Patient 18 BM MN	Patient 25 Total BM	Patient 35 Total BM
Patient 1 BM PN	Patient 9 BM MN	Patient 18 BM PN	Patient 26 Total blood	Patient 36 Total BM
Patient 2 BM MN	Patient 9 blood PN	Patient 18 BM MN	Patient 27 Total blood	Patient 37 Total BM
Patient 2 BM PN	Patient 10 blood PN	Patient 19 blood PN	Patient 28 Total blood	Patient 38 Total BM
Patient 3 BM PN	Patient 11 blood PN	Patient 19 blood MN	Patient 29 Total blood	
Patient 4 BM PN	Patient 12 Total BM	Patient 19 BM PN	Patient 30 Total blood	
Patient 5 BM MN	Patient 13 BM MN	Patient 20 Total BM	Patient 31 Total blood	
Patient 6 BM MN	Patient 14 Total blood	Patient 21 BM MN	Patient 32 Total blood	
Patient 7 BM MN	Patient 14 BM MN	Patient 22 blood MN	Patient 33 Total blood	
Patient 8 BM MN	Patient 15 Total blood	Patient 23 Total BM	Patient 34 Total blood	
	Patient 16 blood PN	Patient 23 Total blood		
	Patient 17 blood PN	Patient 24 Total blood		

Replicates from different sources analysis

With the *MethylKit* method, we analyzed, for mononuclear and polynuclear cells, 17,000 genes for which methylation data for CpGs near initiation site of transcription were available. Only 32 genes were significantly differentially methylated between the two groups. The DMRs analysis also did not highlight any significant signature between mononuclear and polynuclear cells. In consequence, when several sample's sources were available for the same patient, we selected for each comparison the most similar source as in the compared group.

Table 4. Samples and patients feature.

$JAK2^{V617F}$ MDS/MPN-*SF3B1*-T = $JAK2^{V617F}$ mutated myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, TN MDS/MPN-*SF3B1*-T = triple negative myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, $JAK2^{V617F}$ ET = $JAK2^{V617F}$ mutated essential thrombocytemia, MDS-*SF3B1* = myelodysplastic neoplasm with mutated *SF3B1*, BM = healthy bone marrows.

	MDS- <i>SF3B1</i>	TN MDS/MPN- <i>SF3B1</i> -T	$JAK2^{V617F}$ MDS/MPN- <i>SF3B1</i> -T	$JAK2^{V617F}$ TE	BM
N (samples/patients)	10/8	12/10	12/7	10/10	4/4
Age at diagnosis (years)					
Median	70,5	76,5	71	53,5	73
Min-Max	62-88	58-92	62-80	23-79	58-81
Gender, n (%)					
Male	5 (63%)	4 (40%)	6 (86%)	3 (30%)	1 (25%)
Female	3 (37%)	6 (60%)	1 (14%)	7 (70%)	3 (75%)
Samples sources, n (%)					
Bone marrow	10(100%)	5 (42%)	7 (38%)	1 (10%)	4 (100%)
Blood	0 (0%)	7 (38%)	5 (42%)	9 (90%)	0 (0%)
Type of cells samples					
MN	6 (60%)	3 (25%)	5 (42%)	0 (0%)	0 (0%)
PN	4 (40%)	6 (50%)	3 (25%)	0 (0%)	0 (0%)
Total	0 (0%)	3 (25%)	4 (33%)	10 (100%)	4 (100%)
Thrombotic events, n (%)					
Yes	2 (25%)	0	1 (14%)	4 (40%)	0
No	6 (75%)	7 (70%)	5 (72%)	6 (60%)	4 (100%)
Missing data	0	3 (30%)	1 (14%)	0	0
Secondary AML, n (%)					
Yes	1 (12%)	1 (10%)	1 (14%)	0	0
No	7 (88%)	6 (60%)	5 (72%)	10 (100%)	4 (100%)
Missing data	0	3 (30%)	1 (14%)	0	0
Secondary myelofibrosis, n (%)					
Yes	0	0	3 (43%)	0	0
No	8 (100%)	7 (70%)	4 (57%)	10 (100%)	4 (100%)
Missing data	0	3 (30%)	0	0	0
Treatments, n (%)					
ASE	6 (75%)	6 (60%)	4 (57%)	0	0
Transfusions	8 (100%)	7 (70%)	4 (57%)	0	0
Lenalidomide	2 (25%)	1 (10%)	1 (14%)	0	0
Luspatercept	0	0	0	0	0
Hydroxyurea	0	4 (40%)	3 (43%)	4 (40%)	0
azacytidine	3	1 (10%)	0	0	0
interferon	0	0	0	4 (40%)	0
Antiagregants	4 (50%)	5 (50%)	2 (29%)	10 (100%)	0
Anticoagulants	2 (25%)	1 (10%)	2 (29%)	0	0
Missing data	0	3 (30%)	2 (29%)	0	0

Overall methylation comparison

There was no significant difference in the global methylation level between the five groups (**Figure 3**).

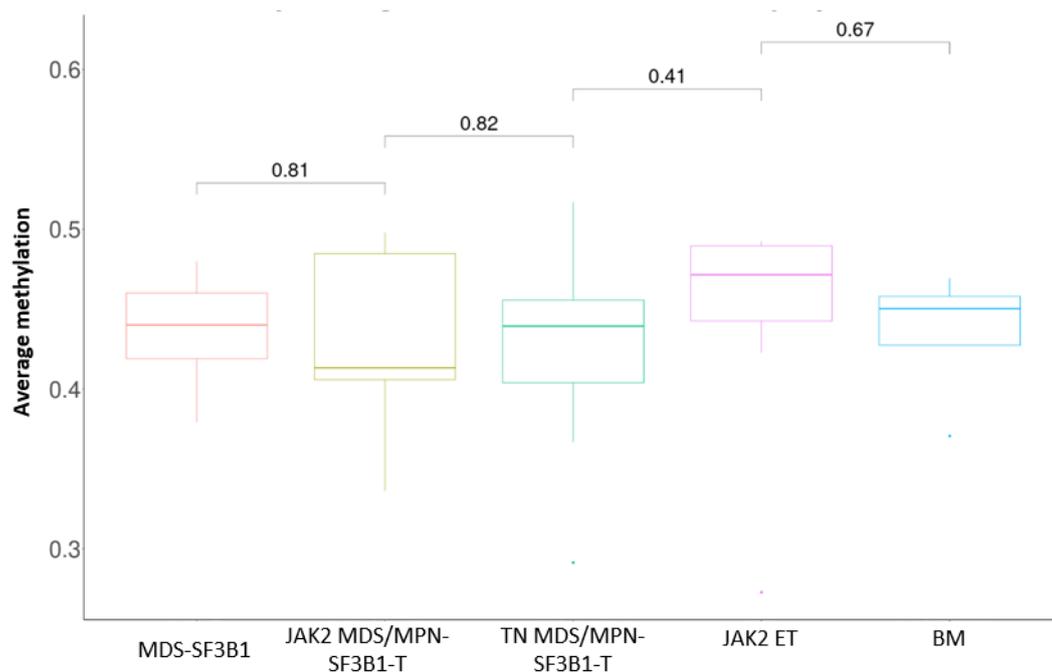


Figure 3. Average global methylation analysis.

JAK2 MDS/MPN-*SF3B1*-T = *JAK2*^{V617F} mutated myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, TN MDS/MPN-*SF3B1*-T = triple negative myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, JAK2 ET = *JAK2*^{V617F} mutated essential thrombocytemia, MDS-*SF3B1* = myelodysplastic neoplasm with mutated *SF3B1*, BM = healthy bone marrows.

CpGs methylation comparison

The MDS/MPN-*SF3B1*-T versus *JAK2*^{V617F} MDS/MPN-*SF3B1*-T comparison identified 81,510 significant CpGs with a methylation difference over 5%. The unsupervised analysis of these 81,510 CpGs by PCA (Principal Component Analysis) did not cluster the samples by their respective groups (**Figure 4a**). In the same way, the unsupervised analysis of the 15,000 most variant CpGs in MDS/MPN-*SF3B1*-T versus *JAK2*^{V617F} MDS/MPN-*SF3B1*-T comparison did not cluster the samples by their respective groups (**Figure 4b**).

After the clinical pertinence score application, the TN MDS/MPN-*SF3B1*-T versus *JAK2*^{V617F} MDS/MPN-*SF3B1*-T comparison identified 959 significant CpGs with a methylation difference over 5% and 113 CpG among the 15,000 most variant CpGs with the maximal score (5/5). The unsupervised analysis of these 959 and 113 CpGs by PCA better clustered the samples by their respective groups and particularly individualized TN MDS/MPN-*SF3B1*-T (**Figure 4c and 4d**).

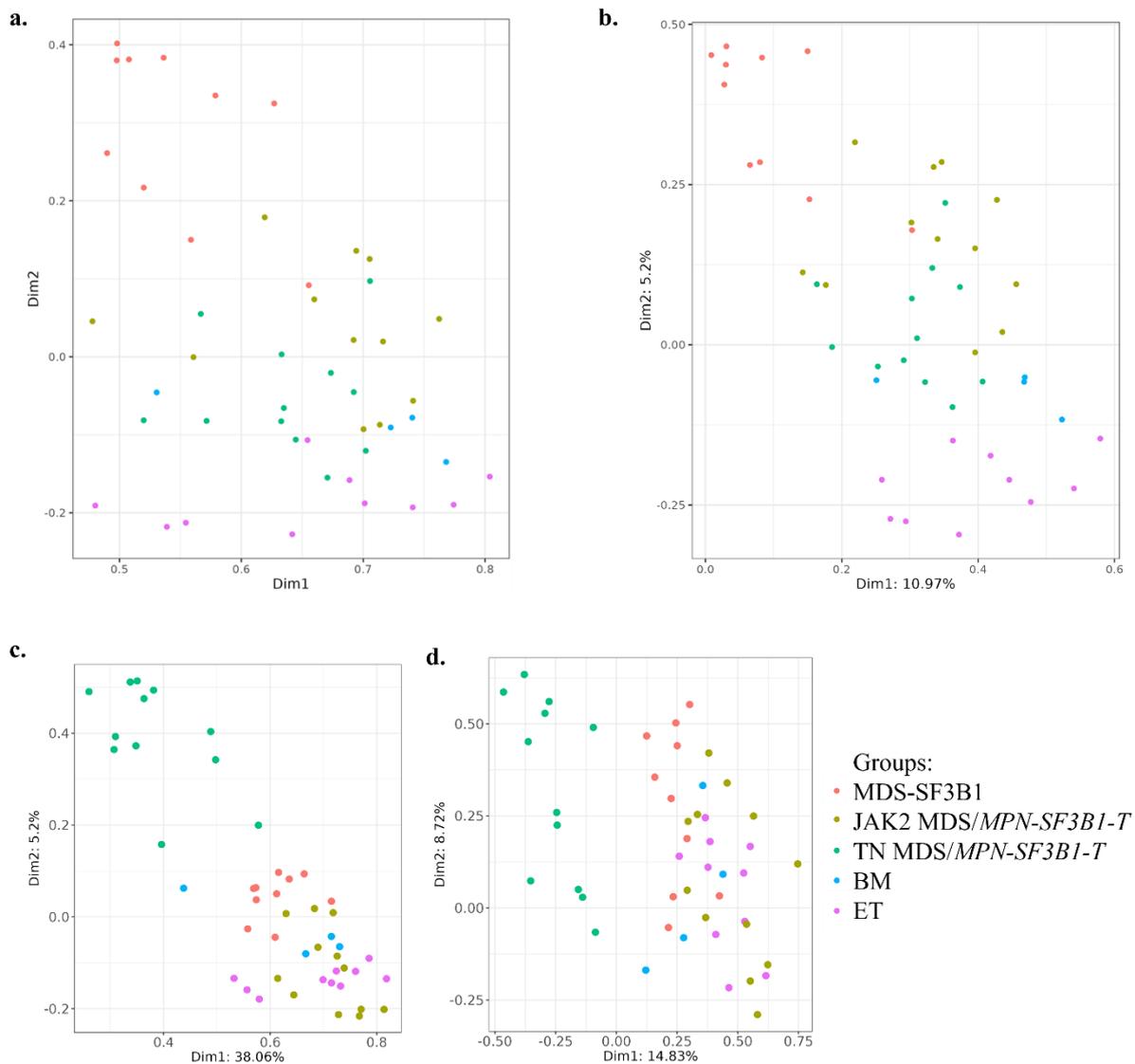


Figure 4. PCA (Principal Component Analysis) unsupervised analysis.

JAK2 MDS/MPN-SF3B1-T = $JAK2^{V617F}$ mutated myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, TN MDS/MPN-SF3B1-T = triple negative myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, JAK2 ET = $JAK2^{V617F}$ mutated essential thrombocytemia, MDS-SF3B1 = myelodysplastic neoplasm with mutated *SF3B1*, BM = healthy bone marrows. **a.** PCA unsupervised analysis of 81 510 CpGs with significant methylation difference over than 5%. **b.** PCA unsupervised analysis of the 15,000 most variant CpGs. **c.** PCA unsupervised analysis of 959 CpGs with significant methylation difference over than 5% and 5/5 clinical pertinence score. **d.** PCA unsupervised analysis of 113 CpGs with 5/5 clinical pertinence score among the 15,000 most variant CpGs.

We identified 6,246 significant differentially methylated CpGs between MDS/MPN-*SF3B1*-T TN and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T (**Figure 5a**): 3, 849 CpGs hypermethylated in TN MDS/MPN-*SF3B1*-T and 2, 297 CpGs hypomethylated in TN MDS/MPN-*SF3B1*-T. The most significant CpGs were associated with the following genes: *G3BP1*, *JAKMIP3*, *SLC6A15*, *CHGA*, *CLMN*, *RAB5IF* and *PPIL2*.

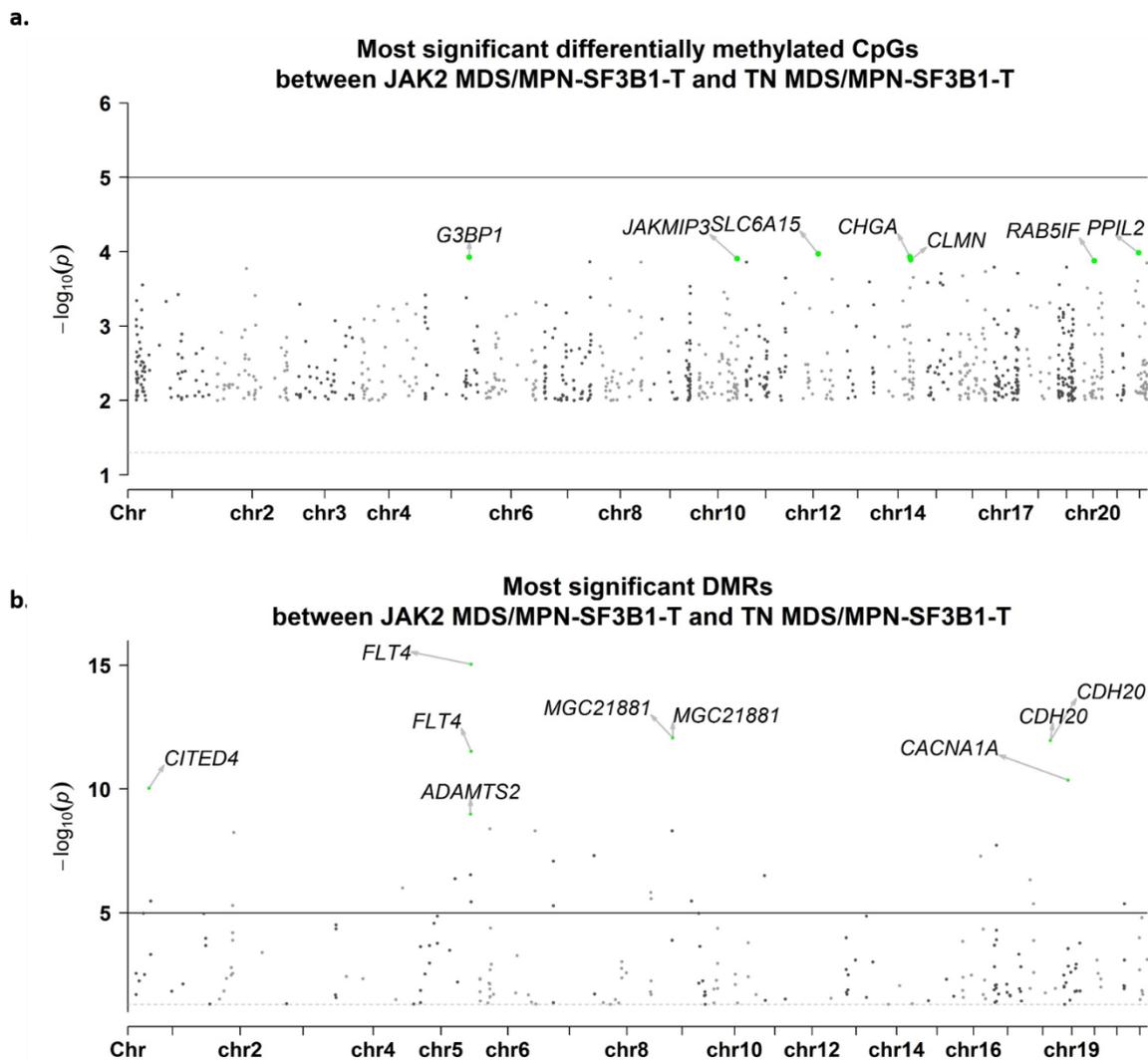


Figure 5. Most significant differentially methylated CpGs and DMRs between TN MDS/MPN-*SF3B1*-T and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T. **a.** Most significant differentially methylated CpG between MDS/MPN-*SF3B1*-T TN and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T. **b.** Most significant DMRs between TN MDS/MPN-*SF3B1*-T and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T.

DMR methylation comparison

We identified 234 significant DMRs between TN MDS/MPN-*SF3B1*-T and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T (**Figure 5b**): 131 DMRs hypermethylated in TN MDS/MPN-*SF3B1*-T and 103 DMRs hypomethylated in TN MDS/MPN-*SF3B1*-T. The most significant DMRs were associated with the following genes: *CITED4*, *FLT4*, *ADAMTS2*, *MGC21881*, *CACNA1A* et *CDH20* (**Figure 5b**).

The application of 20% of methylation difference threshold, background noise elimination and clinical pertinence score identified, among the 234 significant DMRs, 15 DMRs with maximal score and annotated to 11 genes. We also identified 38 DMRs annotated to 28 genes with at least 4/5 maximal pertinence score.

Functional annotations highlighted an enrichment for MAPK signaling pathway genes. We selected 8 relevant candidate genes to be assessed by qPCR: *AATK*, *BLCAP*, *DUSP2*, *SPRY2*, *FLT4*, *IMPA2*, *GPM6A* and *LMX1B*. *AATK*, *BLCAP*, *DUSP2*, *FLT4*, *GPM6A*, *LHX6* et *SPRY2* were hypomethylated in TN MDS/MPN-*SF3B1*-T compared to *JAK2* MDS/MPN-*SF3B1*-T and *LMX1B* and *IMPA2* were hypermethylated (**Table 5**).

Table 5. Methylation and qPCR fold changes of candidate genes.

JAK2^{V617F} MDS/MPN-*SF3B1*-T = *JAK2*^{V617F} mutated myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, TN MDS/MPN-*SF3B1*-T = triple negative myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, *JAK2*^{V617F} ET = *JAK2*^{V617F} mutated essential thrombocytemia, MDS-*SF3B1* = myelodysplastic neoplasm with mutated *SF3B1*, BM = healthy bone marrows.

Gene	Methylation	Expression fold changes		
	TN versus <i>JAK2</i> ^{V617F}	TN MDS/MPN- <i>SF3B1</i> -T versus BM	<i>JAK2</i> ^{V617F} MDS/MPN- <i>SF3B1</i> -T versus BM	MDS- <i>SF3B1</i> versus BM
<i>DUSP2</i>	-24%	18	4	1
<i>SPRY2</i>	-29%	2052	5	0,3
<i>AATK</i>	-31%	25	10	0,03
<i>LMX1B</i>	+23%	272	77	0,18
<i>BLCAP</i>	-31%	32	13	0,2
<i>FLT4</i>	-56%	16	14	<0,01
<i>GPM6A</i>	-32%	7	16	7
<i>IMPA2</i>	+34%	0,06	0,14	< 0,01

qPCR assessment

RNA was available for 4 samples from healthy bone marrows (controls), 1 sample from *JAK2*^{V617F} MDS/MPN-*SF3B1*-T, 1 sample from MDS/MPN-*SF3B1*-T TN and 1 sample from MDS-*SF3B1*. qPCR analysis showed an overexpression of hypomethylated genes *AATK*, *BLCAP*, *DUSP2*, *FLT4* and *SPRY2* in MDS/MPN-*SF3B1*-T TN as compared to *JAK2*^{V617F} MDS/MPN-*SF3B1*-T (**Figure 7, Table 5**). To a lesser extent, these genes are also overexpressed in *JAK2*^{V617F} MDS/MPN-*SF3B1*-T as compared to MDS-*SF3B1* and BM. The hypermethylated gene *LMX1B* was also similarly overexpressed (**Figure 7, Table 5**). Despite its hypomethylation in TN MDS/MPN-*SF3B1*-T, *GPM6A* was underexpressed as compared to *JAK2*^{V617F} MDS/MPN-*SF3B1*-T, but overexpressed as compared to BM (**Figure 7, Table 5**). *IMPA2* was hypermethylated and underexpressed in TN MDS/MPN-*SF3B1*-T as compared to *JAK2*^{V617F} MDS/MPN-*SF3B1*-T. It also was underexpressed in both TN MDS/MPN-*SF3B1*-T and *JAK2*^{V617F} MDS/MPN-*SF3B1*-T as compared to BM.

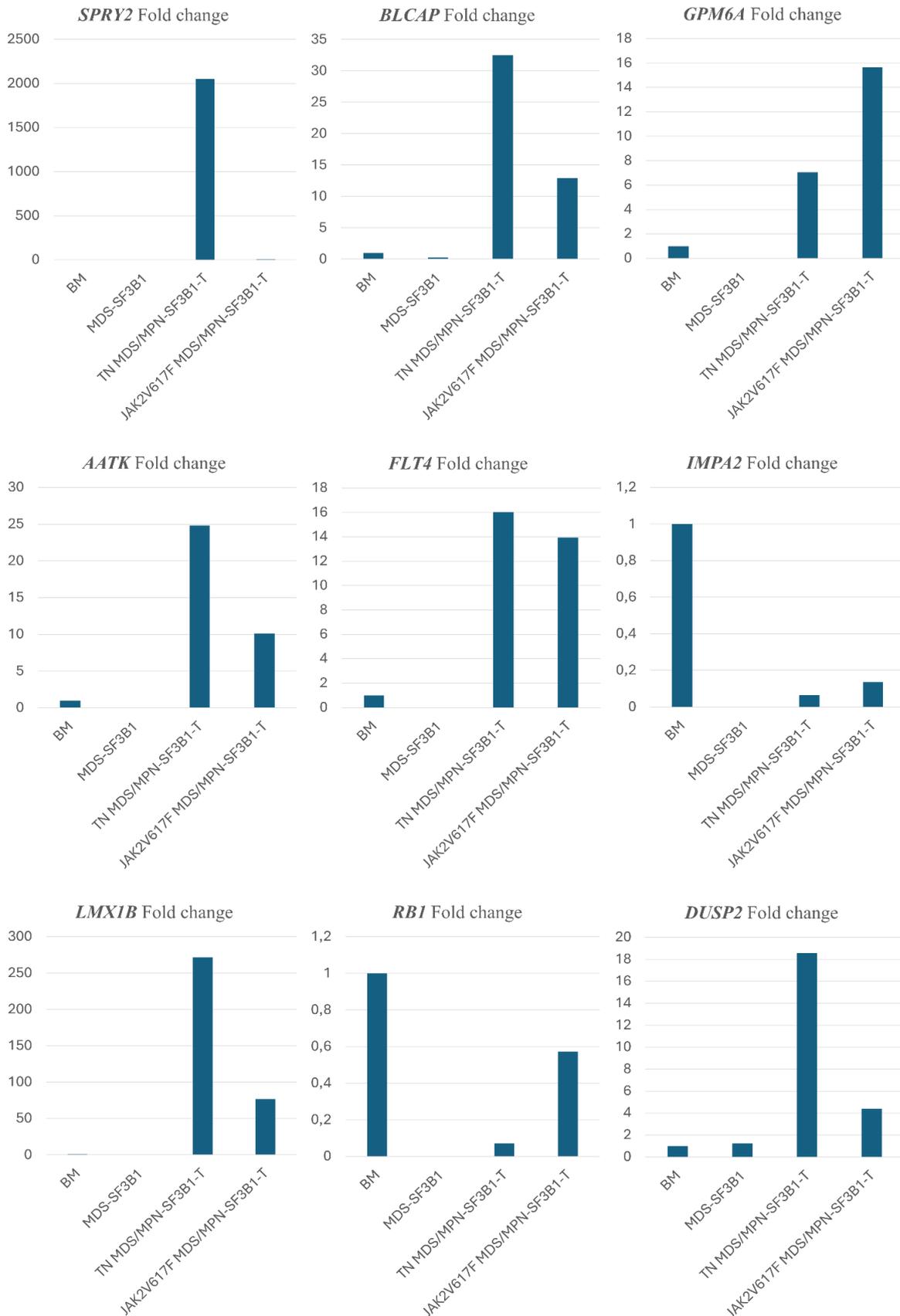


Figure 7. Quantitative polymerase chain reaction for candidate genes RNA.

JAK2 MDS/MPN-*SF3B1*-T = *JAK2*^{V617F} mutated myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, TN MDS/MPN-*SF3B1*-T = triple negative myelodysplastic/myeloproliferative neoplasm with *SF3B1* mutation and thrombocytosis, *JAK2* ET = *JAK2*^{V617F} mutated essential thrombocytemia, MDS-*SF3B1* = myelodysplastic neoplasm with mutated *SF3B1*, BM = healthy bone marrows.

DISCUSSION

Triple negative feature of TN MDS/MPN-SF3B1-T leads to diagnosis difficulties and platelets proliferation pathophysiology misunderstanding. The treatment relies on anemia and thrombosis prevention. Sometimes it requires a cytoreduction, which may be difficult to manage due to anemia aggravation. A better understanding of proliferative mechanisms would help to build more targeted and less toxic new cytoreductive therapeutical strategies.

Here we conducted comparative RRBS in TN MDS/MPN-SF3B1-T, JAK2V617F MDS/MPN-SF3B1-T, JAK2V617F ET, MDS-SF3B1 and healthy bone marrows. We highlighted MAPK implication in triple negative related thrombocytosis MDS/MPN-SF3B1-T and developed a new approach for DMRs selection in RRBS analysis.

It has been demonstrated that cellular proliferation in non-Philadelphia MPN is correlated to a constitutive activation of JAK-STAT pathway and also a cytokines enriched inflammatory environment, with interleukin 3, TPO or GM-CSF especially (18). JAK-STAT pathway is also connected to MAPK pathway via PTPN11 and GRB2 (19). *PTPN11* is frequently mutated in JMML (juvenile myelomonocytic leukemia). In contrast, MAPK pathway is feedbacked in *JAK2V617F* cells with MAPK and TP53 negative regulators overexpression as *DUSP1* (18). This overexpression was not found in *MPL^{W515L}* cells, but it gave a likely justification for JAK2 inhibitors therapeutic escape in *JAK2^{V617F}* cells. Apoptotic threshold rising in HSC leads indeed to their survival under treatment and to the uncomplete destruction of the clone. MPN and MAPK pathway are then closely linked and MPN are impacted by both positive and negative regulation of MAPK pathway.

Our new approach for DMRs detection un RRBS analysis highlighted a MAPK related significant enrichment in TN MDS/MPN-SF3B1-T versus *JAK2^{V617F}* MDS/MPN-SF3B1-T comparison. We also identified 4 MAPK negative regulator genes overexpression: *LMTK1*, *BLCAP*, *DUSP2* (*DUSP1* paralog) and *SPRY2*. *LMTK1* is a tyrosine kinase involved in growth and proliferation of myeloid progenitors (20). It inhibits Akt and ERK phosphorylation and enhances MAPK p38 and JUN kinases. Overexpression of JUN proteins has been documented in ET (21). *BLCAP* and *SPRY2* are tumor suppressor genes (22,23). *SPRY2* is a direct negative regulator of MAPK pathway. It has been described as negatively regulated by hypermethylation in myeloid pathologies as chronic myelomonocytic leukemia. This protein also inhibits GRB2 and SOS in their JAK-STAT and MAPK interconnections.

Compared to controls, this consistent set of clear overexpression (expression fold changes ranged from 16 to 2052) may be the reflection of a reactional consequence to constitutive activation of proliferation pathway in the tumoral cell. It strengthens the hypothesis that MDS/MPN-SF3B1-T - and more widely MPN - are beyond JAK-STAT pathway, a MPAK and proliferation pathways disease. The absence of JAK-STAT pathway mutation in TN MPN would then not impact a similar phenotype emergence and might guide to outside JAK-STAT common therapeutic possibilities.

We also highlighted the overexpression of hypomethylated MAPK pathway positive regulator kinase coding gene *FLT4* (24) and hypermethylated *LMX1B* (17). *LMX1B* is an oncogenic transcriptional factor from LIM-homeodomains family. Its overexpression by hypermethylation is associated with proliferation, migration and radio-resistance in various cancers (25). As described in oncology in literature (26), *IMPA2*, an ERK from MAPK pathway inhibitor coding gene, was hypermethylated and underexpressed (expression fold changes at 0,06 and 0,45 as compared to controls and MDS/MPN-*SF3B1*-T *JAK2*^{V617F} respectively). Recent works has demonstrated the inhibitory function of *GPM6A* on MAPK pathway, particularly for lung cancer in which it is underexpressed by hypermethylation, suggesting *GPM6A* as a new therapeutic target (27). *GPM6A* was here hypomethylated and underexpressed in MDS/MPN-*SF3B1*-T as compared to controls and particularly in TN MDS/MPN-*SF3B1*-T. The expression fold changes of *FLT4*, *GPMA6* and *IMPA2* would then be directly causal in TN MDS/MPN-*SF3B1*-T phenotype. Their epigenetic regulation alterations also might be a consequence of methylation alterations attributable to the dysplastic feature of the disease.

Interestingly, except for *GPM6A*, the overexpression is also observed to a lesser extent in the sample of *JAK2*^{V617F} MDS/MPN-*SF3B1*-T (expression fold changes ranged from 4 to 77). The same genes were rather underexpressed in the MDS-*SF3B1* sample (expression fold changes ranged from 0.03 to 1) compared to controls. This prioritization was also observed in underexpression of *IMPA2*, less expressed in TN MDS/MPN-*SF3B1*-T than in *JAK2*^{V617F} MDS/MPN-*SF3B1*-T. In one hand, it consolidates the hypothesis that the alterations are reactional the four negative regulator genes but also rises the likely assumption that *LMX1B*, *FLT4* and *IMPA2* are only partially implicated in TN MDS/MPN-*SF3B1*-T phenotype.

Methylation data analysis from RRBS requires identification tools as well as *MethylKit*, *Metilene*, *BiSeq* or *DMRfinder* (11,28–31). The *MethylKit* methodology relies on arbitrary creation of regions by a sliding window along the genome. The comparison of these regions between two groups leads to DMRs identification. This strategy is frequently used for RRBS but there is no gold standard for this kind of analysis. According to *MethylKit*, the methylation rate calculation of each DMR is based on the percentage of unconverted cytosines during bisulfite treatment on the entire region. The position information of each CpG inside the region and the differentially methylated rate among the CpGs from a same region are then lost.

We then provided here a new methodology inspired by *MethylKit*, but conserving this information in our comparisons. We compared vectors composed of methylation ratio from each CpG of the region instead of global means.

Data from next generation sequencing are particularly numerous. To select the most relevant results, bioinformatic analysis requires a pipeline application containing filtering stages. Functional annotations are from a great importance but are limited by established knowledge and the capacity to deeply and manually explore many of these candidates. To select only relevant candidates to answer studying issue, we created a clinical pertinence score. It translates the phenotypical coherence of a candidate DMR on a numeric scale and better utilized multiple comparisons performed with overlapping diseases and controls. This strategy helps to eliminate significant DMRs reflecting a difference between two groups but not answering to the studying issue.

However, this strategy has limits. Our method may potentially eliminate relevant DMRs not identified as such. Moreover, the rareness of MDS/MPN-*SF3B1*-T leads to small numbers of samples and a non-exhaustive analysis. We do not control all variables impacting methylation. In particular, we adapted samples to their tissular or cellular origins, but these differences could not be entirely removed.

RRBS is less used for methylation observation. To date, whole genome methylation analysis or methylation chips have a better coverage and tend to discuss our choice. Recent works in bioinformatics have shown a better biological relevance of “*e-values*” to detect DMRs from RRBS using classic tools such as *MethylKit* (31). Despite their large use for statistical significance estimation, *p-values* and corrected *p-values* are increasingly discussed and controverted. *P-values* evaluate the probability to reject a null hypothesis after a statistical test. The lower it is, the more the hypothesis can be rejected. In 2023, Yang et al proposed in RRBS DMRs detection the use of *e-value*, which corresponds to a probability ratio to accept the null hypothesis including the parameter space (31). So it might be interesting to reanalyze our data and to associate *e-values* with the conservation of positional information and application of the clinical pertinence score.

Despite these limits, the obtained results with our new approach in RRBS analysis are in coherence with the actual knowledge and the phenotype of TN MDS/MPN-*SF3B1-T*. We do not identify a precise epigenetic mechanism for platelets proliferation in TN MDS/MPN-*SF3B1-T*, but these results strengthen the idea that the strategy we provided might be applicable to other RRBS analyses.

In conclusion, TN MDS/MPN-*SF3B1-T* is a complex disease. Its phenotype is multifactorial and involve genomic mutations affecting splice and epigenomic alterations. The coexistence of mutation affecting methylation might dysregulate proliferation pathways by methylation of pro-oncogene promoters.

BIBLIOGRAPHIE

1. Asaulenko ZP, Spiridonov IN, Baram DV, Krivolapov YA. [WHO Classification of Tumors of Hematopoietic and Lymphoid Tissues, 2022 (5th edition): Myeloid and Histiocytic Tumors]. *Arkh Patol.* 2023;85(5):36-44.
2. Broseus J, Florensa L, Zipperer E, Schnittger S, Malcovati L, Richebourg S, et al. Clinical features and course of refractory anemia with ring sideroblasts associated with marked thrombocytosis. *Haematologica.* juill 2012;97(7):1036-41.
3. Jeromin S, Haferlach T, Weissmann S, Meggendorfer M, Eder C, Nadarajah N, et al. Refractory anemia with ring sideroblasts and marked thrombocytosis cases harbor mutations in *SF3B1* or other spliceosome genes accompanied by *JAK2V617F* and *ASXL1* mutations. *Haematologica.* avr 2015;100(4):e125-127.
4. Szpurka H, Tiu R, Murugesan G, Aboudola S, Hsi ED, Theil KS, et al. Refractory anemia with ringed sideroblasts associated with marked thrombocytosis (RARS-T), another myeloproliferative condition characterized by *JAK2 V617F* mutation. *Blood.* 1 oct 2006;108(7):2173-81.
5. Klampfl T, Gisslinger H, Harutyunyan AS, Nivarthi H, Rumi E, Milosevic JD, et al. Somatic mutations of calreticulin in myeloproliferative neoplasms. *N Engl J Med.* 19 déc 2013;369(25):2379-90.
6. Nangalia J, Massie CE, Baxter EJ, Nice FL, Gundem G, Wedge DC, et al. Somatic *CALR* mutations in myeloproliferative neoplasms with nonmutated *JAK2*. *N Engl J Med.* 19 déc 2013;369(25):2391-405.
7. Patnaik MM, Belachew A, Finke C, Lasho TL, Hanson CA, Tefferi A. *CALR* mutations are infrequent in WHO-defined refractory anemia with ring sideroblasts. *Leukemia.* juin 2014;28(6):1370-1.
8. Greenfield G, McMullin MF. Epigenetics in myeloproliferative neoplasms. *Front Oncol.* 2023;13:1206965.

9. Hultcrantz M, Björkholm M, Dickman PW, Landgren O, Derolf ÅR, Kristinsson SY, et al. Risk for Arterial and Venous Thrombosis in Patients With Myeloproliferative Neoplasms: A Population-Based Cohort Study. *Ann Intern Med.* 6 mars 2018;168(5):317-25.
10. Barbui T, Tefferi A, Vannucchi AM, Passamonti F, Silver RT, Hoffman R, et al. Philadelphia chromosome-negative classical myeloproliferative neoplasms: revised management recommendations from European LeukemiaNet. *Leukemia.* mai 2018;32(5):1057-69.
11. Akalin A, Kormaksson M, Li S, Garrett-Bakelman FE, Figueroa ME, Melnick A, et al. methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* 3 oct 2012;13(10):R87.
12. Gu H, Bock C, Mikkelsen TS, Jäger N, Smith ZD, Tomazou E, et al. Genome-scale DNA methylation mapping of clinical samples at single-nucleotide resolution. *Nat Methods.* févr 2010;7(2):133-6.
13. Krueger F, Andrews SR. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinforma Oxf Engl.* 1 juin 2011;27(11):1571-2.
14. Braschi B, Denny P, Gray K, Jones T, Seal R, Tweedie S, et al. Genenames.org: the HGNC and VGNC resources in 2019. *Nucleic Acids Res.* 8 janv 2019;47(D1):D786-92.
15. Kuhn RM, Karolchik D, Zweig AS, Wang T, Smith KE, Rosenbloom KR, et al. The UCSC Genome Browser Database: update 2009. *Nucleic Acids Res.* janv 2009;37(Database issue):D755-761.
16. Kennedy JA, Ebert BL. Clinical Implications of Genetic Mutations in Myelodysplastic Syndrome. *J Clin Oncol Off J Am Soc Clin Oncol.* 20 mars 2017;35(9):968-74.
17. Fan L, Zhang A, Deng P. LMX1B mRNA expression and its gene body CpG methylation are valuable prognostic biomarkers for laryngeal squamous cell carcinoma. *Biomed Pharmacother Biomedecine Pharmacother.* sept 2019;117:109174.
18. Kesarwani M, Kincaid Z, Azhar M, Menke J, Schwieterman J, Ansari S, et al. MAPK-negative feedback regulation confers dependence to JAK2V617F signaling. *Leukemia.* août 2023;37(8):1686-97.
19. Ibrahim SSA, Huttunen KM. Orchestrated modulation of rheumatoid arthritis via crosstalking intracellular signaling pathways. *Inflammopharmacology.* août 2021;29(4):965-74.
20. Sun P, Sun X, Zhao W, Ren M, Zhang C, Wang Z, et al. Lemur Tyrosine Kinase-3 Suppresses Growth of Prostate Cancer Via the AKT and MAPK Signaling Pathways. *Cell Physiol Biochem Int J Exp Cell Physiol Biochem Pharmacol.* 2017;42(6):2582-92.
21. Müller CI, Luong QT, Shih LY, Jones LC, Desmond JC, Kawamata N, et al. Identification of marker genes including RUNX3 (AML2) that discriminate between different myeloproliferative neoplasms and normal individuals. *Leukemia.* sept 2008;22(9):1773-8.
22. Han F, Hu M, Zhang L, Fan X, Wang J, Lou Z, et al. A-to-I RNA editing of BLCAP promotes cell proliferation by losing the inhibitory of Rb1 in colorectal cancer. *Exp Cell Res.* 15 août 2022;417(2):113209.
23. Kunitomo H, Meydan C, Nazir A, Whitfield J, Shank K, Rapaport F, et al. Cooperative Epigenetic Remodeling by TET2 Loss and NRAS Mutation Drives Myeloid Transformation and MEK Inhibitor Sensitivity. *Cancer Cell.* 8 janv 2018;33(1):44-59.e8.
24. Liang J, Lei P, Su X, Gao J, Ren B, Zhang Y, et al. Dauricine Inhibits Non-small Cell Lung Cancer Development by Regulating PTEN/AKT/mTOR and Ras/MEK1/2/ERK1/2 Pathways in a FLT4-dependent Manner. *Curr Cancer Drug Targets.* 2024;24(11):1157-68.
25. Zang C, Zhao F, Pu Y. LMX1B involved in the radioresistance, proliferation and migration of esophageal cancer cells. *Biomed Pharmacother Biomedecine Pharmacother.* oct 2019;118:109358.
26. Zhang K, Liu L, Wang M, Yang M, Li X, Xia X, et al. A novel function of IMPA2, plays a tumor-promoting role in cervical cancer. *Cell Death Dis.* 14 mai 2020;11(5):371.

27. Zhang Q, Deng S, Li Q, Wang G, Guo Z, Zhu D. Glycoprotein M6A Suppresses Lung Adenocarcinoma Progression via Inhibition of the PI3K/AKT Pathway. *J Oncol.* 2022;2022:4601501.
28. Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S. metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res.* févr 2016;26(2):256-62.
29. Hebestreit K, Dugas M, Klein HU. Detection of significantly differentially methylated regions in targeted bisulfite sequencing data. *Bioinforma Oxf Engl.* 1 juill 2013;29(13):1647-53.
30. Liu Y, Han Y, Zhou L, Pan X, Sun X, Liu Y, et al. A comprehensive evaluation of computational tools to identify differential methylation regions using RRBS data. *Genomics.* nov 2020;112(6):4567-76.
31. Yang Y, Liu H, Liu Y, Zhou L, Zheng X, Yue R, et al. E-value: a superior alternative to P-value and its adjustments in DNA methylation studies. *Brief Bioinform.* 20 juill 2023;24(4):bbad241.

Annexe 6. Publication – Divoux et al. Efficacy of lenalidomide in MDS/MPN-SF3B1-T and an extreme platelet count.

Received: 22 January 2020 | Revised: 24 April 2020 | Accepted: 10 May 2020

DOI: 10.1002/ccr3.3026

CASE REPORT

Clinical Case Reports Open Access WILEY

Efficacy of lenalidomide in myelodysplastic/myeloproliferative neoplasms with ring sideroblasts and an extreme platelet count

Marion Divoux¹  | Alexia Plocque² | Margaux Sevin³ | Laurent Voillat⁴ |
Pierre Feugier¹ | Agnès Guerci-Bresler¹ | Francois Girodon^{2,3}  | Julien Broséus⁵ 

¹Université de Lorraine, CHRU-Nancy, Service d'Hématologie Clinique, Pôle Spécialités Médicales, France, Nancy, France

²Haematology Laboratory, University Hospital, Dijon, France

³Inserm U1231, University of Bourgogne Franche-Comté, Dijon, France

⁴Haemato-Oncology Department, Hospital of Chalon-sur-Saône, Chalon-sur-Saône, France

⁵Université de Lorraine, CHRU-Nancy, Service d'Hématologie Biologique, Pôle Laboratoires, France, Nancy, France

Correspondence

Julien Broséus, Haematology Laboratory, University Hospital of Nancy, Rue du Morvan, 5400 Vandœuvre-lès-Nancy, France.
Email: Julien.broseus@univ-lorraine.fr

François Girodon, Haematology Laboratory, University Hospital of Dijon, Biology Facility, 2 rue Angélique Ducoudray Dijon France.
Email: francois.girodon@chu-dijon.fr

Abstract

Lenalidomide is efficient in reducing red blood cell transfusion dependency and markedly lowering platelet counts in MDS/MPN-RS-T in the context of major platelet counts.

KEYWORDS

lenalidomide, myelodysplastic syndrome, myelodysplastic/myeloproliferative neoplasms with ring sideroblasts and thrombosis, myeloproliferative syndrome, thrombocytosis

1 | INTRODUCTION

Myelodysplastic/myeloproliferative neoplasms (MDS/MPN) with ring sideroblasts and thrombosis (MDS/MPN-RS-T), previously known as refractory anemia with ring sideroblasts and thrombocytosis (RARS-T), are rare overlapping syndromes associating the dysplastic features of myelodysplastic syndromes with ring sideroblasts (MDS-RS, previously

known as refractory anemia with ring sideroblasts) and the myeloproliferative features of essential thrombocythemia (ET).¹ MDS/MPN-RS-T present with clinical, biological, and prognostic features that differ from those of MDS-RS and ET.² Moreover, MDS/MPN-RS-T are characterized by a particular mutational pattern associating: (a) genomic abnormalities responsible for the myeloproliferative part, such as *JAK2*^{V617F} mutations (40%-50% of cases) or less frequently

Francois Girodon and Julien Broséus participated equally to this work.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Clinical Case Reports* published by John Wiley & Sons Ltd

mutations in exon 10 of *MPL* (myeloproliferative leukemia) or in exon 9 of *CALR* (calreticulin) and (b) a high rate of splicing factor 3B subunit 1 (*SF3B1*) mutations, responsible for the myelodysplastic component of the disease.²⁻⁶ Thus, MDS/MPN-RS-T is now considered as an independent entity.¹

The risk of thrombosis is higher in MDS/MPN-RS-T than in MDS-RS patients without a high platelet count and this thrombotic risk often leads clinicians to use cytoreductive agents known to reduce the platelet count in almost 33% of cases.² However, the use of cytoreductive agents is frequently interrupted, due to the worsening of cytopenia, especially anemia. The management and treatment of this disease are currently based on thrombosis risk stratification. When platelet count is $<1000 \times 10^9/L$, Patnaik and Tefferi⁷ recommend stratifying patients according to two thrombosis risk factors: age > 60 years and prior arterial or venous thrombosis. Patients with no risk factors should be treated with low-dose aspirin or observation alone in *JAK2*^{V617F}-negative diseases with absence of thrombotic risk factors. Patients with 1 or 2 risk factors should be treated with low-dose aspirin. Cytoreductive therapy with hydroxyurea is only recommended in case of high thrombotic risk. Lenalidomide is an immunomodulatory agent frequently used in low-risk myelodysplastic syndromes.⁸ In this context, it is considered as third-line treatment, in case of hydroxyurea failure associated with anemia. In case of a high platelet count $>1000 \times 10^9/L$, aspirin may exacerbate bleeding, cytoreductive therapy often worsens anemia and lenalidomide use is not suggested. Lenalidomide has been tested in published MDS/MPN-RS-T cases, with conflicting results.⁹⁻¹⁵ Here, we report our experience using lenalidomide on two patients with *JAK2*^{V617F}-negative MDS/MPN-RS-T, one of them presenting with a major thrombocytosis.

2 | CASES HISTORY

The first patient was a 78-year-old woman (Patient 1, Table 1), with a complete blood count (CBC) showing hemoglobin level at 85 g/L, persistent thrombocytosis (platelet count: $743 \times 10^9/L$) and a leukocyte count at $6.4 \times 10^9/L$. The bone marrow aspirates revealed erythroid hyperplasia with myelodysplastic features and 64% ring sideroblasts, no blast cells, associated with atypical megakaryocytes, leading to the diagnosis of MDS/MPN-RS-T, according to the revised 2016 world health organization classification's criteria.¹ Bone marrow cytogenetics showed a normal 46XX karyotype. A Lys700Glu *SF3B1* mutation was noted, without *JAK2*^{V617F}, *MPL* exon 10 or *CALR* mutations. Blood transfusions were performed for one year but, in order to avoid relying on red blood cell transfusions, a treatment with lenalidomide (5 mg daily, 21/28 days) was started

TABLE 1 Biological characteristics at diagnosis of the two MDS/MPN-RS-T patients treated with lenalidomide

	Patient 1	Patient 2
Age (y)	78	58
Sex	F	F
Hb level (g/L)	85	114
MCV (fL)	92	97
Platelet count ($10^9/L$)	743	710
Leukocytes ($10^9/L$)	6.4	7.4
Ring sideroblasts (%)	64	24
Erythroid dysplasia	Yes	Yes
Megakaryocytic dysplasia	Marked	Marked
Excess of blasts	No	No
Karyotype	Normal	Normal
<i>SF3B1</i>	Mutated	N/A
<i>JAK2</i> ^{V617F}	No	No
<i>MPL</i> ^{WS15K/L}	No	N/A
<i>CALR</i>	Unmutated	N/A
First treatment	RBC transfusions	Watch and wait
Evolution after 1st treatment	Increased transfusion dependency	Marked increase in the platelet count: $2000 \times 10^9/L$
Second treatment	Lenalidomide 5 mg daily 21 d/28	Hydroxyurea 500 then 1000 mg/d
Evolution after 2nd treatment	Normal platelet count Hb increase to 100 g/L Transfusion frequency reduced	Adverse effects on hemoglobin levels. Irregular elevated platelet counts ranging from $1700 \times 10^9/L$ to $3622 \times 10^9/L$. The decision to stop hydroxyurea and to start lenalidomide was made after a new increase in platelet count at $3106 \times 10^9/L$. EPO was maintained once a week.
Third treatment	Not applicable	Lenalidomide 5 then 10 mg daily 21 d/28
Evolution after 3rd treatment	Not applicable	Major decrease in platelet count Subnormal Hb levels: 118 g/L EPO every 2 wk

Abbreviations: EPO, erythropoietin; F, female; Hb, hemoglobin; M, male; MCV, mean corpuscular volume; N, normal; N/A, nonavailable; RBC, red blood cells; y, years.

resulting in the decrease of platelet count from $686 \times 10^9/L$ (start of treatment) to $150 \times 10^9/L$, associated with an improvement in hemoglobin levels from 80 to 100 g/L over

TABLE 2 Summary of data of the literature regarding MDS/MPN-RS-T case reports

Patient	Age (y)	Gender	Hb (g/L)	MCV (fL)	Platelet count (10 ⁹ /L)	Leukocyte count (10 ⁹ /L)	Ring sideroblasts (%)	Erythroid dysplasia	Megakaryocytic dysplasia	Karyotype	JAK2 ^{V617F} (qPCR)	MPL ^{W515K/L}
1	81	F	79	108	1677	10	86	Yes	Atypical megakaryocytes	Normal	Yes (qPCR) 19% load	N/A
2	60	M	69	88	1592	12.7	98	Yes	Marked Hyperlobulated nuclei	Normal	Yes (qPCR) 32,3% load	N/A
3	47	F	112	N/A	700	N/A	25-45	Yes		Normal	Yes	N/A
4	84	F	77	N	1515	N/A	90	Yes	Numerous atypical megakaryocytes Hypolobulated nuclei	5q- (1 mitosis) Not confirmed by FISH	Yes 22% load	N/A
5	39	F	82	122	1024	5.66	44	Yes	Highly atypical megakaryocytes Hyperlobulated nuclei	t(2;3)(p23;q29)	No (RT-qPCR)	No
6	58	M	98	N/A	1163	N/A	30	Yes	Large hyperlobulated nuclei	Normal FISH neg.	Yes (AS-PCR)	N/A
7	68	F	61	N/A	1257	N/A	Positive	Yes	Megakaryocytic hyperplasia	Normal	No	No
8	49	M	107	93	935	9.2	65	Yes	Yes	Normal	Yes, allele burden 74%	No
9	73	M	67	93.8	669	7.7	25	Yes	Yes	Normal	Yes, allele burden 40%	No
10	85	F	68	88	1203	3	45	N/A	N/A	Normal	No	No
11	78	F	85	92	743	6.4	64	Yes	Marked	Normal	No	No
12	58	F	114	97	710	7.4	24	Yes	Marked	Normal	No	N/A

Abbreviations: AS-PCR, allele-specific polymerase chain reaction; BM, bone marrow; EPO, erythropoietin; F, female; Hb, hemoglobin; M, male; MCV, mean corpuscular volume; N, normal; N/A, non-available; Plt, platelets; RBC, red blood cells; SCT, stem cell transplantation; y, years.

TABLE 2 Continued

CALR	SF3BI	First treatment	Evolution after first treatment	Second treatment	Evolution after second treatment	Third treatment	Evolution after third treatment	Citation
N/A	N/A	EPO	EPO only temporarily successful After 2 y: chronic pulmonary embolism	Hydroxyurea 500 mg 3x/d 2 wk	Reduced platelet count Transfusion dependency	Lenalidomide 5 mg daily	Platelets: $100 \times 10^9/L$ Transfusion independent Hb almost normalized <i>JAK2V617F</i> burden 0.8%	⁹
N/A	N/A	Pyridoxine Anabolic steroids	Transfusion need	Lenalidomide 10 mg daily	3 RBC in 6 mo Plt: $680 \times 10^9/L$ <i>JAK2V617F</i> burden unchanged	N/A	N/A	⁹
N/A	N/A	Hydroxyurea	Minor reduction of spleen size Worsening of anemia, RBC transfusion	Lenalidomide 10 mg daily 21 d/28	Pancytopenia, Increased transfusion requirement Grade 3 BM fibrosis	Allogeneic SCT	Graft lost Transfusion dependency Clonal evolution Death due to sepsis	¹⁰
N/A	Unmutated	Transfusion Lenalidomide 10 mg daily	Platelets: $281 \times 10^9/L$ Transfusion independency BM normalization <i>JAK2V617F</i> burden <2%	N/A	N/A	N/A	N/A	¹¹
N/A	N/A	Hydroxyurea 1000 mg/d α -interferon 3 M units 2x/wk	Reduced platelet count but worse anemia Transfusion dependence (4 units RBC/4 wk)	Hydroxyurea 500 mg 1x/d Pyridoxine Steroids EPO	Failure	Lenalidomide 5 mg daily 7 mo	Platelet drop $363 \times 10^9/L$ Hemoglobin: 90 g/L Transfusion independent Normal BM	¹²
N/A	N/A	Hydroxyurea 500 mg daily	Mild decrease of hemoglobin 83 g/L without efficacy	+ steroids EPO 40 000 units weekly	Inefficacy Transfusion dependency	Lenalidomide 10 mg daily 21 d/28	Transfusion independent; Hb > 9 g/dL Plt < $600 \times 10^9/L$	¹³
Mutated	Mutated	Transfusion Iron supplementation	Symptoms improvement Thrombocytosis persistence Transfusion dependency	Lenalidomide 10 mg daily	Platelets: $497 \times 10^9/L$ Lenalidomide stopped because of severe nausea and anorexia. Platelets: $856 \times 10^9/L$	Lenalidomide 5 mg daily	No toxicity Platelets: $351 \times 10^9/L$ Hemoglobin: 133 g/L No adverse events	¹⁴
No	Yes, allele burden 46%	Lenalidomide 10 mg daily	Decrease of platelet count to $585 \times 10^9/L$. Stop lenalidomide after 8 mo due to loss of response	N/A	N/A	N/A	N/A	¹⁵
No	Yes, allele burden 25%	EPO	Transfusion dependency	Lenalidomide 10 mg daily	Transfusion independency Platelet count lowered to $470 \times 10^9/L$ Stop after 17 mo due to loss of response	N/A	N/A	¹⁵
No	Yes, allele burden 44%	EPO, hydroxyurea, anagrelide	Transfusion dependency Suboptimal response	Lenalidomide 5 mg daily	Platelet count decrease: $558 \times 10^9/L$ Transfusion dependence	N/A	N/A	¹⁵
Unmutated	Mutated	Transfusion 1 y	Increased transfusion dependency	Lenalidomide 5 mg daily 21 d/28	Platelets: $150 \times 10^9/L$ Hb: 80-100 g/L – RBC requirement drastically reduced	N/A	N/A	Current work
N/A	N/A	Watch and wait	Marked increase in the platelet count: $2000 \times 10^9/L$	Hydroxyurea 500 mg/d	Platelet count $3106 \times 10^9/L$ Hemoglobin: 84 g/L EPO 1x/wk Stop hydroxyurea	Lenalidomide 5 then 10 mg daily 21 d/28	Platelets: $760 \times 10^9/L$ Hemoglobin: 118 g/L EPO maintained every 2 wk	Current work

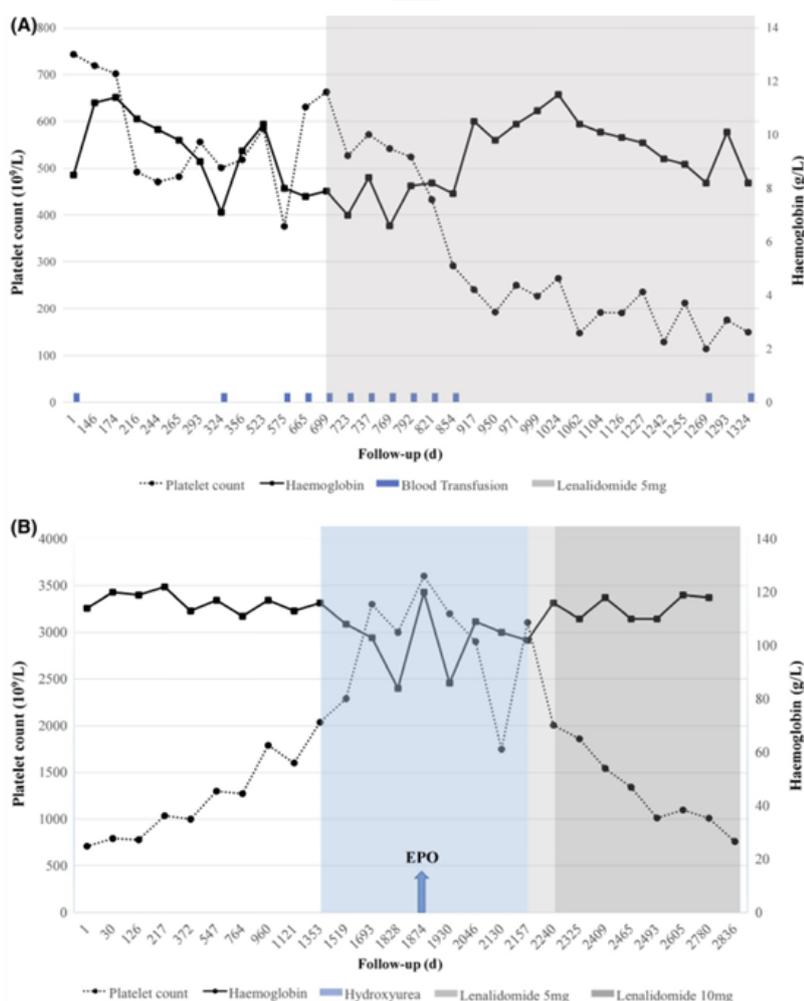


FIGURE 1 Course of blood counts for the two patients treated with lenalidomide: A, Patient 1; B, Patient 2. The left y-axis represents the platelet counts ($10^9/L$), and the right y-axis represents hemoglobin level (g/L). The x-axis represents the follow-up (in days)

the first 28 weeks of lenalidomide treatment. Along these 28 weeks after the beginning of lenalidomide, only 6 red blood cells (RBC) units were transfused (Figure 1A), and no transfusion was required in the subsequent 47 weeks since the hemoglobin level was above 90 g/L. However, due to the subsequent decrease in the hemoglobin level, RBC transfusions were later reinitiated, but with a lower frequency (2 units of RBC every 2 months), in association with lenalidomide. In other words, over the 20 months of treatment with lenalidomide, the RBC requirements were drastically reduced. However, grade IV neutropenia was observed (granulocytes: $0.5 \times 10^9/L$), without any infectious disease, but leading to treatment stop.

The second patient was a 58-year-old woman (Patient 2, Table 1) whose initial CBC showed hemoglobin 114 g/L, platelet count $710 \times 10^9/L$, leukocytes $7.4 \times 10^9/L$. Bone marrow aspirate showed erythroid hyperplasia with myelodysplastic features, no excess of blasts and 24% ring sideroblasts associated with atypical megakaryocytes, leading

to the diagnosis of MDS/MPN-RS-T.¹ Bone marrow cytogenetics showed a normal karyotype. A watch-and-wait strategy was initiated, but due to a marked increase in the platelet count $> 2000 \times 10^9/L$, hydroxyurea (500 mg/d and then 1000 mg/d) was started. It worsened the anemia to 108 g/L and then to 84 g/L, leading to a weekly use of erythropoietin. After 18 months, hydroxyurea was stopped due to (a) its adverse effects on hemoglobin level and (b) the observation of irregular elevated platelet counts ranging from $1700 \times 10^9/L$ to $3622 \times 10^9/L$. After a drop of platelet count at $1700 \times 10^9/L$, a new increase was observed at $3106 \times 10^9/L$ leading to the start of a lenalidomide treatment (5 mg daily, 21/28 days), secondarily increased to 10 mg daily (21/28 days) since very well tolerated. A marked decrease in the platelet count from 3106 to $760 \times 10^9/L$ was noted, while hemoglobin level raised up from 102 to 118 g/L (Figure 1B). However, for this patient, erythropoietin treatment was maintained every 2 weeks in combination with lenalidomide. Four years after the beginning of lenalidomide therapy, platelet

count remains stable around $750 \times 10^9/L$. Neither adverse effects, nor thrombosis or bleeding occurred.

Informed consent for publication was obtained from both patients.

3 | DISCUSSION

In conclusion, this two-case experiment with more than 3-year follow-up shows the efficacy of lenalidomide in normalizing (Patient 1) or markedly reducing (Patient 2) the platelet count and allowing independency from RBC transfusion in MDS/MPN-RS-T. To our best knowledge, ten MDS/MPN-RS-T cases have been published so far (Table 2).⁹⁻¹⁵ The efficacy of lenalidomide was constant in early stages, except for one case of advanced disease with rapid evolution in myelofibrosis and bone marrow failure.¹⁰ Efficacy of lenalidomide in reducing platelet count has been observed in 3 published cases (Table 2, patients 1, 7, and 10),^{9,14,15} in the context of moderately elevated platelet counts.

Our second case is striking, since it shows that lenalidomide induced a significant decrease in platelet count, even starting from very high counts (up to $3106 \times 10^9/L$), along with an increase in hemoglobin level. So far, the efficacy of lenalidomide has never been shown in MDS/MPN-RS-T with such a high platelet count. It highlights the interest of lenalidomide as an alternative treatment for MDS/MPN-RS-T, including when they present with a major platelet count. Prospective trials are still needed to confirm those encouraging results, but because of rarity of the disease, such trials are very difficult to perform.

ACKNOWLEDGMENTS

The authors thank the association "Tulipes contre le Cancer Lions Club Chalon" for the support. Published with written consent of the patient.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

DM: collected data on clinical and biological parameters, analyzed data, produced the figure, and wrote the manuscript. PA: collected data on clinical and biological parameters. SM: collected data on clinical and biological parameters. VL: collected data on clinical and biological parameters. FP: collected data on clinical and biological parameters. G-BA: collected data on clinical and biological parameters. GF: analyzed data, collected data on clinical and biological parameters, and wrote the manuscript. BJ: analyzed data, collected data on clinical and biological parameters, and wrote the manuscript.

ORCID

Marion Divoux  <https://orcid.org/0000-0002-5946-311X>

Francois Girodon  <https://orcid.org/0000-0003-3151-1068>

Julien Broséus  <https://orcid.org/0000-0002-6689-4006>

REFERENCES

1. Arber DA, Orazi A, Hasserjian R, et al. The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*. 2016;127(20):2391-2405.
2. Broséus J, Florensa L, Zipperer E, et al. Clinical features and course of refractory anemia with ring sideroblasts associated with marked thrombocytosis. *Haematologica*. 2012;97(7):1036-1041.
3. Jeromin S, Haferlach T, Weissmann S, et al. Refractory anemia with ring sideroblasts and marked thrombocytosis cases harbor mutations in SF3B1 or other spliceosome genes accompanied by JAK2V617F and ASXL1 mutations. *Haematologica*. 2015;100(4):125-127.
4. Broséus J, Lippert E, Harutyunyan AS, et al. Low rate of calreticulin mutations in refractory anaemia with ring sideroblasts and marked thrombocytosis. *Leukemia*. 2014;28(6):1374-1376.
5. Patnaik MM, Lasho TL, Finke CM, et al. Predictors of survival in refractory anemia with ring sideroblasts and thrombocytosis (RARS-T) and the role of next-generation sequencing. *Am J Hematol*. 2016;91(5):492-498.
6. Broséus J, Alpermann T, Wulfert M, et al. MPN and MPN-EuroNet (COST Action BM0902). Age, JAK2(V617F) and SF3B1 mutations are the main predicting factors for survival in refractory anaemia with ring sideroblasts and marked thrombocytosis. *Leukemia*. 2013;27(9):1826-1831.
7. Patnaik MM, Tefferi A. Refractory anemia with ring sideroblasts (RARS) and RARS with thrombocytosis (RARS-T): 2017 update on diagnosis, risk-stratification, and management. *Am J Hematol*. 2017;92(3):297-310.
8. List A, Kurtin S, Roe DJ, et al. Efficacy of Lenalidomide in myelodysplastic syndromes. *N Engl J Med*. 2005;352(6):549-557.
9. Huls G, Mulder AB, Rosati S, et al. Efficacy of single-agent Lenalidomide in patients with JAK2 (V617F) mutated refractory anemia with ring sideroblasts and thrombocytosis. *Blood*. 2010;116(2):180-182.
10. Taylor G, Culligan D, Vickers MA. Refractory anemia with ring sideroblasts associated with marked thrombocytosis complicated by massive splenomegaly treated with Lenalidomide resulting in resolution of splenomegaly but severe and prolonged pancytopenia. *Case Rep Hematol*. 2013;2013:718480.
11. Caers J, Hafraoui K, Keutgens A, et al. Haematological and molecular responses in refractory anaemia with ring sideroblasts and thrombocytosis treated with Lenalidomide. *Eur J Haematol*. 2014;92(2):179-180.
12. Zhang M, You Y, Li X, et al. Response to Lenalidomide of a patient with t(2;3)(p23;q29) and JAK2 non-mutated refractory anemia with ring sideroblasts and thrombocytosis. *Leuk Lymphoma*. 2013;54(7):1544-1546.
13. Nichele I, Ruggeri M, Rodeghiero F. Effectiveness of Lenalidomide in a patient with refractory anemia with ring sideroblasts and thrombocytosis with JAK2 (V617F) mutation. *Am J Hematol*. 2015;90(8):E148-E149.

14. Keen R, Pantin J, Savage N, et al. Treatment of refractory anemia with ring sideroblasts associated with marked thrombocytosis with lenalidomide in a patient testing negative for 5q deletion and JAK2 V617F and MPL W515K/L mutations. *Hematol Rep.* 2016;8(4):6592.
15. Nicolosi M, Mudireddy M, Vallapureddy R, et al. Lenalidomide therapy in patients with myelodysplastic syndrome/myeloproliferative neoplasm with ring sideroblasts and thrombocytosis (MDS/MPN-RS-T). *Am J Hematol.* 2018;93(1):E27-E30.

How to cite this article: Divoux M, Plocque A, Sevin M, et al. Efficacy of lenalidomide in myelodysplastic/myeloproliferative neoplasms with ring sideroblasts and an extreme platelet count. *Clin Case Rep.* 2020;8:1774–1780. <https://doi.org/10.1002/ccr3.3026>

Annexe 7. Publication en cours de reviewing *American Journal of Haematology*. Divoux et al. Anti-human T lymphocyte globulin Fresenius in graft-versus-host disease prophylaxis for unrelated hematopoietic stem cell transplantation after myeloablative conditioning: a long-term real-life retrospective study.

Long-term real-life impact of ATLG in GvHD prophylaxis for unrelated HSCT after MAC

Anti-human T lymphocyte globulin Fresenius in graft-versus-host disease prophylaxis for unrelated hematopoietic stem cell transplantation after myeloablative conditioning: a long-term real-life retrospective study

Marion Divoux¹, Matthieu Resche-Rigon², David Michonneau³, Aurélien Sutra Del Galy³, Nathalie Dhedin⁴, Alienor Xhaard³, Flore Sicre de Fontbrune³, Marie Robin³, Gérard Socié³, Régis Peffault de Latour³

(1) Service d'hématologie, CHRU de Nancy, Université de Lorraine, Nancy, France.

(2) Service de biostatistiques et d'information médicale, Hôpital Saint-Louis, AP-HP, Paris, France.

(3) Service d'hématologie-greffe, Hôpital Saint-Louis, AP-HP, Paris, France, Université ParisCité, Paris, France, Inserm UMR 976, Paris, France.

(4) Service d'hématologie Adolescent et jeunes adultes, Hôpital Saint-Louis, AP-HP, Paris, France.

marion.divoux@gmail.com ORCID : 0000-0002-5946-311X

regis.peffaultdelatour@aphp.fr

matthieu.resche-rigon@univ-paris-diderot.fr

david.michonneau@aphp.fr

aurelien.sutra-del-galy@curie.fr

nathalie.dhedin@aphp.fr

alienor.xhaard@aphp.fr

flore.sicre-de-fontbrune@aphp.fr

marie.robin@aphp.fr

gerard.socie@aphp.fr

Corresponding authors:

- Marion Divoux:

marion.divoux@gmail.com

Service d'hématologie Clinique

CHRU Nancy Brabois

Rue du morvan, 54500 Vandoeuvre-lès-Nancy

- Régis Peffault de Latour:

regis.peffaultdelatour@aphp.fr

Service d'hématologie-greffe, Hôpital Saint-Louis, AP-HP

1, avenue Claude Vellefaux 75475 Paris

Key words: anti-human T lymphocyte globulin Fresenius; graft-versus-host disease; myeloablative conditioning; unrelated hematopoietic stem cell transplantation

Main text word count: 1,491

Tables: 0

Figures: 1

Supplementary files: 4

No conflict of interest to disclose. The review of medical records was approved by the Institutional Ethical Committee in agreement with the Helsinki Declaration of 1975, revised in 2008.

Annexe 8. Publication – Hergalant, Saurel, Divoux et al. Correlation between DNA Methylation and Cell Proliferation Identifies New Candidate Predictive Markers in Meningioma.



Article

Correlation between DNA Methylation and Cell Proliferation Identifies New Candidate Predictive Markers in Meningioma

Sébastien Hergalant ^{1,†}, Chloé Saurel ^{2,†}, Marion Divoux ¹, Fabien Rech ^{3,4}, Celso Pouget ², Catherine Godfraind ⁵, Pierre Rouyer ¹, Stéphanie Lacomme ⁶, Shyue-Fang Battaglia-Hsu ⁴ and Guillaume Gauchotte ^{1,2,6,*}

- ¹ UMR Inserm 1256 NGERE (Nutrition, Génétique et Exposition aux Risques Environnementaux), Université de Lorraine, 54000 Nancy, France
 - ² Department of Biopathology CHRU-ICL, CHRU, 54500 Nancy, France
 - ³ Department of Neurosurgery, CHRU, 54500 Nancy, France
 - ⁴ CRAN, CNRS, Université de Lorraine, 54000 Nancy, France
 - ⁵ Neuropathology, CHU of Clermont-Ferrand, UMR INSERM/Université d'Auvergne U1071, 63000 Clermont-Ferrand, France
 - ⁶ Centre de Ressources Biologiques, BB-0033-00035, CHRU, 54500 Nancy, France
- * Correspondence: g.gauchotte@chru-nancy.fr; Tel.: +33-3-83-65-60-17
† These authors contributed equally to this work.



Citation: Hergalant, S.; Saurel, C.; Divoux, M.; Rech, F.; Pouget, C.; Godfraind, C.; Rouyer, P.; Lacomme, S.; Battaglia-Hsu, S.-F.; Gauchotte, G. Correlation between DNA Methylation and Cell Proliferation Identifies New Candidate Predictive Markers in Meningioma. *Cancers* **2022**, *14*, 6227. <https://doi.org/10.3390/cancers14246227>

Academic Editor: David Wong

Received: 2 November 2022

Accepted: 14 December 2022

Published: 17 December 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: In adults, meningioma is the most common primary tumor of the brain. It is classified into three clinical grades of aggressiveness. Whereas disease recurrence after surgery and survival are associated with grade, it is worth investigating proliferation at a molecular level to identify markers capable of improving the clinical management of meningioma. In this study, we explore the DNA methylation profiles of 48 tumors of various grades and conduct statistical analyses on several proliferation indices and markers, such as mitotic index, grade, and Ki-67 or MCM6 expression levels. We identify differential methylation profiles between grades, loci highly correlated with cell growth and division, and a specific methylation signature of regulatory regions persistently associated with proliferation indices, grade, and survival. Finally, we report candidate genes under the control of these regions with potential prognostic and therapeutic value and deserving clinical evaluation.

Abstract: Meningiomas are the most common primary tumors of the central nervous system. Based on the 2021 WHO classification, they are classified into three grades reflecting recurrence risk and aggressiveness. However, the WHO's histopathological criteria defining these grades are somewhat subjective. Together with reliable immunohistochemical proliferation indices, other molecular markers such as those studied with genome-wide epigenetics promise to revamp the current prognostic classification. In this study, 48 meningiomas of various grades were randomly included and explored for DNA methylation with the Infinium MethylationEPIC microarray over 850k CpG sites. We conducted differential and correlative analyses on grade and several proliferation indices and markers, such as mitotic index and Ki-67 or MCM6 immunohistochemistry. We also set up Cox proportional hazard models for extensive associations between CpG methylation and survival. We identified loci highly correlated with cell growth and a targeted methylation signature of regulatory regions persistently associated with proliferation, grade, and survival. Candidate genes under the control of these regions include *SMC4*, *ESRRG*, *PAX6*, *DOK7*, *VAV2*, *OTX1*, and *PCDHA-PCDHB-PCDHG*, i.e., the protocadherin gene clusters. This study highlights the crucial role played by epigenetic mechanisms in shaping dysregulated cellular proliferation and provides potential biomarkers bearing prognostic and therapeutic value for the clinical management of meningioma.

Keywords: genome-wide DNA methylation; meningioma; methylome; proliferation signature; biomarkers; survival; Ki-67; MCM6

Annexe 9. Publication. Aubin, Vilas Boas, Daltro De Oliveira, Le Brun, Divoux et al. *CALR*-mutated patients with low allele burden represent a specific subtype of essential thrombocythemia: A study on behalf of FIM and GBMHM.

Received: 10 November 2023 | Revised: 8 February 2024 | Accepted: 13 February 2024

DOI: 10.1002/ajh.27265

CORRESPONDENCE



CALR-mutated patients with low allele burden represent a specific subtype of essential thrombocythemia: A study on behalf of FIM and GBMHM

To the Editor:

BCR::ABL1-negative myeloproliferative neoplasms (MPNs) are clonal disorders characterized by an overproduction of myeloid mature cells, including essential thrombocythemia (ET), primary myelofibrosis (PMF), and polycythemia vera (PV). Somatic gain-of-function mutations in one of the three driver genes (*JAK2*, *CALR* or *MPL*) are the causative events leading to the constitutive activation of the JAK/STAT pathway and are found in almost 80% of MPN cases.¹ The quantification of the *JAK2V617F* allele burden in blood leukocytes, reflecting the size of the mutant clone, is highly variable from 1% to 100% in MPN patients and correlates with an increased risk of thrombosis and myelofibrotic evolution in PV and ET patients with an allele burden >50%. Some MPN patients display a low *JAK2V617F* allele burden in blood leukocytes (i.e. ≤20%). A recent study demonstrated that this is caused by a bias in clonal hematopoiesis with a late expansion of the mutated clone restricted to the erythroid and/or megakaryocytic lineages.² Conversely, the allele burden of *CALR* mutations appeared to be less variable, with values around 40%–50% associated with clonal dominance in hematopoietic stem cells and progenitors compartments.³ In the literature, *CALR*-mutated MPNs with a low allele burden appear to be present in a small number of patients, but scarce data are available concerning the profile of these patients.

We aimed to characterize *CALR*-mutated MPN with a low allele burden (i.e. <20%) by screening our local cohort and subsequently 11 French hospitals for patients with low allele burden. This study was registered by the French authority CNIL (Commission nationale de l'informatique et des libertés, French Data Protection Authority, authorization ar22-0016v0) and a non-opposition of patients was obtained. *CALR* allele burden quantifications were performed at the time of diagnosis and prior to any cytoreductive treatment (details in Data S1). In our local cohort, a low allele burden was encountered in 15.4% of *CALR*-mutated MPN (22/143) (Figure S1A), and a global frequency of 10.4% in the multicentric screening involving a total of 1479 *CALR*-mutated MPNs was found (Figure S1B). Among low allele burden patients, the diagnostic of ET was found in the vast majority of cases (127 of 135 patients, 94%).

We therefore aimed to compare the 127 low allele burden ET with a multicentric control cohort of 207 *CALR*-mutated ET with an

allele burden ≥20% at the time of diagnosis (details in Data S1). All comparisons are shown in Table S1. The median and range of *CALR* allele burden were of 14% [0.4–19.8] and 38% [20.3–64.6] in the low and high allele burden groups, respectively (Figure S2). The majority of patients (54%) with a low allele burden had an allele burden between 5% and 15%, whereas 4.5% of patients had an allele burden below 5%. The sex ratio was reverted, with a female predominance in low allele burden patients compared with a male predominance in patients with *CALR* allele burden ≥20% (M/F: 0.63 vs. 1.49, $p = 0.0002$). No difference was observed for age at the time of diagnosis ($p = 0.473$). ET patients with *CALR* low allele burden had lower platelets and neutrophils counts, and lower LDH levels than *CALR*-mutated ET patients with a *CALR* allele burden ≥20% ($p = 0.0005$, $p = 0.0008$ and $p = 0.009$, respectively; Figure S3). Among *CALR*-mutated ET, the subtype of *CALR* mutation influenced the phenotype with higher platelet counts in type 2/2-like patients.⁴ However, the phenotype observed in low allele burden *CALR*-mutated ET is not linked to the subtype of *CALR* mutation because the distribution of mutation subtypes was similar between the two groups, with a majority of type 1/1-like mutations (53% and 57% of patients for low and high allele burden, respectively).

In terms of prognosis, *CALR* low allele burden ET exhibited a better overall survival than ET patients with a *CALR* allele burden ≥20% (Figure 1A). Multivariable analysis confirmed the better overall survival of *CALR* low allele burden and showed that age >60 years and a previous history of thrombosis were also independently associated with a decreased overall survival (Figure 1B). No differences were found regarding hematological transformations (Figure 1C) and thrombosis-free survival, although *CALR* low allele burden ET patients may have a lower long-term risk of thrombosis (Figure 1D). Nevertheless, the follow-up period for this cohort is limited (with a median duration of 51 months for low allele burden), and a longer follow-up would provide more conclusions regarding the impact of *CALR* allele burden on hematological transformations and thrombosis. Compared with *JAK2V617F* driver mutation, *CALR* mutations are associated with a better prognosis in PMF, but no difference was found in ET in terms of overall survival, thrombotic events, or fibrotic transformation. Thus, *CALR*-mutated ET patients with a low allele burden seemed to represent a subtype of ET with a more favorable prognosis.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *American Journal of Hematology* published by Wiley Periodicals LLC.

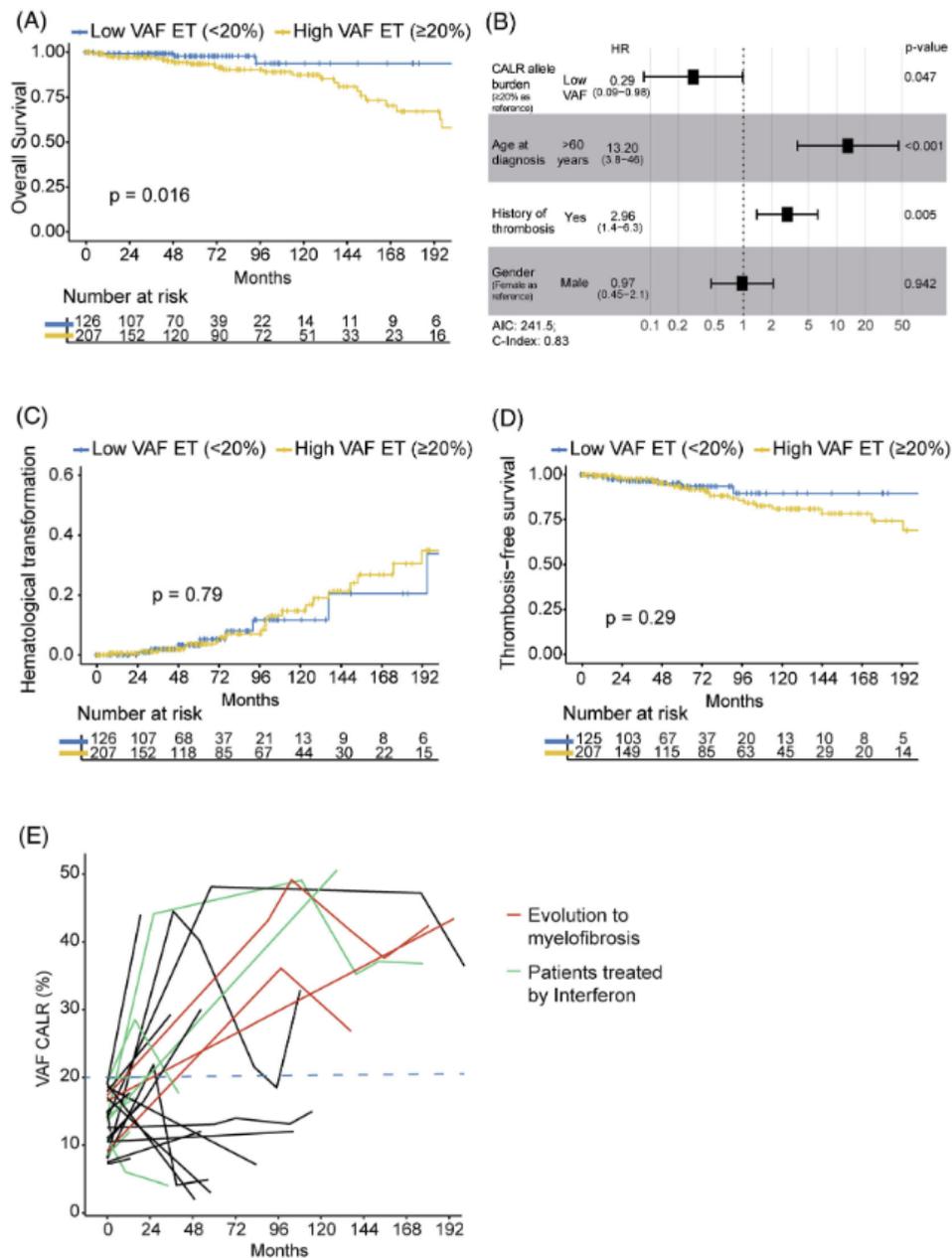


FIGURE 1 Kaplan–Meier curve representing overall survival (A) according to the allele burden of *CALR*-mutated patients at the time of diagnosis. Low and high allele burden groups have been defined using the cutoff of 20%. Forest plot (B) showing the multivariable analysis for overall survival in *CALR*-mutated ET. Kaplan–Meier curves representing the incidence of hematological transformation (C) (either secondary myelofibrosis, myelodysplastic syndrome, or acute leukemia) and thrombosis-free survival (D) according to the allele burden of *CALR*-mutated patients at the time of diagnosis. Trajectories of the allele burden of *CALR* mutations were represented for 26 low allele burden patients (E).

We then aimed to study the evolution of *CALR* allele burden in low allele burden patients. Follow-up samples were available for 26 *CALR* low allele burden ET patients, providing a total of 80 *CALR* measurements with a mean follow-up of 6 years between diagnosis and the last *CALR* quantification. Fifteen of the 26 patients (58%)

remained with a low *CALR* allele burden below 20%, while 11 patients (42%) had an increase of *CALR* allele burden during their follow-up (Figure 1E). Three of the 26 patients progressed to secondary myelofibrosis, and all of them had an increase of their allele burden. This increase was observed at the time of myelofibrotic transformation for

two patients and 40 months before for the third patient. An increase of CALR allele burden could be related to an increase of the clone size and/or to a transition to homozygosity by a loss of heterozygosity (LOH) at chromosome 19. Such homozygosity is a rare event but associated with disease progression.⁵ A previous study also reported the prognostic impact of an increase of CALR allele burden in a cohort of ET patients.⁶ Sixteen of the 26 patients were treated by cytoreductive therapy during follow-up: 6 by hydroxycarbamide, 7 by interferon- α , and 3 by anagrelide. Treatment by interferon- α was not associated with a decrease of CALR allele burden, as 4 out of 7 patients showed an increase during follow-up.

Finally, we aimed to investigate whether a clonal expansion restricted to a lineage exists in CALR MPN, similarly to what has been described in MPN with low allele burden JAK2 mutation.² We quantified CALR mutations in platelets, granulocytes, monocytes, B cells, T cells, NK cells, hematopoietic stem cells (CD34⁺CD38⁻, HSC), and hematopoietic progenitor cells (CD34⁺CD38⁺, HPC) for 19 CALR-mutated patients, including 13 patients undergoing treatment (Table S2 and Figure S4A). Median time between diagnosis and cell sorting was 6 years [2.76;13.48] (Figure S4A). The median allele burden of CALR mutation in total leukocytes was of 38%, and 6 of the 19 patients had low allele burden (i.e., <20%). The allelic burden of CALR mutation was relatively stable among HSC, HPC, and myeloid lineages (Figure S4B). For some patients, we found a lower CALR allele burden in platelets. Among lymphoid lineages, NK cells had the higher level of CALR allele burden, followed by B cells and lastly T cells, which were mostly non-clonal. These results indicate that a skewed expansion of mutated hematopoiesis is not present in CALR-mutated MPN. Future research may help elucidate the mechanisms by which a small clone in all compartments contributes to the observed phenotype.

Overall, our results indicate that CALR-mutated ET patients with a low allele burden exhibit a milder phenotype compared with CALR-mutated ET patients with an allele burden $\geq 20\%$. A low allele burden of CALR mutation at the time of diagnosis also influences the disease course with a better overall survival independently of classical prognostic markers such as age and previous history of thrombosis. In this retrospective real-world study, the diagnoses of ET were determined locally by a multidisciplinary team; however, we cannot exclude the possibility that certain cases may actually represent pre-fibrotic forms. Future large-scale studies with long-term follow-up and systematic monitoring of allele burden are essential for a more comprehensive understanding of the dynamics of evolution of these patients.

AUTHOR CONTRIBUTIONS

V.U., L.C., and D.L.P. conceived the research; L.A., R.V.B., V.L.B., L.C., and D.L.P. performed and analyzed the experiments; R.V.B., R.D.D.O., M.D., J.R., O.M., J.C.I., C.P., A.D., A.M., V.d.M., S.T., F.G., J.S.D., N.B., V.G.M., F.B., C.O., D.R., E.C., L.L.C., C.N., C.R., T.B.L., B.C., P.R., G.B., C.B., K.L., L.B., and E.L. provided clinical and biological data; and L.C. and D.L.P. supervised the project. L.A., L.C. and D.L.P. wrote the original draft and all authors reviewed and approved the manuscript. This paper was written on behalf of the French Intergroup of Myeloproliferative Neoplasms (FIM) and of the 'Groupe des Biologistes Moléculaires des Hémapathies Malignes' (GBMHM).

ACKNOWLEDGMENTS

The authors are indebted to Catherine Guillet from the PACeM facility, who helped with FACS sorting experiments and Rébecca Jouanneau-Courville, who performed CALR quantifications. The authors thank physicians and biologists from the FIM and GBMHM for helpful scientific discussions. The authors also thank all participating biobanks (Centres de Ressources Biologiques) for providing high-quality DNA samples. The authors thank patients, biologists, pathologists, hematologists, and data managers who participated in the FIMBANK project, especially Lea Sureau for data coordination and Elouan Rouméas-Noël for data collection. Finally, the authors thank Willy Vaillant, Marc Muller, Pascale Cornillet-Lefebvre, Tasnim Fethi, Eric Delabesse, Jean Claude Chomel, Nathalie Sorel, Aurélie Chauveau, Vincent Rebière, Françoise Jeanne, and Corinne Souyris for collecting data. This study was supported by grants from the French National Institute of Cancer (INCa BCB 2013 & 2022).

FUNDING INFORMATION

Institut National Du Cancer, Grant/Award Number: FIMBANK-BCB-2013&2022.

CONFLICT OF INTEREST STATEMENT

All authors declare no competing financial interests related to this work.

DATA AVAILABILITY STATEMENT

Data available on request due to privacy/ethical restrictions (individual data cannot be shared without a new regulatory approval).

Laura Aubin^{1,2,3}, Rúben Vilas Boas⁴, Rafael Daltro De Oliveira⁵,
 Valoris Le Brun⁶, Marion Divoux^{7,8}, Jérôme Rey⁹,
 Olivier Mansier^{10,11}, Jean-Christophe Ianotto¹², Cédric Pastoret³,
 Anne Desmares³, Anne Murati¹³, Véronique de Mas¹⁴,
 Suzanne Tavitian¹⁵, François Girodon⁴, Juliette Soret Dulphy⁵,
 Nabih Maslah¹⁶, Valérie Goncalves Monteiro¹⁷, Françoise Boyer¹⁸,
 Corentin Orvain¹⁸, Dana Ranta⁷, Émilie Caussals¹⁹, Lenaig Le Clech²⁰,
 Christophe Nicol²¹, Camille Rottier²², Teresa Botin Lopez²³,
 Brice Castel²⁴, Patrick Rispal²⁵, Guillaume Beziat²⁶,
 Charles Bescond²⁷, Kamel Laribi²⁸, Lina Benajiba⁵, Valérie Ugo¹,
 Eric Lippert^{6,29}, Laurane Cottin¹, Damien Luque Paz¹

¹Univ Angers, Nantes Université, CHU Angers, Inserm, CNRS, CRCI2NA, Angers, France

²Laboratoire d'Hématologie, CH St Malo, St Malo, France

³Laboratoire d'Hématologie, CHU Rennes, Rennes, France

⁴Laboratoire d'Hématologie, CHU Dijon, Dijon, France

⁵INSERM U944/CNRS UMR7212, Hôpital Saint Louis APHP, Université de Paris, Paris, France

⁶Laboratoire d'Hématologie, CHRU Brest, Brest, France

⁷Hématologie Clinique, CHU Nancy, Nancy, France

⁸Laboratoire d'Hématologie, CHU Nancy, Nancy, France

⁹Centre de Recherche en Cancérologie de Marseille (CRCM), Département d'Hématologie, Institut Paoli-Calmettes, CRCM, Inserm, Marseille, France

- ¹⁰Laboratoire d'Hématologie, CHU Bordeaux, Bordeaux, France
¹¹Inserm U1034, Université de Bordeaux, Bordeaux, France
¹²Service d'Hématologie Clinique, CHU Brest, Brest, France
¹³Centre de Recherche en Cancérologie de Marseille (CRCM),
Département de Biopathologie et Département d'Oncologie Prédictive,
Institut Paoli-Calmettes, Inserm, Marseille, France
¹⁴Laboratoire d'Hématologie, Institut Universitaire du Cancer de Toulouse
Oncopole, CHU Toulouse, Toulouse, France
¹⁵Service d'Hématologie, Institut Universitaire du Cancer de Toulouse
Oncopole, CHU Toulouse, Toulouse, France
¹⁶Laboratoire de Biologie Cellulaire, Université Paris Cité, APHP, Hôpital
Saint-Louis, Paris, France
¹⁷Laboratoire d'Hématologie, CHU Reims, Reims, France
¹⁸Service des Maladies du Sang, CHU Angers, Angers, France
¹⁹Service d'Oncologie Hématologique et Thérapie Cellulaire, CHU de
Poitiers, Poitiers, France
²⁰Hématologie Clinique, CH de Cornouaille, Quimper, France
²¹Hématologie Clinique, CH des Pays de Morlaix, Morlaix, France
²²Laboratoire d'oncobiologie moléculaire, CHU Amiens-Picardie,
Amiens, France
²³Hématologie Clinique, CH Castres, Castres, France
²⁴Service de Médecine Interne, CH de Bigorre, Tarbes, France
²⁵Service de Médecine Interne, CH Agen, Agen, France
²⁶Hématologie Clinique, CH Albi, Albi, France
²⁷Hématologie Clinique, CH Cholet, Cholet, France
²⁸Hématologie Clinique, CH Le Mans, Le Mans, France
²⁹INSERM, U1078, Université de Brest, Brest, France

Correspondence

Laurane Cottin and Damien Luque Paz, Laboratoire d'Hématologie,
Institut de Biologie en Santé, CHU ANGERS, 4 Rue Larrey, 49933
ANGERS Cedex 9, France.
Email: laurane.cottin@chu-angers.fr; damien.luquepaz@chu-angers.fr

Laurane Cottin and Damien Luque Paz contributed equally to this work.

ORCID

Cédric Pastoret  <https://orcid.org/0000-0002-5675-3154>
Corentin Orvain  <https://orcid.org/0000-0002-8717-0824>
Eric Lippert  <https://orcid.org/0000-0002-2312-6584>
Damien Luque Paz  <https://orcid.org/0000-0002-4523-675X>

REFERENCES

1. Luque Paz D, Kralovics R, Skoda RC. Genetic basis and molecular profiling in myeloproliferative neoplasms. *Blood*. 2023;141:1909-1921. doi:10.1182/blood.2022017578
2. Nienhold R, Ashcroft P, Zmajkovic J, et al. MPN patients with low mutant JAK2 allele burden show late expansion restricted to erythroid and megakaryocytic lineages. *Blood*. 2020;136:2591-2595. doi:10.1182/blood.2019002943
3. El-Khoury M, Cabagnols X, Mosca M, et al. Different impact of calreticulin mutations on human hematopoiesis in myeloproliferative neoplasms. *Oncogene*. 2020;39:5323-5337. doi:10.1038/s41388-020-1368-3
4. Pietra D, Rumi E, Ferretti VV, et al. Differential clinical effects of different mutation subtypes in CALR-mutant myeloproliferative neoplasms. *Leukemia*. 2016;30:431-438. doi:10.1038/leu.2015.277
5. Guglielmelli P, Maccari C, Sordi B, et al. Phenotypic correlations of CALR mutation variant allele frequency in patients with myelofibrosis. *Blood Cancer J*. 2023;13:21. doi:10.1038/s41408-023-00786-x
6. Cottin L, Riou J, Orvain C, et al. Sequential mutational evaluation of CALR-mutated myeloproliferative neoplasms with thrombocytosis reveals an association between CALR allele burden evolution and disease progression. *Br J Haematol*. 2020;188:935-944. doi:10.1111/bjh.16276

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

RÉSUMÉS

Les néoplasies myélodysplasiques, myéloprolifératives avec mutation de SF3B1 et thrombocytose sont des pathologies clonales de la cellule souche hématopoïétique regroupant des caractéristiques communes aux néoplasies myélodysplasiques (MDS) et aux néoplasies myéloprolifératives (MPN). La mutation JAK2V617F est retrouvée dans seulement 45 à 50% des cas. Les mutations de CALR ou MPL ne sont retrouvées que dans moins de 5% des cas. La prolifération plaquettaire est donc inexplicable pour les 50% restant, conduisant à des difficultés diagnostiques pour ces MDS/MPN-SF3B1-T triple négatifs (MDS/MPN-SF3B1-T TN). Dans le but de comprendre le mécanisme moléculaire expliquant la thrombocytose, nous avons réalisé un séquençage complet d'exome et exploré la méthylation des MDS/MPN-SF3B1-T TN.

Pour l'analyse d'exome, la cohorte de screening était constituée de 12 échantillons de 6 patients atteints de MDS/MPN-SF3B1-T TN : un échantillon cellulaire CD3+ (lymphocytes T, considérés comme tissu sain) et un échantillon CD3- (enrichi en granulocytes, considéré tumoral). L'analyse comparative a permis d'identifier 49 gènes candidats. Il a également mis en évidence un biais de transversion pouvant refléter une atteinte de la réparation BER et la duplication d'une région incluant le gène PTPN12. Le NGS ciblé des gènes candidats sur une cohorte de confirmation a identifié 5 gènes particulièrement pertinents : BRAT1, LMTK3, LTK, POU4F1, SGK223 et STAT5b. L'un des variants de STAT5B induit un épissage alternatif avec saut d'un exon aboutissant à une potentielle protéine tronquée non phosphorylable et pouvant être impliqué dans la prolifération hématopoïétique.

Nous avons également développé ici une nouvelle approche d'analyse du méthylome à partir de données obtenues par RRBS comparatif sur une cohorte composée d'échantillons issus de 7 MDS/MPN-SF3B1-T JAK2V617F, 9 MDS/MPN-SF3B1-T TN, 8 MDS-SF3B1, 10 thrombocytémies essentielles JAK2V617F (TE JAK2V617F) et 4 moelles normales (MO). L'analyse du RRBS a été réalisée avec une nouvelle approche intégrant la position de chaque CpG dans les DMRs (Differentially methylated regions). Cette méthode a permis d'identifier 131 DMRs significativement hypométhylées et 103 DMRs significativement hyperméthylées dans les MDS/MPN-SF3B1-T TN par rapport aux MDS/MPN-SF3B1-T JAK2V617F. Nous avons établi un score clinique de pertinence pour sélectionner les DMRs en lien avec la thrombocytose non reliée à JAK2V617F, identifiant ainsi 15 DMRs annotées à 11 gènes. Neuf gènes candidats étaient reliés à la voie MAPK : 7 gènes hypométhylés dans les MDS/MPN-SF3B1-T TN par rapport aux MDS/MPN-SF3B1-T-JAK2V617F (AATK, BLCAP, DUSP2, FLT4, GPM6A, LHX6 et SPRY2) et 2 gènes hyperméthylés dans les MDS/MPN-SF3B1-T TN (LMX1B et IMPA2). Les transcrits de 8 de ces gènes ont été testés par qPCR dans des échantillons issus du RRBS. La qPCR a confirmé la surexpression de AATK, BLCAP, DUSP2, FLT4, LMX1B et SPRY2 dans les MDS/MPN-SF3B1-T TN par rapport aux MDS/MPN-SF3B1-T-JAK2V617F et par rapport aux MO. IMPA2 était sous-exprimé dans les MDS/MPN-SF3B1-T TN par rapport aux MDS/MPN-SF3B1-T-JAK2V617F et par rapport aux MO. GPM6A était surexprimé dans les MDS/MPN-SF3B1-T TN et les MDS/MPN-SF3B1-T-JAK2V617F par rapport aux MO.

En conclusion, nous avons identifié des candidats génomiques pertinents tels que STAT5B dont les analyses fonctionnelles doivent se poursuivre. Nous avons construit une méthode innovante d'analyse des données de RRBS pour l'identification et la sélection de DMRs potentiellement reliées à la thrombocytose indépendante de JAK2V617F dans les MDS/MPN-SF3B1-T-TN. Nous avons ainsi mis en évidence l'implication de la voie MAPK dans les MDS/MPN-SF3B1-T, en particulier TN. La régulation positive et négative des MAPKs dans les MDS/MPN-SF3B1-T ouvre la voie à de nouvelles thérapies spécifiques et approches diagnostiques dans les MDS/MPN-SF3B1-T et plus généralement dans les MPN non mutés JAK2/CALR/MPL.