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 : ddoc-theses-contact@univ-lorraine.fr

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'UNIVERSITÉ DE LORRAINE

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

par Julien BROSÉUS

Approche génomique des syndromes myéloprolifératifs et des lymphomes B-diffus à grandes cellules en rechute.

Le 12 septembre 2016

Membres du jury :

Rapporteurs :	Florence CYMBALISTA	PU-PH, Université Paris XIII
	Fabrice JARDIN	PU-PH, Université de Rouen
Examinateurs:	Rémi HOULGATTE	DR, Inserm U954, Nancy, Directeur de thèse
	Catherine THIEBLEMONT	PU-PH, Université Paris VII, Co-directrice de thèse
	Pierre FEUGIER	PU-PH, Université de Lorraine
	Jean-Louis MERLIN	PU-PH, Université de Lorraine
	Jean-Noël BASTIE	PU-PH, Université de Bourgogne
	Christian GISSELBRECHT	PU-PH, Université Paris VII

UMR Inserm 954 – Laboratoire de Nutrition génétique et exposition aux risques environnementaux

REMERCIEMENTS

A Monsieur le Docteur Rémi HOULGATTE, Directeur de Thèse, et responsable du groupe « *Bioinformatics and Integrative Genomics* » (BIG) de l'Unité Inserm U954,

Rémi, je t'adresse mes sincères remerciements pour avoir accepté d'être mon Directeur de Thèse et avoir encadré ce travail au cours de ces 4 années. J'ai pu bénéficier de ton expertise dans le domaine de la génomique des lymphomes, et ainsi mieux appréhender ce domaine de recherche. Je souhaite sincèrement que ce travail soit le premier d'une longue série, au sein de l'équipe de recherche dont nous ferons partie dans les années à venir. Je te remercie aussi pour la générosité et la gentillesse dont tu as fait preuve durant ces 4 années. Pour toutes ces raisons, soit assuré de ma gratitude et de mon respect.

A Madame le Professeur Catherine THIEBLEMONT, Co-Directrice de Thèse,

Catherine, je te remercie très chaleureusement de m'avoir confié ce projet de recherche sur la cohorte CORAL. Tu m'as offert un soutien constant et un suivi rigoureux au cours de ces 4 années, ceci malgré l'éloignement géographique et un emploi du temps débordant. J'ai beaucoup appris à tes côtés, dans le domaine des lymphomes, de la collaboration scientifique, mais aussi dans l'art d'écrire un article scientifique. Soit assurée de ma sincère reconnaissance et de mon profond respect.

A Madame le Professeur Florence CYMBALISTA, Rapporteuse de ce travail,

Je suis honoré que vous ayez accepté d'être rapporteuse de ce travail. Votre expertise dans le domaine des hémopathies lymphoïdes sera très précieuse, à la fois pour l'évaluation de ce travail, mais aussi pour la conduite des projets futurs.

A Monsieur le Professeur Fabrice JARDIN, Rapporteur de ce travail,

Je suis honoré que vous ayez accepté d'être rapporteur de ce travail, tant votre compétence dans le domaine de la génomique des lymphomes est reconnue. Il est très important pour moi de bénéficier de votre expertise, à la fois à propos de ce travail, mais aussi à propos des projets à venir.

A Monsieur le Professeur Pierre FEUGIER, membre du groupe « *Bioinformatics and Integrative Genomics* » de l'Unité Inserm U954 et examinateur de ce travail,

Pierre, tu m'as convaincu de venir à Nancy et d'intégrer l'équipe universitaire d'Hématologie de la Faculté de Médecine de Nancy, ainsi que le groupe de recherche « *Bioinformatics and Integrative Genomics* » au sein de l'Unité Inserm U954. Tu m'as toujours fait confiance et soutenu depuis. Je rends hommage à ton enthousiasme et ta volonté dans la construction d'une équipe de recherche sur le lymphome à Nancy. Je suis très heureux que tu fasses partie de ce jury.

A Monsieur le Professeur Jean-Louis MERLIN, examinateur de ce travail,

Je vous remercie vivement d'avoir accepté de faire partie de ce jury. Je suis heureux et honoré de pouvoir vous exposer les travaux réalisés et les projets pressentis. Votre participation à ce jury ouvre la perspective de collaborations futures.

A Monsieur le Professeur Christian GISSELBRECHT, examinateur de ce travail,

A ce jour l'étude CORAL demeure la plus grande étude clinique jamais réalisée sur les lymphomes B-diffus à grandes cellules en rechute. Vous m'avez permis de réaliser ce travail et me faites l'honneur de le juger. Veuillez recevoir l'expression de mon profond respect et de ma sincère reconnaissance.

A Monsieur le Professeur Jean-Noël BASTIE, examinateur de ce travail,

Tu as accompagné mes premiers pas dans la recherche fondamentale à Dijon, d'abord en Master 2 puis au cours d'une première année de thèse. Je suis heureux de pouvoir bénéficier de ton expertise dans les différents domaines de l'Hématologie développés dans ce travail.

A Monsieur le Professeur Jean-Louis GUÉANT, Directeur de l'Unité Inserm U954 N-GERE (Nutrition Génétique et Exposition aux Risques Environnementaux),

Vous m'avez ouvert les portes de votre unité et permis de travailler dans un environnement labellisé et organisé, au contact de professionnels compétents. Sans vous tout ceci n'eut été possible. Veuillez recevoir l'expression de ma profonde reconnaissance et de mon respect.

A Monsieur le Docteur Sébastien HERGALANT, Ingénieur en Bioinformatique, animateur de la plateforme de bioinformatique et membre du groupe « *Bioinformatics and Integrative Genomics* » de l'Unité Inserm U954,

Ton parcours de formation t'a permis de bâtir un large éventail de compétences, alliant la culture et l'appétit scientifique d'un chercheur à une maîtrise approfondie de l'outil informatique. Dans ce domaine, tes compétences en programmation, conception, veille scientifique, maîtrise de l'architecture réseau et des bases de données sont une garantie d'être à jour, malgré l'évolution très rapide des techniques et des connaissances. Ton esprit scientifique, ta capacité de travail en équipe, ton souci du sort des autres et ta grande disponibilité font de toi un moteur de cette équipe. Je te souhaite d'obtenir la reconnaissance institutionnelle que tu mérites. Celle de ceux qui travaillent à tes cotés t'est déjà acquise depuis longtemps.

A Madame le Docteur Aurore PERROT, Maître de Conférences des Universités, Praticien Hospitalier (Université de Loraine et CHRU de Nancy) et membre du groupe « *Bioinformatics and Integrative Genomics* » (BIG) de l'Unité Inserm U954,

Ta grande connaissance de l'ensemble de l'Hématologie, ta haute compétence unanimement reconnue, ton volume de travail vertigineux, ta capacité d'intégration et de synthèse font de toi un exemple pour ceux qui t'entourent. Ta capacité à identifier les bonnes problématiques et à proposer des idées pertinentes rendent ta présence dans cette équipe de recherche à la fois précieuse et indispensable. Notre collaboration, qu'elle soit hospitalière, pédagogique ou scientifique, ouvre de grandes perspectives.

A Madame le Docteur Gaïli CHEN, Docteur de l'Université de Lorraine,

Je te remercie pour ces deux années passées ensemble sur les chemins de la génomique intégrative, dans une ambiance saine et agréable au travail, culturelle et ludique en dehors.

A Madame le Docteur Catherine CHEVALIER, Ingénieure, responsable de la plateforme de génomique intégrative de l'Unité Inserm UMR 1087, Université de Nantes,

Tu m'as accueilli sur ta plateforme à Nantes à plusieurs reprises et permis d'appréhender et de mieux comprendre le versant technique des puces à ADN. Ces séjours Nantais ont été pour moi très précieux. Je te remercie sincèrement pour la gentillesse de ton accueil et la clarté de tes explications.

A Monsieur le Docteur Gérard RAMSTEIN, Maître de Conférences des Universités à l'Université de Nantes, LINA DUKe, UMR 6241,

Je te remercie d'avoir accepté de faire partie de mon comité de suivi de thèse, mais aussi de ton aide et ta participation à nos travaux de recherche. Tes compétences en mathématiques, en statistique et en informatique sont très précieuses, et ton analyse toujours écoutée avec intérêt.

A Monsieur le Docteur Arnaud BIANCHI, Ingénieur de Recherche à l'UMR 7365 CNRS-UL IMOPA (Ingénierie Moléculaire et Physiopathologie Articulaire),

Je vous suis reconnaissant d'avoir accepté de faire partie de mon comité de suivi de thèse et de vous être acquitté de cette tâche avec beaucoup de sérieux.

A Monsieur le Professeur Patrick MENU, Directeur de l'École Doctorale BioSE,

Je vous remercie pour votre disponibilité, votre écoute, votre patience et votre bienveillance.

A Madame Hélène PRIMERANO, secrétaire de l'École Doctorale BioSE,

Je vous remercie pour votre conscience professionnelle, votre connaissance approfondie du fonctionnement de l'École Doctorale et votre grande disponibilité. Aucune de mes questions n'est restée sans réponse.

A tous les membres de l'Unité Inserm U954, pour leur gentillesse, en particulier :

Justine FLAYAC, Dominique GUILLAUME, Jérémy WILLEKENS, Maryvonne MERLET-BILLON, Natacha DREUMONT, Catherine TAVERA, Rémy UMORET, Shyue-Fang HSU-BATTAGLIA, Aurélie ROBERT, Philippe GÉRARD, Céline CHERY, Aline CAZÉ, Pierre ROUYER, Déborah HELLE et Jean-Marc ALBERTO.

A mes cher(e)s interne(s), qui m'ont beaucoup apporté durant ces années, en particulier Julien DECKER, Alexandra JACQUES, Alain DUDA, Loïs REGNAULT, François-Xavier DOUCET, Lucille ALTOUNIAN, Benoît THOMAS, Lucie CHABAS, Samantha SCHULMANN, Brieuc LEFAURE, Alexandre SAULA, Hélène SCHWALLER, Anaëlle CUNTZMANN, Tshinguta Anita MBUYI, Christelle GERVAISE-HENRY, Mathilde TREFFEL, Lauriane FILLIATRE-LEGRAS, Charlotte FELICI, Manon MOUGINOT et Benoît THOMAS.

A Pierre FILHINE-TRESARRIEU, mon camarade de bureau à l'Inserm depuis presque un an maintenant. Et ce n'est qu'un début !

A Delphine GÉRARD et Claire FRANCZAK, deux perles promises à un bel avenir.

A Nicolas MONNIN, confident et ami à la force de caractère prodigieuse, qui m'a tant aidé dans les moments difficiles. Une rencontre comme on en fait peu.

A Monsieur le Docteur Jean DEVIGNES, Chef de Structure du service d'Hématologie biologique,

Vous êtes l'un des principaux artisans de ma venue à Nancy. Depuis mon arrivée dans le service d'Hématologie Biologique, vous m'avez toujours suivi et encouragé. Votre soutien est un des déterminants de mon parcours à Nancy.

A Monsieur le Docteur Jean-François LESESVE,

Nous partageons le même bureau depuis mon arrivée à Nancy en 2012. En tant que cytologiste de référence du service, tu m'as déjà beaucoup appris dans ce domaine, et tu as encore beaucoup à m'apprendre. De notre collaboration est née une vraie complicité. Je suis très heureux que cette collaboration s'étende à la recherche fondamentale, par ton implication dans notre équipe de recherche.

A Madame le Docteur Michèle MARTIN,

Pour ton écoute et ta disponibilité dans les moments difficiles. Mais aussi pour ta compétence en hémostase, à laquelle je fais appel régulièrement. J'ai été très heureux de réaliser mon mémoire de Pédagogie Médicale avec toi. Compte-tenu de tes réelles qualités d'enseignante, j'espère que tu exerceras encore ton talent au sein de la Faculté et que nous y animerons d'autres projets !

A Monsieur le Docteur Sylvain SALIGNAC,

Monsieur Qualité. Je te remercie de m'avoir si bien accueilli dans le service, de m'avoir aidé dans ma formation de cytologiste et pour être toujours disponible en cas de problème ; mais aussi pour nos discussions extra-professionnelles !

A Monsieur le Docteur Julien PERRIN,

Pour ton flegme, ta disponibilité, ton écoute et tes conseils avisés.

A madame le Docteur Véronique LATGER-CANNARD et monsieur le Docteur Frédéric MASSIN

Pour votre accueil et votre propension à animer la plateforme de cytométrie de flux, véritable atout pour notre service.

A tous les membres du service d'Hématologie Biologique du CHRU de Nancy, que j'ai côtoyés durant ces 4 années, en particulier, mesdames les Docteurs FROTSCHER, ESCHWEGE, BRIQUEL et TOUSSAINT-HACQUART

A Madame le Docteur Christine MANENQ, future Assistante Hospitalière et Universitaire dans notre service et dont je suis très fier.

A Monsieur le Professeur Philippe JONVEAUX, Chef du Pôle Laboratoires,

Pour ses conseils et son soutien.

A Madame Véronique SAUNIER,

Pour son aide patiente et répétée dans la finalisation de ce manuscrit !

A Monsieur le Professeur Marc BRAUN, Doyen de la Faculté de Médecine de Nancy,

Marc, je vous remercie de tout cœur pour votre soutien constant et votre bienveillance. J'ai eu la chance de travailler avec vous dans le cadre du DIU de Pédagogie Médicale. Cela a été un réel plaisir et votre tutorat m'a permis de progresser dans ma pratique d'enseignant. Votre disponibilité et votre écoute sont pour moi extrêmement précieuses. Votre force de travail, votre enthousiasme contagieux et votre capacité à rendre les choses limpides constituent pour moi un exemple. Soyez assuré de ma sincère gratitude et de mon profond respect.

A Monsieur le Professeur François GIRODON, Professeur des Universités-Praticien Hospitalier à l'Université de Bourgogne,

François, tu es la première personne qui m'a fait confiance, en me proposant de participer à tes projets de recherche. Nos collaborations scientifiques m'ont apporté beaucoup de satisfaction, ainsi que plusieurs articles qui ont été très importants dans mon parcours. Ton honnêteté, ta droiture, tes qualités humaines ainsi que ta maitrise scientifique et pédagogique constituent des exemples pour moi. Je te remercie encore pour ton soutien inconditionnel. Notre rencontre a été capitale et je souhaite de tout cœur que notre collaboration se poursuive.

A mes parents et à ma sœur,

PUBLICATIONS

Broséus J, Chen G, Hergalant S, Ramstein G, Mounier N, Guéant JL *et al*. Relapsed diffuse large B-cell lymphoma present different genomic profiles between early and late relapses. *Oncotarget*. 2016. *Epub ahead of print*.

Chen G, **Broséus J**, Hergalant S, Donnart A, Chevalier C, Bolaños-Jiménez F *et al*. Identification of master genes involved in liver key functions through transcriptomics and epigenomics of methyl donor deficiency in rat: Relevance to nonalcoholic liver disease. *Mol Nutr Food Res*. 2015; 59(2):293-302.

Broséus J, Park JH, Carillo S, Hermouet S and Girodon F. Presence of calreticulin mutations in JAK2-negative polycythemia vera. *Blood* 2014; 124(26): 3964-6.

Broséus J, Lippert E, Klampfl T, Jeromin S, Zipperer E, Florensa L *et al*. Low rate of calreticulin mutations in refractory anaemia with ring sideroblasts and marked thrombocytosis. *Leukemia.* 2014; 28(6): 1374-6.

Broséus J, Alpermann T, Wulfert M, Brichs LF, Jeromin S, Lippert E *et al*. 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.

SOMMAIRE

INTRODUCTION 28

PARTIE 1 : LES SYNDROMES MYELOPROLIFERATIFS 28

1. ROLE DE JAK2 DANS LA MYELOPOÏESE 29
2. MUTATIONS RECURRENTES DANS LES SMP BCR-ABL NEGATIFS 30
3. LA THROMBOCYTEMIE ESSENTIELLE 32
4. LA POLYGLOBULIE DE VAQUEZ 33

PARTIE 2 : SYNDROMES MYELODYSPLASIQUES. 34

PARTIE 3 : SYNDROMES MIXTES MYELODYSPLASIQUES/ MYELOPROLIFERATIFS 38

LES SYNDROMES MYELOPROLIFERATIFS/MYELODYSPLASIQUES INCLASSABLES 40

L'ANEMIE REFRACTAIRE AVEC SIDEROBLASTES EN COURONNE ET THROMBOCYTOSE 41

PARTIE 4 : SYNDROMES LYMPHOPROLIFERATIFS DE LIGNEE B 51

1. LYMPHOPOÏESE B ET LYMPHOMAGENESE B. 51
   - 1.1. La lymphopoïese B 51
     - 1.1.1. Première étape : l'étape médullaire 52
     - 1.1.2. L'étape ganglionnaire ou immunopoïese 53
   - 1.2. La lymphomagénèse 55
2. CLASSIFICATION DES DIFFERENTS LYMPHOMES MALINS NON HODGKINIENS 56
3. CLASSIFICATION OMS 2008 56
4. LES LYMPHOMES B-DIFFUS A GRANDES CELLULES 59
   - 4.1. Epidémiologie 59
   - 4.2. Clinique 59

4.3. Morphologie/histologie 59
4.3.1. Le variant centroblastique 60
4.3.2. Le variant immunoblastique 60
4.3.3. Le variant anaplasique 60
4.3.4. Immunophénotypage 60
5. PROFIL D'EXPRESSION GENIQUE 61
5.1. Le sous-type Germinal Center B-like (GC) 62
5.2. Le sous-type Activated B-Cell (ABC) 62
6. PRONOSTIC 62
6.1. Pronostic clinique 62
6.2. Apports pronostics du profil d'expression génique 64
6.3. Apports de l'immunohistochimie 64
6.4. Apports de la FISH (Fluorescence In Situ Hybridization) 65
7. THERAPEUTIQUE 66
7.1. Traitements de première ligne 66
7.2. Traitement des sujets jeunes (< 60 ans) 67
7.3. Traitement des sujets âgés (60-80 ans) 68
7.4. Traitement des sujets très âgés (> 80 ans) 68
7.5. Traitement des rechutes 68
8. DEREGULATIONS MOLECULAIRES DANS LES LBDGC 71
8.1. Anomalie associées préférentiellement au sous-type GC 71
8.2. Translocations impliquant le gène *BCL2* 71
8.3. Activation constitutive de BCL6 72
8.4. Activation de la voie PI3K/AKT/mTOR 74
8.5. Anomalies de la régulation épigénétique 75

8.6. Anomalies moléculaires affectant la domiciliation des lymphocytes B 76

9. ANOMALIES ASSOCIEES PREFERENTIELLEMENT AU SOUS-TYPE ABC .. 77

9.1. Mutations entraînant une activation chronique du BCR 79

9.1.1. Mutations de CD79A et CD79B 79

9.1.2. Complexe CARD11-BCL10-MALT1..... 79

9.2. Mutations entraînant une activation chronique du TLR..... 80

9.3. Délétions bi-alléliques de *TNFAIP3/A20*..... 81

9.4. Activation de la voie NOTCH..... 81

9.5. Surexpression de la serine-thréonine kinase PIM1 82

10. ANOMALIES MOLECULAIRES AFFECTANT LA REGULATION DU CYCLE CELLULAIRE 83

10.1. Anomalies du locus CDKN2A/B..... 83

10.2. Anomalies du facteur de transcription MYC..... 83

10.3. Anomalies de TP53 85

11. ANOMALIE CONDUISANT A L'ECHAPPEMENT A LA REPONSE IMMUNE . 87

PROBLEMATIQUE GENERALE ET OBJECTIFS 88

Partie 1 : Génomique des syndromes myéloprolifératifs et des syndromes mixtes myéloprolifératifs / myélodysplasiques. 88

Partie 2 : Génomique des lymphomes B-diffus à grandes cellules en rechute..... 90

MATERIELS ET METHODES 92

1. COHORTE CORAL 92

2. L'IMPACT DU STATUT GC/ABC SUR L'EFFICACITE DU TRAITEMENT DE RATTRAPAGE : L'ETUDE BIOCORAL 95

3. LES PUCES A ADN..... 96

3.1. Technologie..... 96

3.2. Etude du transcriptome 97

3.2.1. Normalisation des données 98

3.2.2. Regroupement des données selon la méthode du *K-means* 100

3.2.3. Regroupement des données selon la méthode de la classification hiérarchique 102

3.2.4. Annotation fonctionnelle des clusters 104

3.3. Etude des anomalies du nombre de copies géniques et des polymorphismes de l'ADN 106

3.3.1. Puce Affymetrix SNP 6.0 106

3.3.2. L'algorithme CRLMM 107

3.3.3. Concept de fenêtre glissante 108

3.3.4. Définition des CNV 109

RESULTATS 110

PARTIE 1 : GENOMIQUE DES SYNDROMES MYELOPROLIFERATIFS ET DES SYNDROMES MIXTES MYELOPROLIFERATIFS / MYELODYSPLASIQUES 110

1. AGE, JAK2^{V617F} AND *SF3B1* MUTATIONS ARE THE MAIN PREDICTING FACTORS FOR SURVIVAL IN REFRACTORY ANAEMIA WITH RING SIDEROBLASTS AND MARKED THROMBOCYTOSIS 110

2. LOW RATE OF CALRETICULIN MUTATIONS IN REFRACTORY ANAEMIA WITH RING SIDEROBLASTS AND MARKED THROMBOCYTOSIS ... 127

3. PRESENCE OF CALRETICULIN MUTATIONS IN JAK2-NEGATIVE POLYCYTHEMIA VERA 137

PARTIE 2 : GENOMIQUE DES LYMPHOMES B-DIFFUS A GRANDES CELLULES EN RECHUTE 144

1. RELAPSED DIFFUSE LARGE B-CELL LYMPHOMA PRESENT DIFFERENT GENOMIC PROFILES BETWEEN EARLY AND LATE RELAPSES 144

DISCUSSION GENERALE 176

CONCLUSION ET PERSPECTIVES 181

1. ETUDE GENETIQUE-GENOMIQUE DES LYMPHOMES 181

2. ETUDE GENOMIQUE DU SYNDROME DE RICHTER 185

REFERENCES BIBLIOGRAPHIQUES 190

LISTE DES FIGURES

Figure 1. Effets de l'activation permanente de l'activité kinase de JAK2 sur les voies de signalisation dépendantes de JAK2. 30
Figure 2. Panorama mutationnel des SMP *BCR-ABL* négatifs. 31
Figure 3. Syndromes myéloprolifératifs, syndromes myélodysplasiques et syndromes mixtes myélodysplasiques/myéloprolifératifs. 39
Figure 4. Physiopathologie des SMP/SMD. 39
Figure 5. Trois modèles pour expliquer l'origine de l'ARS-T. 42
Figure 6. Les acteurs de l'épissage alternatif. 45
Figure 7. Distribution des mutations de *SF3B1*. 46
Figure 8. Composition d'une immunoglobuline. 52
Figure 9. Processus de recombinaison « VDJ ». 53
Figure 10. Lymphomes B assignés à leur équivalent lymphocytaire B normal. 57
Figure 11. Les LBDGC de type GC et ABC sont identifiables par leur signature transcriptomique. 61
Figure 12. Classification pronostique d'Ann Arbor. 63
Figure 13. Classification pronostique selon le score IPI. 63
Figure 14. Les LBDGC de type GC et ABC sont deux entités de pronostic différent. 64
Figure 15. Stratégie de traitement de première ligne des LBDGC. 67
Figure 16. Survie globale chez 7400 patients atteints de LBDGC âgés de 18 à 80 ans, en fonction de leur réponse au traitement de première ligne présentant un LBDGC et inclus dans les protocoles du GELA. 69
Figure 17. Principales voies de régulation impliquées dans la genèse des LBDGC de type GC. 71
Figure 18. Fonctions cellulaires régulées par BCL6 au sein du centre germinatif. 73
Figure 19. Voie NF-κb. 78
Figure 20. Principales voies de régulation cellulaires impliquées dans la genèse des LBDGC de type ABC. 78
Figure 21. Voie du TLR. 80
Figure 22. Voie NOTCH. 82
Figure 23. Proportion de réarrangements et de surexpression protéique de MYC, BCL2 et BCL6 dans les LBDGC. 85
Figure 24. Différents modes d'action de TP53 dans le lymphocyte. 86

Figure 25 : Devenir des patients atteints de LBDGC. 90
Figure 26. Conception de l'étude CORAL. 93
Figure 27. Etude CORAL, survie globale (A) et survie sans progression (B) en fonction du traitement décidé en première randomisation (R-ICE ou R-DHAP). 94
Figure 28. Etude CORAL, survie sans progression et survie globale en fonction du traitement de seconde ligne entrepris : R-DHAP (A,C) ou R-ICE (B,D) et du statut GC/ABC. 95
Figure 29. Structure d'une puce GeneChip® commercialisée par Affymetrix. 96
Figure 30. Obtention de la matrice d'expression des gènes. 98
Figure 31. Normalisation. A. Lowess; B. Statistiques descriptives avant et après normalisation; C. Corrélation de chaque échantillon par rapport au profil médian avant et après normalisation. 100
Figure 32. Exemple d'utilisation de la méthode du *K-means*. 101
Figure 33. Principe de la classification hiérarchique. 102
Figure 34. Exemple de résultat obtenu grâce à la méthode de la classification hiérarchique. D'après Bogard *et al*.[156] 103
Figure 35. Exemples d'ontologies des gènes. 105
Figure 36. Mécanismes physiopathologiques induits par les mutants de *CALR* dans les SMP. 177
Figure 37. Caractéristiques moléculaires des syndromes mixtes myélodysplasiques/myéloprolifératifs. 178
Figure 38. Reconstitution des réseaux de régulation génique grâce à l'approche génétique-génomique. 181
Figure 39. Le concept génétique-génomique et la correspondance entre séquence génique (ADN), niveau d'expression des gènes (ARN) et phénotype. 182
Figure 40. Méthode de croisement systématique des données de génotypage et d'expression. 183
Figure 41. Les déterminants de la transformation de la LLC en SR. 188

LISTE DES TABLEAUX

Tableau I. Liste des syndromes myéloprolifératifs. 29
Tableau II. Critères diagnostiques de la TE. 32
Tableau III. Caractéristiques des dysplasies..... 35
Tableau IV: Causes non clonales de dysmyélopoïèse. 35
Tableau V. Anomalies cytogénétiques les plus fréquentes dans les SMD. 35
Tableau VI. Caractéristiques cytologiques des syndromes myélodysplasiques. 36
Tableau VII. Score pronostique IPSS. 37
Tableau VIII. Syndromes mixtes myéloprolifératifs / myélodysplasiques. 38
Tableau IX. Critères diagnostiques pour les SMP/SMD inclassables. 40
Tableau X. Fréquence des mutations de *SF3B1* dans les pathologies cancéreuses. 48
Tableau XI. Classification OMS 2008 des hémopathies malignes B matures. 58
Tableau XII. Puce Affymetrix *Genome-wide Human SNP Array 6.0* nombre de sondes couvrant chaque chromosome. 107

LISTE DES ANNEXES

ANNEXE 1 Protocole *Agilent Two-Color Microarray-Based Gene Expression.* 202
ANNEXE 2 Protocole d'utilisation de la puce Affymetrix SNP 6.0 204

LISTE DES ABREVIATIONS

ABC	*Activated B-Cell like*
ACVBP	Doxorubicin, Cyclophosphamide, Vindesine, Bleomycin, Prednisone
ADCC	*Antibody derived cell-mediated cytotoxicity*
ADN	Acide Desoxyribo Nucléique
AKT	*Protein Kinase B*
ALK	*Anaplastic Lymphoma Kinase*
AR	Anémie Réfractaire
AREB	Anémie Réfractaire avec Excès de Blastes
ARF	*Alternate open Reading Frame*
ARS	Anémie Réfractaire Sidéroblastique
ARS-T	Anémie Réfractaire avec Sidéroblastes en couronne et Thrombocytose
ASCT	*Autologous Stem-Cell Transplantation*
ATM	*Ataxia Telangiectasia Mutated*
B2M	Beta-2 microglobuline
BAD	*BCL2-associated Agonist of cell Death*
BCL2	*B-Cell/CLL Lymphoma 2*
BCL6	*B-Cell/CLL Lymphoma 6*
BCL10	*B-cell lymphoma leukaemia 10*
BCR	*B-cell receptor*
BCR-ABL1	*Breakpoint Cluster Region – Abelson 1*
BEAM	Carmustine Etoposide Cytarabine Melphalan
BTK	*Bruton Tyrosine Kinase*
CALR	Calréticuline

CARD11	*Caspase recruitment domain containing protein 11*
CDK	*Cyclin-Dependent Kinase*
CDKN2A/B	*Cyclin-Dependent Kinase Inhibitor A/B*
CGH	*Comparative Genomic Hybridization*
CH	*Constant Heavy*
CHAS	*Chromosome Analysis*
CHOEP	Cyclophosphamide, Adriblastine, Oncovin, Etoposide, Prednisone
CHOP	Cyclophosphamide, Adriblastine, Oncovin, Prednisone
CIITA	*Class II, major histocompatibility complex, transactivator*
CL	*Constant Light*
CNV	*Copy Number Variation*
CORAL	*Collaborative Trial in Relapsed Agressive Lymphoma*
CRDM	Cytopénie Réfractaire avec Dysplasie Multilignée
CREB	*C-AMP Response Element-binding protein*
CREBBP	*CREB Binding Protein*
CRF	*Case Report Form*
CSH	Cellule Souche Hématopoïétique
CXCL12	Chemokine (C-X-C Motif) Ligand 12
DAG	*DyAcyl-Glycérol*
DGV	*Database of Genomic Variants*
DHAP	Dexamethasone, Aracytine haute dose, Cisplatine
DRAK1	*DAP kinase-Related Apoptosis-inducing protein Kinase 1*
DSHNHL	*Deutsche Studiengruppe für Hochmaligne Non-Hodgkin-Lymphome*
EBV	*Epstein-Barr Virus*
ECOG	*Eastern Cooperative Oncology Group*
EP300	*E1A binding Protein p300*

EPO	Erythropoïétine
EZH2	*Enhancer of Zeste 2 polycomb repressive complex 2 subunit*
FDR	*False Discovery Rate*
FGFR1	*Fibroblast Growth Factor Receptor 1*
FISH	*Fluorescence In Situ Hybridization*
FLI1	*Friend Leukemia Virus Integration 1*
FOXP1	*Forkhead-box P1*
GC	*Germinal Center B-like*
GELA	Groupe d'Etude des Lymphomes de l'Adulte
GISTIC	*Genomic Identification of Significant Targets in Cancer*
GNAI2	*Guanine Nucleotide binding protein, alpha inhibiting activity polypeptide 2*
GNA13	*Guanine Nucleotide Binding Protein, Alpha 13*
GO	*Gene Ontology*
G-CSF	*Granulocyte-Colony Stimulating Factor*
HEAT	*Huntingtin Elongation factor 3, protein phosphatase 2A, TOR1*
HHV8	*Human Herpes Virus 8*
HLA	*Human Leukocyte Antigen*
HMGXB4	*High Mobility Group Box domain containing 4*
HOX11	*Homeo Box 11*
Ig	Immunoglobuline
IgH	Immunoglobuline Heavy Locus
IgK	Immunoglobuline Kappa Locus
IgL	Immunoglobuline Lambda Locus
IgVH	*Immunogolobuline Variable Heavy*
IgVL	*Immunogolobuline Variable Light*
IPI	*International Prognostic Index*

IPSS	*International Prognostic Scoring system*
IRAK	*IL-1 Associated Kinases*
IRF4	*Interferon Regulatory Factor 4*
JAK2	*Janus Kinase 2*
LAM	Leucémie Aiguë Myéloblastique
LBDGC	Lymphome B Diffus à Grandes Cellules
LLC	Leucémie Lymphoïde Chronique
LMC	Leucémie Myéloïde Chronique
LMMC	Leucémie Myélo-Monocytaire Chronique
LMMJ	Leucémie Myélo-Monocytaire Juvénile
LMP1	*Latent Membrane Protein 1*
LMO2	*LIM domain Only 2*
LNH	Lymphomes non hodgkiniens
LOWESS	*Locally Weight Scatterplot Smoothing*
LRMP	*Lymphoid Restricted Membrane Protein*
MALT	*Mucosae Associated Lymphoid Tissue*
MALT1	*Mucosa Associated Lymphoid Tissue Lymphoma Translocation Protein 1*
MAPK	*Mitogen Activated Protein Kinase*
MDM2	*Mouse Double Minute 2 homolog*
MEF2B	*Myocyte Enhancer Factor 2B*
MFP	Myélofibrose Primitive
MLL2	*Mixed Lineage Leukemia 2*
MPL	*Myeloproliferative Leukemia*
m-TOR	*mammalian Target Of Rapamycin*
MYC	*Avian Myelocytomatosis virus oncogene homolog*
MYD88	*Myeloid Differenciation primary response 88*

NF-κB	*Nuclear Factor kappa B*
OMS	Organisation Mondiale de la Santé
PCR	*Polymerase Chain Reaction*
PDGFRA	*Platelet Derived Growth Factor Receptor Alpha polypeptide*
PDGFRB	*Platelet Derived Growth Factor Receptor Beta polypeptide*
PEST	*Proline, Glutamic Acid, Serine and Threonine*
PET	*Positron Emission Tomography*
PKC	*Protein Kinase C*
PI3K	Phosphatidyl Inositol 4,5-biphosphate 3 Kinase
PIM1	Pim-1 proto-oncogene, serine/threonine kinase
PIM2	Pim-2 proto-oncogene, serine/threonine kinase
PNN	Polynucléaire Neutrophile
PR	*Partial Response*
PRC2	*Polycomb Repressive Complex 2*
PTEN	*Phosphatase and Tensin Homolog*
PV	Polyglobulie de Vaquez
QTL	*Quantitative Trait Loci*
R	Rituximab
R-ACVBP	Rituximab, Doxorubicin, Cyclophosphamide, Vindesine, Bleomycin, Prednisone
R-CHOP	Rituximab-Cyclophosphamide Hydroxydaunorubicine Oncovin Prednisone
R-DHAP	Rituximab-Dexamethasone High dose cytarabine cisPlatine
R-ICE	Rituximab- Ifosfamide Carboplatine Etoposide
RICOVER 60	*Rituximab with CHOP Over Age 60 Years*
S1PR2	*Sphingosine-1-Phosphate Receptor 2*
SAM	*Significance Analysis of Microarray*
SC	Sidéroblaste en couronne

SF3B1	*Splicing Factor 3B subunit 1*
SLP	Syndrome lympho-prolifératif
SMD	Syndrome myélodysplasique
SMP	Syndrome myéloprolifératif
SNP	*Single Nucleotide Polymorphism*
snRNA	*Small nucleolar RiboNucleo Acid*
snRNP	*small nucleolar RiboNucleo Protein*
SRSF2	*Serine/arginine-rich Splicing Factor 2*
STAT	*Signal Transducer and Activator of Transcription*
TE	Thrombocytémie Essentielle
TGFβ	*Tumor Growth Factor Beta*
TNFα	*Tumor Necrosis Factor alpha*
TNFAIP3	*Tumour Necrosis Factor Alpha-Induced Protein 3*
TLR	*Toll-Like Receptor*
TP53	*Tumor Protein p53*
TPO	Thrombopoïétine
U2AF35	*U2 small nucleolar RNA Auxiliary Factor 35*
WNT	*Wingless-Type MMTV Integration Site*
ZRSR2	*Zing finger CCCH type, RNA-binding motif and Serine/arginine rich 2*

INTRODUCTION

La classification 2008 des tumeurs des tissus hématopoïétiques et lymphoïdes proposée par l'Organisation Mondiale de la Santé (OMS) définit entre autres les syndromes myéloprolifératifs (SMP), les syndromes myélodysplasiques (SMD) et les syndromes lymphoprolifératifs (SLP).

PARTIE 1 : LES SYNDROMES MYELOPROLIFERATIFS

Les SMP constituent un ensemble d'hémopathies malignes chroniques acquises et clonales affectant la cellule souche hématopoïétique (CSH). Ils sont caractérisés par l'hyperplasie d'une ou plusieurs lignées myéloïdes, constituée de cellules matures, sans blocage de maturation et sans avortement intra-médullaire. D'après les données du Registre des Hémopathies Malignes de Côte d'Or (période 1980-2004) et celles du Registre de Basse-Normandie (période 1997-2004), les SMP représentent entre 11,6 et 12% des hémopathies malignes. Leur incidence en Côte d'Or est de 3,2 nouveaux cas pour 100 000 habitants et par an. La médiane de survie des SMP est de 91,7 mois.[1, 2]

Ces pathologies sont caractérisées par une moelle de cellularité augmentée avec une maturation hématopoïétique complète, conduisant à une augmentation du nombre de cellules circulantes (granulocytes, érythrocytes et/ou plaquettes). Cliniquement, ces hyperplasies des lignées myéloïdes se traduisent par un syndrome tumoral, avec une splénomégalie et/ou une hépatomégalie, un syndrome d'hyperviscosité et un risque accru de complications thrombotiques ou hémorragiques. Ces pathologies présentent un risque d'évolution vers l'insuffisance médullaire par myélofibrose secondaire, et/ou vers la leucémie aiguë myéloblastique (LAM).[3] La liste des SMP admis par l'OMS figure dans le **tableau I**.

Syndromes myéloprolifératifs
Leucémie myéloïde chronique, *BCR-ABL1* positive
Leucémie neutrophile chronique
Maladie de Vaquez
Myélofibrose primitive
Thrombocytémie essentielle
Leucémie éosinophile chronique
Mastocytose
Syndromes myéloprolifératifs inclassables

Tableau I. Liste des syndromes myéloprolifératifs.
D'après Vardiman *et al*.[3]

Compte-tenu de leur plus forte incidence, la leucémie myéloïde chronique (LMC), la polyglobulie de Vaquez (PV), la thrombocytémie essentielle (TE) et la myélofibrose primitive (MFP) constituent les 4 principaux SMP.

La physiopathologie des SMP s'explique par des anomalies clonales touchant des gènes codant pour des protéines à activité tyrosine kinase. Ces mutations conduisent à une activation constitutive de ces protéines et des voies de transduction qu'elles gouvernent, entraînant une prolifération cellulaire anormale. Le caryotype, indispensable, s'attache à rechercher la présence d'une anomalie cytogénétique particulière, la translocation t(9;22)(q34;q11.2), donnant naissance au gène de fusion *BCR/ABL1* (*Breakpoint Cluster Region – Abelson 1),* spécifique de la LMC. Les 3 autres principaux SMP ne présentent pas cette anomalie, mais sont caractérisés par une forte proportion de mutations affectant le gène *JAK2* (*Janus Kinase 2*).

1. ROLE DE JAK2 DANS LA MYELOPOÏESE

La protéine JAK2 est une protéine à activité tyrosine kinase associée à plusieurs récepteurs cytokiniques de classe I dont plusieurs sont essentiels dans la myélopoïèse, en particulier le récepteur à l'érythropoïétine (EPO) et le récepteur à la thrombopoïétine (TPO). Le récepteur à l'EPO est un homodimère dont la partie cytoplasmique est phosphorylée par la tyrosine kinase JAK2 en présence du ligand, l'EPO. La transduction du signal passe par la phosphorylation croisée de résidus tyrosine des deux protéines JAK2 associées au récepteur. Ceci permet ensuite l'activation du facteur de transcription STAT (*Signal Transducer and Activator*

of Transcription), des voies MAPK (*Mitogen Activated Protein Kinase*) et PI3K (*Phosphatidyl Inositol 3 Kinase*) et donc à la prolifération des cellules hématopoïétiques (**Figure 1**). JAK2 possède un domaine inhibiteur qui interrompt la mise en jeu de la voie d'activation dès lors que le ligand est détaché du récepteur.

La mutation du gène *JAK2*, situé au niveau du bras court du chromosome 9 (9p24), exerce un rôle majeur dans la physiopathologie des SMP *BCR-ABL1* négatifs.[4-7] La plus fréquente de ces mutations, est une substitution d'une guanine par une thymine au sein de l'exon 14 du gène menant au remplacement d'une valine par une phénylalanine en position 617 de la protéine : *JAK2*^{V617F}. Cette mutation conduit à la synthèse d'un récepteur JAK2 dont le domaine inactivateur JH2 est non fonctionnel. Ceci entraîne une hypersensibilité aux cytokines activatrices, une activation constitutive de JAK2 et une mise en jeu prolongée de plusieurs voies de signalisation de récepteurs aux cytokines, dont le récepteur à l'EPO et le récepteur à la TPO, menant à la prolifération des lignées myéloïdes.

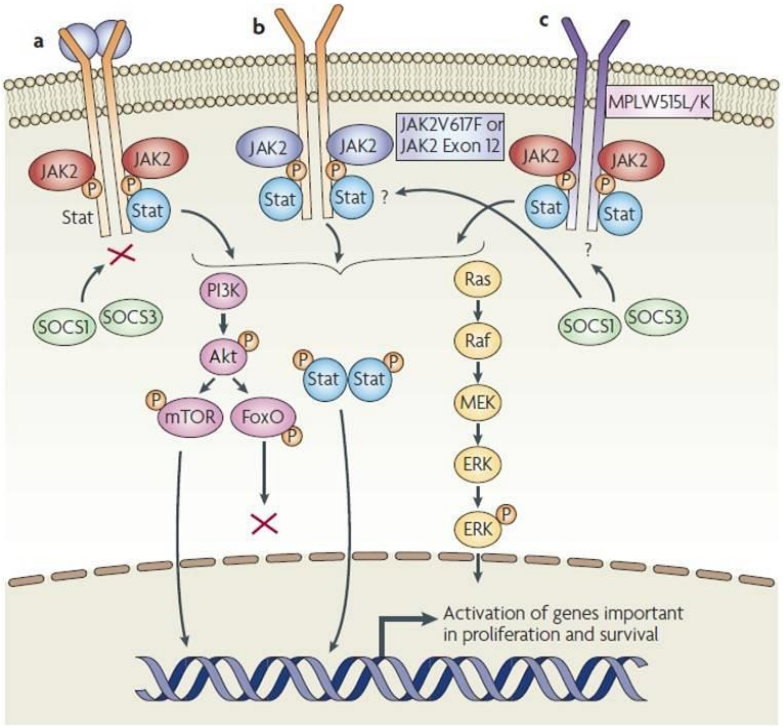


Figure 1. Effets de l'activation permanente de l'activité kinase de JAK2 sur les voies de signalisation dépendantes de JAK2.
D'après Levine *et al*.[7]

2. MUTATIONS RECURRENTES DANS LES SMP BCR-ABL NEGATIFS

La mutation *JAK2*^{V617F} est retrouvée chez 95 à 97 % des patients atteints de PV, 60 % des patients atteints de MFP et 60 % des patients atteints de TE. Chez 1 à 2% patients atteints de la maladie de Vaquez *JAK2*^{V617F} négatifs, une mutation activatrice de *JAK2* peut être retrouvée au niveau de l'exon 12, avec des conséquences fonctionnelles équivalentes à celle de la mutation *JAK2*^{V617F}.

La mutation *JAK2*^{V617F} n'est pas spécifique d'un SMP particulier, de même que son absence n'exclut pas le diagnostic de SMP. De telles mutations ont d'ailleurs été décrites dans des cas de syndromes mixtes myéloprolifératifs/myélodysplasiques (SMP/SMD), en particulier l'Anémie Réfractaire avec Sidéroblastes en couronne et Thrombocytose (ARS-T).[8-16]

Plusieurs mutations activatrices du gène *MPL* (*MyéloProliferative Leukemia)*, codant pour le récepteur de la TPO ou) : *MPL*^{W515L} ou *MPL*^{W515K} peuvent être retrouvées dans 3 à 5 % des cas de TE et 5 à 10% des cas de MFP.[17, 18]

Enfin, il a été montré en 2013 que les mutations se situant dans l'exon 9 du gène codant pour la calréticuline (*CALR*) sont retrouvées dans plus de 2 tiers des cas de TE et de MFP ne présentant ni mutation de *JAK2* ni mutation de *MPL* (**Figure 2**).[19, 20]

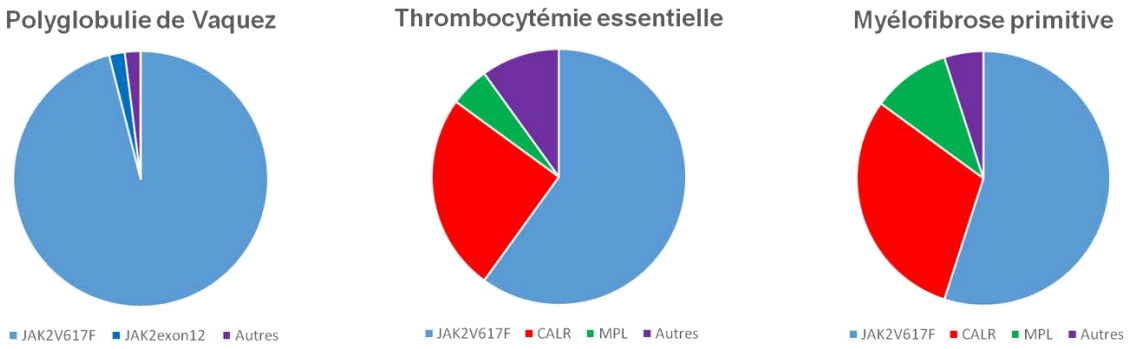


Figure 2. Panorama mutationnel des SMP *BCR-ABL* négatifs.

3. LA THROMBOCYTEMIE ESSENTIELLE

La TE est le plus fréquent des SMP et implique de manière prépondérante la lignée mégacaryocytaire. Elle est caractérisée par une thrombocytose (> 450 x 10^9 plaquettes/L) avec anisocytose plaquettaire, micromégacaryocytes circulants, un nombre augmenté de mégacaryocytes géants, dysmorphiques (noyaux hyperlobés en « bois de cerfs »), une hyperleucocytose modérée avec une discrète myélémie et des épisodes de thrombose et/ou d'hémorragie. L'existence d'une dysérythropoïèse et/ou d'une dysgranulopoïèse marquées doit plutôt faire évoquer le diagnostic de SMD. Le réseau réticulinique est normal ou seulement légèrement augmenté. Les autres causes de thrombocytose doivent être exclues (autre SMP, syndrome inflammatoire, carence en fer). Les critères diagnostiques de la TE sont listés dans le **tableau II**.[21]

Critères diagnostiques de la TE
1. Numération plaquettaire > 450 x 10^9/L
2. Biopsie ostéomédullaire montrant une prolifération touchant principalement la lignée mégacaryocytaire avec un nombre augmenté de grands mégacaryocytes matures. Pas d'augmentation significative de la proportion ni d'anomalie de maturation des précurseurs granuleux ou érythroblastiques.
3. Absence des critères diagnostiques de la maladie de Vaquez, de la myélofibrose primitive, de la leucémie myéloïde chronique *BCR-ABL1* positive, d'un syndrome myélodysplasique ou d'une autre néoplasie myéloïde.
4. Présence de la mutation *JAK2*^{V617F} ou d'un autre marqueur clonal, ou à défaut, pas d'argument en faveur d'une thrombocytose réactionnelle.

Tableau II. Critères diagnostiques de la TE.
D'après Thiele *et al*.[21]

Au sein des SMP, la TE a l'incidence la plus élevée avec 1,2 nouveau cas pour 100 000 habitants et par an.[1] Une partie des patients ne présente pas de point d'appel clinique et le diagnostic est suspecté sur les résultats d'un hémogramme. Pour les autres patients le diagnostic est suspecté à la suite d'une manifestation thrombotique ou hémorragique. Une splénomégalie est retrouvée dans 50% des cas au diagnostic. La mutation *JAK2*^{V617F} est retrouvée dans 60% des cas et permet la démonstration d'une clonalité.[18, 22] De la même façon, l'étude de la pousse spontanée de progéniteurs mégacaryocytaires *in vitro*, bien que non spécifique de la TE, permet d'éliminer une origine réactionnelle.[23]

La TE est une maladie indolente caractérisée par de longs intervalles de survie sans événement interrompus par des événements thrombo-emboliques veineux et artériels liés à la thrombocytose et à des anomalies fonctionnelles plaquettaires, ainsi que par des événements hémorragiques liés à une thrombopathie. Après un certain nombre d'années, la TE peut se compliquer d'une myélofibrose. L'évolution de la TE en LAM est rare (< 5% des patients) et en général due à une thérapie cytotoxique prolongée par hydroxyurée. L'espérance de vie des patients atteints de TE est proche de la normale dans la plupart des cas (médiane de survie de 129 mois).[24]

4. LA POLYGLOBULIE DE VAQUEZ

La PV a une incidence de 0,6 nouveau cas pour 100 000 habitants et par an ce qui la classe en troisième position parmi les SMP en terme d'incidence, derrière la TE et la LMC. Elle est caractérisée par une survie longue équivalente à celle de la TE.

Ce SMP prédomine sur la lignée érythroblastique et se traduit par une augmentation du taux d'hémoglobine, souvent accompagnée d'une polynucléose neutrophile et d'une thrombocytose. Sur le plan clinique, on observe un syndrome d'hyperviscosité sanguine, une splénomégalie, une érythrose faciale. L'évolution est marquée par un risque accru de complications thrombo-emboliques, de transformation myélofibrotique ou en leucémie aiguë secondaire. Le diagnostic positif repose sur des critères majeurs et des critères mineurs.

Le premier critère majeur, indispensable, permet d'objectiver une augmentation de la masse érythrocytaire. Cette augmentation est démontrable soit par : (i) un volume globulaire total > 125% de la normale, ou (ii) un taux d'hémoglobine > 18,5 g/dL chez l'homme ou > 16,5 g/dL chez la femme ou (iii) un hématocrite supérieur au 99ème percentile des valeurs théoriques par l'âge, le sexe et l'altitude, ou (iv) un taux d'hémoglobine > 17 g/dL chez l'homme ou 15 g/dL chez la femme avec élévation de 2 g/dL, sans traitement martial.

Le deuxième critère majeur permet d'objectiver le caractère primitif de la maladie. Il repose sur la démonstration d'une clonalité soit par (i) la mutation *JAK2*^{V617F} dans 95 à 97% des cas, soit par (ii) une mutation située dans l'exon 12 de *JAK2* dans 1 à 2% des cas. Les autres mutations sont plus rares.

Les critères mineurs sont : (i) une hyperplasie myéloïde globale sur la biopsie ostéomédullaire ; (ii) un taux d'EPO sérique inférieur à la normale ; (iii) la pousse spontanée des précurseurs érythroblastiques *in vitro*.

Le diagnostic est posé lorsque les 2 critères majeurs sont réunis, ainsi qu'un critère mineur, ou que l'on a le premier critère majeur et 2 critères mineurs.[25]

PARTIE 2 : SYNDROMES MYELODYSPLASIQUES.

Il s'agit d'un ensemble hétérogène de maladies clonales de la CSH, caractérisées par des cytopénies périphériques, la dysplasie d'une ou plusieurs lignée(s) myéloïde(s), une hématopoïèse inefficace et un risque augmenté d'évolution vers la LAM.[26] Il existe une grande hétérogénéité au sein de cet ensemble, qui regroupe des pathologies indolentes avec une longue espérance de vie, et dont la prise en charge des cytopénies est l'enjeu essentiel, et des pathologies beaucoup plus graves, évoluant rapidement vers l'insuffisance médullaire ou la transformation en LAM. Les SMD représentent 22 % des hémopathies myéloïdes, soit 6,7 % des hémopathies malignes. Leur incidence annuelle en Côte d'Or est de 1,35 nouveau cas pour 100 000 habitants. C'est une maladie de l'adulte âgé, avec une augmentation marquée de l'incidence chez les patients de plus de 65 ans, atteignant 54 nouveaux cas pour 100 000 habitants et par an chez les patients âgés de plus de 84 ans. La médiane de survie des SMD est de 33,8 mois.[1]

Les patients se présentent avec des cytopénies sanguines, définies ainsi : anémie si hémoglobine < 10 g/dL, neutropénie si le nombre de polynucléaires neutrophiles (PNN) est inférieur à 1,8 x 10^9/L et thrombopénie si le compte plaquettaire est inférieur à 100 x 10^9/L [27] L'anémie est beaucoup plus fréquente que la neutropénie ou la thrombopénie. La moelle est hypercellulaire dans 90% des cas, contrastant avec la ou les cytopénie(s) sanguine(s). La myélofibrose est retrouvée dans 10% des SMD. L'examen par microscopie optique des cellules de la moelle révèle la dysplasie d'une ou plusieurs lignée(s) myéloïde(s) (**Tableau III**). Le diagnostic de SMD ne doit être évoqué qu'en l'absence de causes non clonales de dysmyélopoïèse (**Tableau IV**).

Dysérythropoïèse	Dysgranulopoïèse	Dysmégacaryopoïèse
Bourgeonnements nucléaires Ponts internucléaires Caryorrhexis Mutinucléarité Hyperlobulation nucléaire Aspect mégaloblastique Sidéroblastes en couronne Vacuoles cytoplasmiques	Anomalies de taille Hypolobulation nucléaire (pseudo Pelger-Huët) Hypersegmentation irrégulière Hypogranularité, agranularité Granules de pseudo Chediak-Higashi Bâtonnets d'Auer	Micromégacaryocytes Hypolobulation nucléaire Mégacaryocytes multinuclées

Tableau III. Caractéristiques des dysplasies.

Nutritionnelles	Toxiques	Médicamenteuses	Congénitales	Infectieuses
Carence en vitamine B9 et/ou B12	Métaux lourds Arsenic	Cotrimoxazole® Cellcept® Chimiothérapie G-CSF	Dyserytropoïèse congénitale	Parvovirus B19

Tableau IV: Causes non clonales de dysmyélopoïèse.
G-CSF : *Granulocyte-Colony Stimulating Factor*

L'étude cytogénétique permet de confirmer le diagnostic par la mise en évidence d'une ou plusieurs anomalie(s) clonale(s). On retrouve au moins une anomalie génétique clonale dans environ 50% des cas de SMD (**Tableau V**). Les caryotypes comprenant au moins 3 anomalies sont considérés comme des caryotypes complexes. La délétion 5q et l'inversion du chromosome 3 (inv(3)(q21q26.2)) sont associées à des anomalies de la lignée mégacaryocytaire. La ou les anomalies clonales caractéristiques du clone malin ont aussi une valeur pronostique.[27, 28]

Anomalie chromosomique	Fréquence
Trisomie 8*	10%
Délétion 7q	10%
Délétion 5q	10%
Délétion 20q*	5-8%
Perte du chromosome Y*	5%
Inversion 17q	3-5%
Délétion 13q	3%
Délétion 11q	3%
Inv (3)(q21q26.2)	1%

Tableau V. Anomalies cytogénétiques les plus fréquentes dans les SMD.

En fonction de toutes ces données biologiques, le SMD sera classé selon plusieurs critères (**Tableau VI**).[29]

Pathologie	Exploration sanguine	Exploration médullaire
Cytopénies réfractaires avec dysplasie unilignée : Anémie réfractaire (AR) Neutropénie réfractaire Thrombopénie réfractaire	Cytopénie touchant 1 ou 2 lignée(s) < 1% de blastes	Dysplasie d'une seule lignée < 5% de blastes < 15% de sidéroblastes en couronne parmi les précurseurs érythroïdes
Anémie réfractaire avec sidéroblastes en couronne (ARS)	Anémie Absence de blastes	Dysplasie touchant exclusivement la lignée érythroïde > 15% de sidéroblastes en couronne parmi les précurseurs érythroïdes < 5% de blastes
Cytopénie réfractaire avec dysplasie multilignée (CRDM)	Cytopénies < 1% de blastes Pas de bâtonnet d'Auer Monocytes < 1 x 10^9/L	Dysplasie touchant au moins 2 lignées myéloïdes < 5% de blastes Pas de bâtonnet d'Auer +/- 15% de sidéroblastes en couronne
Anémie réfractaire avec excès de blastes (AREB) I	Cytopénies < 5% de blastes Pas de bâtonnet d'Auer Monocytes < 1 x 10^9/L	Dysplasie uni- ou multilignée 5 à 9% de blastes Pas de bâtonnet d'Auer
Anémie réfractaire avec excès de blastes II	Cytopénies 5 à 19% de blastes +/- bâtonnets d'Auer Monocytes < 1 x 10^9/l	Dysplasie uni- ou multilignée 10 à 19% de blastes +/- bâtonnets d'Auer
Syndrome myélodysplasiques inclassables	Cytopénies < 1% de blastes	Dysplasie(s) touchant moins de 10% des cellules dans 1 ou plusieurs lignée(s) myéloïde(s) mais accompagnée(s) d'une anomalie cytogénétique présomptive d'un SMD < 5% de blastes
Syndrome myélodysplasique associé à une délétion 5q isolée	Anémie Compte plaquettaire normal ou augmenté < 1% de blastes	Mégacaryocytes nombreux, avec un noyau hypolobé < 5% de blastes Délétion 5q isolée Absence de bâtonnets d'Auer

Tableau VI. Caractéristiques cytologiques des syndromes myélodysplasiques. D'après Brunning *et al*.[29]. On conclut à la dysplasie d'une lignée si au moins 10% des cellules de cette lignée présentent des signes de dysplasie.

Le pronostic est évalué selon le score IPSS (*International Prognostic Scoring System*) qui définit 4 groupes pronostiques (**Tableau VII**) : un groupe de faible risque (score IPSS = 0) dont la médiane de survie est de 5,7 ans, un groupe de risque intermédiaire 1 (score IPSS compris entre 0,5 et 1) dont la médiane de survie est de 3,5 ans, un groupe de risque intermédiaire 2 (score IPSS compris entre 1,5 et 2) dont la médiane de survie est de 1,2 an, enfin, un groupe de haut risque (score IPSS supérieur ou égal à 2,5) dont la médiane de survie est de 0,4 an.[27]

Score	**0**	**0,5**	**1**	**1,5**
Pourcentage de blastes médullaires	< 5 %	5 à 10 %		11 à 19 %
Caryotype	Favorable	Intermédiaire	Défavorable	
Cytopénie(s)	0 à 1	2 à 3		
Caryotype Favorable : normal, délétion du chromosome Y, délétion 5q, délétion 20q Défavorable : caryotypes complexes (au moins 3 anomalies), anomalies du chromosome 7 Intermédiaire : autres anomalies				
Cytopénies Anémie : hémoglobine < 10 g/dL Neutropénie : polynucléaires neutrophiles < 1,8 x 10^9/L Thrombopénie : plaquettes < 100 x 10^9/L				

Tableau VII. Score pronostique IPSS.
D'après Greenberg *et al*.[27]

Un score IPSS révisé a été proposé en 2012. Ce score repose sur les mêmes critères, mais permet un affinage de l'évaluation pronostique en définissant 5 groupes de risque.[28]

PARTIE 3 : SYNDROMES MIXTES MYELODYSPLASIQUES/ MYELOPROLIFERATIFS

Il s'agit de maladies clonales de la CSH qui, au moment du diagnostic, présentent à la fois des caractéristiques des syndromes myéloprolifératifs et des caractéristiques des syndromes myélodysplasiques (**Tableau VIII ; Figures 3 et 4**).[30, 31] Ils représentent 9,8% des hémopathies myéloïdes, soit environ 3% des hémopathies malignes. Leur incidence est de 0,6 nouveau cas pour 100 000 habitants et par an et leur médiane de survie est de 26,6 mois.[1] Les SMP/SMD regroupent la leucémie myélo-monocytaire chronique (LMMC), qui représente 96% des SMP/SMD, la leucémie myéloïde chronique atypique *BCR-ABL1* négative, la leucémie myélo-monocytaire juvénile (LMMJ), les SMP/SMD inclassables et une entité provisoire : l'ARS-T.[31] Ils sont classiquement caractérisés par une richesse médullaire augmentée suite à la prolifération d'une ou plusieurs lignées myéloïdes et par une augmentation du nombre de cellules matures issues de cette lignée dans le sang. Parallèlement, une ou plusieurs autres lignée(s) myéloïde(s) montre(nt) des signes de dysplasie ainsi qu'une prolifération inefficace, ce qui entraîne un déficit en cellules matures de ces autres lignées dans le sang. Par définition, le pourcentage de blastes médullaires reste inférieur à 20%.

Syndromes mixtes myéloprolifératifs / myélodysplasiques
Leucémie myélo-monocytaire chronique
Leucémie myéloïde chronique atypique *BCR-ABL1* négative
Leucémie myélo-monocytaire juvénile
Syndromes myéloprolifératifs/myélodysplasiques inclassables
Entité provisoire : anémie réfractaire avec sidéroblastes en couronne et thrombocytose

Tableau VIII. Syndromes mixtes myéloprolifératifs / myélodysplasiques. D'après Vardiman *et al*.[31]

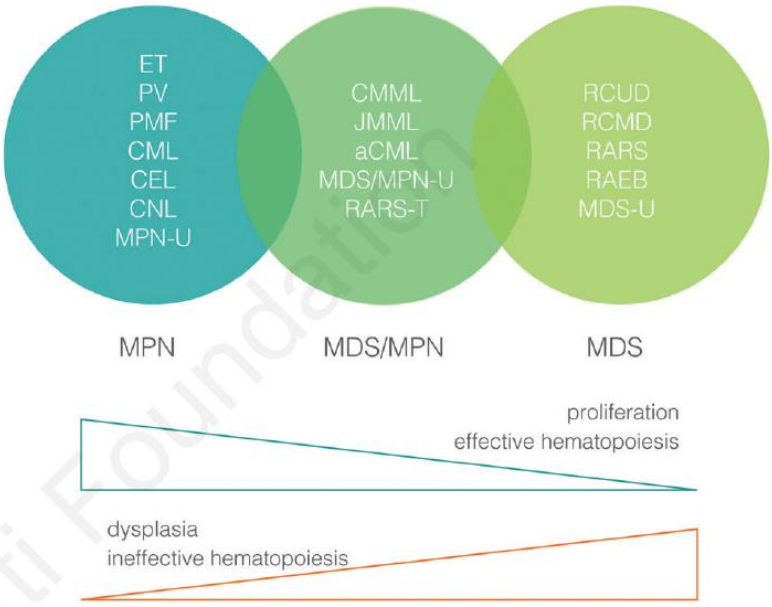


Figure 3. Syndromes myéloprolifératifs, syndromes myélodysplasiques et syndromes mixtes myélodysplasiques/myéloprolifératifs.
D'après Mughal *et al.*[30]

MPN : Myeloproliferative Neoplasms ; ET : Essential Thrombocythemia ; PV : Polycythemia Vera ; PMF : Primary MyeloFibrosis ; CML : Chronic Myeloid Leukemia ; CEL : Chronic Eosinophilic Leukemia ; CNL : Chronic Neutrophilic Leukemia ; MPN-U : Myeloproliferative Neoplasms Unclassified. MDS/MPN : MyeloDysplastic Syndroms/Myeloproliferative Neoplasms ; CMML : Chronic Myelo-Monocytic Leukemia ; JMML : Juvenile Myelo-Monocytic Leukemia ; aCML : atypic Chronic Myeloid Leukemia ; MDS/MPN-U : MyeloDysplastic Syndroms/Myeloproliferative Neoplasms Unclassified ; RARS-T : Refractory Anemia with Ring Sideroblasts and marked Thrombocytosis ; MDS : MyeloDysplastic Syndrom ; RCUD : Refractory Cytopenia with Unilineage Dysplasia ; RCMD : Refractory Cytopenia with Multilineage Dysplasia ; RARS : Refractory Anemia with Ring Sideroblasts ; RAEB : Refractory Anemia with Excess Blasts ; MDS-U : MyeloDysplastic Syndroms-Unclassified.

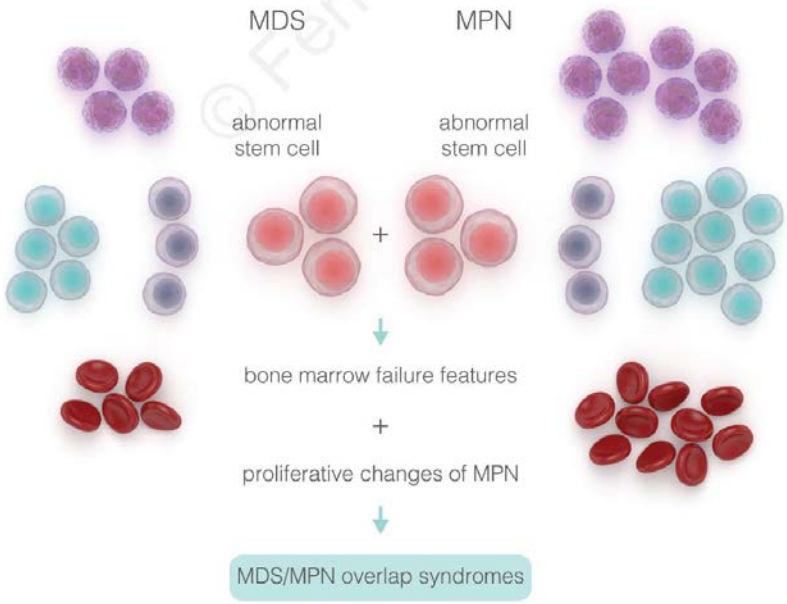


Figure 4. Physiopathologie des SMP/SMD.
D'après Mughal *et al.*[30]

LES SYNDROMES MYELOPROLIFERATIFS/MYELODYSPLASIQUES INCLASSABLES

Ces syndromes présentent au diagnostic à la fois des caractéristiques propres aux syndromes myéloprolifératifs et des caractéristiques propres aux syndromes myélodysplasiques, sans pour autant présenter les critères diagnostiques permettant de les classer parmi les SMP/SMD décrits précédemment (LMMC, LMMJ, LMC atypique) (**Tableau IX**). La découverte du gène de fusion *BCR-ABL1*, de réarrangements géniques impliquant *PDGFRA (Platelet-Derived Growth Factor Receptor Alpha polypeptide)*, *PDGFRB (Platelet-Derived Growth Factor Receptor Beta polypeptide)* ou *FGFR1 (Fibroblast Growth Factor Receptor 1*), d'une délétion 5q isolée ou de l'inversion (3)(q21;q26.2) excluent le diagnostic de SMP/SMD inclassable. Les patients ayant d'abord présenté un SMP bien défini puis développé des myélodysplasies sont exclus. Doivent aussi être exclus de cette entité les SMP/SMD dont l'origine médicamenteuse n'a pu être écartée (médicaments cytotoxiques, facteurs de croissance).[31]

Critères diagnostiques pour les SMP/SMD inclassables
La pathologie a des caractéristiques cliniques, biologiques et morphologiques d'une des catégories des SMD et on compte moins de 20% de blastes médullaires.
ET
On observe des éléments caractéristiques des syndromes myéloprolifératifs comme :
- un compte plaquettaire > 450 x 10^9/L avec prolifération mégacaryocytaire ou
- une leucocytose > 13 x 10^9/L avec ou sans splénomégalie.
ET
Absence d'antécédent de SMP ou de SMD.
Pas d'intoxication récente ou de traitement pouvant expliquer les caractéristiques myéloprolifératives ou myélodysplasiques.
Pas de chromosome Philadelphie ou de gène de fusion *BCR-ABL1*.
Pas réarrangement de *PDGFRA*, *PDGFRB* ou de *FGFR1*.
Pas de délétion 5q isolée ou d'inversion (3)(q21;q26).
OU
Maladie *de novo* avec des caractéristiques myéloprolifératives et myélodysplasiques, classée ni dans les SMP, ni dans les SMD, ni dans une autre catégorie de SMP/SMD**.**

Tableau IX. Critères diagnostiques pour les SMP/SMD inclassables. D'après Vardiman *et al*.[31].

L'un des sous-groupes les plus controversés est l'entité provisoire appelée Anémie Réfractaire avec Sidéroblastes en couronne et Thrombocytose ou ARS-T.

L'ANEMIE REFRACTAIRE AVEC SIDEROBLASTES EN COURONNE ET THROMBOCYTOSE

Dans la classification OMS 2001 des hémopathies malignes, l'ARS-T (appelée à l'époque thrombocytémie essentielle avec sidéroblastes en couronne) a été proposée comme entité provisoire pour regrouper tous les patients présentant des caractéristiques cliniques et cytologiques des SMD, mais aussi une thrombocytose marquée associée à des anomalies mégacaryocytaires similaires à celles observées dans la TE.[32] En 2001, l'ARS-T était définie selon les critères suivants : (i) plaquettes > 600 x 10^9/L, (ii) au moins 15% de sidéroblastes en couronne (SC), (iii) < 5% de blastes médullaires, (iv) prolifération de mégacaryocytes avec une morphologie caractéristique, (v) exclusion de la mutation *BCR-ABL1*, (vi) exclusion de l'inv (3)(q21;q26.2) et de la délétion 5q.

En 2005, Shaw décrit l'entité "sidéroblastes en couronne et thrombocytose", correspondant aux critères suivants : (i) thrombocytose > 500 x 10^9/L durant au moins 6 mois, (ii) ≥ 15% de sidéroblastes en couronne, (iii) pas d'anomalie cytogénétique, (iv) élimination des autres causes de thrombocytose, (v) élimination des causes réversibles d'apparition des sidéroblastes en couronne, (vi) moins de 3% de blastes dans la moelle, (vii) moins de 1% de blastes sanguins.[33]

Dans la classification OMS 2008 des hémopathies malignes, les critères diagnostiques de l'ARS-T sont modifiés : (i) anémie (hémoglobine < 12,5 g/dl chez la femme et < 13,5 g/dl chez l'homme), (ii) < 5% de cellules blastiques médullaires, (iii) dysérythropoïèse touchant au moins 10% des cellules érythroblastiques ; (iv) ≥ 15% des érythroblastes présentent l'aspect de SC ; (v) thrombocytose > 450 x 10^9/L ; (vi) prolifération de grands mégacaryocytes atypiques de morphologie similaire à ceux retrouvés dans les SMP *BCR-ABL1* négatifs ; (vii) exclusion de la mutation *BCR-ABL1* ; (viii) exclusion de l'inv (3)(q21;q26.2) et de la délétion 5q.[31]

Dans la classification OMS de 2008, l'ARS-T est considérée comme une entité provisoire et la question de son existence en tant qu'entité indépendante n'est pas résolue. Plusieurs hypothèses ont été proposées pour expliquer l'origine de l'ARS-T (**Figure 5**).[34]

(i) L'ARS-T pourrait être la coïncidence chez un même patient de deux pathologies : l'ARS et la TE.

(ii) L'ARS-T pourrait n'être qu'un variant d'une autre maladie : soit une ARS particulière ayant acquis une mutation de *JAK2* ou de *MPL* conduisant à une thrombocytose surajoutée, soit une TE qui verrait sa proportion de SC dépasser le seuil de 15% par l'acquisition d'une mutation d'un ou plusieurs gènes impliqués dans le métabolisme mitochondrial du fer.

(iii) L'ARS-T pourrait être une entité distincte de l'ARS et de la TE, mais empruntant certaines caractéristiques à l'une et l'autre.

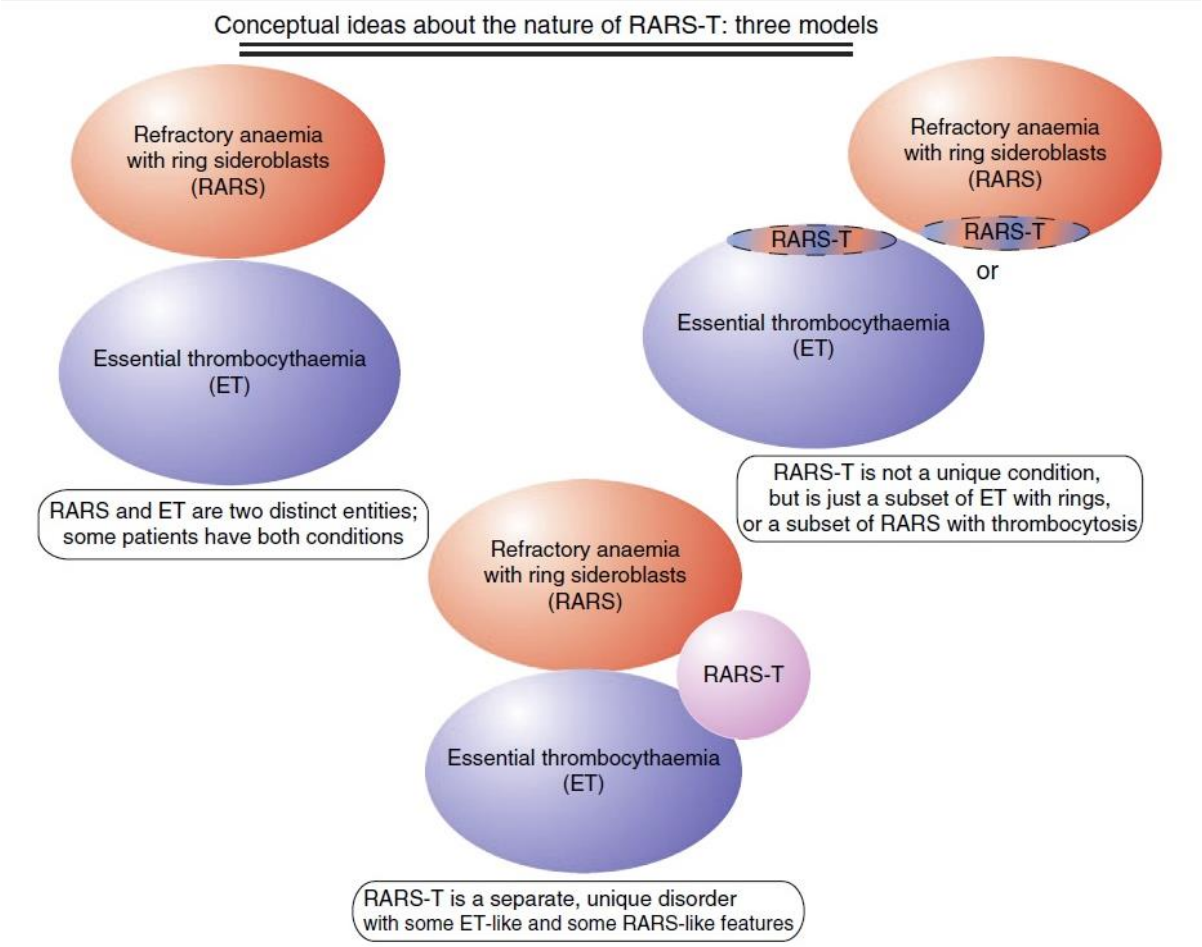


Figure 5. Trois modèles pour expliquer l'origine de l'ARS-T.
D'après Wardrop *et al.*[34]

L'anémie réfractaire avec sidéroblastes en couronne

L'ARS est une maladie indolente, d'évolution lente, présentant un faible risque d'évolution vers une leucémie aiguë.[35] Elle représente 35% des SMD. Son incidence en Côte d'Or est de 0,4 nouveau cas pour 100 000 habitants et par an.[1]

Les patients présentent une anémie normocytaire ou macrocytaire le plus souvent isolée. La cytologie médullaire montre une dysplasie confinée aux précurseurs érythroïdes, dont au moins 15% sont des SC.[36] Les SC sont des précurseurs érythroïdes présentant une accumulation anormale de fer dans leurs mitochondries (au moins 5 granules ferriques encerclant au moins un tiers du noyau) mis en évidence par la coloration de Perls.[37] Les SC ne sont pas spécifiques de

l'ARS. Des pourcentages inférieurs à 15% peuvent être retrouvés dans d'autres SMD,[32, 38] ainsi que dans certains SMP.[13, 39] Il existe des causes non néoplasiques de la présence de sidéroblastes en couronne : alcool, intoxication au plomb, benzène, isoniazide, déficience en cuivre, surcharge en zinc, anémie sidéroblastique congénitale qui doivent être exclues.

Les CSH de patients atteints d'ARS développent des dépôts anormaux de fer à un stade précoce du développement érythroblastique, ce qui suggère une ou plusieurs anomalies génétiques clonales responsable(s) des anomalies du métabolisme du fer dans cette lignée. L'ARS présente une anomalie cytogénétique dans 5 à 20% des cas, impliquant classiquement un seul chromosome. L'évolution en LAM n'est observée que dans 1 à 2% des ARS. La médiane de survie globale varie de 69 à 108 mois selon les études.[40]

L'ARS-T a été proposée comme entité provisoire dans la classification OMS de 2001, puis reconduite comme entité provisoire dans celle de 2008 dans la mesure où il n'a pas été clairement déterminé s'il s'agit d'une pathologie indépendante, s'il s'agit d'une TE ayant acquis une ou plusieurs mutations géniques entraînant l'apparition d'un pourcentage de SC supérieur à 15% ou bien s'il s'agit d'une ARS s'accompagnant d'une thrombocytose en raison de l'acquisition d'une mutation de *JAK2* ou de *MPL*.

Nous avons réuni une cohorte de 200 patients présentant une ARS-T. L'étude de cette cohorte nous a permis de montrer que l'ARS-T se distingue de la TE et de l'ARS du point de vue clinique, biologique, pronostique et moléculaire. La survie sans leucémie de l'ARS-T est inférieure à celle de la TE. Enfin, la survie globale de l'ARS-T est intermédiaire entre celle de l'ARS et celle de la TE.

Ainsi il apparaît clairement que l'ARS-T partage un bon nombre de caractéristiques de l'ARS, mais avec des critères prolifératifs (thrombocytose en particulier). De plus, l'existence de 6 cas bien documentés d'ARS ayant évolué vers une ARS-T de manière concomitante à l'acquisition d'une mutation de *JAK2* ou plus rarement de *MPL*, plaide en faveur de l'hypothèse d'une transformation d'une ARS préexistante.[41]

A partir de 2011, de nouvelles mutations géniques associées aux SMD ont été découvertes. L'association entre mutations somatiques des acteurs-clé de l'épissage alternatif et SMD a ainsi été mise en évidence.[42]

L'épissage alternatif est un mécanisme de régulation transcriptionnelle partiellement connu, qui permet de garantir une grande diversité de protéines fonctionnelles malgré un nombre limité de gènes. Ce mécanisme repose sur un complexe de snRNP (*small nuclear RiboNucleoProtein*) et de leurs nombreuses protéines associées qui forment le "spliceosome". Chaque snRNP est composé d'un snRNA (*small nuclear RiboNucleic Acid*) et de protéines associées et a une fonction spécifique. Ce mécanisme permet la maturation du pré-ARNm en ARNm mature et fonctionnel et contrôle la diversité des variants d'épissage. La qualité de la protéine finale est donc hautement dépendante de la qualité de l'épissage du pré-ARNm et les mutations géniques affectant les acteurs de l'épissage alternatif peuvent mener à la synthèse de protéines anormales et pathogènes.

Les mutations récurrentes récemment découvertes chez les patients atteints de SMD affectent les gènes codant pour des protéines agissant lors de la phase initiale de l'épissage, comme la reconnaissance du site 3' accepteur de l'intron du pré-ARNm. Il s'agit des protéines U2AF35 (*U2 small nuclear RNA Auxiliary Factor 35*), SRSF2 (*Serine/arginine-rich Splicing Factor 2*)), *ZRSR2* (*Zinc finger CCCH type, RNA-binding motif and Serine/arginine rich 2*). SF3B1 (*Splicing Factor 3B subunit 1*) appartient à U2snRNP, l'un des 2 complexes formant le spliceosome chez les eucaryotes, et a un rôle important à la fois dans l'assemblage des différents acteurs de l'épissage, mais aussi dans l'activité catalytique d'U2snRNP. U2snRNP reconnaît le site d'épissage en 3', au niveau des jonctions intron-exon (**Figure 6**).

Ces mutations sont, la plupart du temps, mutuellement exclusives et hautement spécifiques des SMD avec (84,9%) ou sans sidéroblastes en couronne (43,9%). On les retrouve aussi dans la LMMC (54,5%), les LAM post-SMD (25,8%) et les SMP (9,4%). En revanche elles ne sont pas retrouvées dans les LAM *de novo*. La forte association de ces mutations avec les SMD en fait un élément caractéristique de cette pathologie et indique que ces mutations jouent un rôle central dans leur physiopathologie.[42-46]

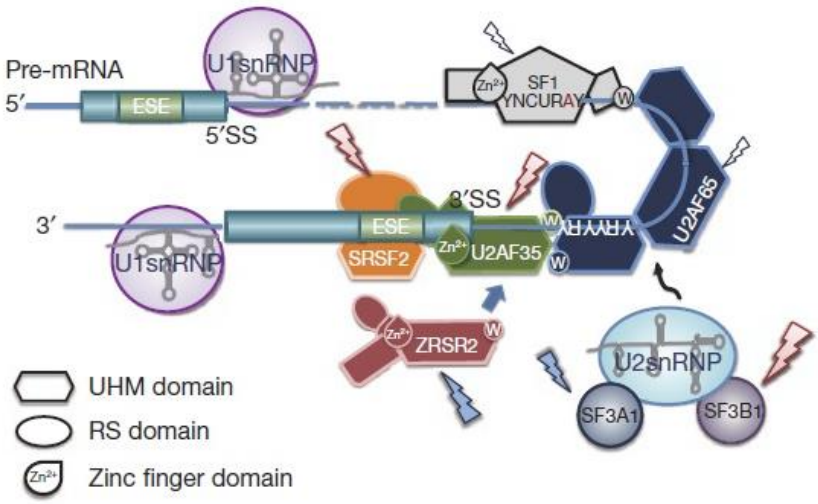


Figure 6. Les acteurs de l'épissage alternatif.
D'après Yoshida *et al*.[42])

Des mutations somatiques récurrentes de *SF3B1* ont été identifiées dans environ 20% des SMD. La fréquence de ces mutations se situe entre 65 et 85% dans les SMD caractérisés par la présence de plus de 15% de SC, en particulier les ARS, les ARS-T. Le domaine N-terminal de la protéine SF3B1 est impliqué dans les interactions protéine-ARN et protéine-protéine, alors que le domaine C-terminal de la protéine se compose de 22 domaines répétés HEAT (*Huntingtin Elongation factor 3, protein phosphatase 2A, TOR1*). Les mutations de *SF3B1* qui ont été décrites sont des substitutions hétérozygotes situées au niveau des exons 12 à 16 du gène et affectant les premiers domaines HEAT de la protéine, en particulier les domaines 3 à 6. Les mutations de *SF3B1* les plus fréquentes modifient respectivement le résidu K700 (57,7% des mutations de *SF3B1*), les résidus E622 (7,1%), R625 (6%), H662 (10,1%), K666 (10,7%) et I704. La mutation la plus fréquente, K700E, est retrouvée dans 50% des ARS et 44% des ARS-T (**Figure 7**).[42, 43, 47-49]

Les mutations somatiques de *SF3B1* répertoriées jusqu'à présent ne sont que des substitutions : on ne retrouve pas de mutation non-sens, de mutations des sites d'épissage, ou de mutations responsables du décalage du cadre de lecture, ce qui suggère que ces mutations altèrent la fonction de la protéine tout en épargnant l'intégrité de sa structure globale. Epissage et transcription sont deux mécanismes contemporains et interdépendants. On peut donc penser que les mutations de *SF3B1* peuvent influencer non seulement le mécanisme d'épissage, mais aussi les différentes étapes de la transcription.

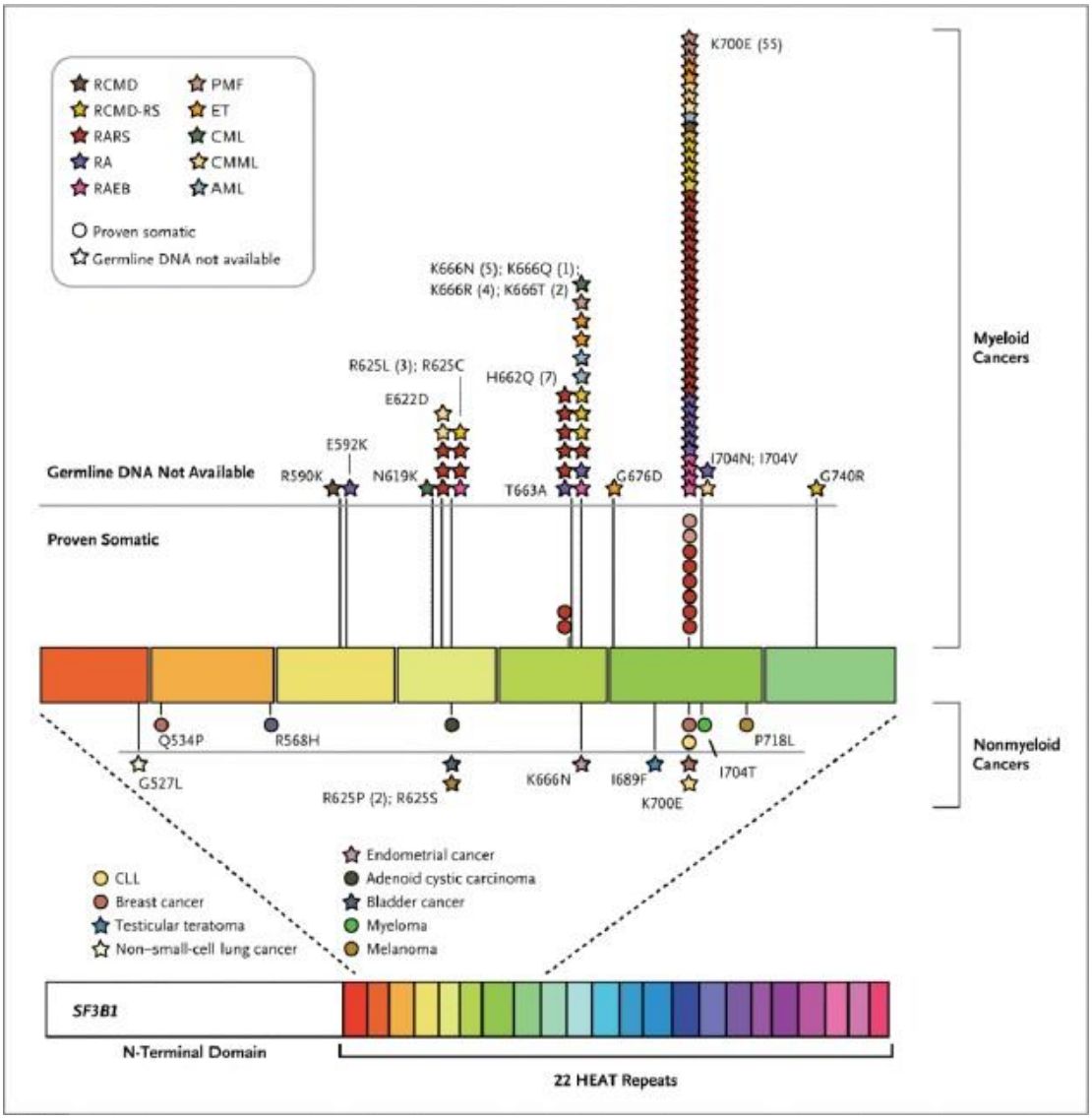


Figure 7. Distribution des mutations de *SF3B1*.
D'après Papaemmanuil *et al*.[43]

AML : Acute Myeloid Leukemia, CLL : Chronic Lymphocytic Leukemia, CML : Chronic Myeloid Leukemia, CMML : Chronic Myelo-Monocytic Leukemia, ET : Essential Thrombocythemia, PMF : Primary MyeloFibrosis,, RA : Refractory Anemia, RARS : Refractory Anemia with Ring Sideroblasrs, RAEB : Refractory Anemia with Excess Blasts, RCMD : Refractory Cytopenia with Multilineage Dysplasia, RCMD-RS : RCMD with Ring Sideroblasts.

L'analyse du profil d'expression génique des cellules CD34+ de patients atteints de SMD et présentant une mutation de *SF3B1* révèle un groupe de 94 gènes dont le niveau d'expression est significativement diminué. Parmi les 50 gènes dont le niveau d'expression est le plus fortement diminué en raison des mutations de *SF3B1*, 7 sont déterminants pour l'exercice des fonctions mitochondriales, ce qui suggère que les mutations de *SF3B1* sont associées à la sous-expression de gènes-clés du fonctionnement mitochondrial et du métabolisme du fer. Le mécanisme exact

par lequel les perturbations du mécanisme d'épissage provoquées par les mutations de *SF3B1* jouent sur la transcription des ARNm et la pathogénie cancéreuse demeure inconnu. Les mutations de *U2AF35*, un autre membre du complexe d'épissage, provoquent, dans les lignées cellulaires où elles ont été introduites, non pas une expansion clonale mais une diminution de la prolifération cellulaire, une augmentation de l'apoptose et des anomalies de la différenciation, ce qui est cohérent avec les caractéristiques cytologiques des SMD, à savoir cytopénies, dysplasies et apoptose accrue. Dans ce modèle, la protéine modifiée agit comme un dominant négatif de la protéine sauvage.[42] Le résultat au niveau moléculaire est une anomalie globale des transcrits cellulaires avec une production augmentée de transcrits matures possédant encore des séquences introniques non excisées.

Le lien entre mutation de *SF3B1* et SC a été confirmé par un travail expérimental récent. *In vitro*, la diminution de 50% du niveau d'expression de SF3B1 dans des cellules K562 entraîne des anomalies qualitatives de l'épissage qui affectent entre autres les transcrits des gènes *ASXL1*, *CBL*, *EZH1* et *RUNX3*, sans modification de leur niveau d'expression. L'inhibition pharmacologique des protéines de la famille SF3B dans des cellules issues de la moelle osseuse de patients sains entraîne l'apparition de SC. *In vivo*, les précurseurs érythroïdes de souris haplo-insuffisantes pour *Sf3b1* présentent un excès de SC par rapport au groupe témoin. Enfin, l'analyse différentielle du transcriptome entre des patients atteints d'ARS et mutés pour *SF3B1* et un groupe de donneurs sains montre une différence d'épissage des transcrits portant sur 423 exons issus de 350 gènes. Pour 56 des gènes testés, la différence porte sur plus d'un exon, et parmi ces gènes, on retrouve *ASXL1*, *CBL*, *EZH1* et *RUNX3* dont le rôle dans la physiopathologie des SMD est connu.[50]

Les mutations de *SF3B1* sont retrouvées dans 20% de l'ensemble des SMD. Elles sont particulièrement fréquentes (64 à 82,6%) dans les SMD avec SC, en particulier chez 66,7 à 72% des patients ARS-T. Par contre, ces mutations de *SF3B1* sont retrouvées à une fréquence beaucoup plus faible dans les autres SMD : 10 à 11,5% des AR, environ 6% des CRDM et 5 à 8,4% des AREB.[42, 46, 47] Les mutations de *SF3B1* sont retrouvées dans d'autres pathologies myéloïdes : environ 5% des LAM, 4% des MFP, 3% des TE, 4 à 6,5% des LMMC, ainsi que dans 5% des leucémies lymphoïdes chroniques (LLC), 3% des myélomes multiples et dans certains cancers solides (**Tableau X**). La fréquence des mutations de *SF3B1* est très différente entre l'ARS-T et la TE. L'ARS-T se distingue donc de la TE par un profil génique différent, ce qui confirme que l'ARS-T est une entité indépendante de la TE. Au sein des SMP/SMD, on note une incidence de *SF3B1*^{mut} très différente entre ARS-T et LMMC. En revanche, les autres gènes codant pour des protéines du spliceosome (*SRSF2*, *USAF35*, *ZRSR2*, *SF3A1*) sont fréquemment

mutés (50% des cas), ce qui indique que les anomalies du complexe d'épissage jouent un rôle majeur dans la physiopathologie de ces maladies.

Ainsi les mutations de *SF3B1* sont hautement spécifiques des pathologies myéloïdes avec SC. Parmi les patients porteurs d'une mutation de *SF3B1*, 90% ont plus de 15% de SC, 7% des patients avaient entre 1 et 14% de SC et l'absence de SC n'a été constatée que chez 3% de ces patients. La valeur prédictive positive du statut *SF3B1*^{mut} pour un diagnostic de SMD caractérisé par la présence de SC, est de 97,7%, alors que la valeur prédictive négative est de 97,8%.[46]

Proportions de mutations de *SF3B1* (%)				
Type tumoral	**Papaemmanuil *et al.***	**Yoshida *et al.***	**Visconte *et al.***	**Malcovati *et al.***
SMD	20,3	20	14,7	28,1
AR	10			11,5
ARS	68	82,6	64	79
CRDM	6			6,3
CRDM avec SC	57	76		57,7
AREB	6			19,7
SMP	**3**	**0**		
PV	0			
TE	3			
MFP	4,4			
LMC	0			
SMP/SMD				**19,3**
LMMC	5	4		6,5
ARS-T			72	66,7
Autres				
LAM	5,3	3,2		5,3
LLC	5			
Myélome multiple	3			
Cancers solides	4			
Cancer du rein	3			

Tableau X. Fréquence des mutations de *SF3B1* dans les pathologies cancéreuses.

La charge allélique médiane chez les patients mutés pour *SF3B1* est d'environ 39,5%, ce qui indique la présence d'un clone dominant hétérozygote pour la mutation. Cette charge allélique est équivalente, que l'on prenne seulement les précurseurs hématopoïétiques CD34+, les cellules granuleuses, ou l'ensemble des cellules médullaires. Cela indique que la mutation naît dans les précurseurs hématopoïétiques, est transmise aux progéniteurs puis aux cellules matures. Dans environ 20% des cas, la faible charge allélique en *SF3B1*^{mut} est plutôt cohérente avec la présence d'un clone minoritaire, ce qui suggère que la transformation myélodysplasique dépend de la mutation d'un autre gène et que la mutation de *SF3B1* est un événement secondaire ici. Aucune corrélation n'a été démontrée entre la charge allélique en *SF3B1*^{mut} et les caractéristiques cliniques, biologiques et pronostiques des patients, en particulier chez les patients ARS-T. En revanche, il y a une corrélation entre charge allélique en *SF3B1*^{mut} et le pourcentage de SC. Enfin, on retrouve beaucoup plus de patients atteints d'un SMD avec SC parmi les patients dont la charge allélique en *SF3B1* est supérieure à 25%. Chez les patients ARS-T, la charge allélique varie de 17 à 62%, soit une médiane de 41%, ce qui est similaire à ce qui est retrouvé dans les autres SMD caractérisés par la présence de SC.[46]

Parmi les patients atteints de SMD, ceux qui arborent une mutation de *SF3B1* ont au diagnostic un plus grand nombre de globules blancs et de plaquettes, une hyperplasie érythroblastique plus marquée (ce qui indique qu'il y a une relation entre mutation de *SF3B1* et érythropoïèse hyperplasique mais inefficace) avec un plus grand nombre de SC, ainsi qu'une plus faible proportion de blastes médullaires.[43, 46] Sur 323 patients atteints de SMD, ceux qui présentent une mutation de *SF3B1* ont une meilleure survie globale et une meilleure survie sans leucémie que les patients *SF3B1*^{WT}, et ce en analyse univariée et en analyse multivariée incluant l'âge, le sexe, le taux d'hémoglobine, le nombre de PNN, le nombre de plaquettes, le risque cytogénétique, le nombre de blastes médullaires et le pourcentage de SC. Au total, la survie sans événements des patients *SF3B1*^{mut} est supérieure à celle des patients *SF3B1*^{WT}, ce qui suggère que cette mutation est associée à des SMD de bon pronostic. En revanche, si l'on considère uniquement les ARS, les patients *SF3B1*^{mut} ont, en analyse univariée, une meilleure survie globale mais une survie sans leucémie équivalente. Cet effet protecteur n'est pas retrouvé en analyse multivariée. Le statut mutationnel de *SF3B1* n'a aucun impact sur les données cliniques et biologiques des ARS-T. Sur une série de 107 SMD avec SC (48 ARS, 43 CRDM avec SC, 11 AREB 1 avec SC et 7 AREB 2 avec SC), le statut *SF3B1*^{mut} est associé à une meilleure survie globale et à une meilleure survie sans leucémie en analyse univariée. Cette association est perdue en analyse multivariée car le statut *SF3B1*^{mut} est fortement associé à des sous-groupes de SMD de bon pronostic.[46]

En d'autres termes, si l'on considère spécifiquement les ARS, les ARS-T et les CRDM avec SC, le statut mutationnel *SF3B1* n'apporte aucune valeur pronostique.[51] Sur une série de 15 SMD, le risque thrombotique artériel et veineux est accru chez les patients porteurs d'une mutation de *SF3B1* (n = 10) : 60% d'entre eux ont présenté une complication thrombotique alors qu'aucune n'a été signalée chez les patients *SF3B1*^{WT}.[47]

Chaque mutation d'un gène codant pour une protéine appartenant au complexe d'épissage est associée de manière significative à la mutation d'un gène codant pour une protéine impliquée dans la régulation épigénétique. La moitié des patients arborant une mutation de *DNMT3A* portent aussi une mutation de *SF3B1.* On note aussi que 33% des patients atteints d'un SMD présentant la mutation *JAK2*^{V617F} présentent aussi une mutation de *SF3B1*. L'impact pronostique du statut mutationnel de *SF3B1* n'est pas encore clairement établi, et les études récentes donnent des résultats contradictoires.[44, 46, 52] En revanche, les patients ayant un statut *SF3B1*^{wt}/*DNMT3A*^{mut} ont une survie plus faible comparativement aux autres combinaisons. Les mutations de *DNMT3A* ont été décrites chez les patients ARS et elles sont retrouvées chez 42% des patients ARS et 17% des ARS-T. Chez les patients ARS-T, les mutations de *SF3B1* sont associées à la mutation *JAK2*^{V617F} dans 11 à 30% des cas et à des mutations de *MPL* dans 6 à 23% des cas.[46, 47]

Environ 25% des SMD avec SC sont *SF3B1*^{WT}, ce qui suggère que d'autres gènes mutants peuvent être impliqués dans la physiopathologie des SC. D'ailleurs des mutations somatiques de *SRSF2* ou *ZRSR2* ont été décrites dans environ 7% des SMD avec SC *SF3B1*^{WT}.[42] D'autres mutations concernant d'autres acteurs de l'épissage associés à SF3B1 (SF3B4 et SF3B14) pourraient expliquer que *SF3B1* ne soit pas retrouvé systématiquement muté en présence de SC.[44]

PARTIE 4 : SYNDROMES LYMPHOPROLIFERATIFS DE LIGNEE B

1. LYMPHOPOÏESE B ET LYMPHOMAGENESE B.

1.1. La lymphopoïese B

Les lymphocytes B sont les acteurs de l'immunité humorale, c'est-à-dire reposant sur la sécrétion d'anticorps spécifiques de l'antigène. Cette reconnaissance de l'antigène se fait grâce à un récepteur de surface, le BCR (*B-Cell Receptor*), composé d'une immunoglobuline (Ig) de surface (**Figure 8**) couplée à des molécules transmembranaires permettant la transduction du signal (CD79a et CD79b). L'Ig de surface est composée de deux chaînes lourdes identiques et de deux chaînes légères identiques liées de manière covalente. Chacune de ces chaînes se compose d'une partie constante (fonction effectrice, c'est-à-dire permettant l'activation secondaire des effecteurs du système immunitaire, en particulier le système du complément dans le cadre de l'ADCC ou *antibody derived cell-mediated cytotoxicity*) et d'une partie variable, spécifique, qui reconnait l'antigène.

Le BCR est donc un composant essentiel du lymphocyte B et sa maturation constitue le fil directeur de la lymphopoïèse B, c'est-à-dire le processus physiologique de fabrication des lymphocytes B. Celle-ci se compose de deux étapes.

La première étape a lieu dans la moelle osseuse et donne naissance à des lymphocytes B matures (c'est-à-dire possédant à leur surface un BCR fonctionnel) mais naïfs (car n'ayant jamais rencontré d'antigène). L'Ig est déterminée de manière aléatoire, par réarrangement génique. Ce processus permet à l'organisme de disposer en permanence d'un répertoire de l'ordre de 10^{18} cellules B naïves différentes, couvrant en principe l'essentiel des antigènes auquel l'organisme peut être exposé. Ne survivent que les cellules B naïves qui rencontrent un antigène pour lequel elles ont une affinité.

La deuxième étape de la lymphopoïèse B a lieu dans les organes lymphoïdes secondaires (ganglions lymphatiques, rate, tissus lymphoïdes associés aux muqueuses) et repose sur la sélection des lymphocytes B par l'antigène. Si le BCR de la cellule B naïve présente une affinité pour l'antigène rencontré, cette cellule B naïve va survivre et proliférer dans le cadre de la réponse immunitaire humorale.

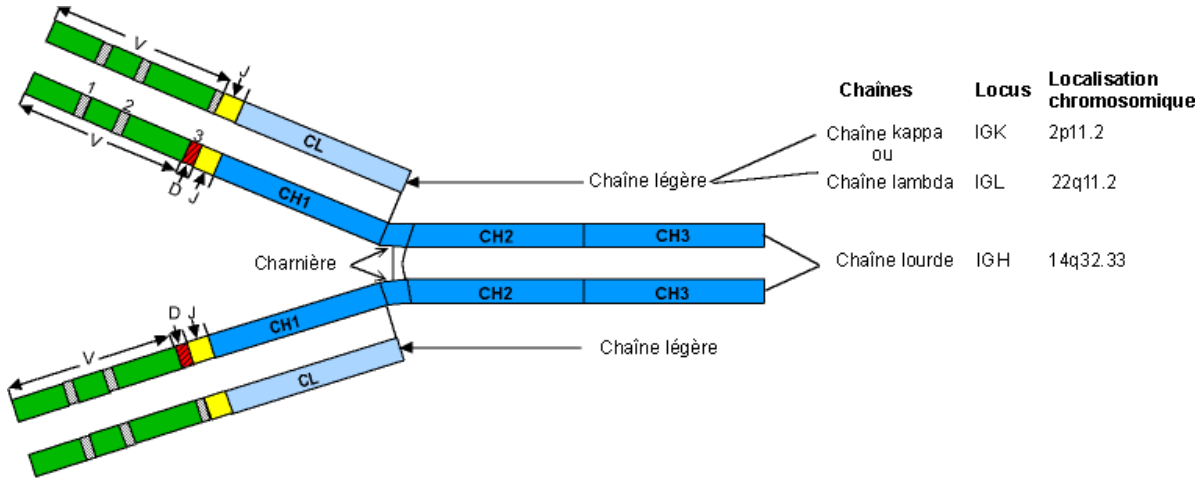


Figure 8. Composition d'une immunoglobuline.

Celle-ci se compose de deux chaines lourdes identiques (bleu foncé) et de deux chaines légères identiques (bleu clair). Les deux chaines lourdes se composent d'une partie constante (CH pour Constant Heavy) et d'une partie légère déterminées par recombinaison VDJ. Les deux chaines légères se composent d'une partie constante (CL pour Constant Light) et d'une partie variable déterminée par recombinaison VJ

1.1.1. Première étape : l'étape médullaire

La finalité de cette première étape est d'obtenir des lymphocytes B exprimant à leur surface un BCR fonctionnel et non auto-réactif. L'assemblage du BCR repose sur la synthèse des chaînes lourdes et des chaînes légères de l'Ig de surface.

La génération des parties variables des chaînes lourdes (IgvH) commence par une succession de recombinaisons géniques aléatoires impliquant trois différents segments géniques du gène codant pour l'IgvH : les segments V, D et J. Cette première étape porte donc le nom de «recombinaison VDJ». L'ADN germinal de ce gène comporte 50 séquences «V», 27 séquences «D» et 6 séquences «J». L'ADN codant pour la chaîne lourde de l'Ig de surface sera un réassemblage d'une séquence V, d'une séquence D et d'une séquence J déterminées de manière aléatoire. On parle de diversité combinatoire. A cette diversité combinatoire s'ajoute une diversité jonctionnelle par ajout ou retrait de nucléotides au niveau des sites de jonction des segments V, D et J. A cette partie variable de la chaîne lourde de l'Ig s'ajoute la partie constante, cμ (**Figure 9**).

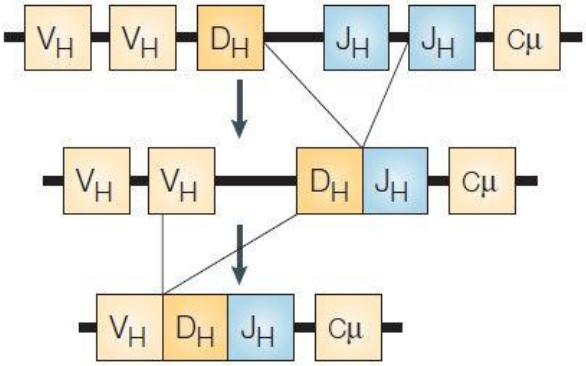


Figure 9. Processus de recombinaison « VDJ ».

L'ADN germinal codant pour l'IgVH comporte plusieurs séquences V, D et J. Une séquence de chaque type est conservée pour les étapes suivantes. Les séquences V, D et J choisies sont assemblées ensuite. A cette partie variable de la chaîne lourde de l'immunoglobuline s'ajoute la partie constante, Cμ, correspondant à la partie constante d'une IgM. D'après Küppers *et al*.[53]

Cette recombinaison du gène codant pour l'IgvH précède la recombinaison impliquant le gène codant pour la partie variable des chaines légères d'Ig (IgVL). Ce dernier gène ne comporte pas de segment «D». Il s'agit donc d'une recombinaison «VJ». Les chaines légères peuvent être de deux types : kappa et lambda.

A la sortie de la moelle osseuse, les lymphocytes B naïfs expriment à leur surface un BCR complet et fonctionnel.

1.1.2. L'étape ganglionnaire ou immunopoïese

Les lymphocytes B naïfs migrent vers les follicules primaires des organes lymphoïdes secondaires. C'est ici que la rencontre avec l'antigène est orchestrée. L'antigène est présenté par les cellules présentatrices d'antigènes (macrophages, cellules dendritiques). L'essentiel des lymphocytes B naïfs qui pénètrent dans le ganglion lymphatique ne rencontreront pas d'antigène complémentaire et mourront par apoptose. *A contrario*, une très faible proportion des lymphocytes B rencontrera un antigène pour lequel il présente une spécificité et pourra recevoir des signaux de survie en provenance des lymphocytes T et migrer dans une zone particulière du ganglion appelée centre germinatif.

Dans le centre germinatif, la prolifération des lymphocytes B sélectionnés provoque la transformation du follicule primaire en un follicule secondaire. Au sein du follicule secondaire, les lymphocytes B vont subir une nouvelle étape de sélection. A ce stade, les lymphocytes B

survivants présentent une faible affinité pour l'antigène. Pour augmenter l'affinité de l'Ig de surface pour l'antigène, l'organisme met en jeu un processus d'hyper-mutation somatique des gènes d'Ig. Il s'agit d'une phase de remaniements géniques impliquant des cassures localisées du double-brin d'ADN, des mutations ponctuelles, des délétions et des duplications impliquant les segment V des gènes réarrangés codant pour l'IgvH.[53] Cette étape suppose l'arrêt transitoire de l'expression du BCR ainsi que la mise au repos de molécules anti-apoptotiques comme BCL2 (*B-Cell CLL/Lymphoma 2*). Le chef d'orchestre au niveau du centre germinatif est BCL6 (*B-Cell CLL/Lymphoma 6*). L'Ig mutée est réexprimée à la surface du lymphocyte B et son affinité pour l'antigène est de nouveau testée par les lymphocytes T *helpers* et les cellules dendritiques du follicule secondaire. Ne seront sélectionnés que les rares lymphocytes B dont l'affinité pour l'antigène a augmenté.

Ceux-ci peuvent alors entrer dans la phase de commutation isotypique qui permet de changer de classe d'Ig par modification de la partie constante de la chaine lourde (passage d'une IgM à une IgG, A ou E) afin d'augmenter son efficacité immunologique, ceci sans modification de la partie variable. Les lymphocytes B expriment alors à nouveau la protéine anti-apoptotique BCL2.

Les lymphocytes sortent ensuite du centre germinatif, certains sous la forme de lymphocytes B mémoire, gardiens de la mémoire immunitaire à long terme, permettant une réponse immunitaire plus rapide et plus efficace en cas de réintroduction de l'antigène dans l'organisme, et d'autres sous la forme de plasmocytes, cellules spécialisées dans la synthèse d'Ig.

La lymphopoïèse B présente donc un risque important de transformation maligne en particulier lors des phases d'hyper-mutation somatique (si elle n'est pas parfaitement contrôlée et affecte d'autres gènes que ceux codant pour l'Ig) et de commutation isotypique qui ont lieu dans le centre germinatif. En effet, même s'il est hautement régulé, ce processus de réarrangement génique présente un risque pour l'intégrité du génome et se conclut parfois par une transformation maligne lymphomateuse.[54-56]

1.2. La lymphomagénèse

A la différence des leucémies aiguës lymphoblastiques qui résultent de la transformation maligne d'un lymphocyte naïf médullaire, les lymphomes sont des tumeurs à point de départ ganglionnaire, liées à la prolifération maligne de lymphocytes matures ayant rencontré l'antigène.

Les apports de la cytogénétique ont permis d'identifier des translocations chromosomiques récurrentes et caractéristiques de certains sous-types de lymphomes.

Ces translocations relèvent de la fusion de la partie promotrice d'un gène avec la partie codante d'un oncogène. Ainsi, la translocation (14;18) met en présence la partie promotrice du gène codant pour la chaine lourde des Ig avec *BCL2* et est associée au lymphome folliculaire alors que la translocation (11;14) fusionne la partie promotrice du même gène avec le gène codant pour la cycline D1, ce qui est à l'origine du lymphome du manteau. Les translocations (8;14) et (14;22) fusionnent le gène codant pour le facteur de transcription MYC (*Avian Myelocytomatosis virus oncogene homolog*) avec la partie promotrice respectivement de la chaine légère kappa et de la chaine légère lambda, ce qui dans les deux cas est à l'origine du lymphome de Burkitt. Etant donné la forte activité promotrice des gènes codant pour les composants de l'Ig au cours de la lymphopoïèse B, ces translocations possèdent un fort pouvoir de dérégulation génique.

Ces translocations chromosomiques n'affectent pas toutes les éléments régulateurs-clé que sont BCL2, BCL6 et MYC : plusieurs centaines de ces translocations ont été identifiées à ce jour. Ces réarrangements sont dus aux processus physiologiques de la lymphopoïèse B qui impliquent et tolèrent des cassures double-brin de l'ADN : les recombinaisons VDJ, les hypermutations somatiques et la commutation de classe.

Dans les deux cas, la lymphomagénèse est la conséquence de la dérégulation de fonctions cellulaires-clé comme le programme de différenciation, la prolifération cellulaire ou la survie.

Pourtant, la transformation tumorale requiert des événements oncogéniques secondaires dans le cadre du modèle actuel d'oncogénèse multi-hit.[57] Dans ce modèle, trois types de mutations secondaires sont décrites. Les mutations défavorables à la tumeur sont présentes au sein du spectre de mutations du clone malin mais n'offrent aucun avantage de prolifération ou de survie. Les mutations positives ou « *driver* » confèrent un avantage de sélection. Les mutations neutres ou « *passenger* » ne confèrent ni avantage ni désavantage mais peuvent se convertir en mutations positives à la faveur d'un changement de la pression de sélection. Parmi les

événements oncogéniques identifiés, on note les mutations ou duplications/délétions affectant des gènes suppresseurs de tumeur comme TP53, CD95, ATM (*Ataxia Telangiectasia Mutated*) et IKbα.

La mise en jeu du BCR est essentiel à la survie cellulaire, son activation constante/constitutive joue un rôle dans la lymphomagénèse. Le rôle de l'activation continue du BCR par le virus d'Epstein-Barr (EBV) dans la genèse du lymphome de Burkitt est un premier exemple. La mise en jeu constitutive de la voie du BCR, c'est-à-dire de tout ou partie des effecteurs d'aval du BCR avec mise en jeu de la voie NF-κB (*Nuclear Factor kappa B)* est un autre exemple.

Enfin, le microenvironnement est tout aussi important dans la lymphomagénèse en ceci qu'il permet la survie et le maintien de la prolifération cellulaire.

2. CLASSIFICATION DES DIFFERENTS LYMPHOMES MALINS NON HODGKINIENS.

Les lymphomes malins sont partagés en deux grands groupes : les lymphomes Hodgkiniens et les lymphomes non Hodgkiniens (LNH). La forme la plus fréquente de LNH est le lymphome B diffus à grandes cellules (LBDGC), qui représente 30 à 40% des LNH de l'adulte.[58]

3. CLASSIFICATION OMS 2008

L'identification d'un type de lymphome se fait principalement en fonction de sa cellule d'origine c'est-à-dire sa contrepartie physiologique dans la lymphopoïèse normale. La plupart des lymphomes dérivent de cellules normales du centre germinatif ce qui confirme le rôle important de cette phase dans la lymphopoïèse.

Par ailleurs, les lymphomes sont aussi identifiés par les différentes zones ganglionnaires aux dépens desquelles ils se développent. Ainsi on trouve des lymphomes de la zone marginale, des lymphomes du manteau, des lymphomes folliculaires (**Figure 10**).

Certains sous-types de lymphomes B naissent aux dépens de tissus lymphoïdes associés aux muqueuses (MALT), ou sont découverts après avoir essaimé à un organe non-lymphoïde (lymphome primitif cutané, LBDGC primitif du système nerveux central).

Enfin, certains sous-types de lymphomes se développent à la suite d'une stimulation antigénique chronique. C'est le cas des LBDGC EBV positifs et des LBDGC secondaires à HHV8 (*Human Herpes Virus 8*).

Cette classification a été considérablement enrichie et diversifiée grâce aux apports de la génétique (**Tableau XI**).

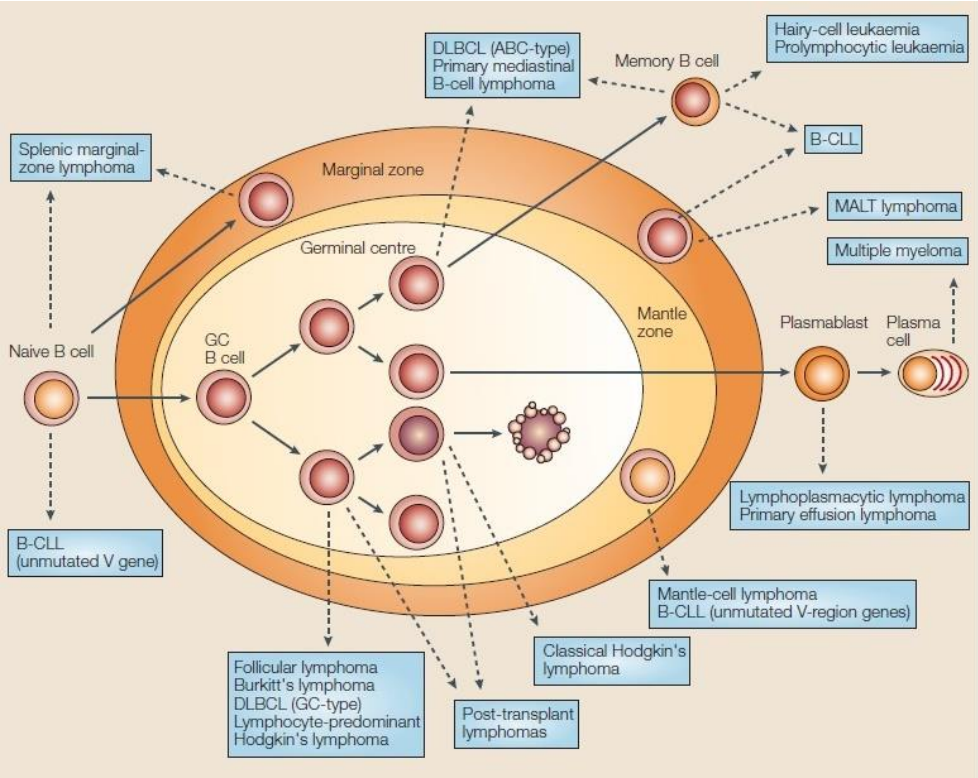


Figure 10. Lymphomes B assignés à leur équivalent lymphocytaire B normal. D'après Küppers *et al*.[53]

Leucémie lymphoïde chronique / lymphome lymphocytique
Leucémie prolymphocytaire B
Lymphome de la zone marginale splénique
Leucémie à tricholeucocytes
Lymphomes/leucémies spléniques non classables
Lymphome lymphoplasmocytaire
Maladies des chaines lourdes
Néoplasies à cellules plasmocytaires
Lymphomes des tissus lymphoïdes associés aux muqueuses (MALT)
Lymphome de la zone marginale ganglionnaire
Lymphome folliculaire
Lymphome primitif cutané
Lymphome à cellules du manteau
Lymphomes B-diffus à grandes cellules (LBDGC)
Lymphome à grandes cellules B riches en cellules T et en histiocytes
Lymphome B-diffus à grandes cellules primitif du système nerveux central
Lymphome B-diffus à grandes cellules primitif de la jambe (*leg type*)
Lymphome B-diffus à grandes cellules EBV positif de la personne âgée
Lymphome B-diffus à grandes cellules associé à une inflammation chronique
Granulomatose lymphomateuse
Lymphome à grandes cellules B primitif du médiastin (thymique)
Lymphome intravasculaire à grandes cellules B
Lymphome à grandes cellules B ALK (*Anaplastic Lymphoma Kinase*) positif
Lymphome plasmoblastique
Lymphome à grandes cellules B associé à HHV8
Epanchement lymphomateux primitif
Lymphome de Burkitt
Lymphome B inclassable avec des caractéristiques intermédiaires entre LBDGC et lymphome de Burkitt
Lymphome B inclassable avec des caractéristiques intermédiaires entre LBDGC et lymphome de Hodgkin

Tableau XI. Classification OMS 2008 des hémopathies malignes B matures. D'après Swerdlow *et al*.[59]

4. LES LYMPHOMES B-DIFFUS A GRANDES CELLULES

4.1. Epidémiologie

Aux Etats-Unis en 2013, l'incidence estimée des LNH était de 69 740 nouveaux cas par an, ce qui représente 4% des nouveaux cas de cancer, et 3% des causes de mortalité par cancer, sans distinction de sexe.[60]

Les LBDGC représentent 30 à 40% des LNH dans les pays industrialisés et un pourcentage supérieur dans les pays en voie de développement.[58]

Il s'agit d'une maladie touchant préférentiellement les sujets âgés, avec une médiane d'âge au diagnostic se situant dans la septième décade de la vie.

La majorité des LBDGC sont primitifs/*de novo* mais dans un certain nombre de cas, le LBDGC constitue une évolution agressive d'un lymphome indolent (LLC, lymphome folliculaire, lymphome de la zone marginale).

4.2. Clinique

Lorsqu'ils sont symptomatiques, les patients se présentent avec un tableau tumoral rapidement évolutif avec atteintes ganglionnaires et extra-ganglionnaires multiples. Par ailleurs, on observe des signes généraux en rapport avec la sécrétion cytokinique : perte de poids, sueurs nocturnes.

4.3. Morphologie/histologie

L'analyse histologique des ganglions lymphatiques tumoraux révèle la présence d'une prolifération monomorphe de grandes cellules lymphoïdes effaçant l'architecture ganglionnaire normale et infiltrant les tissus extra-ganglionnaires.

Les deux principales présentations morphologiques sont les formes centroblastiques, les formes immunoblastiques et les formes anaplasiques.

4.3.1. Le variant centroblastique

Les centroblastes correspondent à des cellules d'une taille deux à trois fois supérieure à un lymphocyte normal. Le noyau est rond ou ovalaire, avec un rapport nucléo-cytoplasmique se situe entre 0,5 et 0,6. La chromatine est fine, le plus souvent nucléolée et la basophilie cytoplasmique est moyenne à forte.

Il s'agit de la présentation morphologique la plus fréquente. Toutefois, dans la plupart des cas elle est hétérogène avec coexistence de centroblastes et d'immunoblastes.

4.3.2. Le variant immunoblastique

Il s'agit d'un variant plus rare, caractérisé par la présence de plus de 90% d'immunoblastes. Les immunoblastes possèdent la plupart des caractéristiques du centroblaste, mais présentent une taille et un rapport nucléocytoplasmique supérieurs et une basophilie plus forte.

4.3.3. Le variant anaplasique

Dans cette troisième présentation morphologique, on trouve des cellules de très grande taille, rondes ou polygonales avec un noyau pléiomorphe, ce qui pose le problème du diagnostic différentiel avec des cellules de Hodgkin, des cellules de Reed-Sternberg ou des cellules de lymphome T anaplasique.

4.3.4. Immunophénotypage

Les cellules néoplasiques expriment les marqueurs B : CD19, CD20, CD22 et CD79a. La présence d'une Ig membranaire (mais sans co-expression de CD38 et de CD138) ou intracellulaire est authentifiée dans 50 à 75% des cas. Le CD30 peut être exprimé dans le variant anaplasique. Le CD5 est retrouvé dans 10% des cas, CD10 dans 30 à 60% des cas, BCL6 dans 60 à 90% des cas et IRF4 dans 35 à 65% des cas avec co-expression de BCL6 et IRF4 (*Interferon Regulatory Factor 4*) dans 50% des cas. L'indice de prolifération Ki67 est élevé dans la plupart des cas.[58]

5. PROFIL D'EXPRESSION GENIQUE

Deux sous-groupes de LBDGC ont été identifiés sur la base d'une étude d'expression génique sur puce à ADN : le sous-type *Germinal Center B-like* (GC) et le sous-type *Activated B-Cell like* (ABC) (**Figure 11**).[61] Ces classifications ne sont pas superposables aux classifications morphologiques. Plus tard, l'analyse des profils d'expression génique a permis d'identifier un troisième sous-type que nous ne traiterons pas ici : le lymphome à grandes cellules B primitif du médiastin.[62]

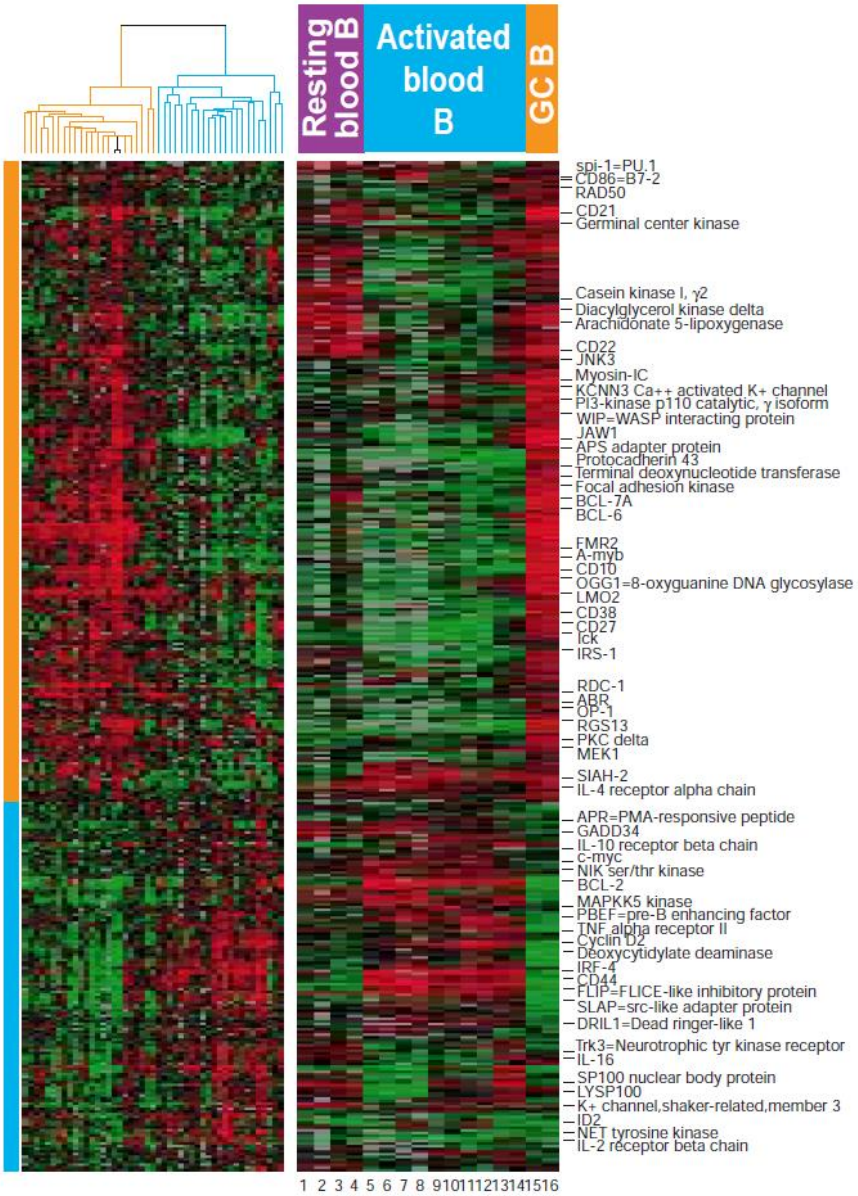


Figure 11. Les LBDGC de type GC et ABC sont identifiables par leur signature transcriptomique. D'après Alizadeh *et al*.[61]

5.1. Le sous-type Germinal Center B-like (GC)

Ces LBDGC présentent un profil d'expression génique correspondant aux cellules du centre germinatif, avec notamment l'expression de *BCL6*, *LRMP (Lymphoid Restricted Membrane Protein)* et *LMO2 (LIM domain Only 2)*.

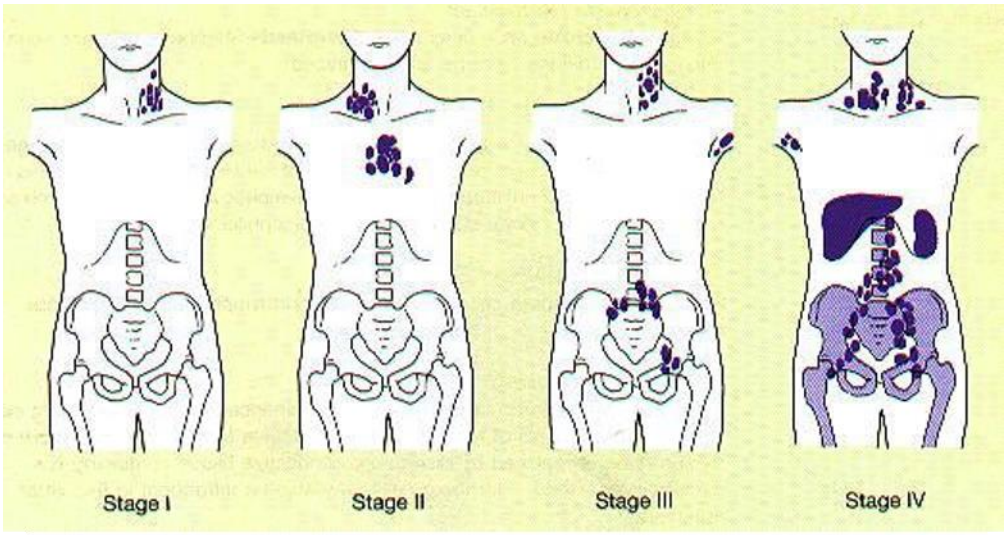
5.2. Le sous-type Activated B-Cell (ABC)

Il correspond au profil d'expression d'une cellule en voie de transformation plasmoblastique, avec notamment l'expression *d'IRF4, PIM2 (*Pim-2 proto-oncogène, serine/thréonine kinase*) et FOXP1 (Forkhead-box P1).*

6. PRONOSTIC

6.1. Pronostic clinique

Sur le plan clinique, la classification d'Ann Arbor,[63] initialement conçue pour le lymphome Hodgkinien tient compte du nombre d'aires ganglionnaires touchées et de la notion d'atteinte localisée (d'un seul côté du diaphragme) ou étendue (des deux côtés du diaphragme) mais n'est pas parfaitement adaptée aux LNH (**Figure 12**).



Stade I Atteinte d'*1 seule aire* ganglionnaire sus- ou sous-diaphragmatique

Stade II Atteinte d'*au moins 2 aires ganglionnaires d'un même côté du diaphragme*

Stade III Atteintes ganglionnaires *de part et d'autre du diaphragme (la rate correspond à un groupe ganglionnaire)*

Stade IV *Atteinte viscérale* à distance d'une aire ganglionnaire atteinte (médullaire, hépatique, pulmonaire...)

Figure 12. Classification pronostique d'Ann Arbor.
D'après Carbone *et al*.[63]

La distinction entre les stades localisés (I et II) et les stades étendus (III et IV) entre dans le calcul du score IPI (*International Prognostic Index*). Ce score est pertinent pour l'évaluation de la gravité des LBDGC et repose sur des critères cliniques et biologiques (**Figure 13**). Ce score permet de classer les malades en 4 groupes pronostiques.[64]

Score IPI (*International Prognostic Index*)
1.Age > 60 ans
2.ECOG *performans status* ≥ 2
3.Stade Ann Arbor III ou IV
4.Lactate deshydrogenases > normal
5.Atteinte > 1 site extra-ganglionnaire

Groupes de risque :
Faible risque : 0 ou 1 point
Risque intermédiaire faible : 2 points
Risque intermédiaire haut : 3 points
Risque élevé : 4 ou 5 points

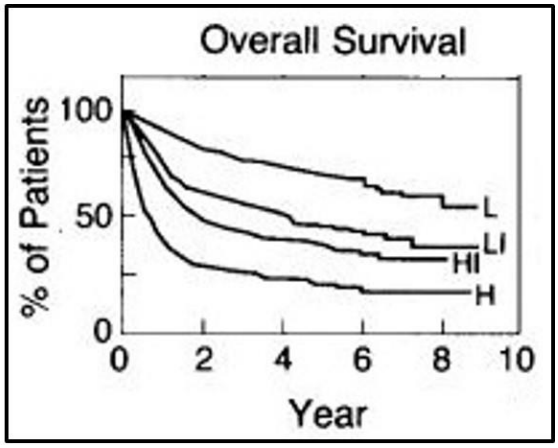


Figure 13. Classification pronostique selon le score IPI.
L: Low risk; LI: Low Intermediate risk; HI: High Intermediate risk; H: High risk. D'après [64]

6.2. Apports pronostics du profil d'expression génique

Les LBDGC sont classés en trois entités différentes : le lymphome primitif du médiastin, le LBDGC de type *Germinal Center B-like* (GC) et le LBDGC de type *Activated B-Cell like* (ABC). Ces deux derniers types de LBDGC ont été identifiés grâce à une étude de profils d'expression génique (**Figure 14**).[61]

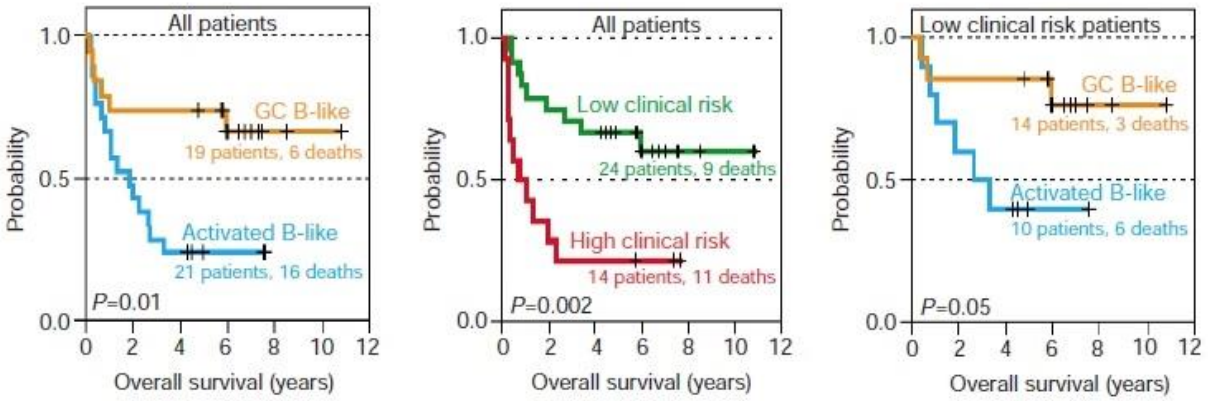


Figure 14. Les LBDGC de type GC et ABC sont deux entités de pronostic différent. D'après Alizadeh *et al*.[61]

Par la suite, plusieurs groupes ont proposé des scores pronostiques fondés sur l'existence de signatures d'expression.[65-68]

6.3. Apports de l'immunohistochimie

La classification des LBDGC en sous-types GC et ABC repose sur l'analyse du profil d'expression génique. Toutefois, cette analyse est trop coûteuse pour être réalisée en pratique courante. Cette classification a été approchée par méthode immunohistochimique, sur tissu en paraffine, en GC et non-GC sur la base de l'expression protéique de BCL6, IRF4 et CD10. Les LBDGC CD10-, BCL6+, IRF4- sont considérés GC et les autres non-GC.[69] Cette classification n'est pas parfaitement superposable avec la classification fondée sur le profil d'expression génique. D'autres algorithmes décisionnels ont été proposés par la suite.[70-74] L'expression protéique de LMO2 est associée au sous-type GC et celle de FOXP1 au sous-type non-GC.[75]

6.4. Apports de la FISH (Fluorescence In Situ Hybridization)

Parmi les LBDGC, l'analyse par FISH permet d'identifier une translocation impliquant *BCL2* (18q21) dans 20% des cas (essentiellement de type GC), une translocation impliquant *BCL6* (3q27) dans 30% des cas (essentiellement de type non-GC) et une translocation impliquant *MYC* (8q24) dans 3% des cas. La combinaison de données immuno-histochimiques et cytogénétiques a permis de construire un score mixte composé de 3 marqueurs associés au sous-type non-GC : (i) expression protéique de FOXP1, expression protéique de MUM1/IRF4 et (iii) translocation impliquant *BCL6*. L'existence d'au moins 2 des 3 paramètres est associée à une diminution de la survie globale à 5 ans, indépendamment du score IPI associé à l'âge.[75]

7. THERAPEUTIQUE

Ce type de lymphome est caractérisé par la prolifération diffuse de grandes cellules lymphoïdes B effaçant l'architecture ganglionnaire normale. Son évolution agressive en fait une maladie spontanément létale en l'absence de traitement.

7.1. Traitements de première ligne

Depuis 1993, la chimiothérapie de référence pour le traitement des LBDGC est le CHOP (Cyclophosphamide, Adriblastine, Oncovin, Prednisone). La survie sans progression et la survie globale à 3 ans sont respectivement de 40 et 50%.[76] Le schéma de référence prévoit une cure tous les 21 jours (CHOP21).

En 2004, le groupe d'étude Allemand « *Deutsche Studiengruppe für Hochmaligne Non-Hodgkin-Lymphome* » (DSHNHL) a voulu comparer le CHOP21 à un schéma d'administration sur 14 jours (CHOP14) ainsi qu'aux mêmes schémas avec l'ajout d'etoposide (CHOEP14 et CHOEP 21). La première étude a montré la supériorité du CHOEP sur le CHOP chez les sujets jeunes de bon pronostic [77] et celle du CHOP14 sur le CHOP21 chez les sujets âgés de plus de 60 ans.[78]

L'avènement de l'immunothérapie par anticorps anti-CD20 (Rituximab) a permis d'améliorer considérablement la réponse au traitement ainsi que la survie globale. Etant donné son efficacité et sa faible toxicité, le Rituximab (R) a pu être ajouté à la chimiothérapie par CHOP, ce qui a porté la survie sans évènement et la survie globale à 5 ans respectivement à 54 et 58%, faisant du R-CHOP le traitement standard de première ligne des LBDGC chez les patients de plus de 60 ans.[79, 80]

Cette supériorité du R-CHOP sur le CHOP est vérifiée aussi bien pour les LBDGC de type GC que pour les type non-GC, les LBDGC de type GC restant de meilleur pronostic que les non-GC.[81]

Les études comparant le R-CHOP14 et le R-CHOP21 ont montré une équivalence des deux schémas en terme de survie sans événement, avec une plus grande toxicité du R-CHOP14.[82, 83] Le schéma de référence aujourd'hui est donc le R-CHOP21.

La stratégie thérapeutique de première ligne est guidée par le score IPI ajusté à l'âge (IPI_{aa}) (**Figure 15**).[64]

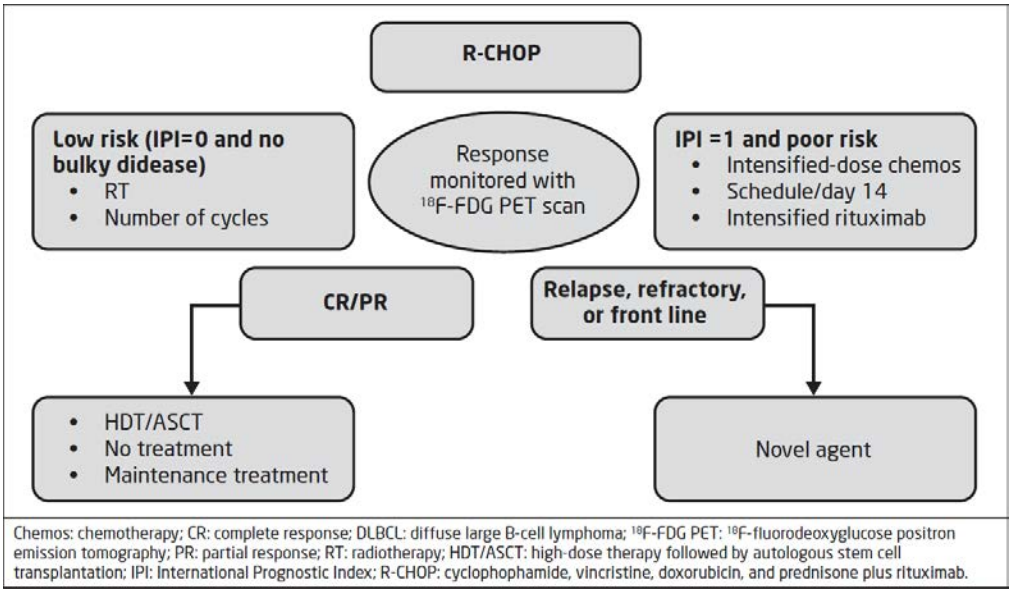


Figure 15. Stratégie de traitement de première ligne des LBDGC.
D'après Thieblemont *et al*.[84]

7.2. Traitement des sujets jeunes (< 60 ans)

Pour les stades localisés, sans masse tumorale importante (5 à 10 cm selon les auteurs) avec un score IPIaa de 0 (groupe de très bon pronostic), la supériorité de 3 cycles d'ACVBP (Doxorubicin, Cyclophosphamide, Vindesine, Bleomycin, Prednisone) suivie d'une consolidation séquentielle sur l'association CHOP + radiothérapie en terme de survie globale et de survie sans événement à 5 ans a été démontrée.[85] L'association R-ACVBP est actuellement un des traitements de référence.

L'approche thérapeutique par 6 cures de R-CHOP est efficace elle-aussi avec une survie sans événement supérieure à 90% et une survie globale à 5 ans de 97%.[86]

L'enjeu chez ces patients de faible risque est de réduire la toxicité de la chimiothérapie tout en maintenant son efficacité.[84] C'est le sens des études actuelles qui visent à évaluer l'efficacité de 4 R-CHOP + 2 R face à 6 R-CHOP.

Lorsqu'il existe une masse tumorale importante mais localisée, l'adjonction d'une radiothérapie sur les aires atteintes permet d'obtenir de meilleurs résultats qu'avec la chimiothérapie seule.[87]

Pour les patients jeunes de risque intermédiaire, à savoir avec un IPI_{aa} de 1, le traitement par 6 cures de R-CHOP est supérieur au traitement par CHOP.[86] Suite au protocole LNH03-2B,

le schéma fondé sur le R-ACVBP suivi d'une consolidation est devenu la référence suite à la démonstration de sa supériorité sur le R-CHOP chez ces patients.[88]

Pour les stades avancés, à savoir un IPIaa de 2 ou 3, le pronostic est plus sombre avec une survie globale à 5 ans inférieure à 50%.[89] Pour ces patients, les traitements efficaces reposent sur une association Rituximab + chimiothérapie suivie d'une intensification thérapeutique de type BEAM (Carmusitne, Etoposide, Aracytine, Melphalan) couplée à une autogreffe de CSH.[90]

7.3. Traitement des sujets âgés (60-80 ans)

Pour les stades localisés sans facteurs de risque (IPIaa =0), l'irradiation des aires ganglionnaires atteintes en sus de la chimiothérapie par CHOP n'apporte pas d'avantage par rapport à la chimiothérapie seule.[91] L'étude RICOVER 60 (*Rituximab with CHOP Over Age 60 Years*) a montré le bénéfice d'ajouter du Rituximab à un schéma de chimiothérapie de type CHOP14.[92] Le protocole LHN03-6B du GELA (Groupe d'Etude des Lymphomes de l'Adulte) a montré l'équivalence du R-CHOP21 et du R-CHOP14 dans cette indication.[83]

7.4. Traitement des sujets très âgés (> 80 ans)

Le schéma basé sur le R-CHOP peut être utilisé chez ces patients.[79] Toutefois, l'accumulation de comorbidités et l'état cardiaque souvent fragile de ces patients empêchent le plus souvent le recours au R-CHOP. Le schéma à doses réduites (R-mini CHOP) constitue une possibilité avec des premiers résultats intéressants en terme de survie et de toxicité.

7.5. Traitement des rechutes

La rechute constitue un événement péjoratif dans l'évolution des LBDGC. Il existe plusieurs profils de patients en rechute.[93] Les LBDGC résistants au traitement de première ligne sont associés à une mortalité très importante, malgré les traitements de rattrapage (**Figure 16**).

Chez les patients réfractaires (ou non répondeurs), la maladie continue de progresser malgré le traitement ou évolue à nouveau dès la fin du traitement (8-10% des malades). Dans le

cas d'une réponse clinique objective mais insuffisante, avec persistance de localisations lymphomateuses au PET-scanner (*Positon Emission Tomography - scanner*), on parle de réponse partielle (8-10% des malades).

Parmi les patients chez lesquels une réponse complète est obtenue, le délai de la rechute après la mise en route du traitement est essentiel (20 à 30% des malades). Si la rechute est précoce et survient dans l'année qui suit la mise en route du traitement de première ligne, le pronostic est similaire à celui des patients réfractaires avec une médiane de survie inférieure à 6 mois.[94, 95] Si la rechute est tardive et survient plus d'un an après la fin du traitement de première ligne, la survie est plus longue (**Figure 16**).

La répartition entre rechute précoce et rechute tardive est équilibrée. Ainsi, les patients en rechute tardive ou en réponse partielle ont le même pronostic et représentent 20% des LBDGC. L'ensemble constitué des patients réfractaires et des patients en rechute précoce correspond aussi à 20% des LBDGC et présentent une faible survie. Environ 60% des malades ne rechutent pas et guérissent de leur maladie.

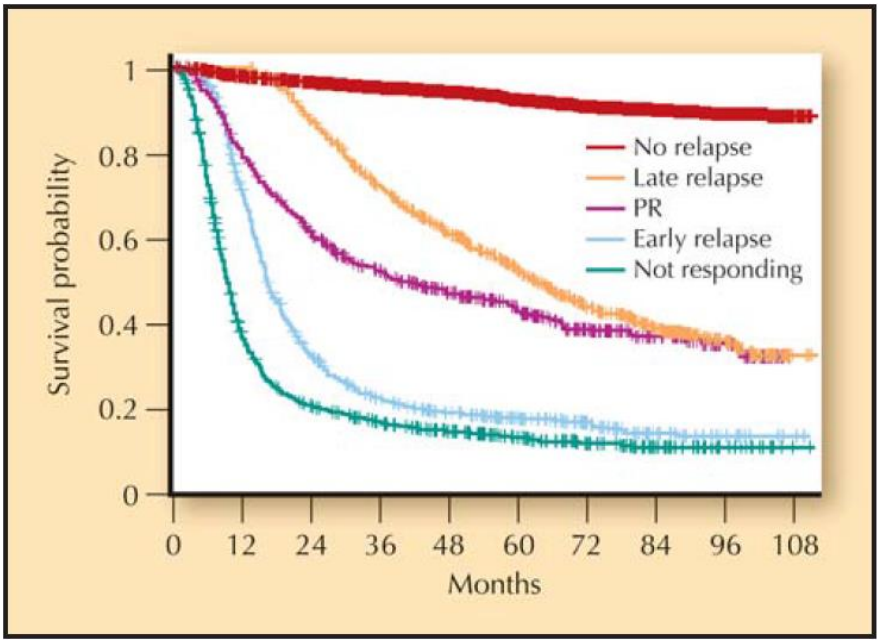


Figure 16. Survie globale chez 7400 patients atteints de LBDGC âgés de 18 à 80 ans, en fonction de leur réponse au traitement de première ligne présentant un LBDGC et inclus dans les protocoles du GELA.

La rechute est associée à une mortalité importante. PR : *Partial Response*. D'après Thieblemont *et al*.[93]

Les patients en rechute bénéficient d'une chimiothérapie de rattrapage, suivie d'une chimiothérapie intensive avec autogreffe de CSH. Le schéma associant 2 cures d'une chimiothérapie de deuxième ligne par DHAP (Dexamethasone, Aracytine haute dose, Cisplatine) puis une chimiothérapie intensive avec autogreffe de CSH offre une meilleure survie sans progression et une meilleure survie globale que la chimiothérapie par DHAP seule.[94]

En traitement de la rechute, l'allogreffe donne des résultats équivalents à l'autogreffe en terme de survie sans progression, mais au prix d'une mortalité due au traitement beaucoup élevée.[96] L'autogreffe est donc préférable à l'allogreffe en traitement de la rechute. Toutefois, l'autogreffe de CSH est un traitement lourd qui n'est accessible qu'aux sujets jeunes sans comorbidité et qui n'est efficace que si la maladie reste chimiosensible.

L'allogreffe peut être indiquée chez les patients en seconde rechute après autogreffe, avec un meilleur pronostic si la première rechute est survenue après plus d'un an.[97]

L'ajout du Rituximab aux chimiothérapies de rechute type DHAP ou ICE permet d'obtenir un meilleur taux de rémission et une meilleure survie sans progression qu'avec la chimiothérapie seule.[98, 99]

8. DEREGULATIONS MOLECULAIRES DANS LES LBDGC

8.1. Anomalie associées préférentiellement au sous-type GC

Les LBDGC de type GC dérivent d'un lymphocyte du centre germinatif et présentent fréquemment des anomalies impliquant des gènes qui gouvernent la mise en jeu du centre germinatif B normal (**Figure 17**).

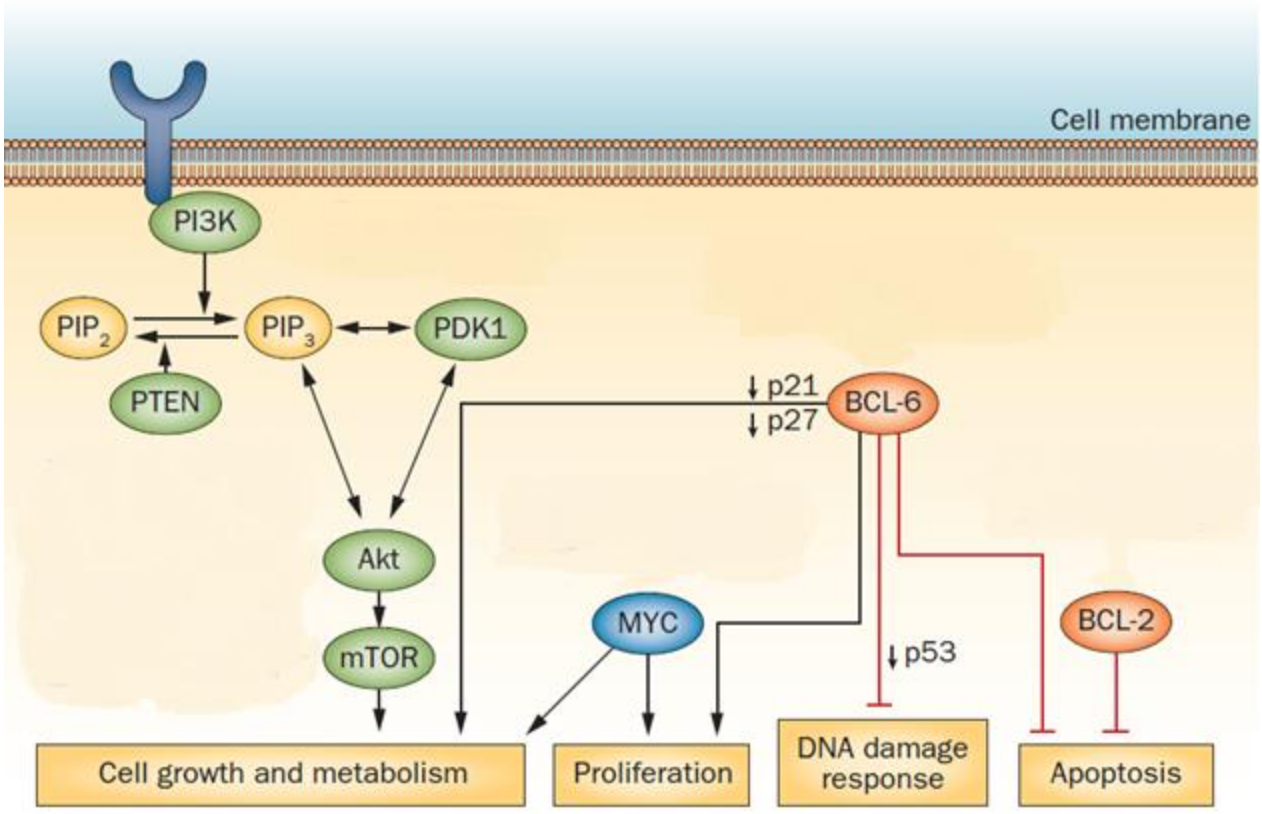


Figure 17. Principales voies de régulation impliquées dans la genèse des LBDGC de type GC. D'après Roschewski *et al*.[100]

8.2. Translocations impliquant le gène *BCL2*.

BCL2 est un inhibiteur de l'apoptose dont l'expression est transitoire au cours de la mise en jeu du centre germinatif. Le gène *BCL2* est situé sur le chromosome 18. La translocation (14;18) (q32;q21) met en présence la partie promotrice du gène codant pour la chaine lourde des immunoglobulines (en 14q32) avec *BCL2* (en 18q21), ce qui a pour conséquence l'activation constitutive de BCL2 qui se comporte alors comme un oncogène. Cette anomalie est

classiquement associée au lymphome folliculaire et est considérée comme l'événement initiateur de la lymphomagénèse dans ce sous-type de lymphome.

Elle est aussi retrouvée dans 17% des LBDGC. La grande majorité des translocations (14;18) (q32;q21) est retrouvée dans le sous-type GC (34% des GC présentent cette translocation) et présente un profil d'expression génique correspondant au sous-type GC. Aucune translocation (14;18) n'a été retrouvée dans le sous-type ABC. La translocation (14;18)(q32;q21) est associée à la surexpression (i) de gènes impliqués dans l'apoptose : BCL2, BAD (*BCL2-associated Agonist of cell Death*), DRAK1 (*DAP Kinase-Related Apoptosis-Inducing Protein Kinase 1*), (ii) de gènes associés à l'adhésion cellulaire : CD62L, SDF1/CXCL12 (*Chemokine (C-X-C Motif Ligand 12*) , CD31 et (iii) de facteurs de transcription : FLI1 (*Friend Leukemia Virus Integration 1*), HOX11 (*Homeo Box 11*). L'étude immuno-histochimique montre que les protéines BCL2, BCL6 et CD10 sont davantage exprimées en cas de translocation (14;18)(q32;q21).[101, 102]

Pour autant, même en l'absence de t(14 ;18)(q32 ;q21) le sous-type ABC présente un haut niveau d'expression protéique et transcriptomique de BCL2, ceci par deux mécanismes : la duplication du gène et l'augmentation de la transcription de *BCL2 via* l'activation constitutive de la voie NF-kB.[103]

Par ailleurs, le promoteur du gène *BCL2* est très fréquemment muté à cause des hyper-mutations somatiques, ce qui entraine sa résistance à l'inhibition par BCL6.[104]

Enfin, *BCL2* est le gène le plus fréquemment muté dans les LBDGC avec 37% des échantillons présentant au moins une mutation ponctuelle dans la partie codante du gène. Ces mutations de *BCL2* sont significativement associées à la translocation (14 ;18)(q32 ;q21) et au sous-type GC.[102]

8.3. Activation constitutive de BCL6

BCL6 est un régulateur essentiel de la réaction ayant lieu dans le centre germinatif. Physiologiquement, il agit comme répresseur transcriptionnel et son expression est restreinte aux cellules B du centre germinatif. Il est recruté par de nombreux promoteurs de gènes, impliqués dans la survie cellulaire (*BCL2*), la prolifération cellulaire (*MYC*), l'apoptose en réponse aux dommages à l'ADN (*TP53,ATR,CHEK1*), le contrôle du cycle cellulaire (*CDKN1A/B*), la synthèse de cytokines (impliquées dans l'activation de la voie JAK-STAT notamment), l'expression et la mise en jeu des récepteurs toll-like, de TGFβ (*Tumor Growth Factor Beta*) et de la voie de signalisation WNT (*Wingless-Type MMTV Integration Site*), qui sont des régulateurs positifs de la

différenciation post-centre germinatif) et enfin la différenciation cellulaire plasmoblastique (*PRDM1/Blimp1*). En somme, BCL6 permet le maintien de la réaction du centre germinatif, c'est-à-dire un haut niveau de division cellulaire et de prolifération centroblastique, ainsi que les hypermutations somatiques des gènes codant pour l'Ig de surface. Le but de cette réaction est d'augmenter l'affinité du BCR pour l'antigène mais se fait au prix d'un risque tumoral augmenté en raison de la tolérance aux nombreuses cassures de l'ADN.[105]

Une fois que la sélection d'un lymphocyte B de haute affinité pour l'antigène a été réalisée, la réaction du centre germinatif doit prendre fin. Au stade centrocytique, la répression de BCL6 se fait par deux mécanismes : répression transcriptionnelle après la stimulation du CD40 lymphocytaire B par le CD40-ligand lymphocytaire T et phosphorylation puis dégradation de la protéine BCL6 par le protéasome après activation du BCR. Cette répression de BCL6 permet l'activation de la voie NF-kB et l'activation d'IRF4, ce qui permet le début de la différenciation plasmoblastique (**Figure 18**).

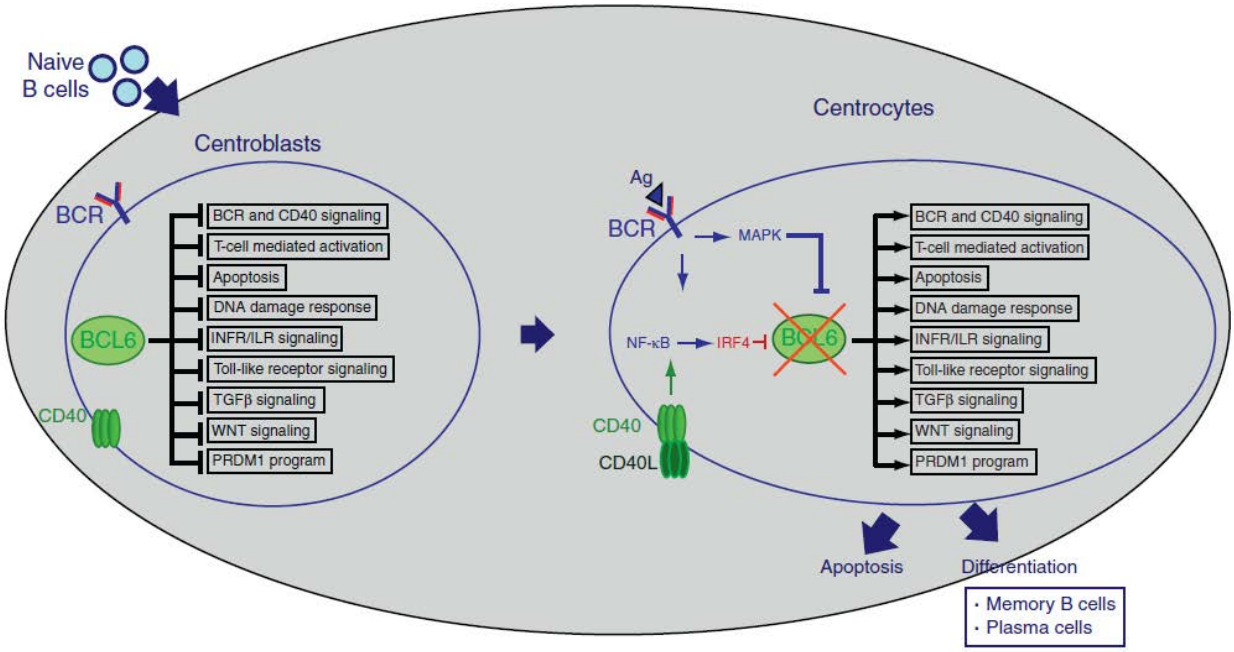


Figure 18. Fonctions cellulaires régulées par BCL6 au sein du centre germinatif. D'après Basso *et al*.[105].

Les translocations impliquant *BCL6* en 3q27 sont impliquées dans 35% des LBDGC.[106] Ces translocations mettent en présence la partie codante de *BCL6* avec des promoteurs constitutionnellement actifs dans les lymphocytes B, issus d'autres chromosomes. Ceci entraîne l'expression constitutive de *BCL6*, au-delà de la réaction du centre germinatif, alors que

physiologiquement cette expression est strictement circonscrite aux lymphocytes B du centre germinatif.

Par ailleurs, dans 61% des LBDGC (70% des GC et 44% des ABC),[107] la transcription constitutive de *BCL6* peut être issue d'un autre mécanisme, les mutations activatrices affectant ses régions régulatrices, en particulier son promoteur, acquises au cours de la phase d'hypermutations somatiques. En effet celles-ci peuvent affecter d'autres gènes que les gènes codant pour l'Ig de surface, en particulier *BCL6*.

Les deux situations aboutissent au maintien de l'expression de *BCL6* au-delà du stade centroblastique, aux stades immunoblastiques et plasmoblastiques. Ceci favorise la prolifération cellulaire à travers la sous-expression de protéines-clé régulant le cycle cellulaire (CDKN1A/B), la résistance à l'apoptose à travers la sous-expression de TP53 et donc la tolérance aux cassures double-brin de l'ADN, mais aussi la répression de la différenciation plasmoblastique.[105, 106]

La forte occurrence de translocations impliquant *BCL6* dans les LBDGC de type ABC indique que ces mécanismes favorisant la lymphomagénèse sont aussi mis en jeu dans les cellules B post-centre germinatif.[107]

8.4. Activation de la voie PI3K/AKT/mTOR

La voie PI3K/AKT/mTOR (*mammalian Target Of Rapamycin*) permet la transduction d'un signal de croissance et de survie cellulaires ainsi que d'angiogenèse. La PI3K phosphoryle le phosphatidyl-inositol-4,5 diphosphate en position 3 pour donner le phosphatidyl-inositol-3,4,5 triphosphate, qui active secondairement AKT en aval de la voie. PTEN *(Phosphatase and Tensin Homolog)* est le principal régulateur négatif de la voie PI3K/AKT/mTOR et agit comme une phosphatase lipidique en déphosphorylant le phosphatidyl-inositol-3,4,5 triphosphate en position 3, permettant ainsi la mise au repos de la voie.

Un grand nombre de LBDGC de type GC sont caractérisés par une activation constitutive de cette voie. Celle-ci se fait par l'inactivation de PTEN et relève de trois mécanismes : la délétion du gène *PTEN*, la perte d'expression de *PTEN* et la surexpression de miR17-92, un inhibiteur de PTEN (**Figure 17**).[100]

La délétion de *PTEN* en 10q23 est observée dans 11% des LBDGC de type GC mais dans aucun cas de LBDGC de type ABC.[108] La perte de l'expression de *PTEN*, identifiée chez 55% des LBDGC de type GC et 14% des LBDGC non-GC est en rapport avec une délétion

monoallélique ou biallélique du gène et/ou une ou plusieurs mutations somatiques entraînant une perte de fonction (15% des cas). Ces mutations sont situées dans la région codant le domaine phosphatase de la protéine (faux-sens, non-sens ou décalage du cadre de lecture). L'expression de *PTEN* est inversement proportionnelle au niveau d'activation de la voie PI3K/AKT/mTOR.[109] Le micro ARN miR17-92 est un inhibiteur de l'expression de *PTEN*. Il est amplifié dans 12,5% des LBDGC de type GC, ce qui induit sa surexpression.[108, 110]

8.5. Anomalies de la régulation épigénétique

Les mutations affectant les régulateurs de la méthylation (histones et non-histone) sont identifiés dans les deux sous-types de LBDGC, avec une prédominance pour le sous-type GC.

EZH2 (*Enhancer of Zeste 2 polycomb repressive complex 2 subunit*) est une histone méthyl-transférase qui constitue la partie catalytique du complexe PRC2 (polycomb-repressive complex 2) dont l'une des fonctions consiste à ajouter un groupement méthyle sur la lysine 27 de H3, ce qui réprime la transcription.[111] La mutation somatique de l'exon 15 du gène *EZH2* a pour conséquence le remplacement d'une tyrosine en position 641 dans le domaine SET de la protéine. Le variant $EZH2^{tyr641}$ agit comme un dominant gain-de-fonction et présente une activité enzymatique augmentée par rapport à la forme non mutée.[112] $EZH2^{Y641}$ agit comme un répresseur transcriptionnel de gènes impliqués dans le contrôle du cycle cellulaire (comme *CDKN1A*) et de l'apoptose.[113, 114] Cette mutation a été décrite dans 22% des LBDGC de type GC mais n'a pas été identifiée dans le sous-type ABC.[115]

MLL2 (*Mixed Lineage Leukemia 2*) est une histone lysine methyltransferase ciblant H3K4, la trimethylation de H3K4 étant une marque d'activation transcriptionnelle régulant en particulier les gènes de la famille HOX. Les mutations ponctuelles biallélques affectant le gène *MLL2* ont été identifiées dans 24 à 32% des LBDGC, répartis de manière homogène entre les sous-types GC et ABC. Il s'agit majoritairement de mutations non-sens entraînant un arrêt de la synthèse protéique ou de petites insertions/délétions entraînant un décalage du cadre de lecture. Par ailleurs, des délétions du gène ont été décrites. Toutes ces anomalies ont pour conséquence la génération d'une protéine MLL2 non fonctionnelle dépourvue de son domaine méthyl-transférase SET.[116, 117]

CREBBP (*CREB Binding Protein*) et EP300 (*E1A binding Protein p300)* sont des lysine acetyl-transferases très proches structurellement et fonctionnellement, dont l'une des cibles est H3K27 et qui ont un effet co-activateur transcriptionnel. Les mutations et les délétions

monoalléliques affectant *CREBBP* sont retrouvées respectivement dans 29% des LBDGC et sont plus souvent retrouvées dans le sous-type GC que dans le sous-type ABC (41,5% contre 17%). Les mutations affectant *CREBBP* sont toutes à l'origine de l'inactivation du domaine catalytique de la protéine. Les délétions incluant CREBBP et les mutations inactivatrices sont mutuellement exclusives dans l'immense majorité des cas, ce qui indique une distribution préférentiellement monoallélique des anomalies et une haplo-insuffisance. L'effet inactivateur de CREBBP est synergique avec celui des mutations activatrices d'*EZH2*. Les mutations et délétions affectant *EP300* sont retrouvées dans 10% des LBDGC et ont les mêmes conséquences fonctionnelles que les anomalies de CREBBP, en particulier l'incapacité d'acétyler BCL6 et TP53, ce qui entraine une activation constitutive de BCL6 et une diminution de l'activité suppresseur de tumeur de TP53. Dans un faible nombre de cas, il existe un très faible taux d'expression protéique de CREBBP et/ou EP300, sans anomalie structurale de ces gènes, ce qui suggère un autre mécanisme d'inactivation.[116]

Enfin, ME2FB (*Myocyte Enhancer Factor 2B*) agit comme cofacteur de CREBBP et EP300. Lorsque la concentration cytoplasmique de calcium est augmentée, MEF2B se lie à CREBBP et EP300 est permet leur recrutement sur ses promoteurs cibles. *ME2FB* est muté dans 11,4% des LBDGC, tous de type GC.[117]

8.6. Anomalies moléculaires affectant la domiciliation des lymphocytes B

GNAI2 (*Guanine Nucleotide binding protein, alpha inhibiting activity polypeptide 2*), S1PR2 (*Sphingosine-1-Phosphate Receptor 2*) et GNA13 (*Guanine Nucleotide Binding Protein, Alpha 13*) codent pour des protéines qui agissent de concert pour réguler la domiciliation des cellules B. S1PR2 code pour S1P2, le récepteur à la cytokine S1P. GNAI2 code pour la sous-unité α d'une protéine G couplée à S1P2. Enfin, GNA13 code pour la protéine Gα13, elle-aussi associée au récepteur S1P2. La régulation de la migration des cellules B et de leur domiciliation est liée à l'affinité du récepteur S1P2 pour la cytokine/chimiokine S1P. La concentration de cette chimiokine est très élevée dans le centre germinatif des ganglions lymphatiques. Par ailleurs, GI2 et Gα13 participent à la transduction du signal médié par S1P2 avec des effets opposés sur l'activation de la voie PI3K/AKT (activatrice pour GI2 et inhibitrice pour Gα13).

Les mutations inactivatrices du gène codant pour la GTPase GNA13 ont été identifiées dans 20% des LBDGC *de novo* essentiellement du sous type GC. Les mutations et délétions

activatrices de GNAI2 et S1PR2 sont uniquement observées dans les LBDGC de type GC. L'effet prédictif de ces mutations est une activation constitutive de la voie PI3K/AKT.[118]

9. ANOMALIES ASSOCIEES PREFERENTIELLEMENT AU SOUS-TYPE ABC

Les LBDGC de type ABC dérivent d'un lymphocyte ayant déjà entamé sa différenciation plasmoblastique et sont caractérisés par l'activation constitutive des voies NF-KB et JAK-STAT. Les LBDGC de type ABC ont un profil d'expression génique comparable à celui des cellules B activées par la mise en jeu de leur BCR, c'est-à-dire à une phase plus tardive de la lymphopoïèse B qui précède la transformation en plasmoblaste. Pour cette raison, ce type de LBDGC exprime peu les gènes de la signature du centre germinatif.

Les facteurs de transcription de la famille NF-κB gouvernent la sécrétion cytokinique, l'inflammation, la prolifération, l'angiogenèse et la mort cellulaire.[119] A l'état basal, ces facteurs sont gardés inactifs par leur association à une protéine inhibitrice : IκBα. Il existe une voie canonique et une voie non-canonique (**Figure 19**).

La voie canonique est activable par le BCR et les récepteurs de la superfamille des TNF (*Tumor Necrosis Factor*). La cascade de signalisation passe par le recrutement de TRAF5 et TRAF6, puis de la serine-thréonine kinase MAP3K7 qui phosphoryle la kinase IKKB. La forme activée d'IKKB phosphoryle IκBα entraînant son ubiquitination puis sa dégradation, permettant la mise en jeu du facteur de transcription RelA/p50, de la famille NF-Kβ. TNFAIP3 est un régulateur négatif de cette voie, qui agit en inactivant IKKβ et TRAF6.[120]

La voie non-canonique est mise en jeu par les récepteurs CD40 et BAFF. Après fixation du ligand, le complexe régulateur négatif TRAF3/MAP3K14-TRAF2/BIRC3 est inactivé permettant le relargage cytoplasmique de MAP3K14, qui active la kinase IKKα qui phosphoryle à son tour RelB/p100, induisant sa protéolyse partielle entraînant la formation d'une protéine p52. La protéine p52 se lie à RelB/p50 puis migre vers le noyau et agit comme régulateur transcriptionnel.[120]

Les LBDGC de type ABC sont caractérisés par l'activation constitutive de la voie NF-κB, qui aboutit à l'expression du facteur de transcription IRF4 qui gouverne la transformation plasmoblastique.[121, 122] Cette activation constitutive de la voie NF-κB est la résultante d'un spectre de mutations affectant les différents déclencheurs de cette voie, qu'il s'agisse de la voie du BCR ou de la voie du TLR (*Toll-Like Receptor*) (**Figure 20**).

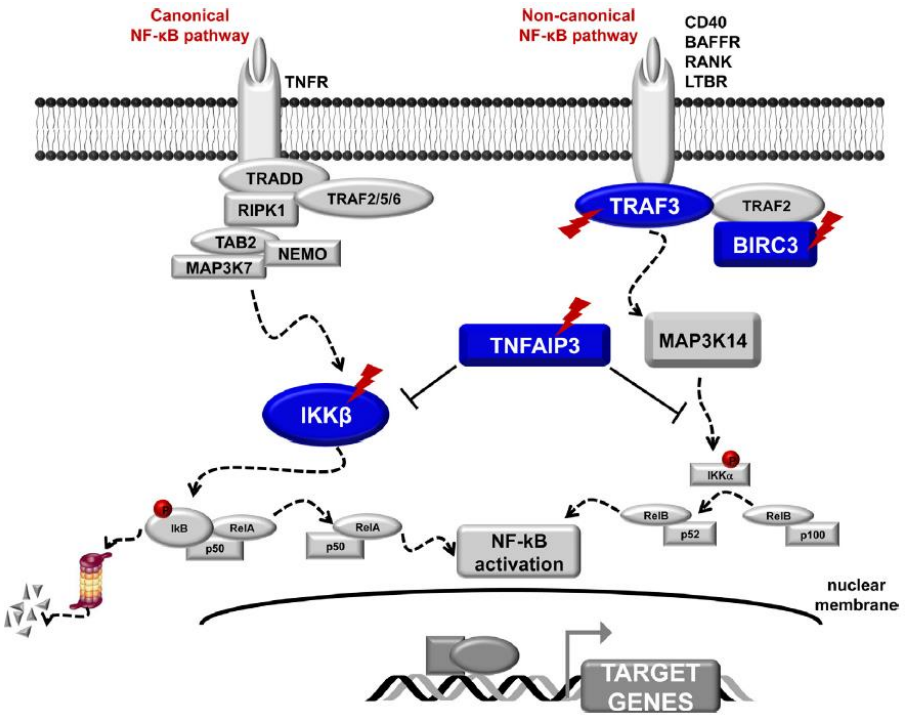


Figure 19. Voie NF-κb.
D'après Rossi *et al*.[120]

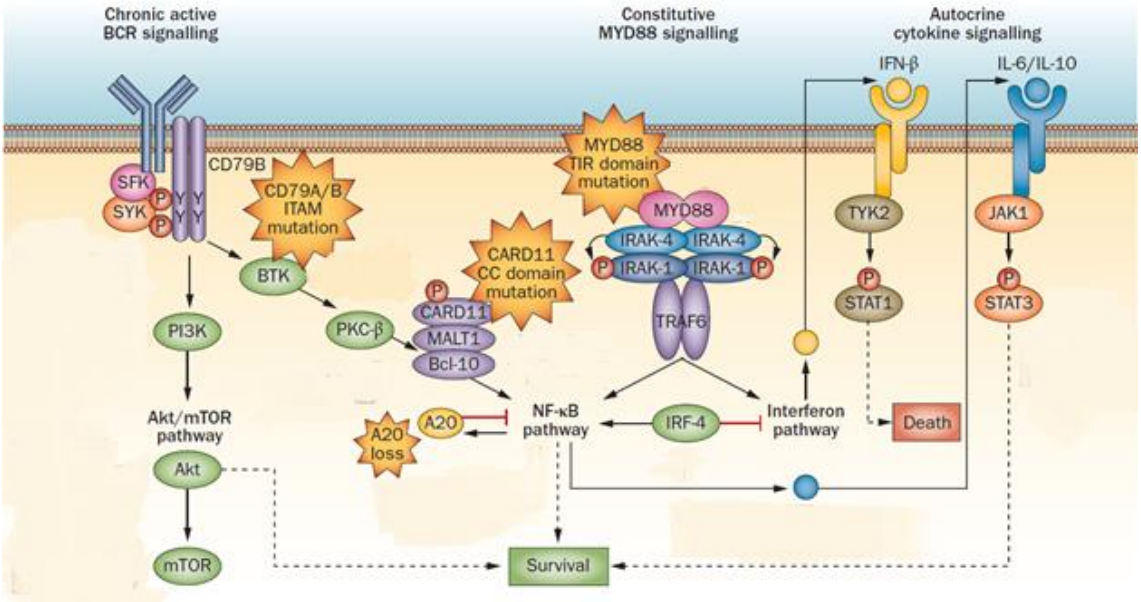


Figure 20. Principales voies de régulation cellulaires impliquées dans la genèse des LBDGC de type ABC.
D'après Roschewski *et al*.[100]

9.1. Mutations entraînant une activation chronique du BCR

Le complexe BCR est composé d'une Ig de surface de type IgM, associée à un hétérodimère CD79A/CD79B. Lorsqu'un antigène se fixe à l'IgM de surface, les résidus tyrosine de la portion cytoplasmique ITAM de CD79A et CD79B sont phosphorylés par les kinases de la famille Src : SYK et LYN. L'activation de SYK par liaison aux résidus tyrosine phosphorylés des domaines ITAM de CD79A et CD79B déclenche une cascade de signalisation impliquant BTK (*Bruton Tyrosine Kinase*). BTK active à son tour la phospholipase PLCγ qui produit alors du DyAcyl-Glycérol (DAG) qui active la protéine kinase Cβ (PKCβ). PKCβ phosphoryle CARD11, ce qui déclenche l'assemblage du complexe multiprotéique CBM, composé de CARD11, BCL10 et MALT1, l'activation d'IKKβ puis de NF-κβ.

L'activation chronique du BCR est caractéristique du sous-type ABC et relève d'un spectre de mutations affectant les différents acteurs de la voie du BCR (**Figure 20**).[120]

9.1.1. Mutations de CD79A et CD79B

Les mutations ponctuelles affectant les résidus tyrosine du domaine ITAM du co-récepteur CD79A/CD79B ont pour conséquence une activation constitutive du BCR et sont identifiées dans 21% des LBDGC de type ABC, alors qu'elles sont rares dans les autres types de lymphome.[120, 123] Ces mutations activatrices ont deux effets : (i) les formes mutées de CD79A et CD79B ne fixent plus la kinase LYN, un régulateur négatif de la voie du BCR, et (ii) les formes mutantes de CD79A et CD79B augmentent l'expression membranaire du BCR par diminution de son internalisation. L'ensemble résulte en l'initiation constitutive de la cascade de transduction impliquant SYK, PKC-β, CBM, IKKβ et NF-κβ.[100, 124, 125]

9.1.2. Complexe CARD11-BCL10-MALT1

Dans environ 10% des LBDGC de type ABC et 3,8% des LBDGC de type GC, l'activation chronique de NF-κβ est due à l'existence de mutations gain de fonction affectant *CARD11*[122, 126].

CARD11 (*Caspase recruitment domain containing protein 11*), BCL10 (*B-cell lymphoma leukaemia 10*) et MALT1 (*Mucosa-Associated Lymphoid Tissue lymphoma translocation protein 1*) forment un complexe qui joue un rôle d'activateur de la voie NF-κβ en aval de la mise en jeu du BCR. CARD11 est la protéine autour de laquelle le complexe s'assemble. A la suite de l'activation du BCR, le domaine inhibiteur de CARD11 est phosphorylé par PKCβ, ce qui donne à

la CARD11 la conformation permettant le recrutement de BCL10 et MALT1 au sein du complexe CBM. Les mutations de *CARD11* décrites dans les LBDGC empêchent le retour de CARD11 à son état inactif et provoquent l'activation chronique de CARD11 indépendamment de l'activation du BCR. Par ailleurs, l'activité protéolytique de MALT1 est indispensable à la survie des LBDGC de type ABC.[127]

9.2. Mutations entraînant une activation chronique du TLR

Dans les lymphocytes B, il existe une voie d'activation indépendante du BCR : la voie du TLR. Le TLR est mis en jeu par la reconnaissance de motifs (lipoprotéines, acides nucléiques) associés aux agents pathogènes, bactéries, virus et champignons.

Après fixation du ligand, les TLR s'associent et recrutent MYD88 (*Myeloid Differentiation primary response 88*) puis IRAK4 et IRAK1, l'activation de TRAF6, l'ubiquitination de la kinase MAP3K7, puis la phosphorylation de IKKβ et enfin l'activation de NF-Kβ **(Figure 21)**.[120]

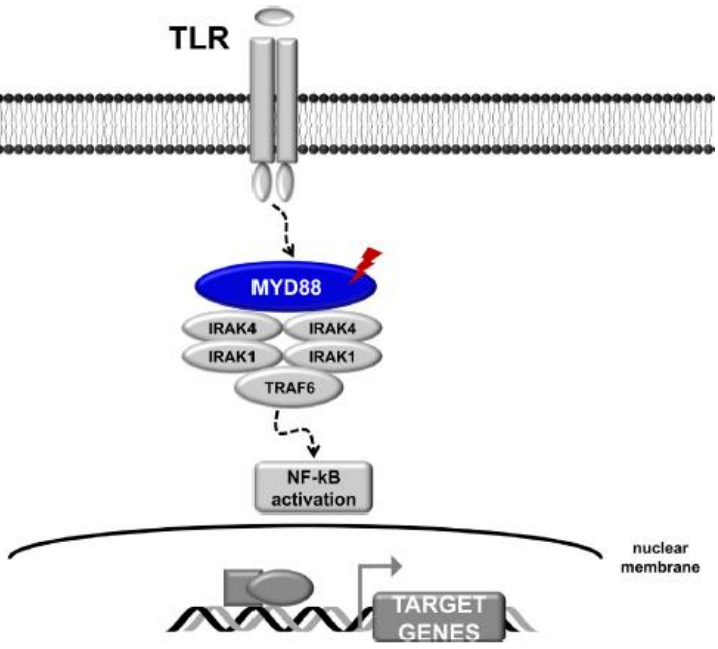


Figure 21. Voie du TLR.
D'après Rossi *et al*.[120]

MYD88 est une protéine adaptatrice permettant la transduction du signal à la suite de l'activation des TLR et du récepteur à l'interleukine 1. L'activation de l'ensemble constitué par la

protéine MYD88 et ses sérine-thréonine kinases associées IRAK (*IL-1-associated Kinases*) 1 et 4 est essentielle pour la survie des LBDGC de type ABC.

Les gains de copies du gène *MYD88* sont observées respectivement dans 38% des LBDGC de type ABC. La mutation gain de fonction *MYD88*^{L265P}, située dans le domaine TIR permettant l'interaction avec le TLR, induit une activation constitutive de la voie TLR et est observée dans 29% des cas de LBDGC de type ABC mais très rarement observée dans le sous-type GC. Cette mutation entraîne l'assemblage du complexe MYD88/IRAK1/IRAK4 et secondairement les voies NF-KB, MAP kinases et JAK-STAT3, ainsi que la synthèse d'IL6, d'IL10 et d'interféron β.[128, 129]

9.3. Délétions bi-alléliques de *TNFAIP3/A20*

TNFAIP3 (*Tumour Necrosis Factor Alpha-Induced Protein 3*) est un régulateur négatif de NF-κB dont la synthèse est engendrée après mise en jeu de TNFα (*Tumor Necrosis Factor alpha*). TNFAIP3 est un régulateur- clé de la réponse immune en ceci qu'il permet d'éviter une activation constitutive de NF-κB (**Figure 20**) en exerçant une fonction d'ubiquitination du récepteur au TNFα, les TLR, le CD40 et la protéine LMP1 (*Latent Membrane Protein 1*).

L'inactivation biallélique de *TNFAIP3* par des mutations somatiques inactivatrices et/ou des délétions bialléliques de *TNFAIP3* en 6p23.3 sont retrouvées dans 32% des LBDGC de type ABC et 2,3% des LBDGC de type GC.[122] Il s'agit de mutations perte de fonction ayant pour effet une inactivation de TNFAIP3 et une activation excessive de la voie NF-κB ce qui entraine une prolifération cellulaire excessive et une résistance à l'apoptose.[130]

9.4. Activation de la voie NOTCH

Les gènes de la famille NOTCH (*NOTCH1 à NOTCH4*) codent pour des protéines hétérodimériques transmembranaires fonctionnant comme des facteurs de transcription après fixation du ligand. Cette fixation entraine le clivage protéolytique de l'ancre membranaire et la partie intracytoplasmique, devenue libre, migre vers le noyau et forme un complexe avec le facteur de transcription RBPJ. En l'absence d'un partenaire de la famille NOTCH, RBPJ agit comme un répresseur transcriptionnel. En revanche, la présence de NOTCH entraîne le départ des répresseurs attachés à RBPJ et le recrutement de co-activateurs comme MAML1, MAML2 et

l'histone acetyl-transferase EP300. Ce complexe ainsi formé permet la transcription de gènes cibles comme *MYC, CCND1, BCL2* et plusieurs composants de la voie NF-κβ (**Figure 22**).[131]

Le principal mécanisme de suppression du signal NOTCH fait appel à sa séquence protéique PEST (*Proline, Glutamic Acid, Serine and Threonine*). Le domaine PEST de la protéine activée est reconnue par l'ubiquitine ligase FBXW7, qui met fin à l'activation de NOTCH par protéolyse. Deux types d'anomalies géniques affectant les gènes *NOTCH1* et *NOTCH2* sont retrouvés dans les LBDGC : (i) un mécanisme d'amplification génique et (ii) des mutations faux-sens, non-sens ou de petites insertions/délétion entrainant un décalage du cadre de lecture, situées principalement dans l'exon 34 du gène (codant pour le domaine PEST). Ces mutations donnent un gain de fonction à la protéine NOTCH puisqu'elles entraînent une diminution de la dégradation de la forme active par le protéasome.[120, 131-134]

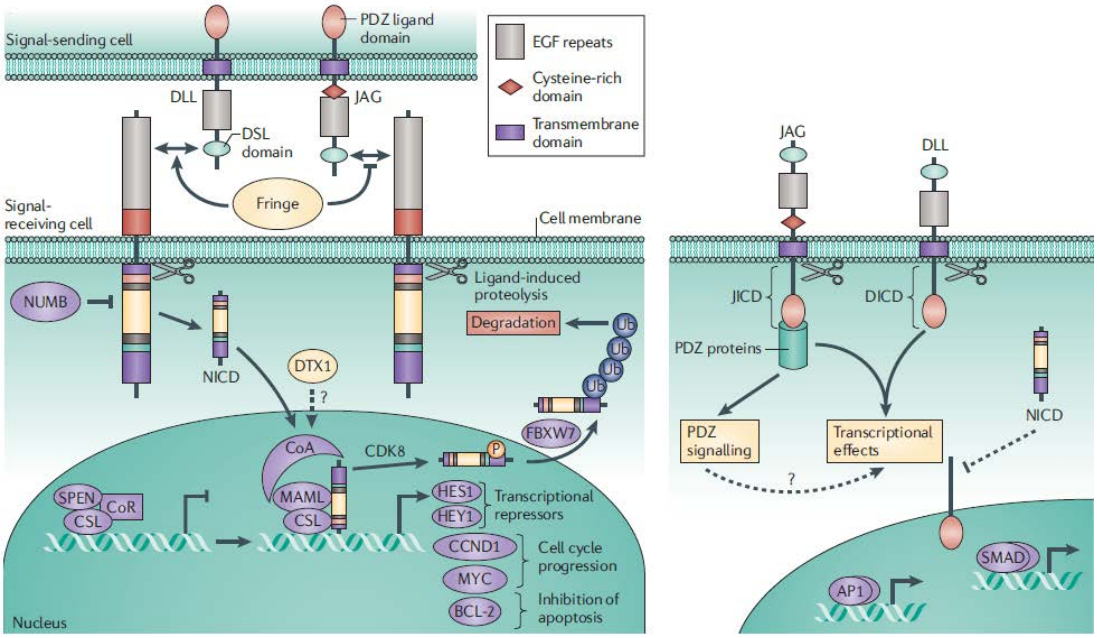


Figure 22. Voie NOTCH.
D'après Ranganathan *et al*.[131]

9.5. Surexpression de la serine-thréonine kinase PIM1

Une augmentation du niveau d'expression de la serine-théonine kinase PIM1 (*Pim-1 proto-oncogène, serine/thréonine kinase*) due à une amplification du locus 6p21 a été décrite dans les LBDGC de type ABC.[61, 135]

10. ANOMALIES MOLECULAIRES AFFECTANT LA REGULATION DU CYCLE CELLULAIRE

10.1. Anomalies du locus CDKN2A/B

Le locus CDKN2A/B (*Cyclin-Dependent Kinase Inhibitor 2A/B*) est situé en 9p21. Le gène *CDKN2A/B* produit au moins trois transcrits alternatifs qui diffèrent par leurs premiers exons et codent trois protéines distinctes. Les deux premières (dont p16/INK4α) possèdent une fonction inhibitrice des kinases CDK (*Cyclin Dependent Kinase*) 4 et 6 qui induisent la poursuite du cycle cellulaire et la division cellulaire. Le troisième transcrit est très différent des deux premiers en raison d'un décalage du cadre de lecture (ARF ou *Alternate open Reading Frame*). Ce troisième variant protéique, p14(ARF), agit comme un stabilisateur de TP53 en séquestrant MDM2 (*Mouse Double Minute 2 homolog*), une protéine responsable de la dégradation de TP53. Les 3 variants protéiques exercent un contrôle du cycle cellulaire, soit lors de la phase G1 du cycle en inhibant CDK4 et 6, soit à la phase G2 *via* TP53 et agissent comme suppresseur de tumeur en empêchant la prolifération cellulaire. Le gène *CDKN2B* code pour p15, un inhibiteur de CDK4 et CDK6 qui permet un contrôle du cycle cellulaire lors de la phase G1 du cycle.

Le locus CDKN2A/B présente des délétions mono- ou biallélique dans 30% des LBDGC de type ABC et 4% des LBDGC de type GC. L'expression de *CDKN2A* est significativement diminuée dans les cas de LBDGC présentant une délétion du locus *CDKN2A/B*.[108, 136]

10.2. Anomalies du facteur de transcription MYC

MYC est un facteur de transcription dont le gène se situe en 8q24. MYC régule la transcription d'environ 15% des gènes, en particulier de gènes impliqués dans la progression du cycle cellulaire. MYC possède aussi des fonctions indépendantes de la prolifération cellulaire : déclenchement l'apoptose en lien avec TP53, rôle dans le renouvellement des cellules souches, la synthèse protéique et le métabolisme.[137] MYC est fréquemment impliqué dans des translocations dans les LNH. Ces translocations mettent le plus souvent le gène *MYC* sous le contrôle du promoteur des gènes *IgH* (*Immunoglobuline Heavy locus*) en14q32, *IgK* (*Immunoglobuline kappa locus*) en 2p11 ou *IgL* (*Immunoglobuline lambda locus*) en 22q11 et ont pour effet la surexpression de *MYC* et une forte prolifération cellulaire. La présence d'une des trois translocations, le plus souvent t(8;14)(q24;q32) mais aussi t(2;8)(p12;q24) ou t(8;22)(q24;q11) sont constantes dans le lymphome de Burkitt, un lymphome particulièrement

agressif en raison d'un fort index mitotique. Alors que ces anomalies sont isolées dans le lymphome de Burkitt, elles sont souvent associées à des caryotypes complexes dans les LBDGC, notamment avec des translocations de *BCL2* et/ou *BCL6*. Par ailleurs, le profil d'expression génique du lymphome de Burkitt est différent de celui des LBDGC.[138]

Deux types d'anomalies de *MYC* ont été décrits dans les LBDGC : les réarrangements géniques et les surexpressions protéiques.

Les réarrangements de *MYC* (MYC+), évalués par FISH interphasique, sont retrouvés dans 7,9 à 11% des LBDGC, majoritairement de sous-type GC. Quel que soit le traitement de première ligne entrepris (CHOP, CHOEP ou R-CHOP), les patients MYC+ ont un plus mauvais pronostic que les patients ne présentant pas de réarrangement de *MYC*.[139-143] Parmi les LBDGC traités par R-CHOP, ceux qui présentent un réarrangement de *MYC* (MYC+) ont une survie globale à 5 ans de 33% alors que ceux qui ne présentent de tel réarrangement ont une survie globale à 5 ans de 72%. En analyse multivariée, les translocations impliquant *MYC* constituent un facteur de risque indépendant de mauvais pronostic. Par ailleurs, les LBDGC MYC+ présentent un risque plus important de rechute au niveau du système nerveux central (17% contre 3%).[140] Les formes MYC+ isolées sont minoritaires (17% des cas) alors que les formes « *double hits* » sont fréquentes avec 54% de LBDGC MYC+/BCL2 et 8% de formes MYC+/BCL6. Les formes « *triple hit* » (MYC+/BCL2/BCL6) représentent 20% des cas de LBDGC MYC+.[142]

Dans les LBDGC en rechute, la proportion de cas MYC+ est de 17%, parmi lesquels 25% de « *simple hit* » et 75% de « *double ou triple hits* » ce qui est concordant avec les biopsies réalisées au diagnostic. Le statut MYC+ est associé à un score IPI plus élevé et à une survie globale diminuée (29% à 4 ans contre 42% pour les formes MYC non réarrangées) et à une mauvaise réponse au traitement de deuxième ligne, qu'il s'agisse de R-DHAP ou de R-ICE. Dans les LBDGC en rechute la notion de « *simple hit* » ou de « *double ou triple hit* » n'avait pas d'impact pronostique.[144]

La surexpression de la protéine MYC est définie par une expression par plus de 40% des cellules de la biopsie, évaluée par immunohistochimie. La surexpression de la protéine MYC est détectée dans 33% des biopsies de LBDGC, ce qui est supérieur à la proportion de cas MYC+ (11% dans l'étude de Johnson *et al.*), et supérieur au nombre de cas présentant une surexpression de l'ARNm MYC (11%), indiquant qu'il existe d'autres mécanismes aboutissant à l'augmentation de l'expression protéique de *MYC*, comme la modification du profil des miARN,[145] ou l'amplification du gène *MYC*.[146] Pour autant, il existe une corrélation entre le statut MYC+ et la proportion de cellules exprimant la protéine MYC (61% [20-100] contre 29% [0-100]). La

présence d'au moins 70% de cellules exprimant la protéine MYC sur la biopsie est systématiquement associée à un statut MYC+.[147] La surexpression protéique de MYC a un impact pronostique négatif si elle est associée à l'expression de la protéine BCL2.[137]

L'étude de Horn *et al*, portait sur les patients de l'essai RICOVER60 (patients âgés de plus de 60 ans, traités par CHOP ou R-CHOP14 en première intention). 31,8% des patients présentaient une surexpression de la protéine MYC et 8,8% présentaient un statut MYC+. 69,2% des échantillons MYC+ présentaient une surexpression de la protéine MYC. 60% des patients MYC présentaient un « *double hit* » ou un « *triple hit* ». Le statut MYC+, la surexpression protéique de MYC, de BCL2 ou de BCL6 constituent des facteurs indépendants de mauvais pronostic (**Figure 23**). Les auteurs ont ainsi pu identifier un sous-groupe de particulièrement mauvais pronostic parmi les patients ayant déjà un score IPI élevé (3-5).[143, 148]

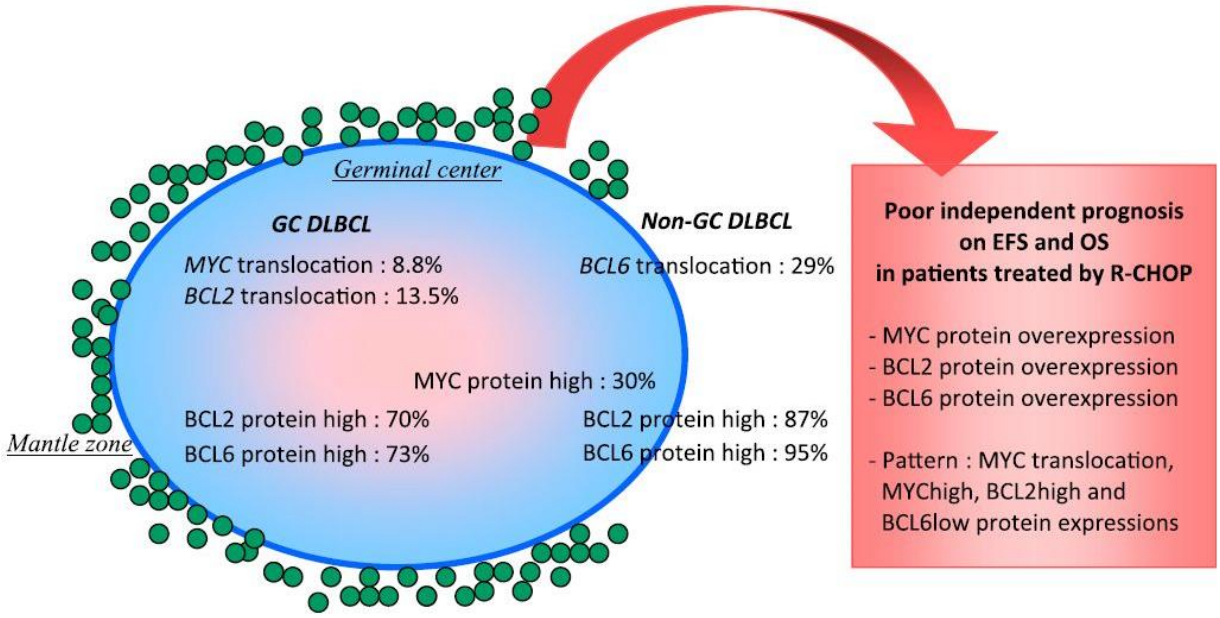


Figure 23. Proportion de réarrangements et de surexpression protéique de MYC, BCL2 et BCL6 dans les LBDGC.
D'après Thieblemont *et al*.[148]

10.3. Anomalies de TP53

Le gène *TP53* (*Tumor Protein p53*) est situé en 17p13. Le principal transcrit est un ARNm de 2586 nucléotides qui code la forme principale de TP53, une protéine de 393 acides aminés. TP53 est exprimé dans tous les tissus et possède une demi-vie de 20 minutes car il est rapidement dégradé par MDM2. Les modifications post-traductionnelles de TP53 (phosphorylation, acétylation et méthylation) activent TP53 et la préserve de la dégradation par MDM2.

La protéine TP53 est un suppresseur de tumeur chargé de préserver la stabilité génomique. Le premier mode d'action de TP53 est transcription-dépendant et régule le niveau d'expression de nombreux gènes impliqués dans le contrôle du cycle cellulaire, la réparation de l'ADN, l'apoptose, les voies de signalisation cellulaires, la transcription et le métabolisme. Le deuxième mode d'action, indépendant de la transcription, provoque directement l'apoptose et l'autophagie (**Figure 24**).

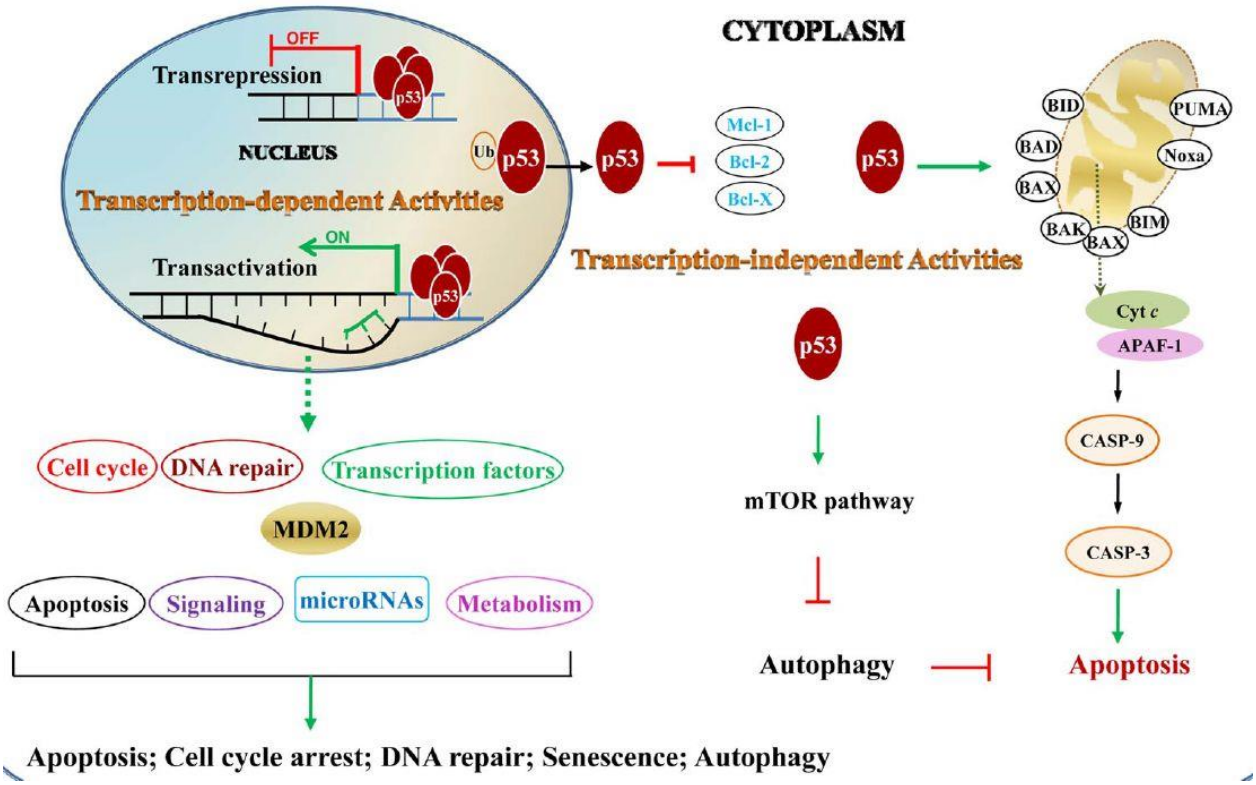


Figure 24. Différents modes d'action de TP53 dans le lymphocyte.
D'après Xu-Monette *et al.*[149]

Des mutations ponctuelles et des délétions monoalléliques aboutissant à la perte de fonction de TP53 sont fréquemment décrites dans les cancers, en particulier les LBDGC et représentent un facteur de mauvais pronostic.[149]

11. ANOMALIE CONDUISANT A L'ECHAPPEMENT A LA REPONSE IMMUNE

Plusieurs anomalies géniques acquises permettent au clone malin d'échapper à la réponse immunitaire T.

Tout d'abord, les inactivations bialléliques de *B2M (Beta-2 Microglobuline)* soit par mutations perte de fonction, soit par délétion focal, ont pour conséquence une diminution de l'expression membranaire des molécules HLA (*Human Leukocyte Antigen)* de classe I et par conséquence, de la reconnaissance de l'antigène par les lymphocytes T cytotoxiques.(150)

Par ailleurs, CD58 est un ligand de la protéine CD2 qui est portée par les lymphocytes T et NK. Les délétions et les mutations non-sens affectant le locus CD58 sont responsables de la synthèse d'une protéine tronquée dans 20% des LBDGC.(150)

Enfin, la sous-expression des molécules HLA de classe II, due à des mutations du gène *CIITA (class II, major histocompatibility complex, transactivator)* sont observés dans 38% des lymphomes B primitifs du médiastin.(151)

PROBLEMATIQUE GENERALE ET OBJECTIFS

Partie 1 : Génomique des syndromes myéloprolifératifs et des syndromes mixtes myéloprolifératifs / myélodysplasiques.

Lorsque ces travaux ont débuté, la reconnaissance de l'ARS-T en tant qu'entité indépendante était contestée. Une première publication, réalisée sur la première grande cohorte d'ARS-T (n=200) a permis de démontrer que sur le plan clinique, biologique et moléculaire, l'ARS-T pouvait être considérée comme une entité distincte de la TE et de l'ARS.[41] Toutefois, la caractérisation moléculaire de l'ARS-T n'était pas définie de manière précise.

Concomitamment, plusieurs groupes ont montré que les SMD avec excès de SC présentent fréquemment une mutation affectant un gène codant pour un acteur de l'épissage alternatif. En particulier, les mutations de *SF3B1* sont identifiées dans 66,7 à 72% des cas d'ARS-T, mais cette proportion a été évaluée sur un petit nombre de cas.[42, 43, 47]

Enfin, deux groupes ont identifié des mutations se situant dans l'exon 9 du gène CALR dans plus de 2 tiers des cas de TE et de MFP ne présentant ni mutation de JAK2 ni mutation de MPL.[19, 20]

Dans ce contexte scientifique, voici les problématiques que nous avons voulu résoudre et les hypothèses que nous avons formulées :

Quelle(s) mutation(s) est (sont) responsable(s) du versant dysplasique de l'ARS-T ? Notre hypothèse est que l'ARS-T présente fréquemment une mutation de *SF3B1* et nous pensons que cela sera confirmé sur notre grande cohorte.

- Quelle(s) mutation(s) est (sont) responsable(s) du versant prolifératif des ARS-T JAK2 non-mutées et MPL non-mutées ? Notre hypothèse est que CALR est un bon candidat.
- Quelle(s) mutation(s) est (sont) associée(s) aux très rares PV JAKwt ? Là aussi, notre hypothèse est que CALR est un bon candidat.
- Quelle est la combinaison de mutations responsables du double caractère myéloprolifératif et myélodysplasique de l'ARS-T ? Notre hypothèse est (i) que le caractère dysplasique de

l'ARS-T est lié dans l'immense majorité des cas à une mutation somatique affectant un régulateur de l'épissage-alternatif, (ii) que le caractère prolifératif est lié dans environ la moitié des cas à une mutation de JAK2, dans environ 5% des cas à une mutation de MPL et dans un grand nombre de cas restants à une mutation de CALR, et (iii) que l'ARS-T relève de cette combinaison mutationnelle qui la définit comme entité pathologique indépendante.

- Quelle la valeur pronostique de ces différentes mutations ? Notre hypothèse est que le statut mutationnel JAK2 et le statut mutationnel SF3B1 ont une valeur pronostique.

Partie 2 : Génomique des lymphomes B-diffus à grandes cellules en rechute.

Les LBDGC représentent 30 à 40% des LNH de l'adulte. Le traitement de référence est fondé sur une immuno-chimiothérapie : le R-CHOP. Ce traitement permet d'obtenir une guérison dans 2 tiers des cas. Par contre, dans 1 tiers des cas, les patients rechutent après ce traitement de première ligne. A la rechute, les traitements de deuxième ligne, ce qui constitue un défi thérapeutique : (i) la moitié de ces patients ne sont pas éligibles au traitement de rattrapage par chimiothérapie puis greffe de CSH ; et (ii) parmi les patients éligibles, seulement 20% seront guéris. Cela signifie que parmi les patients présentant un LBDGC en rechute, 90% décèdent de leur maladie (**Figure 25**).[152]

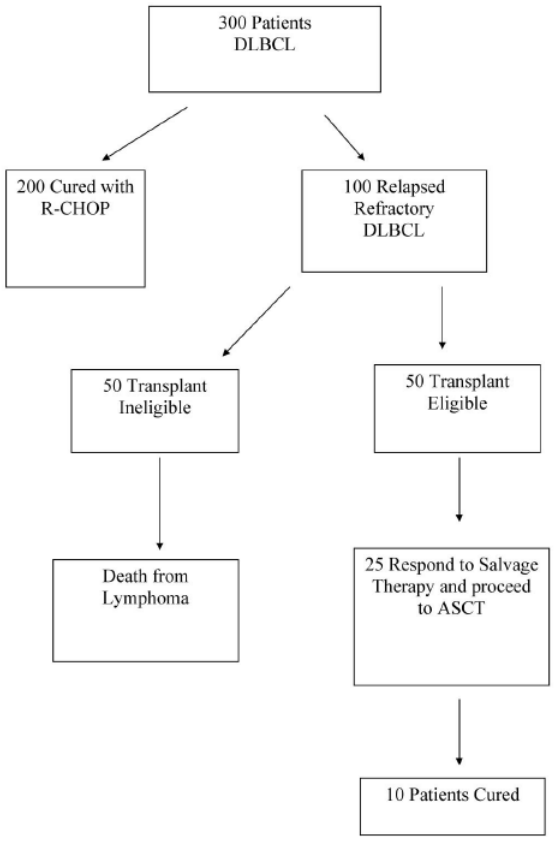


Figure 25 : Devenir des patients atteints de LBDGC.
D'après Friedberg *et al.* [152]

Parmi les patients chez lesquels une réponse complète est obtenue, le délai de la rechute après la fin du traitement est essentiel et a une valeur pronostique. Si la rechute est précoce et survient dans l'année qui suit la fin du traitement de première ligne, le pronostic est similaire à celui des patients réfractaires avec une médiane de survie inférieure à 6 mois.[94, 95] Si la rechute est tardive et survient plus d'un an après la fin du traitement, la survie est plus longue, comparable à celle observée en cas de réponse partielle. La répartition entre rechute précoce et rechute tardive est équilibrée.[93]

L'enjeu actuel est de pouvoir proposer à ces malades un traitement ciblé et nos efforts vont se concentrer sur la compréhension des caractéristiques biologiques de ces formes résistantes de LBDGC.

Nous disposons de 39 échantillons issus d'une cohorte unique, homogène, très bien caractérisée, la cohorte CORAL (*Collaborative Trial in Relapsed Aggressive Lymphoma*), qui est la plus grande cohorte de patients présentant un LBDGC en rechute, réunie à ce jour.[153, 154] Pour ces 39 échantillons, nous disposons de données de SNP (*Single Nucleotide Polymorphism*) et de CNV (*Copy Number variations*), et pour 30 échantillons parmi ces 39, nous disposons des données de transcriptomique.

A l'aide de ces données, nous avons voulu répondre aux questions suivantes :

(i) Quelle est la complexité génomique des LBDGC en rechute ? Est –elle différente si l'on compare rechutes précoces et rechutes tardives ?

(ii) A la rechute, les CNVs identifiés affectent-ils les acteurs de la lymphomagénèse décrits dans la littérature ? Le cas échéant, existe-t-il des anomalies différentielles entre rechutes précoces et rechutes tardives ?

(iii) Si l'on compare les CNV entre rechutes précoces et rechutes tardives, existe-t-il des spectres de CNV spécifiques pour chacun des deux ensembles ? Quelles sont leurs conséquences sur le transcriptome ?

(iv) Quelles sont les fonctions cellulaires affectées par ces anomalies ?

MATERIELS ET METHODES

1. COHORTE CORAL

L'étude CORAL a été bâtie pour mener un essai thérapeutique prospectif dont l'objectif était de comparer l'efficacité des traitements de deuxième ligne des LBDGC réfractaires ou à la première rechute, éligibles à la chimiothérapie intensive de type BEAM associée à la greffe de CSH, c'est-à-dire des sujets jeunes, de 18 à 65 ans, sans comorbidité ni contre-indication à la greffe.[153]

Le premier objectif de cette étude est de comparer les deux protocoles les plus courants de chimiothérapie de rattrapage associée au Rituximab : le R-DHAP (Rituximab, Dexamethasone, Cytarabine haute dose, Cisplatine) et le R-ICE (Rituximab, Ifosfamide, Carboplatine, Etoposide).

Les 396 patients initiaux, âgés de 18 à 65 ans, ont été suivis entre 2003 et 2007. Après la première randomisation, 197 patients ont été inclus dans le bras R-ICE et 191 ont été inclus dans le bras R-DHAP. Au final, 206 patients ont bénéficié de l'ensemble du traitement, c'est-à-dire 3 cures de R-ICE ou de R-DHAP puis un conditionnement par BEAM et une autogreffe de CSH. 196 patients ont fait l'objet d'une deuxième randomisation, 98 d'entre eux ont bénéficié d'un traitement d'entretien par Rituximab, les 98 autres faisant l'objet d'une simple surveillance clinique (**Figure 26**).

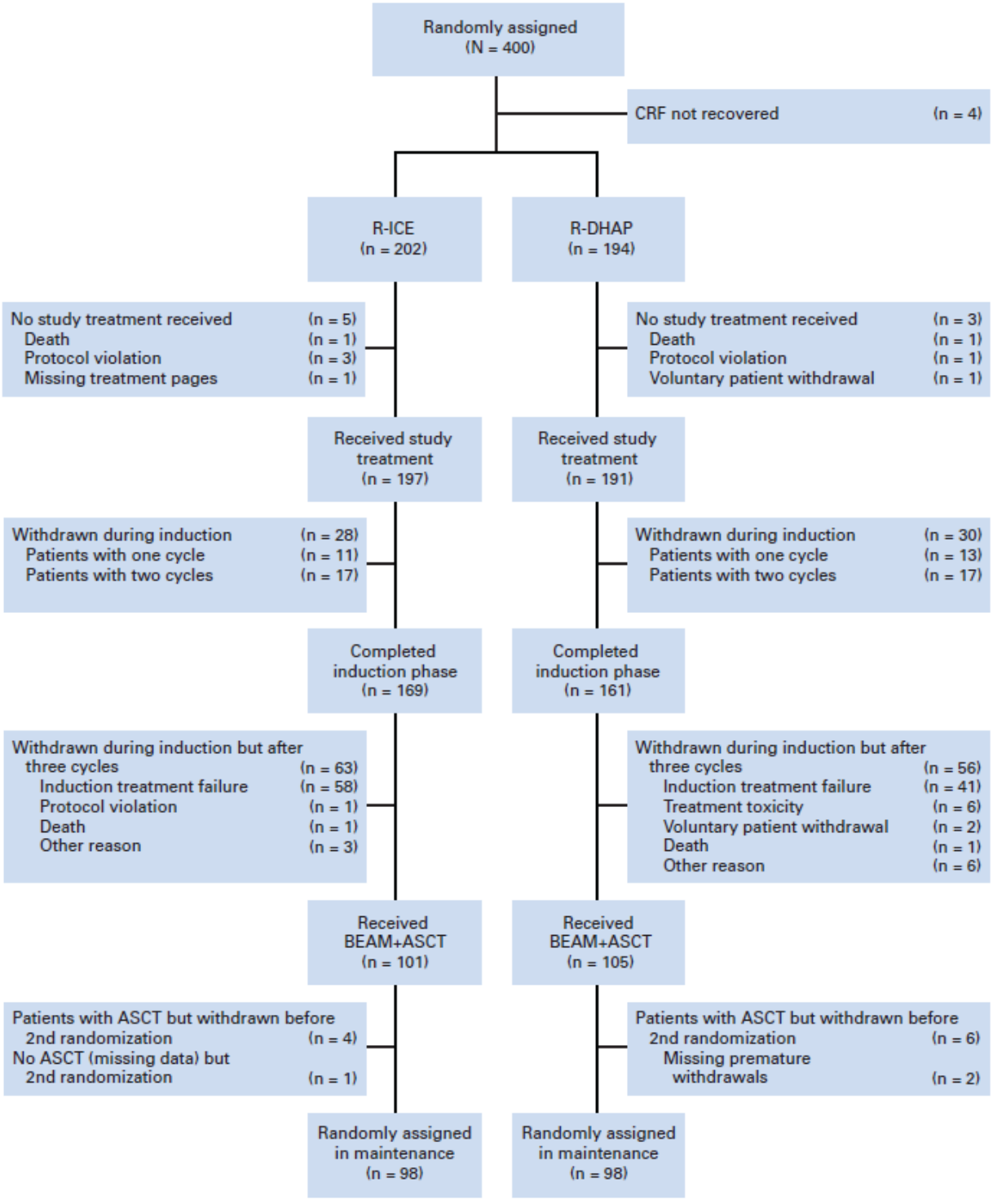


Figure 26. Conception de l'étude CORAL.
D'après Gisselbrecht *et al.* [153])

CRF : Case Report Form; R-ICE : Rituximab, Ifosfamide, Carboplatin, Etoposide; R-DHAP : Rituximab, Dexamethasone, Cytarabine, Cisplatin; BEAM : Carmustine, Etoposide, Cytarabine, Melphalan ; ASCT : Autologous Stem-Cell Transplantation.

Cette étude a démontré que malgré une toxicité plus importante du R-DHAP, les deux protocoles sont équivalents en terme de taux de réponse, de survie sans progression, de survie sans événement et de survie globale, tout ceci à 3 ans (**Figure 27**).

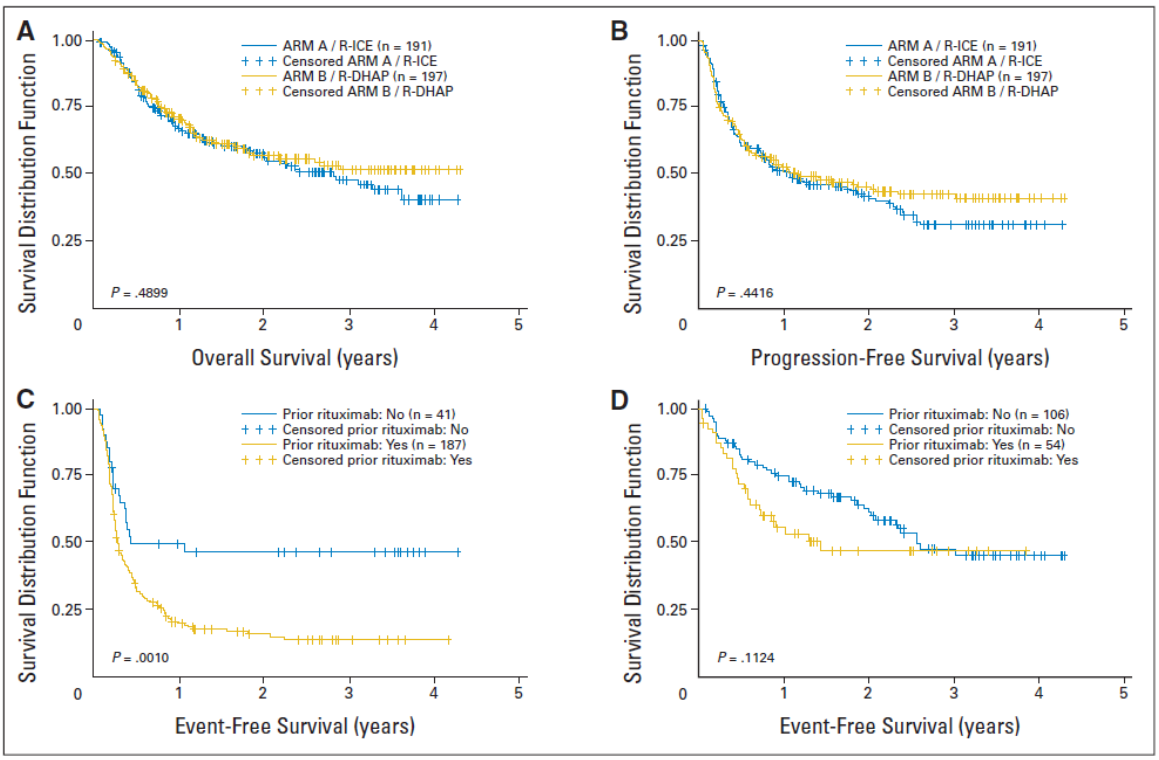


Figure 27. Etude CORAL, survie globale (A) et survie sans progression (B) en fonction du traitement décidé en première randomisation (R-ICE ou R-DHAP).
D'après Gisselbrecht *et al.* [153])

Survie sans événement en fonction de l'adjonction ou non de Rituximab au traitement de première ligne chez les patients ayant rechuté (C) moins de 12 mois ou (D) plus de 12 mois après le début du traitement de première ligne. R-ICE : Rituximab, Ifosfamide, Carboplatin, Etoposide; R-DHAP : Rituximab, Dexamethasone, high-dose cytarabine, Cisplatin.

Cette étude a permis d'identifier plusieurs facteurs pronostiques chez les patients en rechute : l'inclusion du Rituximab dans le traitement de première ligne, le délai de rechute supérieur à 12 mois et un score IPI secondaire de 0 ou 1 sont des facteurs de bon pronostic.

2. L'IMPACT DU STATUT GC/ABC SUR L'EFFICACITE DU TRAITEMENT DE RATTRAPAGE : L'ETUDE BIOCORAL

L'objectif de cette étude était d'évaluer l'impact pronostique du statut GC ou ABC dans les lymphomes B-diffus à grandes cellules en rechute. L'étude Bio-CORAL a démontré la meilleure efficacité du traitement de rattrapage par R-DHAP sur les LBDGC de type GC (**Figure 28**).[154] Le traitement par R-ICE donne des résultats équivalents pour les LBDGC de type GC ou non-GC.

La présence de réarrangements impliquant le gène *MYC*, le phénotype non-GC, l'adjonction de Rituximab au traitement de première ligne, le score IPI ajusté à l'âge et le niveau d'expression de *FOXP1* ont été identifiés comme facteurs de mauvais pronostic.

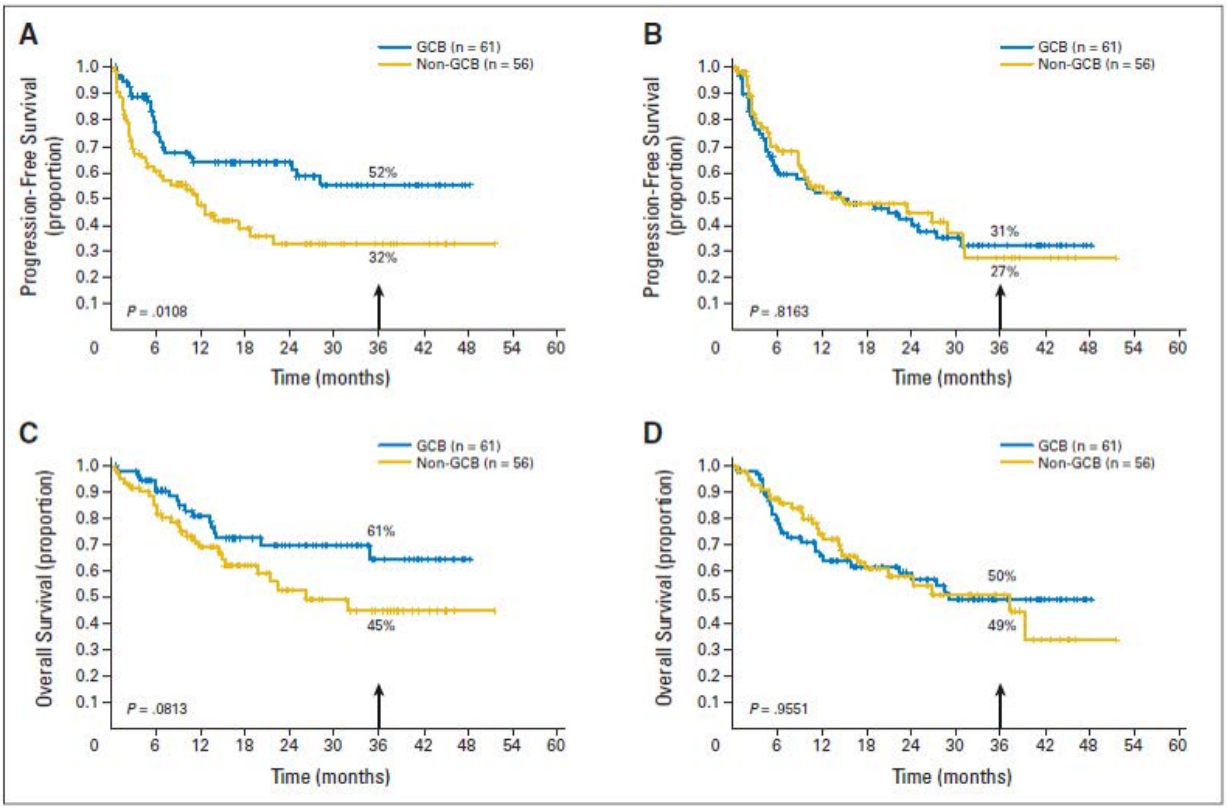


Figure 28. Etude CORAL, survie sans progression et survie globale en fonction du traitement de seconde ligne entrepris : R-DHAP (A,C) ou R-ICE (B,D) et du statut GC/ABC. D'après Thieblemont *et al*. [154]

C'est précisément sur les échantillons issus de cette étude que nous travaillerons pour identifier les processus biologiques à l'origine des mécanismes de résistance des LBDGC.

3. LES PUCES A ADN

3.1. Technologie

Les puces à ADN (*microarray*) sont des réseaux denses (plusieurs millions de sondes) réguliers (*spots* distincts et régulièrement répartis) et ordonnés (on sait ce que l'on fixe et où on le fixe) de sondes d'acides nucléiques déposées sur un support solide, le plus souvent une lame de verre. Les sondes sont des fragments d'acides nucléiques de synthèse, courts (environ 25 nucléotides). Chaque sonde est spécifique d'un gène donné : les séquences sont conçues pour éviter les hybridations croisées (**Figure 29**). La puce présente donc à sa surface, une librairie de séquences nucléotidiques représentatives de gènes.[155, 156]

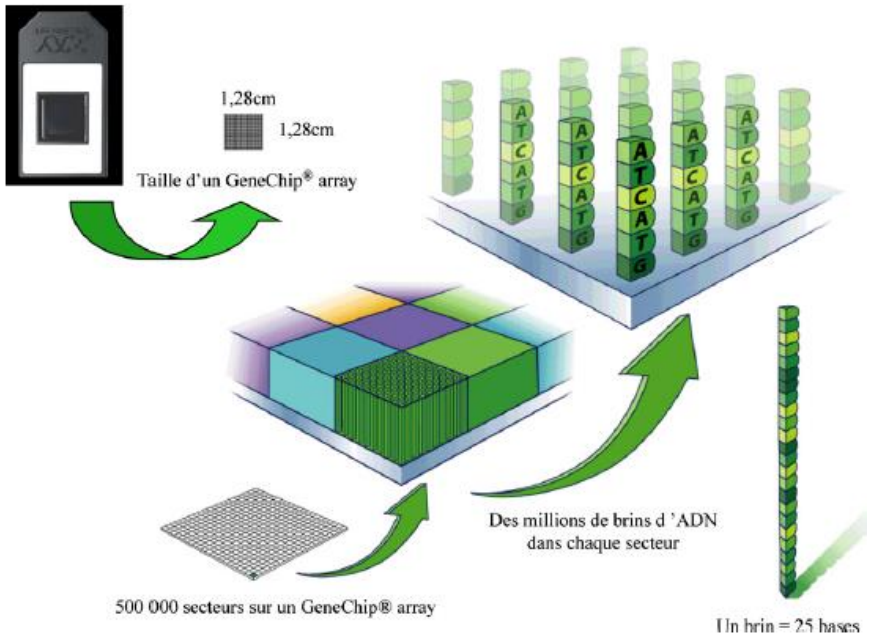


Figure 29. Structure d'une puce GeneChip® commercialisée par Affymetrix. D'après Bogard *et al*.[156]

La technique repose sur le principe d'une hybridation hautement spécifique à grande échelle sur un format réduit, permettant la détection et quantification relative des acides nucléiques par mesure de la fluorescence émise. Les acides nucléiques de l'échantillon (ADNc ou ADN génomique) s'hybrident aux sondes/oligonucléotides présents à la surface de la puce. Or ces derniers sont préalablement marqués par des fluorochromes : la cyanine 3 (qui émet dans le vert) ou la Cyanine 5 (qui émet dans le rouge). Après hybridation, les signaux d'hybridation sont détectés à l'aide d'un scanner de fluorescence. L'utilisation de deux fluorochromes différents

permet la détection simultanée des signaux d'hybridation de deux échantillons distincts au cours d'une même expérience (**Annexe 1**).

3.2. Etude du transcriptome

Une des principales applications de cette technologie est l'analyse du transcriptome, c'est-à-dire la mesure simultanée, à un instant t, de l'ensemble des ARN messagers présents dans un échantillon biologique. La détection et la quantification relative des acides nucléiques est possible grâce à leur marquage par un fluorochrome : le signal mesuré sur chaque sonde est proportionnel au nombre d'ARN messagers et donc à l'expression du gène.

Nous avons utilisé la puce *Agilent Whole Human Genome Microarray 4x44K* (Agilent technologies, Santa Clara, CA, États-Unis d'Amérique). Cette puce à ADN contient 4 emplacements de 44 000 sets de sondes destinées à mesurer le niveau d'expression de 43 370 gènes et transcrits humains.

Le scanner *Agilent* possède une résolution de l'ordre de 1 µm. Le laser excitateur frappe les fluorochromes et les photomultiplicateurs mesurent la fluorescence émise en retour. Les images traitées par un logiciel spécifique et les différentes intensités de fluorescence sont décomposées, mesurées et attribuées aux différents échantillons. Après soustraction du bruit de fond, vérification des contrôles internes et traitement logarithmique du signal, on obtient une matrice mettant en correspondance la liste de l'ensemble des transcrits mesurés et l'intensité de fluorescence correspondante. Les plus souvent, les gènes sont disposés verticalement et les échantillons horizontalement (**Figure 30**).

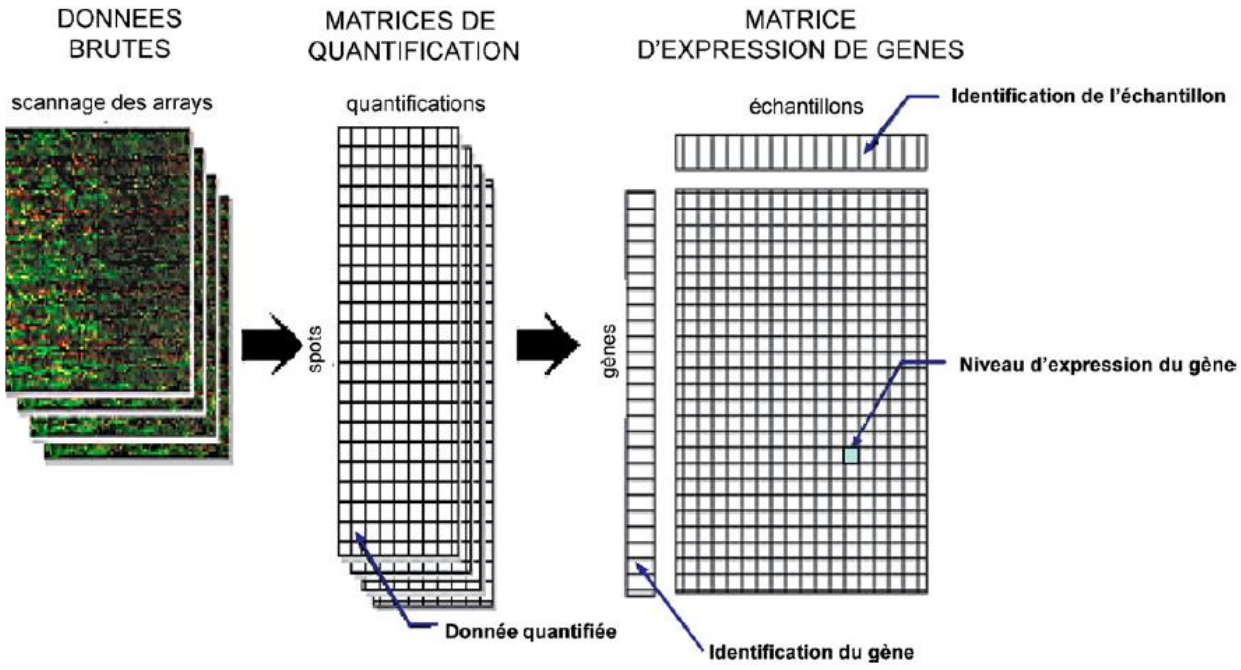


Figure 30. Obtention de la matrice d'expression des gènes.
D'après Bogard *et al*.[156]

3.2.1. Normalisation des données

La normalisation des données est une étape indispensable pour pouvoir comparer les données issues de plusieurs *arrays* en supprimant toute influence non biologique ou biais expérimental : amplification inégale, hybridation inégale, marquage d'efficacité différente (taux d'incorporation des différents fluorochromes), quantité et qualité de l'ARN, intensité différente des fluorochromes (différence de comportement).

Ces méthodes reposent sur deux hypothèses principales: la majorité des gènes n'ont pas d'expression différentielle entre les deux conditions étudiées, et les effets systématiques observés ne sont pas des effets biologiques. La méthode que nous avons utilisé est la méthode du LOWESS (*Locally weight scatterplot smoothing*).[157] Il s'agit d'une méthode de régression pondérée locale permettant d'éliminer les biais non linéaires. Chaque échantillon va être normalisé par rapport à un profil de référence, comme par exemple le profil médian, à l'aide d'une fenêtre glissante. Une fonction de régression est calculée pour chaque fenêtre, ce qui va permettre de construire une courbe d'ajustement global pour normaliser les données (**Figure 31**).

La Log-transformation préalable des données ainsi que leur centrage sur la valeur médiane : log(mesure)-log(médiane) = log(mesure/médiane), permet de donner le même poids à toutes les valeurs, en particulier en diminuant le poids des valeurs extrêmes. Après cette transformation, la distribution des données se rapproche d'une Gaussienne.

Pour permettre une meilleure visualisation d'ensemble et une meilleure compréhension des données, les intensités de fluorescence composant la matrice sont converties en couleurs. Cette matrice de couleurs correspond à une représentation relative du niveau d'expression des gènes comparativement à l'échantillon médian, qui est construit en calculant pour chaque gène la médiane du niveau de fluorescence. Par convention, un transcrit surexprimé par rapport à l'échantillon médian apparaitra en rouge, l'intensité du rouge étant proportionnelle au degré de surexpression. De la même façon, un transcrit sous-exprimé par rapport à l'échantillon médian apparaitra en vert, l'intensité du vert étant proportionnelle au degré de sous-expression.

L'étape suivante est de regrouper les données (gènes et échantillons) par similitude pour regrouper des échantillons ou des gènes ayant un comportement identique. Nous utilisons la méthode des *K-means* et la classification hiérarchique,[158] qui sont des méthodes d'analyse non supervisées.

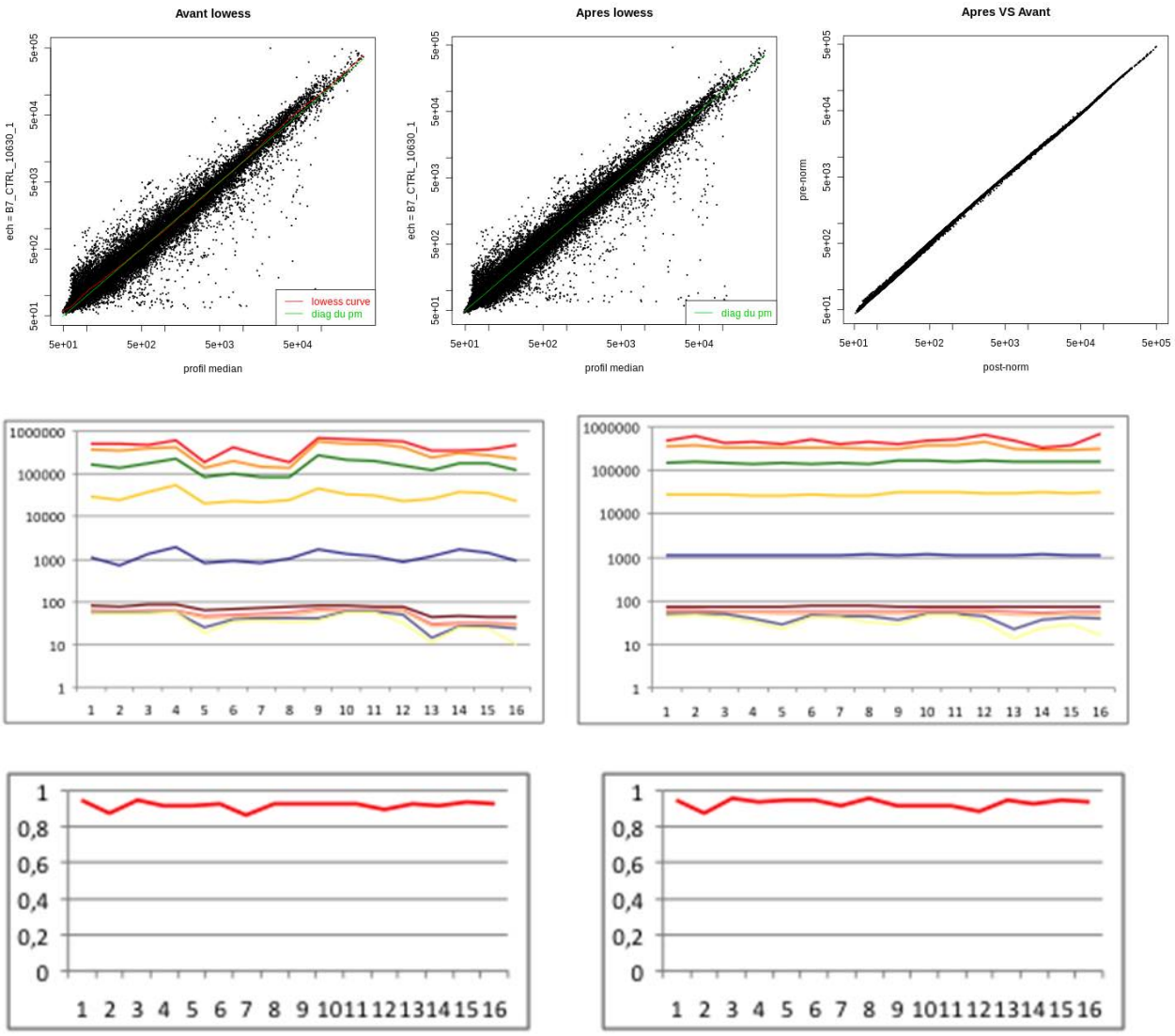


Figure 31. Normalisation. A. Lowess; B. Statistiques descriptives avant et après normalisation; C. Corrélation de chaque échantillon par rapport au profil médian avant et après normalisation.

3.2.2. Regroupement des données selon la méthode du *K-means*

Il s'agit d'une méthode itérative et dynamique qui suppose la définition d'un nombre x de classes parmi lesquelles les échantillons seront classés. Le nombre p de permutations et le nombre de classes sont définis initialement. Les données sont préalablement normalisées, log-transformées et log-centrées. La mesure de similarité utilisée est la corrélation de Pearson. Chaque classe est définie par la valeur médiane des échantillons qui la composent. Les x premières valeurs définissent les x premières classes, puis les valeurs suivantes sont incorporées dans celle des x classes dont elles sont le plus proche du centre de gravité, correspondant à la

valeur médiane. Chaque incorporation d'un nouvel échantillon dans une classe modifie automatiquement sa valeur médiane.

Au final, les gènes sont classés au sein d'une matrice de x classes ou *clusters*. Les clusters d'intérêts sont ensuite étudiés. La classification choisie est celle qui est la plus souvent retrouvée au terme de p permutations (**Figure 32**).

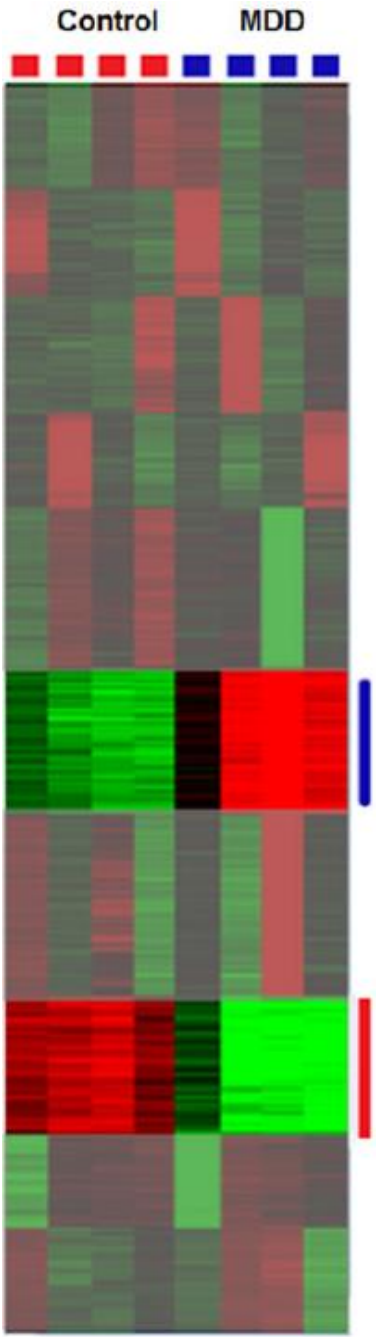


Figure 32. Exemple d'utilisation de la méthode du *K-means*.
D'après Chen *et al*.[159] *MDD : Methyl Donor Deficient*

3.2.3. Regroupement des données selon la méthode de la classification hiérarchique

La classification hiérarchique permet le regroupement des gènes en fonction de leur niveau d'expression et des échantillons selon leur profil d'expression. Le principe de cette méthode est de calculer une distance entre toutes les paires de gènes et d'échantillons possibles afin de regrouper de manière ascendante les gènes et les échantillons les plus similaires entre eux. L'objectif de ce regroupement est d'avoir des distances inter-groupes fortes, et des distances intra-groupes faibles (**Figure 33**).

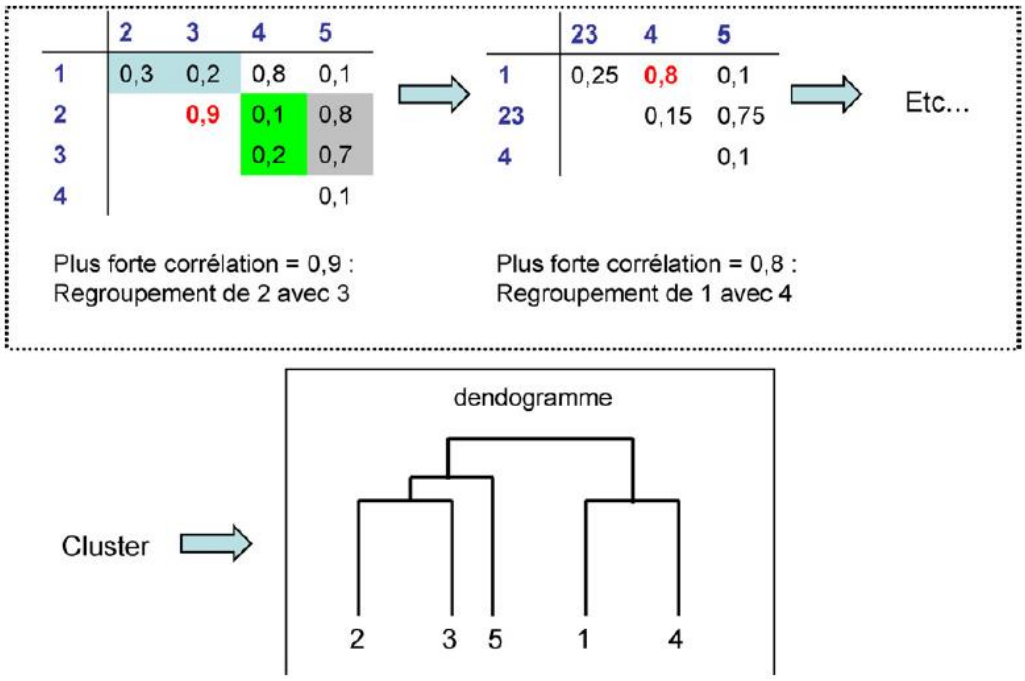


Figure 33. Principe de la classification hiérarchique.
D'après Bogard *et al.*[156]

Il a été démontré que chaque cluster identifié par le *clustering* hiérarchique contient des gènes co-régulés, dont l'expression est corrélée, qui participent à une fonction biologique. Cette méthode permet donc d'identifier des signatures d'expression spécifiques d'un type d'échantillon, offrant un intérêt physiopathologique (**Figure 34**).

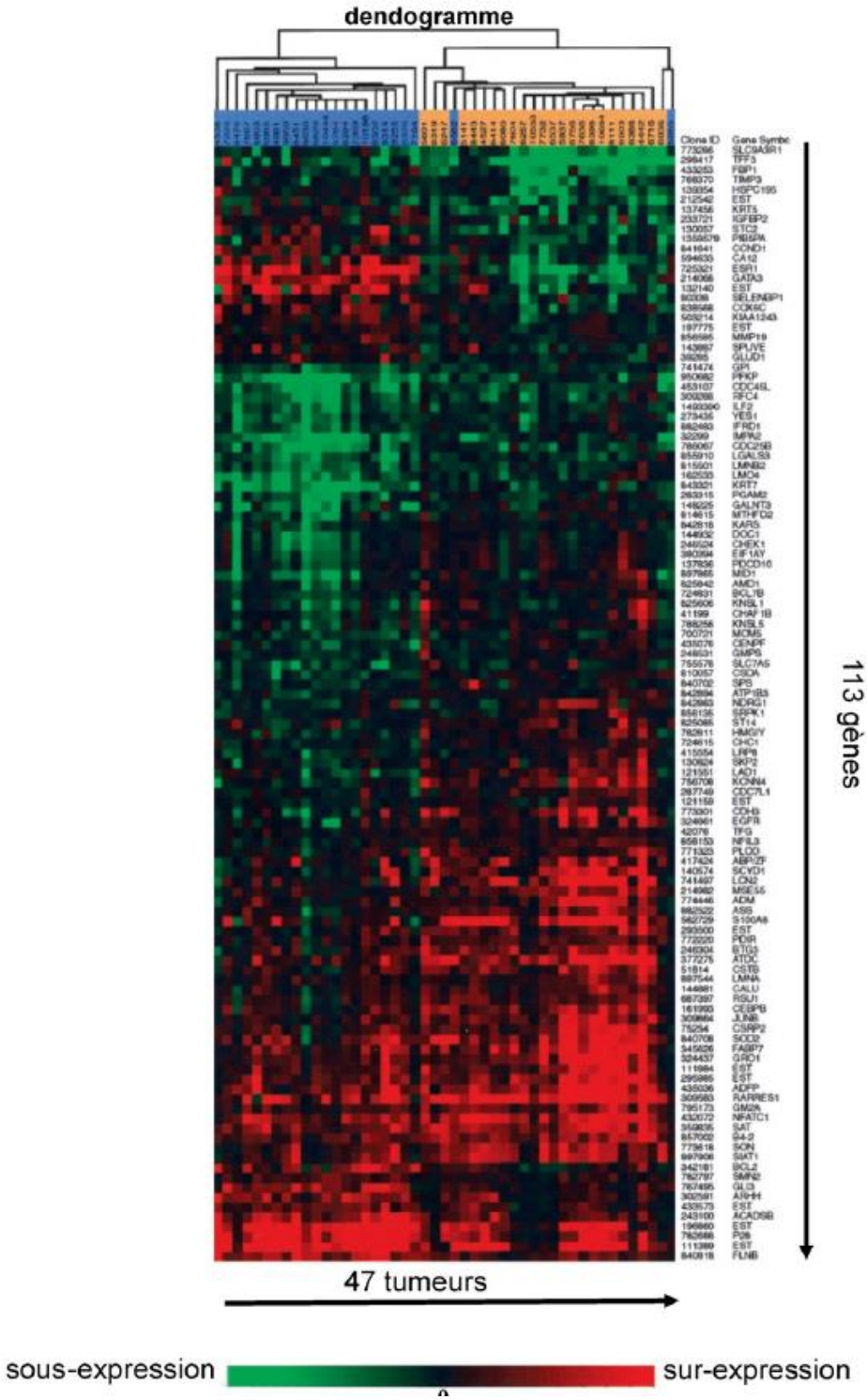


Figure 34. Exemple de résultat obtenu grâce à la méthode de la classification hiérarchique. D'après Bogard *et al*.[156]

3.2.4. Annotation fonctionnelle des clusters

L'étape suivante est d'identifier les fonctions biologiques correspondant aux différents *clusters* de gènes obtenus. Il s'agit de l'annotation fonctionnelle des gènes. Pour cela, il est nécessaire d'utiliser un vocabulaire structuré et contrôlé pour qualifier la fonction des gènes. Le langage utilisé est le *Gene Ontology* (GO).[160] Son objectif est d'établir un vocabulaire structuré, contrôlé (ontologie) décrivant les gènes ou produits de gènes selon le processus biologique, la localisation cellulaire ou les fonctions moléculaires. Processus biologique, fonction moléculaire et composant cellulaire sont les trois catégories de GO. « Processus biologique » se réfère à un objet biologique auquel le produit du gène contribue, « fonction moléculaire » correspond à l'activité biochimique d'un produit du gène et « composant cellulaire » se réfère à la zone géographique de la cellule où le produit du gène est actif (**Figure 35**). Plusieurs outils d'annotation ont été développés à partir des termes GO pour identifier des fonctions biologiques dans un groupe de gènes. Le principe de ces outils est d'identifier des termes GO statistiquement sur- ou sous-représentés dans le groupe de gènes étudié par rapport à l'ensemble des gènes considérés dans l'étude.[161] A partir de la liste des gènes d'un groupe, et de la liste complète des gènes obtenus après normalisation sur la puce à ADN, cet outil fournit une liste d'annotations fonctionnelles, chacune associée à une p-valeur. En plus de cette p-valeur, nous effectuons un calcul d'enrichissement, afin de tenir compte du nombre de gènes ayant l'annotation considérée. En effet, selon la fonction, le nombre de gènes concernés peut être très variable et le calcul de l'enrichissement permet en plus de savoir si l'annotation fonctionnelle est sur- ou sous-représentée dans le cluster, ce que n'indique pas la p-valeur.

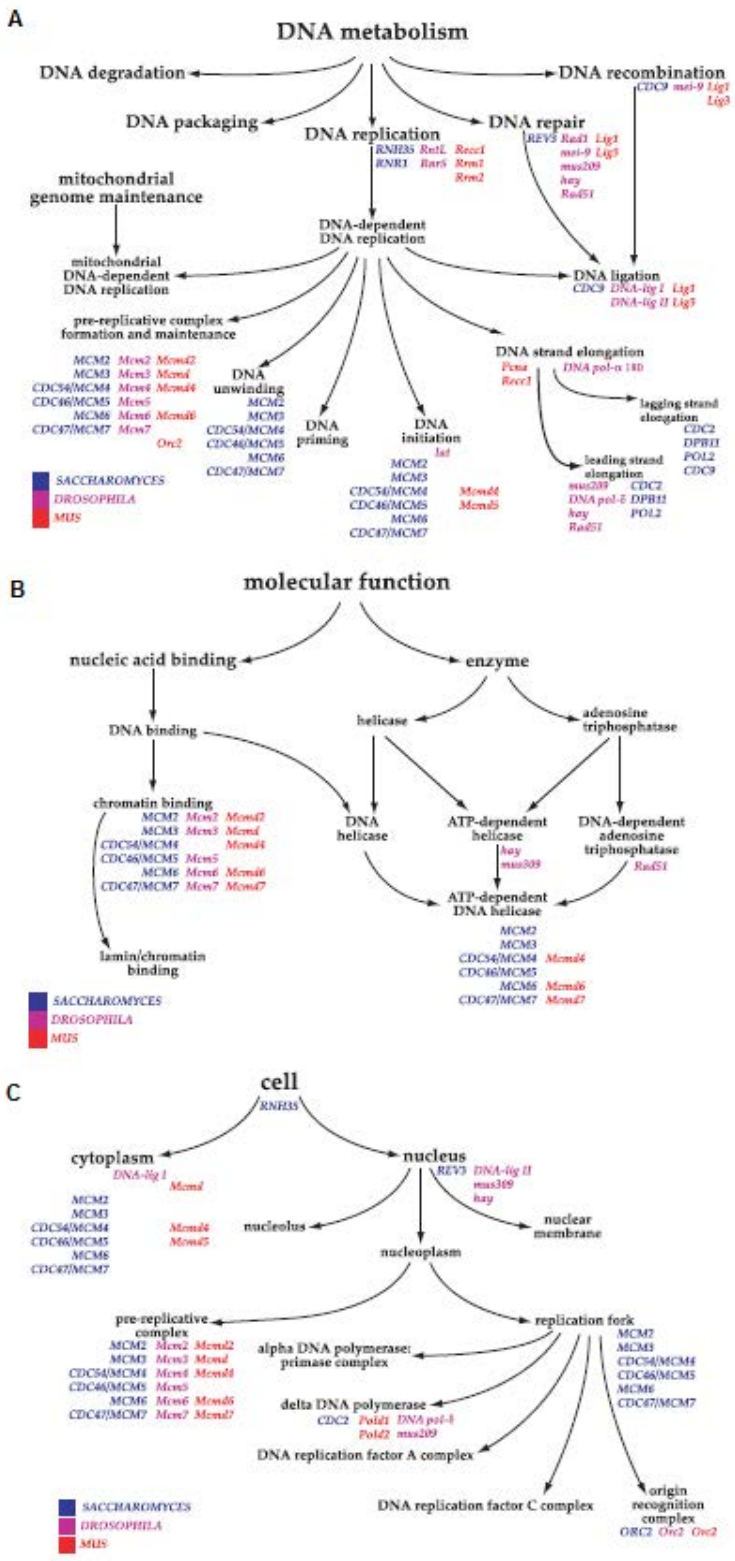


Figure 35. Exemples d'ontologies des gènes.

3.3. Etude des anomalies du nombre de copies géniques et des polymorphismes de l'ADN

Une autre application des puces est l'étude de l'ADN. L'ensemble du génome est couvert au moyen de sondes régulièrement réparties tout au long du génome. Il est possible d'hybrider de l'ADN génomique pour déterminer l'existence d'anomalies de nombre de copies : *Copy Number Variations* (CNV). Cela permet d'identifier des délétions ou des duplications de régions chromosomiques ou même de chromosomes entiers. Après hybridation, l'intensité du signal émis par chaque sonde est proportionnelle au nombre de copies. Un signal diminué, correspondant à une seule copie sur un grand nombre de sondes consécutives, signe une délétion d'une copie du gène.

Il est possible d'analyser le polymorphisme de la séquence d'ADN, c'est-à-dire de faire le génotypage de plusieurs centaines de milliers de variants connus.

3.3.1. Puce Affymetrix SNP 6.0

Cette puce à ADN contient plus d' 1 800 000 sondes, incluant (**Tableau XII**):

(i) 906 600 sondes polymorphes dédiées à l'analyse des polymorphismes simple base (*Single Nucleotide Polymorphism* – SNP). Chacune de ces sondes permet d'explorer pour chaque SNP les deux génotypes les plus fréquents.

(ii) 946 000 sondes non-polymorphes s'ajoutant aux premières pour détecter les anomalies de nombre de copies. Parmi ces sondes, 744 000 ont été sélectionnées pour leur couverture du génome et 202 000 à titre de sondes-test permettant d'explorer des CNVs connus et répertoriés dans le Toronto *Database of Genomic Variants* (DGV).

Ainsi, la distance moyenne entre deux sondes consécutives est de 700 bases, ce qui permet une analyse haute résolution par *SNP-array*. Les performances de la puce Affymetrix *Genome-wide Human SNP Array* 6.0 ont été validées sur 270 échantillons du projet HapMap.[162]

3.3.2. L'algorithme CRLMM

CRLMM ou "*Genotype Calling and copy number analysis tool for Affymetrix SNP 5.0 and 6.0 and Illumina arrays*" [163] est un algorithme opérant sous l'environnement R et permettant d'extraire les valeurs brutes de nombre de copie(s) pour chacun des 1 800 000 sondes de la puce Affymetrix SNP 6.0, organisées par chromosome.

Cet algorithme nous a permis d'obtenir pour chaque chromosome une matrice contenant un échantillon par colonne et une sonde par ligne. Les sondes sont organisées du haut vers le bas consécutivement de l'extrémité du bras p vers l'extrémité du bras q. Chaque case contient une valeur du nombre total de copies relatif à un échantillon et à une sonde. Les valeurs du nombre de copies sont réparties de 0 à 10.

Chromosome	Nombre de sondes	Chromosome	Nombre de sondes
1	144293	12	86044
2	151642	13	65213
3	126136	14	56234
4	118766	15	52728
5	114147	16	53237
6	111261	17	45903
7	99460	18	51409
8	96857	19	29859
9	80911	20	43000
10	92199	21	24723
11	88180	22	23894
Total	1 756 096		

Tableau XII. Puce Affymetrix *Genome-wide Human SNP Array 6.0* nombre de sondes couvrant chaque chromosome.

Le fabricant propose pour chaque sonde une annotation selon la version Hg19 du génome humain. A savoir pour chacune des sondes : (i) un identifiant Affymetrix unique, (ii) sa position sur le chromosome en partant de l'extrémité du bras p, (iii) sa situation ; exon, intron, aval, amont ou

extrémité 3' UTR d'un gène, (iv) sa distance au gène le plus proche si elle n'est pas située dans un exon ou un intron, (v) son identifiant ENST, (vi) le gène dans lequel elle se situe ou dont elle est le plus proche.

Nous avons interrogé la base de données *ENSEMBL Genome Browser* (http://www.ensembl.org/) à l'aide des identifiants ENST pour obtenir l'annotation la plus récente pour chaque sonde. Lorsque l'annotation Affymetrix était satisfaisante, nous l'avons conservée. Dans le cas contraire, nous avons adopté l'annotation la plus récente proposée par *ENSEMBL*.

3.3.3. Concept de fenêtre glissante

Un premier examen des valeurs brutes de nombre de copies données par l'algorithme CRLMM a fait apparaitre deux principaux problèmes :

L'espacement moyen entre deux sondes consécutives est de 700 bases. Malgré cela, les valeurs de nombre de copies données par deux sondes consécutives pouvaient être très différentes. Il nous fallait donc calculer un paramètre permettant d'examiner les CNV sur une zone plus large.

La répartition des valeurs de nombre total de copies données par CRLMM se situant entre 0 et 10. Ainsi, nous avions un biais en faveur des gains puisque un certain nombre de valeurs de nombre de copies étaient supérieures à 4.

Pour répondre à ces deux problèmes, nous avons choisi de calculer pour chaque sonde la médiane des 11 sondes consécutives centrées sur cette dernière, à savoir la médiane des 5 sondes d'amont, de la sonde centrale et des 5 sondes d'aval.

Pour chaque sonde de la puce, la valeur brute de nombre de copies était remplacée par la médiane des valeurs des 11 sondes consécutives centrées sur cette sonde. Cette fenêtre glissante était décalée de sonde en sonde. Ceci permet de disposer d'une valeur de nombre de copies qui soit représentative d'une zone plus large (environ 7 kb) mais aussi de supprimer le biais dû aux valeurs comprises entre 4 et 10 puisque la médiane est peu sensible aux valeurs extrêmes.

L'utilisation d'une fenêtre glissante de 11 sondes consécutives permet d'étudier les anomalies de nombre de copies à l'échelle du gène.

Lorsque nous avons voulu faire l'inventaire des anomalies de grande taille, nous avons utilisé une fenêtre glissante d'une taille supérieure, 300 sondes consécutives, pour fusionner les événements de petite taille.

3.3.4. Définition des CNV

Nous avons choisi de définir les différentes anomalies de nombre de copies en nous inspirant d'une méthode publiée[108] : (i) entre 1,5 et 2,5 : nombre normal de copies, (ii) au-delà de 2,5 copies : gain ; (iii) en dessous de 1,5 copie : perte.

RESULTATS

PARTIE 1 : GENOMIQUE DES SYNDROMES MYELOPROLIFERATIFS ET DES SYNDROMES MIXTES MYELOPROLIFERATIFS / MYELODYSPLASIQUES

1. AGE, JAK2^{V617F} AND *SF3B1* MUTATIONS ARE THE MAIN PREDICTING FACTORS FOR SURVIVAL IN REFRACTORY ANAEMIA WITH RING SIDEROBLASTS AND MARKED THROMBOCYTOSIS

Contexte

L'anémie réfractaire sidéroblastique avec thrombocytose est une entité provisoire dans la classification 2008 de l'OMS. Il s'agit d'une pathologie rare, appartenant à la famille des syndromes mixtes myéloprolifératifs/myélodysplasiques, partageant à la fois la présentation dysplasique de l'ARS et le caractère prolifératif de la TE. Nous avons montré que cette entité peut prétendre au statut de pathologie indépendante compte-tenu de ses particularités biologiques, moléculaires et pronostiques. En particulier, la proportion de mutations *JAK2*^{V617F} est de 42,9% dans l'ARS-T, alors qu'elle est de moins de 5% dans l'ARS, et de 70% dans la TE.[41]

Notre hypothèse est que l'ARS-T est une entité indépendante, caractérisée sur le plan moléculaire par la co-occurrence d'une mutation donnant un avantage prolifératif et d'une mutation responsable du caractère dysplasique. Cette hypothèse s'est trouvée renforcée par la démonstration d'une forte proportion de mutations de gènes codant pour des protéines impliquées dans l'épissage alternatif, dans les SMD avec sidéroblastes en couronne, en particulier *SF3B1*. Ainsi, le gène *SF3B1* est muté dans 64 à 82,6% des ARS[42-44] et 72% des ARS-T.[46] Ces mutations sont rares dans la TE (0-3%). L'ARS-T associerait donc une forte proportion de mutations *JAK2*^{V617F} et de mutations de *SF3B1*, confirmant son caractère indépendant de l'ARS et de la TE.[164]

Objectifs

Le premier objectif de ce travail multicentrique européen est d'évaluer la fréquence des mutations de *SF3B1* dans une grande cohorte de patients atteints d'ARS-T, ainsi que d'explorer l'impact pronostique de ces mutations.

Le second objectif est de mieux caractériser l'ARS-T sur le plan moléculaire.

Méthodologie

Grâce à un recrutement à l'échelle européenne, nous avons réuni une cohorte de 111 malades réunissant les critères OMS 2008 définissant l'ARS-T. Les mutations de *SF3B1* ont été recherchées par séquençage haut débit, avec une sensibilité < 1%). La recherche de la mutation *JAK2*^{V617F} a été recherchée par PCR allèle-spécifique avec une sensibilité de 1%.[22] La recherche des mutations situées dans l'exon 12 de *JAK2* et les mutations de MPL ont aussi été recherchées.[165, 166]

Principaux résultats

Nous avons réuni 111 ARS-T et 33 ARS. Les mutations de *SF3B1* sont retrouvées dans 96/111 cas d'ARS-T (86,5%) et 28/33 (84,8%) ARS. La plupart des mutations de *SF3B1* se situent dans les exons 14 et 15 du gène. Cette étude confirme par ailleurs la forte proportion de mutations *JAK2*^{V617F} dans l'ARS-T (48,6%) et la faible fréquence de cette mutation dans les ARS (0%).

Parmi les ARS-T, les patients SF3B1 mutés avaient une médiane de survie de 6,9 ans contre 3,3 pour ceux qui ne portaient pas la mutation (p = 0,003). L'analyse multivariée révèle 3 facteurs de mauvais pronostic : un âge > 80 ans, le statut *SF3B1* non-muté et le statut *JAK2*^{WT}. Nous avons bâti un score pronostique accordant un point à chacun de ces 3 facteurs de risque, et permettant de différencier 2 groupes pronostiques différents. Les patients de faible risque (score à 0 ou 1) ont une médiane de survie de 8 ans, alors que les patients de haut risque (score 2 ou 3) ont une médiane de survie de 1,6 an.

Conclusion

Jusqu'à présent les études portaient sur de petites séries d'ARS-T. Notre cohorte européenne est la plus importante, avec 111 ARS-T. Notre étude confirme la forte proportion de mutations de *SF3B1* dans les ARS-T. Si une mutation affectant un des acteurs de l'épissage alternatif est présente dans presque tous les cas, expliquant le caractère dysplasique de la maladie, la proportion de mutations *JAK2*^{V617F} n'est que de 50%. L'enjeu à présent, est de rechercher la ou les mutations responsables du caractère myéloprolifératif dans les ARS-T *JAK2*non-mutées.

Age, *JAK2*^{V617F} and *SF3B1* mutations are the main predicting factors for survival in Refractory Anaemia with Ring Sideroblasts and marked Thrombocytosis

Running title: Impact of *SF3B1* mutations on the prognosis of RARS-T.

Julien Broséus[1] (MD), Tamara Alpermann[2], Michael Wulfert[3] (PhD), Lourdes Florensa Brichs[4*] (MD), Sabine Jeromin[2] (PhD), Eric Lippert[5] (MD, PhD), Maria Rozman[6*], Francois Lifermann[7] (MD), Vera Grossmann[2] (PhD), Torsten Haferlach[2] (MD), Ulrich Germing[3] (MD), Elisa Luño[8*] (MD), Francois Girodon[9] (MD, PhD) and Susanne Schnittger[2] (PhD) for the MPN&MPNr-EuroNet (COST Action BM0902)

[1]Haematology Laboratory, University Hospital, Nancy, France; [2]MLL Munich Leukemia Laboratory, Munich, Germany; [3]Departement of Haematology, Oncology and Clinical Immunology, Heinrich-Heine-Universität, Düsseldorf, Germany; [4]Laboratorio Citologia Hematològica, Servicio Patologia, Hospital del Mar, Barcelona, Spain; [5]Haematology laboratory, University Hospital, Bordeaux, France; [6]Unidad de Hematopatología, Departamento de Patología, Hospital Clínic, IDIBAPS, Barcelona, Spain; [7]Department of Internal Medicine, Hospital of Dax, France; [8]Servicio de Hematología, Servicio de Salud del Principado de Asturias, Oviedo, Spain; [9]Haematology Laboratory, University Hospital, Dijon, France.

(*) On behalf of the Spanish Group of Hematological Cytology (GECH)

Corresponding authors:

Pr François Girodon
Haematology laboratory, University Hospital, Dijon, France
Plateau technique de biologie, 2 rue Angélique Ducoudray 21079 Dijon cedex
Tel: (+33) 3 80 29 57 10 ; Fax: (+33) 3 80 29 36 60
francois.girodon@chu-dijon.fr

PD Dr. rer. nat. Susanne Schnittger
MLL Munich Leukemia Laboratory
Tel: (+49) 89-990-17-300 ; Fax: (+49) 89-990-17-309
susanne.schnittger@mll.com

Grants and supports to disclose:

This work was supported by grants from the association "Tulipes contre le cancer" (Châlon s/Saône, Burgundy, France) and from FEHH (Spain) and 2009 SGR 541 (Generalitat de Catalunya).

Keywords: refractory anaemia with ring sideroblasts and marked thrombocytosis, *SF3B1*, prognostic impact, survival.

Abstract:

Refractory Anaemia with Ring Sideroblasts and marked Thrombocytosis (RARS-T) is a provisional entity in the WHO 2008 classification and has previously been shown to have a high proportion of *JAK2*^{V617F} and *SF3B1* mutations. The purpose of the present study was to analyse the frequency of *SF3B1* mutations in a large cohort of 111 patients with RARS-T and 33 patients with RARS and to explore the prognostic impact of *SF3B1* mutational status on RARS-T.

The frequency of *SF3B1* mutations in RARS-T (96/111, 86.5%) and RARS (28/33, 84.8%) was similar. In RARS-T, median survival was better in *SF3B1* mutated patients than in *SF3B1* non-mutated patients (6.9 and 3.3 years respectively, p= 0.003). RARS can be differentiated from RARS-T by the frequency of *JAK2*^{V617F} (0% vs. 48.6%). In RARS-T patients, *SF3B1* mutations (p=0.021) and *JAK2* mutations (p=0.016) were independent factors for a better prognosis. Altogether, our results confirm that RARS-T is an independent entity that should be recognized by the next WHO classification. The assessment of *SF3B1* mutations is of prognostic interest in RARS-T patients. Younger age, *JAK2*^{V617F} and *SF3B1* mutations are the main predicting factors for survival in RARS-T.

Introduction:

Refractory Anaemia with Ring Sideroblasts (RARS) and marked Thrombocytosis (RARS-T) has been proposed in the WHO 2001 classification of tumours of haematopoietic and lymphoid tissues and retained as a provisional entity in the classification of 2008.[1] RARS-T is classified in the myeloproliferative/myelodysplastic (MDS/MPN) disorders group because it presents with the dysplastic features of RARS[2] and the myeloproliferative features of Essential Thrombocythemia (ET).[3] RARS-T is characterized by a high rate of *JAK2*^{V617F} (*Janus Kinase 2*) mutations[4-15] and/or the presence of the mutation *MPL*$^{W515L/R}$ (*MyeloProliferative Leukemia*).[16, 17] The classification of RARS-T as an entity that is independent from RARS or ET is currently a matter of debate. Several specialists favour the hypothesis that RARS-T is a form of ET with more than 15% of ring sideroblasts in the bone marrow[18] while others think that RARS-T develops from RARS with secondary thrombocytosis accompanying the acquisition of the *JAK2*^{V617F} mutation. [19]

Recent publication from our group in a European retrospective multicentre collaborative study demonstrated that RARS-T was independent from RARS and ET from a clinical and biological as well as prognostic point of view.[20]

Our results have recently been strengthened by the discovery of the association between myelodysplastic syndromes (MDS) and mutations involving components of the RNA splicing

machinery, including *U2AF35* (*U2 small nuclear RNA Auxiliary Factor 35*), *ZRSR2* (*Zinc finger CCCH type, RNA-binding motif and Serine/aRginine rich 2*), *SRSF2* (*Serine/aRginine-rich Splicing Factor 2*) and *SF3B1 (Splicing Factor 3B subunit 1)*. *SF3B1* mutations (*SF3B1*^{mut}) are found in about 20% of total MDS, and correlate strongly with the presence of more than 15% of ring sideroblasts (MDS-RS) (64 to 82.6% in RARS, 57 to 76% in Refractory Cytopenia with Multilineage Dysplasia and Ring Sideroblasts (RCMD-RS) and 66.7 to 72% in RARS-T).[21-27] On the other hand, mutations of *SF3B1* are found at a lower frequency in MDS with less than 15% ring sideroblasts, which confirms the specificity of *SF3B1*^{mut} in MDS-RS. *SF3B1* mutations are rare in myeloproliferative neoplasms (MPN) and particularly in ET (0 to 3%).[22, 28] The high frequency of *SF3B1* mutations suggests that these mutations play a pathophysiological role in these diseases, probably through perturbations of RNA splicing. The link between *SF3B1* mutated status and ring sideroblasts has been confirmed in a recent experimental study on murine models.[29] About one quarter of MDS-RS are *SF3B1*^{wt} and somatic mutations of *SRSF2* or *ZRSR2* have been described in about 7% of MDS-RS,[21] which suggests that other mutant genes may play a role in the appearance of ring sideroblasts. Furthermore, a recent study showed that RARS-T presented with a particular genetic pattern with a high frequency of *JAK2*^{V617F} and *SF3B1* mutations, confirming the classification of RARS-T in the category of myelodysplastic/myeloproliferative neoplasms.[30]

Finally, since precedent studies have been performed on little RARS-T cohorts, the prognostic impact of *SF3B1*^{mut} status remains controversial,[22, 24, 25, 27, 31] and there is a need for a study on a larger cohort. Our purpose was to analyse the frequency of *SF3B1* mutations in a large cohort of 111 RARS-T and to explore the prognostic impact of *SF3B1*^{mut} status in this disorder.

Patients and methods:

Patient selection

According to the WHO 2008 classification, patients were diagnosed with RARS-T if they fulfilled the following criteria: i) anaemia (haemoglobin level <125 g/L for females and <135 g/L for males) with erythroid dysplasia and more than 15% ring sideroblasts; ii) thrombocytosis of more than $450x10^9$ platelets/L; iii) less than 5% blast cells in the bone marrow; iv) presence of large atypical megakaryocytes similar to those observed in *BCR-ABL1* negative MPN; v) no secondary cause of ring sideroblasts; and vi) no karyotype abnormalities such as del(5q), t(3;3)(q21;q26) or inv(3)(q21q26).[1]

To obtain a comprehensive data set of this rare entity, samples from seven European centres in 3 European countries were collected. The total cohort comprised 111 cases of RARS-T and 33 cases of RARS.

Data record

For each patient, demographic (gender, age at diagnosis, date of death or last follow-up) and biological data (blood cell count, bone marrow exploration, ring sideroblasts, karyotype, molecular explorations) were collected. The *SF3B1* mutations were analysed with a sensitive next-generation amplicon deep-sequencing assay (454 Life Sciences, Branford, CT, USA) with a median coverage of 500 reads. This approach was able to detect mutations with a sensitivity below 1%. The *JAK2*^{V617F} mutation was analysed by allele-specific real-time PCR to estimate allele burden according to methods published by Lippert *et al*.[32] and Schnittger *et al.*[33] with a sensitivity of 1%. *JAK2*^{exon12} analysis was performed according to the method of Schnittger *et al*.[34] and the *MPL*^{W515} mutations were analysed by high resolution melting curve analyses followed by Sanger sequencing if positive, as previously published by Schnittger *et al.*[35]

Statistical analyses

Standardized overall survival was estimated by the Kaplan-Meier method and based on the time from diagnosis to death or last contact. Survival curves for the different groups were compared using the log rank test. Multivariate analysis was performed using Cox's proportional hazards model. Approval for the study was obtained from the ethics committee of each institution and the procedures were carried out in accordance with the Helsinki Declaration of 1975, as revised in 2000.

Results:

Demographic and biological features

A total of 144 cases (111 RARS-T and 33 RARS including 72 males and 72 females) were recorded in the study. Median age at diagnosis was 73.9 y (years) (range: 44.4-96.1 years). The median follow-up was 37.5 and 55.2 months for the RARS-T and RARS cohort, respectively (table 1). Survival data was available in 130 (97 RARS-T and 33 RARS) of the 144 patients.

Frequencies and characterization of mutations

A karyotype was available in 128 cases. One hundred and ten (85.9%) patients carried a normal karyotype, whereas 18 (14.1%) patients showed aberrant karyotypes, which was equally distributed between RARS-T and RARS patients. Even if the IPSS (International Prognostic

Scoring System) score can only be applied to MDS *de novo*, we calculated it to check if we had a homogeneous group of patients. Most of the patients of the total cohort (133 out of 144) had an IPSS score of 0.

A *SF3B1*^{mut} was noted in 124 out of the 144 patients (86.1%). A total of 127 *SF3B1* mutations were detected in these 124 patients (28 RARS and 96 RARS-T). Three RARS-T cases carried 2 different mutations. With the exception of one p.Arg549Cys in exon 12 and two in exon 16, all mutations were located in exons 14 and 15. All but one del/ins mutations (p.Met784_Lys785del/insIle) were missense mutations. In detail, the most frequent mutation was p.Lys700Glu (68/127 53.5%), followed by p.Lys666Glu/Thr/Asp/Asn mutations (n=18, 14.2%), p.His662Asp/Gln (n=13, 10.2%), p.Arg625Cys/Leu (n=10, 7.9%), p.Glu622Asp (n=9, 7.1%), and p.Thr663Pro (n=4, 3.1%). Five further mutations were detected in single cases only. Frequencies and positions of mutations are illustrated in table 1 and figure 1. Median mutation/wildtype load was 40% (range: 15-78%). Small subclones with *SF3B1*^{mut} were not detected.

Frequency of mutations in RARS and RARS-T

The frequency of *SF3B1* mutations in RARS-T (96/111, 86.5%) was similar to that in RARS (28/33, 84.8%). In contrast, both entities differed by the presence of the *JAK2*^{V617F} mutation, which was detected in 54/144 (37.5%) in the total cohort but in 54/111 (48.6%) in RARS-T compared with none (0/33) in RARS (p<0.001). Among the RARS-T *SF3B1*^{mut}, 46/96 (47.9%) harbored a *JAK2*^{V617F} mutation. *JAK2*^{V617F} allele burden was very heterogeneous with a median of 49% (range: 1-100%). No *JAK2*^{exon12} mutation (111 tested) was observed, whereas only one case with the *MPL*^{W515L} mutation was noted in a RARS-T (102 tested; table 1 and figure 2).

Biological association

The presence of *SF3B1*^{mut} was analysed with respect to age, sex, white blood cell count (WBC), haemoglobin levels, platelet counts, blast counts, percentage of ring sideroblasts, karyotype and *JAK2*^{V617F} allele burden. In RARS-T, *SF3B1* mutations were more frequent in females (56/59, 94.9%) than in males (40/52, 76.9%) (p=0.010), and mean ring sideroblast counts were higher in *SF3B1*^{mut} than in *SF3B1*^{wt} (55% vs. 38%) (p=0.007). No further correlations were detected for these parameters.

Impact of mutations on outcome

The difference in survival between RARS-T and RARS was not statistically significant (median survival 10.7 *vs.* 6.2 y, respectively, p>0.05). On the other hand, in the total cohort,

patients with *SF3B1*^{mut} had longer survival than those with *SF3B1*^{wt} (8.0 *vs.* 1.8 years, respectively, p<0.001; figure 3a). When restricted to RARS-T (85 *SF3B1*^{mut} and 12 *SF3B1*^{wt}), median overall survival was 6.9 y in *SF3B1*^{mut} *vs.* 3.3 y in *SF3B1*^{wt} (p=0.003; figure 3b). In RARS, survival was 10.7 years in the *SF3B1*^{mut} (n=28) and 1.8 years in *SF3B1*^{wt} patients (n=5; p=0.001; figure 3c). In RARS-T patients, the survival rates at 2, 4 and 6 years within the *JAK2*^{V617F} sub-cohort were 94.9, 84.3 and 67.4% respectively, while within the *JAK2*^{wt} sub-cohort, they were 79.7, 69.6 and 32.2% respectively. *JAK2*^{V617F} (n=50) was then associated with a more favorable outcome compared to *JAK2*^{wt} (n=47; p=0.019; figure 3d).

Cox regression analysis

In the total cohort including RARS-T and RARS cases in univariate analysis age (p=0.020), ring sideroblast count (p=0.008), and *SF3B1* mutational status (p<0.001) were prognostically significant, but *SF3B1* mutational status was the only independent prognostic factor (p=0.001) in multivariable analysis.

In the RARS-T cohort in univariate analysis age (p=0.038), ring sideroblast count (p=0.014), *JAK2*^{V617F} (p=0.024) and *SF3B1*^{mut} (p=0.006) were prognostically significant but only age (p=0.044), *JAK2*^{V617F} (p=0.016) and *SF3B1*^{mut} (p=0.021) were independent prognostic parameters in multivariable analysis. Survival was better in patients with age < 80 years, *JAK2*^{V617F} and *SF3B1* mutations. Taking into account these three prognostic factors, a model for survival in RARS-T patients was constructed in which each risk factor (i.e. age > 80 years, *SF3B1*^{wt} and *JAK2*^{wt}) was worth 1 point, and defined two groups of risk: high risk for patients with a score of two or three (n=13) and low risk when the score is 0 or 1 (n=84). Median survival in the high risk group was 1.6 years *vs*. 8.0 years in the low risk group (p<0.001; corresponding survival rates at 2, 4, and 6 years are 31.4%, 31.4%, and 0.0% for high risk patients *vs*. 95.8%, 84.4%, and 62.8% for low risk patients; figure 4).

Discussion:

Up to now, *SF3B1* mutations have been studied in small series of RARS-T patients. In this study, we provide data on *SF3B1* mutations in a large cohort of RARS-T patients, which is, to the best of our knowledge, the largest published to date.

SF3B1 mutations were observed with a high frequency in both RARS-T and RARS patients (86.5% and 84.8%, respectively). These proportions are slightly higher than those already published, since *SF3B1* mutations have been found in 64 to 82.6% of RARS[21-24] and 66.7 to 72% in RARS-T.[23, 24] The slightly higher frequency of *SF3B1* mutations in the current study may be

due to the larger cohort of RARS-T patients than in other studies. There may also be differences in methodology, for example direct sequencing versus non-sequencing based screening strategies.

However, RARS differs from RARS-T in that there are no *JAK2*^{V617F} mutations in RARS[4, 14] whereas there is a high frequency in RARS-T.[4-13, 15, 19] We recently showed that RARS-T differed from RARS and ET from a clinical, biological and prognostic point of view, suggesting that RARS-T could be considered a unique entity.[20] The results of our current study showing the presence of both *SF3B1* and *JAK2*^{V617F} mutations in a high proportion of RARS-T confirm that RARS-T is a unique entity. Indeed, RARS-T was associated with high rates of *SF3B1* mutations (86.5%) and *JAK2*^{V617F} mutations while ET patients have a low frequency of *SF3B1* mutations, and in RARS, no *JAK2*^{V617F} mutations were detected.

A minority of RARS-T patients (13.5%) presented without *SF3B1* mutations. This could be due to another mutation affecting components of the RNA splicing machinery since mutations of *SRSF2*, *U2AF35, and ZRSR2* have already been described in 12.4%, 7.3%, and 3.1% of MDS respectively,[36] or due to mutations of proteins associated with SF3B1 (SF3B4 and SF3B14).

These results are in line with our hypothesis that RARS-T is an independent entity characterized by a particular molecular combination associating mutations that give a myeloproliferative advantage (*JAK2*^{V617F}, *MPL*$^{W515R/L}$ or other unknown mutations) and mutations of components of the splicing machinery responsible for myelodysplastic features (*SF3B1* in most cases, and possibly *SRSF2*, *U2AF35* or *ZRSF2* in the remaining cases).

Conflicting results on the prognostic impact of *SF3B1* mutations in MDS have been reported since several studies noted a good prognostic impact in MDS, whereas others hypothesized that *SF3B1* mutations were associated with good-prognosis MDS subgroups, but lost their prognostic impact in RARS and in RCMD-RS.[27]

In our large cohort of RARS-T patients, *SF3B1* mutations were associated with female sex, higher ring sideroblast counts and a longer overall survival than in *SF3B1*^{wt} patients. In a multivariable analysis, age > 80 years at diagnosis, *SF3B1*^{wt} as well as *JAK2*^{wt} were independent factors of a worse prognosis. Based on these three independent parameters, a prognostic score for RARS-T patients was created to define two risk groups: high risk when there were 2 or 3 risk factors, low when there was only one or no risk factor. Median survival was 1.6 *vs.* 8.0 years in the high and low-risk group respectively, underlying the relevance of such score in RARS-T patients.

Exploring *SF3B1* mutations in MDS associated with ring sideroblasts is of interest from a prognostic point of view, in particular since specific treatment will be available.[37, 38] Also, allele-

specific PCR have been designed, and these could be useful for monitoring minimal residual disease in MDS.[39]

In summary, this study confirms that RARS-T should be considered an independent entity. In RARS-T patients, age < 80 years at diagnosis, *SF3B1* and *JAK2* mutations are independent factors for better survival and may be used to stratify patients.

Acknowledgments

The authors would like to thank the Medical Doctors from the haematology department and laboratory of the university hospital of Dijon, the Spanish Group of Hematological Cytology (GECH), as well as Philip Bastable for revising the manuscript. EL and FL are grateful to the Tumor Bank of the CHU of Bordeaux.

Conflicts of interest

Susanne Schnittger and Torsten Haferlach declare part ownership of the MLL Munich Leukemia Laboratory. Tamara Alpermann, Sabine Jeromin and Vera Grossmann are employed by the MLL Munich Leukemia Laboratory.

All other authors have no conflict of interest to declare.

Table 1

Demographic,and biological characteristics of RARS-T and RARS patients. RARS-T: Refractory anaemia with Ring Sideroblasts and marked Thrombocytosis; RARS: Refractory Anaemia with Ring Sideroblasts; WBC: White Blood Cells; *SF3B1*: Splicing Factor 3B subunit 1; *JAK2*: Janus Kinase 2; *MPL*: MyeloProliferative Leukaemia; IPSS: International Prognostic Scoring System.

	RARS-T patients	RARS patients
n	111	33
Male (%)	46.8	60.6
Age at diagnosis (years)		
Median	74.3	71.1
Range	44.4-92.1	48.4-96.1
20-50 years n (%)	4 (3.6)	2 (6.1)
50-70 years n (%)	36 (32.5)	9 (27.2)
> 70 years n (%)	71 (63.9)	22 (66.7)
Available survival data (n)	97	33
Median Follow-up (years)	3.1	4.6
WBC (x 10^9/L)		
Median	7.6	5.2
Range	2.1-60.0	1.6-17.3
Hb (g/L)		
Median	96.5	91.0
Range	51.0-131.0	69.0-128.0
Platelets (x 10^9/L)		
Median	646	314
Range	452-1500	61-444
450-600 n (%)	54 (48.6)	
> 600 n (%)	57 (51.4)	
Ring sideroblasts (%)		
Median	52	40
Range	16-97	19-85
SF3B1 mutations (%)	86.5	84.8
Tested (n)	111	33
Mutated (n)	96	28
p.Lys700Glu	51	16
p.Lys666Glu/Thr/Asp/Asn	16	2
p.His662Asp/Gln	11	2
p.Glu622Asp	7	2
p.Arg625Gys/Leu/Gly	5	3
p.Thr663Pro	2	2
p.Met784_Lys785delinsIle	1	0
p.Asp781Gly	0	1
Two different mutations	3	0
JAK2^{V617F} mutations (%)	48.6	0
Tested (n)	111	33
MPL mutations (%)	1	0
Tested (n)	102	27
IPSS		
0	104	29
0.5	5	3
1	2	1

References:

1. Vardiman JW, Bennett JM, Bain BJ, Baumann I, Thiele J, Orazi A. Myelodysplastic/myeloproliferative neoplasms, unclassifiable. In: Swerdlow SH, Campo E, Lee Harris N, Jaffe ES, Pileri SA, Stein H *et al.* (eds). *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissue*, 4th edn. IARC: Lyon, France, 2008, pp 85-86.

2. Hasserjian RP, Gatterman N, Bennett JM, Brunning RD, Thiele J. Refractory anemia with ringed sideroblasts. In: Swerdlow SH, Campo E, Lee Harris N, Jaffe ES, Pileri SA, Stein H *et al.* (eds). *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissue,* 4th edn. IARC: Lyon, France, 2008, pp 96-97.

3. Thiele J, Kvasnicka HM, Orazi A, Tefferi A, Gisslinger H. Essential thrombocythaemia. In: Swerdlow SH, Campo E, Lee Harris N, Jaffe ES, Pileri SA, Stein H *et al.* (eds). *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissue,* 4th edn. IARC: Lyon, France, 2008, pp 48-50.

4. Ceesay MM, Lea NC, Ingram W, Westwood NB, Gaken J, Mohamedali A, *et al.* The JAK2 V617F mutation is rare in RARS but common in RARS-T. *Leukemia* 2006; **20**:2060-1.

5. 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* 2006; **108**:1781-2.

6. Flach J, Dicker F, Schnittger S, Kohlmann A, Haferlach T, Haferlach C. Mutations of JAK2 and TET2, but not CBL are detectable in a high portion of patients with refractory anemia with ring sideroblasts and thrombocytosis. *Haematologica* 2010; **95**:518-9.

7. 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* 2007; **109**:1334-5.

8. Hellstrom-Lindberg E, Cazzola M. The role of JAK2 mutations in RARS and other MDS. Hematology Am Soc Hematol Educ Program 2008:52-9.

9. Raya JM, Arenillas L, Domingo A, Bellosillo B, Gutierrez G, Luno 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.* 2008; **88**:387-95.

10. Remacha AF, Nomdedeu 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* 2006; **91**:719-20.

11. 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* 2006; **20**:2067-70.

12. Schmitt-Graeff AH, Teo SS, Olschewski M, Schaub F, Haxelmans S, Kirn A, *et al.* JAK2V617F mutation status identifies subtypes of refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Haematologica* 2008; **93**:34-40.

13. 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* 2006; **108**:2173-81.

14. Steensma DP, Tefferi A. JAK2 V617F and ringed sideroblasts: not necessarily RARS-T. *Blood* 2008; **111**:1748.

15. 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* 2006; **20**:1641-4.

16. 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* 2006; **108**:3472-6.

17. Schnittger S, Bacher U, Haferlach C, Dengler R, Krober 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* 2008; **22**:453-5.

18. Wardrop D, Steensma DP. Is refractory anaemia with ring sideroblasts and thrombocytosis (RARS-T) a necessary or useful diagnostic category? *Br J Haematol.* 2009; **144**:809-17.

19. Malcovati L, Della Porta MG, Pietra D, Boveri E, Pellagatti A, Galli A, *et al.* Molecular and clinical features of refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Blood* 2009; **114**:3538-45.

20. 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* 2012; **97**:1036-41.

21. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, *et al.* Frequent pathway mutations of splicing machinery in myelodysplasia. *Nature* 2011; **478**:64-9.

22. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, *et al.* Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med.* 2011; **365**:1384-95.

23. Visconte V, Makishima H, Jankowska A, Szpurka H, Traina F, Jerez A, *et al.* SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. *Leukemia* 2012; **26**:542-5.

24. Malcovati L, Papaemmanuil E, Bowen DT, Boultwood J, Della Porta MG, Pascutto C, *et al.* Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. *Blood* 2011; **118**:6239-46.

25. Damm F, Thol F, Kosmider O, Kade S, Loffeld P, Dreyfus F, *et al.* SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. *Leukemia* 2012; **26**:1137-40.

26. Damm F, Kosmider O, Gelsi-Boyer V, Renneville A, Carbuccia N, Hidalgo-Curtis C, *et al.* Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. *Blood* 2012; **119**:3211-8.

27. Patnaik MM, Lasho TL, Hodnefield JM, Knudson RA, Ketterling RP, Garcia-Manero G, *et al.* SF3B1 mutations are prevalent in myelodysplastic syndromes with ring sideroblasts but do not hold independent prognostic value. *Blood* 2012; **119**:569-72.

28. Visconte V, Makishima H, Maciejewski JP, Tiu RV. Emerging roles of the spliceosomal machinery in myelodysplastic syndromes and other hematological disorders. *Leukemia* 2012; **26**:2447-54.

29. Visconte V, Rogers HJ, Singh J, Barnard J, Bupathi M, Traina F, *et al.* SF3B1 haploinsufficiency leads to formation of ring sideroblasts in myelodysplastic syndromes. *Blood* 2012; **120**:3173-86.

30. Jeromin S, Haferlach T, Grossmann V, Alpermann T, Kowarsch A, Haferlach C, *et al.* High frequencies of SF3B1 and JAK2 mutations in refractory anemia with ring sideroblasts associated with marked thrombocytosis strengthen the assignment to the category of myelodysplastic/myeloproliferative neoplasms. *Haematologica* 2013; **98**:15-17.

31. Cazzola M, Rossi M, Malcovati L. Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms. *Blood* 2013; **121**:260-9.

32. Lippert E, Boissinot M, Kralovics R, Girodon F, Dobo I, Praloran V, *et al.* The JAK2-V617F mutation is frequently present at diagnosis in patients with essential thrombocythemia and polycythemia vera. *Blood* 2006; **108**:1865-7.

33. Schnittger S, Bacher U, Kern W, Schroder M, Haferlach T, Schoch C. Report on two novel nucleotide exchanges in the JAK2 pseudokinase domain: D620E and E627E. *Leukemia* 2006; **20**:2195-7.

34. Schnittger S, Bacher U, Haferlach C, Geer T, Muller P, Mittermuller J, *et al.* Detection of JAK2 exon 12 mutations in 15 patients with JAK2V617F negative polycythemia vera. *Haematologica* 2009; **94**:414-8.

35. Schnittger S, Bacher U, Haferlach C, Beelen D, Bojko P, Burkle D, *et al.* Characterization of 35 new cases with four different MPLW515 mutations and essential thrombocytosis or primary myelofibrosis. *Haematologica* 2009; **94**:141-4.

36. Thol F, Kade S, Schlarmann C, Loffeld P, Morgan M, Krauter J, *et al.* Frequency and prognostic impact of mutations in SRSF2, U2AF1, and ZRSR2 in patients with myelodysplastic syndromes. *Blood* 2012; **119**:3578-84.

37. Rymond B. Targeting the spliceosome. *Nat Chem Biol.* 2007; **3**:533-5.

38. Webb TR, Joyner AS, Potter PM. The development and application of small molecule modulators of SF3b as therapeutic agents for cancer. *Drug Discov today* 2013; **18**:43-9.

39. Matsuda K, Ishida F, Ito T, Nakazawa H, Miura S, Taira C, *et al.* Spliceosome-related gene mutations in myelodysplastic syndrome can be used as stable markers for monitoring minimal residual disease during follow-up. *Leuk Res.* 2012; **36**:1393-7.

Figures

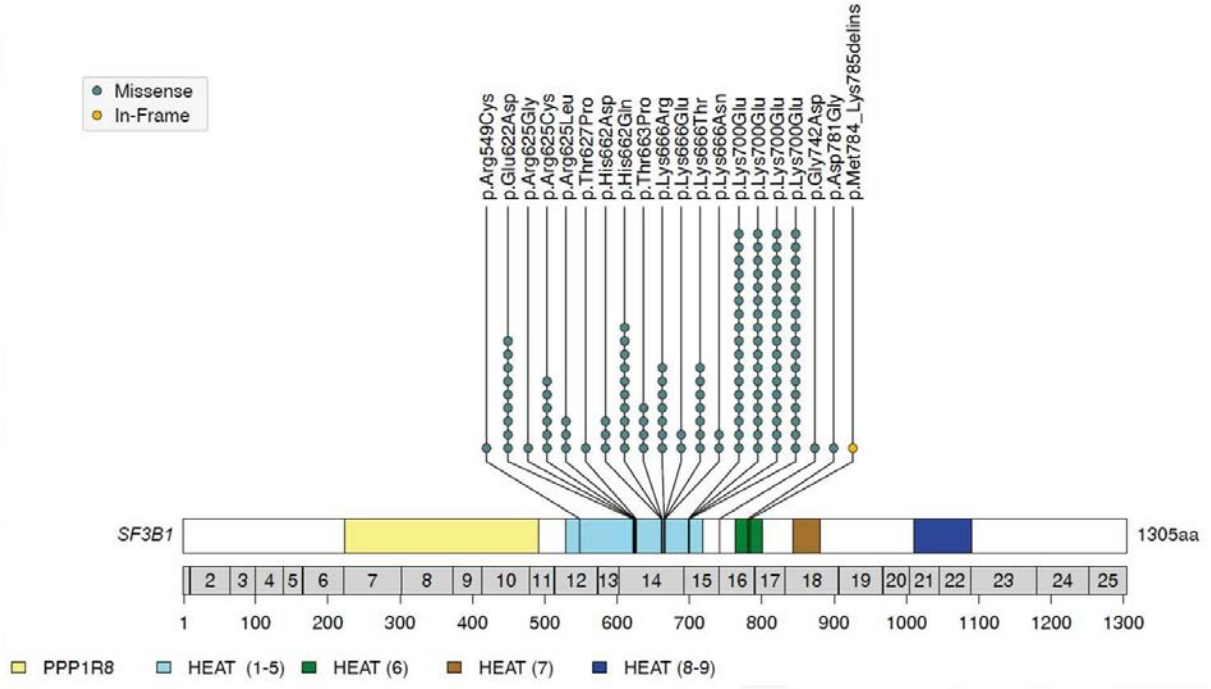


Figure 1: Positions and characterization of mutations in the HEAT domains 1-6 (exons 12-16) of the *SF3B1* gene. Missense mutations are indicated in green and the rare ins/del mutation in yellow. HEAT: Huntingtin, Elongation factor 3, protein phosphatase 2A, Tor1); PPP1R8: protein phosphatase 1, regulatory subunit 8; aa: amino acid.



Figure 2: Distribution of *SF3B1* mutations and *JAK2*^{V617F} and *MPL*^{W515L} mutations in RARS-T and RARS.

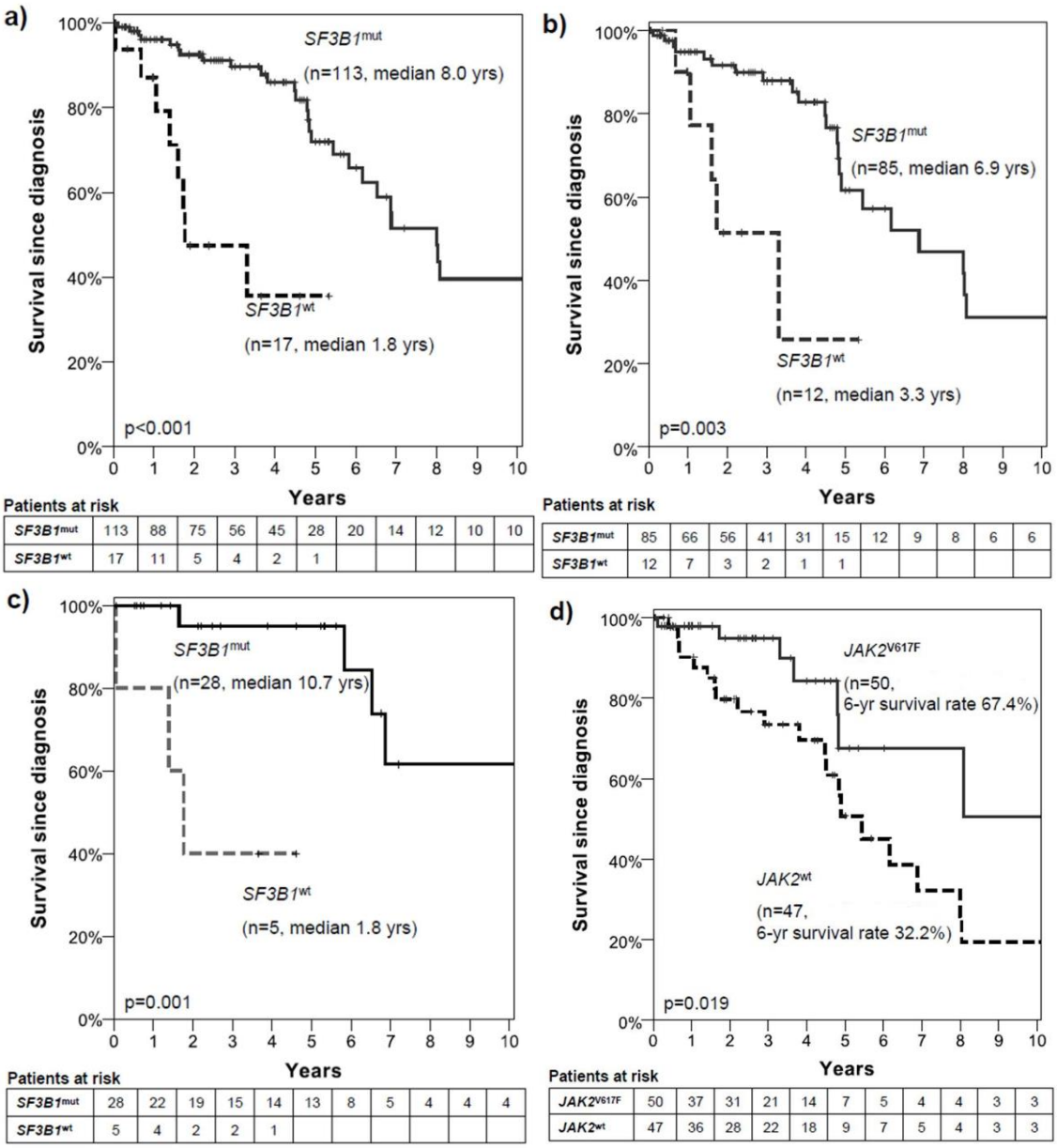
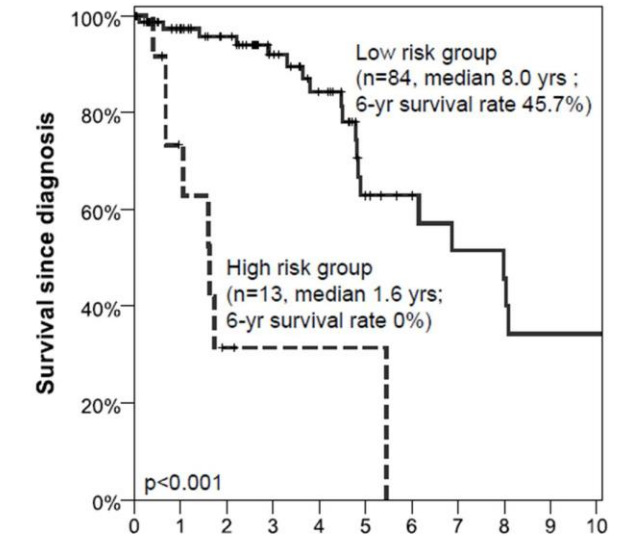


Figure 3: Kaplan Meier analysis for survival since diagnosis according to *SF3B1* and *JAK2*^{V617F} mutation status. *SF3B1* in a) total cohort, b) RARS-T, c) RARS, d) *JAK2*^{V617F} in RARS-T.



Patients at risk

Low risk	84	66	57	42	31	15	12	9	8	6	6
High risk	13	7	2	1	1	1					

Figure 4: Kaplan Meier analysis for survival according to 2 risk groups. High risk comprising patient with at least two of the following risk factors: age > 80 years, *SF3B1*^{wt}, *JAK2*^{wt}; and low risk: one or less of the risk factors.

2. LOW RATE OF CALRETICULIN MUTATIONS IN REFRACTORY ANAEMIA WITH RING SIDEROBLASTS AND MARKED THROMBOCYTOSIS

Contexte

L'ARS-T demeure une entité provisoire du point de vue de la classification OMS 2008. Pourtant son caractère indépendant repose sur des caractéristiques biologiques, pronostiques et moléculaires obtenues sur une grande cohorte européenne. Pour autant, les déterminants de ses caractéristiques myéloprolifératives ne sont pas totalement élucidés. En effet, environ 50% des ARS-T sont porteuses de la mutation *JAK2*^{V617F} et 1% sont porteuses de la mutation *MPL*^{W515L}. Ce qui signifie que près de 50% des ARS-T ne sont pas porteuses d'une mutation connue pour être responsable du caractère prolifératif. La problématique est la même pour les SMP *JAK2* et *MPL* non-mutés, en particulier la TE et la MFP.

Or il a récemment été démontré que les mutations affectant *CALR*, le gène codant pour la calréticuline, étaient retrouvées dans 67 à 71% des TE et 56 à 88% des MFP *JAK2* non-mutées. La calréticuline est une protéine résidant dans la lumière du réticulum endoplasmique et agissant comme une protéine chaperonne pour de nombreuses glycoprotéines, permettant leur compaction. Par ailleurs, le domaine C-terminal de la protéine participe à la régulation de l'homéostasie calcique et donc la signalisation cellulaire. Plus de 50 types de mutations de CALR ont été décrites. Il s'agit d'insertions ou de délétions situées dans l'exon 9 du gène, responsables d'un décalage du cadre de lecture. Plus de 80% de ces mutations sont des délétions de 52 paires de bases (type 1) ou des insertions de 5 paires de bases (type 2). Les mutations de *CALR*, *JAK2* et *MPL* ont été décrites comme mutuellement exclusives. Ainsi, il existe désormais un marqueur de clonalité dans près 90% des TE et des MFP.[19, 20] Cette mutation a été recherchée dans d'autres pathologies proches des syndromes myéloprolifératifs, en particulier l'ARS-T. Les mutations de *CALR* sont retrouvées dans 0 à 12,5% des ARS-T, mais les cohortes explorées étaient de petite taille (24 malades au maximum).

Objectifs

Notre objectif est de rechercher les mutations de *CALR* dans une cohorte d'ARS-T plus importante. Notre hypothèse est qu'il existe un fort taux de mutations de ce gène dans les ARS-T, expliquant le caractère myéloprolifératif dans les cas *JAK2* non-mutés.

Méthodologie

Nous avons réuni une cohorte de 95 ARS-T, réunissant tous les critères OMS. Les mutations de *SF3B1* et les mutations *JAK2*^{V617F} et *JAK2*^{exon12} ont été recherchées selon les méthodes précédemment décrites.[167] Les mutations de l'exon 9 de *CALR* ont été recherchées par analyse de fragments et séquençage selon Sanger, selon la méthode publiée.[19]

Principaux résultats

82,7% des ARS et 88,3% des ARS-T étaient porteuses d'une mutation de *SF3B1*. 49,4% des ARS-T étaient porteuses de la mutation *JAK2*^{V617F}. Aucune mutation de l'exon 12 de *JAK2* n'a été retrouvée. Enfin, seul une ARS-T portait la mutation *MPL*^{W515L} (sur 88 testées).

Seule une ARS-T portait une mutation de l'exon 9 de *CALR*. Il s'agissait d'une délétion de 10 paires de bases, responsable d'un décalage du cadre de lecture. Par ailleurs, cette ARS-T portait la mutation *JAK2*^{V617F} - avec une charge allélique de 4% - et une mutation de *SF3B1*. Le statut mutationnel a été vérifié dans deux laboratoires indépendants, sur deux échantillons d'ARN différents.

Conclusion

En ce qui concerne les mutations de *JAK2*, *MPL* et *SF3B1*, nos résultats sont concordants avec les données de la littérature.

En revanche notre prévalence de mutations de *CALR* est de 1%. Les prévalences décrites varient de 0 à 12,5%, mais ont été décrites sur des cohortes d'effectifs plus faibles. Par ailleurs, dans une de ces cohortes, le taux de mutations de *SF3B1* était plus faible, faisant craindre l'inclusion de plusieurs cas de TE.[19]

Nous décrivons pour la première fois la co-occurrence d'une mutation de *JAK2* et de *CALR* dans une ARS-T. Un cas de MFP double mutant *JAK2* et *CALR* a été décrit.[168]

Le caractère myéloprolifératif reste inexpliqué pour environ 50% d'ARS-T, qui sont triple négatives pour *JAK2*, *MPL* et *CALR*.

Low rate of Calreticulin mutations in Refractory Anaemia with Ring Sideroblasts and marked Thrombocytosis

Running title: *Calreticulin* mutations in RARS-T.

Julien Broséus,[1,2,12,13,14] Eric Lippert,[3,4,12,13,14] Thorsten Klampfl,[5] Sabine Jeromin,[6] Esther Zipperer,[7] Lourdes Florensa,[8,15] Ashot S. Harutyunyan,[5] Jelena D. Milosevic,[5] Torsten Haferlach,[6] Ulrich Germing,[7] Elisa Luño,[9,15] Susanne Schnittger,[6,13] Robert Kralovics,[5,13] and Francois Girodon[10,11,13,14‡]

[1]Service d'Hématologie Biologique, University Hospital, Nancy, France
[2]Inserm U954, faculty of medicine, Nancy, France
[3]Haematology laboratory, University Hospital, Bordeaux, France
[4]Inserm U1035-Université Bordeaux Segalen, France
[5]CeMM Research Center for Molecular Medicine of the Austria Academy of Sciences, Austria
[6]MLL Munich Leukemia Laboratory, Munich, Germany;
[7]Departement of Haematology, Oncology and Clinical Immunology, Heinrich-Heine-Universität, Düsseldorf, Germany;
[8]Laboratorio Citologia Hematològica, Servicio Patologia, Hospital del Mar, Barcelona, Spain;
[9]Servicio de Hematología, Servicio de Salud del Principado de Asturias, Oviedo, Spain;
[10]Haematology Laboratory, University Hospital, Dijon, France;
[11]Inserm U866, faculty of medicine Dijon, France;
[12]These authors contributed equally to this work
[13]On behalf of MPN&MPNr-EuroNet (COST Action BM0902).
[14]On behalf of the French Intergroup of MPNs (FIM)
[15]On behalf of the Spanish Group of Hematological Cytology (GECH).

‡Corresponding author:

Pr François Girodon
Haematology laboratory, University Hospital, Dijon, France
Plateau technique de biologie, 2 rue Angélique Ducoudray 21079 Dijon cedex
francois.girodon@chu-dijon.fr
Tel: (+33) 3 80 29 57 10 ; Fax: (+33) 3 80 29 36 60

Grants and supports to disclose:

This work was supported by grants from the association "Tulipes contre le cancer" (Châlon s/Saône, Burgundy, France) and from FEHH (Spain) and 2009 SGR 541 (Generalitat de Catalunya).

Acknowledgments

The authors would like to thank the Medical Doctors from the haematology department and laboratory of the university hospital of Dijon, the Spanish Group of Hematological Cytology (GECH). EL is grateful to Ms Marina Migeon for expert technical help and to the Tumor Bank of the CHU of Bordeaux. We thank Sylvie Hermouet, medical doctor at the university hospital of Nantes, for reading the manuscript as well as Philip Bastable for revising the manuscript.

Conflicts of interest

Susanne Schnittger and Torsten Haferlach declare part ownership of the MLL Munich Leukemia Laboratory. Sabine Jeromin is employed by the MLL Munich Leukemia Laboratory. All other authors have no conflict of interest to declare.

Refractory Anaemia with Ring Sideroblasts (RARS) and marked Thrombocytosis (RARS-T) was proposed in the World Health Organisation 2001 classification of tumours of haematopoietic and lymphoid tissues and retained as a provisional entity in the 2008 version.[1] RARS-T is characterized by a high rate of *JAK2*^{V617F} mutations[2-4] or the presence of mutations in exon 10 of *MPL* (*MyeloProliferative Leukaemia*).[5,6] The existence of RARS-T as an entity independent from RARS and Essential Thrombocythemia (ET) was a matter of debate, as certain specialists favoured the hypothesis that RARS-T was a form of ET with more than 15% of ring sideroblasts in the bone marrow, while others thought that RARS-T derived from RARS with secondary thrombocytosis developing through the acquisition of the *JAK2*^{V617F} mutation.

We recently demonstrated that RARS-T differed from RARS and ET from a clinical, biological and prognostic point of view.[4] The presence of high rates of Splicing Factor 3B subunit 1 (*SF3B1*) mutations in RARS-T and the absence of these mutations in ET strengthened the hypothesis that RARS-T was a distinct entity.[7-9]

However, the myeloproliferative features of RARS-T are not totally explained by *JAK2*^{V617F} and *MPL*^{W515L} mutations since these account for only 50% and 1% of RARS-T, respectively. The question remains about other causative mutations responsible for the myeloproliferative part of RARS-T. The challenge is the same in myeloproliferative neoplasms (MPN) since 50 to 60% of ET and Primary Myelofibrosis (PMF) do not present with any *JAK2* or *MPL* mutations.

The recent discovery of a high rate of *Calreticulin* (*CALR*) mutations in *JAK2*-non-mutated MPNs was a great step towards better understanding of the molecular pattern of these diseases. Mutations in exon 9 of *CALR* have been reported as mutually exclusive of *JAK2* and *MPL* mutations, and found in 67 to 71% of ET and 56 to 88% of PMF with wild type *JAK2* or *MPL*.[10,11] These mutations are insertions or deletions leading to a frameshift responsible for modifying the C-terminal part of the protein. Following these modifications, the C-terminal part of the protein becomes positively charged, and the reticulum targeting KDEL sequence is abrogated, which disturbs its cellular localisation.

Up to now, *CALR* mutations have only been explored in small RARS-T cohorts, with a mutation frequency ranging from 0 to 12.5%.[10,11] The purpose of our study was to analyse the frequency of *CALR* mutations in a large RARS-T cohort in order to determine whether *CALR* mutations may also be responsible for thrombocytosis in this disease. We therefore analysed a large cohort of 95 RARS-T patients, with 29 RARS as the control.

According to the WHO 2008 classification, patients are diagnosed with RARS-T if they fulfil the following criteria: i) anaemia (haemoglobin level <125 g/L for females and <135 g/L for males) with erythroid dysplasia and more than 15% ring sideroblasts; ii) thrombocytosis with more than

450 x10^9 platelets/L; iii) less than 5% blast cells in the bone marrow; iv) presence of large atypical megakaryocytes similar to those observed in *BCR-ABL1* negative MPN; v) no secondary cause of ring sideroblasts; and vi) no karyotype abnormalities such as del(5q), t(3;3)(q21;q26) or inv(3)(q21q26).[1] One hundred and twenty-four samples including 95 RARS-T and 29 RARS from seven European centres in three European countries were collected and tested. This cohort has been published previously.[4, 9]

For each patient, demographic (gender, age at diagnosis) and biological data (blood cell count, bone marrow exploration, ring sideroblasts, karyotype, molecular explorations) were collected. The *SF3B1* mutations were analysed with a sensitive next-generation amplicon deep-sequencing assay (454 Life Sciences, Branford, CT, USA) with a median coverage of 500 reads. This approach was able to detect mutations with a sensitivity below 1%.[9] The *JAK2*^{V617F} mutation was analysed by allele-specific real-time PCR to estimate allele burden according to method published by Lippert *et al*.[12] with a sensitivity of 1%. *JAK2* exon12 analysis was performed according to the method of Schnittger *et al*.[13] and the *MPL*^{W515} mutations were analysed by high-resolution melting curve analyses followed by Sanger sequencing if positive, as previously published by Schnittger *et al.*[14] The *CALR* exon 9 mutations were screened by fragment analysis and Sanger sequencing according to the method of Klampfl *et al*.[10]

Approval for the study was obtained from the ethics committee of each institution and the procedures were carried out in accordance with the Helsinki Declaration of 1975, as revised in 2000.

A total of 124 cases with 95 RARS-T and 29 RARS (including 62 males and 62 females) were recorded in the study, which is, to our best knowledge, the largest series of Myelodysplastic Syndromes with Ring Sideroblasts (MDS-RS) studied for *CALR* mutations. The median age at diagnosis was 74 and 73 years for the RARS-T and RARS cohort, respectively (table 1). A karyotype was available in 112 cases (87 RARS-T and 25 RARS). Seventy-five (86.2%) RARS-T patients carried a normal karyotype, whereas 12 (13.8%) patients showed aberrant karyotypes. The rates in RARS-T and RARS patients were comparable.

An *SF3B1* mutation was noted in 24 (82.7%) RARS and 83 (88.3%) RARS-T confirming that this mutation is highly represented in MDS-RS.

A *JAK2*^{V617F} mutation was noted in 47 of the 95 RARS-T patients (49.4%) but in no RARS patients. The *JAK2*^{V617F} allele burden was very heterogeneous with a median of 33.3% (range: 1-92%). No *JAK2* exon12 mutation (19 tested) was observed, whereas only one *MPL*^{W515L} mutation was noted in a patient with RARS-T (88 tested; table 1 and figure 1).

A *CALR* mutation was noted in only one (1%) RARS-T and one (3.4%) RARS patient. The *CALR* mutation found in the RARS-T patient was a 10 bp (1129-1138) deletion located in exon 9. Surprisingly, this patient also presented a *JAK2*^{V617F} mutation with a low allele burden (4%) and an *SF3B1* mutation. The mutational status of this patient was confirmed by two independent laboratories on two different DNA samples. This patient is a 73-year old women diagnosed in 2010. At diagnosis, her blood cell count revealed an anaemia (haemoglobin 106 g/L) and a marked thrombocytosis (992 x10^9 platelets/L). She is currently under cytoreductive and anti-platelet treatment and her last blood count revealed a mild anaemia (haemoglobin 104 g/L) and a platelet count lowered to 600 x10^9/L.

Among the 29 RARS, we also identified a single case with a *CALR* mutation. This unique *CALR*-positive RARS case was a 70-year-old man who had thrombocytopenia and carried an *SF3B1* mutation.

In the present study on *JAK2* and *MPL* mutations, our results are in accordance with the usually reported prevalence of *JAK2*^{V617F} in half of RARS-T and the low frequency of *MPL* mutations.[2-6]

Recent publications have reported a variable frequency of *CALR* mutations in MDS-RS: 0/6 to 3/24 in RARS-T, 3/27 in RARS and 0/3 in Refractory Cytopenia with Multilineage Dysplasia and Ring Sideroblasts.[10,11] However, these results were obtained in small series of patients, thus leading to potential biases. Of note, a high proportion (12.5%) of *CALR* mutations in RARS-T in one series of 24 patients contrasts with our results where only 1 positive case in 95 patients was noted. In particular, one of the 3 previously reported *CALR*-positive RARS-T cases did not harbour the *SF3B1* mutation, a frequent hallmark of RARS-T. Moreover, in this series, the frequency of the *SF3B1* mutation was much lower than in ours (66% *vs.* 88%) suggesting that this *CALR*-positive *SF3B1*-negative RARS-T case could be MPN rather than RARS-T, due to the difficulty of asserting the diagnosis in some cases. Finally, the low rate of *CALR* mutations in the present RARS-T series is similar (0 to 3%) to that reported in chronic myelomonocytic leukaemia, another MDS/MPN.[10,11]

Among the 29 RARS, we also identified a single case with a *CALR* mutation, which is consistent with a previous report.[11]

Surprisingly, the *CALR*-positive RARS-T patient also carried the *JAK2*^{V617F} mutation, which is to date, the first double mutant *JAK2*/*CALR* ever described in RARS-T. Indeed, the *CALR* and *JAK2*^{V617F} mutations were described as mutually exclusive in the two recent studies of more than

2 300 hematologic cancers.[10,11] However, another group recently described a case of a *JAK2*/*CALR* double mutation in primary myelofibrosis,[15] which is in keeping with our results and proves that these mutations coexist in rare cases. One can hypothesize that this RARS-T patient with a *JAK2*/*CALR* mutation had two different clones. However, due to absence of frozen progenitor cells, no clonogenic assay was feasible.

The two mayor types of CALR mutations are type 1 mutations (52 bp deletion; c.1192_1143del) and type 2 mutations (5 bp insertion; c.1154_1155insTTGTC).[10] The *CARL* mutation described in our RARS-T is a 10 bp deletion (1129-1138del), which is a rare type of *CALR* exon 9 mutation that shifts the reading frame in a similar way to all reported *CALR* mutations.

Finally our study showed that *CALR* mutations are rare in MDS-RS and thus do not explain the myeloproliferative part of RARS-T. Approximately 40 to 50% of RARS-T are triple-negative for *JAK2*, *MPL* and *CALR* mutations, indicating that further research is necessary to explain the myeloproliferation in this rare MDS/MPN.

References:

1. Vardiman JW, Bennet JM, Bain BJ, Baumann I, Thiele J, Orazi A. Myelodysplastic/myeloproliferative neoplasms, unclassifiable. In: Swerdlow SH, Campo E, Lee Harris N, Jaffe ES, Pileri SA, Stein H, eds. WHO classification of tumours of haematopoietic and lymphoid tissues, 4th edn. Lyon, France: IARC; 2008. pp. 85-6.

2. Renneville A, Quesnel B, Charpentier A, Terriou L, Crinquette A, Laï JL, *et al*. High occurrence of JAK2 V617 mutation in refractory anemia with ringed sideroblasts associated with marked thrombocytosis. *Leukemia* 2006; **20**: 2067-70.

3. Steensma DP, Caudill JS, Pardanani A, McClure RF, Lasho TL, Tefferi A. MPL W515 and JAK2 V617 mutation analysis in patients with refractory anemia with ringed sideroblasts and an elevated platelet count. *Haematologica* 2006; **91**(12 Suppl): ECR57.

4. 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* 2012; **97**: 1036-41.

5. 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* 2006; **108**: 3472-6.

6. 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* 2008; **22**: 453-5.

7. Visconte V, Makishima H, Jankowska A, Szpurka H, Traina F, Jerez A, *et al*. SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. *Leukemia* 2012; **26**: 542-5.

8. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, *et al*. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. *N Engl J Med* 2011; **365**: 1384-95.

9. Broseus J, Alpermann T, Wulfert M, Florensa Brichs L, Jeromin S, Lippert E, *et al*. 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**: 1826-31.

10. 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* 2013; **369**: 2379-2390.

11. 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* 2013; **369**: 2391-2405.

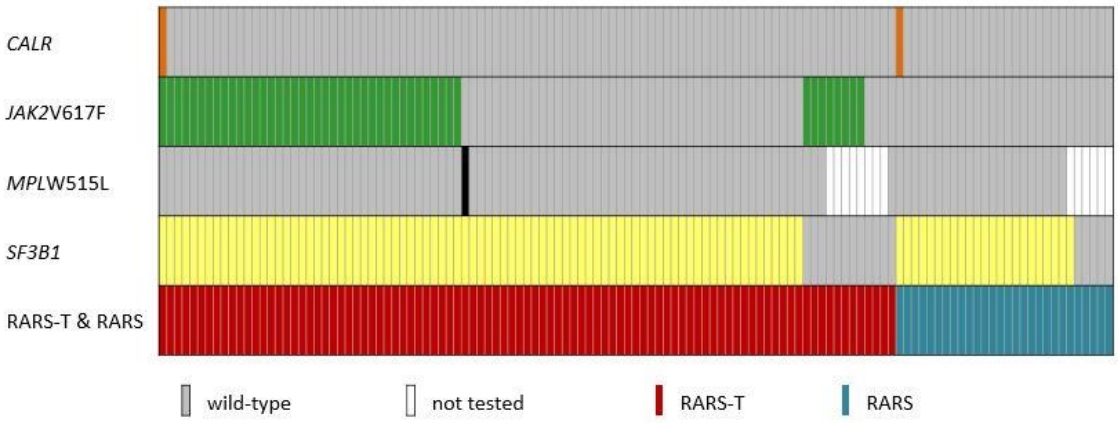
12. Lippert E, Boissinot M, Kralovics R, Girodon F, Dobo I, Praloran V, *et al*. The JAK2-V617F mutation is frequently present at diagnosis in patients with essential thrombocythemia and polycythemia vera. *Blood* 2006; **108**: 1865-7.

13. Schnittger S, Bacher U, Haferlach C, Geer T, Müller P, Mittermüller J, *et al*. Detection of JAK2 exon 12 mutations in 15 patients with JAK2V617F negative polycythemia vera. *Haematologica* 2009; **94**: 414-8.

14. Schnittger S, Bacher U, Haferlach C, Beelen D, Bojko P, Bürkle D, *et al*. Characterization of 35 new cases with four different MPLW515 mutations and essential thrombocytosis or primary myelofibrosis. *Haematologica* 2009; **94**: 141-4.

15. Tefferi A, Lasho TL, Finke CM, Knudson RA, Ketterling R, Hanson CH, *et al*. CALR vs JAK2 vs MPL mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. *Leukemia* 2014; e-pub ahead of print 9 january 2014; doi: 10.1038/leu.2014.3.

Figure 1.



Distribution of mutations in *SF3B1* , *JAK2*^{V617F} , *MPL*^{W515L} and *CALR* mutations in 95 Refractory Anaemia with Ring Sideroblasts and marked Thrombocytosis (RARS-T) and 29 Refractory Anaemia with Ring Sideroblasts (RARS)**.** In RARS-T patients, *SF3B1*, *JAK2*^{V617F}, *MPL* and *CALR* mutations were observed in 88.3%, 49.4%, 1.1% and 1% respectively. In RARS patients, *SF3B1*, *JAK2*^{V617F}, *MPL* and *CALR* mutations were observed in 82.7%, 0%, 0% and 3.4% respectively.

Table legends:

Table 1: Demographic and biological characteristics of RARS-T and RARS patients.

	RARS-T patients	RARS patients
n	95	29
Male (%)	45.2	65.5
Age at diagnosis (years)		
Median	74.0	73.0
Range	44.4-88.1	48.4-93.0
20-50 years (n)	4	2
50-70 years (n)	30	8
> 70 years (n)	61	19
WBC (x 10^9/L)		
Median	7.6	5.5
Range	2.1-60	1.9-17.3
Hb (g/L)		
Median	9.6	9.1
Range	5.1-12.9	6.9-12.8
Platelets (x 10^9/L)		
Median	655	274
Range	455-1500	61-444
450-600 (n)	45	
> 600 (n)	50	
Ring sideroblasts (%)		
Median	53.5	46.5
Range	16-97	19-85
Karyotype		
Tested(n)	87	25
Abnormal (n)	12	6
SF3B1 mutations (%)	88.3	82.7
Tested (n)	94	29
JAK2^{V617F} mutations (%)	49.4	0.0
Tested (n)	95	29
MPL mutations (%)	1.1	0.0
Tested (n)	88	23
CALR mutations (%)	1	3.4
Tested (n)	95	29

Table 1: Demographic and biological characteristics of RARS-T and RARS patients. RARS-T: Refractory anaemia with Ring Sideroblasts and marked Thrombocytosis; RARS: Refractory Anaemia with Ring Sideroblasts; WBC: White Blood Cells; *SF3B1*: Splicing Factor 3B subunit 1; *JAK2*: Janus Kinase 2; *MPL*: MyeloProliferative Leukaemia; *CALR*: Calreticulin.

3. PRESENCE OF CALRETICULIN MUTATIONS IN JAK2-NEGATIVE POLYCYTHEMIA VERA

Contexte

Les mutations de l'exon 9 de *CALR* ont été décrites dans 67 à 71% des TE et 56 à 88% des MFP *JAK2* et *MPL* non-mutées. Les mutations de *JAK2*, *MPL* et *CALR* sont mutuellement exclusives dans l'immense majorité des cas, mais de rares cas de doubles mutants ont été décrits. Par contre, aucune mutation de *CALR* n'a été décrite dans la PV, où les mutations de *JAK2* sont présentes dans 96 à 99% des cas.[18-20, 168, 169]

Objectifs

Il existe de rares cas de PV ne présentant ni mutation *JAK2*^{V617F}, ni mutation de l'exon 12 de *JAK2*. Nous avons recherché une mutation de l'exon 9 de *CALR* dans ces rares cas.

Méthodologie

Nous avons recherché dans notre base de données les PV ne présentant aucune mutation de *JAK2* ou de *MPL*. Les mutations de l'exon 9 de *CALR*, de *JAK2* (exons 12, 13 et 14) et de *MPL* ont été recherchées selon des méthodes précédemment décrites.[22, 170, 171] La recherche de mutations de *CALR* a été réalisée à la fois sur les cellules nucléées du sang circulant et sur colonies érythroblastiques obtenues à partir de cellules du sang circulant.

Principaux résultats

Nous avons retrouvé une mutation de l'exon 9 de *CALR* pour deux échantillons de PV *JAK2* non-muté selon les critères de l'OMS. Dans les deux cas, il s'agissait d'une délétion de 52 paires de bases, correspondant au type de mutation le plus fréquent (type 1). Dans un des deux cas, nous avons pu obtenir des colonies érythroblastiques à partir du sang circulant. La mutation identifiée dans les cellules granuleuses du sang circulant a été retrouvée à l'état hétérozygote dans 30 des 37 colonies obtenues et à l'état homozygote dans 1 colonie. Aucune mutation de *JAK2* (exons 12, 13 et 14) n'a été retrouvée.

Conclusion

Ces résultats ouvrent l'hypothèse qu'une mutation de *CALR* peut être à l'origine d'une PV. Il convient toutefois de rester prudent car des cas de SMP biclonaux *JAK2*^{V617F} et *CALR* mutés ont été décrits.[172] Nous ne pouvons donc pas éliminer l'apparition possible de mutations de *CALR* au cours de l'évolution clonale, qui seraient précédées par une autre mutation, encore inconnue, et responsable de la maladie. En revanche, nous pouvons considérer que dans les très rares cas de PV *JAK2* non-mutées, la recherche de mutations de l'exon 9 de *CALR* peut être utile, à titre de marqueur de clonalité.

Presence of Calreticulin Mutations in *JAK2*-negative Polycythemia Vera

Short title: CALR mutations in polycythemia vera

Julien Broséus[1-2], Ji-Hye Park[3], Serge Carillo[4-5], Sylvie Hermouet[6-7] and François Girodon[3-8, 9]

[1]Hematology Laboratory, University Hospital, Nancy, France;
[2]Inserm U954, University of Lorraine, Nancy, France;
[3]Hematology Laboratory, University Hospital, Dijon, France;
[4]Hematology Laboratory, University Hospital, Nîmes, France;
[5]UMR CNRS 5247, University of Montpellier, Montpellier, France;
[6]Hematology Laboratory, University Hospital, Nantes, France;
[7]Inserm U892, Institut de Recherche Thérapeutique, Nantes, France;
[8]Inserm U866, University of Burgundy, Dijon, France
[9] CRB Ferdinand Cabanne, Dijon, France

Key Points:

CALR mutations may rarely be observed in Polycythemia Vera negative for *JAK2* mutations.

Acknowledgments:

We thank Mr. Philip Bastable (University Hospital, Dijon, France) for revising the manuscript; and Mrs. Martine Courtois, Dominique Bouchot and Mathilde Bas (University Hospital, Dijon, France) for excellent technical assistance. This work was supported by a grant from the association 'Tulipes contre le cancer' (Châlon s/Saône, Burgundy, France).

Authorship:

Contribution: JB, SH and FG analyzed the data and wrote the paper. JHP performed the analyses. FG designed research. SC performed research and analyzed the data.

Conflict-of-interest disclosure: the authors have nothing to disclose.

Correspondence: François Girodon, Hematology Laboratory, PTB 2 rue Angelique

Ducoudray, CHU Dijon, 21079 Dijon, France. E-mail: francois.girodon@chu-dijon.fr

Phone: + 33 3 80 29 57 10; Fax : + 33 3 80 29 36 60

Abstract

Calreticulin (CALR) mutations have been reported in *JAK2 (Janus Kinase 2)*- and *MPL (Myeloproliferative Leukemia)*-negative essential thrombocythemia and primary myelofibrosis. On the contrary, so far, no *CALR* mutations were reported in cases of Polycythemia Vera (PV). Consequently, *CALR* mutations were considered exclusive to *JAK2* and *MPL* mutations. Here, we describe two *JAK2*^{V617F}- *JAK2*^{exon12}- negative PV patients who presented with a *CALR* mutation on peripheral granulocytes at the time of diagnosis. In both cases, the *CALR* mutation was a del 52 bp. At the same time, single BFU-E from one of the patients were grown *in vitro* and secondary genotyped: the same *CALR* del 52 bp was noted in 21 out of the 27 colonies tested. In spite of a not-yet-known mutation leading to the PV clone cannot be ruled out, our results suggest that *CALR* mutations may be rarely associated to *JAK2*-negative PV.

Introduction

Calreticulin (CALR) mutations have recently been reported in *JAK2*- and *MPL*-negative Myeloproliferative Neoplasms (MPN), particularly essential thrombocythemia (ET) and primary myelofibrosis (PMF).[1,2] The clinical course of sporadic *CALR*-mutated patients seems to be more indolent than that of *JAK2*-mutated patients.[1,3] On the contrary, so far, no *CALR* mutations were found in the 647 published cases of Polycythemia Vera (PV) patients tested. Consequently, *CALR* mutations were considered exclusive to *JAK2* and *MPL* mutations. Since 98% of PV patients harbor a *JAK2* mutation (mostly the V617F mutation in exon 14 and more rarely, in exon 12), the absence of *CALR* mutations in PV seemed logical. However, we and others have demonstrated that *JAK2*^{V617F} and *CALR* mutations can coexist in rare cases of refractory anemia with ring sideroblasts and thrombocytosis,[4] PMF,[5] or ET.[6,7] Here, we describe two *JAK2*^{V617F}-negative PV patients who presented with a *CALR* mutation at the time of diagnosis.

Patients and Methods

The procedures followed were in accordance with the Helsinki Declaration, and samples were obtained with patients' written informed consent. Two *JAK2* (exons 12, 13 and 14) - and MPL (exon 10)-negative PV patients were identified in our database and tested for a *CALR* mutation. Purification of granulocytes and peripheral mononuclear cells (PBMCs) and extraction of DNA were performed as described.[8] The *JAK2*^{V617F} mutation was analysed by allele-specific real-time

PCR to estimate the allele burden according to the method published by Lippert et al[8] with a sensitivity of 1%. *JAK2* exons 12, 13 and 14 analyses were performed according to the method by Carillo et al.[9] and the MPL mutations were analysed by High Resolution Melting (HRM) curve analyses followed by Sanger sequencing if positive, as previously published by Boyd et al.[10] The *CALR* exon 9 mutations were screened either by HRM (Carillo, personal communication), product sizing analysis and Sanger sequencing according to the methods by Klampfl et al.[1] Single burst-forming units–erythroid (BFU-E) colony assays were performed using PBMCs from patient #1 with methylcellulose (no. 5112) or collagen (no. 5411)-based media containing erythropoietin from Stem Alpha (Saint Denis l'Argentière, France). To ensure that plates with an optimal density were obtained so that colonies can be picked without contamination by cells from neighbouring colonies, PBMCs were plated at 4 different concentrations (20,000; 40,000; 100,000 and 200,000 per mL).

Results and discussion

Patient # 1 had hemoglobin at 168 g/L, hematocrit at 51.3%, and increased red cell mass (RCM) at 128% associated with a normal eythropoietin level. The bone marrow biopsy showed hypercellularity for age, panmyelosis associated with normal megakaryocytes and rare isolated abnormal enlarged forms. Using reticulin stain, no myelofibrosis was noted. Patient # 2 had hemoglobin at 194 g/L, hematocrit at 53% and low erythropoietin level without any dehydration (Table 1). Both had moderately elevated platelet counts (658 and 575 $x10^9$/L respectively) with normal leukocyte counts. They were negative for *BCR-ABL*. No mutation was found in *JAK2* exons 12, 13 and 14 by HRM and allele-specific real-time PCR[8, 9] or in *MPL* exon 10. Using HRM analyses, *CALR* mutations were suspected in both patients and therefore confirmed using Sanger sequencing and product sizing analyses : *CALR* mutations were in both patients type 1 deletions (52-bp deletion; c.1092_1143del) with estimated allele burden using product sizing analysis of 64% and Y% in patients #1 and #2, respectively. To complete genomic tests made on peripheral blood granulocytes, we performed colony assays in methylcellulose and in collagen, picked single BFU-E colonies grown after 14 days in the presence of erythropoietin, and genotyped each colony individually for *CALR*. On the 27 colonies genotyped, 6 had no PCR amplification and 21 harboured the same *CALR* mutation observed on peripheral blood granulocytes, i.e 52-bp deletion; c.1092_1143del. The BFU-E were found heterozygous for *CALR*, with a mean allele burden of 49%. To our knowledge, these patients are the first cases of *CALR*-mutated PV to be reported. However, due to a biclonal *JAK2*V617F and *CALR* MPN case recently reported,[6] we cannot rule out the possibility of a biclonal disease involving a not-yet-known mutation associated

with a *CALR* mutation. On the other hand, the presence of a *CALR* mutation both in peripheral granulocytes and in BFU-E suggests a potential role of the *CALR* mutation in the polycythemia phenotype. Our observations highlight the fact that in the absence of a *JAK2* mutation*, CALR* mutations can also be associated with PV. Of note, it has been reported that when RCM was measured in patients initially diagnosed with ET, it could in fact reveal masked PV with normal hematocrit and hemoglobin levels in 46.5% of cases.[11] Iron deficiency and/or splenomegaly may explain, in part, artificially normal hemoglobin levels in such PV patients. One can therefore imagine that the large series of published cases of *CALR*-positive ET may have included a few patients with undetected increased RCM, leading to an erroneous diagnosis of ET. In conclusion, our data indicate that testing *JAK2*-negative PV patients for *CALR* mutations may be useful.

References

1. 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. 2013.

2. 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. 2013.

3. 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. 2014.

4. 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. 2014.

5. Tefferi A, Lasho TL, Finke CM, Knudson RA, Ketterling R, Hanson CH, et al. CALR vs JAK2 vs MPL mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. Leukemia. 2014.

6. Lundberg P, Karow A, Nienhold R, Looser R, Hao-Shen H, Nissen I, et al. Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. Blood. 2014;123(14):2220-8.

7. McGaffin G, Harper K, Stirling D, McLintock L. JAK2 V617F and CALR mutations are not mutually exclusive; findings from retrospective analysis of a small patient cohort. Br J Haematol. 2014.

8. Lippert E, Boissinot M, Kralovics R, Girodon F, Dobo I, Praloran V, et al. The JAK2-V617F mutation is frequently present at diagnosis in patients with essential thrombocythemia and polycythemia vera. Blood. 2006;108(6):1865-7.

9. Carillo S, Henry L, Lippert E, Girodon F, Guiraud I, Richard C, et al. Nested high-resolution melting curve analysis a highly sensitive, reliable, and simple method for detection of JAK2 exon 12 mutations--clinical relevance in the monitoring of polycythemia. J Mol Diagn. 2011;13(3):263-70.

10. Boyd EM, Bench AJ, Goday-Fernández A, Anand S, Vaghela KJ, Beer P, et al. Clinical utility of routine MPL exon 10 analysis in the diagnosis of essential thrombocythaemia and primary myelofibrosis. Br J Haematol. 2010;149(2):250-7.

11. Cassinat B, Laguillier C, Gardin C, de Beco V, Burcheri S, Fenaux P, et al. Classification of myeloproliferative disorders in the JAK2 era: is there a role for red cell mass? Leukemia. 2008;22(2):452-3.

	Patient 1	Patient 2
Age, years	**69**	**83**
Sex, M/F	**M**	**M**
Hemoglobin level, g/L	**168**	**194**
Hematocrit, %	**51.3**	**53**
Platelet count, 10^9/L	**658**	**575**
White blood cells, 10^9/L	**7.6**	**7.5**
Erythropoietin level, mUI/mL	**5.9**	**2.8**
Bone marrow biopsy	**PV**	**N/A**
Red cell mass	**increased**	**N/A**
***JAK2*^{V617F} mutation**	**Negative**	**Negative**
***JAK2* exon 12 mutation**	**Negative**	**Negative**
***JAK2* exon 13 mutation**	**Negative**	**N/A**
***JAK2* exon 14 mutation**	**Negative**	**N/A**
***MPL* mutation**	**Negative**	**Negative**
***BCR-ABL1* mutation**	**Negative**	**Negative**
***CALR* mutation**	**Positive**	**Positive**

Table 1: characteristics of the two patients at diagnosis. N/A: Non Available; MPN: Myeloproliferative Neoplasm; PV: Polycythemia Vera; *JAK2*: Janus Kinase 2; *MPL*: Myeloproliferative Leukemia; *BCR-ABL*: Breakpoint Cluster Region - Abelson; *CALR*: Calreticulin.

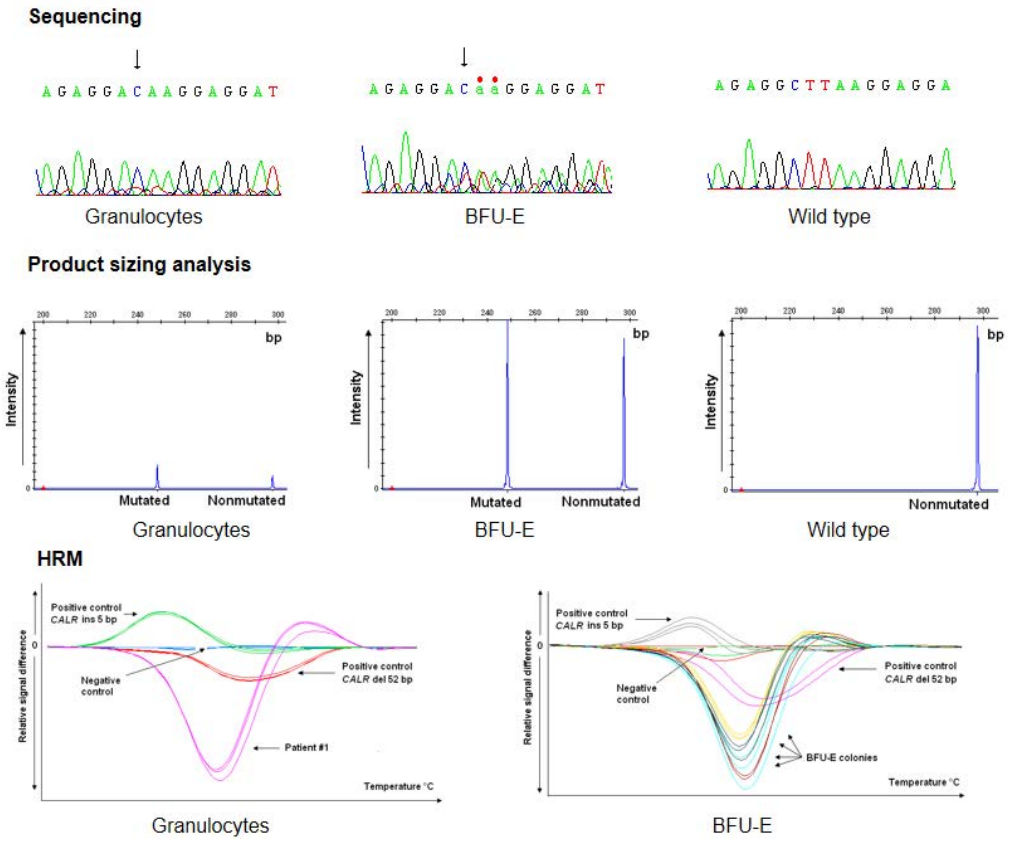


Figure 1: *CALR* mutations determined using sequencing, product sizing analysis and High Resolution Melting (HRM) curve analyses in peripheral blood granulocytes and BFU-E in patient #1.

PARTIE 2 : GENOMIQUE DES LYMPHOMES B-DIFFUS A GRANDES CELLULES EN RECHUTE

1. RELAPSED DIFFUSE LARGE B-CELL LYMPHOMA PRESENT DIFFERENT GENOMIC PROFILES BETWEEN EARLY AND LATE RELAPSES

Contexte

Près de 40% des patients atteints de LBDGC rechutent après le traitement de première ligne. Chez ces patients, le pronostic est réservé en raison de la faible efficacité des traitements de rattrapage.[93, 153] Les rechutes sont réparties équitablement entre rechutes précoces, c'est-à-dire se déclarant moins d'un an après la fin du traitement de première ligne, et rechutes tardives, c'est-à-dire se déclarant plus d'un an après la fin du traitement. Le délai de la rechute a une valeur pronostique.[91, 154] Le pronostic est particulièrement sombre pour les patients en rechute précoce, avec une survie de 17% à 3 ans comparativement aux rechutes tardives qui ont une survie de 50% à 3 ans.(93) Sur le plan génomique, les LBDGC sont très bien caractérisés au diagnostic,[100, 173] mais peu étudiés à la rechute.[174] Les anomalies du fonctionnement cellulaire déterminantes dans les LBDGC en rechute sont peu connues.

Objectifs

Notre objectif est de réaliser une étude génomique comparative des LBDGC en rechute précoce et des LBDGC en rechute tardive pour identifier les anomalies génétiques associées à la précocité de la rechute.

Méthodologie

Pour 39 patients issus de la cohorte CORAL,[153] nous disposons de données génomiques issues de la puce Affymetrix SNP 6.0. Cette puce contient 1 800 000 sondes couvrant le génome, à raison d'une sonde toutes les 1800 bases en moyenne. Cet outil permet donc la recherche d'anomalies de nombre de copies à haute résolution.[162] Pour garantir la significativité des résultats, nous avons mis au point une méthode originale, adaptée de la méthode SAM (*Significance Analysis of Microarray*).[175] Cette méthode est fondée sur la permutation des données ou *bootstrap*, et permet de réduire le nombre de faux positifs dus aux tests multiples.

Cette méthode a été conçue initialement pour analyser les données d'expression génique. Dans ce travail, nous l'avons appliquée à l'analyse des anomalies de nombre de copies. Nous avons lissé les données brutes issues de la puce en attribuant à une sonde donnée un nombre de copies correspondant à la médiane des 11 sondes consécutives encadrant cette sonde. Les gains sont définis comme un nombre de copies > 2,5 et les pertes comme un nombre de copies < 1,5. Une valeur comprise en 1,5 et 2,5 correspond à un nombre normal de copies. Pour chaque gène, nous avons comparé la distribution des anomalies entre nos deux groupes, à savoir les LBDGC en rechute précoce (n=19) et les LBDGC en rechute tardive (n=20). Nous n'avons retenu que les plus significatifs, avec un *False Discovery Rate* (FDR) < 5%.

Principaux résultats

Les anomalies de grande taille (≥ 2Mb) sont retrouvées en quantités équivalentes entre rechutes précoces et tardives. L'analyse d'une sélection de 36 gènes connus pour être associés aux LBDGC ne montre pas de différence entre rechutes précoces et tardives en terme de distribution des anomalies de nombre de copies.

Notre méthode d'analyse a permis d'identifier une liste de 56 gènes codants pour une protéine et 27 longs ARN non-codants dont les anomalies de nombre de copies présentent une distribution significativement différente entre rechute précoce et rechute tardive, avec un FDR < 0,05. Les rechutes précoces sont associées à des anomalies géniques affectant les régulateurs du cycle cellulaire, de l'apoptose et de la transcription. Les rechutes tardives sont associées à des anomalies affectant la réponse immune et la prolifération cellulaire.

La classification hiérarchique des anomalies de nombre de copies ainsi découvertes a permis de révéler 3 groupes de patients présentant un profil d'anomalies corrélées. Le premier groupe était composé exclusivement de rechutes précoces, le deuxième groupe exclusivement de rechutes tardives et le troisième groupe comprenait à la fois des rechutes précoces et des rechutes tardives, sans anomalie récurrente.

Lorsque les données d'expression étaient disponibles pour un nombre suffisant d'échantillons, nous avons pu démontrer une conséquence des anomalies de nombre de copies sur le niveau d'expression du gène concerné.

Conclusion

Même si les LBDGC constituent un ensemble hétérogène, nous avons pu identifier des anomalies récurrentes dans les rechutes précoces et dans les rechutes tardives, ceci en appliquant une méthode statistique originale dérivée du SAM. Nous avons identifié deux signatures, correspondant respectivement aux rechutes précoces et aux rechutes tardives. La signature des rechutes précoces comprend des anomalies affectant le cycle cellulaire, l'apoptose et la transcription. Cette signature peut expliquer l'avantage sélectif conféré au clone malin, ainsi que sa chimiorésistance.

Relapsed diffuse large B-cell lymphoma present different genomic profiles between early and late relapses

Broséus J,[1-2] Chen G,[3#] Hergalant S,[1#] Ramstein G,[4] Mounier N,[5] Guéant JL,[1,6] Feugier P,[1,7] Gisselbrecht C,[8] Thieblemont C[8,9*] and Houlgatte R[1,10*]

[1]Inserm U954, Faculty of Medicine, Nancy, France
[2]Hematology, Laboratory department, University Hospital of Nancy, Nancy, France
[3]ZhongNan Hospital of Wuhan University, Wuhan, China
[4]LINA DUKe, UMR 6241, Université de Nantes, Nantes, France
[5]Hemato-oncology, University Hospital of l'Archet, Nice, France
[6]Biochemistry, Laboratory department, University Hospital of Nancy, Nancy, France
[7]Hematology department, University Hospital of Nancy, Nancy, France
[8]APHP, Saint-Louis Hospital, Hemato-oncology department; Paris, France
[9]Paris Diderot University-Sorbonne Paris-Cité, Paris, France
[10]DRCI, University Hospital of Nancy, Nancy, France

[#]These authors contributed equally to this work
*Last authors equally contributed to this work.

Corresponding authors:

Rémi Houlgatte, Inserm U954, Faculty of Medicine, Avenue de la Forêt de Haye 54500 Vandoeuvre-lès-Nancy, France.
Tel: +33 3 83 68 32 92; Fax: +33 3 83 68 32 79
E-mail: remi.houlgatte@inserm.fr

Catherine Thieblemont, APHP, Saint-Louis Hospital, Hemato-oncology department, 1 avenue Claude Vellefaux, 75010 Paris, France
Tel: +33 1 42 38 51 34; Fax: +33 1 42 49 96 41
E-mail: catherine.thieblemont@aphp.fr

Keywords: diffuse large B-cell lymphoma, early relapse, late relapse, genomics, copy number variations

Acknowledgments

The authors would like to thank fondation Nelia et Amedeo Barletta, Paris, France.
We thank the LYSARC for providing the clinical data; Fabienne Morand, and the project leaders from the different countries; the patients and their families; all investigators and pathologists.

Conflicts of interest

Professor C. Gisselbrecht: Research Funding by Roche. The other authors declare no competing financial interests.

Grant support

The authors would like to thank fondation Nelia et Amedeo Barletta, Paris, France.
We thank the Regional Council of Lorraine for funding through the "Chercheurs d'Excellence" program.

Abstract:

Despite major advances in first-line treatment, a significant proportion of patients with diffuse large B-cell lymphoma (DLBCL) will experience treatment failure. Prognosis is particularly poor for relapses occurring less than one year after the end of first-line treatment (early relapses/ER) compared to those occurring more than one year after (late relapses/LR). To better understand genomic alterations underlying the delay of relapse, we identified copy number variations (CNVs) on 39 tumor samples from a homogeneous series of patients included in the Collaborative Trial in Relapsed Aggressive Lymphoma (CORAL) prospective study. To identify CNVs associated with ER or LR, we devised an original method based on Significance Analysis of Microarrays, a permutation-based method which allows control of false positives due to multiple testing. Deletions of *CDKN2A/B* (28%) and *IBTK* (23%) were frequent events in relapsed DLBCLs. We identified 56 protein-coding genes and 25 long non-coding RNAs with significantly differential CNVs distribution between ER and LR DLBCLs, with a false discovery rate < 0.05. In ER DLBCLs, CNVs were related to transcription regulation, cell cycle and apoptosis, with duplications of histone *H1T* (31%), deletions of *DIABLO* (26%), *PTMS* (21%) and *CK2B* (15%). In LR DLBCLs, CNVs were related to immune response, with deletions of *B2M* (20%) and *CD58* (10%), cell proliferation regulation, with duplications of *HES1* (25%) and *DVL3* (20%), and transcription regulation, with *MTERF4* deletions (20%). This study provides new insights into the genetic aberrations in relapsed DLBCLs and suggest pathway-targeted therapies in ER and LR DLBCLs.

Introduction:

Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma in adults [1]. The addition of the anti-CD20 monoclonal antibody rituximab to the chemotherapy CHOP (cyclophosphamide, adriamycin, vincristine and prednisone) has dramatically improved the outcome of patients with DLBCL [2, 3]. Despite this major advance a significant number of patients will experience treatment failure. At such a time, treatment is a real challenge and options are to be discussed regarding the eligibility or not for transplant and prognostic factors at relapse [4, 5]. One of the most important prognostic parameters in this context is the time of relapse [6, 7]. Two groups of patients are identified. The first group of early relapse (ER) is defined as relapse occurring within 1 year after the initial treatment. The second group of late relapse (LR) is defined as relapse occurring more than 1 year after the initial treatment. In the ER group, the overall survival (OS) is estimated at less than 17% at 3 years, while in the LR group, the OS is estimated at about 50% at 5 years [4].

The heterogeneous prognosis of patients with DLBCL reflects a deep biological diversity. DLBCL is composed of at least 2 molecular subtypes that differ in gene expression profile (GEP): germinal center B-cell like (GCB) and activated B-cell like (ABC), each with different prognosis [8]. The prognostic impact of this molecular classification has been reported in several independent series and different types of treatment at diagnosis as well as at relapse [9-11].

Recent advances in genomic technology have provided further understanding in the biology of DLBCL. Constitutive activation of the NF-κB pathway, owing to oncogenic mutations activating BCR (B-Cell Receptor) and toll-like receptor signaling, leads to cell proliferation and resistance to apoptosis. These constitutive activations, characterizing ABC DLBCLs, are related to somatic gain of function mutations and/or duplications affecting positive regulators of this pathway such as *CD79a* and *CD79b* [12], *MYD88* [13], *CARD11* [14], *TRAF2*, *TRAF5*, *MAP3K7/TAK1* and *TNFRSF11A/RANK* [15] or somatic inactivating mutations and/or deletions of negative regulators such as *TNFAIP3/A20* [15,16]. The GCB subtype is associated with abnormalities of epigenetic modifiers, with frequent bi-allelic point mutations leading to the loss of function of the histone lysine methyl-transferase *MLL2* [17,18], high frequency of gain-of-function mutations in the histone lysine methyl-transferase *EZH2* [19], inactivating mutations and deletions of the histone lysine acetyltransferases *CREBBP* and *EP300* [18], and mutations of their cofactor, *ME2FB* [17]. Constitutive activation of the PI3K/AKT/mTOR signaling pathway through *PTEN* deletions, *MiR17-92* locus amplifications and mutations of *PIK3CD, PIK3R1* and *mTOR* is another oncogenic mechanism commonly observed in GCB DLBCLs [20, 21]. Dysregulation of cell cycle is a common event in all DLBCLs, potentially owing to an increased expression of the anti-apoptotic factor *BCL2* [22-24], mono and bi-allelic deletions affecting the *CDKN2A/B* locus [22,25], point mutations and allele deletions leading to the loss of function of TP53 [26], or dysfunction of the transcription factor MYC, which has been described in 5 to 11% of *de novo* DLBCLs [27] and 17% of relapsed DLBCLs [28]. Constitutive activation of the NOTCH pathway through *NOTCH1* and *NOTCH2* somatic gain-of-function mutations or gain of copy numbers [29-31] cause an increased cell proliferation in B-cell lymphoma. *BCL6* expression beyond the germinal center reaction leads to increased tolerance to DNA breaks and blockade of plasmablastic differentiation [32, 33]. Importantly, escape to immune response through biallelic inactivations of *B2M* leading to defects of surface HLA class I molecules assembly [34], downregulation of surface HLA class II molecules expression due to breakpoints and mutations of *CIITA* [35] or deletions and truncating mutations of the *CD58* locus [34] represent another major player in lymphomagenesis in DLBCLs.

All these genomic studies have been conducted on *de novo* DLBCLs at diagnosis. Cellular functions and signaling pathway disruptions associated with relapsed DLBCLs remain largely unknown. Here we describe the number and the type of genetic alterations present in relapsed

DLBCLs and make a comparison of Copy Number Variations (CNVs) between ER and LR DLBCLs to identify genetic alterations that are associated with the delay of relapse. To achieve this outcome, we designed an original method allowing the statistical assessment with the high level of confidence required when dealing with a very large set of candidates. Finally we evaluated the impact of relevant CNVs on gene expression.

Results:

Landscape of large copy number variations in relapsed DLBCL

CNVs ≥ 2 megabases (Mb) were counted to evaluate the complexity of structural aberrations of relapsed DLBCL (Figure 1). For the whole group, average CNVs per sample was 15.50 (ranging from 0 to 67). We found an equivalent number of duplications and deletions with an average 8.30 duplications per sample (0-43) and 7.20 deletions per sample (0-34; p=0.60, Student's t-test). Chromosomes 1, 2, 3, 6, 12 and 18 were the most frequently altered, with 1.38 CNV per chromosome per sample on average while other chromosomes presented with an average of 0.45 CNV per chromosome per sample ($p=6x10^{-4}$).

In the ER group, the average CNVs per sample was 14.89. There was a wide heterogeneity within this group as samples presented with 0 to 67 CNVs ≥ 2 Mb. Proportions of duplications and deletions were comparable with an average 8.15 duplications (0-34) and 6.73 deletions per sample (0-33; p=0.61). Chromosomes 1, 2, 6, 8, 12 and 18 were the most frequently altered, with an average of 1,58 CNV per chromosome per sample, whereas other chromosomes presented with an average of 0.34 CNV per chromosome per sample ($p=6x10^{-3}$).

In the LR group, the average CNVs per sample was 16.1. CNVs varied from 0 to 50 per sample within this group. Duplications and deletions were present in equivalent proportions with an average of 8.45 duplications per sample (0-43) and 7.65 deletions per sample (0-34), respectively (p=0.81). Chromosomes 2, 3, 6, 7, 15, 18 were the most frequently altered, with an average of 1.55 CNV per chromosome, whereas remaining chromosomes presented with 0.42 CNV per chromosome on average ($p=7x10^{-4}$).

ER and LR DLBCLs samples were affected by an equivalent total CNVs number with an average of 14.89 and 16.11 CNVs per sample, respectively (p=0.81). Comparing ER and LR for each chromosome revealed differential signatures on chromosome 1 only, with an average of 2.63 CNVs per sample in the ER group and 0.70 CNV per sample in the LR group ($p=6x10^{-3}$).

Identification of genes differentially altered among the ER and the LR subgroups

After smoothing all the copy number probes using the sliding window technique, we observed that 0.38% of the values were below 1 and 1.53% above 3. This means that almost all losses and gains of copy numbers corresponded to heterozygous deletions and duplications, respectively.

The use of Significance Analysis of Microarrays (SAM), a permutation-based method built on discriminating scores and their significance [36] identified 148 genes with significantly differential CNVs distribution between ER and LR DLBCLs with a false discovery rate (FDR) < 0.05. Among these genes were 56 protein coding genes (Table 1; Figure 2), 25 Long non-coding RNA (LncRNA), 22 antisense RNAs, 2 Small Nucleolar RNAs, 1 sense overlapping RNA, 1 miscellaneous RNA and 41 pseudogenes (Supplemental table 1).

ER DLBCLs were characterized by frequent duplications affecting *HIST1H1T* (6/19 samples, 31%), *HIST1H2BC* (5/19 samples, 26%), *HIST1H2AK* (5/19 samples, 26%), and *PTMS* (4/19 samples, 21%), implicated in histone-related regulation of transcription. We found a high number of duplications of *AIRN* (4/19 samples, 21%), a LncRNA interacting with the chromatin remodeling complex G9a. Deletions of *HNRNPU* (4/19 samples, 21%), implicated in pre-mRNA processing and alternative splicing, were also present. In addition, deletions of programmed-cell-death implicated genes *DIABLO* (5/19 samples, 26%) and *CSNK2B/CK2B* (3/19 samples, 15%), were identified. Duplications of *LGALS9C* (9/19 samples, 47%), deletions of *BTNL2* (5/19 samples, 26%) and *HLA-DRA* (4/19 samples, 21%) were also revealed. A significant fraction of samples was affected by duplications of *CCHCR1* (6/19 samples, 31%), a regulator of keratinocyte proliferation.

LR DLBCLs were characterized by a high number of duplications affecting *DVL3* (4/20 samples, 20%), a regulator of cell proliferation. We found frequent loss of copy number for *LGALS9C* (9/20 samples, 45%), *GALNT5* (4/20 samples, 21%), implicated in O-linked oligosaccharide biosynthesis, *ADSL* (3/20 samples, 15%), involved in *de novo* purine biosynthesis and *ART5* (4/20 samples, 20%), an ADP-ribosyltransferase. Genes implicated in BCR (*IgHV*, *IgKV*) and T-Cell Receptor (TCR) assembly *(TRAJ* and *TRAV*) and transcriptional regulation, such as *NFE2l2*, *MTERF4*, *HES1*, *EXOSC8*, *NARFL* and *ZNF414*, were affected by either duplication or deletion in at least 20% of LR DLBCL samples.

Among the subset of 36 genes selected for their implication in lymphomagenesis, the 10 most frequently altered were *CDKN2A/B* (43%)*, IBTK* (33%)*, PRDM1* (33%)*, FOXO1* (25%)*, BCL2* (25%)*, TP53* (20%)*, TNFAIP3* (20%)*, XPO1* (15%)*, B2M* (15%) and *CD58* (12%)(Figure 3). Cell cycle regulation and apoptosis were the most represented of affected pathways, with a high

number of alterations encompassing *CDKN2A/B*, *PRDM1*, *BCL2*, *TP53* and *XPO1*. Deletions of *CDKNA2/B* and *IBTK* were identified in 28% and 23% of DLBCL samples, respectively.

ER DLBCLs harbored frequent deletions of *CDKN2A/B* (8/19 ER samples, 42% *versus* 3/20 LR samples, 15%), *PIM1* (3/19 ER samples, 15% *versus* 0/20 LR sample, 0%), *MLL2* (4/19 ER samples, 21% *versus* 1/20 LR sample, 5%) and *CREBBP* (2/19 ER samples, 10% *versus* 0 LR sample). In LR DLBCLs, we identified CNVs affecting *B2M* (5/20 LR samples, 25% *versus* 1/19 ER samples, 5.26%) (p=0.1), deletions of *B2M* being restricted to LR DLBCLs, and *CD58* (deletions in 2/20 LR samples, 10% *versus* duplications in 3/19 ER samples, 15.7%) (p=0.07).

When we compared CNVs distribution among ER and LR samples, we did not find any statistically significant difference. We also did not observe any correlation between CNVs distribution and the GC/ABC subtype (data not shown).

Evidence on recurrent events in relapsed DLBCLs

We performed a hierarchical clustering on 56 protein-coding genes and 25 LncRNA CNVs to identify events recurrently associated with ER and LR DLBCLs (Figure 4). The resulting heatmap shows recurrent CNVs in DLBCLs and defines three groups of patients.

The first group is homogeneous and comprises 7 ER samples presenting with correlated amplifications of *HIST1H2BC*, *HIST1H1T*, *HIS1H2AK*, *CSNK2B/CK2B*, *KAAG1*, *CCHCR1*, SERPINB1 and *IgKV,* as well as correlated deletions of *HRNPU* and four LncRNAs: *RP11-527D7.1*, *MTND4P10*, *MTND5P19* and *RP4-580N22.2*.

The second group is composed of 9 LR DLBCLs showing correlated amplifications of *AMY2A*, *AADAC*, *MUC20*, *DVL3*, *SENP5*, *CPN2*, and *HES1* as well as correlated deletions of *MTERF4*, *NFE2L2*, *GALNT5*, *ADSL* and three LncRNAs: *AC007879.6*, *AC092625.1*, *and AC019186.1*.

Remaining ER and LR samples constituted a third group without specific CNV profiles and no recurrent events.

Impact of CNVs on gene expression

We examined the impact of CNVs on average gene expression level (AGEL) (Figure 5).

In ER DLBCLs, deletions of *DIABLO* were related to a decrease in its AGEL, with a fold change of 0.58 between deleted and normal samples ($p=5x10^{-4}$, Student's t-Test). We observed the same correlation between deletions of the genes and decrease in their expression level for *PTMS, HNRPU, and CSNK2B*. Deletions of these three genes resulted in a decrease in their respective AGELs, with a fold change of 0.47, 0.71, and 0.53, respectively (p=0.01, 0.02 and 0.03).

In LR DLBCLs, deletions of *MTERF4* were related to a decrease in its AGEL, with a fold change of 0.51 between deleted and normal samples ($p=8x10^{-7}$). Accordingly, deletions of *LGALS9C* led to a decrease in AGEL compared to normal samples (fold change=0.54; p=0.05).

In the set of 36 genes, deletions of *CDKN2A/B* were related to a decrease in AGEL (fold change=0.31; p=0.07). Deletions of *CD58* were also related to a decrease in AGEL (fold change=0.65; p=0.08).

Discussion:

Understanding the relationship between tumor biology and outcome is important for identifying molecular targets and may yield more effective therapies for DLBCLs. Almost all genomic studies reported to date were performed on *de novo* DLBCLs at diagnosis and little is known about structural aberrations of refractory/relapsed DLBCLs. Here we analyzed one a the few series of relapsed DLBCLs published to date [37, 38], composed of 39 samples from the Collaborative Trial in Relapsed Aggressive Lymphoma (CORAL) trial, a unique and well-documented prospective study on relapsed DLBCLs [7]. Studying this thirty-nine-patient cohort allowed us to make strong statements since we devised and used for the first time a robust, comprehensive, unsupervised analysis method to analyze CNVs. This method makes use of SAM, a statistical technique originally designed for gene expression data analysis [36]. This permutation-based method allowed the discovery of a list a gene containing less than 5% of false positives despite multiple testing, independent from data distribution. Applied to oligonucleotide microarrays covering the entire genome, the method allowed us to delineate statistically validated CNVs with great precision and locate structural aberrations linked with ER and LR DLBCLs. Other published algorithms such as GISTIC [39], GEDI [40], or ComFocal [25] are designed to identify recurrent regions of deletions or duplications, which is not the purpose of this work. Here we aimed at identifying CNVs strongly associated with ER or LR. Which is the main reason motivating the development of this original method.

ER and LR DLBCLs display a comparable CNV landscape, with a high degree of heterogeneity within each group. In relapsed DLBCLs, the most frequently altered genes were related to cell cycle regulation and apoptosis, regulation of transcription, immune response and NF-κB pathway (Figure 6). Among genes associated with lymphomagenesis, CNVs affecting *CDKN2A/B* locus (42% of samples) were the most frequents, with a majority of deletions (28% of samples). CDKN2A is a cycline-dependent kinase restraining cellular proliferation. These results are consistent with previous reports describing 30-35% of *CDKN2A* deletions in primary DLBCLs, predominantly in ABC subtype, with subsequent increase in cell proliferation and poorer outcome [22, 41, 42]. We also observed a high rate of *IBTK* aberrations (33% of samples), mainly deletions

(23% of samples). IBTK (Inhibitor of Burton Tyrosine Kinase) is a BTK binding protein which inhibits BTK's kinase activity and disrupts BTK-mediated calcium mobilization leading to a decrease in BTK-induced activation of the anti-apoptotic NF-κB pathway [43]. A recent study reported a high rate of complete or partial responses and long-lasting remissions with ibrutinib treatment in relapsed/refractory DLBCLs, essentially in ABC subtype [44]. In this work, we did not observe any correlation between CNVs and the GC/ABC subtype. These results are in keeping with a recently published study showing (i) an absence of correlation between known driver mutations in lymphomagenesis and response to therapy, and (ii) an equivalent distribution of genomic events predicting survival across GC and ABC subtypes [37].

When we compared ER and LR DLBCLs at the gene scale, aberrant genomic profiles were distinct on 56 protein-coding genes, with a subsequent effect on gene expression levels. In some cases, we were unable to demonstrate a statistically significant effect on gene expression. This can be explained by intronic localization of CNVs.

First, structural aberrations of cell cycle and apoptosis regulators were rather associated with ER DLBCLs, with deletions of *DIABLO/Smac*, a promoter of apoptosis which allows activation of caspases by binding to inhibitor of apoptosis proteins [45], and deletions of *CSNK2B/CK2B*, the regulatory subunit of protein kinase CK2, involved in a wide number of cellular processes, such as regulation of programmed cell death, cell division and proliferation, DNA damage repair, gene transcription and protein translation. CK2 increased protein expression and enzymatic activity has been described to positively regulate the PI3K pathway and STAT3 and NF-κB signaling in hematological malignancies [46]. We hypothesize that deletions of the *CSNK2B/CK2B* regulatory subunit with subsequent decrease in gene expression level result in increased enzymatic activity. LR DLBCLs were associated with duplications of *DVL3*, a member of the disheveled protein family which has been described to promote cell growth in ALK-positive anaplastic large cell lymphoma [47].

Second, CNVs encompassing genes related to immune response included duplications of *LGALS9C* identified in 47% of ER DLBCLs. LR DLBCLs were associated with partial deletions and duplications of *TRAJ* and *TRAV*, both implicated in TCR assembly. This suggests infiltration of the tumor by T-cell lymphocytes as part of the immune response. Deletions of *B2M* (20%) and *CD58* (10%) were restricted to LR DLBCLs. Deletions and duplications of HLA related genes, were observed frequently in our cohort and are most likely related to HLA system polymorphism. Losses in *IgHV* and *IgKV* clusters are presumably a sign of BCR rearrangement and cannot be considered as recurrent lesion in tumor samples. We made the choice to consider all CNVs, including those that are described in databases of genomic variants. Limiting our analysis to

acquired lesions may avoid constitutional abnormalities that are prone to foster lymphoma development.

Third, dysregulation of transcription through altering events in genes encoding regulators of chromatin compaction and DNA replication were associated with ER DLBCLs. Duplications of the linker histone H1 gene and/or deletions of its regulatory partner *PTMS* [48] were found in 21 to 31% of cases. This observation is consistent with a previous report, where mutations of the linker histone H1 gene have been described as a key evolutionary process governing transformation of follicular lymphoma into DLBCL [49]. LR DLBCLs were linked with deletions of *MTERF4*, a mitochondrial transcription regulator [50], in 20% of cases. LncRNA are nucleic acids often longer than 2 kb involved in the regulation of transcription through epigenetic modifications by recruiting chromatin remodeling complexes to various genomic loci [51]. *AIRN (*Air noncoding RNA) was of particular interest as it was deleted in 20% of ER DLBCLs. This LncRNA plays a role in epigenetic regulation of transcription by cooperating with the histone methyl transferase G9a [52]. Duplications of *HES1*, a downstream effector of the NOTCH pathway associated with Fanconi anemia [53] and blast crisis transformation in chronic myelogenous leukemia [54] were restricted to LR samples (25%).

Even though DLBCL is a heterogeneous disease, our method allowed the identification of structural aberrations that are statistically associated with ER and LR. Moreover, hierarchical clustering of CNVs identified three groups of patients. Two groups (the first one composed of ER samples and the second one of LR samples) were well separated in homogeneous clusters with common deletions and duplications. The remaining samples formed a mixed group of ER and LR samples with no recurrent CNV. This result gives evidence of a continuum between ER and LR DLBCLs. Identifying two groups of samples with correlated CNVs profiles demonstrates the statistical significance of our results, since these correlated groups could not be identified by chance.

Our original method unraveled genomic aberrations typical of the "LR signature", which includes genomic aberrations encompassing genes related to immune response, cell proliferation and regulation of transcription. The "ER signature" includes events leading to disruption of cell cycle, apoptosis and gene transcription and provides a structural basis for selective growth advantage and chemotherapy resistance leading to a worse prognosis. This is of great importance since it may be a starting point for further studies and supports the rationale for targeted therapy. Even if this study needs to be confirmed on a larger cohort of patients, it provides new insights into the genetic complexity and dysregulated pathways in relapsed DLBCLs, suggesting the possibility of individualized targeted therapy for those patients with poor prognosis.

Materials and methods:

Patient cohort

The phase III CORAL prospective study included 396 patients presenting with relapsed/refractory DLBCL after first-line treatment with R-CHOP (rituximab, cyclophosphamide, adriamycin, vincristine and prednisone). These patients were randomly assigned to receive a salvage chemotherapy regimen, either rituximab, ifosfamide, carboplatin and etoposide (R-ICE) or rituximab, dexamethasone, cytarabine, and cisplatin (R-DHAP). Patients who responded to the chemotherapy were submitted to high dose therapy and autologous stem cell transplant (ASCT). After ASCT, the trial compared rituximab treatment every 2 months for 1 year with observation alone. The initial results revealed no significant difference in outcome. Several factors did affect survival, including ER, the international prognostic index at relapse, and prior exposure to rituximab [7].

From these 396 patients with relapsed/refractory DLBCL, frozen lymph node biopsies were available from a homogeneous group of 39 patients. A part of tumors were sampled at diagnosis, before first-line treatment, and the others were sampled at relapse, before salvage chemotherapy. Thus, among the 19 tumor samples composing the ER group, 11 were sampled at diagnosis, and 8 at relapse. Among the 20 tumor samples from the LR group, 9 were sampled at diagnosis and 11 at relapse. Molecular characterization regarding GCB DLBCLs and ABC DLBCLs could be determined with GEP in 30 cases. Characteristics of the patients are listed in table 2. Procedures used were in accordance with the Helsinki Declaration of 1975 (as revised in 1983).

Copy number analysis

CNVs of the 39 samples were determined using the Affymetrix SNP 6.0 platform (Affymetrix, Santa Clara, CA, USA), which interrogates 1 800 000 copy number probes distributed evenly along the human genome (mean interval: 1.8 kilobases) [55]. The 1 800 000 probes were annotated with home-made tools querying the Ensembl 77th version with Human Genome HG37 MYSQL database available at ftp.ensembl.org/pub/release-77/mysql/. Total copy numbers were computed using the Bioconductor crlmm package [56]. A sliding window (length=11 probes) representing the signal median value was used to smooth artifactual data.

This strategy allowed CNV discovery at high resolution and identification of genes frequently altered in relapsed DLBCLs. Duplications were defined by a copy number ≥2.5, deletions were defined by a copy number ≤1.5 and copy numbers between 1.5 and 2.5 were considered normal,

according to an already published method [22]. We used Java Treeview 3.0 [57] to visualize the results at a genomic scale.

Identification of differential genomic alterations between ER and LR

We aimed at identifying genes affected by differentially distributed CNVs between ER and LR DLBCL patients, *ie* cases where deletions and duplications are associated preferentially with ER or LR. Given the large number of statistical tests, we used a permutation-based analysis to reduce false-positive results due to multiple testing. Our method was based on the widely used technique called Significance Analysis of Microarrays (SAM) [36]. SAM is a non-parametric test that avoids unjustified assumptions over gene distribution and gene independence. SAM was originally designed to determine whether changes in gene expression are experimentally significant. As claimed by the authors, SAM can be adapted for other experimental types of data. For our purpose, one needs to define a measure of the strength of the relationship between the CNV distribution (loss of copy number, normal copy number and gain of copy number) and ER or LR. Many statistics such as Pearson's chi-squared test, and other measures such as the Kullback–Leibler divergence, are only defined in case of non-null frequencies. Therefore we preferred to select the Hellinger distance $H(P,Q)$ which quantifies the difference between probability distributions $P=(p_1,\ldots, p_k)$ and $Q=(q_1,\ldots, q_k)$. The Hellinger distance is defined as:

$$H(P,Q)=\frac{1}{\sqrt{2}}\sqrt{\sum_{i=1}^{k}\left(\sqrt{pi}-\sqrt{qi}\right)^2}$$

The lower bound of $H(P,Q) = 0$ corresponds to $P=Q$ and its upper bound $H(P,Q) = 1$ implies that $p_iq_i = 0 \ \forall \ i \in [1,k]$. In our method, P and Q represent the probability distributions over the ER and LR DLBCL copy numbers, expressed in three discrete states (deletion, normal, duplication).

Genes were considered positive when the FDR was inferior to 5%. For a FDR of 5%, our algorithm can be stated as: (i) compute the observed Hellinger distances for all the probes mapping to gene g_i. The distance associated to g_i is defined as the mean of the probe distances. (ii) form an observed vector H_i by ordering the observed gene distances. (iii) perform $B = 20$ random permutations over the patient labels and compute the corresponding expected null gene distances. Order the B distance vectors and compute the corresponding mean expected vector $H_{Ei.}$ (iv) plot the observed vector against the expected null vector (Supplemental Figure 1). (v) set a threshold $\Delta = \max(H_i - H_{Ei})$ (vi) define the significant genes verifying $H_i - H_{Ei} \geq \Delta$ (vii) estimate

the FDR. This value can be computed as the average of genes called significant from all the *B* permutations. (viii) if FDR > 0.05, decrease the threshold Δ and repeat steps 6 and 7.

Then, the aim of this method is not to provide an individual p-value for each gene of the genome but to identify of a list of genes which CNVs distribution is differential between ER and LR, with systematic correction of multiple testing and less than 5% of false positives.

SNP microarray data acquired on Affymetrix SNP 6.0 platform can be accessed online on Gene Expression Omnibus (GSE73791).

Selected set of genes implicated in lymphomagenesis

We also explored a subset of 36 genes selected from literature and already known to play a role in lymphomagenesis through the dysregulation of key cellular pathways [58]. *CDKN2A, CDKN2B, PRDM1, TP53, BCL2, XPO1, MYC, GNA13, MFHAS1* and *KLF4* were chosen for their implication in cell cycle and apoptosis; *MLL2, CREBBP, EZH2* and *EP300* for their role in epigenetic regulation; *B2M, CD58, TNFSF14* and *CIIT*A for their implication in dysregulation of immunity and escape to immune response; *BCL6* for its role in germinal center reaction; *IBTK, FOXO1, IRF4, ITPKB, CD79b, TCF3, ID3* and *SYK* for their implication in the BCR pathway; *TNFAIP3, MYD88, CARD11* and *PIM1,* for their implication in constitutive activation of NF-κB pathway; *NOTCH1* and *NOTCH2* for their association with NOTCH pathway disruption; *STAT6* and SOCS1 implicated in JAK-STAT pathway; *BRAF* for its role in MAP kinase pathway.

CNV impact on gene expression through integrative genomics

Gene expression profiles were available for 30 out of the 39 patients. These data were acquired on an Agilent Whole Human Genome microarray 4x44K platform (Agilent technologies, Santa Clara, CA, USA). Gene expression levels for each sample were integrated with CNVs. The impact of deletions and duplications of the selected genes on their own expression level was evaluated by comparing the average gene expression levels of deleted, duplicated and normal samples using Student's t-test.

References

1.Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013; 63(1): 11-30. doi: 10.3322/caac.21166.

2.Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, Morel P, Van Den Neste E, Salles G, Gaulard P, Reyes F, Lederlin P and Gisselbrecht C. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med. 2002; 346(4): 235-42.

3.Feugier P, Van Hoof A, Sebban C, Solal-Celigny P, Bouabdallah R, Fermé C, Christian B, Lepage E, Tilly H, Morschhauser F, Gaulard P, Salles G, Bosly A et al. Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol. 2005; 23(18):4117-26.

4.Thieblemont C, Gisselbrecht C. Second-line treatment paradigms for diffuse large B-cell lymphomas. Curr Oncol Rep. 2009; 11(5):386-93.

5.Sehn LH, Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. *Blood* 2015; 125(1):22-32. doi: 10.1182/blood-2014-05-577189.

6.Guglielmi C, Gomez F, Philip T, Hagenbeek A, Martelli M, Sebban C, Milpied N, Bron D, Cahn JY, Somers R, Sonneveld P, Gisselbrecht C, Van Der Lelie H et al. Time to relapse has prognostic value in patients with aggressive lymphoma enrolled onto the Parma trial. J Clin Oncol. 1998; 16(10): 3264-9.

7.Gisselbrecht C, Glass B, Mounier N, Singh Gill D, Linch DC, Trneny M, Bosly A, Ketterer N, Shpilberg O, Haqberg H, Ma D, Brière J, Moskowitz CH and Schmitz N. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. J Clin Oncol. 2010; 28(27): 4184-90. doi: 10.1200/JCO.2010.28.1618

8.Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000; 403(6769): 503-11.

9.Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltnane JM, Hurt EM, Zhao H, Averett L, *et al*. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med. 2002; 346(25): 1937-47.

10.Thieblemont C, Briere J, Mounier N, Voelker HU, Cuccuini W, Hirchaud E, Rosenwald A, Jack A, Sundstrom C, Coqliatti S, Trougouboff P, Boudova L, Ysebaert L, et al. The germinal center/activated B-cell subclassification has a prognostic impact for response to salvage therapy in relapsed/refractory diffuse large B-cell lymphoma: a bio-CORAL study. J Clin Oncol. 2011; 29(31):4079-87. doi: 10.1200/JCO.2011.35.4423.

11.Molina TJ, Canioni D, Copie-Bergman C, Recher C, Brière J, Haioun C, Berger F, Fermé C, Copin MC, Casasnovas O, Thieblemont C, Petrella T, Leroy K, et al. Young patients with non-germinal center B-cell-like diffuse large B-cell lymphoma benefit from intensified chemotherapy with ACVBP plus rituximab compared with CHOP plus rituximab: analysis of data from the Groupe d'Etudes des Lymphomes de l'Adulte/lymphoma study association phase III trial LNH 03-2B. J Clin Oncol. 2014; 32(35): 3996-4003. doi: 10.1200/JCO.2013.54.9493.

12.Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, Kohlhammer H, Lamy L, Zhao H, Yang Y, Xu W, Shaffer AL, Wright G, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. Nature. 2010; 463(7277): 88-92. doi: 10.1038/nature08638.

13.Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, Kohlhammer H, Xu W, Yang Y, Zhao H, Shaffer AL, Romesser P, Wright G, et al. Oncogenically active MYD88 mutations in human lymphoma. Nature. 2011; 470(7332): 115-9. doi: 10.1038/nature09671.

14.Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW, Dave SS, Zhao H, Xu W, Rosenwald A, Ott G, Muller-Hermelink HK, Gascoyne RD, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. Science. 2008; 319 (5870): 1676-9. doi: 10.1126/science.1153629.

15.Compagno M, Lim WK, Grunn A, Nandula SV, Brahmachary M, Shen Q, Bertoni F, Ponzoni M, Scandurra M, Califano A, Bhagat G, Chadburn A, Dalla-Favera R, et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. Nature. 2009; 459(7247): 717-21. doi: 10.1038/nature07968.

16.Kato M, Sanada M, Kato I, Sato Y, Takita J, Takeuchi K, Niwa A, Chen Y, Nakazaki K, Nomoto J, Asakura Y, Muto S, Tamura A, et al. Frequent inactivation of A20 in B-cell lymphomas. Nature. 2009; 459(7247): 712-6. doi: 10.1038/nature07969.

17.Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, Johnson NA, Severson TM, Chiu R, Field M, Jackman S, Krzywinski M, Scott DW, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. 2011; 476(7360): 298-303. doi: 10.1038/nature10351.

18.Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, Kasper LH, Lerach S, Tang H, Ma J, Rossi D, Chadburn A, Murty VV, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature. 2011; 471(7337): 189-95. doi: 10.1038/nature09730.

19.Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, Paul JE, Boyle M, Woolcock BW, Kuchenbauer F, Yap D, Humphries RK, Griffith OL, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet 2010; 42(2): 181-5. doi: 10.1038/ng.518.

20.Roschewski M, Staudt LM, Wilson WH. Diffuse large B-cell lymphoma-treatment approaches in the molecular era. Nat Rev Clin Oncol. 2014; 11(1): 12-23. doi: 10.1038/nrclinonc.2013.197.

21.Zhang J, Grubir V, Love CL, Banerjee A, Richards KL, Mieczkowski PA, Dunphy C, Choi W, Au WY, Srivastava G, Lugar PL, Rizzieri DA, Lagoo AS, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. Proc Natl Acad Sci USA 2013; 110(4): 1398-1403. doi: 10.1073/pnas.1205299110.

22.Lenz G, Wright GW, Emre NC, Kohlhammer H, Dave SS, Davis RE, Carty S, Lam LT, Shaffer AL, Xiao W, Powell J, Rosenwald A, Ott G, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proc Natl Acad Sci USA 2008; 105(36): 13520-5. doi: 10.1073/pnas.0804295105.

23.Schuetz JM, Johnson NA, Morin RD, Scott DW, Tan K, Ben-Nierah S, Boyle M, Slack GW, Marra MA, Connors JM, Brooks-Wilson AR and Gascoyne R. BCL2 mutations in diffuse large B-cell lymphoma. Leukemia. 2012; 26(6): 1383-90. doi: 10.1038/leu.2011.378.

24.Iqbal J, Neppalli VT, Wright G, Dave BJ, Horsman DE, Rosenwald A, Lynch J, Hans CP, Weisenburger DD, Greiner TC, Gascoyne R, Campo E, Ott G, et al. BCL2 expression is a prognostic marker for the activated B-cell-like type of diffuse large B-cell lymphoma. J Clin Oncol. 2006; 24(6): 961-8.

25.Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, Wells VA, Grunn A, Messina M, Elliot O, Chan J, Bhagat G, Chadburn A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet. 2011; 43(9): 830-7. doi: 10.1038/ng.892.

26.Xu-Monette ZY, Wu L, Visco C, Tai YC, Tzankov A, Liu WM, Montes-Moreno S, Dybkaer K, Chiu A, Orazi A, Zu Y, Bhagat G, Richards KL, et al. Mutational profile and prognostic significance of TP53 in diffuse large B-cell lymphoma patients treated with R-CHOP: report from an International DLBCL Rituximab-CHOP Consortium Program Study. Blood. 2012; 120(19): 3986-96. doi: 10.1182/blood-2012-05-433334.

27.Savage KJ, Johnson NA, Ben-Neriah S, Connors JM, Sehn LH, Farinha P, Horsman DE and Gascoyne RD. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. Blood. 2009;114(17):3533-7. doi: 10.1182/blood-2009-05-220095.

28.Cuccuini W, Briere J, Mounier N, Voelker HU, Rosenwald A, Sundstrom C, Cogliatti S, Hirchaud E, Ysebaert L, Bron D, Soulier J, Gaulard F, Houlgatte R, et al. MYC+ diffuse large B-cell lymphoma is not salvaged by classical R-ICE or R-DHAP followed by BEAM plus autologous stem cell transplantation. Blood. 2012; **119(20)**: 4619-24. doi: 10.1182/blood-2012-01-406033.

29.Lohr J, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C, Cruz-Gordillo P, Knoechel B, Asmann YW, Slager SL, Novak AJ, Dogan A, Ansell SM, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci USA. 2012; 109(10): 3879-84. doi: 10.1073/pnas.1121343109.

30.Rossi D, Trifonov V, Fangazio M, Bruscaggin A, Rasi S, Spina V, Monti S, Vaisitti T, Arruga F, Famà R, Ciardullo C, Greco M, Cresta S et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. J Exp Med. 2012; 209(9): 1537-51. doi: 10.1084/jem.20120904.

31.Lee SY, Kumano K, Nakazaki K, Sanada M, Matsumoto A, Yamamoto G, Nannya Y, Suzuki R, Ota S, Ota Y, Izutsu K, Sakata-Yanagimoto M, Hangaishi A, et al. Gain-of-function mutations and copy number increases of Notch2 in diffuse large B-cell lymphoma. Cancer Sci. 2009; 100(5): 920-6.

32.Pasqualucci L. The genetic basis of diffuse large B-cell lymphoma. Curr Opin Hematol. 2013; 20(4): 336-44. doi: 10.1097/MOH.0b013e3283623d7f.

33.Iqbal J, Greiner TC, Patel K, Dave BJ, Smith L, Ji J, Wright G, Sanger WG, Pickering DL, Jain S, Horsman DE, ShenY, Fu K, et al. Distinctive patterns of BCL6 molecular alterations and thier functionnal consequences in different subgroups of diffuse large B-cell lymphoma. Leukemia. 2007; 21(11): 2332-43.

34.Challa-Malladi M, Lieu YK, Califano O, Holmes AB, Bhagat G, Murty VV, Dominguez-Sola D, Pasqualucci L, and Dalla-Favera R. Combined genetic inactivation of β2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Cancer Cell. 2011; 20(6): 728-40. doi: 10.1016/j.ccr.2011.11.006.

35.Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P, Jonhson NA, Zhao Y, Telenius A, Neriah SB, PcPherson A, Meissner B, Okoye UC, et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature. 2011; 471(7338): 377-81. doi: 10.1038/nature09754.

36.Tusher VG, Tibshirani R and Chu G. Significance analysis of microarray applied to the ionizing radiation response. Proc Natl Acad Sci USA. 2001; 98(9): 5116-5121.

37.Novak AJ, Asmann YW, Maurer MJ, Wang C, Slager SL, Hodge LS, Manske M, Price-Troska T, Yang ZZ, Zimmermann MT, Nowakowski GS, Ansell SM, Witzig TE et al. Whole-exome analysis reveals novel somatic genomic alterations associated with outcome in immunochemotherapy-treated diffuse large B-cell lymphoma. Blood Cancer J. 2015; 28; 5:e346. doi: 10.1038/bcj.2015.69.

38. Xian RR, Crane GM, Haley LM, Gocke CD, Lin MT, Pallavajjalla A, Borowitz MJ, Swinnen LJ and Duffield AS. Late-Relapse Diffuse Large B-Cell Lymphoma Frequently Represents Recurrence of the Original Disease, and Demonstrates Evidence of Superimposed Clonal Heterogeneity and Clonal Evolution. Blood. 2014; 124(21):2955.

39. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhim R and Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol. 2011; 12(4): R41. doi:10.1186/gb-2011-12-4-r41.

40.Fujita A, Sato JR, Ferreira CE and Sogayar MC. GEDI: a user-friendly toolbox for analysis of large-scale gene expression data. BMC Bioinformatics 2007(19); 8:457.

41.Monti S, Chapuy B, Takeyama K, Rodig SJ, Hao Y, Yeda KT et al. Integrative analysis reveals an outcome-associated and targetable pattern of p53 and cell cycle deregulation in diffuse large B-cell lymphoma. Cancer Cell. 2012; 22(3): 359-372. doi: 10.1016/j.ccr.2012.07.014.

42.Jardin F, Jais JP, Molina TJ, Parmentier F, Picquenot JM, Ruminy P, Tilly H, Bastard C, Salles GA, Feugier P, Thieblemont C, Gisselbrecht C, De Reynies A, et al. Diffuse large B-cell lymphomas with CDKN2A deletion have a distinct gene expression signature and a poor prognosis under R-CHOP treatment: a GELA study. Blood. 2010; 116(7): 1092-104. doi: 10.1182/blood-2009-10-247122.

43.Janda E, Palmieri C, Pisano A, Pontoriero M, Iaccino E, Falcone C, Fiume G, Gaspari M, Nevolo M, Di Salle E, Rossi A, De Laurentiis A, Greco A, et al. Btk regulation in human and mouse B cells via protein kinase C phosphorylation of IBtkγ. Blood. 2011; 117(24): 6520-31. doi: 10.1182/blood-2010-09-308080.

44.Wilson WH, Young RM, Schmitz R, Yang Y, Pittaluga S, Wright G, Lih CJ, Williams PM, Shaffer AL, Gerecitano J, De Vos S, Goy A, Kenkre VP, et al. Targeting B cell receptor signaling with ibrutinib in diffuse large B cell lymphoma. Nat Med. 2015; 21(8): 922-6. doi: 10.1038/nm.3884.

45.Du C, Fang M, Li Y, Li L and Wang X. Smac, a mitochondrial protein that promotes cytochrome c-dependent caspase activation by eliminating IAP inhibition. Cell. 2000; 102(1): 33-42.

46.Piazza F, Manni S, Ruzzene M, Pinna LA, Gurrieri C and Semenzato G. Protein kinase CK2 in hematologic malignancies: reliance on a pivotal cell survival regulator by oncogenic signaling pathways. Leukemia. 2012; 26(6): 1174-1179. doi: 10.1038/leu.2011.385.

47.Hegazy SA, Alshareef A, Gelebart P, Anand M, Armanious H, Ingham RJ and Lai R. Disheveled proteins promote cell growth and tumorigenicity in ALK-positive anaplastic large cell lymphoma. Cell Signal. 2013; 25(1): 295-307. doi: 10.1016/j.cellsig.2012.09.027.

48.Martic G, Karetsou Z, Kefala K, Politou AS, Clapier CR, Straub T and Papamarcaki T. Parathymosin affects the binding of linker histone H1 to nucleosomes and remodels chromatin structure. J Biol Chem. 2005; 280(16): 16143-50.

49.Okosun J, Bödör C, Wang J, Araf S, Yang CY, Pan C, Boller S, Cittaro D, Bozek M, Iqbal S, Matthews J, Wrench D, Marzec J et al. Integrated genomics identifies recurrent mutations and evolution patterns driving the initiation and progression of follicular lymphoma. Nat Genet. 2014; 46(2): 176-81. doi: 10.1038/ng.2856.

50.Cámara Y, Asin-Cayuela J, Park CB, Metodiev MD, Shi Y, Ruzzenente B, Kukat C, Habermann B, Wibom R, Hultenby K, Franz T, Erdjument-Bromage H, Tempst P, *et al*. MTERF4 regulates translation by targeting the methyltransferase NSUN4 to the mammalian mitochondrial ribosome. Cell Metab 2011; 13(5): 527-39. doi: 10.1016/j.cmet.2011.04.002.

51.Saxena A and Carninci P. Long non-coding RNA modifies chromatin: epigenetic silencing by long non-coding RNAs. Bioessays. 2011; 33(11): 830-9. doi: 10.1002/bies.201100084.

52.Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R and Fraser P. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. Science. 2008; 322(5908): 1717-20. doi: 10.1126/science.1163802.

53.Tremblay CS, Huang FF, Habi O, Huard CC, Godin C, Lévesque G and Carreau M. HES1 is a novel interactor of the Fanconi anemia core complex. Blood. 2008; 112(5): 2062-70. doi: 10.1182/blood-2008-04-152710.

54.Nakahara F, Sakata-Yanagimoto M, Komeno Y, Kato N, Uchida T, Haragushi K, Kumano K, Harada Y, Harada H, Kitaura J, Ogawa S, Kurokawa M, Kitamura T, et al. Hes1 immortalizes committed progenitors and plays a role in blast crisis transition in chronic myelogenous leukemia. Blood. 2010; 115(14): 2872-81. doi: 10.1182/blood-2009-05-222836.

55.McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, De Bakker PI, Maller JB, Kirby A, Elliott AL, Parkin M, Hubbel E, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet. 2008; 40(10**)**: 1166-74. doi: 10.1038/ng.238.

56.Scharpf RB, Irizarry RA, Ritchie ME, Carvalho B and Ruczinski I. Using the R Package crlmm for Genotyping and Copy Number Estimation. J Stat Softw. 2011; 40(12): 1-32.

57.Saldanha AJ. Java Treeview-extensible visualization of microarray data. Bioinformatics. 2004; 20(17): 3246-8.

58.Jardin F. Next generation sequencing and the management of diffuse large B-cell lymphoma: from whole exome analysis to targeted therapy. Discov Med. 2014; **18**: 51-65.

Figures

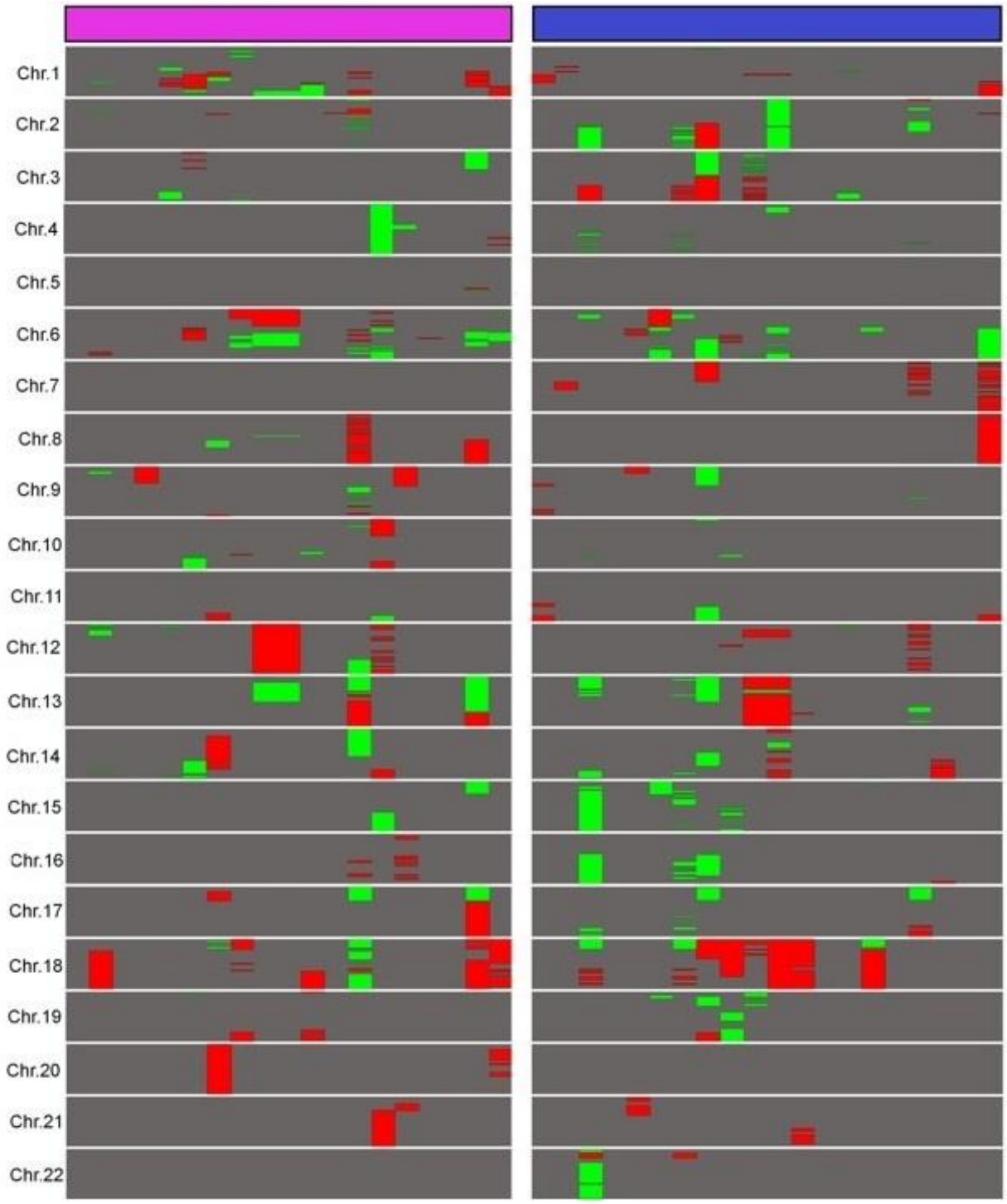


Figure 1: Landscape of large copy number variations in relapsed DLBCL. A comparative representation of abnormalities longer than 2 Mb in ER and LR DLBCLs was built. Duplications are represented in red and deletions in green. ER DLBCLs (n=19) are represented in purple and LR DLBCLs (n=20) in blue. Chromosomes X and Y are not showed because of gender differences in copy numbers. This figure was obtained using Java Treeview 3.0.

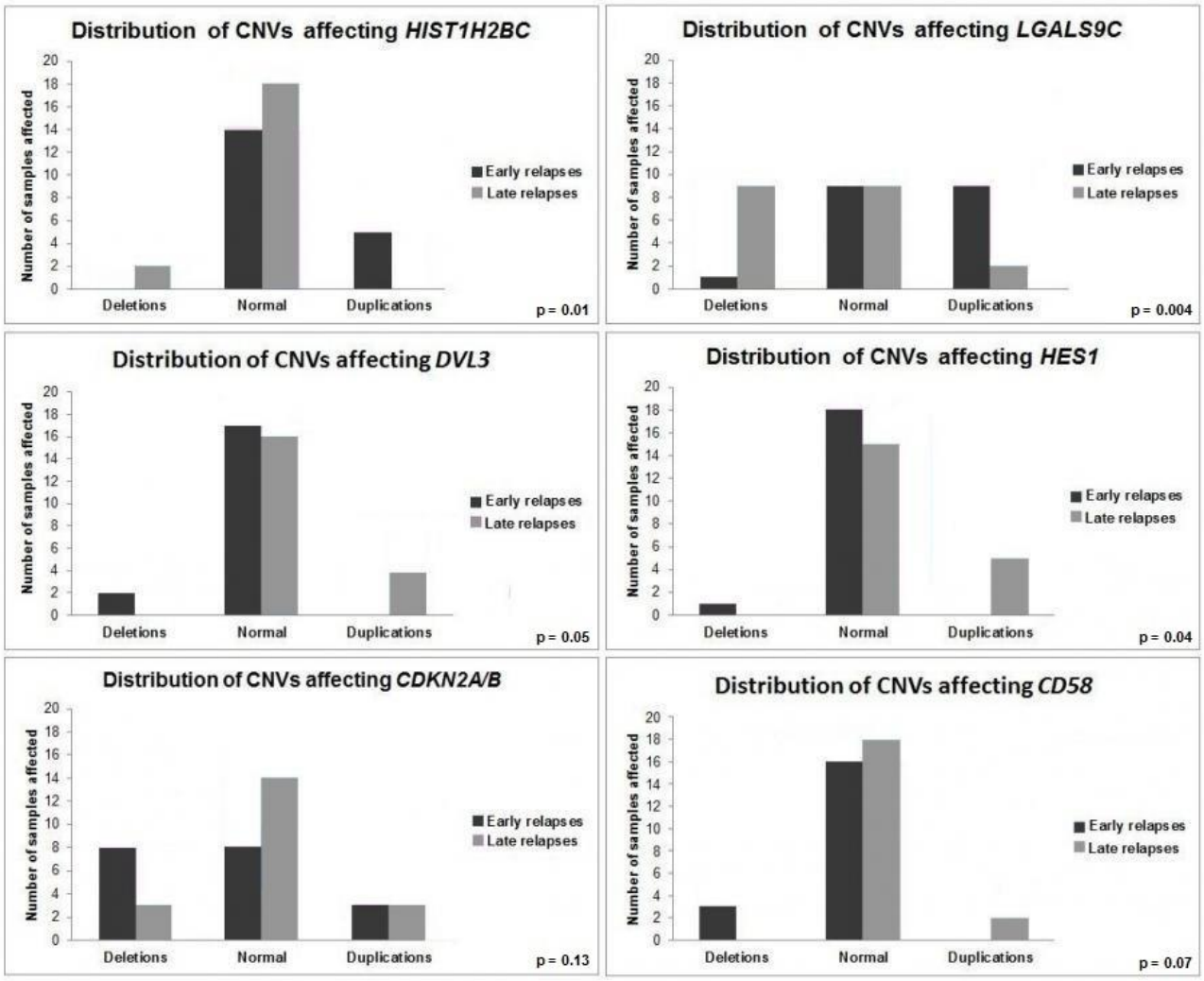


Figure 2: Genes differentially altered between ER DLBCLs and LR DLBCLs. Comparison of CNV profiles between ER and LR DLBCLs has been performed at a genomic scale. *HIST1H2BC*, *LGALS9C, DVL3, HES1*, *CDKN2A/B* and *CD58* are genes with structural aberrations differentially distributed between ER and LR DLBCLs**.** ER DLBCLs (n=19) are represented in dark grey and LR DLBCLs (n=20) in light grey.

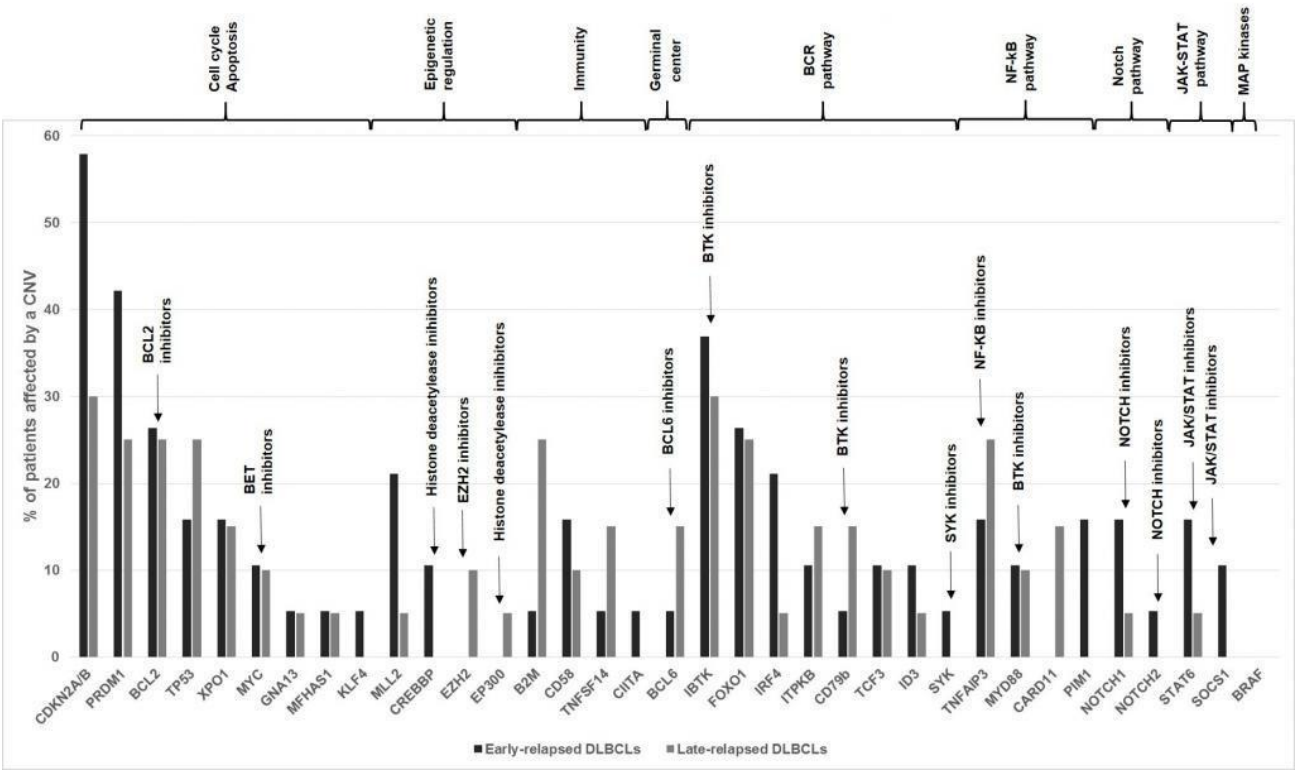


Figure 3: Genomic alterations of the 36 genes set involved in lymphomagenesis. Comparative distribution of CNVs among ER and LR DLBCLs is shown. Genes are grouped according to key dysregulated pathways in DLBCLs. ER DLBCLs (n=19) are colored in dark grey and LR DLBCLs (n=20) in light grey. CNV: Copy Number Variation.

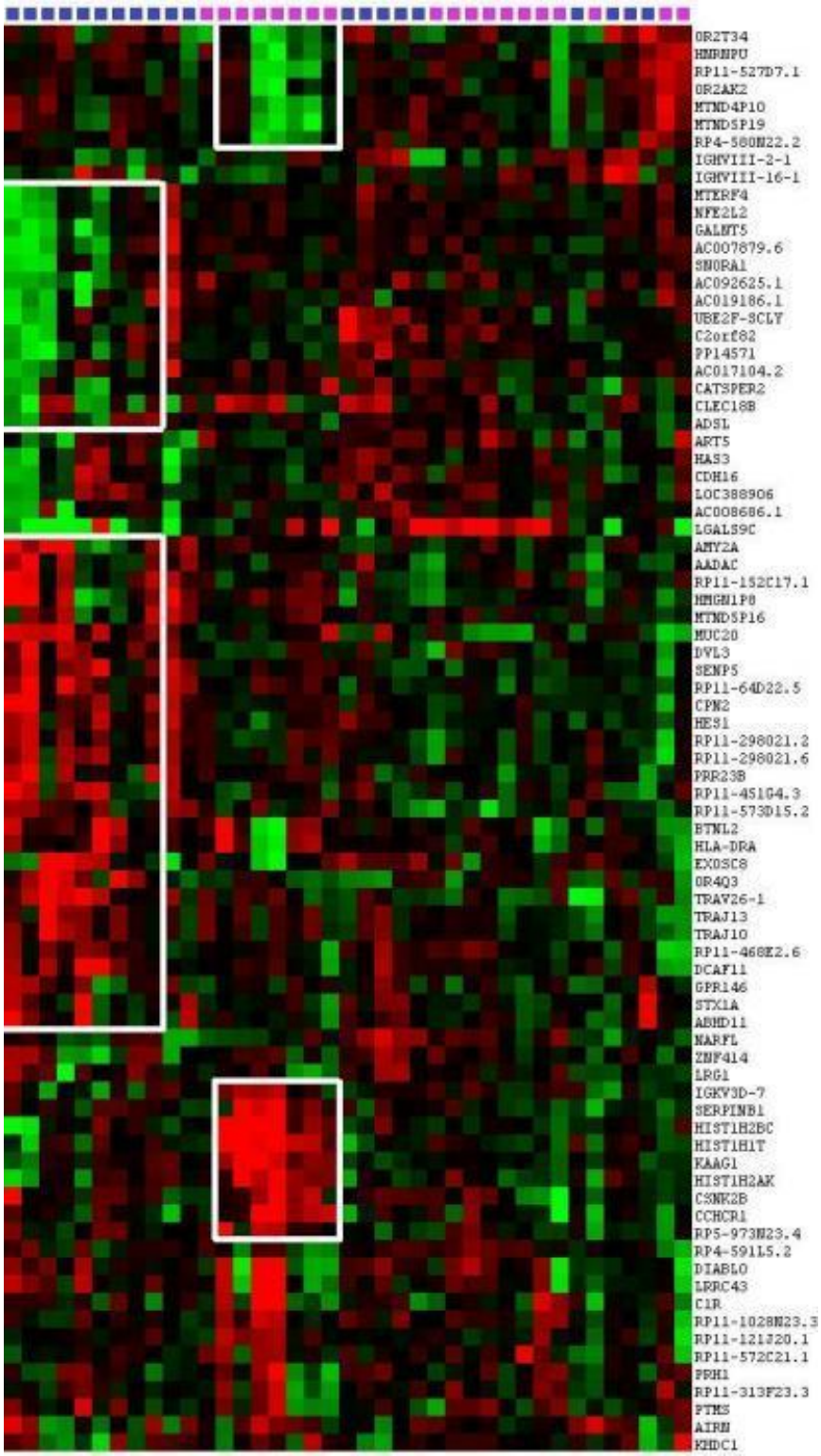


Figure 4: Evidence for recurrent events in relapsed DLBCLs. Hierarchical clustering of 56 protein-coding genes CNVs and 25 LncRNAs CNVs identifies events associated with ER and LR DLBCLs. The heatmap shows recurrent CNVs in DLBCL patients and gives evidence of genetic aberration occurrences in ER and LR DLBCLs. In ER DLBCLs, *HIST1H2BC*, *HIST1H1T*, *HIS1H2AK*, *CSNK2B/CK2B*, *KAAG1*, *CCHCR1*, and *IgKV3D-7* are co-amplified in the same patients. Accordingly, in LR DLBCLs, *AMY2A*, *AADAC*, *MUC20*, *DVL3*, *SENP5*, *CPN2*, and *HES1* are co-amplified in the same patients. This figure was obtained using Cluster 3.0 and Java Treeview 3.0.

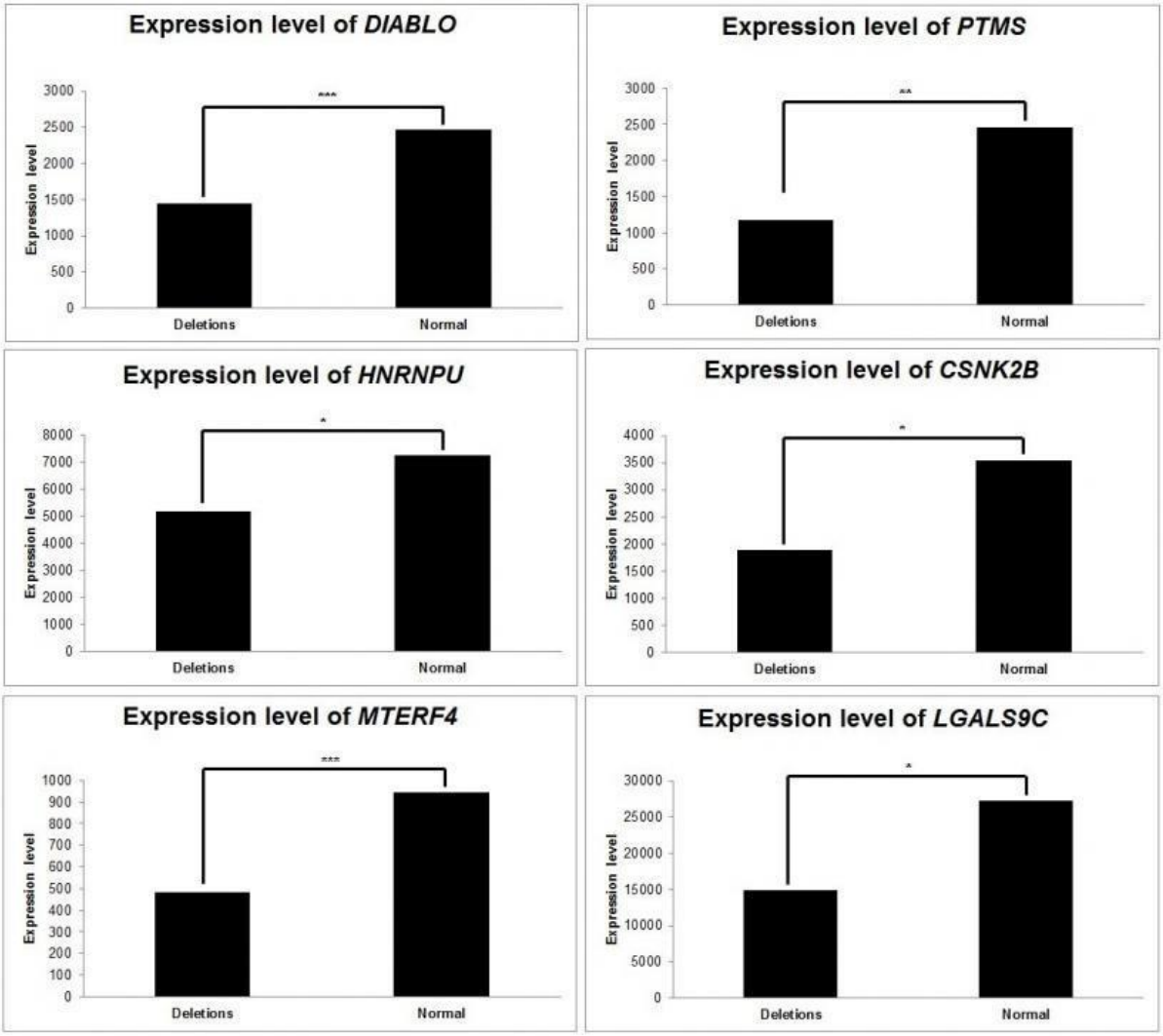


Figure 5: **Consequences of relevant copy number variations on gene expression**. Average gene expression was computed, according to relevant CNVs. CNVs affecting *DIABLO, PTMS, HNRNPU, CSNK2B, MTERF4 and LGLAS9C* have subsequent consequences on the expression levels of those genes. *: $p < 5 \times 10^{-2}$; **: $p < 10^{-2}$; ***: $p < 10^{-3}$.

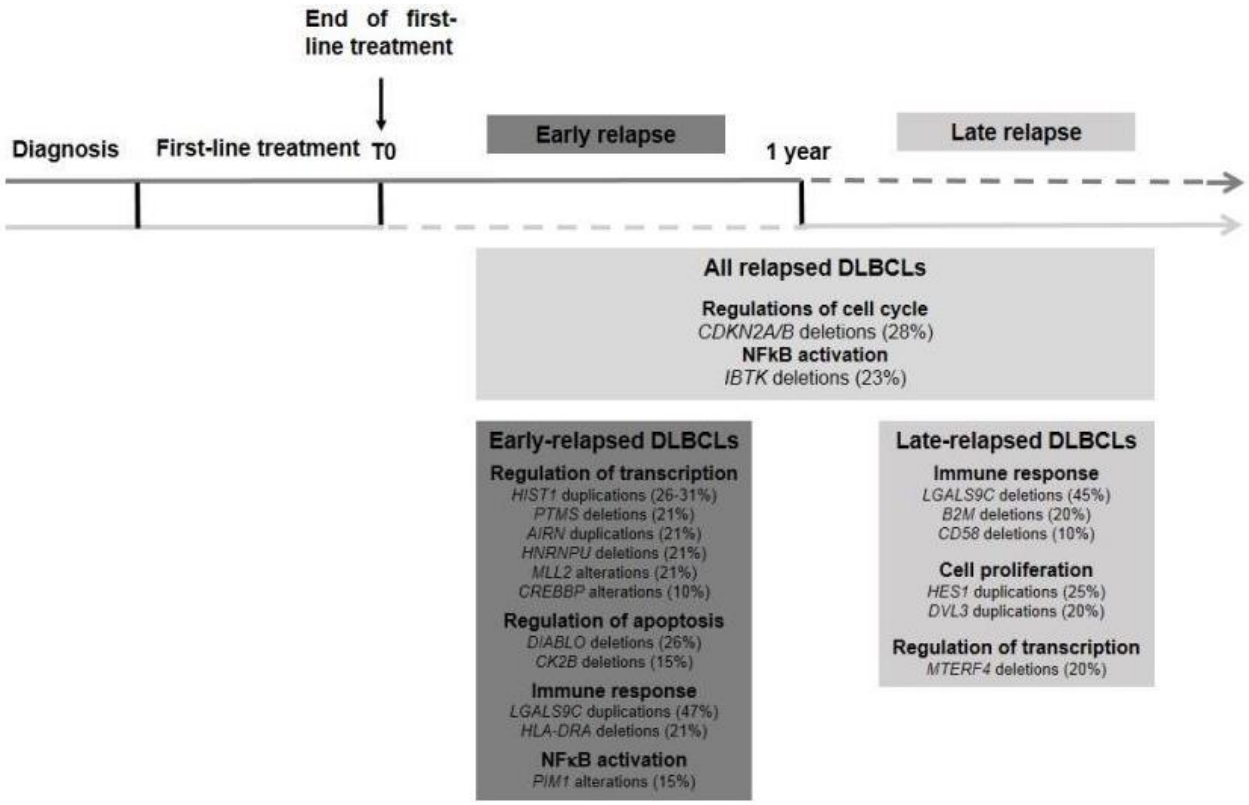


Figure 6: Key oncogenic pathways in relapsed DLBCLs. Structural abnormalities related with relapsed DLBCLs are shown, with distinction between ER and LR DLBCLs. ER DLBCLs are preferentially associated with dysregulation of gene transcription, cell cycle and apoptosis. LR DLBCLs are rather associated with dysregulation of immune response, cell proliferation and regulation of transcription.

Tables

Table 1: Detailed list of the 56 protein coding genes presenting with differential CNVs distribution between ER and LR. The SAM method allowed identification of 148 genes with differentially distributed genomic alterations between ER and LR DLBCLs. The 56 protein coding genes are presented in this table.

Chr. Band	Gene	Gene full name	Description/function
1p21	*AMY2A*	Amylase, Alpha 2A	Member of the alpha-amylase family
1q44	*HNRPU*	Heterogeneous Nuclear Ribonucleoprotein U	pre-mRNA processing
1q44	*OR2T34*	Olfactory receptor, family 2, subfam. T, mb 34	G-protein-coupled olfactory receptor
1q44	*OR2AK2*	Olfactory receptor, family 2, subfam. AK, mb 2	G-protein-coupled olfactory receptor
2p11	*IGKV3D-7*	Immunoglobulin kappa variable 3D-7	BCR assembly
2q24	*GALNT5*	N-Acetylgalactosaminyltransferase 5	O-linked oligosaccharide biosynthesis
2q31	*NFE2L2*	Nuclear factor (erythroid-derived 2)-like 2	Transcriptionnal regulator
2q33	*SNORA1*	Small nucleolar RNA, H/ACA box 1	Novel snoRNA/Small nucleolar RNA
2q37	*MTERF4*	Mitochondrial Transcription Termination Factor 4	Regulator of mitochondrial gene expression
2q37	*UBE2F-SCLY*	UBE2F-SCLY Readthrough	MHC class I mediated antigen processing
2q37	*C2orf82*	Chromosome 2 Open Reading Frame 82	---
3q22	*PRR23B*	Proline Rich 23B	---
3q25	*AADAC*	Arylacetamide deacetylase	Microsomal arylacetamide deacetylase
3q27	*DVL3*	Dishevelled, dsh homolog 3	Regulation of cell proliferation
3q29	*CPN2*	Carboxypeptidase N, Polypeptide 2	Carboxypeptidase
3q29	*SENP5*	SUMO1/sentrin specific peptidase 5	Post-translational modification of proteins
3q29	*MUC20*	Mucin 20, Cell Surface Associated	Glycoprotein of mucous barrier
3q29	*HES1*	Hairy enhancer of split 1	Transcription factor
6p21	*CSNK2B/CK2B*	Casein kinase 2, beta subunit	Regulation of programmed cell death
6p21	*BTNL2*	Butyrophilin-Like 2	Negative regulation of T-cell proliferation
6p21	*CCHCR1*	Coiled-Coil Alpha-Helical Rod Protein 1	Regulation of keratinocyte proliferation
6p21	*HIST1H1T*	Histone Cluster 1, H1t	Epigenetic regulation of transcription
6p21	*HLA-DRA*	MHC complex, Class II, DR Alpha	Antigen presentation
6p22	*HIST1H2BC*	Histone Cluster 1, H2bc	Epigenetic regulation of transcription
6p22	*HIS1H2AK*	Histone Cluster 1, H2ak	Epigenetic regulation of transcription
6p22	*KAAG1*	Kidney Associated Antigen 1	---
6p25	*SERPINB1*	Serpin Peptidase Inhibitor, Class B , Mb1	Proteinase inhibition
6q13	*KHDC1*	KH homology domain containing 1	Predicted membrane protein
7p22	*GPR146*	G Protein-Coupled Receptor 146	G protein coupled receptor
7q11	*ABHD11*	Abhydrolase domain containing 11	---
7q11	*STX1A*	Syntaxin 1A	Docking of synaptic vesicles
11p15	*ART5*	ADP-Ribosyltransferase 5	ARG-specific ADP-ribosyltransferase
12p13	*PRH1*	Proline-Rich Protein HaeIII Subfamily 1	Inhibition of crystal calcium phosphates

12p13	*C1R*	Complement Component 1, R Subcomponent	Complement activation
12p13	*PTMS*	Parathymosin	Inhibition of linkage of H1 to chromatin
12q24	*DIABLO*	Diablo, IAP-Binding Mitochondrial Protein	Apoptosis promotion
12q24	*LRRC43*	Leucine Rich Repeat Containing 43	---
13q13	*EXOSC8*	Exosome Component 8	Transcriptionnal regulation
14q11	*OR4Q3*	Olfactory receptor, family 4, subfam. Q, mb 3	G-protein-coupled olfactory receptor
14q11	*TRAV-26-1*	T Cell Receptor Alpha Variable 26-1	TCR assembly
14q11	*TRAJ10*	T Cell Receptor Alpha Joining 10	TCR assembly
14q11	*TRAJ13*	T Cell Receptor Alpha Joining 13	TCR assembly
14q12	*RP11-468E2.6*	---	---
14q32	*IGHVIII-2-1*	Immunoglobulin Heavy Variable (III)-2-1	BCR assembly
14q32	*IGHVIII-16-1*	Immunoglobulin Heavy Variable (III)-16-1	BCR assembly
14q32	*DCAF11*	DDB1 And CUL4 Associated Factor 11	Assembly of the DDB1–CUL4A ubiquitin ligase
15q15	*CATSPER2*	Cation Channel, Sperm Associated 2	Sperm motility and male fertility
16p13	*NARFL*	Nuclear prelamin A recognition factor-like	Modulation of hypoxia-inducible factor-1*α* activity
16q22	*HAS3*	Hyaluronan Synthase 3	Synthesis of glycosaminoglycan hyaluronan
16q22	*CDH16*	Cadherin 16	Ca2+-dependent glycoprotein
16q22	*CLEC18B*	C-Type Lectin Domain Family 18, Member B	---
17p11	*LGALS9C*	Lectin, Galactoside-Binding, Soluble, 9C	Regulation of immune response
19p13	*LRG1*	Leucine-rich alpha-2-glycoprotein 1	Signal transduction, and cell adhesion
19p13	*ZNF414*	Zinc Finger Protein 414	Transcriptional regulation
19p13	*AC008686.1*	---	Uncharacterized protein
22q13	*ADSL*	Adenylosuccinate lyase	*De novo* purine biosynthesis

Table 2: characteristics of the cohort of 39 relapsed DLBCLs.
IPI: International prognostic index; CHOP: cyclophosphamide, adriamycin, vincristine and prednisone; ACVB: adriamycin, cyclophosphamide, vindesine, bleomycin; ER: Early Relapse; LR: Late Relapse; GC: Germinal Center B; ABC: Activated B-cell; NA: Not Available; NP: Non Predictable; 1: Yes; 2: No.

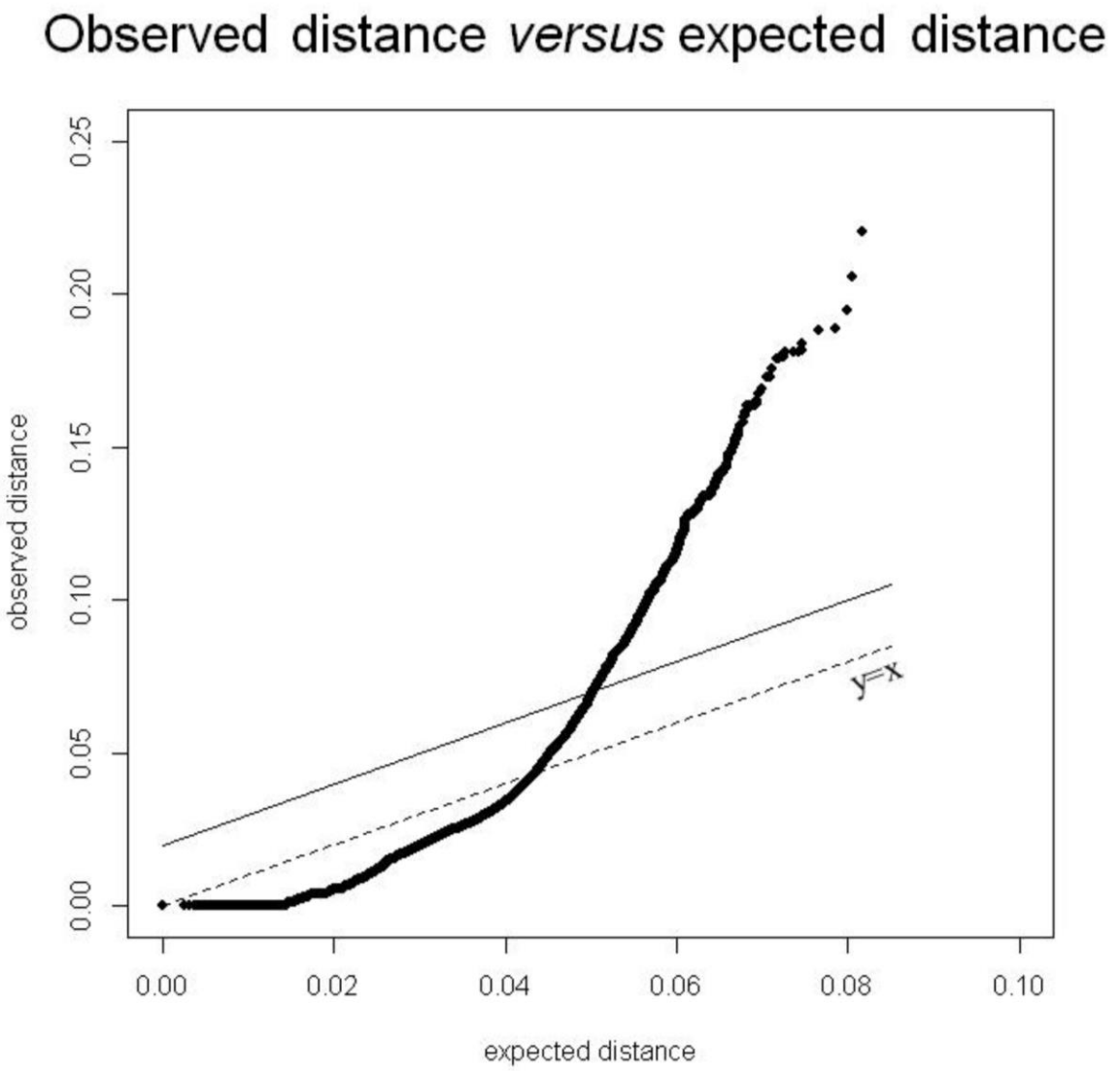
N°	Sex	Age at diagnosis	IPI at diagnosis	First-line treatment	Delay of relapse (months)	ER/ LR	Subtype	Rearrangements			Sample
								MYC	*BCL2*	*BCL6*	
1	Male	62	3	CHOP	11	ER	GC	No	1	2	Diagnosis
2	Female	20	1	CHOP	6	ER	ABC	No	2	2	Diagnosis
3	Female	60	4	CHOP	6	ER	ABC	No	1	2	Diagnosis
4	Female	60	4	ACVB	8	ER	NA	No	2	2	Diagnosis
5	Male	59	0	CHOP	3	ER	GC	No	1	2	Relapse
6	Male	28	0	ACVB	9	ER	NA	No	2	2	Relapse
7	Male	43	1	CHOP	6	ER	ABC	NA	NA	NA	Diagnosis
8	Male	41	1	CHOP	11	ER	NP	NA	NA	NA	Relapse
9	Female	38	3	CHOP	5	ER	ABC	No	2	2	Diagnosis
10	Male	46	2	CHOP	0	ER	GC	No	1	NA	Diagnosis
11	Male	34	2	CHOP	2	ER	GC	NA	NA	NA	Relapse
12	Male	54	0	ACVB	9	ER	GC	No	1	2	Relapse
13	Male	57	1	CHOP	0	ER	GC	Yes	1	2	Diagnosis
14	Male	57	1	CHOP	0	ER	GC	Yes	1	2	Relapse
15	Male	57	2	ACVB	2	ER	NP	No	2	NA	Diagnosis
16	Male	32	2	CHOP	0	ER	GC	No	2	2	Relapse
17	Female	35	2	ACVB	7	ER	ABC	NA	NA	NA	Diagnosis
18	Female	35	2	ACVB	7	ER	ABC	No	2	2	Relapse
19	Male	33	0	ACVB	11	ER	GC	No	2	2	Diagnosis
20	Male	29	2	ACVB	82	LR	GC	NA	NA	NA	Relapse
21	Male	41	0	CHOP	148	LR	NA	NA	NA	NA	Relapse
22	Male	42	2	ACVB	29	LR	NA	Yes	2	2	Diagnosis
23	Male	41	1	ACVB	155	LR	ABC	No	2	2	Relapse
24	Male	59	0	ACVB	20	LR	GC	No	1	2	Relapse
25	Male	53	0	ACVB	50	LR	GC	No	2	2	Diagnosis
26	Male	53	0	ACVB	50	LR	GC	No	2	2	Relapse
27	Female	58	2	ACVB	60	LR	ABC	NA	1	1	Relapse
28	Male	46	0	ACVB	40	LR	ABC	NA	NA	NA	Diagnosis
29	Male	33	0	ACVB	109	LR	ABC	No	2	2	Relapse
30	Male	43	0	ACVB	103	LR	ABC	No	2	2	Diagnosis
31	Female	38	1	CHOP	62	LR	NA	No	2	NA	Diagnosis
32	Male	43	0	CHOP	158	LR	GC	Yes	1	NA	Diagnosis
33	Male	62	2	CHOP	21	LR	NA	No	1	2	Diagnosis
34	Male	56	0	ACVB	29	LR	GC	No	2	2	Diagnosis
35	Male	56	0	ACVB	29	LR	NA	No	2	2	Relapse
36	Male	45	0	CHOP	153	LR	GC	No	1	2	Relapse
37	Male	26	0	CHOP	116	LR	GC	No	2	2	Relapse
38	Male	58	2	CHOP	19	LR	NA	NA	NA	NA	Diagnosis
39	Male	NA	NA	NA	NA	LR	NA	No	2	1	Relapse

Supplemental table 1: List of the 148 genes with differential CNVs distribution among early-relapsed and late-relapsed DLBCLs.

Chr. Band	Gene	Gene full name	Description/function
1p21	*AMY2A*	Amylase, Alpha 2A	Member of the alpha-amylase family
1p35	*RP4-591L5.2*	---	LncRNA
1p35	*RP1-65J11.5*	---	Antisense RNA
1q23	*RP11-367J7.4*	---	Pseudogene
1q32	*RP11-168O16.1*	---	Antisense RNA
1q42	*RP4-580N22.2*	---	LncRNA
1q42	*MTND5P19*	MT-ND5 Pseudogene 19	LncRNA
1q42	*MTND4P10*	MT-ND4 Pseudogene 10	LncRNA
1q42	*MTND6P14*	MT-ND6 Pseudogene 14	Pseudogene
1q42	*RP5-827C21.2*	---	Antisense RNA
1q43	*RP11-193H5.2*	---	Pseudogene
1q43	*RP11-567G24.1*	---	Pseudogene
1q43	*RP11-527D7.1*	---	LncRNA
1q43	*RNU6-1139P*	---	snRNA
1q44	*HNRNPU-AS1*	HNRNPU Antisense RNA 1	Antisense RNA
1q44	*FABP7P1*	Fatty acid binding protein 7, brain pseudogene 1	Pseudogene
1q44	*RNU6-947P*	RNA, U6 Small Nuclear 947, Pseudogene	Pseudogene
1q44	*HNRPU*	Heterogeneous Nuclear Ribonucleoprotein U	pre-mRNA processing
1q44	*OR2T34*	Olfactory receptor, family 2, subfam. T, mb 34	G-protein-coupled olfactory receptor
1q44	*OR2AK2*	Olfactory receptor, family 2, subfam. AK, mb 2	G-protein-coupled olfactory receptor
2p11	*IGKV3D-7*	Immunoglobulin kappa variable 3D-7	BCR assembly
2q23	*AC079790.2*	---	Antisense RNA
2q24	*GALNT5*	N-Acetylgalactosaminyltransferase 5	O-linked oligosaccharide biosynthesis
2q24	*AC092625.1*	---	LncRNA
2q24	*AC019186.1*	---	LncRNA
2q24	*AC009227.2*	---	Antisense RNA
2q31	*EXTL2P1*	Exostoses (multiple)-like 2 pseudogene 1	Pseudogene
2q31	*NFE2L2*	Nuclear factor (erythroid-derived 2)-like 2	Transcriptionnal regulator
2q32	*RN7SL267P*	RNA, 7SL, cytoplasmic 267	Pseudogene
2q33	*RNU6-31P*	RNA, U6 Small Nuclear 31	snRNA
2q33	*AC007163.9*	---	Pseudogene
2q33	*MTND4P29*	MT-ND4 Pseudogene 29	Pseudogene
2q33	*MTND5P31*	MT-ND5 Pseudogene 31	Pseudogene
2q33	*SNORA1*	Small nucleolar RNA, H/ACA box 1	Novel snoRNA
2q33	*AC007879.6*	---	LncRNA
2q34	*MEAF6P1*	MYST/Esa1-associated factor 6 pseudogene 1	Pseudogene
2q35	*AC053503.11*	---	Antisense RNA
2q35	*AC010136.2*	---	Antisense RNA
2q35	*CXCR2P1*	Chemokine Receptor 2 Pseudogene 1	Pseudogene
2q37	*AC017104.2*	---	LncRNA
2q37	*PP14571*	---	LncRNA
2q37	*MTERF4*	Mitochondrial Transcription Termination Factor 4	Regulator of mitochondrial gene expression
2q37	*UBE2F-SCLY*	UBE2F-SCLY Readthrough	MHC class I mediated antigen processing
2q37	*C2orf82*	Chromosome 2 Open Reading Frame 82	---
2q37	*AC104667.3*	---	Antisense RNA
3p22	*AC011816.1*	---	Pseudogene
3q13	*RP11-10G15.4*	---	Pseudogene
3q13	*MTND5P16*	MT-ND5 pseudogene 16	LncRNA
3q21	*GS1-388B5.8*	---	Pseudogene
3q21	*RNA5SP58*	RNA, 5S Ribosomal Pseudogene 57	Pseudogene
3q22	*HSPA8P19*	Heat Shock 70kDa Protein 8 Pseudogene 19	Pseudogene
3q22	*PRR23B*	Proline Rich 23B	---
3q23	*ACTG1P1*	Actin Gamma 1 Pseudogene 1	Pseudogene
3q23	*RNU1-100P*	RNA, U1 Small Nuclear 100	Pseudogene
3q23	*RP11-68L1.2*	---	Antisense RNA
3q23	*CLSTN2-AS1*	CLSTN2 antisense RNA 1	Antisense RNA
3q25	*RP11-64D22.5*	---	LncRNA
3q25	*RP11-451G4.3*	---	LncRNA

3q25	*AADAC*	Arylacetamide deacetylase	Microsomal arylacetamide deacetylase
3q25	*PA2G4P4*	Proliferation-Associated 2G4 Pseudogene 4	Pseudogene
3q25	*RPL21P42*	ribosomal protein L21 pseudogene 42	Pseudogene
3q25	*PLCH1-AS1*	PLCH1 antisense RNA 1	Antisense RNA
3q25	*RP11-103G8.2*	ERICH6 antisense RNA 1	Antisense RNA
3q25	*IQCJ-SCHIP1-AS1*	IQCJ-SCHIP1 Readthrough Antisense RNA 1	Antisense RNA
3q26	*HMGN1P8*	high mobility group nucleosome binding domain 1 pseudogene 8	LncRNA
3q26	*RP11-152C17.1*	---	LncRNA
3q26	*RP11-298O21.2*	---	LncRNA
3q26	*RP11-298O21.6*	---	LncRNA
3q26	*RP11-816B4.1*	---	Pseudogene
3q26	*AC007620.3*	---	Antisense RNA
3q27	*RNU6-1105P*	RNA, U6 Small Nuclear 1105	Pseudogene
3q27	*RP11-573D15.2*	---	LncRNA
3q27	*DVL3*	Dishevelled, dsh homolog 3	Regulation of cell proliferation
3q27	*EIF2B5-AS1*	EIF2B5 antisense RNA 1	Antisense RNA
3q28	*GCNT1P3*	glucosaminyl transferase 1, core2 pseudogene 3	Pseudogene
3q29	*CPN2*	Carboxypeptidase N, Polypeptide 2	Carboxypeptidase
3q29	*SENP5*	SUMO1/sentrin specific peptidase 5	Post-translational modification of proteins
3q29	*MUC20*	Mucin 20, Cell Surface Associated	Glycoprotein of mucous barrier
3q29	*HES1*	Hairy enhancer of split 1	Transcription factor
3q29	*RP11-135A1.2*	---	Antisense RNA
4p13	*RN7SL691P*	RNA, 7SL, Cytoplasmic 691	Pseudogene
6p21	*RP5-973N23.4*	---	LncRNA
6p21	*CSNK2B/CK2B*	Casein kinase 2, beta subunit	Regulation of programmed cell death
6p21	*BTNL2*	Butyrophilin-Like 2	Negative regulation of T-cell proliferation
6p21	*CCHCR1*	Coiled-Coil Alpha-Helical Rod Protein 1	Regulation of keratinocyte proliferation
6p21	*HIST1H1T*	Histone Cluster 1, H1t	Epigenetic regulation of transcription
6p21	*HLA-DRA*	MHC complex, Class II, DR Alpha	Antigen presentation
6p22	*RP3-369A17.2*	---	Pseudogene
6p22	*VN1R11P*	Vomeronasal 1 Receptor 11 Pseudogene	Pseudogene
6p22	*MCCD1P1*	Mitochondrial coiled-coil domain 1 pseudogene 1	Pseudogene
6p22	*HIST1H2BC*	Histone Cluster 1, H2bc	Epigenetic regulation of transcription
6p22	*HIS1H2AK*	Histone Cluster 1, H2ak	Epigenetic regulation of transcription
6p22	*KAAG1*	Kidney Associated Antigen 1	---
6p24	*RP11-69L16.4*	---	Antisense RNA
6p25	*SERPINB1*	Serpin Peptidase Inhibitor, Class B , Mb1	Proteinase inhibition
6q12	*RP11-74E24.2*	Zinc finger CCCH-Type Domain-Containing-Like	Pseudogene
6q13	*KHDC1*	KH homology domain containing 1	Predicted membrane protein
6q21	*RP11-249L21.4*	---	Pseudogene
6q25	*AIRN*	Antisense of IGF2R Non-Protein Coding RNA	Regulation of transcription
6q25	*RP3-403L10.3*	---	Antisense RNA
7p22	*GPR146*	G Protein-Coupled Receptor 146	G protein coupled receptor
7p22	*AC073343.13*	---	Antisense RNA
7q11	*ABHD11*	Abhydrolase domain containing 11	---
7q11	*STX1A*	Syntaxin 1A	Docking of synaptic vesicles
7q21	*AC079781.7*	---	Pseudogene
11p15	*ART5*	ADP-Ribosyltransferase 5	ARG-specific ADP-ribosyltransferase
12p13	*PRH1*	Proline-Rich Protein HaeIII Subfamily 1	Inhibition of crystal calcium phosphates
12p13	*C1R*	Complement Component 1, R Subcomponent	Complement activation
12p13	*PTMS*	Parathymosin	Inhibition of linkage of H1 to chromatin
12q12	*RP11-313F23.3*	---	LncRNA
12q12	*CLUHP8*	clustered mitochondria homolog pseudogene 8	Pseudogene
12q14	*RAB11AP2*	RAB11A, member RAS oncogene family pseudogene 2	Pseudogene
12q22	*PGAM1P5*	Phosphoglycerate Mutase 1 Pseudogene 5	Pseudogene
12q23	*RNA5SP371*	RNA, 5S Ribosomal Pseudogene 371	Pseudogene
12q24	*RP11-968O1.5*	---	Sense overlapping
12q24	*RP11-428I12.1*	---	Pseudogene
12q24	*RP11-572C21.1*	---	LncRNA
12q24	*RP11-1028N23.3*	---	LncRNA
12q24	*RP11-121J20.1*	---	LncRNA
12q24	*DIABLO*	Diablo, IAP-Binding Mitochondrial Protein	Apoptosis promotion
12q24	*LRRC43*	Leucine Rich Repeat Containing 43	---
12q24	*RP11-64B16.4*	---	Antisense RNA
13q13	*EXOSC8*	Exosome Component 8	Transcriptional regulation
14q11	*AE000660.4*	---	Pseudogene
14q11	*OR4Q3*	Olfactory receptor, family 4, subfam. Q, mb 3	G-protein-coupled olfactory receptor
14q11	*TRAV-26-1*	T Cell Receptor Alpha Variable 26-1	TCR assembly
14q11	*TRAJ10*	T Cell Receptor Alpha Joining 10	TCR assembly

14q11	*TRAJ13*	T Cell Receptor Alpha Joining 13	TCR assembly
14q12	*RP11-468E2.6*	---	---
14q32	*HOMER2P1*	Homer Scaffolding Protein 2 Pseudogene 1	Pseudogene
14q32	*IGHVIII-2-1*	Immunoglobulin Heavy Variable (III)-2-1	BCR assembly
14q32	*IGHVIII-16-1*	Immunoglobulin Heavy Variable (III)-16-1	BCR assembly
14q32	*DCAF11*	DDB1 And CUL4 Associated Factor 11	Assembly of the DDB1–CUL4A ubiquitin ligase
15q15	*CATSPER2*	Cation Channel, Sperm Associated 2	Sperm motility and male fertility
15q25	*RP11-182J1.14*	---	Pseudogene
16p13	*NARFL*	Nuclear prelamin A recognition factor-like	Modulation of hypoxia-inducible factor-1α activity
16q22	*HAS3*	Hyaluronan Synthase 3	Synthesis of glycosaminoglycan hyaluronan
16q22	*CDH16*	Cadherin 16	Ca2+-dependent glycoprotein
16q22	*CLEC18B*	C-Type Lectin Domain Family 18, Member B	---
17p11	*LGALS9C*	Lectin, Galactoside-Binding, Soluble, 9C	Regulation of immune response
19p13	*LRG1*	Leucine-rich alpha-2-glycoprotein 1	Signal transduction, and cell adhesion
19p13	*ZNF414*	Zinc Finger Protein 414	Transcriptional regulation
19p13	*AC008686.1*	---	Uncharacterized protein
22q12	*RP5-1170K4.7*	---	Antisense RNA
22q13	*RN7SKP252*	---	misc RNA
22q13	*RPL35P8*	Ribosomal Protein L35 Pseudogene 8	Pseudogene
22q13	*OGFRP1*	opioid growth factor receptor pseudogene 1	LncRNA
22q13	*ADSL*	Adenylosuccinate lyase	*De novo* purine biosynthesis



Supplemental figure 1: Observed distance versus expected distance using the permutation-based SAM method. One hundred and forty-eight genes presented with an observed distance superior to the expected distance, with a false discovery rate < 0.05.

DISCUSSION GENERALE

La première partie de ce travail a été réalisée à Dijon, sous la direction du Professeur François GIRODON. Nous avons cherché à démontrer que l'ARS-T peut être considérée comme une entité indépendante de l'ARS et de la TE. Pour cela, nous avons réuni une cohorte européenne, que nous avons étudiée sur les plans épidémiologique, clinique, cellulaire et moléculaire.

Au sein de la classification OMS 2008, l'ARS-T est considérée comme une entité provisoire. Son existence en tant qu'entité indépendante demeure controversée, car certains spécialistes considèrent qu'il s'agit d'une forme particulière de TE présentant un excès de sidéroblastes en couronne, alors que d'autres avancent l'hypothèse qu'il s'agit d'une forme d'ARS présentant secondairement une thrombocytose à la faveur de l'acquisition d'une mutation *JAK2V617F*. En 2012, au terme d'une première étude précédant ce travail, nous avions montré, en réunissant une grande cohorte européenne, que l'ARS-T était une entité indépendante de l'ARS et de la TE sur le plan biologique, pronostique et moléculaire. En particulier, la forte proportion de mutations *JAK2V617F* (environ 50% des cas) distingue clairement ARS-T et ARS.[41]

La découverte d'une forte proportion de mutations de *SF3B1* dans les SMD avec sidéroblastes en couronne nous a amené à rechercher ces mutations sur une grande cohorte d'ARS-T. Une mutation de *SF3B1* a été mise en évidence dans 86,5% à 88,3% des cas.[167, 169] Les mutations affectant d'autres acteurs de l'épissage alternatif comme *U2AF35*, *ZRSR2* et *SRSF2* sont présentes dans 7% des SMD avec sidéroblastes en couronne *SF3B1* non-mutés.[42] Nous n'avons pas recherché ces mutations dans notre cohorte, mais, cumulées, les mutations de *SF3B1*, *U2AF35*, *ZRSR2* et *SRSF2* expliqueraient le caractère myélodysplasique dans plus de 90% des ARS-T. La forte proportion de mutations de *SF3B1* dans l'ARS-T permet de démontrer que l'ARS-T est indépendante de la TE.

Seules 50% des ARS-T présentent une mutation de *JAK2* ou de *MPL*. L'enjeu est donc d'identifier la ou les mutation(s) responsable(s) du caractère myéloprolifératif des ARS-T double négatives pour *JAK2* et *MPL*. Cet enjeu existe aussi pour les SMP.

Les mutations de *CALR* ont été découvertes dans 67 à 71% des TE et 56 à 88% des MFP *JAK2* et *MPL* non-mutées. Ces mutations ont été décrites comme mutuellement exclusives. Désormais, plus de 90% des TE et des MFP sont associées à un marqueur clonal expliquant le caractère myéloprolifératif : *JAK2*^{V617F}, mutations de *MPL* et mutations de *CALR*. Aucune mutation

de *CALR* n'a été retrouvée dans la PV.[19, 20] Nos travaux ont permis d'une part de démontrer qu'il existe de très rares cas de PV présentant une mutation de *CALR*,[176] et que les mutations de *CALR* sont retrouvées dans environ 1% des ARS-T.[177]

Sur le plan phénotypique, l'utilisation de souris irradiées à doses létales et transplantées avec de la moelle osseuse infectée par un rétrovirus exprimant *CALR*^{mut} montre que les mutants de type 1 (del 52pb) sont responsables d'un tableau de thrombocytose marquée et d'une progression rapide vers un tableau de MFP, alors que les mutants de type 2 (ins 5pb) sont responsables d'une TE stable.[178, 179] Les mutants de *CALR* activent constitutivement la voie JAK-STAT par le biais de MPL, le récepteur à la TPO, induisant une expansion de la lignée mégacaryocytaire.[179-181]

Sur le plan physiopathologique, (i) il existerait au sein de l'appareil de Golgi une interaction anormalement stable entre la calréticuline mutante et MPL, puis (ii) le récepteur MPL exprimé à la surface de la CSH reste couplé à la calréticuline mutante, (iii) MPL est activé indépendamment de la présence de TPO, ce qui a pour effet la dimérisation puis l'activation de JAK2. La survie et la croissance des CSH dépendent de la TPO. Au stade de précurseur hématopoïétique, seuls les mégacaryocytes demeurent sensibles à la TPO. L'expression d'une protéine calréticuline mutante a pour conséquence une expansion clonale MPL-induite des CSH, la production excessive de plaquettes par les mégacaryocytes puis une fibrose médullaire (**Figure 36**).[179-182]

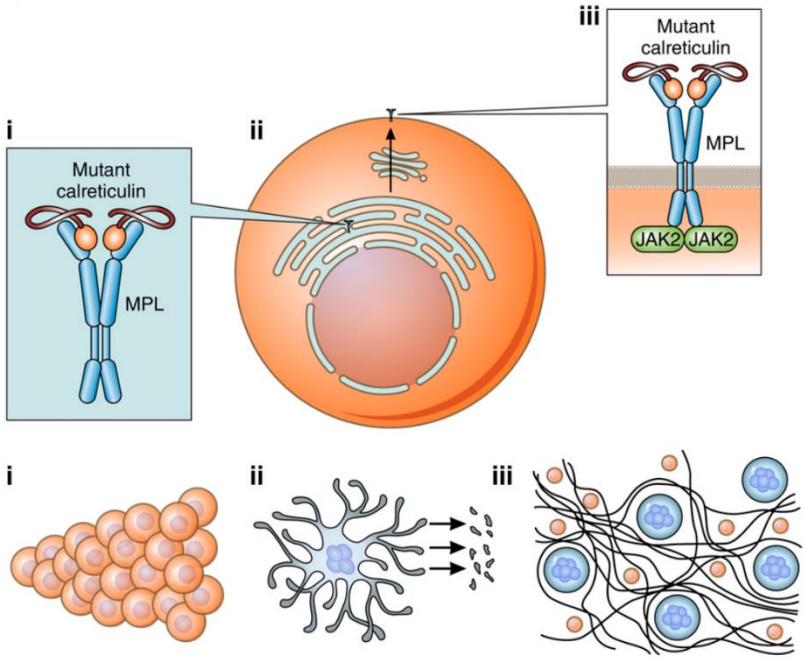


Figure 36. Mécanismes physiopathologiques induits par les mutants de *CALR* dans les SMP. D'après Cazzola *et al*.[182]

L'existence d'une mutation de *CALR* permet d'expliquer la thrombocytose et la myélofibrose dans les TE et les MFP *JAK2* non-mutées. Par contre, la très faible proportion de ces mutations dans l'ARS-T ne permet pas d'expliquer la thrombocytose dans les cas *JAK2* non-mutés.

Au final, l'ARS-T constitue une entité indépendante, caractérisée sur le plan moléculaire par une mutation de *SF3B1* dans près de 90% des cas, ce qui la distingue de la TE, et par une forte proportion de mutations de *JAK2* ou de *MPL*, ce qui la distingue de l'ARS (**Figure 37**).[30]

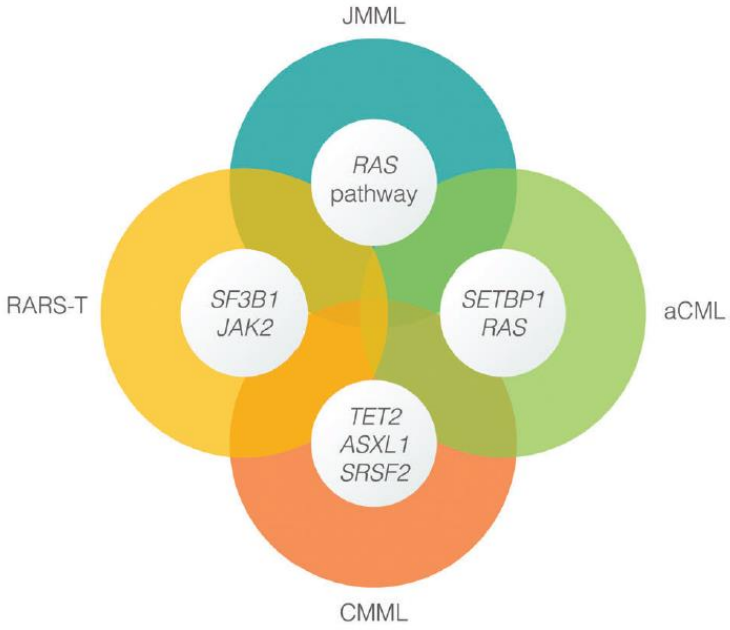


Figure 37. Caractéristiques moléculaires des syndromes mixtes myélodysplasiques/myéloprolifératifs.
D'après Mughal *et al.* [30]

Toutefois, près de 50% des ARS-T ne présentent ni mutation de *JAK2,* ni mutation de *MPL*, ni mutation de *CALR*. L'enjeu actuel est donc d'identifier la ou les mutation(s) responsable(s) de la thrombocytose dans ces cas d'ARS-T. La suite logique serait de réaliser le séquençage de l'exome de nos cas d'ARS-T *JAK2* non-mutées.

Mon arrivée à Nancy s'est accompagnée d'un changement d'équipe de recherche. Au sein de l'unité Inserm U954, notre thématique de recherche porte sur la génomique des LBDGC. Mon apprentissage de la génomique s'est fait en collaboration avec le docteur Gaïli CHEN.[159]

Dans la deuxième partie de ce travail, nous avons voulu analyser le génome des LBDGC en rechute. En particulier, notre objectif était de déterminer quelles anomalies étaient associées aux deux différents types de rechutes, à savoir rechutes précoces (survenant moins de 12 mois après la fin du traitement de première ligne) et rechutes tardives (survenant plus de 12 mois après la fin du traitement de première ligne).[183]

L'étude CORAL avait pour but de comparer les deux traitements de rattrapage les plus utilisés : R-ICE et R-DHAP. Cette étude a identifié plusieurs facteurs de mauvais pronostic à la rechute, dont le délai de la rechute, avec un seuil fixé à 12 mois. Nous avons eu la possibilité d'étudier 39 échantillons issus de cette étude, répartis de manière équilibrée entre rechutes précoces et rechutes tardives.

Ces échantillons ont été analysés par *SNP-array*. Cette méthode permet l'analyse à haute résolution des anomalies de nombre de copies géniques. Notre objectif était de déterminer s'il existe des délétions ou des duplications récurrentes associées à l'un ou l'autre sous-type de rechute. Pour cela, il est nécessaire de réaliser un grand nombre de tests, ce qui expose à un grand nombre de faux positifs. Pour éviter cet écueil, nous avons créé une méthode originale, dérivée de la méthode SAM, qui est basée sur la permutation des données ou *bootstrap*. Ceci permet de limiter le taux de faux positifs à 5%, si l'on compare le résultat de chaque test à la moyenne des résultats obtenus à la suite de 20 permutations.

Cette méthode nous a permis d'identifier une liste de 56 gènes codant pour des protéines et 27 longs ARN non-codants dont le profil de répartition des délétions et duplications est différentiel entre rechutes précoces et rechutes tardives. La classification hiérarchique des anomalies de nombre de copies détectées a montré qu'il existait des signatures regroupant des anomalies récurrentes définissant 3 groupes de malades dont 2 sont entièrement composés de rechutes précoces ou de rechutes tardives. Nous avons pu mettre en évidence des profils spécifiques, ce qui est une première étape de caractérisation du substrat génétique qui sous-tend la chimiorésistance et le délai de rechute.

La chimiorésistance élevée des LBDGC en rechute précoce pourrait être expliquée par des anomalies affectant les régulateurs de l'apoptose et du cycle cellulaire en particulier (i) des délétions de *CDKN2A* dans 42% des cas, ce qui est supérieur à ce qui a été décrit dans les LBDGC *de novo*,[184] (ii) des délétions de *DIABLO* dans 26% des cas et (iii) des délétions de *CK2B*,

codant pour la sous-unité régulatrice de la kinase CK2. La forte proportion de délétions de *CDKN2A* dans les LBDGC en rechute précoce ouvre l'hypothèse d'un lien entre les deux.

De plus, les LBDGC en rechute précoce présentent de nombreuses anomalies de régulateurs de la transcription des gènes, agissant en particulier sur le degré de compaction des histones. Les anomalies affectant les gènes codant pour l'histone H1 ou son régulateur PTMS sont fréquentes (21 à 31% des cas) et fréquemment présentes chez les mêmes malades. La dérégulation épigénétique pourrait donc être en lien avec la précocité de la rechute. Par ailleurs, nous avons trouvé de nombreuses anomalies géniques affectant de longs ARN non-codants, qui entrent, eux-aussi, dans la régulation épigénétique de la transcription. AIRN (*Air noncoding RNA*) est l'objet d'une délétion dans 20% des LBDGC en rechute précoce. Or ces anomalies ne sont retrouvées que les LBDGC en rechute précoce, ce qui renforce l'hypothèse d'un rôle de cette dérégulation dans la précocité de la rechute.

Les LBDGC en rechute tardive sont d'avantage liés à des anomalies des acteurs de la réponse immune, avec de fréquentes anomalies affectant B2M et CD58. Devant ce résultat, nous pouvons formuler l'hypothèse d'une adaptation du clone malin face à son environnement et d'une ré-évolution tumorale plus lente, à la faveur d'une tolérance immune.

Notre analyse a permis d'isoler trois groupes, dont deux très homogènes, composés respectivement de rechutes précoces et de rechutes tardives, et présentant des anomalies récurrentes. Le troisième groupe est composé de rechutes précoces et de rechutes tardives, sans anomalies caractéristiques. Une étude sur un plus grand nombre d'échantillons permettrait d'affiner ces résultats.

CONCLUSION ET PERSPECTIVES

Parmi les ouvertures offertes par ce travail, nous fixons pour objectif : (i) de réaliser une étude génétique-génomique sur les LBDGC en rechute et sur les LZM et (ii) de réaliser une étude génomique du syndrome de Richter.

1. ETUDE GENETIQUE-GENOMIQUE DES LYMPHOMES

Le concept génétique génomique a été proposé pour la première fois en 2001.[185] Le développement des techniques de génomique, en particulier les puces à ADN a permis d'obtenir de grandes quantité de données sur les séquences géniques (polymorphisme de l'ADN, anomalies de nombre de copies) et sur le niveau d'expression des gènes (transcriptome). La question se posait alors d'une intégration de ces données pour mieux comprendre les systèmes biologiques et les voies de régulation impliquées dans les pathologies. Le concept génétique-génomique consiste à évaluer l'influence du polymorphisme génétique sur le niveau d'expression des gènes. On peut par exemple mesurer l'influence d'une mutation d'une base de l'ADN (*Single Nucleotide Polymorphism*/SNP) sur le niveau d'expression de tous les gènes, en comparant de manière systématique leur niveau d'expression selon le génotype. La même démarche peut aussi être appliquée dans les situations où il existe des anomalies de nombre de copies géniques.

L'objectif est de tirer parti des anomalies géniques pour objectiver l'ascendance de certains gènes sur d'autres gènes et de reconstruire des réseaux de régulation (**Figure 38**).

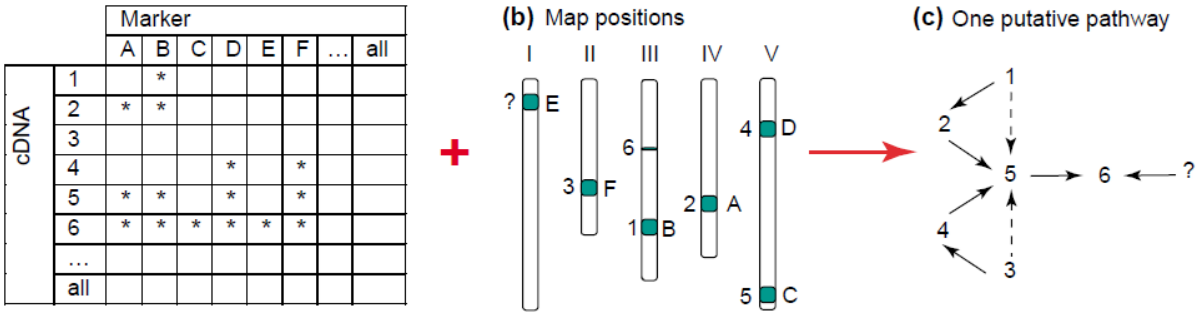


Figure 38. Reconstitution des réseaux de régulation génique grâce à l'approche génétique-génomique.
D'après Jansen *et al.*[185].

La première application a porté sur la plante *Arabidopsis*, un système biologique simple, dont la séquence génique était parfaitement connue et accessible à un traitement statistique peu complexe, compte-tenu du faible nombre de gènes exprimés.(185)

Le concept génétique-génomique s'est progressivement développé. Grâce aux progrès de la bioinformatique, la puissance de calcul ouvre la possibilité d'une application du concept à des systèmes biologiques plus complexes.(186) L'objectif demeure de faire le lien entre les anomalies de séquence et leurs conséquences phénotypiques, en particulier pathologiques. L'intermédiaire entre les deux est le niveau d'expression des gènes, c'est-à-dire le transcriptome. Le niveau d'expression de chaque gène est donc considéré ici comme un trait quantitatif ou « *quantitative trait loci »* (QTL) gouverné par des variations au niveau de la séquence génique (*Expression Quantitative Trait Loci* / E-QTL) (**Figure 39**). Le croisement systématique des données issues de l'analyse de la séquence d'ADN et des données du transcriptome permet (i) d'identifier des « *master regulators* » qui affectent le niveau d'expression de centaines de gènes, puis (ii) de reconstruire des réseaux de régulation génique, et (iii) de mieux comprendre la régulation du niveau d'expression des gènes, en particulier en identifiant ses séquences régulatrices qu'elles soient situées à proximité de la séquence codante (« *Cis-regulation* ») ou à distance (« *Trans-regulation* »).

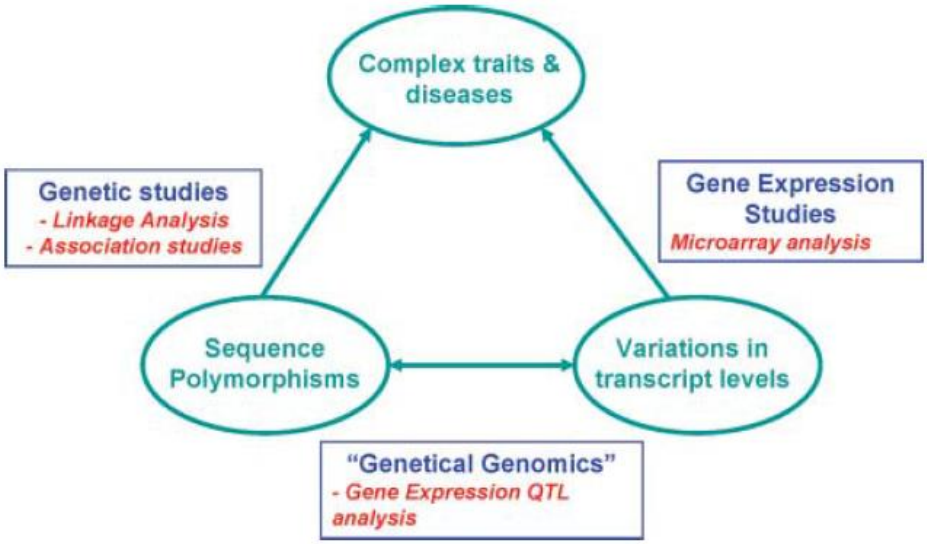


Figure 39. Le concept génétique-génomique et la correspondance entre séquence génique (ADN), niveau d'expression des gènes (ARN) et phénotype.
D'après Li *et al*. (186).

Jusqu'à présent, des études ont appliqué le concept génétique-génomique à la Souris,(187, 188) au Maïs,(187) au Rat,(189) et à la Levure.(190, 191) Deux études ont porté sur des lignées lymphoblastiques Humaines mais ont obtenu des résultats discordants, les outils et les méthodes

statistiques utilisées n'étant pas adaptés.[192, 193] La question se pose à présent de répondre à la complexité biologique et statistique d'une étude génétique-génomique appliquée à l'Homme.

Notre objectif est d'utiliser le concept génétique-génomique pour mieux appréhender la problématique des LBDGC en rechute. Pour 30 de nos échantillons issus de l'étude CORAL, nous disposons à la fois de données d'expression (acquises sur puce *Agilent Whole Human Genome Microarray 4x44K*®) et de données de génotypage pour 936 000 SNP. Notre méthode consiste à croiser de manière systématique et à très grande échelle les données de génotypage (*SNP-array*, exome, génome complet) et les données d'expression (puces à ADN, séquençage de l'ARN) pour extraire les résultats les plus significatifs et reconstituer les réseaux de régulation perturbés dans les LBDGC en rechute (**Figure 40**).

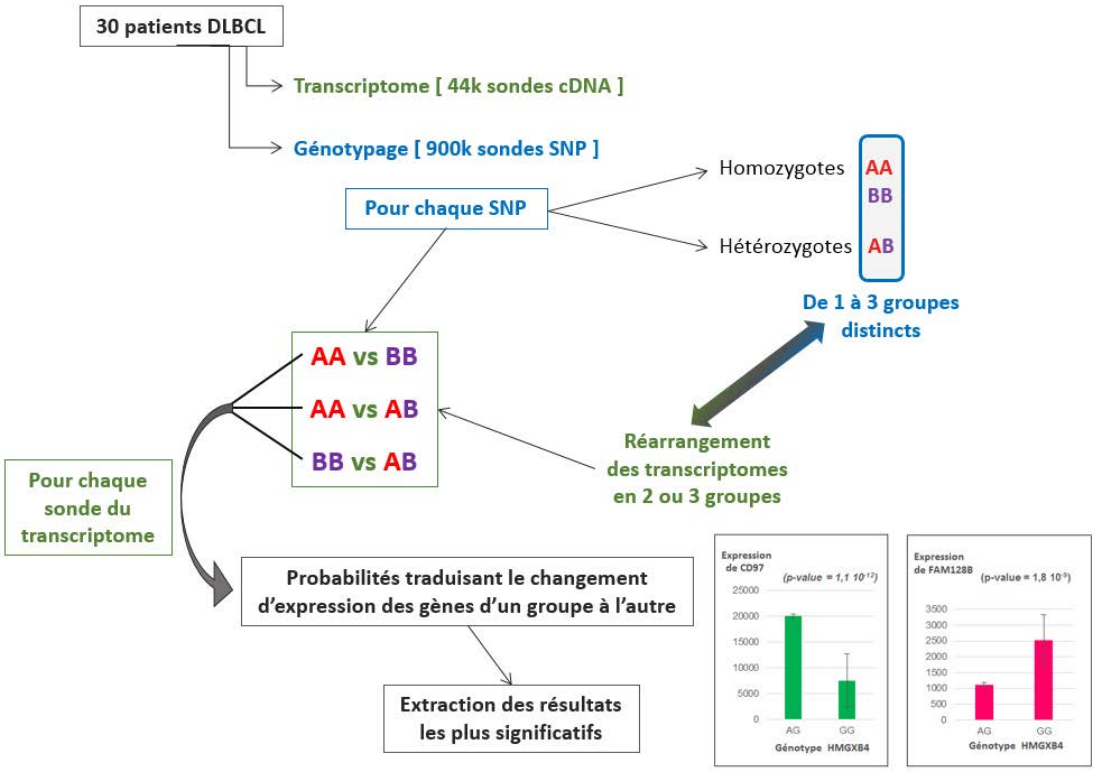


Figure 40. Méthode de croisement systématique des données de génotypage et d'expression. D'après Sébastien Hergalant.

Par ailleurs, nous appliquerons la même méthode pour identifier les perturbations des réseaux de régulation (i) dans les LBDGC selon leur statut mutationnel du gène *MYC* et (ii) dans les lymphomes de la zone marginale. Dans les deux cas, nous disposons de données d'expression et de données de séquençage de l'exome. Ce travail nécessitera la réalisation de milliards de tests et la mise au point d'une méthode statistique qui l'extraction des résultats les

plus significatifs. Parmi les difficultés que nous avons déjà identifiées, la problématique des regroupements de plusieurs SNP au sein d'un même haplotype a retenu notre attention. Dans cette situation plusieurs variants nucléotidiques sont identifiés simultanément chez le même malade. Ainsi, il est difficile de déterminer lequel de ces variants est lié aux variations d'expression des gènes. Par contre, le mode de régulation des gènes sera plus aisé à décrire car celui-ci se fait toujours dans le même sens, le gène présentant un polymorphisme de séquence régulant ceux dont le niveau d'expression change, et non l'inverse.

2. ETUDE GENOMIQUE DU SYNDROME DE RICHTER

La leucémie lymphoïde chronique (LLC) est la leucémie la plus fréquente dans les pays occidentaux avec 4000 nouveaux cas en 2012 en France. Elle est considérée comme une hémopathie lymphoïde B de bas grade dans la classification 2008 de l'Organisation Mondiale de la Santé (OMS).[194] Toutefois, 2,3 à 9% des LLC vont se transformer un lymphome de haut grade se présentant dans 90% des cas avec une histologie de lymphome B-diffus à grandes cellules (LBDGC) : le syndrome de Richter (SR). Dans plus de 90% des cas, le SR présente le profil immunohistochimique d'un LBDGC de phénotype « *non-Germinal Center-like* ».[195-198]

Sur le plan clinique, la transformation d'une LLC en SR s'accompagne généralement de signes généraux, de l'apparition de tuméfactions extra-ganglionnaires et/ou d'une adénopathie >5 cm avec augmentation rapide du volume ganglionnaire.

L'incidence des SR est évaluée à 0,5 à 1% par an mais est directement dépendante de la politique de réalisation des biopsies ganglionnaires, qui est variable d'un centre à l'autre.[199] Le taux de transformation est variable selon les séries, la durée de suivi et la réalisation systématique ou non de biopsies à but diagnostique dans des situations cliniques évocatrices de transformation. L'incidence cumulative est estimée à 2,1% à 5 ans et 4,8% à 10 ans, soit un risque de transformation de 0,5% par an.

Le délai entre le diagnostic de LLC et la transformation en SR est court. La plupart des transformations surviennent dans les 4 ans, avec un délai médian de 23 mois. Le délai de transformation le plus long jamais observé est de 82 mois.[197-201]

Le SR est un LBDGC secondaire constituant une évolution dramatique de la LLC, de par son agressivité particulière, conjuguant forte évolutivité tumorale et chimiorésistance. La survie est plus faible que dans les LBDGC *de novo*. Ces observations indiquent une pathogénèse différente dans le SR.

Il n'existe pas d'étude randomisée comparant différentes approches thérapeutiques dans le SR mais plusieurs études ont testé chacune un traitement spécifique. Le protocole de première ligne le plus répandu est le R-CHOP (Rituximab, Cyclophosphamide, Hydroxydaunorubicine, Oncovin, Prednisone), qui permet, lorsqu'il peut être proposé, d'obtenir une médiane de survie de 15 mois.[202, 203] Des protocoles de chimiothérapie plus intensifs type OFAR (Oxiplatine, Fludarabine, Aracytine, Rituximab) n'ont permis d'obtenir qu'un taux de réponses complètes à 6,5% et une médiane de survie de 6 à 8 mois.[204]

Ces protocoles de première ligne permettent dans 10 à 15% des cas d'obtenir une réponse partielle ou une réponse complète, mais de courte durée, n'offrant qu'une médiane de survie de 12 mois (27% de survie à 3 ans). La consolidation par une auto- ou une allogreffe de cellules souches hématopoïétiques chez un petit nombre de patients sélectionnés par leur réponse à la chimiothérapie permet d'obtenir une survie plus longue, de l'ordre de 75% à 3 ans.[196, 205]

Les caractéristiques moléculaires du SR sont peu connues et les données établies ne permettent pas une caractérisation physiopathologique précise ni la conception de nouvelles molécules thérapeutiques efficaces. Les anomalies identifiées à ce jour affectent les « régulateurs généraux » de la vie cellulaire. Dans le modèle moléculaire actuel, l'évolution du clone LLC en SR est associée à des dérégulations de la prolifération cellulaire, de l'apoptose et du cycle cellulaire, principalement en raison d'anomalies de *TP53*, *NOTCH1*, *MYC* et *CDKN2A*.[206]

L'influence du traitement de la LLC sur le risque de transformation en SR n'est pas établie. Dans une étude de 2008 portant sur des patients atteints de LLC, 82/185 (44,3%) des malades ont été traités. Le nombre de lignes de traitement était très hétérogène dans la mesure où 36/82 malades (43,9%) ont bénéficié d'une seule ligne de traitement, 22/82 malades (26,8%) ont bénéficié de deux lignes de traitement et 25/82 (30,4%) malades ont bénéficié d'au moins 3 lignes de traitement. Les agents alkylants ont été proposés à une majorité de malades : 63/82 (76,8%). Les combinaisons contenant de la Fludarabine ont bénéficié à 47/82 (57,3%) des malades, les combinaisons contenant du Rituximab à 40/82 (48,7%) des malades, ou de l'Alemtuzumab : 11/82 (13,4%).[197] L'ensemble illustre la grande hétérogénéité de prise en charge des malades, rendant très complexe l'attribution d'un risque accru de transformation à un traitement en particulier.

Une étude publiée en 2013 et portant sur 1641 malades a montré que la transformation survient dans la moitié des cas sur un terrain de LLC non traitée et dans l'autre moitié des cas sur une LLC traitée. Le traitement de la LLC est associé à un risque de transformation en SR de 1% par an (5% à 5 ans, 15,2% à 10 ans). Les combinaisons d'analogues des purines, qui empêchent la réparation de l'Acide DesoxyriboNucléique, avec des agents alkylants, qui induisent des cassures de l'ADN, augmentent le risque de transformation à 1,5% par an. Mais l'utilisation des différentes molécules en monothérapie (analogue des purines, agent alkylant ou immunothérapie) n'augmente pas le risque de transformation en SR.[198]

A l'heure actuelle, la combinaison Rituximab-Fludarabine-Cyclophosphamide (RFC) est le traitement de référence en première intention chez les patients en bon état général et atteints d'une LLC dont le clone tumoral ne présente pas de délétion 17p.[207]

Au stade de LLC, la relation entre (i) les caractéristiques biologiques du clone malin, (ii) l'ancienneté de la maladie, (iii) la nature des traitements entrepris pour soigner la LLC et (iv) le risque de développement d'un SR, est mal comprise. Il s'agit pourtant d'un point capital car la capacité à prévoir un risque de transformation en SR dès le stade de LLC permettrait de proposer aux malades concernés une prise en charge adaptée et de leur offrir une meilleure survie.

Les caractéristiques moléculaires des SR sont peu connues. L'analyse de la relation clonale entre le clone LLC et le clone transformé en SR est fondée sur l'étude du réarrangement du gène *IgH* VDJ CDR3. En cas de relation clonale, la séquence *IgH* VDJ sera identique au stade de LLC et au stade de SR.[200] Selon les études, 80 à 90% des SR présentent un réarrangement *IgH* VDJ CDR3 correspondant à celui identifié chez le clone LLC initial, signant une filiation et donc une évolution linéaire du clone malin. Ces SR présentent une médiane de survie de 14,2 mois et sont considérés comme les « véritables SR ». A l'opposé, les 10 à 20% de SR ne présentant pas de relation clonale avec la LLC initiale ont une médiane de survie très proche de celle des LBDGC *de novo* (62,5 mois) et sont considérés par la plupart des auteurs comme des néoplasies indépendantes. La notion de relation clonale entre le clone LLC et le clone transformé est une dimension essentielle à l'heure actuelle. La relation clonale est considérée comme le facteur pronostic le plus probant et la vérification de la relation clonale LLC/SR, pourtant non réalisée en pratique courante, est encouragée par certains auteurs.[201]

Sur le plan moléculaire, les anomalies identifiées à ce jour affectent les « régulateurs généraux » de la vie cellulaire. Dans le modèle moléculaire actuel, l'évolution du clone LLC en SR est associée à des dérégulations de la prolifération cellulaire, de l'apoptose et du cycle cellulaire, principalement en raison d'anomalies de *TP53*, *NOTCH1*, *MYC* et *CDKN2A*. Seuls 10% des RS ne présentent aucune de ces 4 anomalies. 50% des cas de transformation en SR sont associés à l'acquisition d'une inactivation de TP53, d'une activation de C-MYC ou d'une délétion de *CDKN2A*. Dans 30% des cas, la transformation est associée à une trisomie 12 et une mutation de *NOTCH1* et ces anomalies sont mutuellement exclusives avec les inactivations de *TP53* ou *CDKN2A*. Les 20% des transformations restantes sont associées à d'autres anomalies génétiques (**Figure 41**).[202, 206, 208, 209]

Contrairement aux LBDGC *de novo*, dont la pathogénie moléculaire est très bien connue, on ne connait pas à l'heure actuelle de « lésion(s) moléculaire(s) » unique(s) ou combinée(s), ni de mise en jeu constitutive d'une ou plusieurs voie(s) d'activation cellulaire(s) pouvant être responsable(s) de la transformation en SR ou de l'agressivité particulière de cette maladie.

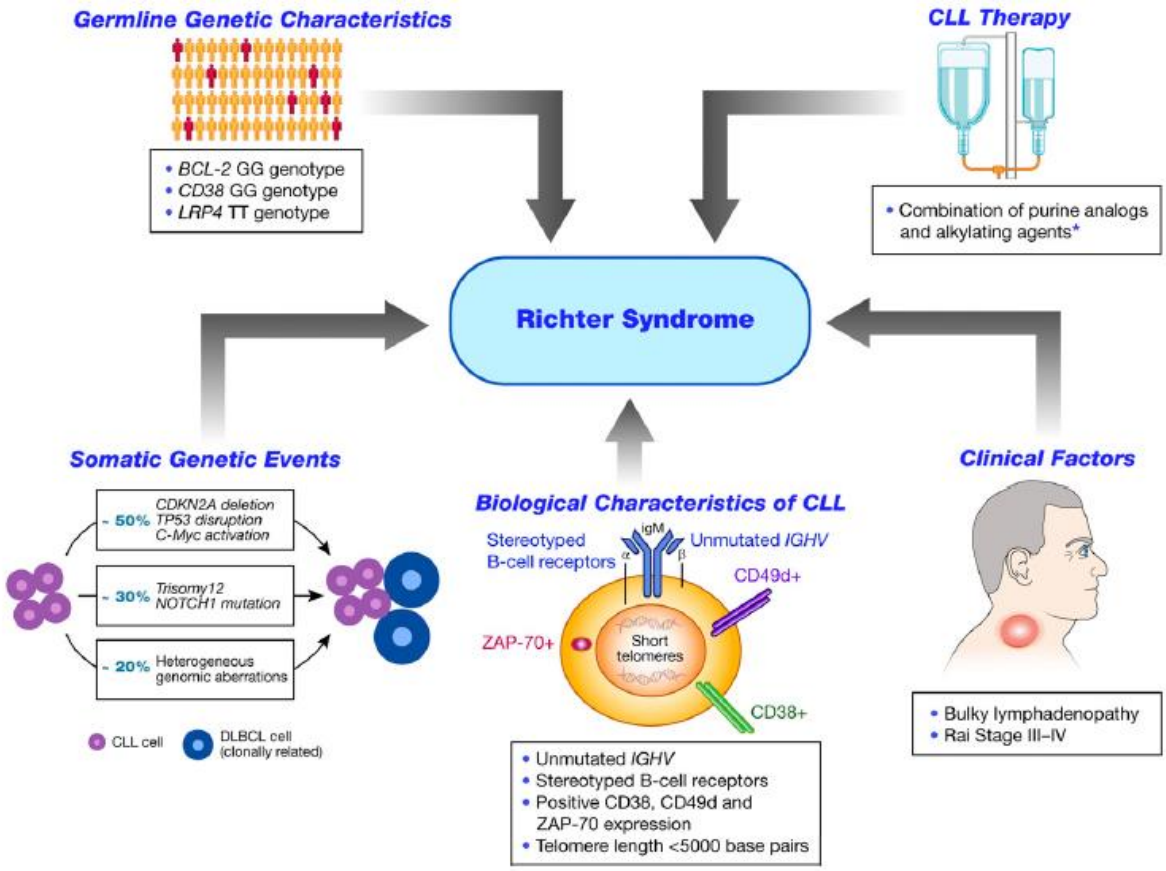


Figure 41. Les déterminants de la transformation de la LLC en SR.
D'après Parikh *et al*.[202]

Aucune de ces anomalies ne peut définir un SR, que ce soit à elle-seule ou sous forme combinée. Les bases moléculaires du développement du SR sont donc très méconnues. Une meilleure connaissance de ces mécanismes permettrait le développement de molécules thérapeutiques efficaces.

Aucune étude transcriptomique du SR n'a été réalisée à ce jour. Le potentiel d'une telle étude est pourtant réel. Une étude par séquençage de l'ARN a permis de caractériser le panorama transcriptionnel de la LLC à la fois sur le plan du niveau d'expression des gènes, mais aussi du point de vue de l'épissage alternatif et des différents isoformes transcriptionnels qu'il génère. Ainsi le séquençage de l'ARN a pu identifier : (i) des profils d'expression géniques spécifiques de deux sous-types de LLC, (ii) des modifications de la répartition des différents transcrits alternatifs physiologiques de plusieurs gènes, dont *BCL2* et *SF3B1*, (iii) l'existence de transcrits nouveaux, spécifiques des cellules tumorales et (iv) la dérégulation de différentes voies d'activation cellulaires comme la voie du BCR et la voie JAK-STAT, associée à une prolifération tumorale accrue.[210] Cette étude n'évoque pas la transformation de la LLC en SR.

Nous souhaitons réaliser une étude comparative par séquençage de l'ARN entre LLC non évolutives, futures LLC transformées et SR, pour identifier : (i) les gènes dont le niveau d'expression est différentiel (ii) les différences de répartition des différents transcrits alternatifs physiologiques, (iii) l'existence de transcrits spécifiques des différentes conditions comparées et *in fine* (iv) les perturbations de fonctions cellulaires à l'origine de la transformation de certaines LLC en SR, et de l'agressivité et la chimiorésistance particulières du SR.

Notre première hypothèse est que le SR présente des anomalies transcriptionnelles à l'origine de dérégulations cellulaires spécifiques, expliquant son caractère particulièrement agressif. La caractérisation de ces anomalies permettrait la conception de thérapeutiques ciblées efficaces dans le SR.

Notre deuxième hypothèse est que les déterminants génétiques de la transformation en SR sont présents dès le diagnostic de LLC. Une étude par séquençage de l'ARN nous permettrait d'identifier les signatures transcriptomiques et les anomalies de l'épissage alternatif caractéristiques des futures LLC transformées. La connaissance des déterminants de la transformation dès le stade de LLC permettrait de proposer aux malades concernés une surveillance adaptée, de faire le diagnostic de transformation plus tôt et d'offrir aux malades une meilleure survie.

REFERENCES BIBLIOGRAPHIQUES

1. Maynadié M, Girodon F, Manivet-Janoray I, Mounier M, Mugneret F, Bailly F, et al. Twenty-five years of epidemiological recording on myeloid malignancies: data from the specialized registry of hematologic malignancies of Cote d'Or (Burgundy, France). Haematologica. 2011;96(1):55-61.
2. Troussard X MM, Duchenet V, Mouchel D, Chèze S, Collignon A. Épidémiologie des syndromes myélodysplasiques (SMD) et des syndromes myélodysplasiques/syndromes myéloprolifératifs (SMD/SMP). Revue Francophone des Laboratoires. 2009;413(Juin 2009):25-9.
3. Vardiman J, Brunning R, Arber D, Le Beau M, Porwit A, Tefferi A, et al. Introduction and overview of the classification of the myeloid neoplasms. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of tumours of haematopoietic and lymphoid tissues 4th ed. Lyon, France: IARC; 2008. p. 18-30.
4. 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. 2005;434(7037):1144-8.
5. 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. 2005;352(17):1779-90.
6. Baxter EJ, Scott LM, Campbell PJ, East C, Fourouclas N, Swanton S, et al. Acquired mutation of the tyrosine kinase JAK2 in human myeloproliferative disorders. Lancet. 2005;365(9464):1054-61.
7. Levine RL, Wadleigh M, Cools J, Ebert BL, Wernig G, Huntly BJ, et al. Activating mutation in the tyrosine kinase JAK2 in polycythemia vera, essential thrombocythemia, and myeloid metaplasia with myelofibrosis. Cancer Cell. 2005;7(4):387-97.
8. 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. 2006;108(5):1781-2.
9. 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. 2006;20(11):2060-1.
10. 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. 2006;91(5):719-20.
11. Renneville A, Quesnel B, Charpentier A, Terriou L, Crinquette A, Laï JL, et al. High occurrence of JAK2 V617 mutation in refractory anemia with ringed sideroblasts associated with marked thrombocytosis. Leukemia. 2006;20(11):2067-70.
12. 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. 2008;22(2):453-5.
13. Steensma DP, Tefferi A. JAK2 V617F and ringed sideroblasts: not necessarily RARS-T. Blood. 2008;111(3):1748.
14. 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. 2006;108(7):2173-81.
15. 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. 2006;20(9):1641-4.
16. 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. 2007;109(3):1334-5.

17. 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. 2006;108(10):3472-6.
18. Cross NC. Genetic and epigenetic complexity in myeloproliferative neoplasms. Hematology Am Soc Hematol Educ Program. 2011;2011:208-14.
19. 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. 2013.
20. 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. 2013.
21. Thiele J, Kvasnicka H, Orazi A, Tefferi A, Gisslinger H. Essential thrombocythaemia. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of tumours of haematopoietic and lymphoid tissues, 4th ed. Lyon, France: IARC; 2008. p. 48-50.
22. Lippert E, Boissinot M, Kralovics R, Girodon F, Dobo I, Praloran V, et al. The JAK2-V617F mutation is frequently present at diagnosis in patients with essential thrombocythemia and polycythemia vera. Blood. 2006;108(6):1865-7.
23. Dobo I, Boiret N, Lippert E, Girodon F, Mossuz P, Donnard M, et al. A standardized endogenous megakaryocytic erythroid colony assay for the diagnosis of essential thrombocythemia. Haematologica. 2004;89(10):1207-12.
24. Girodon F, Dutrillaux F, Broséus J, Mounier M, Goussot V, Bardonnaud P, et al. Leukocytosis is associated with poor survival but not with increased risk of thrombosis in essential thrombocythemia: a population-based study of 311 patients. Leukemia. 2010;24(4):900-3.
25. Thiele J, Kvasnicka H, Orazi A, Tefferi A, Birgegard G. Polycythaemia Vera. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of tumours of haematopoietic and lymphoid tissues, 4th Ed. Lyon, France: IARC; 2008. p. 40-3.
26. Cazzola M, Malcovati L. Myelodysplastic syndromes--coping with ineffective hematopoiesis. N Engl J Med. 2005;352(6):536-8.
27. Greenberg P, Cox C, LeBeau MM, Fenaux P, Morel P, Sanz G, et al. International scoring system for evaluating prognosis in myelodysplastic syndromes. Blood. 1997;89(6):2079-88.
28. Greenberg PL, Tuechler H, Schanz J, Sanz G, Garcia-Manero G, Solé F, et al. Revised international prognostic scoring system for myelodysplastic syndromes. Blood. 2012;120(12):2454-65.
29. Brunning R, Orazi A, Germing U, Le Beau M, Porwit A, Baumann I, et al. Myelodysplastic syndromes/neoplasms, overview. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Classification of tumours oh heamatopoietic and lymphoid tissues, 4th Ed. Lyon, France: IARC; 2008. p. 88-93.
30. Mughal TI, Cross NC, Padron E, Tiu RV, Savona M, Malcovati L, et al. An International MDS/MPN Working Group's perspective and recommendations on molecular pathogenesis, diagnosis and clinical characterization of myelodysplastic/myeloproliferative neoplasms. Haematologica. 2015;100(9):1117-30.
31. Vardiman J, Bennett J, Bain B, Baumann I, Thiele J, Orazi A. Myelodysplastic/myeloproliferative neoplasms, unclassifiable. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of tumours of haematopoietic and lymphoid tissues, 4th ed. Lyon, France: IARC; 2008. p. 85-6.
32. Juneja SK, Imbert M, Sigaux F, Jouault H, Sultan C. Prevalence and distribution of ringed sideroblasts in primary myelodysplastic syndromes. J Clin Pathol. 1983;36(5):566-9.
33. Shaw GR. Ringed sideroblasts with thrombocytosis: an uncommon mixed myelodysplastic/myeloproliferative disease of older adults. Br J Haematol. 2005;131(2):180-4.
34. Wardrop D, Steensma DP. Is refractory anaemia with ring sideroblasts and thrombocytosis (RARS-T) a necessary or useful diagnostic category? Br J Haematol. 2009;144(6):809-17.
35. Malcovati L, Germing U, Kuendgen A, Della Porta MG, Pascutto C, Invernizzi R, et al. Time-dependent prognostic scoring system for predicting survival and leukemic evolution in myelodysplastic syndromes. J Clin Oncol. 2007;25(23):3503-10.

36. Hasserjian R, Gatterman N, Bennett J, Brunning R, Thiele J. Refractory anemia with ring sideroblasts. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of tumours of haematopoietic and lymphoid tissue, 4th edn. Lyon, France: IARC; 2008. p. 96-7.
37. Cazzola M, Invernizzi R, Bergamaschi G, Levi S, Corsi B, Travaglino E, et al. Mitochondrial ferritin expression in erythroid cells from patients with sideroblastic anemia. Blood. 2003;101(5):1996-2000.
38. Cazzola M, Invernizzi R. Ring sideroblasts and sideroblastic anemias. Haematologica. 2011;96(6):789-92.
39. 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. 2002;87(4):392-9.
40. Germing U, Gattermann N, Strupp C, Aivado M, Aul C. Validation of the WHO proposals for a new classification of primary myelodysplastic syndromes: a retrospective analysis of 1600 patients. Leuk Res. 2000;24(12):983-92.
41. 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. 2012;97(7):1036-41.
42. Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al. Frequent pathway mutations of splicing machinery in myelodysplasia. Nature. 2011;478(7367):64-9.
43. Papaemmanuil E, Cazzola M, Boultwood J, Malcovati L, Vyas P, Bowen D, et al. Somatic SF3B1 mutation in myelodysplasia with ring sideroblasts. N Engl J Med. 2011;365(15):1384-95.
44. Damm F, Thol F, Kosmider O, Kade S, Löffeld P, Dreyfus F, et al. SF3B1 mutations in myelodysplastic syndromes: clinical associations and prognostic implications. Leukemia. 2012;26(5):1137-40.
45. Damm F, Kosmider O, Gelsi-Boyer V, Renneville A, Carbuccia N, Hidalgo-Curtis C, et al. Mutations affecting mRNA splicing define distinct clinical phenotypes and correlate with patient outcome in myelodysplastic syndromes. Blood. 2012;119(14):3211-8.
46. Malcovati L, Papaemmanuil E, Bowen DT, Boultwood J, Della Porta MG, Pascutto C, et al. Clinical significance of SF3B1 mutations in myelodysplastic syndromes and myelodysplastic/myeloproliferative neoplasms. Blood. 2011;118(24):6239-46.
47. Visconte V, Makishima H, Jankowska A, Szpurka H, Traina F, Jerez A, et al. SF3B1, a splicing factor is frequently mutated in refractory anemia with ring sideroblasts. Leukemia. 2012;26(3):542-5.
48. Visconte V, Makishima H, Maciejewski JP, Tiu RV. Emerging roles of the spliceosomal machinery in myelodysplastic syndromes and other hematological disorders. Leukemia. 2012;26(12):2447-54.
49. Cazzola M, Rossi M, Malcovati L, Mieloproliferative AIplRsCGIM. Biologic and clinical significance of somatic mutations of SF3B1 in myeloid and lymphoid neoplasms. Blood. 2013;121(2):260-9.
50. Visconte V, Rogers HJ, Singh J, Barnard J, Bupathi M, Traina F, et al. SF3B1 haploinsufficiency leads to formation of ring sideroblasts in myelodysplastic syndromes. Blood. 2012;120(16):3173-86.
51. Patnaik MM, Hanson CA, Sulai NH, Hodnefield JM, Knudson RA, Ketterling RP, et al. Prognostic irrelevance of ring sideroblast percentage in World Health Organization-defined myelodysplastic syndromes without excess blasts. Blood. 2012;119(24):5674-7.
52. Patnaik MM, Lasho TL, Hodnefield JM, Knudson RA, Ketterling RP, Garcia-Manero G, et al. SF3B1 mutations are prevalent in myelodysplastic syndromes with ring sideroblasts but do not hold independent prognostic value. Blood. 2012;119(2):569-72.
53. Küppers R. Mechanisms of B-cell lymphoma pathogenesis. Nat Rev Cancer. 2005;5(4):251-62.
54. Sagaert X, Sprangers B, De Wolf-Peeters C. The dynamics of the B follicle: understanding the normal counterpart of B-cell-derived malignancies. Leukemia. 2007;21(7):1378-86.
55. Lenz G, Staudt LM. Aggressive lymphomas. N Engl J Med. 2010;362(15):1417-29.
56. Nogai H, Dörken B, Lenz G. Pathogenesis of non-Hodgkin's lymphoma. J Clin Oncol. 2011;29(14):1803-11.
57. Stratton MR, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009;458(7239):719-24.

58. Stein H, Warnke R, Chan W, Jaffe E, Chan J, Gatter K, et al. Diffuse large B-cell lymphoma, not otherwise specified. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of haematopoietic and lymphoid tissues, 4th Ed. Lyon, France: IARC; 2008. p. 233-7.
59. Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al. Mature B-cell Neoplasms. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, et al., editors. World Health Organization Classification of haematopoietic and lymphoid tissues, 4th Ed2008. p. 179-268.
60. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2013. CA Cancer J Clin. 2013;63(1):11-30.
61. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, et al. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000;403(6769):503-11.
62. Savage KJ, Monti S, Kutok JL, Cattoretti G, Neuberg D, De Leval L, et al. The molecular signature of mediastinal large B-cell lymphoma differs from that of other diffuse large B-cell lymphomas and shares features with classical Hodgkin lymphoma. Blood. 2003;102(12):3871-9.
63. Carbone PP, Kaplan HS, Musshoff K, Smithers DW, Tubiana M. Report of the Committee on Hodgkin's Disease Staging Classification. Cancer Res. 1971;31(11):1860-1.
64. A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. N Engl J Med. 1993;329(14):987-94.
65. Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, et al. The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. N Engl J Med. 2002;346(25):1937-47.
66. Shipp MA, Ross KN, Tamayo P, Weng AP, Kutok JL, Aguiar RC, et al. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. Nat Med. 2002;8(1):68-74.
67. Monti S, Savage KJ, Kutok JL, Feuerhake F, Kurtin P, Mihm M, et al. Molecular profiling of diffuse large B-cell lymphoma identifies robust subtypes including one characterized by host inflammatory response. Blood. 2005;105(5):1851-61.
68. Lenz G, Wright G, Dave SS, Xiao W, Powell J, Zhao H, et al. Stromal gene signatures in large-B-cell lymphomas. N Engl J Med. 2008;359(22):2313-23.
69. Hans CP, Weisenburger DD, Greiner TC, Gascoyne RD, Delabie J, Ott G, et al. Confirmation of the molecular classification of diffuse large B-cell lymphoma by immunohistochemistry using a tissue microarray. Blood. 2004;103(1):275-82.
70. Muris JJ, Meijer CJ, Vos W, van Krieken JH, Jiwa NM, Ossenkoppele GJ, et al. Immunohistochemical profiling based on Bcl-2, CD10 and MUM1 expression improves risk stratification in patients with primary nodal diffuse large B cell lymphoma. J Pathol. 2006;208(5):714-23.
71. Malumbres R, Chen J, Tibshirani R, Johnson NA, Sehn LH, Natkunam Y, et al. Paraffin-based 6-gene model predicts outcome in diffuse large B-cell lymphoma patients treated with R-CHOP. Blood. 2008;111(12):5509-14.
72. Nyman H, Jerkeman M, Karjalainen-Lindsberg ML, Banham AH, Leppä S. Prognostic impact of activated B-cell focused classification in diffuse large B-cell lymphoma patients treated with R-CHOP. Mod Pathol. 2009;22(8):1094-101.
73. Choi WW, Weisenburger DD, Greiner TC, Piris MA, Banham AH, Delabie J, et al. A new immunostain algorithm classifies diffuse large B-cell lymphoma into molecular subtypes with high accuracy. Clin Cancer Res. 2009;15(17):5494-502.
74. Green TM, Young KH, Visco C, Xu-Monette ZY, Orazi A, Go RS, et al. Immunohistochemical double-hit score is a strong predictor of outcome in patients with diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol. 2012;30(28):3460-7.
75. Copie-Bergman C, Gaulard P, Leroy K, Briere J, Baia M, Jais JP, et al. Immuno-fluorescence in situ hybridization index predicts survival in patients with diffuse large B-cell lymphoma treated with R-CHOP: a GELA study. J Clin Oncol. 2009;27(33):5573-9.

76. Fisher RI, Gaynor ER, Dahlberg S, Oken MM, Grogan TM, Mize EM, et al. Comparison of a standard regimen (CHOP) with three intensive chemotherapy regimens for advanced non-Hodgkin's lymphoma. N Engl J Med. 1993;328(14):1002-6.
77. Pfreundschuh M, Trümper L, Kloess M, Schmits R, Feller AC, Rudolph C, et al. Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of young patients with good-prognosis (normal LDH) aggressive lymphomas: results of the NHL-B1 trial of the DSHNHL. Blood. 2004;104(3):626-33.
78. Pfreundschuh M, Trümper L, Kloess M, Schmits R, Feller AC, Rübe C, et al. Two-weekly or 3-weekly CHOP chemotherapy with or without etoposide for the treatment of elderly patients with aggressive lymphomas: results of the NHL-B2 trial of the DSHNHL. Blood. 2004;104(3):634-41.
79. Coiffier B, Lepage E, Briere J, Herbrecht R, Tilly H, Bouabdallah R, et al. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. N Engl J Med. 2002;346(4):235-42.
80. Feugier P, Van Hoof A, Sebban C, Solal-Celigny P, Bouabdallah R, Fermé C, et al. Long-term results of the R-CHOP study in the treatment of elderly patients with diffuse large B-cell lymphoma: a study by the Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol. 2005;23(18):4117-26.
81. Fu K, Weisenburger DD, Choi WW, Perry KD, Smith LM, Shi X, et al. Addition of rituximab to standard chemotherapy improves the survival of both the germinal center B-cell-like and non-germinal center B-cell-like subtypes of diffuse large B-cell lymphoma. J Clin Oncol. 2008;26(28):4587-94.
82. Cunningham D, Hawkes EA, Jack A, Qian W, Smith P, Mouncey P, et al. Rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisolone in patients with newly diagnosed diffuse large B-cell non-Hodgkin lymphoma: a phase 3 comparison of dose intensification with 14-day versus 21-day cycles. Lancet. 2013;381(9880):1817-26.
83. Delarue R, Tilly H, Mounier N, Petrella T, Salles G, Thieblemont C, et al. Dose-dense rituximab-CHOP compared with standard rituximab-CHOP in elderly patients with diffuse large B-cell lymphoma (the LNH03-6B study): a randomised phase 3 trial. Lancet Oncol. 2013;14(6):525-33.
84. Thieblemont C, Gisselbrecht C. Treatment strategies for diffuse large B-cell lymphoma. Leuk Lymphoma. 2010;18(3):57-69.
85. Reyes F, Lepage E, Ganem G, Molina TJ, Brice P, Coiffier B, et al. ACVBP versus CHOP plus radiotherapy for localized aggressive lymphoma. N Engl J Med. 2005;352(12):1197-205.
86. Pfreundschuh M, Trümper L, Osterborg A, Pettengell R, Trneny M, Imrie K, et al. CHOP-like chemotherapy plus rituximab versus CHOP-like chemotherapy alone in young patients with good-prognosis diffuse large-B-cell lymphoma: a randomised controlled trial by the MabThera International Trial (MInT) Group. Lancet Oncol. 2006;7(5):379-91.
87. Phan J, Mazloom A, Medeiros LJ, Zreik TG, Wogan C, Shihadeh F, et al. Benefit of consolidative radiation therapy in patients with diffuse large B-cell lymphoma treated with R-CHOP chemotherapy. J Clin Oncol. 2010;28(27):4170-6.
88. Récher C, Coiffier B, Haioun C, Molina TJ, Fermé C, Casasnovas O, et al. Intensified chemotherapy with ACVBP plus rituximab versus standard CHOP plus rituximab for the treatment of diffuse large B-cell lymphoma (LNH03-2B): an open-label randomised phase 3 trial. Lancet. 2011;378(9806):1858-67.
89. Sehn LH, Berry B, Chhanabhai M, Fitzgerald C, Gill K, Hoskins P, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. Blood. 2007;109(5):1857-61.
90. Tarella C, Zanni M, Di Nicola M, Patti C, Calvi R, Pescarollo A, et al. Prolonged survival in poor-risk diffuse large B-cell lymphoma following front-line treatment with rituximab-supplemented, early-intensified chemotherapy with multiple autologous hematopoietic stem cell support: a multicenter study by GITIL (Gruppo Italiano Terapie Innovative nei Linfomi). Leukemia. 2007;21(8):1802-11.
91. Bonnet C, Fillet G, Mounier N, Ganem G, Molina TJ, Thiéblemont C, et al. CHOP alone compared with CHOP plus radiotherapy for localized aggressive lymphoma in elderly patients: a study by the Groupe d'Etude des Lymphomes de l'Adulte. J Clin Oncol. 2007;25(7):787-92.

92. Pfreundschuh M, Schubert J, Ziepert M, Schmits R, Mohren M, Lengfelder E, et al. Six versus eight cycles of bi-weekly CHOP-14 with or without rituximab in elderly patients with aggressive CD20+ B-cell lymphomas: a randomised controlled trial (RICOVER-60). Lancet Oncol. 2008;9(2):105-16.
93. Thieblemont C, Gisselbrecht C. Second-line treatment paradigms for diffuse large B-cell lymphomas. Curr Oncol Rep. 2009;11(5):386-93.
94. Philip T, Guglielmi C, Hagenbeek A, Somers R, Van der Lelie H, Bron D, et al. Autologous bone marrow transplantation as compared with salvage chemotherapy in relapses of chemotherapy-sensitive non-Hodgkin's lymphoma. N Engl J Med. 1995;333(23):1540-5.
95. Guglielmi C, Gomez F, Philip T, Hagenbeek A, Martelli M, Sebban C, et al. Time to relapse has prognostic value in patients with aggressive lymphoma enrolled onto the Parma trial. J Clin Oncol. 1998;16(10):3264-9.
96. Lazarus HM, Zhang MJ, Carreras J, Hayes-Lattin BM, Ataergin AS, Bitran JD, et al. A comparison of HLA-identical sibling allogeneic versus autologous transplantation for diffuse large B cell lymphoma: a report from the CIBMTR. Biol Blood Marrow Transplant. 2010;16(1):35-45.
97. van Kampen RJ, Canals C, Schouten HC, Nagler A, Thomson KJ, Vernant JP, et al. Allogeneic stem-cell transplantation as salvage therapy for patients with diffuse large B-cell non-Hodgkin's lymphoma relapsing after an autologous stem-cell transplantation: an analysis of the European Group for Blood and Marrow Transplantation Registry. J Clin Oncol. 2011;29(10):1342-8.
98. Kewalramani T, Zelenetz AD, Nimer SD, Portlock C, Straus D, Noy A, et al. Rituximab and ICE as second-line therapy before autologous stem cell transplantation for relapsed or primary refractory diffuse large B-cell lymphoma. Blood. 2004;103(10):3684-8.
99. Vellenga E, van Putten WL, van 't Veer MB, Zijlstra JM, Fibbe WE, van Oers MH, et al. Rituximab improves the treatment results of DHAP-VIM-DHAP and ASCT in relapsed/progressive aggressive CD20+ NHL: a prospective randomized HOVON trial. Blood. 2008;111(2):537-43.
100. Roschewski M, Staudt LM, Wilson WH. Diffuse large B-cell lymphoma-treatment approaches in the molecular era. Nat Rev Clin Oncol. 2014;11(1):12-23.
101. Iqbal J, Sanger WG, Horsman DE, Rosenwald A, Pickering DL, Dave B, et al. BCL2 translocation defines a unique tumor subset within the germinal center B-cell-like diffuse large B-cell lymphoma. Am J Pathol. 2004;165(1):159-66.
102. Schuetz JM, Johnson NA, Morin RD, Scott DW, Tan K, Ben-Nierah S, et al. BCL2 mutations in diffuse large B-cell lymphoma. Leukemia. 2012;26(6):1383-90.
103. Iqbal J, Neppalli VT, Wright G, Dave BJ, Horsman DE, Rosenwald A, et al. BCL2 expression is a prognostic marker for the activated B-cell-like type of diffuse large B-cell lymphoma. J Clin Oncol. 2006;24(6):961-8.
104. Saito M, Novak U, Piovan E, Basso K, Sumazin P, Schneider C, et al. BCL6 suppression of BCL2 via Miz1 and its disruption in diffuse large B cell lymphoma. Proc Natl Acad Sci U S A. 2009;106(27):11294-9.
105. Basso K, Dalla-Favera R. BCL6: master regulator of the germinal center reaction and key oncogene in B cell lymphomagenesis. Adv Immunol. 2010;105:193-210.
106. Pasqualucci L. The genetic basis of diffuse large B-cell lymphoma. Curr Opin Hematol. 2013;20(4):336-44.
107. Iqbal J, Greiner TC, Patel K, Dave BJ, Smith L, Ji J, et al. Distinctive patterns of BCL6 molecular alterations and their functional consequences in different subgroups of diffuse large B-cell lymphoma. Leukemia. 2007;21(11):2332-43.
108. Lenz G, Wright GW, Emre NC, Kohlhammer H, Dave SS, Davis RE, et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. Proc Natl Acad Sci U S A. 2008;105(36):13520-5.
109. Pfeifer M, Grau M, Lenze D, Wenzel SS, Wolf A, Wollert-Wulf B, et al. PTEN loss defines a PI3K/AKT pathway-dependent germinal center subtype of diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A. 2013;110(30):12420-5.

110. Zhang J, Grubor V, Love CL, Banerjee A, Richards KL, Mieczkowski PA, et al. Genetic heterogeneity of diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A. 2013;110(4):1398-403.
111. Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R, et al. Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. Genes Dev. 2004;18(13):1592-605.
112. Yap DB, Chu J, Berg T, Schapira M, Cheng SW, Moradian A, et al. Somatic mutations at EZH2 Y641 act dominantly through a mechanism of selectively altered PRC2 catalytic activity, to increase H3K27 trimethylation. Blood. 2011;117(8):2451-9.
113. Béguelin W, Popovic R, Teater M, Jiang Y, Bunting KL, Rosen M, et al. EZH2 is required for germinal center formation and somatic EZH2 mutations promote lymphoid transformation. Cancer Cell. 2013;23(5):677-92.
114. Velichutina I, Shaknovich R, Geng H, Johnson NA, Gascoyne RD, Melnick AM, et al. EZH2-mediated epigenetic silencing in germinal center B cells contributes to proliferation and lymphomagenesis. Blood. 2010;116(24):5247-55.
115. Morin RD, Johnson NA, Severson TM, Mungall AJ, An J, Goya R, et al. Somatic mutations altering EZH2 (Tyr641) in follicular and diffuse large B-cell lymphomas of germinal-center origin. Nat Genet. 2010;42(2):181-5.
116. Pasqualucci L, Dominguez-Sola D, Chiarenza A, Fabbri G, Grunn A, Trifonov V, et al. Inactivating mutations of acetyltransferase genes in B-cell lymphoma. Nature. 2011;471(7337):189-95.
117. Morin RD, Mendez-Lago M, Mungall AJ, Goya R, Mungall KL, Corbett RD, et al. Frequent mutation of histone-modifying genes in non-Hodgkin lymphoma. Nature. 2011;476(7360):298-303.
118. Morin RD, Mungall K, Pleasance E, Mungall AJ, Goya R, Huff RD, et al. Mutational and structural analysis of diffuse large B-cell lymphoma using whole-genome sequencing. Blood. 2013;122(7):1256-65.
119. Lim KH, Yang Y, Staudt LM. Pathogenetic importance and therapeutic implications of NF-κB in lymphoid malignancies. Immunol Rev. 2012;246(1):359-78.
120. Rossi D, Ciardullo C, Gaidano G. Genetic aberrations of signaling pathways in lymphomagenesis: revelations from next generation sequencing studies. Semin Cancer Biol. 2013;23(6):422-30.
121. Davis RE, Brown KD, Siebenlist U, Staudt LM. Constitutive nuclear factor kappaB activity is required for survival of activated B cell-like diffuse large B cell lymphoma cells. J Exp Med. 2001;194(12):1861-74.
122. Compagno M, Lim WK, Grunn A, Nandula SV, Brahmachary M, Shen Q, et al. Mutations of multiple genes cause deregulation of NF-kappaB in diffuse large B-cell lymphoma. Nature. 2009;459(7247):717-21.
123. Davis RE, Ngo VN, Lenz G, Tolar P, Young RM, Romesser PB, et al. Chronic active B-cell-receptor signalling in diffuse large B-cell lymphoma. Nature. 2010;463(7277):88-92.
124. Gauld SB, Dal Porto JM, Cambier JC. B cell antigen receptor signaling: roles in cell development and disease. Science. 2002;296(5573):1641-2.
125. Srinivasan L, Sasaki Y, Calado DP, Zhang B, Paik JH, DePinho RA, et al. PI3 kinase signals BCR-dependent mature B cell survival. Cell. 2009;139(3):573-86.
126. Lenz G, Davis RE, Ngo VN, Lam L, George TC, Wright GW, et al. Oncogenic CARD11 mutations in human diffuse large B cell lymphoma. Science. 2008;319(5870):1676-9.
127. Hailfinger S, Lenz G, Ngo V, Posvitz-Fejfar A, Rebeaud F, Guzzardi M, et al. Essential role of MALT1 protease activity in activated B cell-like diffuse large B-cell lymphoma. Proc Natl Acad Sci U S A. 2009;106(47):19946-51.
128. Ngo VN, Young RM, Schmitz R, Jhavar S, Xiao W, Lim KH, et al. Oncogenically active MYD88 mutations in human lymphoma. Nature. 2011;470(7332):115-9.
129. Lam LT, Wright G, Davis RE, Lenz G, Farinha P, Dang L, et al. Cooperative signaling through the signal transducer and activator of transcription 3 and nuclear factor-{kappa}B pathways in subtypes of diffuse large B-cell lymphoma. Blood. 2008;111(7):3701-13.
130. Kato M, Sanada M, Kato I, Sato Y, Takita J, Takeuchi K, et al. Frequent inactivation of A20 in B-cell lymphomas. Nature. 2009;459(7247):712-6.

131. Ranganathan P, Weaver KL, Capobianco AJ. Notch signalling in solid tumours: a little bit of everything but not all the time. Nat Rev Cancer. 2011;11(5):338-51.
132. Lohr JG, Stojanov P, Lawrence MS, Auclair D, Chapuy B, Sougnez C, et al. Discovery and prioritization of somatic mutations in diffuse large B-cell lymphoma (DLBCL) by whole-exome sequencing. Proc Natl Acad Sci U S A. 2012;109(10):3879-84.
133. Rossi D, Trifonov V, Fangazio M, Bruscaggin A, Rasi S, Spina V, et al. The coding genome of splenic marginal zone lymphoma: activation of NOTCH2 and other pathways regulating marginal zone development. J Exp Med. 2012;209(9):1537-51.
134. Lee SY, Kumano K, Nakazaki K, Sanada M, Matsumoto A, Yamamoto G, et al. Gain-of-function mutations and copy number increases of Notch2 in diffuse large B-cell lymphoma. Cancer Sci. 2009;100(5):920-6.
135. Sivertsen EA, Galteland E, Mu D, Holte H, Meza-Zepeda L, Myklebost O, et al. Gain of chromosome 6p is an infrequent cause of increased PIM1 expression in B-cell non-Hodgkin's lymphomas. Leukemia. 2006;20(3):539-42.
136. Pasqualucci L, Trifonov V, Fabbri G, Ma J, Rossi D, Chiarenza A, et al. Analysis of the coding genome of diffuse large B-cell lymphoma. Nat Genet. 2011;43(9):830-7.
137. Johnson NA, Slack GW, Savage KJ, Connors JM, Ben-Neriah S, Rogic S, et al. Concurrent expression of MYC and BCL2 in diffuse large B-cell lymphoma treated with rituximab plus cyclophosphamide, doxorubicin, vincristine, and prednisone. J Clin Oncol. 2012;30(28):3452-9.
138. Dave SS, Fu K, Wright GW, Lam LT, Kluin P, Boerma EJ, et al. Molecular diagnosis of Burkitt's lymphoma. N Engl J Med. 2006;354(23):2431-42.
139. Klapper W, Stoecklein H, Zeynalova S, Ott G, Kosari F, Rosenwald A, et al. Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). Leukemia. 2008;22(12):2226-9.
140. Savage KJ, Johnson NA, Ben-Neriah S, Connors JM, Sehn LH, Farinha P, et al. MYC gene rearrangements are associated with a poor prognosis in diffuse large B-cell lymphoma patients treated with R-CHOP chemotherapy. Blood. 2009;114(17):3533-7.
141. Nitsu N, Okamoto M, Miura I, Hirano M. Clinical significance of 8q24/c-MYC translocation in diffuse large B-cell lymphoma. Cancer Sci. 2009;100(2):233-7.
142. Barrans S, Crouch S, Smith A, Turner K, Owen R, Patmore R, et al. Rearrangement of MYC is associated with poor prognosis in patients with diffuse large B-cell lymphoma treated in the era of rituximab. J Clin Oncol. 2010;28(20):3360-5.
143. Horn H, Ziepert M, Becher C, Barth TF, Bernd HW, Feller AC, et al. MYC status in concert with BCL2 and BCL6 expression predicts outcome in diffuse large B-cell lymphoma. Blood. 2013;121(12):2253-63.
144. Cuccuini W, Briere J, Mounier N, Voelker HU, Rosenwald A, Sundstrom C, et al. MYC+ diffuse large B-cell lymphoma is not salvaged by classical R-ICE or R-DHAP followed by BEAM plus autologous stem cell transplantation. Blood. 2012;119(20):4619-24.
145. Leucci E, Cocco M, Onnis A, De Falco G, van Cleef P, Bellan C, et al. MYC translocation-negative classical Burkitt lymphoma cases: an alternative pathogenetic mechanism involving miRNA deregulation. J Pathol. 2008;216(4):440-50.
146. Stasik CJ, Nitta H, Zhang W, Mosher CH, Cook JR, Tubbs RR, et al. Increased MYC gene copy number correlates with increased mRNA levels in diffuse large B-cell lymphoma. Haematologica. 2010;95(4):597-603.
147. Green TM, Nielsen O, de Stricker K, Xu-Monette ZY, Young KH, Møller MB. High levels of nuclear MYC protein predict the presence of MYC rearrangement in diffuse large B-cell lymphoma. Am J Surg Pathol. 2012;36(4):612-9.
148. Thieblemont C, Brière J. MYC, BCL2, BCL6 in DLBCL: impact for clinics in the future? Blood. 2013;121(12):2165-6.

149. Xu-Monette ZY, Medeiros LJ, Li Y, Orlowski RZ, Andreeff M, Bueso-Ramos CE, et al. Dysfunction of the TP53 tumor suppressor gene in lymphoid malignancies. Blood. 2012;119(16):3668-83.
150. Challa-Malladi M, Lieu YK, Califano O, Holmes AB, Bhagat G, Murty VV, et al. Combined genetic inactivation of β2-Microglobulin and CD58 reveals frequent escape from immune recognition in diffuse large B cell lymphoma. Cancer Cell. 2011;20(6):728-40.
151. Steidl C, Shah SP, Woolcock BW, Rui L, Kawahara M, Farinha P, et al. MHC class II transactivator CIITA is a recurrent gene fusion partner in lymphoid cancers. Nature. 2011;471(7338):377-81.
152. Friedberg JW. Relapsed/refractory diffuse large B-cell lymphoma. Hematology Am Soc Hematol Educ Program. 2011;2011:498-505.
153. Gisselbrecht C, Glass B, Mounier N, Singh Gill D, Linch DC, Trneny M, et al. Salvage regimens with autologous transplantation for relapsed large B-cell lymphoma in the rituximab era. J Clin Oncol. 2010;28(27):4184-90.
154. Thieblemont C, Briere J, Mounier N, Voelker HU, Cuccuini W, Hirchaud E, et al. The germinal center/activated B-cell subclassification has a prognostic impact for response to salvage therapy in relapsed/refractory diffuse large B-cell lymphoma: a bio-CORAL study. J Clin Oncol. 2011;29(31):4079-87.
155. Baron D, Raharijaona M, Houlgatte R. Puces à ADN : Revue générale. ITBM-RBM. 2007;28:210-5.
156. Bogard M, Ameziane N, Lamoril J. Microarray d'ADN et profils d'expression des gènes. Première partie : concept, fabrication et mise en œuvre. Immuno-analyse et biologie spécialisée. 2008;23:71-88.
157. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res. 2002;30(4):e15.
158. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998;95(25):14863-8.
159. Chen G, Broséus J, Hergalant S, Donnart A, Chevalier C, Bolaños-Jiménez F, et al. Identification of master genes involved in liver key functions through transcriptomics and epigenomics of methyl donor deficiency in rat: relevance to nonalcoholic liver disease. Mol Nutr Food Res. 2015;59(2):293-302.
160. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000;25(1):25-9.
161. Zeeberg BR, Qin H, Narasimhan S, Sunshine M, Cao H, Kane DW, et al. High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). BMC Bioinformatics. 2005;6:168.
162. McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, et al. Integrated detection and population-genetic analysis of SNPs and copy number variation. Nat Genet. 2008;40(10):1166-74.
163. Scharpf RB, Irizarry RA, Ritchie ME, Carvalho B, Ruczinski I. Using the R Package crlmm for Genotyping and Copy Number Estimation. J Stat Softw. 2011;40(12):1-32.
164. Jeromin S, Haferlach T, Grossmann V, Alpermann T, Kowarsch A, Haferlach C, et al. High frequencies of SF3B1 and JAK2 mutations in refractory anemia with ring sideroblasts associated with marked thrombocytosis strengthen the assignment to the category of myelodysplastic/myeloproliferative neoplasms. Haematologica. 2013;98(2):e15-7.
165. Schnittger S, Bacher U, Haferlach C, Geer T, Müller P, Mittermüller J, et al. Detection of JAK2 exon 12 mutations in 15 patients with JAK2V617F negative polycythemia vera. Haematologica. 2009;94(3):414-8.
166. Schnittger S, Bacher U, Haferlach C, Beelen D, Bojko P, Bürkle D, et al. Characterization of 35 new cases with four different MPLW515 mutations and essential thrombocytosis or primary myelofibrosis. Haematologica. 2009;94(1):141-4.
167. Broséus J, Alpermann T, Wulfert M, Florensa Brichs L, Jeromin S, Lippert E, et al. 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):1826-31.

168. Tefferi A, Lasho TL, Finke CM, Knudson RA, Ketterling R, Hanson CH, et al. CALR vs JAK2 vs MPL mutated or triple-negative myelofibrosis: clinical, cytogenetic and molecular comparisons. Leukemia. 2014.
169. 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. 2014;28(6):1374-6.
170. Carillo S, Henry L, Lippert E, Girodon F, Guiraud I, Richard C, et al. Nested high-resolution melting curve analysis a highly sensitive, reliable, and simple method for detection of JAK2 exon 12 mutations--clinical relevance in the monitoring of polycythemia. J Mol Diagn. 2011;13(3):263-70.
171. Boyd EM, Bench AJ, Goday-Fernández A, Anand S, Vaghela KJ, Beer P, et al. Clinical utility of routine MPL exon 10 analysis in the diagnosis of essential thrombocythaemia and primary myelofibrosis. Br J Haematol. 2010;149(2):250-7.
172. Lundberg P, Karow A, Nienhold R, Looser R, Hao-Shen H, Nissen I, et al. Clonal evolution and clinical correlates of somatic mutations in myeloproliferative neoplasms. Blood. 2014;123(14):2220-8.
173. Sehn LH, Gascoyne RD. Diffuse large B-cell lymphoma: optimizing outcome in the context of clinical and biologic heterogeneity. Blood. 2015;125(1):22-32.
174. Novak AJ, Asmann YW, Maurer MJ, Wang C, Slager SL, Hodge LS, et al. Whole-exome analysis reveals novel somatic genomic alterations associated with outcome in immunochemotherapy-treated diffuse large B-cell lymphoma. Blood Cancer J. 2015;5:e346.
175. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A. 2001;98(9):5116-21.
176. Broséus J, Park JH, Carillo S, Hermouet S, Girodon F. Presence of calreticulin mutations in JAK2-negative polycythemia vera. Blood. 2014;124(26):3964-6.
177. 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. 2014.
178. Pietra D, Rumi E, Ferretti VV, Buduo CA, Milanesi C, Cavalloni C, et al. Differential clinical effects of different mutation subtypes in CALR-mutant myeloproliferative neoplasms. Leukemia. 2016;30(2):431-8.
179. Marty C, Pecquet C, Nivarthi H, El-Khoury M, Chachoua I, Tulliez M, et al. Calreticulin mutants in mice induce an MPL-dependent thrombocytosis with frequent progression to myelofibrosis. Blood. 2016;127(10):1317-24.
180. Araki M, Yang Y, Masubuchi N, Hironaka Y, Takei H, Morishita S, et al. Activation of the thrombopoietin receptor by mutant calreticulin in CALR-mutant myeloproliferative neoplasms. Blood. 2016;127(10):1307-16.
181. Chachoua I, Pecquet C, El-Khoury M, Nivarthi H, Albu RI, Marty C, et al. Thrombopoietin receptor activation by myeloproliferative neoplasm associated calreticulin mutants. Blood. 2016;127(10):1325-35.
182. Cazzola M. Mutant calreticulin: when a chaperone becomes intrusive. Blood. 2016;127(10):1219-21.
183. Broséus J, Chen G, Hergalant S, Ramstein G, Mounier N, Guéant J, et al. Relapsed diffuse large B-cell lymphoma present different genomic profiles between early and late relapses. Oncotarget. 2016;In press.
184. Jardin F, Jais JP, Molina TJ, Parmentier F, Picquenot JM, Ruminy P, et al. Diffuse large B-cell lymphomas with CDKN2A deletion have a distinct gene expression signature and a poor prognosis under R-CHOP treatment: a GELA study. Blood. 2010;116(7):1092-104.
185. Jansen RC, Nap JP. Genetical genomics: the added value from segregation. Trends Genet. 2001;17(7):388-91.
186. Li J, Burmeister M. Genetical genomics: combining genetics with gene expression analysis. Hum Mol Genet. 2005;14 Spec No. 2:R163-9.
187. Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, et al. Genetics of gene expression surveyed in maize, mouse and man. Nature. 2003;422(6929):297-302.

188. Bystrykh L, Weersing E, Dontje B, Sutton S, Pletcher MT, Wiltshire T, et al. Uncovering regulatory pathways that affect hematopoietic stem cell function using 'genetical genomics'. Nat Genet. 2005;37(3):225-32.
189. Hubner N, Wallace CA, Zimdahl H, Petretto E, Schulz H, Maciver F, et al. Integrated transcriptional profiling and linkage analysis for identification of genes underlying disease. Nat Genet. 2005;37(3):243-53.
190. Brem RB, Yvert G, Clinton R, Kruglyak L. Genetic dissection of transcriptional regulation in budding yeast. Science. 2002;296(5568):752-5.
191. Gagneur J, Stegle O, Zhu C, Jakob P, Tekkedil MM, Aiyar RS, et al. Genotype-environment interactions reveal causal pathways that mediate genetic effects on phenotype. PLoS Genet. 2013;9(9):e1003803.
192. Monks SA, Leonardson A, Zhu H, Cundiff P, Pietrusiak P, Edwards S, et al. Genetic inheritance of gene expression in human cell lines. Am J Hum Genet. 2004;75(6):1094-105.
193. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, et al. Genetic analysis of genome-wide variation in human gene expression. Nature. 2004;430(7001):743-7.
194. Müller-Hermelink H, Montserrat E, Catovsky D, Campo E, Harris N, Stein H. Chronic lymphocytic leukemia/small lymphocytic lymphoma. In: Swerdlow S, Campo E, Lee Harris N, Jaffe E, Pileri S, Stein H, editors. WHO classification of tumours of haematopoietic and lymphoid tissues, 4th edn. Lyon, France: IARC; 2008. p. 180-2.
195. Tsimberidou AM, Keating MJ. Richter syndrome: biology, incidence, and therapeutic strategies. Cancer. 2005;103(2):216-28.
196. Tsimberidou AM, O'Brien S, Khouri I, Giles FJ, Kantarjian HM, Champlin R, et al. Clinical outcomes and prognostic factors in patients with Richter's syndrome treated with chemotherapy or chemoimmunotherapy with or without stem-cell transplantation. J Clin Oncol. 2006;24(15):2343-51.
197. Rossi D, Cerri M, Capello D, Deambrogi C, Rossi FM, Zucchetto A, et al. Biological and clinical risk factors of chronic lymphocytic leukaemia transformation to Richter syndrome. Br J Haematol. 2008;142(2):202-15.
198. Parikh SA, Rabe KG, Call TG, Zent CS, Habermann TM, Ding W, et al. Diffuse large B-cell lymphoma (Richter syndrome) in patients with chronic lymphocytic leukaemia (CLL): a cohort study of newly diagnosed patients. Br J Haematol. 2013;162(6):774-82.
199. Rossi D, Spina V, Deambrogi C, Rasi S, Laurenti L, Stamatopoulos K, et al. The genetics of Richter syndrome reveals disease heterogeneity and predicts survival after transformation. Blood. 2011;117(12):3391-401.
200. Rossi D, Spina V, Cerri M, Rasi S, Deambrogi C, De Paoli L, et al. Stereotyped B-cell receptor is an independent risk factor of chronic lymphocytic leukemia transformation to Richter syndrome. Clin Cancer Res. 2009;15(13):4415-22.
201. Jamroziak K, Tadmor T, Robak T, Polliack A. Richter syndrome in chronic lymphocytic leukemia: updates on biology, clinical features and therapy. Leuk Lymphoma. 2015;56(7):1949-58.
202. Parikh SA, Kay NE, Shanafelt TD. How we treat Richter syndrome. Blood. 2014;123(11):1647-57.
203. Langerbeins P, Busch R, Anheier N, Dürig J, Bergmann M, Goebeler ME, et al. Poor efficacy and tolerability of R-CHOP in relapsed/refractory chronic lymphocytic leukemia and Richter transformation. Am J Hematol. 2014;89(12):E239-43.
204. Tsimberidou AM, Wierda WG, Plunkett W, Kurzrock R, O'Brien S, Wen S, et al. Phase I-II study of oxaliplatin, fludarabine, cytarabine, and rituximab combination therapy in patients with Richter's syndrome or fludarabine-refractory chronic lymphocytic leukemia. J Clin Oncol. 2008;26(2):196-203.
205. Cwynarski K, van Biezen A, de Wreede L, Stilgenbauer S, Bunjes D, Metzner B, et al. Autologous and allogeneic stem-cell transplantation for transformed chronic lymphocytic leukemia (Richter's syndrome): A retrospective analysis from the chronic lymphocytic leukemia subcommittee of the chronic leukemia working party and lymphoma working party of the European Group for Blood and Marrow Transplantation. J Clin Oncol. 2012;30(18):2211-7.

206. Rossi D. XIII. Molecular pathogenesis of transformed lymphomas. Hematol Oncol. 2015;33 Suppl 1:70-4.
207. Hallek M, Fischer K, Fingerle-Rowson G, Fink AM, Busch R, Mayer J, et al. Addition of rituximab to fludarabine and cyclophosphamide in patients with chronic lymphocytic leukaemia: a randomised, open-label, phase 3 trial. Lancet. 2010;376(9747):1164-74.
208. Chigrinova E, Rinaldi A, Kwee I, Rossi D, Rancoita PM, Strefford JC, et al. Two main genetic pathways lead to the transformation of chronic lymphocytic leukemia to Richter syndrome. Blood. 2013;122(15):2673-82.
209. Fabbri G, Khiabanian H, Holmes AB, Wang J, Messina M, Mullighan CG, et al. Genetic lesions associated with chronic lymphocytic leukemia transformation to Richter syndrome. J Exp Med. 2013;210(11):2273-88.
210. Ferreira PG, Jares P, Rico D, Gómez-López G, Martínez-Trillos A, Villamor N, et al. Transcriptome characterization by RNA sequencing identifies a major molecular and clinical subdivision in chronic lymphocytic leukemia. Genome Res. 2014;24(2):212-26.

ANNEXE 1

Protocole *Agilent Two-Color Microarray-Based Gene Expression.*

Etape 1 : préparation des tampons Spike A et B

Etape 2 : Synthèse de l'ADN complémentaire (ADNc)

La quantité d'ARN de départ est de 250 ng par échantillon.

Amplification linéaire et marquage par intégration de Cyanine 3 (Cy 3) dCTP ou de Cyanine 5 (Cy 5) dCTP.

Etape 3 : purification de l'ARNc

Purification du cRNA à l'aide d'une « Rneasy spin column » issue du kit Qiagen RNAeasy Mini Kit.

Mesure de la concentration d'ARNc et de l'activité spécifique (en pmol de Cy3 ou 5 par µg d'ARNc) à l'aide d'un spectrophotomètre Nanodrop ND 1000.

Etape 4 : préparation des échantillons à hybrider

Chaque échantillon à hybrider sur la puce se compose de 825 µg d'ARNc d'un premier patient hybridé avec de la Cy3 et de 825 µg d'ARNc d'un deuxième patient, hybridé avec de la Cy5. L'étape de fragmentation se fait selon le protocole Agilent Les échantillons sont ensuite dilués au 1/2 dans du tampon d'hybridation Agilent GE 2x.

Etape 5 : Hybridation

Les échantillons (contenant les ARNc de deux malades différents chacun marqué soit par Cy3 soit par Cy5) sont hybridés sur un des 4 emplacements composant la lame. Chacun de ces emplacements contient un réseau de 44 000 sondes. Les lames sont incubées 17h à 65°C dans un four rotatif.

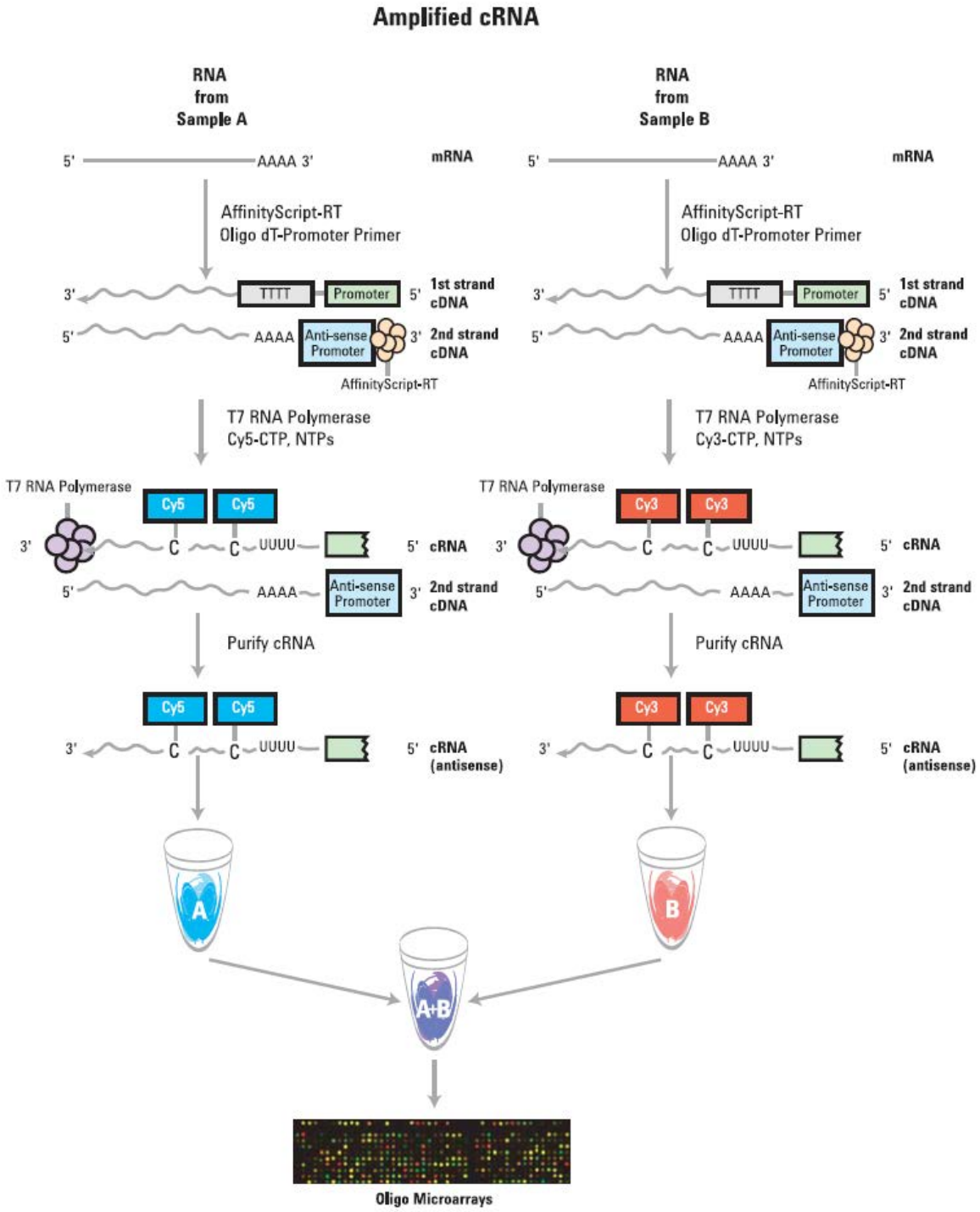
Etape 6 : lavages

Microarrays were subsequently washed and dried with Agilent Wash Buffer (as described in the manufacturer's protocol and already published1

Etape 7 : acquisition des données

Les lames sont analysées par un scanner Agilent G2505B dont la résolution est de 5µm.Les données sont extraites à l'aide du logiciel *Agilent Feature Extraction* version 10.5.1.1.

Les intensités médianes du signal fluorescent sont utilisées pour l'analyse



Protocole Agilent Two-Color Microarray-Based Gene Expression

Source :http://affymetrix.com/support/technical/datasheets

ANNEXE 2

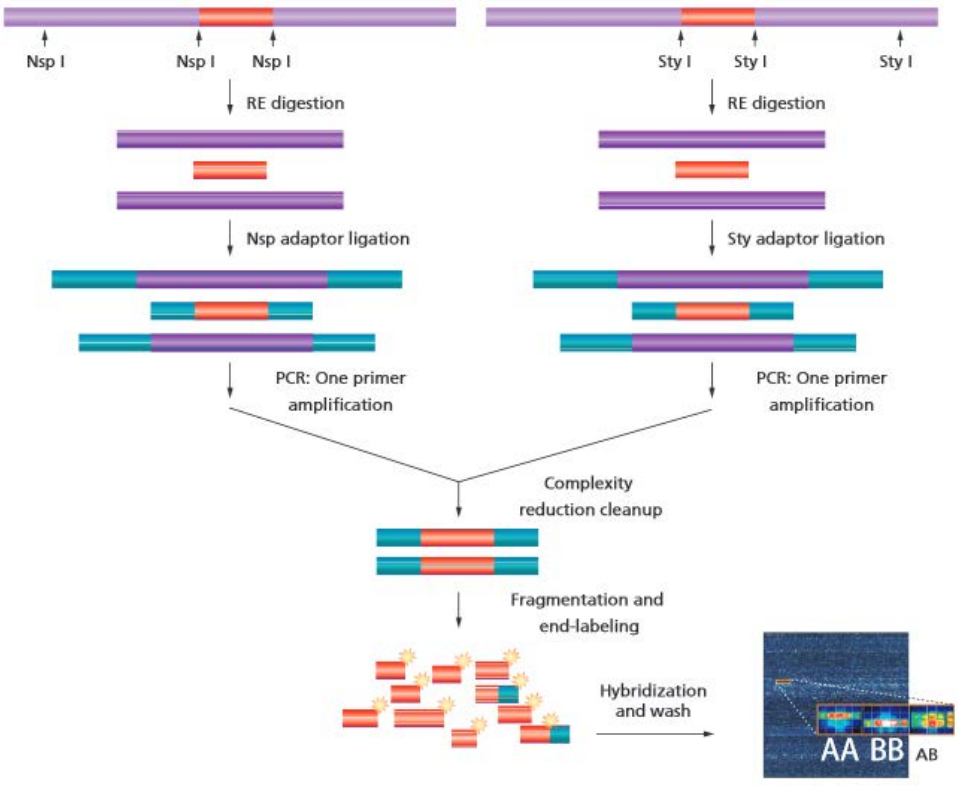
Protocole d'utilisation de la puce Affymetrix SNP 6.0

(i) 500 nanogrammes d'ADN génomique sont digérés séparément avec les enzymes de restriction Nsp I et Sty I.

(ii) Des molécules adaptatrices sont ajoutées aux extrémités de chaque fragment généré par la digestion quelle que soit leur taille.

(iii) Des amorces de PCR (*Polymerase Chain Reaction*) génériques reconnaissant les séquences adaptatrices sont ajoutées puis une réaction de PCR est déclenchée. Les conditions de cette PCR favorisent l'amplification des fragments dont la taille est comprise entre 200 et 1100 paires de bases.

(iv) Les produits de PCR issus de chaque digestion sont mélangés et purifiés à l'aide de billes de polystyrène puis fragmentés, marqués et hybridés sur la puce.



Protocole d'utilisation de la puce *Affymetrix Genome-wide Human SNP Array* 6.0
Source :http://affymetrix.com/support/technical/datasheets/genomewide_snp6_datasheet.pdf.

Résumé:

L'outil génomique a considérablement modifié notre connaissance des hémopathies malignes, que ce soit sur le plan physiopathologique, diagnostique, pronostique ou thérapeutique.

Les syndromes myéloprolifératifs (SMP) et les syndromes mixtes myéloprolifératifs/myélodysplasiques (SMP/SMD) sont des hémopathies malignes caractérisées par des anomalies moléculaires récurrentes. Dans la première partie de ce travail, nous avons travaillé sur une grande cohorte européenne d'anémies réfractaires sidéroblastiques avec thrombocytose (ARS-T), une catégorie rare de SMP/SMD. Nous avons démontré qu'il s'agissait d'une entité indépendante, avec une présentation moléculaire particulière associant (i) des mutations de *SF3B1* dans plus de 85% des cas, expliquant son versant myélodysplasique et (ii) des anomalies de *JAK2*, expliquant son versant prolifératif dans plus de 50% des cas. Les mutations de *CALR*, décrites dans une forte proportion de SMP *JAK2*^{WT}, sont retrouvées dans des proportions trop faibles pour expliquer le caractère myéloprolifératif des ARS-T *JAK2*^{WT}. En revanche, nous avons montré qu'il existait de très rares cas de Polyglobulies de Vaquez (PV) *JAK2*^{WT} présentant une mutation de *CALR*. L'identification de cette mutation à la fois dans les cellules matures et dans leurs précurseurs renforce l'hypothèse d'un rôle physiopathologique des mutations de *CALR* dans ces très rares cas de PV *JAK2*^{WT}. La perspective de cette première partie est d'identifier la ou les mutation(s) responsables du caractère myéloprolifératif dans les ARS-T *JAK2*^{WT}.

Les lymphomes B-diffus à grandes cellules (LBDGC) représentent les lymphomes malins non-Hodgkiniens les plus fréquents chez l'adulte. Près de 40% des LBDGC rechutent après le traitement de première ligne, menant à un décès dans plus de 90% des cas. Le délai de la rechute constitue un élément pronostique majeur, les rechutes précoces (moins d'un an après la fin du traitement de première ligne) présentant un pronostic plus réservé que les rechutes tardives (plus d'un an après la fin du traitement de première ligne). La caractérisation des anomalies génomiques associées à la rechute des LBDGC est au cœur des préoccupations actuelles. Dans la deuxième partie de ce travail, nous avons réalisé l'analyse par *SNP-array* d'une série homogène d'échantillons issus de la cohorte CORAL, une étude prospective internationale portant sur les LBDGC en rechute. Notre série était composée de 39 échantillons ganglionnaires, répartis équitablement entre rechutes précoces et rechutes tardives. Notre objectif était d'identifier les anomalies de nombre de copies (ANC) associées à chacun des deux types de rechutes. L'application d'une méthode statistique originale, inspirée de la méthode SAM (*Significance Analysis of Microarrays*) nous a permis d'identifier 56 gènes et 27 longs ARN non codants dont le profil de répartition des ANC était différentiel entre rechutes précoces et tardives. La classification hiérarchique de ces anomalies nous a permis d'identifier des profils d'anomalies récurrentes dans chacun des deux types de rechutes. Les rechutes précoces sont associées à une forte proportion d'anomalies affectant les régulateurs du cycle cellulaire, de l'apoptose et de la transcription. Les rechutes tardives sont associées à des anomalies affectant les régulateurs de l'immunité et de la prolifération cellulaire. Cette étude permet de mieux comprendre les déterminants de la rechute dans les LBDGC et ouvre de nouvelles perspectives thérapeutiques.

Mots-clés : génomique, syndromes myéloprolifératifs, anémie réfractaire sidéroblastique avec sidéroblastes en couronne et thrombocytose, *CALR,* lymphome B-diffus à grandes cellules, rechute.

Abstract:

Genomics provided new insights in our knowledge of pathophysiology, diagnostic approach, prognosis and therapeutic perspectives in hematological malignancies.

Myeloproliferative neoplasms (MPN) and Myeloproliferative/Myelodysplastic neoplasms (MPN/MDS) are hematological malignancies characterized by recurrent molecular abnormalities. In the first part of this work, we studied a large European cohort of Refractory Anemia with Ring sideroblasts and marked Thrombocytosis (RARS-T), a rare category of MPN/MDS. We demonstrated that RARS-T can be considered as an independent entity, with a specific molecular pattern, associating : (i) *SF3B1* mutations in more than 85% of cases, accounting for its myelodysplastic aspect and (ii) *JAK2* mutations, accounting for its myeloproliferative aspect in more than 50% of cases. *CALR* mutations, described in a large proportion of MPN *JAK2*^{WT}, are very rare in RARS-T, and cannot explain its myeloproliferative part in *JAK2*^{WT} cases. On the other hand, we demonstrated that *CALR* mutations may be identified in very rare cases of Polycythemia Vera (PV) *JAK2*WT. These mutations have been identified in mature blood cells as well as in their precursors, which strengthens the hypothesis of a pathophysiological role of *CALR* mutations in these *JAK2*^{WT} PV. Future prospects of the first part of this work is to identify (the) mutation(s) responsible for the myeloproliferative part of *JAK2*^{WT} RARS-T.

Diffuse large B-cell lymphoma (DLBCL) is the most common subtype of non-Hodgkin lymphoma in adults. Around 40% of DLBCLs relapse after first-line treatment, leading to death in 90% of cases. One of the most important prognostic parameter in this context is the delay of relapse. Late relapse (LR) is defined as relapse occurring more than one year after the initial treatment. Prognosis is worse in early relapse (ER), defined as relapse occurring within 1 year after the initial treatment. Genomic abnormalities associated with relapsed DLBCLs remain largely unknown. In the second part of this work, we performed SNP-array analysis of a homogeneous series of samples from the CORAL cohort, an international prognostic study on relapsed DLBCLs. We gathered 39 lymph node samples equally distributed among ER and LR. Our purpose was to identify Copy Number Variations (CNV) associated ER or LR. We devised an original method based on SAM (Significance Analysis of Microarrays) and identified 56 protein-coding genes and 27 long non-coding RNA with significantly differential CNV distribution between ER and LR DLBCLs. Hierarchical clustering of CNVs allowed the identification of homogeneous clusters of ER and LR samples, with recurrent abnormalities. ER DLBCLs are associated with high rates of CNVs affecting regulators of cell cycle, apoptosis and transcription. In LR DLBCLs, CNVs are related to immune response and cell proliferation. This study provides new insights into the genetic aberrations in relapsed DLBCLs and open up new therapeutic perspectives.

Keywords: genomics, myeloproliferative neoplasms, refractory anemia with ring sideroblasts and marked thrombocytosis, *CALR,* diffuse large B-cell lymphoma, relapse.