



HAL
open science

Identification et implication des gènes DMD et RCBTB1 dans la progression tumorale des sarcomes à génétique complexe

Olivier Mauduit

► **To cite this version:**

Olivier Mauduit. Identification et implication des gènes DMD et RCBTB1 dans la progression tumorale des sarcomes à génétique complexe. Cancer. Université de Lyon, 2017. Français. NNT : 2017LYSE1059 . tel-01710462

HAL Id: tel-01710462

<https://theses.hal.science/tel-01710462>

Submitted on 16 Feb 2018

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

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

N°d'ordre NNT : xxx



THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 340
Biologie Moléculaire Intégrative et Cellulaire

Spécialité de doctorat : Biologie Moléculaire

Discipline : Cancérologie

Soutenue publiquement le 14/04/2017, par :

Olivier MAUDUIT

Identification et implication des gènes *DMD* et *RCBTB1* dans la progression tumorale des sarcomes à génétique complexe

Devant le jury composé de :

Françoise Rédini, Laurent Le Cam,	DR1 à l'université de Nantes DR2 à l'université de Montpellier	Rapporteuse Rapporteur
Marie Castets, Sophie Le Guellec,	CR1 à l'université de Lyon 1 PH à l'université de Toulouse	Examinatrice Examinatrice
Jean-Yves Blay, Frédéric Chibon,	PU-PH au Centre Léon Bérard DR2 à l'institut Bergonié	Directeur de thèse Co-directeur de thèse

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université

Président du Conseil Académique

Vice-président du Conseil d'Administration

Vice-président du Conseil Formation et Vie Universitaire

Vice-président de la Commission Recherche

Directeur Général des Services

M. le Professeur Frédéric FLEURY

M. le Professeur Hamda BEN HADID

M. le Professeur Didier REVEL

M. le Professeur Philippe CHEVALIER

M. Fabrice VALLÉE

M. Alain HELLEU

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard

Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux

Faculté d'Odontologie

Institut des Sciences Pharmaceutiques et Biologiques

Institut des Sciences et Techniques de la Réadaptation

Département de formation et Centre de Recherche en Biologie Humaine

Directeur : M. le Professeur J. ETIENNE

Directeur : Mme la Professeure C. BURILLON

Directeur : M. le Professeur D. BOURGEOIS

Directeur : Mme la Professeure C. VINCIGUERRA

Directeur : M. le Professeur Y. MATILLON

Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies

Département Biologie

Département Chimie Biochimie

Département GEP

Département Informatique

Département Mathématiques

Département Mécanique

Département Physique

UFR Sciences et Techniques des Activités Physiques et Sportives

Observatoire des Sciences de l'Univers de Lyon

Polytech Lyon

Ecole Supérieure de Chimie Physique Electronique

Institut Universitaire de Technologie de Lyon 1

Ecole Supérieure du Professorat et de l'Education

Institut de Science Financière et d'Assurances

Directeur : M. F. DE MARCHI

Directeur : M. le Professeur F. THEVENARD

Directeur : Mme C. FELIX

Directeur : M. Hassan HAMMOURI

Directeur : M. le Professeur S. AKKOUCHE

Directeur : M. le Professeur G. TOMANOV

Directeur : M. le Professeur H. BEN HADID

Directeur : M. le Professeur J-C PLENET

Directeur : M. Y. VANPOULLE

Directeur : M. B. GUIDERDONI

Directeur : M. le Professeur E. PERRIN

Directeur : M. G. PIGNAULT

Directeur : M. le Professeur C. VITON

Directeur : M. le Professeur A. MOUGNIOTTE

Directeur : M. N. LEBOISNE

« Tout travail scientifique est une réalisation communautaire et non pas l'œuvre d'un cavalier seul. Qui sait où Darwin se serait égaré en 1837 sans Gould, sans Owen, et sans la vie scientifique très active de Londres et de Cambridge ? »

Stephen Jay Gould –

Le sourire du flamant rose (1988)

Remerciements

C'est très probablement le plus long moment où je suis resté devant une page blanche... Repensant à ces dernières années, au nombre incalculable de personnes qui m'ont permis d'en arriver là où je suis, j'espère n'oublier personne !

Fred, tu as fait tellement de choses pour moi, que je ne te remercierai jamais assez ! Tu as toujours cru en moi et m'as fait confiance. Grâce à toi, j'ai énormément progressé d'un point de vue scientifique, mais également dans ma vie personnelle. Un jour tu m'as dit « Derrière chaque grand homme, il y a une femme », je ne sais pas si je suis un grand homme, mais je souhaiterais compléter cette citation par « il y a également un grand chef ». Sans toi je n'en serais pas là ! Merci infiniment.

Jean-Yves, sans votre soutien, je ne serais pas ici non plus. Je vous remercie de m'avoir fait confiance pour réaliser cette thèse. Vos conseils pertinents m'ont aidé à aller toujours plus loin dans ma réflexion. Malgré toutes vos responsabilités et votre emploi du temps très chargé, vous avez toujours pris le temps de m'aider avec bonne humeur. Pour tout cela, je vous suis très reconnaissant.

Aux membres de mon jury de thèse : Pr. Françoise Rédini, Dr. Laurent Lecam, Dr. Marie Castets et le Dr. Sophie Leguellec, je tiens à vous remercier pour avoir accepté et pris le temps de juger mon travail de thèse ce qui permettra, j'en suis certain, d'enrichir ce travail.

Tom :

```
0100101001100101001000000111001101100001011010010111001100100000011100010111010101100101001000000110101001100101001000000111010
000100000011000010110100100100000011000100110010101100001011101010110001011011110111010101110000001000000111010001110010011011
110111000000100000011100110110111011101010111011001100101011011100111010000100000011001010110101100010011001010111010001100
10100100000011000010111011001100101011000110010000001101101011001010111001100100000011000101110101011001010111001101101000110
100101101110101110011100110010000001100110111010001100001011101000110100101110011011101000110100101110001011010011101010100101011
10011001000000101110010111000100000010101000111010100100000110110100100000110001011100110010000001100001011100110010000001110100011011101010
110101001101110110101011100100111001100100000011000010110100101100100011001010010000001100101011101000010000001101101010100101
01101101011001010010000001100100011001010111001100100000011001100101110110100101110011001000000110101001110101011100110111000
101110101001000000110000100100000010000001110100011100100110010101110011001000000111010001100001011100100110010000101110001011
100010111000100000010100000110111101110101011100100010000001110100011011110111010101110100001000000110001101100001001000000111
0100010000001001101010001010100100100001101001001001000000100001
```

(Oui, tu vas devoir attendre la version numérique pour traduire :p)

Lydia, je tiens à te remercier pour tous tes conseils aussi bien techniques que scientifiques. Tu as toujours été là pour m'aider quand je galérais à avancer... Il est certain que mes projets n'en seraient pas là sans toi. Je tiens à te remercier pour le

temps que tu m'as accordé, ton écoute, ta gentillesse et de m'avoir hurlé dessus dans les couloirs ;)

Mickael, nos discussions sur l'avenir vont me manquer... J'essaierai d'appliquer tes conseils en matière de bébé et de mariage du mieux que je pourrai :D

Gaëlle, c'est toi qui m'as formé techniquement, tu m'as appris la rigueur et pour ça je t'en remercie. Je tiens également à te remercier de m'avoir fait venir à 5h00 du matin pour rien :p (Tu croyais que j'allais oublier ? :p).

Jess, merci à toi, reine de la trotinette, pour ta bonne humeur constante, ton aide et nos discussions toujours sympathiques.

Nina, tu as été ma première stagiaire et je pense que je n'aurais pas pu tomber sur mieux. Merci pour ta motivation et ta détermination, surtout lorsque l'on connaît RCBTB1 qui est très capricieux :p Je te souhaite tout le courage et la chance nécessaires pour la suite ;)

Candichauve ! A défaut d'avoir un gâteau pour m'avoir aidé cet été à l'animalerie, je te laisse une place dans les remerciements ;) Merci pour cet été :p Je te souhaite bon courage pour la suite, tu y arriveras sans problème ;)

Je remercie également toutes les personnes du laboratoire que j'ai eu l'occasion de côtoyer durant ces années.

Marie, Sam, Agnès et Reynald : Depuis quatre ans où je travaille au même endroit, j'ai créé certains liens avec des personnes qui sont devenues plus que des collègues. Je vous remercie d'avoir été là et d'avoir fait de mon lieu de travail, un lieu de bonne humeur et d'amitiés. Vous me manquerez énormément.

Adrien, Emilie, Elise, Laurianne, Pierre et Elodie : A mes amis Bordelais, pour toutes ces soirées où l'on s'est bien marré, je vous souhaite plein de bonnes choses pour l'avenir !

Greg et Yann : 23 ans environ que nous nous connaissons... Ces 4 dernières années n'ont pas été très faciles pour nous 3. Nous avons été beaucoup occupés et nous

nous sommes peu vus. Néanmoins, cela ne change en rien notre amitié et c'est cela qui est génial. Merci d'avoir été là quand j'avais besoin de vous ☺

Matthieu : Le 1^{er} scientifique de la famille... Peut être est-ce toi qui m'a donné ce goût pour la science... Je me rappelle les jours où tu m'expliquais les maths :p Maintenant, nous travaillons ensemble avec nos connaissances spécifiques pour réaliser quelque chose de nouveau ☺ Merci de faire tout ton possible pour satisfaire mes exigences de biologiste envers l'informatique ;)

Maman et Marie : Je vous dois tout, merci de m'avoir donné toutes les chances pour réaliser ces longues études. Merci d'avoir été là dans tous les moments où j'avais besoin de vous. Merci de m'avoir soutenu, épaulé.... Et supporté ! Je ne pourrai jamais vous remercier suffisamment mais sachez que je vous en serai éternellement reconnaissant.

Vanessa : La meilleure pour la fin... Je ne m'étendrai pas ici car j'en aurais pour 180 pages de plus ;) Si je « synthétise » : tu es quelqu'un d'exceptionnel ! Tu es une fille brillante, intelligente, gentille et attentionnée. J'ai énormément de chance de t'avoir. Sans toi, je ne serais pas allé aussi loin... Merci pour ton aide et ton soutien constant pendant toutes ces années.

Identification et implication des gènes *DMD* et *RCBTB1*
dans la progression tumorale des sarcomes à génétique
complexe

Résumé

Les sarcomes sont des tumeurs rares (environ 1% des cancers) des tissus mous. Ils forment un groupe de tumeurs très hétérogène puisqu'il existe plus de 50 types de tumeurs malignes d'origine mésenchymateuse. Leur rareté et leur hétérogénéité clinique rendent la prise en charge des patients très difficile. Ainsi, il est nécessaire d'identifier et de comprendre les mécanismes de l'oncogenèse de ces tumeurs.

Pour cela, notre équipe a analysé 106 sarcomes par CGH-array et puces d'expression afin d'identifier des gènes d'intérêt pour expliquer la biologie de ces sarcomes. Grâce à des analyses intégratives bio-informatiques, deux gènes candidats majeurs ont été identifiés : les gènes *RCBTB1* et *DMD*, dont la perte génomique et la diminution d'expression sont significativement associées à l'évolution métastatique. Le gène *RCBTB1* est composé de 6 domaines RCC1 (*Regulator of Chromosome Condensation 1*) et de deux domaines BTB (permettant les interactions protéine/protéine). Il a déjà été décrit comme étant un potentiel gène suppresseur de tumeur impliqué dans les leucémies lymphoïdes chroniques. Le gène *DMD* (gène codant pour la dystrophine) est, quant à lui, notamment connu pour provoquer les myopathies musculaires de Becker et de Duchenne lorsqu'il est muté. Dans notre cohorte de sarcomes, il est délété dans 24% des synoviosarcomes, 21% des GISTs (*Gastro-Intestinal Stromal Tumors*) et 15% des sarcomes à génétique complexe.

Dans ce contexte, après la validation exhaustive des données de génomique à haut débit, nous avons créé des modèles cellulaires (synoviosarcomes et léiomyosarcomes) soit inhibant de façon inductible l'expression de ces gènes par shARN, soit sur-exprimant par vecteur lentiviral l'ADNc de ces gènes. Ainsi, nous avons pu évaluer l'impact de l'expression de ces gènes au niveau phénotypique par étude de la prolifération, de la résistance à l'apoptose, de la migration, de l'invasion cellulaires et de la clonogénicité.

Nos résultats montrent que la délétion de *RCBTB1* est pronostique de l'évolution métastatique. Il semblerait que cela soit dû à la sensibilisation des cellules

à certaines chimiothérapies, plus particulièrement le docétaxel, lorsque *RCBTBI* est surexprimé. En effet, nous avons pu montrer qu'une surexpression de ce gène augmentait le pourcentage de cellules en mitose or, le docétaxel agissant uniquement en mitose, ceci explique la sensibilisation des cellules au docétaxel. Enfin, nous avons vérifié les effets de sensibilisation au docétaxel *in vivo*, en étudiant la croissance tumorale après traitement en fonction de l'expression de *RCBTBI*.

Concernant *DMD*, nos résultats obtenus par RNA-seq montrent que seules deux isoformes sont exprimées dans les tumeurs d'origine myogénique : Dp427 et Dp71. Cependant, dans environ 20% des tumeurs, nous observons une délétion ciblant spécifiquement l'isoforme Dp427. L'étude d'une lignée de LMS dans laquelle j'ai inhibé spécifiquement Dp427 grâce au système CRISPr/Cas9 montre que son inactivation ne modifie ni la prolifération cellulaire ni la clonogénicité de la lignée. Etant donné que la délétion de Dp427 n'a pas d'effet sur l'agressivité des cellules, j'ai testé l'hypothèse selon laquelle Dp71 aurait un rôle oncogénique. Ainsi, l'inhibition de son expression par shARN dans deux lignées de LMS et une lignée de synoviosarcome révèle que Dp71 est nécessaire à la croissance cellulaire et à la formation de clones, car son inhibition entraîne un blocage du cycle cellulaire en phase G2/M. J'ai alors voulu tester le rôle oncogénique de Dp71 dans une lignée de myoblaste. La surexpression de Dp71 n'a pas eu d'impact sur la prolifération et la clonogénicité de ces cellules, toutefois la forte expression basale de cette protéine peut expliquer l'absence d'effet notable. En effet, les cellules musculaires immatures expriment Dp71, qui est indispensable pour leur renouvellement, et ne perdent son expression au profit de celle de Dp427 qu'au cours de la différenciation musculaire.

Sommaire

Introduction	16
Chapitre 1 : Les sarcomes des tissus mous	17
I. Généralités	17
a) Incidence	17
b) Survie.....	18
c) Prise en charge thérapeutique.....	19
d) Etiologie	23
e) Classification	24
f) Estimation du risque métastatique	26
II. Sarcomes présentant une altération spécifique	28
a) Mutations.....	28
b) Translocations	29
c) Remaniements spécifiques	32
III. Sarcomes à remaniements complexes	32
Chapitre 2 : Identification de gènes impliqués dans l'oncogenèse ou la progression tumorale des sarcomes à génétique complexe	38
I- Introduction	38
II- Matériel et méthode	39
a) Les gènes altérés en génomique	39
b) Les gènes altérés en transcriptomique.....	40
c) Les voies biologiques	40
III- Résultats	41
a) Les voies biologiques communes	41
b) Les gènes spécifiques de sous-groupes dans les voies biologiques communes	45
IV- Etude du chromosome X	46
Chapitre 3 : <i>RCBTB1</i> : un gène méconnu.....	50
I – Description générale de <i>RCBTB1</i>	50
a) Le gène <i>RCBTB1</i> : structure, expression et régulation.....	50
b) La protéine <i>RCBTB1</i> : structure et localisation	51
II – Rôle physiologique de <i>RCBTB1</i>	55
a) Mises en évidences du rôle cellulaire de <i>RCBTB1 in vitro</i>	56
b) Fonctions possibles impliquant le domaine RCC1	57
c) Fonctions possibles impliquant les domaines BTB.....	58

III- Rôle de <i>RCBTB1</i> dans les pathologies	61
a) <i>RCBTB1</i> et cancer.....	61
b) Maladies cardiovasculaires et rénales.....	62
c) Rétinopathies.....	63
Chapitre 4 : <i>DMD</i> : un gène complexe	64
I- Le gène <i>DMD</i> : Découverte et description	64
II- Les protéines de <i>DMD</i>	65
a) Dp427	66
b) Dp260	67
c) Dp140	67
d) Dp116.....	67
e) Dp71	67
f) Dp45 (ou Dp40).....	68
III- Rôles physiologiques de Dp427 et Dp71	68
a) Dp427	68
b) Dp71	69
IV- Rôle dans les pathologies	74
a) Myopathies de Duchenne et de Becker	74
b) Retard mental	75
c) Rétinopathie.....	76
d) Cancer.....	76
Objectifs :	78
Matériels et Méthodes	79
Résultats	82
I- <i>RCBTB1</i> : Rôle dans la progression tumorale	83
A) Contexte.....	83
B) Résultats	84
II- Rôle de la délétion 5' du gène <i>DMD</i> dans l'évolution métastatique	125
A) Contexte.....	125
B) Résultats	125
Discussion	135
I- <i>RCBTB1</i> : un nouveau marqueur biologique ?	137
II- <i>DMD</i> : un gène complexe associé à la progression tumorale des sarcomes ?	141
III- Discussion générale	148

Références	153
Annexes.....	177
I- Matériel et Méthode du projet RCBTB1	178
II- Matériel et Méthode du projet <i>DMD</i>	181

Table des illustrations

Figure 1 : Localisation des STM	17
Figure 2 : Espérance de vie des patients atteints de sarcomes en fonction de l'âge	18
Figure 3 : Chimiothérapies recommandées pour le traitement des STS avancés	22
Figure 4: Taxonomie des sarcomes des tissus mous	25
Figure 5: Prédiction de l'évolution métastatique en fonction de CINSARC et du grade	27
Figure 6 : Schéma du mécanisme de translocation aboutissant à une protéine de fusion	31
Figure 7 : Etalement chromosomique et caryotype d'un sarcome à génétique complexe	33
Figure 8 : Schéma représentant les différents filtres appliqués	39
Figure 9 : Classification hiérarchique des 41 voies biologiques selon Gene Ontology	43
Figure 10 "Hallmarks of cancer", les grandes caractéristiques de la cellule tumorale	44
Figure 11 : Fréquence des altérations des gènes en fonction des annotations cliniques	46
Figure 12 : Différents types de délétion de DMD observés par CGH array	47
Figure 13: Expression transcriptomique de DMD en fonction de son statut génomique	47
Figure 14 : Kaplan Meier de la survie sans métastase en fonction de la délétion de DMD	48
Figure 15: Délétion ciblée du gène DMD par CGH array	49
Figure 16 Représentation schématique des membres de la superfamille de protéines RCC1	52
Figure 17 Marquage par immunofluorescence de RCBTB1 (vert) dans des lignées cellulaires	53
Figure 18 Marquage immunohistochimique de RCBTB1 dans du tissu sain : rein (A) et foie (B)	54
Figure 19 Marquage immunohistochimique de RCBTB1 dans du tissu tumoral hépatique	54
Figure 20 Localisation de la protéine de fusion RCBTB1-GFP dans les lignées HeLa et Vero	55
Figure 21 La CUL3, E2 et une protéine BTB pour la reconnaissance et l'ubiquitination du substrat ...	59
Figure 22 TIAM1 est dégradé suite à la coopération entre GABARAP et le complexe CUL3	61
Figure 23: Schéma du gène DMD donnant lieu aux différentes isoformes	65
Figure 24: Schéma des domaines protéiques de Dp427	66
Figure 25 : Schéma du complexe DCG	69
Figure 26 Régulation de l'expression de Dp71 lors de la différenciation myogénique et neuronale ...	69
Figure 27: Schéma représentant les différents rôles de Dp71	71
Figure 29 : Inhibition pour RCBTB1 en fonction des différents shARN après 48h	119
Figure 30 : Visualisation des foyers γ H2AX dans les fibroblastes inhibés ou non pour RCBTB1 ...	121
Figure 31 : Profils génomiques des lignées obtenus suite à l'induction à la doxycycline	122
Figure 32 : Poids des tumeurs récupérées 30 jours après le début du traitement	123
Figure 33 : Expression par RNA-seq des isoformes du gène DMD dans 145 sarcomes	126
Figure 34 : Expression des isoformes Dp427 et Dp71 dans des lignées de sarcomes	127
Figure 35 : Expression de Dp427 et Dp71 dans les lignées par Western Blot	127
Figure 36 : Séquençage Sanger de la lignée IB133 après infection	128
Figure 37 : Validation de l'inhibition de Dp427 par CRISPr	128
Figure 38: Tests de prolifération et de clonogénicité	129
Figure 39 : Expression de Dp71 avec les shARN ciblant Dp71 dans les 3 lignées étudiées	130
Figure 40: Courbes de prolifération en fonction de l'inhibition de Dp71	130
Figure 41 : Test de clonogénicité	131
Figure 42 : Etude du cycle cellulaire par cytométrie en flux	132
Figure 43: Western Blot de validation de la sur-expression de Dp71	133
Figure 44 : Tests de prolifération et de clonogénicité en fonction de la sur-expression de Dp71	133
Figure 44 : Schéma du plasmide Tet-pLKO-puro et séquences des shARN utilisés	179

Liste des tableaux

Tableau 1 : Liste des différents sarcomes présentant une ou plusieurs translocations spécifiques	30
Tableau 2 : Liste des 41 voies biologiques alétries dans les 106 sarcomes	42
Tableau 3 : Délétion de DMD en fonction de la métastase et des histotypes.....	48

Abréviations

β -DG : β -dxytroglycane
ABS : *Actin Binding Site*
ADN : Acide désoxyribonucléique
ADNc : Acide désoxyribonucléique complémentaire
ARN : Acide ribonucléique
ATP : Adénosine triphosphate
BMD : *Becker Muscular Dystrophy*
BTB : *Broad complex, Tramtrack and Bric-a-brac*
CaMBS : *Calmodulin Binding Site*
CGH-array : *Comparative Genomic Hybridization array*
CINSARC : *Complexity INdex in SARComas*
CNV : *Copy Number Variation*
DAP : *Dystrophin-Associated Protein*
DCG : *dystrophin glycoprotein complex*
Dp : *Dystrophin protein*
ESMO : *European Society for Medical Oncology*
FISH : *Fluorescent In Situ Hybridization*
FNCLCC : Fédération Nationale de Lutte Contre le Cancer
GISTs : *Gastro-Intestinal Stromal Tumors*
GFP : *Green fluorescent protein*
HPF : *High Power Field*
LMS : Léiomyosarcomes
LOH : *Loss Of Heterozygosity*
LPSP : Liposarcome Pléomorphe
LPSDD : Liposarcome Dédifférencié
MFS : Myxofibrosarcomes
MPNST : *Malignant Peripheral Nerve Sheath Tumors*
NLS : *Nuclear Localization Sequence*
QI : Quotient Intellectuel
Ran : *Ras-related nuclear protein*
RCC1 : *Regulator of chromosome condensation 1*
RCP : Réunion de Concertation Pluridisciplinaire
RLD : *RCC1-like domain*
RMSP : Rhabdomyosarcome pléomorphe
RT-qPCR : *Reverse transcription quantitative polymerase chain reaction*
shARN : *Short Hairpin ARN*
shNT : *Short Hairpin Non Targeting*
STM : Sarcomes des Tissus Mous
UPS : *Undifferentiated Pleomorphic Sarcoma*

Introduction

Chapitre 1 : Les sarcomes des tissus mous

I. Généralités

Les sarcomes des tissus mous (STM), sont des tumeurs malignes d'origine mésenchymateuse et très hétérogènes¹. Plus de 50 types histologiques sont décrits selon la morphologie et la différenciation cellulaire des tumeurs. Ils se développent au niveau des tissus mous et des tissus conjonctifs de manière ubiquitaire. Ces tissus comprennent les tendons, les muscles, la graisse, les vaisseaux sanguins, les nerfs et les tissus synoviaux.

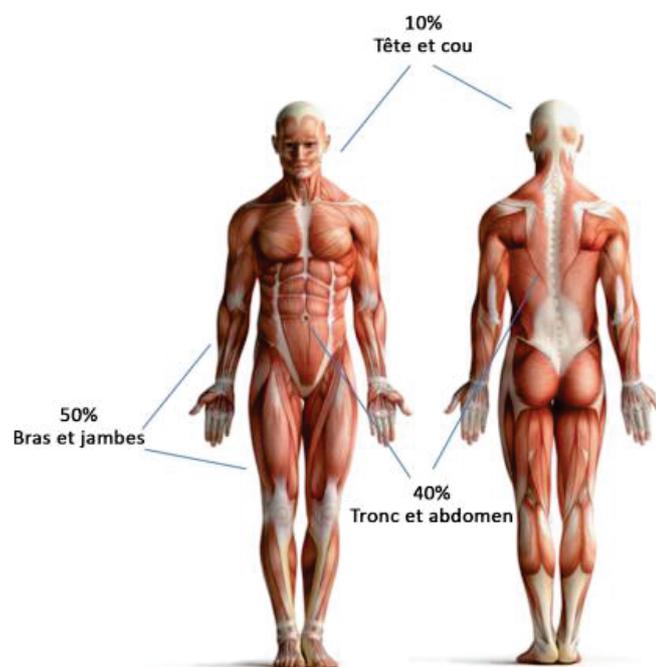


Figure 1 : Localisation des STM

(Fonds Anticancer, 2016)

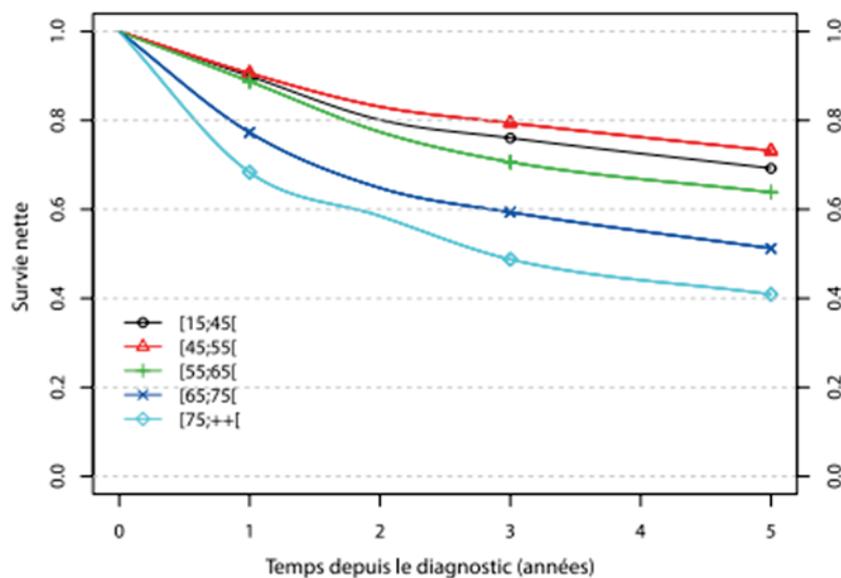
Les STM sont susceptibles de se développer dans presque tout l'organisme mais sont néanmoins plus fréquents au niveau des extrémités (50%) qu'au niveau du tronc et de l'abdomen (40%). Les sarcomes de la tête et du cou ne représentent que 10% des cas.

a) Incidence

Les STM sont des tumeurs rares, ils représentent moins de 1% des cas de cancer chez l'adulte² et de 4 à 8% des cas chez l'enfant³. L'incidence en France est de 2.8 à 3.3 cas pour 100 000 habitants par an. Leur rareté et leur hétérogénéité en font donc des tumeurs difficiles à diagnostiquer et à traiter.

b) *Survie*

La survie des patients présentant un sarcome est faible : la survie à 5 ans est de 57% et varie selon l'âge : 69% pour les personnes âgées de 15 à 45 ans, contre 41% pour les patients de plus de 75 ans.



(Amadeo *et al.*, 2016)

Figure 2 : *Espérance de vie des patients atteints de sarcomes en fonction de l'âge*

Par ailleurs, la survie dépend également d'autres facteurs tels que le type histologique. A titre d'exemple, les léiomyosarcomes sont beaucoup plus agressifs que les fibromyxosarcomes. Par ailleurs, le niveau de différenciation des cellules tumorales joue également sur la survie : dans la majorité des cas, les sarcomes peu différenciés sont plus agressifs que les sarcomes bien différenciés. De plus, il a été montré que l'instabilité génomique est associée à une plus faible survie sans métastase⁴. La localisation a également un impact puisque les tumeurs des extrémités sont associées à un meilleur pronostic que les tumeurs rétro-péritonéales.

c) Prise en charge thérapeutique

➤ La chirurgie

Lorsque le sarcome est localisé, la chirurgie reste le traitement standard avec une exérèse large, c'est-à-dire une marge de 1 à 2 cm autour de la tumeur dans tous les plans. La qualité de l'exérèse est déterminée par confrontation des données chirurgicales et anatomopathologiques :

- R0 : marge microscopique saine
- R1 : existence d'un résidu tumoral microscopique
- R2 : existence d'un résidu macroscopique

En fonction de la qualité de l'exérèse, le patient aura une probabilité plus ou moins forte de faire une rechute. Si la marge réalisée est de type R1 ou R2, une nouvelle discussion en RCP (Réunion de Concertation Pluridisciplinaire) a lieu pour une éventuelle reprise d'exérèse.

➤ La radiothérapie

Elle peut être réalisée avant ou après l'opération chirurgicale. Lorsque la radiothérapie est pré-opératoire, le taux de complications post-chirurgicales est de 35% concernant la cicatrisation, contre 17% avec une radiothérapie post-opératoire⁵. C'est pour cela que la radiothérapie pré-opératoire n'est prescrite que dans certains cas où la tumeur est particulièrement volumineuse, avec un risque de chirurgie mutilante et de marge R1 ou R2. La dose délivrée est généralement de 50 Gy⁶.

Concernant la radiothérapie post-opératoire, elle doit être envisagée lorsque :

- La tumeur est de grade 2 ou 3
- La taille de la tumeur est supérieure ou égale à 5cm
- La tumeur est profonde

- La reprise chirurgicale mutilante

Lorsque la radiothérapie est prescrite, la dose délivrée dépend de la qualité de l'exérèse selon les recommandations suivantes⁶ :

- R0 : 50 Gy
- R1 : 54-60 Gy
- R2 : 60-66 Gy

➤ *La chimiothérapie*

Dans tous les cas, l'administration de chimiothérapie doit être discutée en RCP. De manière générale, la chimiothérapie administrée en 1^{ère} ligne pour le traitement des sarcomes doit comporter un anthracycline (ex : doxorubicine). Il a été montré que la combinaison de plusieurs chimiothérapies augmente le taux de réponses objectives sans que la survie globale n'augmente. Pour le traitement des sarcomes non métastatiques, nous pouvons distinguer deux indications pour le traitement par chimiothérapie : soit en néo-adjuvant (avant exérèse) soit en adjuvant (après exérèse).

- Chimiothérapie néo-adjuvante :

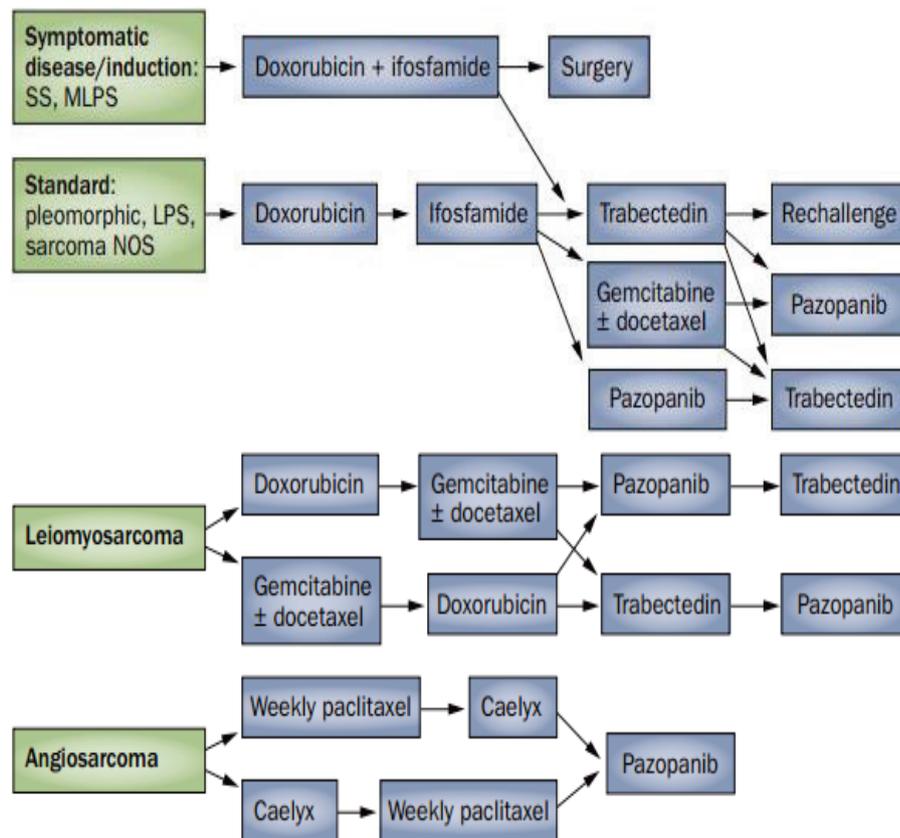
Il n'a pas été montré que la chimiothérapie néo-adjuvante améliore la survie globale ou la survie sans récurrence lorsque la tumeur est directement opérable avec des marges d'exérèse R0. Néanmoins, actuellement, deux essais cliniques européens pas encore publiés semblent trouver un bénéfice à la chimiothérapie néo-adjuvante. Elle va alors peut-être devenir un nouveau standard dans les prochaines années.

Elle est donc indiquée dans le traitement des sarcomes difficilement opérables avec des marges satisfaisantes. Il a été montré que la chimiothérapie néo-adjuvante peut induire une diminution du volume tumoral et ainsi faciliter le geste chirurgical. Actuellement, il n'existe pas de traitement standard pour les chimiothérapies néo-adjuvantes, mais il est conseillé de privilégier les combinaisons de chimiothérapies (par exemple, la doxorubicine associée à l'ifosfamide) afin d'améliorer le taux de réponse.

- Chimiothérapie adjuvante :

La prescription systématique de chimiothérapies dans le traitement des STM reste controversée et il n'existe donc pas de standard. Certaines études ont montré que la chimiothérapie pourrait légèrement améliorer ou au moins retarder la rechute locale et l'apparition de métastase dans le cas des patients à haut risque^{7,8}. Elle est donc proposée aux patients à haut risque, c'est-à-dire avec une tumeur de grade 2 ou 3 et une taille supérieure à 5 cm. Dans ces cas de figure, la combinaison d'anthracycline (comme la doxorubicine) avec l'ifosfamide est classiquement administrée⁶ à doses optimales (4 à 6 cycles).

Concernant les STM métastatiques, la chimiothérapie sera palliative pour les patients auxquels aucun traitement à visée curative ne pourra être proposé. Le plus souvent, une mono-chimiothérapie standard sera proposée, car les combinaisons de plusieurs chimiothérapies n'améliorent pas la survie globale mais présentent une toxicité plus importante. Il sera également envisagé d'intégrer le patient à des essais thérapeutiques. La chimiothérapie est le principal traitement des STM métastatiques, car les médicaments administrés diffusent par la circulation sanguine et peuvent atteindre les cellules tumorales disséminées dans l'organisme.



(Adapté de Linch M. *et al.*, 2014)

Ce schéma ne représente pas le standard pour le traitement des STM puisqu'il n'y a pas de règle générale et que la décision est prise au cas par cas en RCP. Il a pour seul but d'indiquer les principales chimiothérapies utilisées en fonction des types histologiques.

➤ Thérapies ciblées

Il n'existe que deux thérapies ciblées approuvées par L'ESMO (European Society for Medical Oncology) pour le traitement des STM. Il s'agit du pazopanib et de l'imatinib.

Le pazopanib est indiqué dans le traitement des STM métastatiques, excepté ceux d'origine adipocytaire⁹. Il est administré en 2^{ème} ligne chez les patients ayant reçu en

1^{ère} ligne des anthracyclines¹⁰. C'est un inhibiteur des récepteurs à tyrosine kinase pour le facteur de croissance VEGFR2. Il bloque la croissance des tumeurs par l'inhibition de l'angiogenèse. Pour cette raison, il peut interférer dans le processus de guérison et son utilisation doit être considérée avec prudence chez les patients subissant des traitements invasifs.

L'imatinib est utilisé dans le traitement des GISTs (*Gastro-Intestinal Stromal Tumors*) localement avancés et inopérables ou potentiellement métastatiques¹¹. C'est une molécule qui entre en compétition avec l'ATP au niveau des récepteurs à tyrosine kinase tels que CD117 (codé par le gène cKIT) et PDGFR qui sont spécifiquement mutés dans les GISTs¹². En entrant en compétition avec l'ATP, l'imatinib inhibe donc la signalisation de ces récepteurs et tend à provoquer l'apoptose.

d) *Etiologie*

L'étiologie des sarcomes est encore inconnue. Cependant, dans de rares cas, certains facteurs ont été décrits comme associés au développement de tumeurs malignes des tissus mous :

- **environnementaux** :

Il a été rapporté une augmentation de l'incidence des STM après une exposition à certains herbicides et chlorophénols chez les agriculteurs^{13,14}.

- **irradiation** :

Il a été montré une augmentation de l'incidence des sarcomes chez les patients ayant reçu des irradiations supérieures à 50 Gy¹⁵.

- **infection virale et déficience immunitaire** :

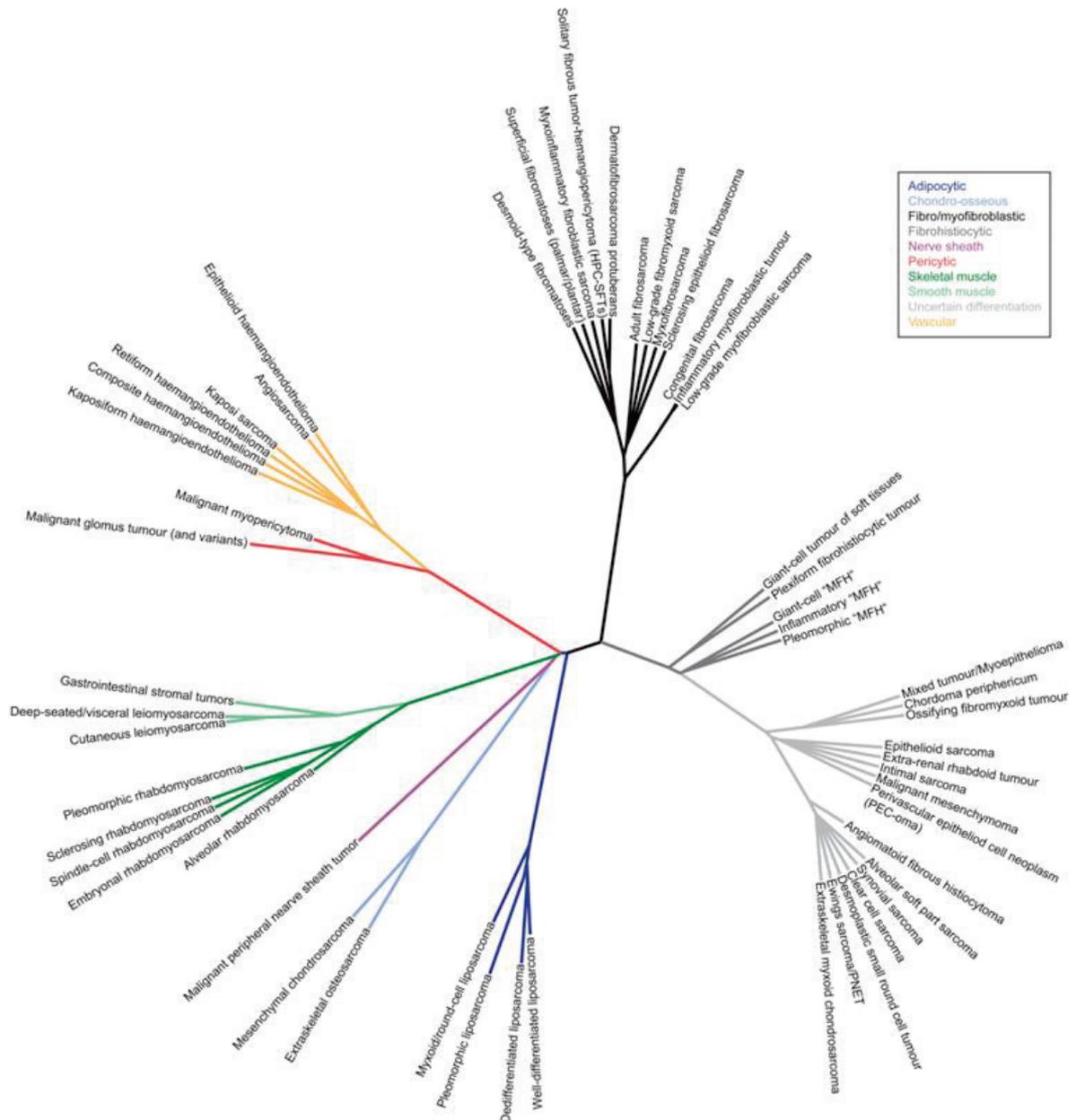
Plusieurs virus ont été décrits comme ayant un rôle dans le développement de sarcomes. Par exemple, le virus humain de l'herpès 8 a été démontré comme impliqué dans le développement du sarcome de Kaposi¹⁶, pour lequel l'évolution clinique dépend de l'efficacité du système immunitaire du patient¹⁷.

- prédisposition génétique :

Environ 15% des tumeurs desmoïdes surviennent chez des patients atteints d'une maladie génétique : la polypose adénomateuse familiale. La neurofibromatose est également associée à des tumeurs survenant au niveau des gaines périphériques du système nerveux. De plus, la mutation du gène *TP53* qui cause le syndrome de Li-Fraumeni^{18,19} est associée à l'apparition de sarcomes principalement au niveau des muscles (rhabdomyosarcomes). Par ailleurs, notre laboratoire a montré que *TP53* était inactivé, de manière directe ou indirecte, dans 100% des sarcomes²⁰. Enfin, la fréquente altération du gène *RBI* dans ces tumeurs est également à souligner. Cependant, la mutation des gènes *TP53* et *RBI* n'est pas spécifique des sarcomes²¹. Il semble donc que la majorité des sarcomes apparaissent de manière spontanée sans cause apparente¹.

e) Classification

Au niveau anatomopathologique, les sarcomes sont classés en fonction de leur localisation, de leur morphologie et d'une ligne de différenciation éventuelle selon des critères immunohistochimiques (Figure 4).



(Tailor *et al.*, 2011)

Figure 4: Taxonomie des sarcomes des tissus mous

Les sarcomes des tissus mous de l'adulte les plus fréquents et présentant une ligne de différenciation sont le liposarcome, le léiomyosarcome et le synovialosarcome. Toutefois, 30% des sarcomes des tissus mous de l'adulte ne montrent pas de ligne de différenciation identifiable : ils sont alors regroupés sous l'appellation « sarcome indifférencié ».

f) Estimation du risque métastatique

L'évolution métastatique des sarcomes reste aujourd'hui le facteur péjoratif majeur dans la prise en charge des patients et entraîne, dans la majorité des cas, le décès du patient. Toutes les tumeurs ne possèdent pas le même potentiel métastatique. De grandes disparités existent d'un type tumoral à un autre, mais également d'une tumeur à une autre pour un même type histologique.

Selon la Fédération Nationale de Lutte Contre le Cancer (FNLCC), trois critères ont été établis pour classer les tumeurs selon la probabilité d'évolution métastatique :

➤ **La différenciation tumorale :**

Score 1 : sarcomes ressemblant à un tissu adulte normal

Score 2 : sarcomes pour lesquels le diagnostic histologique est certain

Score 3 : sarcomes embryonnaires, synoviosarcomes, sarcomes épithélioïdes, sarcomes à cellules claires, sarcomes alvéolaires des parties molles, sarcomes indifférenciés et sarcomes pour lesquels le type histologique est incertain.

➤ **L'index mitotique (surface du champ (=HPF) : 0,174 mm²) :**

Score 1 : 0 à 9 mitoses pour 10 HPF (pour 10 champs)

Score 2 : 10 à 19 mitoses pour 10 HPF

Score 3 : Plus de 19 mitoses pour 10 HPF

➤ **La nécrose tumorale :**

Score 0 : Absence de nécrose

Score 1 : < 50 % de nécrose tumorale

Score 2 : > 50 % de nécrose tumorale

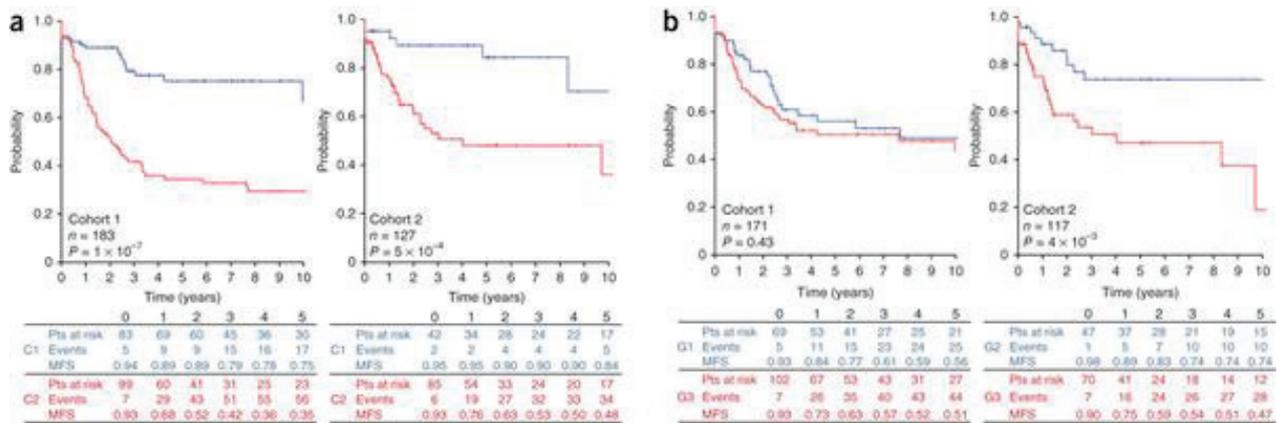
En additionnant le score obtenu pour les trois critères précédents, il est possible de déterminer le grade de la tumeur :

Grade 1 : somme des scores = 2-3 (risque métastatique faible)

Grade 2 : somme des scores = 4-5 (risque métastatique réel)

Grade 3 : somme des scores = 6-8 (risque métastatique élevé)

Nous pouvons noter que ce système de grade encore utilisé actuellement ne tient compte d'aucune donnée moléculaire or, notre laboratoire a montré que l'expression de certains gènes est plus prédictive du devenir des patients que ne l'est le grade FNCLCC²² : il s'agit de la signature CINSARC⁴.



(Chibon *et al.*, 2010)

Figure 5: Prédiction de l'évolution métastatique en fonction de CINSARC et du grade

a. Probabilité de survie sans métastase (MFS) au cours du temps après le diagnostic, pour des sarcomes divisés en deux groupes (C1 et C2) selon la classification CINSARC. Les patients avec une faible expression de la signature CINSARC sont en bleu et ceux avec une forte expression sont en rouge. **b.** Probabilité de survie sans métastase dans les deux groupes de sarcomes selon le grade de la FNCLCC. Les patients avec un sarcome de grade 1 ou 2 sont en bleu, ceux avec un sarcome de grade 3 sont en rouge.

La signature CINSARC est composée de 67 gènes significativement associés à la complexité génomique des tumeurs et à leur évolution métastatique.

Depuis les années 1990, le développement de la biologie moléculaire associé aux techniques de cytogénétique ont permis de caractériser en détail le génome des sarcomes grâce notamment aux techniques pangénomiques (caryotype 24 couleurs et CGH-array), mais également grâce aux techniques permettant d'étudier avec précision un locus (FISH et PCR). Ainsi, les sarcomes ont pu être classés en fonction de leurs anomalies génomiques. A titre d'exemple, en 2003, 20% des sarcomes indifférenciés ont pu être re-classés en liposarcomes différenciés car ils présentaient une amplification des gènes *MDM2* et *CDK4*²³, spécifique des tumeurs présentant une différenciation adipocytaire. Il est donc intéressant de diviser les sarcomes en deux principaux groupes : ceux présentant une altération spécifique (telle que des mutations

ou des translocations) et les sarcomes à remaniements complexes (pour lesquels aucune altération spécifique n'a été mise en évidence).

II. Sarcomes présentant une altération spécifique

a) Mutations

Les tumeurs les plus fréquentes présentant une mutation spécifique sont les GISTs avec une mutation des gènes *KIT* ou *PDGFRA* dans 90% des cas. La mutation activatrice peut avoir lieu dans les exons 9,11, 13 et 17 de *KIT* et dans les exons 12, 14 et 18 de *PDGFRA*¹². Ces tumeurs se développent principalement au niveau de l'estomac et de l'intestin grêle. Le pronostic est variable, les tumeurs de petite taille avec peu de mitoses ont le plus souvent un comportement bénin (guérison définitive après l'exérèse), alors que les tumeurs volumineuses et riches en mitoses ont une probabilité plus importante de devenir métastatiques. De plus, avec l'apport de la biologie moléculaire, on peut prédire la résistance secondaire vis-à-vis de l'imatinib. Ainsi, l'effet de l'imatinib sur les tumeurs ayant une mutation dans l'exon 18 de *PDGFRA* sera faible²⁴.

Le second groupe de STM présentant une mutation spécifique sont les tumeurs rhabdoïdes. Elles sont caractérisées par une mutation inactivatrice du gène *SMARCB1*²² (anciennement appelé *INI1*). Les tumeurs rhabdoïdes affectent principalement les jeunes enfants. Les individus ayant une altération somatique hétérozygote pour *SMARCB1* sont donc prédisposés à développer des tumeurs rhabdoïdes du cerveau, du rein et des tissus mous, et peuvent présenter plusieurs tumeurs primaires²⁵⁻²⁷. Ces enfants sont la plupart du temps diagnostiqués durant la première année de leur vie et ont un mauvais pronostic²⁸. La perte de *SMARCB1* provoque une dérégulation du cycle cellulaire par la sous-expression de p16^{INK4a} et par la surexpression des gènes de la famille E2F et de la Cycline D1^{29,30}.

Les tumeurs desmoïdes ou fibromatoses profondes sont caractérisées par une mutation activatrice de l'exon 3 de la β -caténine (*CTNNB1*)³¹ dans 85 % des formes sporadiques. Les 15% restants apparaissent chez les patients atteints d'une maladie

génétique, la polypose adénomateuse familiale (FAP)³² causée par une mutation du gène *APC*. Les tumeurs desmoïdes ont une incidence de 8% à 12% dans le cas des patients atteints de FAP. Ces deux protéines agissent dans la voie Wnt. En absence du ligand Wnt, les protéines *APC* (*adenomatous polyposis coli*)³³, Axine, *GSK3β* (*Glycogen Synthase 3β*) et CK1 (caséine kinase 1) forment un complexe qui phosphoryle la β -caténine, entraînant ainsi son ubiquitination puis sa dégradation. La fixation de Wnt à son récepteur recrute le complexe et libère ainsi la β -caténine, qui va être transloquée au noyau pour y assurer son rôle de facteur de transcription. Le dérèglement de cette voie de signalisation est impliqué dans le développement de nombreux cancers chez l'homme³⁴.

En conclusion, l'identification de mutations spécifiques dans les sarcomes a permis, d'une part, de faciliter le diagnostic et donc d'améliorer le temps de prise en charge des patients, et d'autre part, de comprendre la biologie de ces tumeurs, ceci ayant pour conséquence d'améliorer le traitement des patients comme c'est le cas pour les GISTs avec l'imatinib.

b) Translocations

La première identification d'une translocation réciproque spécifique dans une tumeur solide remonte à 1983 par Alain Aurias³⁵ et collaborateurs. Cette translocation entre les gènes *EWS* et *FLII* ou *EWS* et *ERG* constitue actuellement le principal critère de définition du sarcome d'Ewing³⁶.

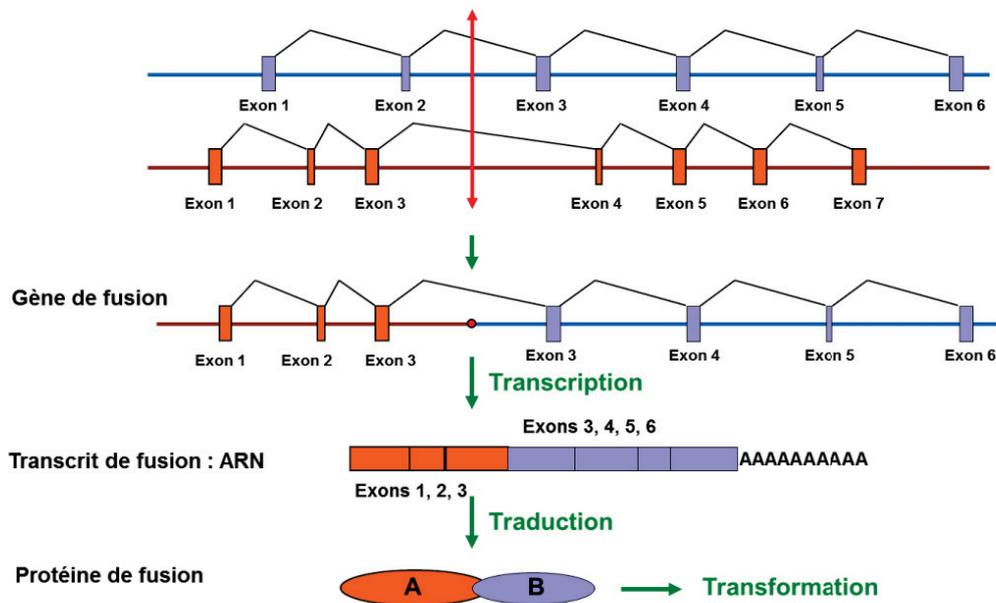
Les sarcomes à translocation spécifique représentent 30% des sarcomes³⁷. Ils sont listés dans le tableau suivant. Les sarcomes les plus fréquents sont les liposarcomes myxoïdes, les sarcomes d'Ewing et les synoviosarcomes. A noter que même si ces sarcomes sont caractérisés par une translocation spécifique, leur génome peut présenter d'autres altérations secondaires comme dans le cas des synoviosarcomes et du sarcome d'Ewing³⁷.

Tableau 1 : Liste des différents sarcomes présentant une ou plusieurs translocations spécifiques

Type tumoral	Translocations	Gènes de fusion	Fréquences
Tumeur d'Ewing ; pPNET	t(11;22)(q24;q12)	EWS/FLI1	85-95 %
	t(21;22)(q22;q12)	EWS/ERG	5-10 %
	t(2;22)(q33;q12)	EWS/FEV	Rare
	t(7;22)(p22;q12)	EWS/ETV1	Rare
	t(17;22)(q12;q12)	EWS/ETV4 (E1AF)	Rare
	t(1;22)(p36;q12)	EWS/ZSG	Rare
Tumeur desmoplastique à petites cellules rondes	t(16;21)(p11;q22)	FUS/ERG	Rare
	t(11;22)(p13;q12)	EWS/WT1	>90 %
Mélanome des tissus mous ; sarcome à cellules claires	t(12;22)(q13;q12)	EWS/ATF1	>90 %
	t(2;22)(q32;q12)	EWS/CREB1	Rare
Liposarcome myxoïde/à cellules rondes	t(12;16)(q13;p11)	TLS (FUS)/CHOP	95 %
	t(12;22)(q13;q12)	EWS/CHOP	Rare
Chondrosarcome myxoïde extrasquelettique	t(9;22)(q22;q12)	EWS/TEC (CHN)	75 %
	t(9;17)(q22;q11)	TAFI168	25 %
	t(9;15)(q22;q21)	(RBP56)/TEC (CHN) TCF12/TEC (CHN)	Rare
Rhabdomyosarcome alvéolaire	t(2;13)(q35;q14)	PAX3/FKHR	60-80 %
	t(1;13)(p36;q14)	PAX7/FKHR	10-20 %
	t(X;2)(q13;q35)	PAX3/AFX	Rare
	t(2;2)(q35;p23)	PAX3/NCOA1	Rare
Fibrosarcome congénital Néphrome mésoblastique	t(12;15)(p13;q25)	ETV6 (TEL)/NTRK3 (TRKC)	80-90 %
Synovialosarcome	t(X;18)(p11.2;q11.2)	SYT/SSX1	65 %
		SYT/SSX2	35 %
		SYT/SSX4	Rare
Dermatofibrosarcome	t(17;22)(q22;q13)	COL1A1/PDGFB	>90 %
Sarcome stromal endometrial	t(7;17)(p15;q21)	JAZF1/JJAZ1	
Tumeur myofibroblastique inflammatoire	T(2;19)(p23;p13.1)	TPM4/ALK	
	t(1;2)(q21.2;p23)	TPM3/ALK	
	t(2;17)(p23;q23)	CLTC/ALK	
	t(2;11)(p23;p15.5)	CARS/ALK	
	t(2;2)(p23;q13)	RANBP2/ALK	
	Remaniements en 2p23	ALK/partenaire	
Sarcome fibromyxoïde de bas grade	t(7;16)(q33;p11)	FUS/CREB3L2	90 %
	t(11;16)(p11;p11)	FUS/CREB3L1	10 %
Histiocytofibrome angiomatoïde	t(12;16)(q13;p11)	FUS/ATF1	2 cas
	t(12;22)(q13;q12)	EWS/ATF1	1 cas
Sarcome alvéolaire des parties molles	t(X;17)(p11;q25)	TFE3/ASPL	>90 %

(Chibon *et al.*, 2007)

Les translocations sont des anomalies génétiques caractérisées par l'échange de matériel chromosomique entre des chromosomes. Deux causes d'oncogenèse principales ont été mises en évidence suite à une translocation : la dérégulation de l'expression de gènes et la formation d'un gène de fusion (*Figure 6*).



(Thèse d'Emilie ANGOT, 2014)

Figure 6 : Schéma du mécanisme de translocation aboutissant à une protéine de fusion

➤ Dérégulation de l'expression d'un gène

C'est le cas de la translocation $t(17;22)$ dans les dermatofibrosarcomes impliquant les gènes *COL1A1* et *PDGFB*. Le gène *COL1A1* codant pour le collagène de type 1A1 est fortement exprimé en raison de son promoteur.³⁸ La translocation a pour conséquence de placer le gène *PDGFB* sous le contrôle du promoteur de *COL1A1* et donc le sur-exprimer par rapport aux conditions physiologiques normales. *PDGFB* est impliqué dans les voies de signalisation des MAP kinases et de PI3K/AKT/mTOR, stimulant ainsi la croissance cellulaire, la différenciation et la migration^{39,40}. Cette translocation a alors pour effet, *via* la surexpression de *PDGFB*, de transformer des cellules normales en cellules néoplasiques^{38,41}, en augmentant leur prolifération et leurs capacités migratoires, qui sont des caractéristiques des cellules tumorales.

➤ Formation de protéines de fusion

Lorsque la translocation a lieu entre deux gènes et que le cadre de lecture n'aboutit pas à l'apparition d'un codon STOP, cela mène à la formation d'une nouvelle protéine ayant une nouvelle fonction. C'est le cas pour le gène le plus fréquemment

impliqué dans les sarcomes à translocations : *EWS* (Tableau 1). *EWS* est une protéine permettant l'activation de la traduction (domaine en N-terminus) en se fixant à l'ARN (domaine en C-terminus). Après translocation, ce gène conserve son domaine d'activation mais son domaine de fixation à l'ARN est remplacé par celui des gènes *FLII*, *WT1*, *ERG* ou *ATF1* qui permet la fixation à l'ADN. Ceci confère à la protéine de fusion un rôle de facteur de transcription régulant l'expression de certains gènes qui aboutit à la transformation tumorale⁴².

Ainsi l'identification des translocations spécifiques dans les sarcomes a eu un impact majeur sur le diagnostic grâce à leur utilisation comme marqueurs moléculaires.

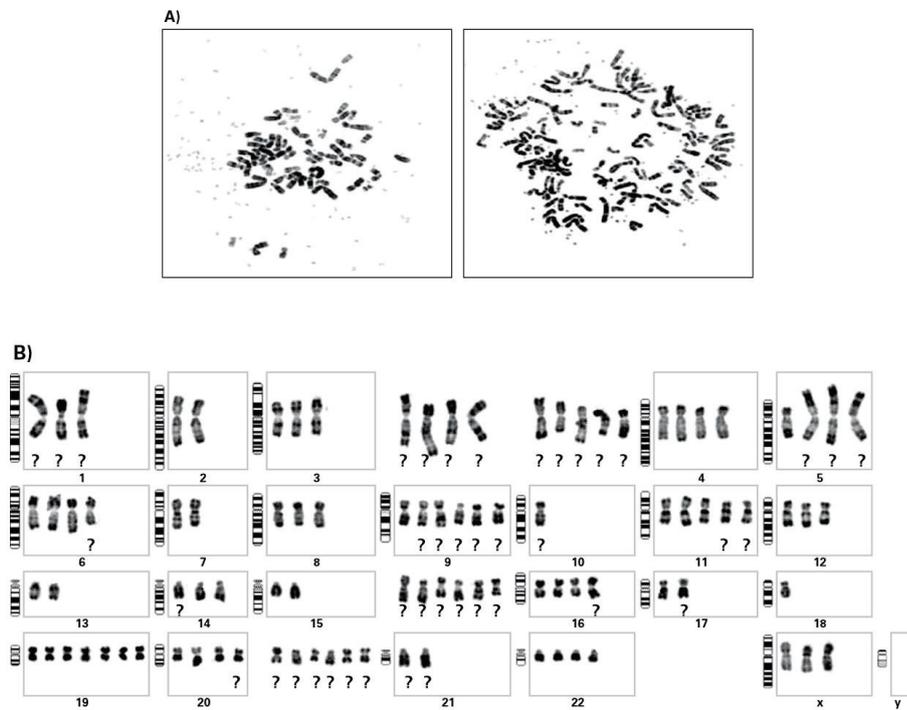
c) Remaniements spécifiques

Le dernier groupe de sarcomes présentant une anomalie spécifique sont ceux ayant une altération quantitative récurrente et relativement spécifique mais pas totalement. L'exemple le plus marquant est celui des liposarcomes bien différenciés et dédifférenciés présentant une amplification du gène *MDM2* dans tous les cas et du gène *CDK4* dans la quasi-totalité des cas (90%). Ainsi, cette anomalie détectable par FISH (Figure 7) a permis de faciliter le diagnostic de ces tumeurs et de re-classifier 20% des sarcomes indifférenciés en liposarcomes dédifférenciés⁴³. Cependant, en raison de la spécificité qui n'est pas complète, cette recherche d'amplification par FISH dans un contexte diagnostique doit nécessairement être associée à une analyse histologique et immunohistochimique pour éliminer formellement un éventuel diagnostic différentiel possédant également cette amplification.

III. Sarcomes à remaniements complexes

Environ 50% des STM ne comportent pas d'anomalie génomique spécifique et sont dits à génomique complexe en raison de leur génome présentant une très grande instabilité. Ces tumeurs sont le sujet de mon travail de thèse ici rapporté. Ce groupe est principalement composé des léiomyosarcomes (LMS), des myxofibrosarcomes (MFS), des sarcomes à cellules fusiformes et pléomorphes indifférenciés (UPS), des

rhabdomyosarcomes pléomorphes (RMSP), des MPNST (*Malignant Peripheral Nerve Sheath Tumors*) et des liposarcomes pléomorphes (LPSP)⁴⁴.



(Chibon et al., 2007)

Figure 7 : Étalement chromosomique et caryotype d'un sarcome à génétique complexe

Les sarcomes à génétique complexe sont caractérisés par une variabilité intra-tumorale importante comme représenté dans la Figure 7A. Deux étalements chromosomiques de deux cellules issues de la même tumeur ont été comparés. On peut notamment constater que ces deux cellules ne présentent pas le même nombre de chromosomes.

Afin d'illustrer l'instabilité chromosomique, le caryotype d'une cellule de sarcome à génétique complexe a été établi (Figure 7B). On peut y dénombrer 97 chromosomes contre 46 pour une cellule normale. De plus, certains chromosomes (notés « ? ») ne sont pas totalement identifiables et sont parfois même inclassables.

➤ Les léiomyosarcomes

Les LMS comptent pour environ 8% à 10% des STM^{1,45}. Ces tumeurs malignes peuvent se développer n'importe où dans l'organisme et présentent des degrés de différenciation de cellule musculaire lisse variables.

D'un point de vue génomique, les LMS présentent un caryotype complexe qui varie d'une tumeur à l'autre⁴⁶⁻⁴⁸. Cependant, des régions fréquemment altérées ont été identifiées comme gagnées (7q36.3, 7q33-q35, 12q13-q15, et 12q23.3), alors que d'autres sont fréquemment perdues (1p21.1, 2p22.2, 6p11.2, 9p21.1, 9p22.1, 14q32.33, 14q32.33 qter et 17pter)⁴³. Nous pouvons également noter de fréquentes amplifications (17p)⁴⁸. Il a été montré que les plus fréquentes altérations telles que la perte du 13q14 (ciblant la région incluant *RBI*)⁴⁹ ou du 10q (ciblant *PTEN*)⁵⁰ étaient corrélées à un plus mauvais pronostic. Nous pouvons également noter que l'activation de la voie PI3K/mTOR, inhibée par PTEN, pourrait jouer un rôle dans le développement des LMS⁵¹.

D'un point de vue clinique, les LMS sont rares chez les enfants et ont tendance à apparaître chez les personnes âgées¹. Les caractéristiques cliniques des LMS dépendent en grande partie de la localisation de ces tumeurs. Par exemple, les LMS se développant au niveau du derme sont souvent petits (<2cm) et sont de bon pronostic alors que ceux se développant au niveau des muscles sont plus grands (> 6cm). Le risque de développement métastatique est de 40 à 45% avec un taux de survie à 5 ans de 64%⁵². Dans le groupe des LMS, nous pouvons également distinguer ceux se développant au niveau du rétropéritoine et de la cavité abdominale qui sont beaucoup plus agressifs, avec un décès des patients dans 70 à 90% des cas à 5 ans⁵².

➤ Les myxofibrosarcomes

Ces sarcomes se développent essentiellement chez les sujets âgés (médiane à 60 ans)^{1,53,54}. Ils partagent de nombreuses altérations avec les LMS et les UPS^{55,56}. Comme précédemment, la perte de *RBI* est très fréquente. Pour ces raisons, mais également du point de vue anatomopathologique, il est difficile de différencier les MFS des UPS. Plus récemment, une étude a montré que deux gènes étaient surexprimés dans les MFS par rapport aux UPS : il s'agit de *GPR64* et *TNXB*⁵⁶.

D'un point de vue clinique, le taux de survie à 5 ans est de 60 à 70% et une rechute locale survient dans environ 50% des cas⁵⁶.

➤ Les sarcomes pléomorphes indifférenciés

Ce groupe rassemble l'ensemble des sarcomes des tissus mous à génomique complexe pour lesquels aucune ligne de différenciation cellulaire n'a été mise en évidence. Ils sont donc diagnostiqués par exclusion^{57,58}. Les remaniements sont semblables à ceux des LMS avec notamment la perte de *RBI* qui est l'altération la plus fréquente dans les UPS^{21,59}. Les altérations de *TP53* sont également récurrentes⁶⁰.

D'un point de vue clinique, les UPS sont des tumeurs agressives avec un taux de décès des patients à 5 ans après le diagnostic de 30 à 50%.

➤ Les rhabdomyosarcomes pléomorphes

Ils représentent 5% des sarcomes pléomorphes et comme les autres sarcomes à génétique complexe, ils présentent de nombreux remaniements non spécifiques. Ils présentent une ligne de différenciation de muscle squelettique.^{61,62} Au niveau clinique, ils sont très agressifs, de grade 3 et ont tendance à être métastatiques^{62,63}. Environ 70% des patients décèdent dans les deux années suivant le diagnostic^{61,62,64}.

➤ Liposarcomes pléomorphes

Ce sont les liposarcomes les plus rares (environ 5%) et ils apparaissent préférentiellement chez les personnes plus âgées (médiane à 55-65 ans)⁶⁵⁻⁶⁷. D'un point de vue morphologique, ils ressemblent beaucoup aux liposarcomes dédifférenciés. Cependant, d'un point de vue moléculaire, la distinction est faite grâce à l'amplification de *MDM2/CDK4* présente uniquement dans les liposarcomes dédifférenciés^{68,69}. Comme pour les autres sarcomes pléomorphes, de fréquentes mutations de *TP53* et délétions de *RBI* (60%) sont observées⁷⁰. D'un point de vue clinique, les LPSP sont des tumeurs de grade 3 qui ont une évolution métastatique dans 30 à 40% des cas et un taux de survie à 5 ans de 40 à 63%^{65,67}.

➤ Les tumeurs malignes des gaines des nerfs périphériques (MPNST)

Ces tumeurs sont définies par leur différenciation ressemblant aux gaines des nerfs (différenciation Schwannienne). Ils représentent 5 à 8% des sarcomes des tissus mous et la moitié d'entre eux apparaît chez des patients atteints de neurofibromatose de type

1. Il n'y a pas d'anomalie spécifique de ces tumeurs, bien que comme précédemment, on observe fréquemment l'inactivation de *RBI* et de *TP53*^{71,72}. La plupart des MPNST sont de haut grade et les patients n'ont un taux de survie à 5 ans que de 34%.

IV. Conclusion

Lorsque les sarcomes présentent une altération spécifique et que les connaissances biologiques permettent de comprendre l'impact de cette altération sur la biologie de la tumeur, cela a une répercussion directe sur le diagnostic ainsi que sur l'efficacité de la prise en charge du patient.

Depuis une vingtaine d'années, la biologie moléculaire ainsi que la cytogénétique ont grandement facilité la démarche du pathologiste dans le diagnostic de certaines tumeurs grâce au développement de marqueurs moléculaires, comme l'amplification *MDM2/CDK4* dans le cas des LPSDD. Néanmoins, au vu de l'absence d'altérations récurrentes dans les sous-groupes des sarcomes à génétique complexe, la compréhension des mécanismes moléculaires ne peut être dissociée de l'analyse anatomopathologique. En effet, l'utilisation de ces anomalies moléculaires doit toujours se faire en complément de la classification histologique car la tumeur est le résultat à la fois des altérations génétiques subies par la cellule d'origine, de sa ligne de différenciation, mais également des pressions de l'environnement exercées sur elle. L'intégration de tous ces aspects permet ainsi d'orienter de façon pertinente à la fois le diagnostic et le traitement qui en découle, mais également les nouvelles pistes de recherche sur le cancer. Il semble que, même si les sarcomes à génétique complexe présentent des lignes de différenciation différentes, les mécanismes moléculaires à l'origine de l'instabilité génomique observée soient communs.

Bien que nous puissions constater l'inactivation quasi-systématique des voies *TP53* et *RBI*, l'inactivation de ces gènes dans des modèles murins provoque l'apparition de sarcomes d'un point de vue histologique mais pas d'un point de vue génétique, puisqu'ils ne présentent pas de remaniements aussi complexes que ceux observés dans les tumeurs humaines. Il semble donc que l'inactivation de ces deux voies soit nécessaire au développement de ces tumeurs mais que l'événement à

l'origine de l'instabilité génomique est indépendante des voies *TP53* et *RBI*. L'instabilité génomique étant associée à un mauvais pronostic, l'identification de l'élément déclencheur du remaniement de la cellule originelle est donc un des défis majeurs à relever dans ce domaine de recherche.

L'objectif de notre équipe est alors de découvrir de nouveaux gènes « driver », soit de l'oncogenèse, soit de la progression tumorale afin de mieux comprendre et mieux traiter ces tumeurs.

Précédemment, les travaux du laboratoire ont permis d'identifier une signature moléculaire de 67 gènes pronostique de l'évolution métastatique des sarcomes : la signature CINSARC. Bien qu'identifiée dans les sarcomes, elle semble également applicable à d'autres cancers, illustrant ainsi le caractère potentiellement universel des mécanismes génétiques et cellulaires de la dissémination métastatique. L'analyse des voies biologiques dans lesquelles les 67 gènes sont impliqués a mis en évidence deux processus biologiques fondamentaux : le contrôle de la mitose et l'intégrité des chromosomes. Il apparaît donc que la dérégulation de ces voies biologiques constitue l'altération commune aux tumeurs métastatiques.

A partir de ce postulat, notre équipe a posé l'hypothèse qu'à défaut de partager des gènes altérés communs, les sarcomes à génétique complexe possèdent des voies biologiques communément dérégulées. A partir de cette hypothèse, nous avons alors développé des stratégies visant à identifier de nouveaux acteurs de l'oncogenèse des sarcomes.

La méthodologie de cette identification a été développée avant mon arrivée en thèse par Céline BRULARD et a été réalisée sur une cohorte de 106 sarcomes à génétique complexe. Pour cette raison, je présente cette méthodologie et les résultats obtenus dans le deuxième chapitre de cette introduction.

Chapitre 2 : Identification de gènes impliqués dans l'oncogénèse ou la progression tumorale des sarcomes à génétique complexe

I- Introduction

Précédemment, notre équipe a identifié une signature (CINSARC) composée de 67 gènes tous significativement associés à la complexité génomique, à l'agressivité et à l'évolution métastatique des tumeurs. L'analyse de ces 67 gènes à l'aide de la base de données Gene Ontology a montré que l'ensemble des gènes annotés été impliqués dans les mêmes processus biologiques : le contrôle de la mitose et l'intégrité des chromosomes. Cette étude avait donc démontré la récurrence de mécanismes biologiques.

Notre objectif est d'identifier dans les sarcomes à génétique complexe de nouveaux gènes d'intérêt impliqués soit dans l'oncogénèse, soit dans la progression tumorale. Or, ces sarcomes présentent de nombreux remaniements sans qu'aucun gène ne soit altéré dans l'ensemble de ces sarcomes. En effet, dans notre cohorte de 106 sarcomes, 2960 gènes en moyenne sont altérés. Par ailleurs, dans ces tumeurs, nous observons de larges régions chromosomiques gagnées ou perdues, comprenant plusieurs gènes. Ainsi, il semblerait que ces larges altérations entraînent des gènes cibles « driver » mais également des gènes proches, dits « passager », qui n'ont probablement pas d'effet dans l'oncogénèse ou la progression tumorale. Pour éliminer les faux positifs, c'est-à-dire les gènes altérés en même temps que le gène d'intérêt, nous avons filtré ces gènes grâce à l'étude de voies biologiques.

L'hypothèse est que, à défaut de partager des gènes cibles dérégulés de manière commune, les sarcomes à génétique complexe partagent des voies biologiques altérées qui peuvent être soit communes à tous les sarcomes, soit communes à des sous-groupes constitués grâce aux annotations cliniques, comme par exemple : le type histologique, la localisation de la tumeur, l'évolution métastatique et le grade FNCLCC. Notre objectif a donc été, dans un premier temps, d'identifier les voies

biologiques dérégulées dans 106 sarcomes pour lesquels nous disposons des données génomiques et transcriptomiques. Puis, dans un second temps, nous avons récupéré l'ensemble des gènes de ces voies biologiques associés aux caractéristiques cliniques.



Figure 8 : Schéma représentant les différents filtres appliqués

À partir des données cliniques, génomiques et transcriptomiques, ce travail bio-informatique repose sur cinq étapes : le traitement préliminaire des données, les filtres génomiques puis transcriptomiques pour sélectionner les gènes altérés, l'identification des voies biologiques dérégulées et enfin la sélection des gènes impliqués. La sélection des gènes altérés est basée sur la présence d'altérations génomiques telles que les pertes ou les gains, identifiés par CGH-array et associés à une répercussion sur l'expression des gènes. Ensuite, les gènes sélectionnés fournissent pour chaque échantillon une liste de voies biologiques altérées. On peut ainsi définir les voies qui sont altérées de façon commune dans l'ensemble de ces sarcomes. Enfin, nous avons pu identifier les gènes « driver », c'est-à-dire impliqués dans au moins une des voies biologiques altérées dans 100% des sarcomes à génétique complexe et dont l'altération est à la fois récurrente et associée à une ou plusieurs annotations cliniques.

II- Matériel et méthode

Les analyses ont été réalisées sur 106 sarcomes à génétique complexe pour lesquels nous disposons des données génomiques, transcriptomiques et cliniques. Cinq échantillons de tissu sain ont été utilisés comme contrôle génomique pour éliminer les régions de CNV (*Copy Number Variation*) constitutionnelles.

a) Les gènes altérés en génomique

Les puces génomiques SNP6 utilisées sont composées de 1,8 millions de sondes qui donnent une information sur le nombre de copies et sur les polymorphismes

génétiqes. Les sondes qui s'hybrident aux chromosomes sexuels, qui sont dans des régions de variations constitutionnelles du nombre de copies, qui ne sont pas situées dans un gène ou qui sont altérées dans au moins deux échantillons de tissu sain ne sont pas conservées. Ainsi, nous avons conservé environ 500 000 sondes correspondants à 12 124 gènes. Le statut de chacune des sondes est analysé afin d'obtenir un statut unique par gène. Au final, 3 statuts sont alors possibles : normal, gagné ou perdu.

b) Les gènes altérés en transcriptomique

Sur les 12 124 gènes étudiés sur le plan génomique, 11 447 gènes figurent sur la puce transcriptomique (U133 plus 2.0). Les 677 gènes ne pouvant être analysés en transcriptomique sont, par défaut, conservés dans la liste des gènes altérés lorsqu'ils ont une altération génomique. La définition d'une surexpression ou d'une sous-expression repose sur une comparaison avec un groupe contrôle. Pour l'ensemble des 11 447 gènes, les groupes contrôles sont composés d'au moins 30 échantillons et sont ceux pour lesquels le statut du gène est dit « normal ». Pour qu'un gène « gagné » soit conservé après le filtre transcriptomique, sa valeur d'expression doit être supérieure à la valeur du troisième quartile des valeurs d'expression des échantillons contrôles. Réciproquement, la valeur d'expression d'un gène qui est « perdu » doit être inférieure à la valeur du premier quartile de la valeur d'expression des échantillons contrôles pour que ce gène soit sélectionné.

c) Les voies biologiques

Une voie biologique est considérée comme altérée dans un échantillon à partir du moment où au moins un des gènes de cette voie biologique est altéré. Pour cette partie, nous avons utilisé les voies biologiques rassemblées dans la base de données Gene Ontology. Ainsi, les 8527 gènes sélectionnés suite aux filtres génomique et transcriptomique de notre étude correspondent à un total de 228 voies biologiques.

d) Les gènes « driver »

Une fois les voies biologiques communément altérées dans tous les sarcomes définies, la liste des gènes de ces voies biologiques est récupérée et seuls les gènes

altérés dans les 106 sarcomes sont sélectionnés. Ensuite, l'enrichissement de l'altération de ces gènes est testé en fonction des caractéristiques cliniques suivantes : métastase, grade FNCLCC, signature CINSARC, type histologique et localisation.

III- Résultats

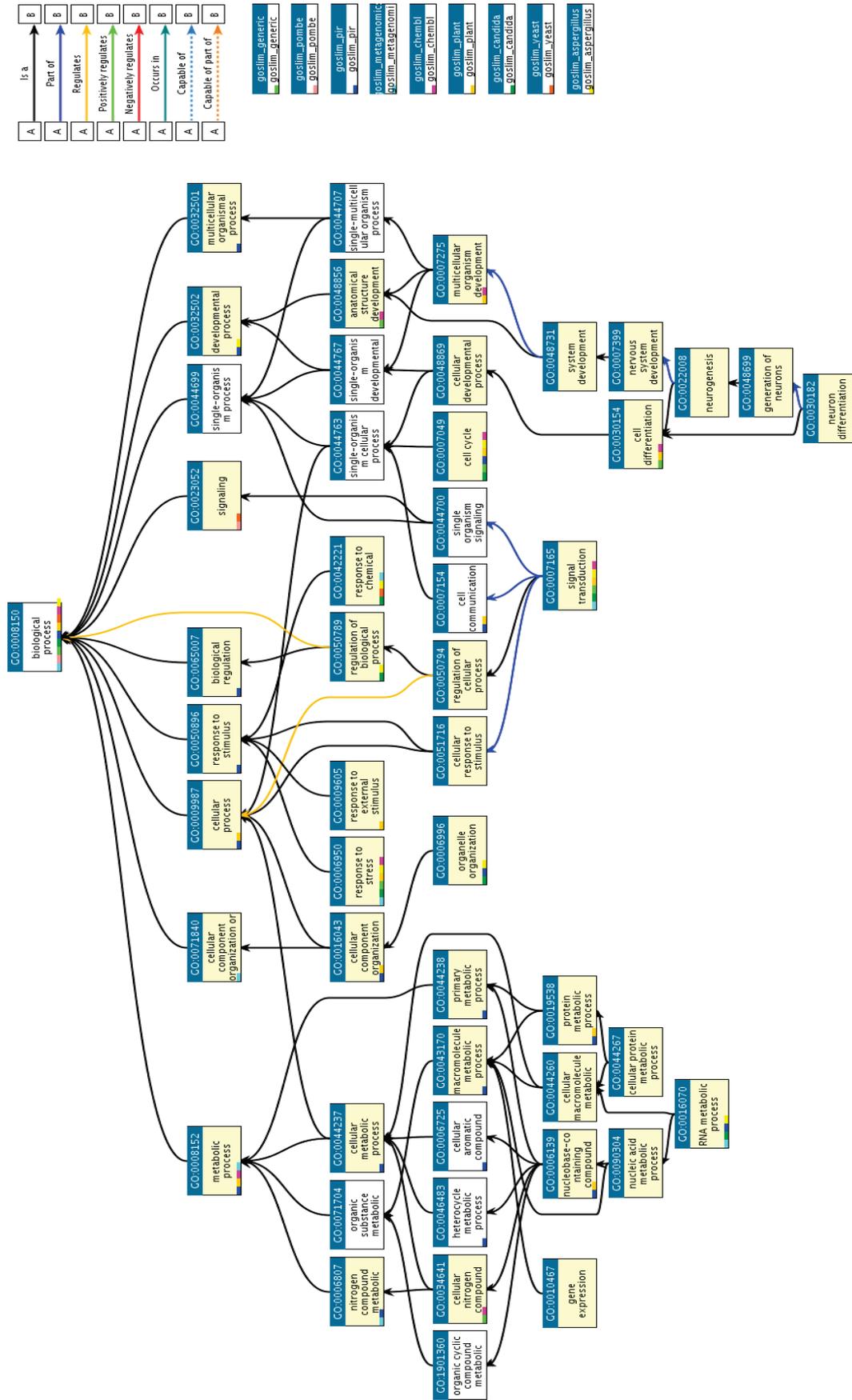
D'un point de vue génomique, les échantillons possèdent entre 30 et 7254 gènes altérés, pour une moyenne de 2960 gènes altérés. Ces tumeurs ne favorisent pas les gains ou les pertes puisque 47% des altérations correspondent à des gènes perdus, alors que 53% des altérations sont des gènes gagnés. En ne sélectionnant que les gènes altérés dont l'expression est modifiée, nous conservons 40% de ces gènes (48% de gènes perdus contre 52% de gènes gagnés), avec une moyenne de 1231 gènes altérés (les tumeurs présentant entre 8 et 3527 gènes altérés). Le test d'égalité des moyennes (test de Student) n'a pas révélé de corrélation significative entre les annotations cliniques et la quantité de gènes altérés dans chacun des échantillons. Les régions chromosomiques 1q21, 19p13,19q13 et 5p13 sont les plus gagnées et les régions 13q14, 10q26, 2q37 et 10q25 sont les plus fréquemment perdues.

a) Les voies biologiques communes

L'analyse des voies biologiques altérées dans l'ensemble des 106 sarcomes à génétique complexe a mis en évidence 41 voies biologiques.

Tableau 2 : Liste des 41 voies biologiques altérées dans les 106 sarcomes

ID	Biological Pathways of GO database :	P-value
GO:0006996	organelle organization	1.26e-88
GO:0007049	cell cycle	8.87e-81
GO:0016070	RNA metabolic process	3.52e-80
GO:0044267	cellular protein metabolic process	6.96e-78
GO:0007399	nervous system development	1.21e-77
GO:0090304	nucleic acid metabolic process	2.05e-77
GO:0030154	cell differentiation	1.29e-76
GO:0006950	response to stress	1.50e-76
GO:0048869	cellular developmental process	1.55e-76
GO:0010467	gene expression	3.13e-76
GO:0048731	system development	7.40e-75
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	5.63e-73
GO:0022008	neurogenesis	1.75e-72
GO:0009605	response to external stimulus	2.09e-72
GO:0007275	multicellular organismal development	6.04e-72
GO:0019538	protein metabolic process	8.07e-72
GO:0048699	generation of neurons	8.60e-72
GO:0048856	anatomical structure development	1.20e-70
GO:0071842	cellular component organization at cellular level	3.35e-70
GO:0016043	cellular component organization	3.35e-70
GO:0030182	neuron differentiation	7.44e-70
GO:0071841	cellular component organization or biogenesis at cellular level	1.31e-69
GO:0071840	cellular component organization or biogenesis	1.31e-69
GO:0042221	response to chemical stimulus	2.71e-69
GO:0034641	cellular nitrogen compound metabolic process	4.16e-66
GO:0032502	developmental process	1.18e-65
GO:0006807	nitrogen compound metabolic process	4.30e-63
GO:0007165	signal transduction	1.13e-60
GO:0023052	signaling	2.52e-57
GO:0051716	cellular response to stimulus	2.00e-52
GO:0032501	multicellular organismal process	6.63e-51
GO:0044260	cellular macromolecule metabolic process	1.92e-47
GO:0050896	response to stimulus	7.55e-41
GO:0043170	macromolecule metabolic process	2.50e-40
GO:0044237	cellular metabolic process	1.79e-30
GO:0044238	primary metabolic process	2.15e-29
GO:0050794	regulation of cellular process	2.97e-27
GO:0050789	regulation of biological process	4.02e-23
GO:0065007	biological regulation	1.43e-19
GO:0008152	metabolic process	1.32e-17
GO:0009987	cellular process	0.04



QuickGO - <http://www.ebi.ac.uk/QuickGO>

Figure 9 : Classification hiérarchique des 41 voies biologiques selon Gene Ontology

Ces 41 voies biologiques peuvent être classées selon la hiérarchie de Gene Ontology. Elles représentent huit groupes majeurs (Figure 9) parmi lesquels le processus cellulaire, le processus métabolique, le processus de développement et la signalisation/régulation. Nous pouvons remarquer que ces voies biologiques sont assez vastes et regroupent de nombreuses fonctions biologiques au niveau cellulaire comme la signalisation, la différenciation et le cycle cellulaire. Ces voies figurent parmi les grandes caractéristiques de la cellule tumorale proposées par Hanahan et Weinberg en 2011 ⁷³ (Figure 10). Les résultats de l'analyse des voies sont donc pertinents quant à la reprogrammation oncogénique de la cellule et semblent concorder avec le fait que seules les voies biologiques avec un ordre hiérarchique important sont altérées de manière commune dans l'ensemble des 106 sarcomes à génétique complexe.



(D'après Hanahan et Weinberg, 2011)

Figure 10 "Hallmarks of cancer", les grandes caractéristiques de la cellule tumorale

b) Les gènes spécifiques de sous-groupes dans les voies biologiques communes

La totalité des gènes des 41 voies biologiques altérées dans l'ensemble des 106 sarcomes a été comparée en termes d'enrichissement en fonction des cinq critères d'annotation clinique :

- Métastase *vs* non métastase
- Grade FNCLCC 2 *vs* grade FNCLCC 3
- Bon pronostic CINSARC *vs* mauvais pronostic CINSARC
- LMS *vs* autres types histologiques
- Localisation interne *vs* externe

Ainsi, nous avons représenté graphiquement l'ensemble des gènes altérés significativement associés aux caractéristiques cliniques précédentes en fonction de leur localisation chromosomique et de la fréquence de leur altération (Figure 11). Chaque barre représente un gène altéré dont la couleur indique à quelle caractéristique clinique il est associé.

Grâce à cette analyse, nous avons pu identifier un gène dont la délétion est significativement associée à deux annotations cliniques : la métastase et les léiomyosarcomes. Il s'agit du gène *RCBTB1* qui est significativement perdu et sous-exprimé dans 62.5% des tumeurs métastatiques et dans 57.6% des LMS. Il est situé au locus 13q14 à proximité du gène *RBI*.

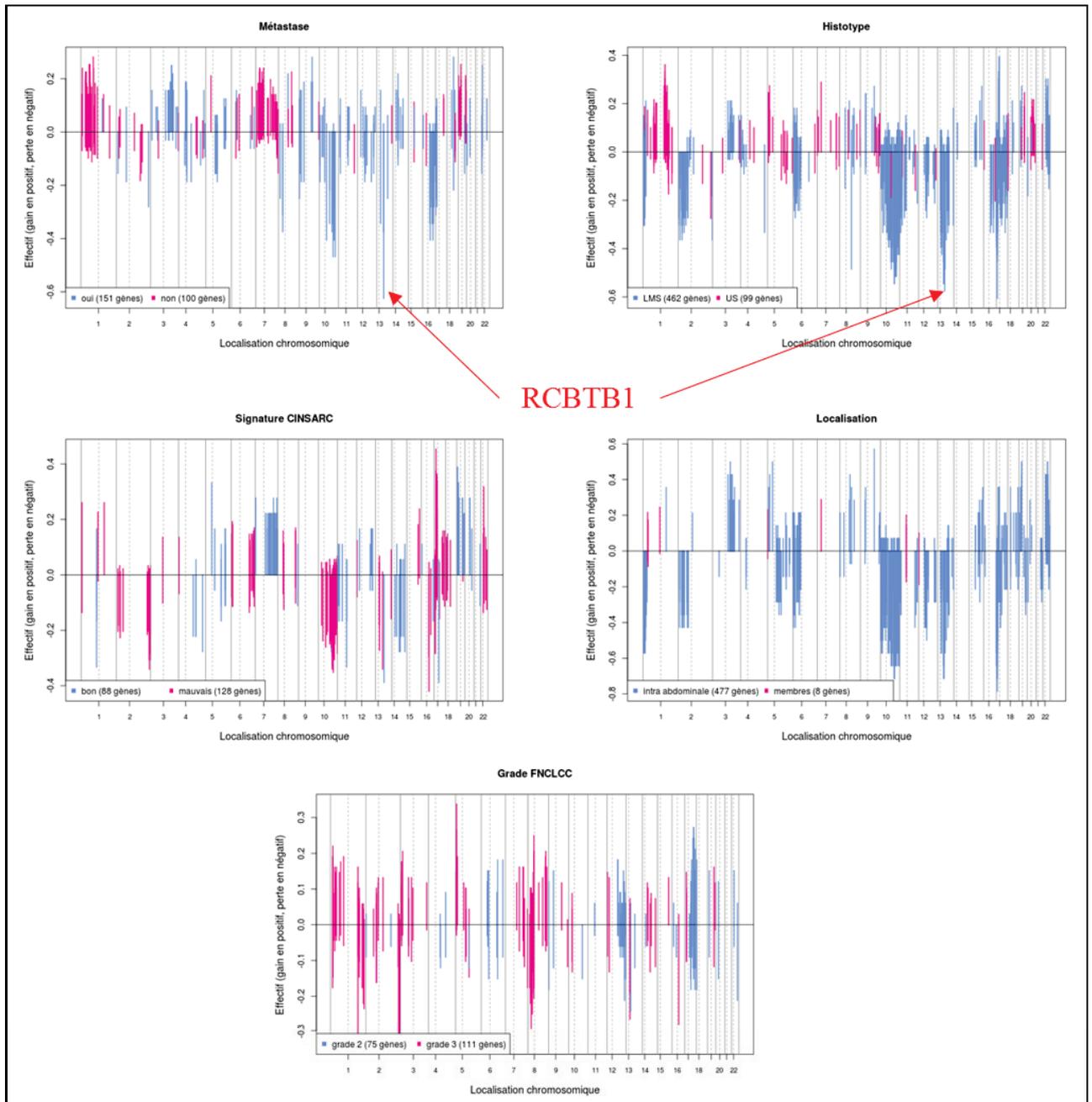


Figure 11 : Fréquence des altérations des gènes en fonction des annotations cliniques

Notre hypothèse est alors que la délétion de RCBTB1 pourrait donc jouer un rôle majeur dans le développement métastatique des LMS.

IV- Etude du chromosome X

Lors de l'analyse précédente, les chromosomes X et Y ont été écartés en raison de la normalisation de l'ensemble des données. En effet, pour automatiser l'analyse, il aurait fallu connaître le sexe de l'ADN témoin contre lequel l'ADN tumoral a été

hybridé. Le chromosome Y a été écarté car, recherchant un gène « driver » de l'oncogenèse des sarcomes chez l'homme comme chez la femme, nous avons fait le postulat qu'il ne pouvait être sur le Y. Nous avons alors entrepris l'étude du chromosome X par CGH-array afin de d'identifier l'altération récurrente d'un gène. Sur les 106 sarcomes à génétique complexe, nous avons ainsi observé une délétion du gène *DMD* dans 17 tumeurs.

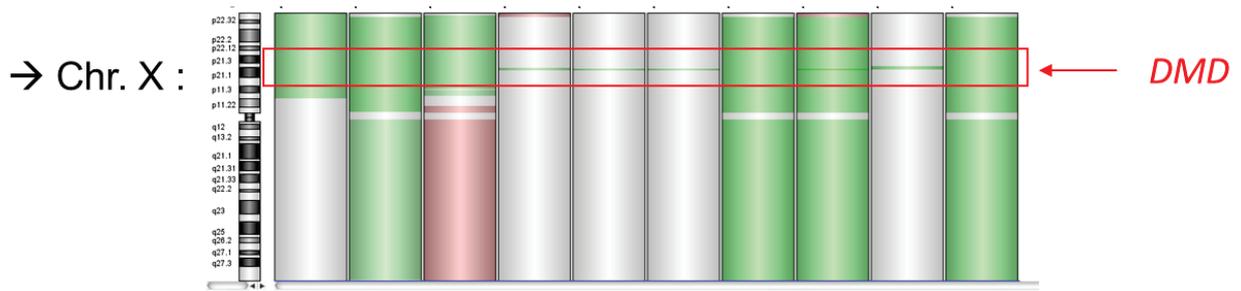


Figure 12 : Différents types de délétion de *DMD* observés par CGH array

Les délétions (représentées en vert) sur la Figure 11, sont de trois types :

- Affectant tout le chromosome X
- Ciblant le bras court du chromosome X
- Ciblant uniquement le gène *DMD*

Nous avons ensuite validé que la délétion du gène *DMD* entraînait bien une diminution de son expression par Taqman (Figure 12), renforçant ainsi l'hypothèse selon laquelle *DMD* pourrait être impliqué dans l'oncogenèse ou la progression tumorale.

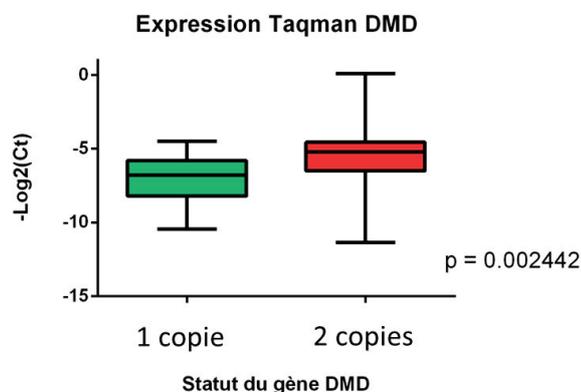


Figure 13: Expression transcriptomique de *DMD* en fonction de son statut génomique

Pour confirmer l'impact de la délétion de *DMD* sur l'oncogenèse ou la progression tumorale, comme précédemment, nous avons cherché si cette altération était associée à une annotation clinique. Nous avons donc dans un premier temps remarqué que cette délétion n'est pas spécifique d'un type histologique particulier, puisqu'elle est présente dans différents sarcomes à génétique complexe et qu'elle est également présente en proportions équivalentes dans les synoviosarcomes et dans les GISTs (Tableau 3).

Tableau 3 : Délétion de *DMD* en fonction de la métastase et des types histologiques

Histotype	Nombre de tumeurs	Non métastatique (212)	Métastatique (105)
GISTs	127	6/89	12/38
Complexes	106	8/71	9/35
SSX	88	10/55	9/33
Total	321	24/215 (11,16%)	30/106 (28,30%)

Nous pouvons également noter que la délétion de *DMD* est significativement associée aux tumeurs métastatiques (Chi2 ; p-value = 0.0002).

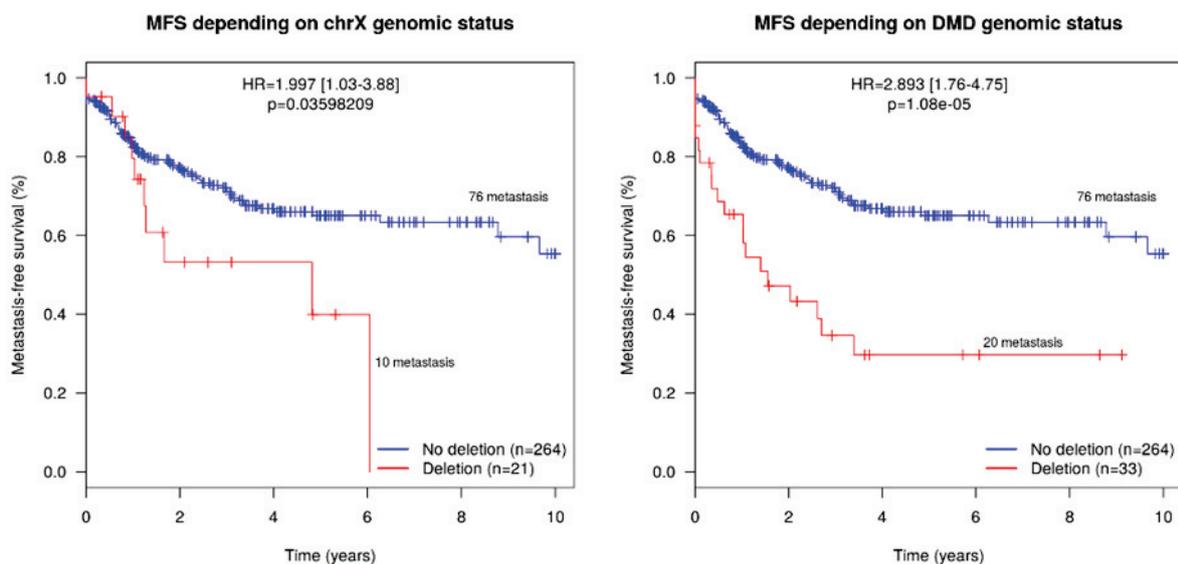


Figure 14 : Kaplan Meier de la survie sans métastase en fonction de la délétion de *DMD*

Nous avons également montré que la délétion de *DMD* est un facteur pronostique de l'apparition de métastase (Figure 13), alors que la délétion du chromosome X ne l'est pas.

En se focalisant sur les délétions ciblant uniquement le gène *DMD*, nous avons observé que seule la région 5' de *DMD* est perdue (Figure 14).

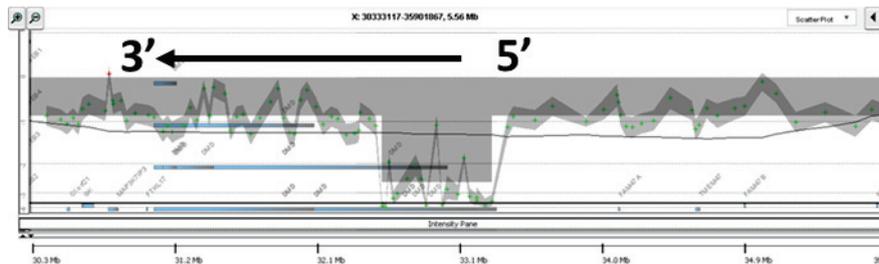


Figure 15: Délétion ciblée du gène *DMD* par CGH array

L'objectif de ma thèse était donc de déterminer le rôle fonctionnel de *RCBTB1* et de *DMD* dans les sarcomes à génétique complexe afin de comprendre comment et pourquoi ces gènes sont associés à l'apparition de métastase.

Pour cette raison, les deux chapitres suivants sont consacrés à l'étude bibliographique de ces deux gènes afin d'appréhender leur rôle au niveau cellulaire permettant ainsi de poser les hypothèses expliquant leur rôle dans l'évolution métastatique.

Chapitre 3 : *RCBTB1* : un gène méconnu

Le gène *RCBTB1* (Regulator of Chromosome Condensation (RCC1) and BTB (POZ) domain containing protein 1), aussi appelé *CLLD7*, *CLLL7*, *E4.5*⁷⁴ et *GLP*⁷⁵, est un membre de la famille RCC1. Selon la base de données *GeneCards*, il s'agit d'un gène très conservé parmi les Vertébrés dont l'expression est ubiquitaire dans le corps humain, ce qui suggère un rôle essentiel de *RCBTB1* dans la cellule animale. De plus, des mutations et des délétions du gène *RCBTB1* sont associées à des pathologies telles que les rétinopathies⁷⁶ et le cancer⁷⁷. Bien que le rôle physiologique de *RCBTB1* demeure mystérieux, du fait du nombre réduit d'études sur le sujet, cette protéine serait impliquée *in vitro* dans la régulation de l'apoptose, la réparation de l'ADN et la signalisation induite par la fixation d'un ligand sur son récepteur. Les deux principaux domaines RCC1 et BTB qui composent *RCBTB1* participent à différentes interactions, permettant ainsi de l'intégrer à de multiples voies biologiques.

I – Description générale de RCBTB1

a) Le gène *RCBTB1* : structure, expression et régulation

Le gène *RCBTB1* a été découvert en 2001 par Mabuchi et collaborateurs, grâce à l'étude de la région génomique située sur le chromosome 13q14, fréquemment délétée dans les leucémies lymphoïdes chroniques. C'est ainsi qu'ils identifièrent *RCBTB1*, qu'ils appelèrent alors *CLLD7*⁷⁸.

RCBTB1 est un gène de 53 638 bp situé en 13q14.2 et composé de 13 exons. Officiellement, il ne code que pour un unique transcrit (NM_018191) de 4033 bp. Bien qu'un autre transcrit soit décrit dans la base de données Ensembl, ce dernier n'est pas mentionné dans la base de données UCSC et de plus, son existence n'a pas été vérifiée de façon expérimentale.

Plusieurs orthologues du gène humain *RCBTB1* ont été décrits parmi les Mammifères, comme par exemple chez le chimpanzé, le singe Rhésus, la souris, le rat, le bœuf, mais également dans d'autres classes de Vertébrés : les oiseaux (Aves), les reptiles (Reptilia) et les poissons à nageoires rayonnées (Actinopterygii). Le fait que ce

gène ait été conservé au fil de l'évolution suggère qu'il assure un rôle fondamental pour la cellule animale.

Le gène humain *RCBTB1* partage 67% d'homologie avec son paralogue *RCBTB2* (aussi appelé *CHCIL*). Le rôle biologique de *RCBTB2* n'a jamais été démontré expérimentalement. Toutefois, des taux d'expression significativement réduits par rapport aux cellules saines ont été observés dans les patients atteints de myélomes⁷⁹ et dans les tumeurs de la prostate⁸⁰.

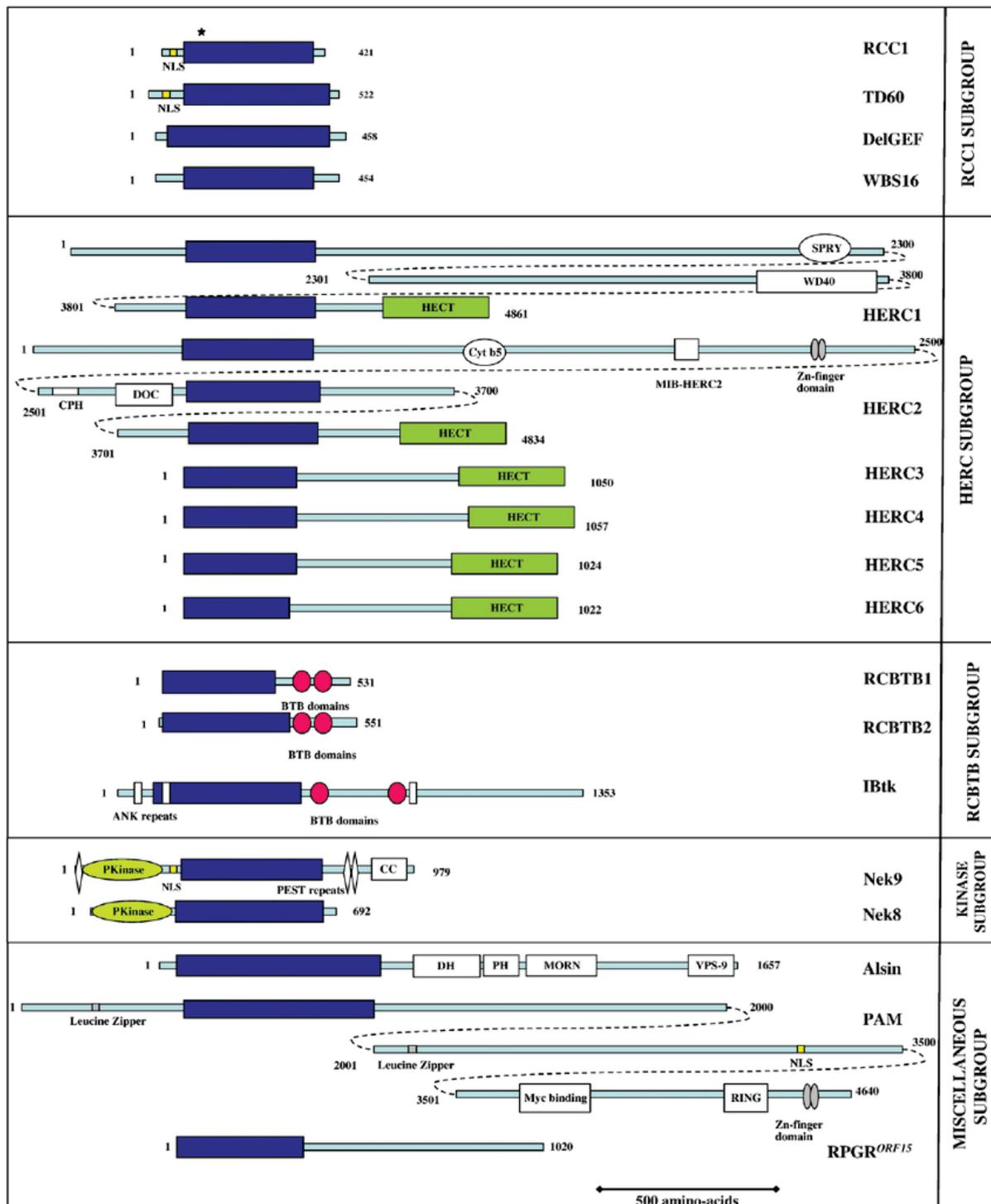
Chez l'homme, *RCBTB1* est exprimé de façon ubiquitaire. On peut constater, grâce aux marquages immuno-histochimiques présents dans la base de données Human Protein Atlas⁸¹, que la protéine RCBTB1 est particulièrement abondante dans la vésicule biliaire et le rein. Un marquage plus modéré est observé dans les tissus tels que le tractus gastro-intestinal, les hépatocytes, le pancréas exocrine, les glandes surrénales et la prostate. Les travaux de Guo *et al.* (2004) ont également montré que l'ARNm de *RCBTB1* est particulièrement abondant dans le rein, le pancréas et le cœur, et que son expression est stimulée par l'Angiotensine II dans les cellules de tubules rénaux immortalisées⁷⁵. Il s'agit de l'unique étude décrivant un mécanisme physiologique de régulation de l'expression de *RCBTB1*, bien que la base de données Genecards⁸² répertorie plusieurs éléments régulant son expression, que nous ne détaillerons pas ici car n'ont jamais été validés fonctionnellement.

b) La protéine RCBTB1 : structure et localisation

La protéine RCBTB1 est composée de six domaines de répétitions RCC1 en N-terminus et de deux domaines BTB (Broad complex, Tramtrack and Bric-a-brac) en C-terminus.

RCBTB1 fait partie de la superfamille RCC1⁸³, qui rassemble les protéines contenant un domaine de 350 à 500 résidus nommé « RCC1-like domain » (RLD) découvert pour la première fois dans la protéine RCC1 (Figure 16). Le domaine RLD de RCC1 est composé de sept répétitions canoniques de 51 à 68 acides aminés chacune. Bien que certains membres de la famille RCC1, comme RCBTB1, ne

possèdent pas le domaine canonique, leurs RLDs correspondent en séquence à chacune des sept répétitions présentes dans la protéine RCC1.

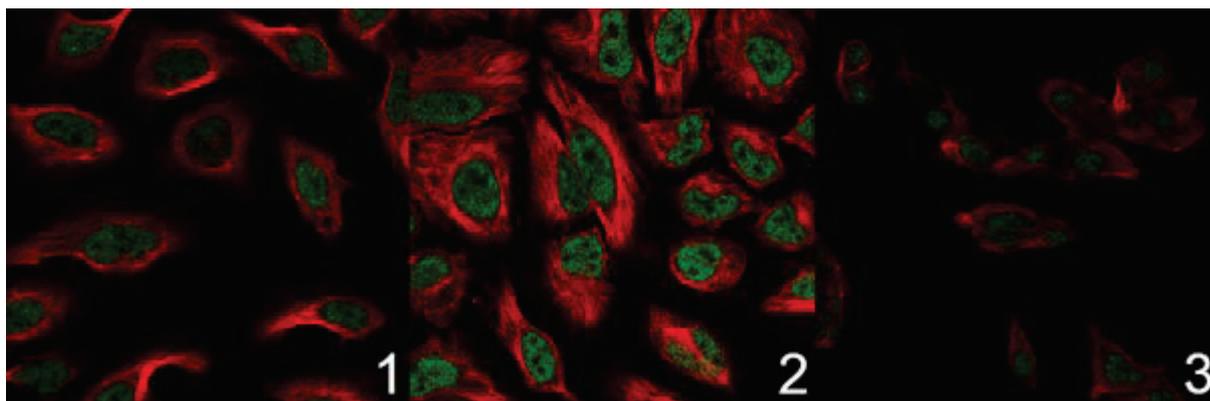


(Hadjebi *et al.*, 2008)

Figure 16 Représentation schématique des membres de la superfamille de protéines RCC1

Le domaine BTB ou POZ (Pox virus and Zinc finger), quant à lui, est présent dans une famille de protéines des poxvirus et dans des protéines contenant le motif Kelch, dans lesquelles il jouerait le rôle d'adaptateur pour les substrats de l'ubiquitination réalisée par la culline 3^{84,85}. Dans la famille RCC1, seul le sous-groupe RCBTB, comprenant RCBTB1, RCBTB2 et IBtk, possède des domaines BTB. Le centre du domaine est constitué d'hélices α entourées de courts feuilletts β en haut et en bas de la molécule. Le domaine BTB/POZ est impliqué dans les interactions protéine/protéine : les dimérisations homomériques et dans certains cas hétéromériques⁸³. Ainsi, l'interaction entre RCBTB1 et RCBTB2 a par ailleurs été validée⁸⁶. De plus, les domaines BTB/POZ de plusieurs protéines en doigt de zinc ont également été démontrés comme impliqués dans la répression de la transcription et comme interagissant avec les composants des complexes co-répresseurs des histones désacétylases⁸⁷.

D'après les annotations électroniques, la protéine RCBTB1 serait localisée au noyau (GO : 0005634 de Gene Ontology). Dans la base de données Human Protein Atlas, certaines lignées de cellules cancéreuses (U2OS, HeLa et HepG2) présentent une localisation nucléaire de RCBTB1 (Figure 17).

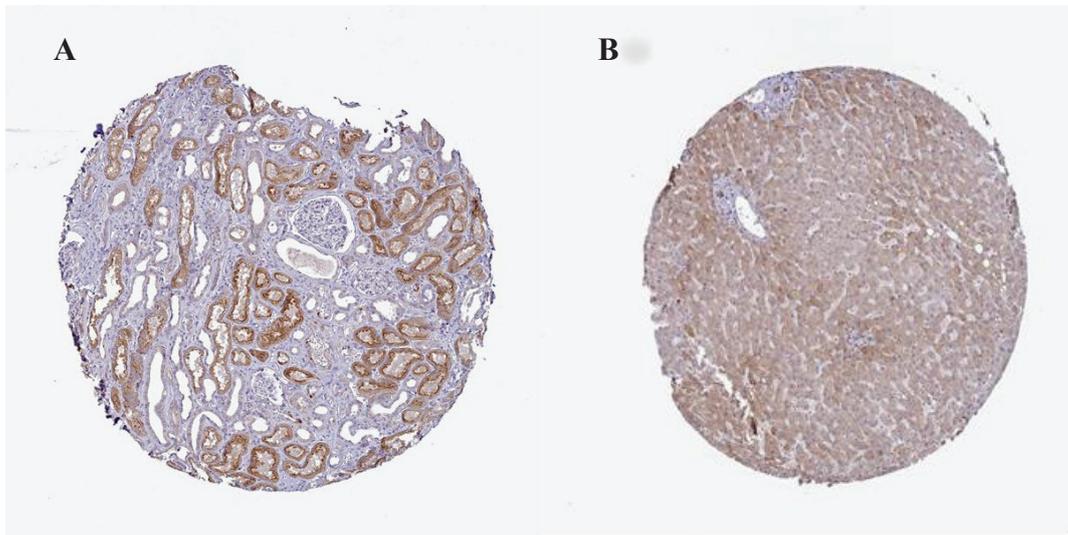


(Human Protein Atlas)

Figure 17 Marquage par immunofluorescence de RCBTB1 (vert) dans des lignées cellulaires cancéreuses : (1) U2OS ; (2) HeLa ; (3) HepG2

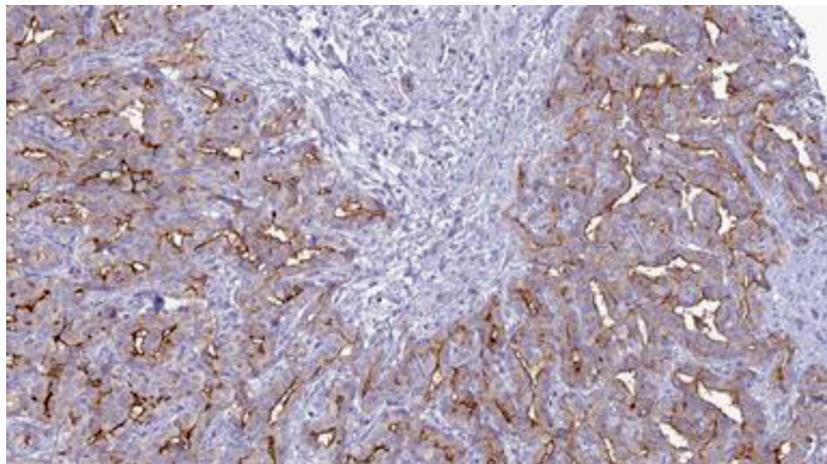
Cependant, il n'existe pas de séquence NLS (Nuclear Localization Sequence) dans la séquence de RCBTB1. De plus, les marquages par immunohistochimie de

tissus sains (Figure 18) ou tumoraux (Figure 19) accessibles sur la base de données Human Protein Atlas montrent une localisation cytoplasmique et sub-membranaire.



(Human Protein Atlas)

Figure 18 Marquage immunohistochimique de RCBTB1 dans du tissu sain : rein (A) et foie (B)



(Human Protein Atlas)

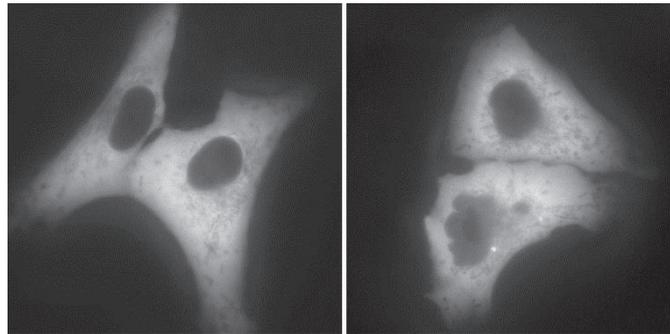
Figure 19 Marquage immunohistochimique de RCBTB1 dans du tissu tumoral hépatique

Par ailleurs, dans les travaux de Guo *et al.* (2004), des cellules HEK-293 exprimant une protéine de fusion RCBTB1-GFP et observées par microscopie à fluorescence suggèrent que RCBTB1-GFP est localisée dans des petites vésicules du

cytoplasme, au niveau de la membrane plasmique, mais qu'elle est absente du noyau

75

Enfin, la localisation cytoplasmique de RCBTB1 est confortée par la base de données LIFEdb^{88,89}, dans laquelle des cellules HeLa et des cellules Vero (cellules rénales de singe) ont été transfectées avec l'ADNc de RCBTB1 fusionné avec celui de la GFP et qui rapporte que dans les deux lignées, RCBTB1 est cytoplasmique (Figure 20).



(LIFEdb)

Figure 20 Localisation de la protéine de fusion RCBTB1-GFP dans les lignées HeLa (A) et Vero (B) observée par microscopie à fluorescence

En conclusion, la plupart des expériences réalisées rapportent une localisation cytoplasmique de RCBTB1 ce qui est concordant avec l'absence de NLS, bien qu'une translocation au noyau puisse avoir lieu dans certains contextes ou modèles.

II – Rôle physiologique de RCBTB1

Dans la base de données Gene Ontology, de par les annotations électroniques, *RCBTB1* figure dans six voies biologiques distinctes :

- Transcription, matrice d'ADN (GO : 0006351) ;
- Régulation de la transcription, matrice d'ADN (GO : 0006355) ;
- Cycle cellulaire (GO : 0007049) ;
- Modification covalente de la chromatine (GO : 0016569) ;
- Noyau (GO : 0005634)

- Cytoplasme (GO : 0005737), qui est la seule issue d'une expérimentation directe, réalisée par LIFEdb.

Selon Gene Ontology, RCBTB1 serait donc impliquée dans la régulation de la transcription et de la condensation de la chromatine. Toutefois, aucune preuve expérimentale ne permet d'étayer ces prédictions bio-informatiques. Néanmoins, ces hypothèses sont appuyées par l'étude de Zhou *et al.* (2010) qui a montré que l'inhibition de l'expression de *RCBTB1* dans des cellules épithéliales était associée à une augmentation des cassures double-brin de l'ADN et que sa ré-expression dans une lignée d'ostéosarcome altérait l'expression de gènes impliqués dans la réparation de l'ADN⁷⁷.

a) Mises en évidences du rôle cellulaire de RCBTB1 *in vitro*

Parmi les fonctions biologiques de RCBTB1 supposées par la nature de ses partenaires, certaines ont pu être vérifiées expérimentalement. Par exemple, des preuves ont été amenées quant à l'interaction de RCBTB1 avec le récepteur de l'Angiotensine II de type 1 (AT1R)⁷⁵, qui est impliqué dans la prolifération cellulaire et l'angiogenèse. Dans leur étude⁷⁵, Guo *et al.* montrent que RCBTB1 régule l'hypertrophie cellulaire induite par l'Angiotensine II non pas en régulant l'internalisation ou la désensibilisation du récepteur AT1R, mais en modulant la voie de signalisation en aval. De plus, les mécanismes moléculaires semblent être dépendants du type cellulaire étant donné que la surexpression de RCBTB1 d'une part active la protéine Akt dans les cellules musculaires lisses vasculaires et d'autre part augmente l'expression de p27^{kip1} uniquement dans les cellules du tube proximal rénal immortalisées, suggérant que dans ces dernières l'effet hypertrophique de RCBTB1 serait dépendant de la régulation du cycle cellulaire⁹⁰.

De plus, il a été montré que RCBTB1 intervient dans la régulation de l'apoptose. En effet, l'inhibition de *RCBTB1* par siRNA dans des cellules épithéliales les protège de la mort induite par la cisplatine⁷⁷. Un rôle de RCBTB1 dans la régulation de l'apoptose suggère que son altération pourrait être impliquée dans certaines maladies.

b) Fonctions possibles impliquant le domaine RCC1

La protéine RCC1 est un facteur d'échange du nucléotide guanine pour la protéine GTP-ase Ran (Ras-related nuclear protein)⁹¹ et de ce fait, elle a un rôle essentiel dans la formation du fuseau mitotique puis dans la reformation de l'enveloppe nucléaire^{92,93}. En tant que régulateur de l'état nucléotidique de Ran au noyau, RCC1 est également impliquée dans le transport nucléoplasmique⁹⁴. Ainsi, il a été supposé que dans RCBTB1, le domaine RCC1 situé en N-terminus assurerait le même rôle. Cependant, les marquages cytoplasmiques et membranaires de RCBTB1 dans les tissus de la base de données Human Protein Atlas, ainsi que l'absence de localisation nucléaire rapportée par Guo *et al.* (2004)⁷⁵ et Zhou *et al.* (2010)⁷⁷, suggèrent que le rôle majeur de RCBTB1 serait cytoplasmique. Il n'est toutefois pas exclu qu'il puisse avoir une activité se rapprochant de celle de RCC1 durant la mitose, après rupture de l'enveloppe nucléaire. Néanmoins, ces hypothèses demeurent invérifiées expérimentalement dans les cellules de Mammifères.

Plus que le partage d'une fonction commune, les protéines humaines possédant un domaine RLD partagent une structure commune qui, en fait, permet une importante versatilité fonctionnelle. En effet, les domaines RLDs participent à de nombreuses interactions qui, par conséquent, les impliquent dans de nombreuses fonctions⁸³. Ces interactions peuvent avoir lieu :

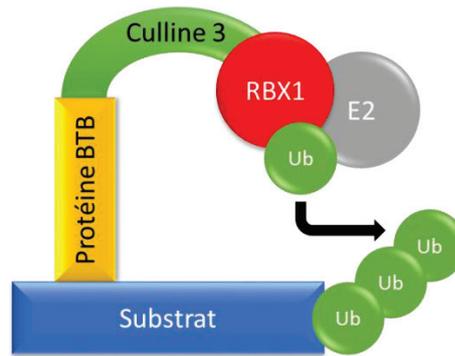
- avec des petites protéines liées au GTP, l'ADN ou encore les histones dans le cas de RCC1 ;
- avec des protéines cibles qui dictent la localisation intracellulaire dans le cas de RPGR par exemple ;
- avec des enzymes comme les adénylates cyclases qui sont inhibées par le domaine RLD de la protéine PAM ;
- ou encore avec des lipides comme l'illustre le domaine RLD1 de la protéine HERC1 qui se lie au phosphatidyl-inositol-bisphosphate (PI(4,5)P₂) afin d'agir comme facteur de libération du GTP (GRF, Guanine nucleotide Release Factor) pour les petites protéines liant le GTP des familles de protéines ARF et Rab.

En conclusion, des études biochimiques et biologiques sont nécessaires pour élucider le rôle du domaine RLD de RCBTB1.

c) Fonctions possibles impliquant les domaines BTB

Les domaines BTB de la partie C-terminale de RCBTB1 réalisent les interactions protéine-protéine et servent très probablement d'adaptateurs pour les substrats des complexes d'ubiquitine-ligase recrutés par la culline 3 ^{86,95-97}.

L'ubiquitination de protéines cibles est réalisée par une cascade de trois enzymes comprenant une enzyme activatrice (E1), une enzyme conjugante (E2) et une ubiquitine-ligase (E3). La plus importante famille d'ubiquitines-ligases E3 est celle des complexes recrutés par les cullines, appelés CRLs (Cullin-RING E3 Ligases), qui sont responsables de l'ubiquitination de 20% des protéines dégradées par le protéasome ⁹⁸. Les CRLs ont été fortement conservés au cours de l'évolution eucaryote et régulent un nombre important de processus biologiques ⁹⁹. Il n'est donc pas surprenant de constater que les CRLs sont fréquemment dérégulés dans certaines maladies, comme le cancer par exemple ^{98,100}. Classiquement, les CRLs sont des complexes protéiques comprenant une culline, une protéine à domaine RING-finger et une ou plusieurs protéines adaptatrices chargées de recruter le substrat pour l'ubiquitination. Contrairement aux sept autres cullines exprimées chez l'homme qui interagissent avec un unique adaptateur de substrat, la culline 3 (CUL3) recrute ses substrats grâce à des polypeptides adaptateurs contenant un domaine BTB ⁹⁶ (Figure 21), dont RCBTB1 ⁸⁶. Les complexes recrutés par CUL3 appelés CRL3 ont été décrits comme ayant un rôle suppresseur de tumeur en régulant la dégradation de substrats impliqués dans le stress oxydatif et le cycle cellulaire ^{98,100}.



(D'après Chen Z. *et al.*, 2015)

Figure 21 La CUL3 recrute une enzyme E2 et une protéine contenant un domaine BTB pour la reconnaissance et l'ubiquitination du substrat

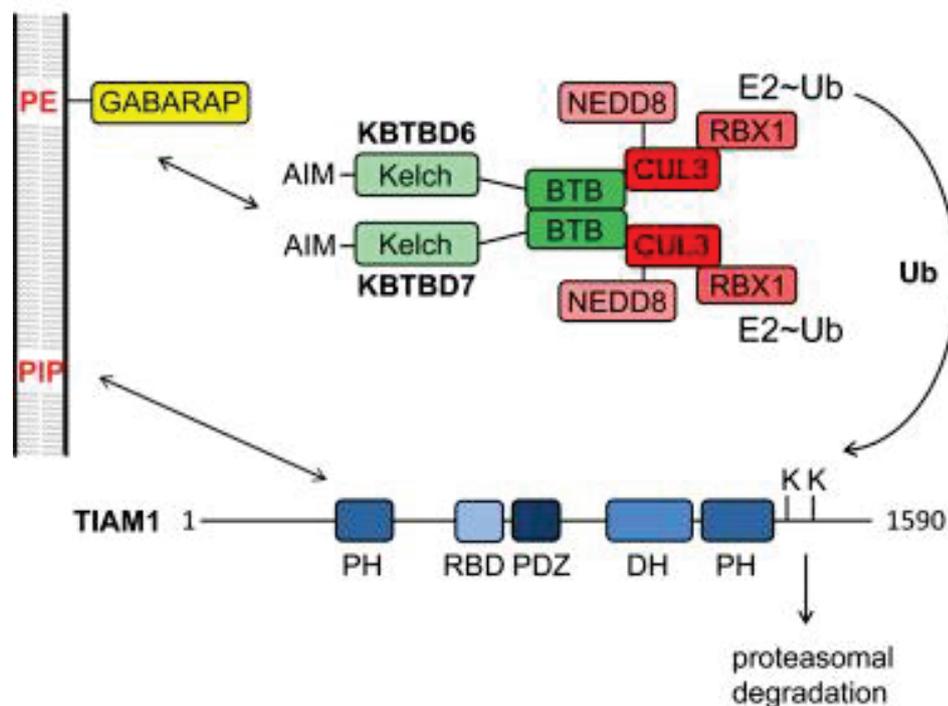
En 2009, Plafker et ses collaborateurs ⁸⁶ ont démontré que CUL3 et UbcM2, l'enzyme conjugante (E2) murine 100% identique à son homologue humain UBE2E3, interagissent toutes deux d'une manière non-canonique : en effet, leur interaction ne requiert pas de protéine à domaine RING-finger mais une protéine adaptatrice à domaine BTB, RCBTB1 notamment. Dans cette même étude, grâce à des expériences de co-immunoprécipitation entre UbcM2 ou CUL3 et des formes mutées ou tronquées de RCBTB1, ils ont pu mettre en évidence que les 87 acides aminés en C-terminus de RCBTB1 permettent son association avec UbcM2 et CUL3 d'une façon non compétitive, et que le domaine BTB soit contribue directement à l'interaction, soit est nécessaire pour la stabiliser. En revanche, la région N-terminale de RCBTB1 semble réguler négativement ces associations. En outre, ils ont montré que RCBTB2, le paralogue de RCBTB1, interagit avec CUL3 et UbcM2 de façon similaire. Au vu de ces observations, les auteurs suggèrent que la perte fréquente de *RCBTB1* constatée dans les cancers hématologiques pourrait engendrer l'accumulation délétère de substrats.

Les protéines ubiquitinylées ne sont pas forcément destinées à être dégradées par le protéasome. Le motif d'ubiquitination participe également à de nombreuses fonctions telles que l'activation de la transcription, le tri endocytaire et la réparation de l'ADN ¹⁰¹. L'association d'une culline avec une enzyme E2 de classe III activée telle que UBE2E3 mène à leur accumulation dans le noyau grâce à la protéine

importine-11¹⁰². Etant donné que l'interaction avec une protéine adaptatrice telle que RCBTB1 est nécessaire pour la formation ces complexes, cela pourrait fournir une explication quant à la localisation nucléaire de RCBTB1 observée dans certaines lignées (Human Protein Atlas) et quant à son rôle dans la réparation de l'ADN proposé par Zhou *et al* (2010)⁷⁷.

Néanmoins, l'identification des substrats ciblés par RCBTB1 a échoué⁸⁶. Ceci pourrait être expliqué par le fait que les phénomènes d'homo- et hétéro-dimérisation sont des caractéristiques communes aux adaptateurs des substrats des complexes CRLs. En effet, celles-ci sont réalisées grâce aux domaines BTB/POZ et cette coopération est supposée importante pour la reconnaissance du substrat⁹⁷. Etant donné que RCBTB1 et RCBTB2 interagissent ensemble et participent aux mêmes complexes protéiques⁸⁶, le fait qu'aucune co-expression des deux protéines n'ait été réalisée pour l'identification des substrats de RCBTB1 peut expliquer cet échec.

La coopération fonctionnelle entre RCBTB1 et RCBTB2 est une hypothèse confortée par plusieurs observations concernant d'autres protéines contenant un domaine BTB. Les protéines BTB sont impliquées dans de nombreux processus biologiques dont la régulation de la dynamique des microtubules et des microfilaments, la transcription, l'apoptose et le cycle cellulaire. Ainsi, Sumara *et al* (2007) ont montré que les protéines KLHL9 et KLHL13 interagissent avec CUL3 pour réguler l'activité de la kinase AuroraB sur les chromosomes mitotiques¹⁰³. Ceci permet la progression fidèle de la mitose et l'aboutissement de la cytokinèse grâce à l'alignement correct des chromosomes lors de la métaphase. Un autre exemple est fourni par les travaux de Genau et ses collaborateurs¹⁰⁴ qui ont montré que les protéines KBTBD6 et KBTBD7 coopèrent pour former un complexe avec CUL3, permettant ainsi l'ubiquitination puis la dégradation du facteur d'échange de guanine (GEF) TIAM1 qui active Rac1, une protéine impliquée dans l'invasion cellulaire et donc dans l'évolution métastatique. Cette dégradation ciblée de TIAM1 nécessite l'intervention de la protéine GABARAP, qui permet l'association du complexe CUL3-KBTBD6/KBTBD7 avec les vésicules adressées à la membrane là où TIAM1 est localisée, permettant ainsi un contrôle localisé de l'activité de Rac1¹⁰⁵ (Figure 22).



(Genau HM. *et al.*, 2015)

Figure 22 *TIAM1* est dégradé suite à la coopération entre *GABARAP* et le complexe *CUL3-KBTBD6/KBTBD7*

De façon intéressante, des interactions entre *RCBTB1* et des protéines de la famille de *GABARAP* sont également citées dans la base de données IntAct¹⁰⁶, renforçant ainsi l'idée selon laquelle les domaines BTB de *RCBTB1* pourraient être impliqués dans la dégradation spécifique de substrats et la modulation de la signalisation intracellulaire.

III- Rôle de *RCBTB1* dans les pathologies

a) *RCBTB1* et cancer

Le gène *RCBTB1* est situé sur le chromosome 13, sur le locus 13q14. Cette région est fréquemment perdue dans plusieurs cancers, incluant les leucémies lymphoïdes chroniques^{107,108}, les lymphomes non-Hodgkiniens^{109,110}, les leucémies lymphoblastiques aiguës¹¹¹, les cancers du poumon^{112,113}, de la prostate^{114,115}, de l'œsophage^{116,117}, de la tête et du cou¹¹⁸ et également dans les sarcomes⁴⁹. Cette région génomique comprend aussi *RBI*, le célèbre gène suppresseur de tumeur, et la

perte d'hétérozygotie (LOH, Loss Of Heterozygosity) de *RCBTBI* est presque aussi fréquente que celle de *RBI*⁷⁷. Ceci est certainement la raison pour laquelle jusqu'à maintenant, l'impact de l'altération de *RCBTBI* sur la transformation oncogénique a été négligé. Cependant, dans le domaine de l'oncologie, l'intérêt est grandissant pour les délétions récurrentes affectant des gènes dits « passagers », car ils offrent une opportunité thérapeutique pour cibler les cellules cancéreuses du fait de leur absence tout en préservant les cellules saines¹¹⁹.

Plusieurs évidences quant au potentiel thérapeutique de *RCBTBI* ont été mises en lumière. La première est apportée par Zhou *et al.*, (2010) qui ont montré que *RCBTBI* est fréquemment sous-exprimé dans les lignées cellulaires cancéreuses et que l'induction de son expression dans une lignée d'ostéosarcome, d'une part, réduit la croissance et la viabilité cellulaires et d'autre part, altère l'expression des gènes impliqués dans la réparation de l'ADN⁷⁷. De plus, Lee *et al* (2013) ont montré que l'inhibition de *RCBTBI* par shARN augmente la résistance des cellules de liposarcome à l'apoptose tandis que sa surexpression la réduit¹²⁰, lors d'un traitement avec le nocodazole, un agent dépolymérisant des microtubules. Dans la même étude, ils signalent que *RCBTBI* est une cible de miR-26a-2, un microARN de mauvais pronostic pour les patients atteints de liposarcome. *RCBTBI* est également régulé par miR-21-3p dans le cancer ovarien¹²¹. En conclusion, toutes ces observations suggèrent que *RCBTBI* est un gène suppresseur de tumeur et qu'il est impliqué dans l'apoptose induite par les chimiothérapies. Bien que les mécanismes moléculaires sous-jacents n'aient pas encore été élucidés, son rôle pourrait être lié aux complexes recrutés par CUL3. En effet, les cullines sont fréquemment dérégulées dans les cancers et apparaissent comme des cibles thérapeutiques attractives⁹⁸ du fait de leur implication dans des processus biologiques essentiels tels que le développement, la régulation des réponses au stress et la mort cellulaire¹²². De façon intéressante, *RCBTB2* est aussi décrit comme un suppresseur de tumeur dans le cancer de la prostate⁸⁰.

b) Maladies cardiovasculaires et rénales

Diverses maladies cardiovasculaires et rénales sont associées à une croissance anormale du tissu (hypertrophie ou prolifération cellulaires) et l'inflammation. La

dérégulation de la signalisation de l'Angiotensine II est impliquée dans ces maladies¹²³. Guo *et al.* (2004) ont démontré que la partie C-terminale de RCBTB1 est essentielle pour l'hypertrophie induite par l'Angiotensine II dans les cellules musculaires lisses vasculaires de rat⁷⁵. Par conséquent, une forte expression de *RCBTB1* pourrait être associée à une signalisation hyperactivée de l'Angiotensine II et donc aux maladies qui y sont liées.

c) Rétinopathies

Les familles souffrant de dystrophies rétiniennes héréditaires, qui sont des maladies génétiques aux manifestations variées qui aboutissent à la perte de la vue par dégénérescence des photorécepteurs et/ou de l'épithélium rétinien, présentent, parmi de nombreuses altérations génétiques, des homozygoties pour des mutations faux-sens du gène *RCBTB1*¹²⁴. D'autre part, les individus affectés ayant des yeux ressemblant à ceux de patients souffrant de la maladie de Coats ou de vitréorétinopathie, qui sont caractérisés par une vascularisation de la rétine altérée, possèdent des mutations hétérozygotes dans le gène *RCBTB1*⁷⁶. Bien que les mécanismes soient encore inconnus, ces observations peuvent être reliées au fait que chez le poisson-zèbre, RCBTB1 est un régulateur de la signalisation de la β -caténine au cours du développement vasculaire de la rétine⁷⁶. Les auteurs suggèrent donc que l'ubiquitination par RCBTB1 pourrait être le mécanisme d'action, étant donné que l'enzyme conjuguante (E2) UbcM2 est abondante dans la rétine¹²⁵ et impliquée dans son développement¹²⁶.

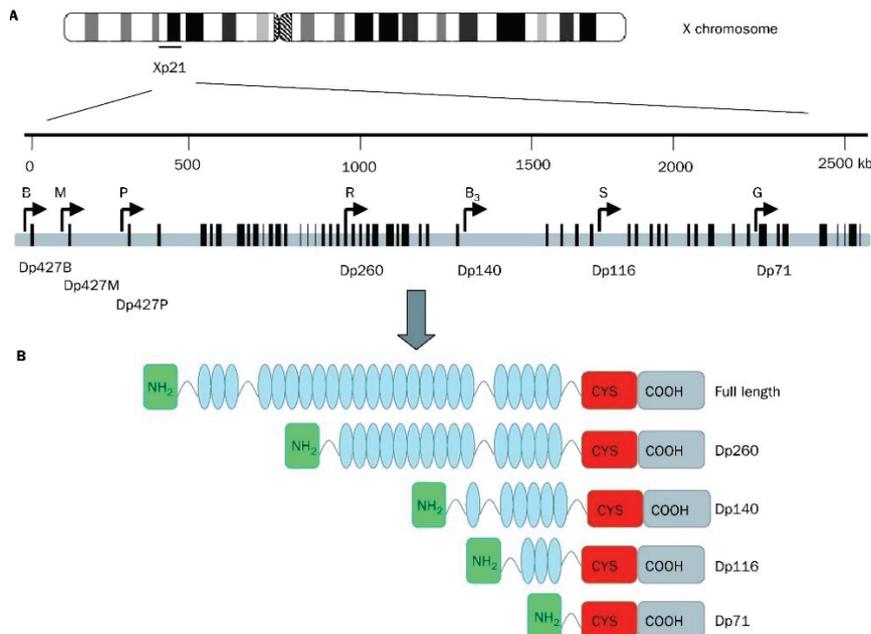
Chapitre 4 : *DMD* : un gène complexe

Dans ce quatrième chapitre, je présenterai les connaissances actuelles concernant *DMD* d'un point de vue génomique puis protéomique. Dans un premier temps, je vais décrire de façon détaillée les domaines de l'isoforme Dp427, qui est la plus grande isoforme codée par ce gène, afin de comprendre sa fonction dans la cellule. Enfin, je m'attarderai en particulier sur la description d'une deuxième isoforme, Dp71, car il s'agit de la seule isoforme épargnée par la délétion ciblant *DMD* (Figure 15) dans notre cohorte de tumeurs et qui, d'après mes travaux, semble jouer un rôle important dans les cellules tumorales.

I- Le gène DMD : Découverte et description

D'une longueur de 2,2Mb, le gène *DMD* est le plus long gène humain. Ce gène doit son nom à la maladie dont il est responsable lorsqu'il est muté : la Dystrophie Musculaire de Duchenne.

Historiquement, c'est en 1868 que le Dr Guillaume Duchenne de Boulogne, médecin français, réalisa la première véritable description clinique de cette maladie. Plusieurs cas avaient été préalablement décrits à partir de 1829, mais c'est le Dr Duchenne qui en donna la description la plus précise, grâce notamment à la première biopsie qu'il a pu réaliser. Plus récemment, en 1955, Dr Peter Emil Becker a décrit une maladie apparentée mais moins sévère, que l'on appelle désormais la dystrophie musculaire de Becker. En 1886, William Richard Gowers constata que dans les familles touchées par cette dystrophie, l'hérédité provenait du côté maternel, suggérant alors une maladie génétique liée au chromosome X. En 1984 et 1985, deux études ont été publiées, montrant que les patients atteints de cette myopathie présentaient des translocations¹²⁷ ou des délétions¹²⁸. Le gène *DMD*, quant à lui, a été découvert en 1987 par l'équipe du Pr. Kunkel par génétique inverse¹²⁹. Ce gène est porté par le chromosome X au niveau de la bande Xp21. La séquence codante, composée de 79 exons, est très complexe car il existe plusieurs promoteurs à l'origine de cinq isoformes principales avec également plusieurs épissages alternatifs (Figure 23).



(Muntoni *et al.* 2003)

Figure 23: Schéma du gène *DMD* codant pour différentes isoformes

(A) Les flèches indiquent la position des promoteurs les mieux décrits le long du gène *DMD* et les barres noires celle des exons ; (B) Représentation des isoformes les plus décrites.

II- Les protéines de *DMD*

La présence de plusieurs promoteurs ainsi que l'épissage alternatif des pré-ARNm conduisent à la production de plusieurs isoformes de dystrophine (Figure 23). Le gène de la dystrophine est à l'origine de protéines de différentes longueurs, ceci grâce à trois mécanismes :

- L'activation de promoteurs alternatifs
- Des phénomènes d'épissage alternatif
- Des sites de terminaison de la transcription alternatifs.

Outre leurs différentes tailles, ces protéines présentent une expression spatiotemporelle variable et sont exprimées dans différents types cellulaires. On utilise l'abréviation Dp pour "dystrophin protein", suivie du poids moléculaire de la protéine et éventuellement d'une lettre pour préciser son expression tissulaire principale.

a) Dp427

C'est le transcrit le plus long du gène *DMD*. Il existe trois promoteurs différents pour l'isoforme Dp427 avec un premier exon unique. Ils sont différenciellement exprimés selon le type cellulaire : Dp427B dans le cerveau, Dp427M dans le muscle et DP427P dans les cellules de Purkinje. Il a également été décrit un autre promoteur noté Dp427L, situé 500Kb en amont de Dp427M, dans les cellules lymphoblastoïdes¹³⁰.

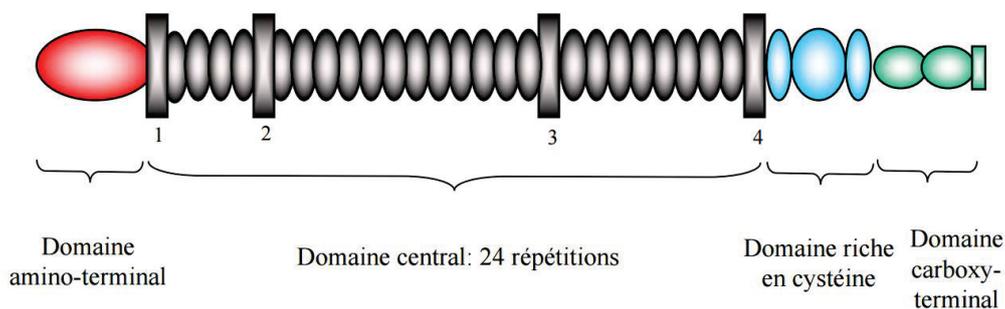


Figure 24: Schéma des domaines protéiques de Dp427

L'isoforme Dp427 est composée de 4 domaines protéiques fonctionnels :

➤ Domaine amino-terminal

Le domaine amino-terminal présente une première séquence d'acides aminés (AA) spécifique dépendant du promoteur qui l'exprime et possède des homologies de séquence avec la région de l' α -actinine¹³¹ et de la β -spectrine qui ont la capacité de se lier à l'actine. Il contient trois sous-unités notées : ABS1, 2 et 3 (pour *Actin Binding Site*), qui sont capables de se fixer à l'actine du cytosquelette membranaire. Le domaine N-terminal contient également deux sites d'interaction avec la calmoduline notés : CaMBS1 et 2 (pour *Calmodulin Binding Site*)¹³².

➤ Domaine central :

Le domaine central (également appelé *Rod Domain*) constitue le corps de la dystrophine. Il correspond à 24 séquences répétées d'une centaine d'AA¹³³. Nous pouvons également noter quatre zones riches en prolines (noté de 1 à 4 sur le schéma), qui confèrent sa flexibilité à la dystrophine¹³⁴.

➤ Le domaine riche en cystéines

Le domaine riche en cystéines comprend 280 acides aminés dont 15 cystéines. Il présente trois motifs distincts : le motif WW, le motif EF et le motif ZZ. L'ensemble de ces trois domaines permet à la dystrophine de se lier au β -dystroglycane¹³⁵.

➤ Domaine Carboxy-terminal

La partie carboxy-terminale de la dystrophine est composée de 420 AA. Ce domaine contient le site de liaison à la dystrobrevine et aux syntrophines¹³⁵.

b) Dp260

L'isoforme Dp260 est transcrite à partir de l'exon 29 de la dystrophine. D'abord identifiée dans la rétine où son altération provoquait des anomalies de neurotransmission¹³⁶, elle a également été mise en évidence dans le cerveau et les intestins¹³⁷.

c) Dp140

Ce transcrit est sous le contrôle d'un promoteur situé en amont de l'exon 45 et produisant un transcrit de 7,5 kb¹³⁸. Ce transcrit est traduit en une protéine de 140kDa qui est exprimée dans le cerveau, les yeux, les muscles squelettiques et l'estomac¹². Son rôle n'est pas clairement identifié, mais il semblerait que son absence dans le cerveau soit associée à une diminution des capacités cognitives chez l'homme¹³⁹.

d) Dp116

Son promoteur situé dans l'intron 55 permet l'expression d'un transcrit de 5,2 kb¹⁴⁰. Cette isoforme a été décrite comme étant exprimée uniquement dans les cellules du système nerveux périphérique¹⁴¹. Elle est localisée au niveau de la membrane des cellules de Schwann¹⁴².

e) Dp71

Dp71 est sous le contrôle d'un promoteur de type « housekeeping », ce qui implique une expression ubiquitaire¹⁴³. Cette protéine garde la capacité de se lier à l'actine¹⁴⁴.

f) Dp45 (ou Dp40)

Ce transcrit de 2,2 kb est sous le contrôle du même promoteur que Dp71. Initialement décrit sous le nom de Dp45, il a été renommé dans les nouvelles études par Dp40. Encore peu décrit, il semble avoir un rôle important au niveau du cerveau^{145,146}.

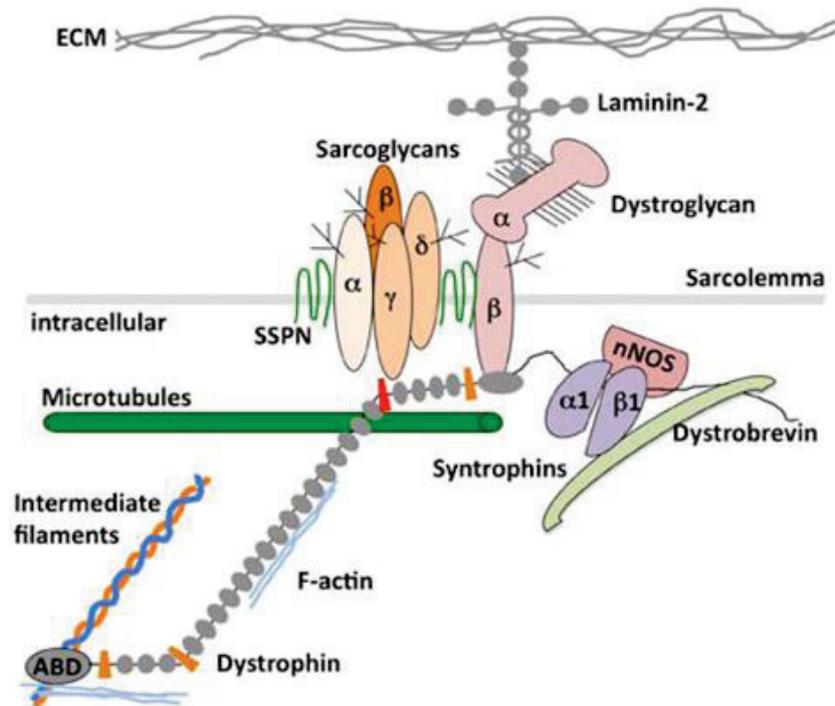
III- Rôles physiologiques de Dp427 et Dp71

a) Dp427

Ce sont les capacités d'interaction de ces quatre domaines qui permettent à la dystrophine d'assurer son rôle dans la cellule. Dans la cellule musculaire, la dystrophine est située sur la partie interne de la membrane, à laquelle elle est ancrée par un complexe (nommé DCG pour *Dystrophin-Glycoprotein Complex*) composé de nombreuses autres protéines (Figure 25)¹⁴⁷.

Le DCG peut être divisé en trois groupes protéiques en fonction de leur localisation :

- Extracellulaire (α -dystroglycane)
- Transmembranaire (β -dystroglycane, sarcoglycane, sarcospane (=SSPN))
- Cytoplasmique (dystrophine, dystrobrevine, syntrophine)



(Gao *et al.*, 2015)

Figure 25 : Schéma du complexe DCG

Les α -dystroglycanes agissent comme récepteurs des ligands extracellulaires tels que la laminine, conférant ainsi au DCG un rôle dans l'adhésion à la lame basale¹⁴⁸. Les α -dystroglycanes sont associés aux β -dystroglycanes par des liaisons covalentes, eux-mêmes liés à la dystrophine par son domaine riche en cystéine¹⁴¹. Au niveau du sarcolemme (membrane plasmique de la cellule musculaire), les sarcoglycanes sont liés aux β -dystroglycanes. Pendant la contraction musculaire, la présence de la dystrophine et de son complexe est importante pour la stabilité mécanique de la membrane cellulaire et la résistance des fibres musculaires à l'étirement.

b) Dp71

➤ Expression

Comme précédemment mentionné, l'expression de Dp71 est ubiquitaire sauf dans le muscle strié squelettique. Néanmoins, le niveau d'expression de Dp71 varie d'un tissu à l'autre et également durant la différenciation. Il a été montré que l'expression de Dp71 augmente durant la différenciation des astrocytes chez le rat¹⁴⁹ et lors de la différenciation des cellules PC12, indiquant clairement le rôle de Dp71 dans le processus de différenciation^{150,151}.

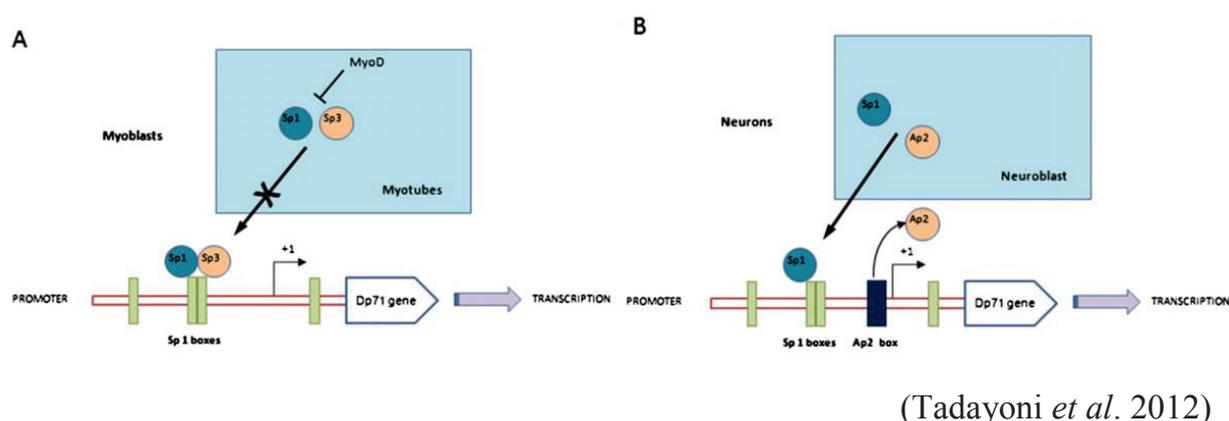


Figure 26 Régulation de l'expression de Dp71 lors des différenciations myogénique et neuronale

Pour comprendre les mécanismes moléculaires à l'origine des différences d'expression entre le système nerveux et le muscle, des chercheurs ont caractérisé en détail le promoteur de Dp71 (Figure 26). Quatre sites de fixation de la protéine Sp1 (en

vert) et un domaine AP2 (en bleu) ont été identifiés. Il a été montré que la région proximale du promoteur contrôle la répression de Dp71 durant la myogenèse aussi bien que son induction lors de la différenciation des cellules neuronales^{152,153}.

Ainsi, dans les myoblastes, l'expression basale de Dp71 est assurée grâce aux facteurs de transcription Sp1 et Sp3. Lors de la myogenèse, le facteur de différenciation myogénique MyoD réprime l'expression de Sp1 et Sp3, entraînant la perte d'expression de Dp71 dans les cellules musculaires matures¹⁵² (Figure 26A). La régulation de Dp71 durant la myogenèse est indispensable car Dp71 participe au remodelage du cytosquelette des myoblastes¹⁵⁴, mais dans les fibres musculaires matures son expression doit être réprimée pour permettre à Dp427 d'être le transcrit prédominant du gène *DMD* et éviter la compétition entre ces deux isoformes pour les sites de liaison au complexe DAPC. En effet, la surexpression de Dp71 dans des cellules musculaires squelettiques de souris donne lieu à un phénotype de dystrophie¹⁵⁵.

Dans les neurones indifférenciés, l'expression basale de Dp71 est régulée positivement par Sp1 d'une part, et négativement par le facteur de transcription AP2 α d'autre part. Dans les neurones différenciés, l'expression de AP2 α est réprimée, laissant ainsi Sp1 activer seule la transcription de Dp71¹⁵³ (Figure 26B). La régulation de l'expression de Dp71 assure sa présence dans des types cellulaires définis et / ou des stades de développement, permettant le bon fonctionnement des processus cellulaires dans lesquels Dp71 intervient. En outre, il semble que le gène *DMD* doive être strictement régulé pour éviter la redondance et la concurrence nuisible entre les différentes isoformes.

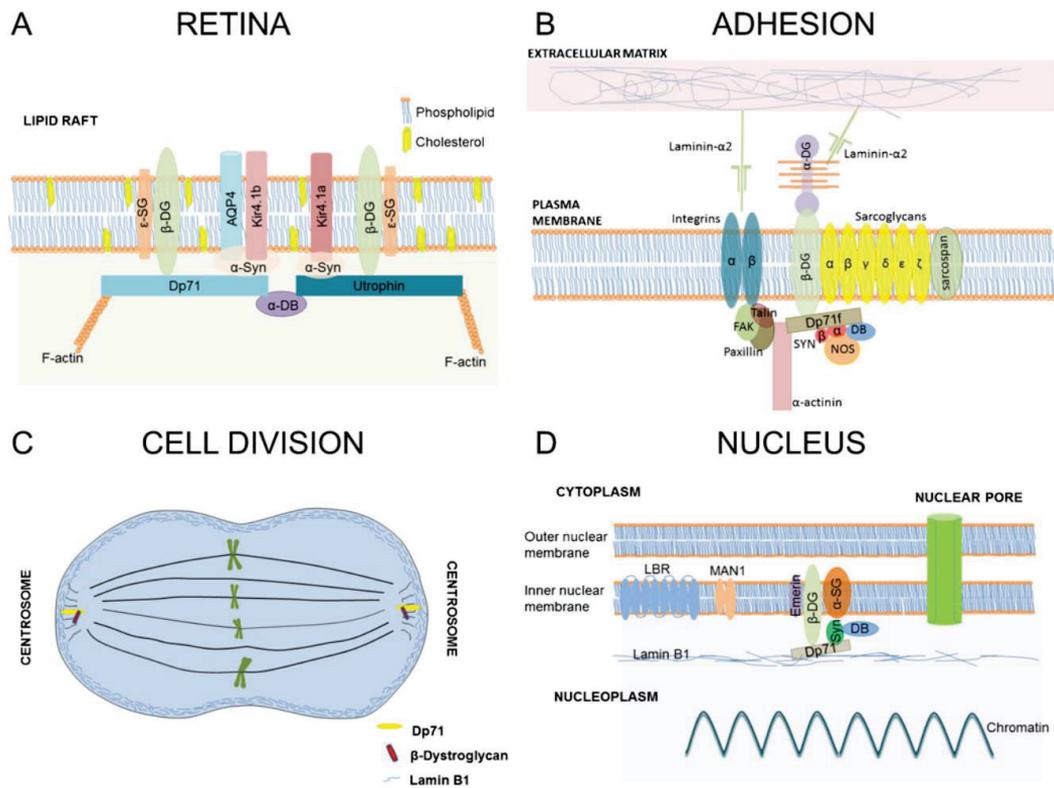


Figure 27: Schéma représentant les différents rôles de Dp71

(Tadayoni *et al.* 2012)

La protéine Dp71, de par sa localisation à la membrane et son association au β -dytroglycane (β -DG) se lie à la matrice extracellulaire via l' α -dytroglycane et permet le maintien de la structure de la membrane cytoplasmique (Figure 27A). Dp71 semble avoir au moins quatre rôles majeurs dans la cellule¹⁵¹ :

➤ Dans la rétine

Dans les cellules gliales de la rétine, les cellules gliales de Müller, Dp71 et le complexe DAP contrôlent l'expression et la localisation des canaux potassiques Kir4.1 et de l'aquaporine AQP4, qui tous deux coopèrent pour la régulation de l'homéostasie de l'eau et du tamponnage des ions potassiques, essentielle pour le bon fonctionnement des cellules gliales de la rétine et le maintien de la barrière hémato-rétinienne¹⁵¹.

➤ Rôle au niveau de l'adhésion

Il a été montré que Dp71 est situé aux frontières des lamellipodes et du complexe DAPC dans les cellules U-373 MG, dans lesquelles il co-localise avec les β 1-intégrine, α -actinine et vinculine. Plus tard, la participation du DAPC contenant Dp71 dans le remodelage des plaquettes par l'actine et le cytosquelette a été révélée¹⁵⁶, favorisant le changement de la forme, l'adhésion, l'agrégation et la centralisation des granules. Ces auteurs ont également révélé que le processus d'adhérence des plaquettes nécessite la participation des microtubules et de l'actine, et que l'interaction entre les deux systèmes cytosquelettiques est réalisée par des protéines β -DG et d'adhérence focale¹⁵⁷. Soutenant le rôle de Dp71 dans les plaquettes, il a été rapporté que l'adhésion au collagène en réponse à la thrombine était significativement diminuée dans les plaquettes isolées à partir de souris *mdx3cv* (altérées pour toutes les isoformes de *DMD*)¹⁵⁸.

Les cellules PC12 ont été utilisées pour définir en détail l'implication de Dp71 dans l'adhésion cellulaire. Il a été démontré par immunofluorescence et des tests d'immunoprécipitation que Dp71 s'associe à différentes intégrines et à l'actine^{159,160}. La liaison de Dp71 au complexe d'adhérence semble être pertinente, car l'inhibition de Dp71 entraîne une défaillance de l'adhérence¹⁶⁰. À l'heure actuelle, il est difficile d'établir comment Dp71 interagit avec le complexe d'adhésion. Puisque l' α -actinine se lie directement à la région terminale C de la dystrophine¹⁶¹, un domaine protéique partagé avec Dp71, il est tentant de proposer que l' α -actinine interagisse avec Dp71 dans les cellules PC12, établissant ainsi une communication croisée entre DAPC contenant Dp71 et la β 1-intégrine (Figure 27B).

Par conséquent, Dp71 semble être un composant structurel du complexe d'adhérence des cellules PC12 qui module l'adhésion. Comme l'adhérence cellulaire médiée par l'intégrine β 1 module la migration neuronale et la synaptogenèse au cours du développement du système nerveux central¹⁶², et parce que les nécropsies de cerveaux de patients atteints de *DMD* présentent des hétérotopies neuronales (présence de neurones en-dehors du cortex)¹⁶³, la participation de Dp71 à la modulation de

l'adhésion cellulaire pourrait expliquer la déficience cognitive associée à la myopathie de Duchenne.

➤ Rôle dans la division cellulaire

Le cycle cellulaire est un processus ordonné, strictement réglementé qui implique de multiples points de contrôle qui évaluent les signaux de croissance extracellulaires, la taille des cellules et l'intégrité de l'ADN. Il a été observé que les cellules PC12 présentant une inhibition de Dp71 présentent un retard marqué dans la croissance cellulaire sans modification de la mort cellulaire, ce qui suggère une défaillance dans le cycle cellulaire¹⁶⁴. L'analyse du cycle cellulaire a révélé un comportement altéré des cellules inhibant Dp71, qui consiste en un retard dans la transition G0/G1. Il a été découvert par microscopie confocale que Dp71 est localisée aux pôles du fuseau mitotique, où il est co-localisé avec la lamine B1 et le β -dystroglycane¹⁶⁵. Ces deux protéines ont été impliquées dans des aspects de la division cellulaire^{166,167}. À la lumière de ces données, les chercheurs ont postulé que Dp71 est une composante du fuseau mitotique et pourrait moduler le cycle cellulaire en se liant à lamine B1 et le β -dystroglycane (Figure 27C). Le rôle de Dp71 dans la division cellulaire PC12 ouvre une nouvelle voie dans l'étude de Dp71¹⁶³.

➤ Dp71 dans le noyau

Comme mentionné précédemment, l'isoforme Dp71 se localise dans le noyau de cellules neuronales cultivées et de différentes lignées cellulaires, formant un complexe nucléaire de type DAPC. Dans ce scénario, les chercheurs du laboratoire du Dr Cisneros ont développé une étude à long terme visant à déterminer la fonction du DAPC nucléaire contenant Dp71. Auparavant, il a été découvert que Dp71 et DAPC sont récupérés dans les fractions de matrice et d'enveloppe nucléaires des cellules HeLa¹⁶⁸ et C2C12¹⁶⁹. L'enveloppe nucléaire est composée de membranes nucléaires internes et externes et du complexe de pores nucléaires. La lamine est une structure fibreuse associée à la membrane nucléaire interne par interaction avec des protéines membranaires et est la composante principale de la matrice nucléaire, qui est un réseau permanent de filaments centraux sous-jacents à des fibres épaisses.

Il a été révélé que la fraction de Dp71 associée à la matrice nucléaire est augmentée au cours de la différenciation neuronale induite par le NGF des cellules PC12¹⁷⁰ et que le knockdown de l'expression de Dp71 provoque des altérations des protéines de l'enveloppe nucléaire comme la lamine B1 et de la localisation de l'emérine¹⁶⁵. Il semblerait que la localisation nucléaire de Dp71 soit dictée par les épissages alternatifs¹⁷¹ ou encore que son transport au noyau soit facilité par sa phosphorylation par la CaMKII (*Ca²⁺ Calmodulin-dependent Kinase II*)¹⁷². Sa translocation au noyau pourrait avoir lieu grâce à son interaction avec le β -dystroglycane, dans lequel une séquence NLS (*Nuclear Localization Sequence*) fonctionnelle a été identifiée¹⁷³.

Il a été proposé que le complexe DAPC nucléaire soit fonctionnellement relié à l'enveloppe nucléaire via l'interaction entre Dp71 et la lamine B1 et, surtout, que Dp71 fonctionne comme une protéine de structure nucléaire qui serait critique pour la stabilité et la localisation appropriée des protéines de l'enveloppe nucléaire. Enfin, il existe un réseau continu entre la membrane plasmique et l'enveloppe nucléaire, auquel le complexe DAPC contenant Dp71 pourrait donc participer afin de permettre la transduction de signaux biochimiques induisant une réorganisation nucléaire.

IV- Rôle dans les pathologies

a) Myopathies de Duchenne et de Becker

La myopathie de Duchenne (DMD) se manifeste très tôt dans l'enfance, entre 3 et 6 ans, par une lordose pathologique et une faiblesse musculaire, notamment au niveau des muscles proximaux et de la partie inférieure du corps. Entre 6 et 11 ans, cette faiblesse s'étend aux muscles du torse et des jambes, amenuisant considérablement et rapidement les capacités de l'enfant à marcher. Le décès des patients survient autour de l'âge de 20 ans, avec trois causes principales : insuffisances cardiaques (30%), respiratoires (25%) et infections pulmonaires (18.6%) (*Thèse de Joe Sarkis, 2012*). La myopathie de Becker (BMD), quant à elle, est associée aux mêmes symptômes que ceux de la myopathie de Duchenne. Toutefois, ils apparaissent plus

tardivement avec moins de sévérité. L'évolution de la maladie est plus lente, avec des patients encore en mesure de se déplacer au-delà de leurs 40 ans.

Les myopathies musculaires de Duchenne et de Becker sont provoquées par des mutations du gène *DMD*, dont un tiers sont sporadiques et environ 65% sont des délétions¹⁷⁴. En général, les délétions affectant le cadre de lecture du gène entraînent une myopathie du type DMD tandis que celles le préservant concernent plutôt les patients BMD, pour lesquels les symptômes sont moins sévères¹⁷⁵. Ainsi, malgré quelques exceptions¹⁷⁶, les patients atteints de DMD ne présentent plus d'expression de la dystrophine tandis que pour ceux de type BMD, la dystrophine est exprimée plus faiblement et/ou sous une forme tronquée avec une fonction altérée. Toutefois, du fait de la complexité de ce gène, une même mutation au sein d'une famille peut avoir des conséquences différentes. Ceci peut notamment s'expliquer par la capacité de certaines cellules à inactiver ou non le chromosome X porteur de la mutation, phénomène ayant été décrit comme influant sur la sévérité des symptômes¹⁷⁷.

b) Retard mental

Le retard cognitif, qui se traduit notamment par un quotient intellectuel (QI) inférieur à celui de la moyenne de la population, est un autre symptôme fréquent chez les patients atteints de DMD. Cette déficience observée chez les enfants est à la fois verbale et non-verbale¹⁷⁸ et serait associée à des mutations ou des réarrangement situés dans la seconde partie du gène et affectant les isoformes Dp140 et Dp71.

Le rôle de Dp71 a été confirmé dans la cognition et son absence se traduirait par un nombre réduit de synapses matures et une morphologie altérée de la zone post-synaptique comme décrit dans des modèles de souris *knock-out* pour Dp71, présentant des déficiences cognitives¹⁷⁹. Une autre observation décisive a pu être réalisée grâce à la génération de souris chimériques¹⁸⁰ : le premier exon de Dp71 a été remplacé par une séquence codant pour une protéine de fusion *β geo* dans des cellules souches de souris, permettant ainsi de suivre l'activité endogène du promoteur de Dp71 au cours du développement grâce au gène rapporteur de la β -galactosidase. Ainsi, il a été montré que Dp71 intervient dans la morphogenèse et la différenciation terminale de

plusieurs tissus murins, incluant notamment le système nerveux périphérique et le prosencéphale, qui donne naissance à l'hippocampe.

Dans le cerveau, le complexe DAPC contenant Dp71, de par son rôle d'échafaudage protéique, interviendrait dans la stabilisation de la membrane plasmique et la signalisation transmembranaire. Ceci expliquerait les altérations de la fonction synaptique et de la plasticité cérébrale qui participent au retard mental observé dans les cas de myopathie. Son rôle décrit dans l'homéostasie de l'eau et des ions potassiques dans les cellules gliales de la rétine s'étend également aux astrocytes périvasculaires, l'impliquant ainsi dans la barrière hémato-encéphalique¹⁸¹ et indirectement dans la fonction neuronale.

c) Rétinopathie

Dans les souris ayant le gène rapporteur de la β -galactosidase sous le contrôle du promoteur de Dp71¹⁸⁰, un autre organe est également marqué au cours de leur développement : les yeux. Il a aussi été démontré qu'en l'absence de Dp71, la mort neuronale est augmentée suite à un épisode ischémique¹⁸². Ainsi, le rôle de Dp71 au niveau de la barrière hémato-rétinienne présenté dans la partie précédente explique les électrorétinogrammes anormaux¹⁸³, un autre symptôme non-musculaire des patients atteints de DMD¹⁵¹ qui présentent dans 88% des cas une perturbation de la neurotransmission rétinienne¹⁸⁴.

d) Cancer

Au-delà de leur implication dans les myopathies, les produits du gène *DMD* intéressent de plus en plus les scientifiques dans le domaine de la cancérologie. En effet, bien que les cas rapportés chez l'homme soient rares¹⁸⁵, il a été démontré que les modèles de souris pour la myopathie de Duchenne (souris *mdx*, dans lesquelles *dmd* est inactivé par mutation dans l'exon 23) développaient spontanément des rhabdomyosarcomes¹⁸⁶ et des sarcomes mixtes^{187,188}. L'impact de l'inactivation de *DMD* sur l'oncogenèse des sarcomes myogéniques a été confirmée et étayée par Wang et ses collaborateurs en 2014¹⁸⁹, qui constatent une délétion intragénique de *DMD* dans 63% des tumeurs myogéniques de haut-grade (29 GISTs, 4 RMS et 7 LMS). Dans

cette même étude, ils montrent que l'expression d'une construction raccourcie de Dp427, qu'ils nomment *miniDMD* et à qui il manque les exons 17 à 48, réduit la migration cellulaire, l'invasion cellulaire, la formation d'invadopodes et l'indépendance d'ancrage. Les délétions observées ciblent le chromosome X actif et préservent l'expression de Dp71, dont l'inhibition a pour conséquence *in vitro* l'arrêt de la croissance des lignées cellulaires. Toutefois, bien que les dystrophines semblent avoir un rôle prépondérant dans les sarcomes myogéniques, leur impact sur le développement tumoral ne semble pas se limiter à ces derniers. En effet, le rôle suppresseur de tumeur de Dp427 a également été démontré dans le mélanome : 5.5% des lignées de mélanomes présentent une délétion de *DMD* et son expression est totalement perdue dans 69% d'entre elles¹⁹⁰. Dans ce modèle également, l'inhibition de *DMD* augmente les capacités invasives des cellules d'une part, et sa réexpression réduit les capacités de migration et induit la sénescence d'autre part. La perte d'expression de Dp427 a été constatée dans d'autres modèles de cancer comme les tumeurs cérébrales¹⁹¹, dans lesquelles l'expression de *DMD* a été étudiée du fait de son appartenance à un site fragile commun (CFS, *Common Fragile Site*). Un CFS est une région caractérisée par une profonde instabilité génomique et donc fréquemment sujette aux cassures d'ADN. Ces sites sont décrits comme des points chauds des réarrangements chromosomiques lors des phases initiales de l'oncogenèse, qui aboutissent à l'amplification d'oncogènes ou à l'inactivation de gènes suppresseurs de tumeurs¹⁹². Une étude bio-informatique exhaustive menée sur les données de puces d'expression et de RNA-seq des bases de données GEOdatabase et cBioPortal a montré que les altérations de *DMD* étaient aussi fréquentes que celles de certains gènes suppresseurs de tumeurs comme *RBI* et donc pertinentes dans des tumeurs humaines non myogéniques¹⁹³. De plus, la mutation de *DMD* est de mauvais pronostic dans certains cancers comme ceux du sein et de l'utérus¹⁹³.

Toutefois, parallèlement à ces observations, ils ont également pu constater que *DMD* était surexprimé dans les leucémies lymphoïdes aiguës, les carcinomes rénaux, les astrocytomes et les épémdymomes¹⁹³. Bien que les mécanismes sous-jacents soient inconnus, cela peut être expliqué par la complexité du gène et des différentes isoformes pour lesquelles il code. En effet, Dp71 semble avoir une fonction versatile

dans les cancers. Une faible expression de cette isoforme est corrélée à un mauvais pronostic dans les adénocarcinomes gastriques¹⁹⁴, tandis que son inhibition réduit l'agressivité de la lignée d'adénocarcinome pulmonaire A549 *in vitro* et *in vivo*¹⁹⁵ et réduit la prolifération de cellules de RMS¹⁸⁹.

Au vu de l'ensemble de ces observations, et étant donné nos résultats de CGH-array et de RNA-seq, le but de ma thèse est de déterminer les rôles respectifs de chacune des deux isoformes Dp427 et Dp71 dans l'oncogenèse des sarcomes à génétique complexe. En effet, nous avons pu constater que dans nos cohortes, les altérations de *DMD* ne se limitent pas aux sarcomes myogéniques mais s'étendent à plusieurs sous-types histologiques, suggérant un mécanisme commun parmi les sarcomes des tissus mous fortement remaniés au niveau génomique.

Objectifs :

Lorsque j'ai débuté ma thèse, les délétions des gènes *DMD* et *RCBTB1* avaient été identifiées comme fréquentes et significativement associées à l'évolution métastatique dans les sarcomes à génétique complexe. Mon objectif principal était alors de comprendre par quels mécanismes l'altération de ces gènes peut favoriser la progression tumorale.

La bibliographie nous indique pour *RCBTB1* des rôles éventuels dans le cycle cellulaire et la condensation de la chromatine. Ainsi, je me suis attaché à étudier l'effet de la surexpression de *RCBTB1* dans des lignées de LMS sur la prolifération, l'apoptose et l'instabilité chromosomique pour comprendre en quoi sa perte pouvait être associée avec l'évolution métastatique.

Concernant le gène *DMD*, son altération semble cibler particulièrement l'isoforme Dp427 et préserver l'isoforme Dp71. Mon objectif était alors de comprendre le rôle de la délétion 5' du gène *DMD* dans l'évolution métastatique. Pour répondre à cette question, j'ai supposé que Dp427 pourrait jouer un rôle de gène suppresseur de tumeur et/ou que l'isoforme Dp71 aurait un rôle oncogène. Ainsi, je me suis attaché à déterminer le rôle spécifique de ces deux isoformes soit dans l'oncogenèse, soit dans la progression tumorale des sarcomes d'origine myogénique.

Matériels et Méthodes

Les protocoles et méthodes d'analyse que j'ai utilisés pour réaliser mes expérimentations dans le cadre du projet sur *RCBTBI* sont présentés dans la section consacrée à l'article soumis, des pages 89 à 95, et en annexe, des pages 178 à 181.

Voici un tableau décrivant les cohortes utilisées pour le projet *RCBTBI* :

	Cohort #1 (106 sarcomas)	Cohort #2 (204 sarcomas)
Median age at diagnosis [95% CI]	63 [59 - 66]	63 [61 - 67]
Median follow-up in months [95% CI]	20.40 [13.80 - 27.07]	29.04 [22.68 - 35.28]
Histotypes (N, %)		
• Leiomyosarcomas	• 33 (31.13)	• 60 (29.41)
• UPS	• 38 (35.85)	• 82 (40.20)
• Myxofibrosarcoma	• 23 (21.70)	• 0
• Pleomorphic liposarcomas	• 7 (6.60)	• 0
• Dedifferentiated liposarcomas	• 0	• 62 (30.39)
• Others	• 5 (4.72)	• 0
Gender (N,%)		
• Male	• 55 (51.89)	• 103 (50.49)
• Female	• 51 (48.11)	• 101 (49.51)
Metastasis (N,%)	32 (30.19)	83 (40.69)
Local recurrence (N, %)	23 (21.70)	49 (24.02)
FNCLCC grades (N, %)		
• Grade 1	• 8 (7.55)	• 12 (5.88)
• Grade 2	• 25 (23.58)	• 69 (33.82)
• Grade 3	• 68 (64.15)	• 103 (50.49)
• Unknown	• 5 (4.72)	• 20 (9.80)
Tumor size (cm) [95% CI]	9 [8 - 10] (n=103)	10 [10 - 11] (n=192)
Location (%)		
External Trunk	• 88 (83.02)	• 151 (74.02)
• Trunk wall	• 25 (23.59)	• 31 (15.20)
• Head and neck	• 1 (0.94)	• 2 (0.98)
• Extremities	• 62 (58.49)	• 118 (57.84)
Internal Trunk	• 15 (14.15)	• 53 (25.98)
Unknown	• 3 (2.83)	• 0 (0)

La partie « matériel et méthode » correspondant au projet *DMD* est développée en annexe, des pages 181 à 185.

La cohorte de GISTs et de synoviosarcomes utilisée dans le projet *DMD* pour déterminer le pourcentage de délétion est présentée ci-dessous :

	Cohort #1 (127 sarcomas)	Cohort #2 (88 sarcomas)
Median follow-up in year [95% CI]	3.64 [3.11 – 4.17]	29.04 [22.68 - 35.28]
Histotypes (N, %)		
• GISTs	• 127 (100)	• 0 (0)
• Synoviosarcomas	• 0 (0)	• 88 (100)
Gender (N,%)		
• Male	• 64 (50.39)	• 48 (54.54)
• Female	• 63 (49.61)	• 40 (45.46)
Metastasis (N,%)	38 (29.92)	33 (37.50)
Local recurrence (N, %)	13 (10.23)	16 (18.18)

La cohorte analysée par RNA-seq est décrite dans le tableau suivant :

	Cohort #1 (145 sarcomas)
Median follow-up in months [95% CI]	50.92 [29.96 – 66.60]
Median age at diagnosis in years [95% CI]	65 [63 –67]
Histotypes (N, %)	
• GISTs	• 12 (8.28)
• LMS	• 37 (25.51)
• LPSDD	• 20 (13.79)
• UPS	• 43 (29.66)
• Others	• 33 (22.76)
Gender (N,%)	
• Male	• 75 (51.73)
• Female	• 70 (48.27)
Metastasis (N,%)	49 (33.79)
Location (%)	
External Trunk	• 101 (69.65)
• Trunk wall	• 26 (17.93)
• Extremities	• 75 (51.72)
Internal Trunk	• 44 (30.35)

Résultats

I- *RCBTB1* : Rôle dans la progression tumorale

A) Contexte

Comme je l'ai indiqué dans l'introduction, les connaissances moléculaires sur certains types histologiques de sarcomes ont permis d'améliorer le diagnostic et le traitement de ces tumeurs. A l'heure actuelle, aucune altération spécifique n'a été identifiée dans les sarcomes à génétique complexe et ils forment l'un des sous-groupes ayant le plus mauvais pronostic. Nous nous sommes donc attachés à identifier des gènes « driver » de ces tumeurs afin soit de comprendre l'oncogenèse de ces tumeurs, soit d'identifier de nouvelles cibles thérapeutiques. Dans cet objectif, grâce à une nouvelle approche bio-informatique intégrant l'analyse des voies biologiques altérées dans l'ensemble des sarcomes à génétique complexe de notre étude (voir introduction page 45), nous avons identifié le gène *RCBTB1*.

La fonction précise de *RCBTB1* au niveau cellulaire n'est pas encore clairement établie. Au vu de nos résultats préliminaires, c'est-à-dire l'association significative entre la délétion de *RCBTB1* et l'évolution métastatique, nous pouvons poser trois hypothèses non exclusives :

- La perte de *RCBTB1* confère à la cellule des capacités d'agressivité accrues pour le développement tumoral telles que la prolifération, la migration cellulaire ou encore l'invasion.
- *RCBTB1* jouerait un rôle au niveau de la sensibilisation à l'apoptose induite par certaines chimiothérapies. Ainsi, les tumeurs ayant une perte de *RCBTB1* répondraient moins bien aux chimiothérapies et développeraient donc des métastases.
- *RCBTB1* joue un rôle au niveau de la condensation de l'ADN par son domaine RCC1 et sa perte pourrait provoquer une instabilité génomique qui serait donc associée à un mauvais pronostic.

Lors de ma thèse, je me suis attaché à tester ces hypothèses et donc à déterminer le rôle de *RCBTB1* dans des lignées cellulaires de fibroblastes et de sarcomes afin de comprendre comment ce gène peut être associé à la métastase. Les résultats concernant

les deux premières hypothèses ont fait l'objet d'une publication actuellement en revue dans le journal *Clinical Cancer Research* et m'ont amené à réaliser des xenogreffes chez la souris, afin de les confirmer *in vivo*. Les résultats sont présentés en dernière partie de ces résultats.

Concernant la dernière hypothèse, selon laquelle la perte de *RCBTBI* jouerait sur l'instabilité génomique, ce travail n'a pas été publié mais est présenté à la suite de l'article.

Dans le papier suivant, nous fournissons des éléments montrant que la plus fréquente des altérations dans ces tumeurs (perte partielle ou totale du chromosome 13 dans 50% des cas), ne cible pas uniquement le gène *RBI* mais également le gène *RCBTBI*. Nos résultats montrent que la délétion de *RCBTBI* est pronostique de l'évolution métastatique. Il semblerait que cela soit dû à la sensibilisation des cellules à certaines chimiothérapies, plus particulièrement le docétaxel, lorsque *RCBTBI* est surexprimé. En effet, nous avons pu montrer qu'une surexpression de ce gène augmentait le pourcentage de cellules en mitose or, le docétaxel agissant uniquement en mitose, ceci explique la sensibilisation des cellules au docétaxel. Enfin, nous avons vérifié les effets de la sensibilisation au docétaxel *in vivo*, en étudiant la croissance tumorale après traitement en fonction de l'expression de *RCBTBI*.

B) Résultats

***RCBTB1*: A novel prognostic factor for Sarcomas metastatic evolution acting through treatment-induced apoptosis**

Mauduit Olivier^{1,2,3}, Brulard Céline^{1,3}, Lesluyes Tom^{1,4}, Delcroix Vanessa^{1,4}, Pérot Gaëlle^{1,3}, Choublier Nina^{1,5}, Michaud Mickael^{1,3}, Baud Jessica^{1,3}, Lagarde Pauline¹, Aurias Alain¹, Coindre Jean-Michel^{1,3,4}, Lartigue Lydia^{1,4}, Blay Jean-Yves⁶, Chibon Frédéric^{1,3}.

1 Inserm U1218, Bergonié Cancer Institute, F-33076 Bordeaux, France.

2 Claude Bernard Lyon 1 University, F-69622 Villeurbanne, France.

3 Department of Pathology, Bergonié Cancer Institute, F-33076 Bordeaux, France.

4 University of Bordeaux, F-33000 Bordeaux, France.

5 University of Orléans, F-45100 Orléans, France

6 Department of Pathology, Léon Bérard Center, F-69003 Lyon, France.

CORRESPONDING AUTHOR: Frédéric Chibon (f.chibon@bordeaux.unicancer.fr)

Inserm U1218, Institut Bergonié, 229 cours de l'Argonne, F-33076 Bordeaux

Phone: +33 556 330 443

KEYWORDS: Leiomyosarcoma, Sarcoma with complex genomics, *RCBTB1*, Docetaxel, cancer

RUNNING TITLE: *RCBTB1* expression is prognostic for metastatic evolution in sarcomas

FINANCIALSUPPORT: This work was supported by INSERM and Integrated cancer research site (SIRIC) of Bordeaux (BRIO-Bordeaux Oncology research) and the association "Au fil d'Oriane"

CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

WORD COUNT:

NUMBER OF FIGURES: 4 – **NUMBER OF TABLES:** 1 – **NUMBER OF REFERENCES:** 40

Abstract

Purpose:

Sarcomas are rare malignant tumors that form a heterogeneous group with numerous histotypes and molecular subtypes. Among them, sarcomas with complex genomic profiles displaying no specific and recurrent genetic alterations represent 50% of all soft tissue sarcomas, and poorly respond to systemic treatment. It is therefore needed to improve our understanding of the mechanisms involved in response to chemotherapy and ultimately find new therapeutic targets.

Experimental design:

One hundred and six sarcoma samples were analyzed by CGH and cDNA arrays in order to establish both genomic and transcriptomic data from each sample. Given that gene alteration recurrences are very rare, we hypothesized that oncogenesis is driven by recurrent pathway alterations among the tumors. Consequently, aCGH profiles of these 106 sarcomas were analyzed to identify pathways with altered genes (gains and losses) in each sample, consecutively filtered out according to matching expression patterns. This analysis highlighted *RCBTB1* gene which is lost and down-regulated in 62.5% of metastatic tumors. We have assessed the impact of *RCBTB1* overexpression on the biology and phenotype of LMS cell lines.

Results:

Response to chemotherapy according to *RCBTB1* expression was confirmed in a retrospective study of two cohorts (with or without chemotherapy) of 40 and 115 patients with sarcomas with complex genomics. The study revealed that high expression of *RCBTB1* is significantly associated with better metastasis-free survival only in the cohort that received chemotherapy ($P=3.072e-02$). *In vitro*, overexpression of *RCBTB1* in LMS cell lines specifically sensitized to docetaxel-induced apoptosis. This observation is explained by a higher proportion of mitotic cells when *RCBTB1* is overexpressed, especially after 24h of docetaxel exposure, with mitotic cells representing 5.4% and 18% in IB112 and IB136 control cell lines, respectively vs 19.1% and 27.2% in *RCBTB1*-expressing cells.

Conclusions:

RCBTBI is commonly altered in sarcomas with complex genetic profiles, especially LMS, and this is associated with a higher risk for developing metastasis while overexpression correlates to a better response to cytotoxic treatment.

Introduction

Sarcomas are a rare group of malignant tumors that arise from mesenchymal tissue. Molecular approaches have described three main genetics in these tumors: reciprocal translocations, specific mutations and complex genomic profiles¹⁹⁶. Soft tissue sarcomas (STS) with complex genomic profiles (that represent 50% of STS) display no specific and recurrent genetic alterations (2). These tumors form a heterogeneous group with numerous histotypes such as leiomyosarcomas (LMS), undifferentiated pleomorphic sarcomas (UPS), myxofibrosarcomas, pleomorphic liposarcomas, pleomorphic rhabdomyosarcomas etc. Patients' outcome is mainly driven by the intrinsic tumor biology and aggressiveness and may be improved by tumor response to systemic chemotherapy. LMS and UPS are frequent histotypes of sarcoma representing 10 to 15% of all sarcoma each. The LMS subgroup, accounting for 8-10% of all STS are tumors associated with poor prognosis, a metastasis risk of about 40–45% and a 5-year survival rate ranging from 10-64% depending on their location¹⁹⁷. UPS account for 15% of STS and are currently diagnosed by exclusion¹⁹⁷. These tumors have a 5-year survival rate ranging between 50-70%, and a metastasis risk of 30–50%¹⁹⁷. The response rate to chemotherapy is in the range from 20% to 60% tumor control rate during first line chemotherapy with doxorubicin, the standard first-line treatment since 40 years¹⁹⁸. For metastatic sarcomas, the combination of

gemcitabine and docetaxel provides a median OS of 18 months (3). Other molecules such as pazopanib¹⁹⁹, regorafenib²⁰⁰ and trabectedine²⁰¹ have also shown efficacy after failure of standard chemotherapy. Since surgery is the current best treatment available against these neoplasms, it is necessary to enhance the efficiency of chemotherapy by improving our understanding of the mechanisms involved in the response to chemotherapy or by finding new therapeutic targets.

These tumors are characterized by a complex karyotype. At the genomic level, there are numerous gains and losses of chromosomes or chromosome regions, mostly variable across tumors¹⁹⁶. Some recurring genetic alterations have been identified in these tumors: the 13q14-21 region is frequently lost in LMS and UPS leading to *RBI* deletion or inactivation²¹. But no specifically altered driver gene has been identified for the oncogenesis of these tumors.

In order to identify potential candidate genes involved in sarcoma oncogenesis, we have developed a bioinformatic approach integrating clinical, genomic and transcriptomic data from 106 sarcomas samples²⁰². Since sarcomas with complex genetic profiles have many alterations, it is difficult to identify genes specifically involved in oncogenesis or chemotherapy resistance. We hypothesized that these tumors may possess altered biological pathways common to all sarcomas or specific to a sub-group (such as histotype, localization, prognosis, clinical evolution). Based on this hypothesis, we only selected altered genes belonging to biological pathways dysregulated in 100% of sarcomas and identified one gene associated with poor patients outcome likely through its role in chemotherapy resistance.

Materials and Methods

Ethics statement

The samples used in this study are part of the Biological Resources Center of Institut Bergonié (CRB-IB). In accordance with the French Public Health Code (articles L. 1243-4 and R. 1243-61), the CRB-IB has received the agreement from the French authorities to deliver samples for scientific research (number AC-2008-812). The project was approved by the Institut Bergonié ethics committee (scientific advisory board). Every case was histologically reviewed by the pathologist subgroup of the French Sarcoma Group and classified according to the 2013 World Health Organization classification by histology, immunohistochemistry and molecular genetics and cytogenetics when needed. (Cohort description in table 1)

1) Chemicals

Doxorubicin (Adriamycin, Pfizer) and gemcitabine (Gemzar; Eli Lilly and Company) were obtained from the pharmacy of the Institut Bergonié (Bordeaux).

Docetaxel (Taxotere; Sanofi-Aventis) was obtained from Sigma Aldrich (01885) as a lyophilized powder that was reconstituted with DMSO (D8418, Sigma Aldrich) at 10mM.

The drugs concentration used during experiments was the IC50 for each cell lines previously determined.

2) Cell Lines

Cell lines were established as previously described ²⁰³. Authentication of cell line has been performed by array CGH and compared with the corresponding original

tumor. LMS cell lines IB112 and IB136 were cultured in RPMI 1640 + GlutaMAX I (Life Technologies, Inc.) supplemented with 10% FCS and 1% penicillin-streptomycin (Life Technologies). HEK-293T cell line was cultured in DMEM + GlutaMAX I (Life Technologies) supplemented with 10% FCS and 1% penicillin-streptomycin (Life Technologies). Cells were grown at 37°C in a humidified atmosphere containing 5% CO₂.

IB112 and IB136 were infected with a lentiviral vector containing the cDNA sequence of *RCBTB1* coupled to HA-tag under the control of CMV promoter (EX-Z7772-Lv120, GeneCopoeia). The control cell line was established with lentiviral transduction of an empty vector (pReceiver-Lv105, GeneCopoeia). VSV-G-pseudotyped lentiviral particles were produced by co-transfection of 293T cells with previous vectors and the compatible packaging plasmids psPAX2 and pVSVg. Cell lines were incubated overnight with lentiviral supernatants in presence of 8 µg/ml polybrene (Sigma H9268). Stably transduced cells were selected with the addition of puromycin (2µg/mL, Sigma P9620) into culture medium. Overexpression was verified by western blotting.

3) Bioinformatics analysis pipeline for genomic and transcriptomic data

Genomic DNA was extracted using the standard phenol-chloroform extraction protocol²⁰⁴. Affymetrix SNP array 6.0 (Affymetrix, Santa Clara, Calif., USA) was used according to the manufacturer's instructions. One hundred and six samples were normalized with the Genotyping console 2.0 software (Affymetrix).

Total RNAs were extracted as described previously²⁰². RNA quality was checked on an Agilent 2100 Bioanalyzer (Agilent Technologies). Samples were then analyzed on Human Genome U133 Plus 2.0 array (Affymetrix), according to the manufacturer's procedures.

One hundred and six samples for which we had genomic and transcriptomic data available were selected for further analysis with home-made Perl (v5.10) or R scripts (2.14) and with the bioconductor package "GO.db".

On the genomic array, probes meeting the following criteria were excluded: probes present on CNV (Copy Number Variations) regions, probes targeting sexual chromosomes, probes that are not localized in a gene and probes that are altered in control samples. In all, 512 055 probes were analyzed, representing 12 124 genes. A unique status was assigned to each gene based on the information gathered by all probes: normal, lost or gained. Among the 12 124 genes studied at the genomic level, 11 447 genes were present on the Affymetrix U133 Plus 2.0 array. For each gene, only one probe was selected and considered as representative of the expression of the gene, based on the Maximum interquartile range (IQRmax) of all probes targeting the same gene. For further analysis, we estimated that a gene is altered when its genomic status is lost or gained and when its expression is under the first quartile or above the third quartile of control group expression, which is calculated on at least 30 samples for which the genomic status of a given gene is normal.

Among the 11 447 genes studied at the transcriptomic level, 8 527 genes are listed in the Gene Ontology (GO). We considered that GO biological pathway is

involved in the oncogenesis process if at least one gene of the pathway is altered. The enrichment of a biological pathway or of an altered gene in a subgroup is tested with a Fisher's exact test. A gene is considered as a candidate "driver" gene if both its biological pathway and its alteration are enriched in a subgroup.

4) RNA sequencing

The process from RNA extraction to final BAM files was previously described²⁰⁵. We used SAMtools and BCFtools (v0.1.19)^{206,207} with custom depths (at least two alternate bases and five total bases) for reporting a candidate variant. These variants were then annotated by ANNOVAR (October 2013)²⁰⁸ with hg19 genome version, transcriptome annotations (November 2013) and following databases of observed variants: dbSNP (v138)²⁰⁹, 1000G project (April 2012)²¹⁰, ClinVar (September 2014)²¹¹ and COSMIC (v70)²¹².

5) Western Blot

Cells were rinsed with ice-cold PBS and lysed for 30 minutes at 4°C in RIPA lysis and extraction buffer (R0278, Sigma) supplemented with a protease/phosphatase inhibitor cocktail (11697498001, Roche). Lysates were pelleted for 10 min at 13 000 x g at 4°C and supernatants were collected for protein quantitation (DC protein assay kit, Biorad). The total proteins of each sample (40 µg) were loaded on gels and separated by SDS-PAGE. Following a transfer onto a PVDF membrane using a dry transfer system, membranes were blocked in non-fat dry milk in PBS-Tween 0.1% and then incubated with the primary antibody: rabbit anti-HA (sc-805 Santa Cruz Biotechnology) at 4°C overnight. After washing, the blots were incubated for 1 h with

a horseradish peroxidase-linked anti-rabbit antibody (Amersham) and processed for chemiluminescent substrate (Amersham ECL Prime detection reagent kit, Sigma) according to the manufacturer's instructions. Signal was detected using Fusion Fx7 (Thermo Fisher Scientific) imaging system. β -actin (A5316, Sigma) was used as a loading control.

6) Apoptosis assay

To perform the assay, 75 000 cells were seeded into 12-well plates with complete culture medium. After 24 h, the medium was replaced and drugs were added to the fresh medium. After 72 h incubation with each drug, cell death was measured by dual Annexin V-FITC / PI staining (kit 556547, BD Bioscience) according to manufacturer's recommendations. Doxorubicin was an exception since its red coloration impaired PI staining, so we analyzed only AnnexinV-FITC staining. Subsequently, 10 000 cells per condition were analyzed by flow cytometry (BD FASCCalibur). Data analysis was performed with FlowJo v10.1 (FlowJo LLC, OR) and Prism6 v6.01 (GraphPad software Inc, La Jolla, CA) softwares.

7) Immunofluorescence

Cells were seeded into 96-well imaging plates (CLS3603, Sigma), incubated with drugs as indicated and fixed with 3.7% formaldehyde for 20 min at room temperature (RT). Then, cells were washed three times with PBS and permeabilized in 0.5% Triton X-100 for 5 min. Nonspecific binding was avoided by blocking with 0.2% gelatin/PBS for 30 min.

a) iCys

Staining was performed using a specific primary anti-HA (sc-805 Santa Cruz Biotechnology) antibody and a fluorescent secondary antibody conjugate (Alexa Fluor 488, anti-rabbit, Molecular Probes) for 1 h at RT. Nuclei were stained with Hoechst 33258 dye (Molecular Probes). Image acquisition was done using the iCysTM laser scanning cytometer (LSC, CompuCyte). LSC generates histograms for cell cycle phases. Chromatin condensation is measured by the max pixel feature, thus allowing discrimination of cells in G2-phase or in mitosis²¹³.

b) Epifluorescence Microscopy

Staining was performed using a specific primary anti-HA (1/200, sc-805 Santa Cruz Biotechnology) antibody and anti-pMPM2 antibody (1/200, 05-368, Merck Millipore) overnight at 4°C and a fluorescent secondary antibody (1/400) conjugate (Alexa Fluor 488, anti-mouse and Alexa Fluor 594, anti-rabbit, Molecular Probes) for 1 hour at RT. Nuclei were stained with Hoechst 33258 dye (Molecular Probes). Cells were analyzed using a Leica DMI8 epifluorescent microscope with appropriate filters. Pictures were captured using a Hamamatsu C11440 CCD camera and the software Leica Application Suite X.

8) Statistical analysis

Metastasis-free survival (MFS) was defined as the interval between diagnosis and the time of distant recurrence or the last follow-up. Survival rates were estimated using the Kaplan–Meier method and compared using the log-rank test and hazard ratios (HR). Descriptive statistics were used to show the distribution of variables in the

population. All statistical tests were two-sided, and $P < 0.05$ indicated statistical significance. All statistical analyses were carried out using the R statistical environment (R version 3.3.2).

Results

Candidate gene identification

One hundred and six sarcoma samples (Table 1) were analyzed by both CGH and cDNA microarray for each sample. We hypothesized that oncogenesis is driven by recurrent pathway alterations shared by subsets of these sarcomas. In order to identify pathways, aCGH profiles of these 106 sarcomas were first analyzed. We identified an average of 2 960 altered genes (with tumor specific CNV) per sample, with similar proportion of gains and losses in tumors: 53% and 47%, respectively. In order to identify driver genes, only an average of 1231 genes per sample (ranging from 8 to 3527 genes) whose expression was significantly modified according to their genomic alteration (upregulation for gains and downregulation for losses) were retained for further analysis.

Using this gene list and GO database²¹⁴, we considered that GO biological pathway is altered in tumor if at least one gene of the pathway is altered. We then selected the 41 biological pathways (Supplementary Table 2) that were altered in all 106 tumors. They consisted of three major groups: cellular process, metabolic process and developmental process which are involved in a variety of cellular functions: signaling, cell organization, acid nucleic metabolism, differentiation and cell cycle (Supplementary Table 1).

All the selected genes involved in these 41 biological pathways were analyzed in light of clinical annotations: metastasis, grade FNCLCC (Fédération Nationale des Centres de Lutte Contre le Cancer)²¹⁵, CINSARC (Complexity INdex in SARComas) signature²⁰², histotype and tumor location. This analysis highlighted *RCBTB1* as the most significantly gene associated with metastatic evolution (Chi2; $P=6.44e-03$) and lost in 62.5% of metastatic tumors (Supplementary Table 2) against 33.8% of non-metastatic tumors (Fig. 1A). *RCBTB1* deletion is also the most frequent alteration significantly (Chi2; $P=3.64e-02$) associated with oncogenesis of LMS (57.6%) compared to others (35.7%) (Fig. 1B).

We tested whether *RCBTB1* may be a prognostic factor for tumor progression. We have correlated the deletion of *RCBTB1* with a significant decrease ($P=1.47e-17$) of its expression (Fig. 1C). The intersection between the curves showing the distribution of *RCBTB1* expression depending on the genomic status (deleted *versus* normal) (Fig.1 D) defines the value of 6.82 that is the threshold between low and high expression of *RCBTB1*. Both *RCBTB1* deletion and low expression levels (Fig. 1E-F) correlate with a significant higher risk for developing metastasis ($P=0.005$; HR=2.69 [1.31-5.55] and $P=0.021$; HR=2.31 [1.11-4.81], respectively). These results on the expression of *RCBTB1* were validated in a second independent cohort (Fig. 1G) of 204 sarcoma patients (cohort #2 described in Table 1), ($P=0.046$; HR=1.61 [1-2.59]) and in 145 sarcomas for which *RCBTB1* expression was determined by RNA-sequencing (Cohort described in Lesluyes *et al.*, 2016)²⁰⁵ (Fig. 1H) ($P=0.037$; HR=1.78 [1.02-3.09]). Thus, in three cohorts: lower *RCBTB1* expression is significantly associated with an increased risk to develop metastasis.

Hypothesizing that *RCBTBI* could act as a tumor suppressor gene we searched for second event mutations, that could occur in addition to deletions detected by aCGH. Among the whole sarcoma RNA-seq cohort (SRP057793), 12 candidate variants were reported in exonic regions. Eleven candidate variants were synonymous and one was a missense variant (NM_018191:p.A24V). However the latter is already known as a common single nucleotide polymorphism since it is already described in the dbSNP database (rs4942848) and in the 1000G project (allele frequency: 0.59). So, we conclude that *RCBTBI* is exclusively down-regulated by deletion mechanisms.

Given that *RCBTBI* deletion is significantly associated with metastatic spread (Fig. 1), we formulated two hypotheses. First, *RCBTBI* deletion could play a role in the acquisition of aggressive tumoral phenotypes such as cell proliferation, invasion or migration, and second, as *RCBTBI* expression is associated with a better response to chemotherapy, we looked for a potential sensitization effect of *RCBTBI* to several chemotherapies. Of note the latter is strengthened by our observation that in the 204 patients cohort, *RCBTBI* prognostic value is indeed restricted to patients receiving chemotherapy (Fig. 2; $P=0.03071$; HR=4.65 [1.01-19.67]).

We thus assessed the role of *RCBTBI* in LMS cell lines in order to determine if *RCBTBI* modifies aggressiveness phenotypes or/and sensitizes to chemotherapy.

One first LMS cell line not expressing *RCBTBI* was selected: IB112, which has a homozygous deletion of *RCBTBI*. *RCBTBI* re-expression (HA-tagged) was induced in this cell line by lentiviral transduction, validated by Western blotting (Fig. 3A) and its effect was tested *in vitro* on cell proliferation, clonogenicity and cell migration. No

significant effect of *RCBTBI* overexpression was noticed on cell proliferation, clonogenicity, or migration (Supplementary Figure 1). These results seem to exclude the hypothesis that *RCBTBI* deletion could play a role in the acquisition of aggressive tumoral phenotypes concerning this cell line.

In order to assess the impact of *RCBTBI* deletion on response to cytotoxics, we compared apoptosis induced by chemotherapeutic agents in LMS cell line overexpressing *RCBTBI* and control cell line. Chemoterapy is not detailed in the cohort used in Fig. 2, to circumvent this lack of information, current drugs for LMS treatment ¹⁹⁸ (doxorubicin, gemcitabine and docetaxel) were screened on IB112 cell line.

Overexpression of *RCBTBI* in IB112 cell line had no significant effect on cell death induced by doxorubicin and gemcitabine, as compared to control (Fig. 3B). However, overexpression of *RCBTBI* sensitized IB112 cells to apoptosis induced by docetaxel ($P < 0.01$) (Fig. 3C). Indeed, in IB112 cell line overexpressing *RCBTBI*, dual Annexin V-FITC / PI staining occurs in 35.8 % (SD \pm 4.4) of cells after 72 h exposure to docetaxel at 10 nM, whereas control cell line shows 22.6% (SD \pm 2.5) of cell death. This result was confirmed in another LMS cell line: IB136 (which has a heterozygous deletion for *RCBTBI*), for which *RCBTBI* overexpression (Fig. 3A) increases average cell death from 29.7% (SD \pm 4.0) to 39.4% (SD \pm 1.9) ($P < 0.01$) after incubation with docetaxel at 500 nM during 72 h (Fig. 3C).

Whereas gemcitabine and doxorubicin interfere with DNA replication during S-phase, docetaxel inhibits microtubule depolymerization during mitosis. Thus,

RCBTB1 appears to be involved in apoptosis induced by microtubule dynamics blockade during mitosis. In order to confirm the link between apoptosis and mitosis when *RCBTB1* is overexpressed, we focused our study on this phase of cell cycle with or without docetaxel treatment. To that purpose, we first used iCys technology that provides a direct measurement of chromatin condensation feature (Fig. 4A), enabling to identify mitotic cells from G2 cells, independently of a specific mitotic marker²¹³. RCBTB1-positive cells were identified due to a green label with a HA-probe. Considering the whole IB112 cell line, we observed 2.8% of mitotic cells without treatment and 8.6% when cells are treated with docetaxel (Fig. 4B). Interestingly, when we focused on the subpopulation expressing high amounts of RCBTB1 (more intense green-labeling defined in R2 in green), we noticed 4.9% and 23.3% of mitosis in control and docetaxel-treated cells, respectively. Therefore, after docetaxel treatment, the percentage of mitotic cells significantly increased ($P<0.001$) in the population that highly expresses RCBTB1 ($P<0.001$). This result was confirmed in IB136 cell line (Fig. 4B). Indeed, the percentage of IB136 mitotic cells is about 1.3% for the global population and 2.5% for RCBTB1-positive cells under control conditions. When docetaxel is added to the medium, we notice 4.5% and 8.6% of mitotic cells for the entire population and cells that highly express RCBTB1, respectively ($P<0.05$). We thus conclude that *RCBTB1* overexpression sensitizes these sarcoma cells to the apoptosis induced by docetaxel treatment by increasing the proportion of mitotic cells. To confirm this result, LMS cell lines expressing an empty vector or RCBTB1-HA were stained with an antibody targeting phosphorylated MPM2 (pMPM-2) (Fig. 4C), a well-known marker of mitosis²¹⁶. In both LMS cell

lines, RCBTB1 (labeled with HA-probe) expression is significantly associated with a higher percentage of mitotic cells ($P<0.01$) (Fig. 4D). Furthermore, this difference is conserved during docetaxel treatment ($P<0.001$). Indeed, after 24 h of docetaxel exposure, mitotic cells represents 5.4% and 18% in IB112 and IB136 control cell lines, respectively vs 19.1% and 27.2% in RCBTB1 overexpressing in IB112 and IB136 cell lines, respectively ($P<0.001$). This suggests that the sensitization effect of RCBTB1 on docetaxel-induced apoptosis is likely due to a higher proportion of mitotic cells.

Discussion

Sarcomas with complex genomic profiles represent a large proportion of sarcomas, including UPS, LMS, myxofibrosarcomas, pleomorphic liposarcomas, dedifferentiated liposarcomas and pleomorphic rhabdomyosarcomas. These rare and aggressive tumors are characterized by a high risk of metastasis, and consequently a poor overall survival. In order to improve *patient care*, it would be important to understand the mechanisms involved in their oncogenesis. For now, genetic alteration analyses are often focused on recurrently altered genes. In turn, single candidate genes are highlighted, but without being integrated into functional biological pathways that would underline the biology of these heterogeneous tumors. For all these reasons, we first identified all the quantitatively altered genes and retained only those that are part of the biological pathways altered in 100% of sarcomas with complex genomics.

This bioinformatic approach, based on pathways instead of genes, allowed us to identify a gene significantly associated with higher risk of metastatic evolution in this group of sarcomas with complex genomics. Here, we report that among the several genes located in the vicinity of the frequently deleted 13q14 band, *RCBTB1* was

filtered out for its impact on the cell cycle pathway (GO:0007049), which is altered in all 106 sarcomas with complex genomic profiles. This *RCBTB1* deletion occurs in about 60% of LMS and is associated with metastatic evolution in sarcomas with complex genetic profiles. Furthermore, we show that *RCBTB1* down-regulation is also significantly associated with metastatic evolution in sarcomas with complex genomics. aCGH and RNA sequencing of *RCBTB1* revealed that this gene is exclusively inactivated by a deletion mechanism. *RCBTB1* is located on 13q14^{78,217} that corresponds to one of the most frequently deleted regions in sarcomas with complex genetic profiles²¹ and comprises *RBI*, which is a well-known tumor suppressor gene among the most frequently altered in those sarcomas²¹⁸. Interestingly, in the cohort of 106 sarcomas, 50% of patients have a deletion that encompasses both *RCBTB1* and *RBI*, but 7% are deleted only for *RBI* whereas 13% carry a deletion that only affects *RCBTB1* (Data not shown). This suggests that *RBI* may not be the only target of the frequent deletions in this chromosomal region. Deletion of this and other chromosomal regions could indeed participate in a larger oncogenic program based on copy number variations that simultaneously affect multiple weak drivers to exert a cumulative effect equivalent to a single potent driver²¹⁹. Yet, there is a lack of evidences about the physiologic role of *RCBTB1*, although experimental evidences are accumulating for a cytoplasmic role of *RCBTB1* related to regulation of cell death^{217,220} and receptor signaling such as AT1R-induced hypertrophy⁹⁰.

Regarding the role of *RCBTB1* *in vitro*, we demonstrated its impact on docetaxel-induced apoptosis in LMS cell lines. Indeed, we observed a significant increase of apoptotic cells following 72 h docetaxel-exposure when *RCBTB1* is

overexpressed. In IB112 cell line, *RCBTB1* deletion is homozygous as opposed to heterozygous in IB136 cell line. Therefore, although very low *RCBTB1* mRNA level was detected in IB136 cell line, it is possible that the residual *RCBTB1* expression explains the lower effect observed in IB136-RCBTB1-HA.

Since docetaxel acts during mitosis by inhibiting microtubule depolymerization²²¹, we next tested if the sensitization was due to an effect of *RCBTB1* on cell cycle and especially on mitosis. So, we observed that high expression of *RCBTB1* increases the proportion of mitotic cells. The *RCBTB1* expression impact on docetaxel-induced apoptosis is consistent with Lee *et al.* study which demonstrated that *RCBTB1* inhibition protects liposarcoma cells from nocodazole-induced apoptosis²²⁰, an inhibitor of microtubule polymerization. In conclusion, the fact that *RCBTB1*-expressing cells were found to be more sensitive to mitotic catastrophe-inducing agents may be explained by the impact of RCBTB1 on mitosis. The next step will be to explore the predictive role of RCBTB1 expression on response to gemcitabine and docetaxel (GD) combination in the randomized study comparing gemcitabine alone (G) vs GD²²². One can hypothesize that GD and G will be equivalent in sarcomas harboring *RCBTB1* deletion whereas GD will be more effective on sarcomas in which *RCBTB1* expression compared to G.

Little is known about RCBTB1 function and partners. This protein is composed of two functional domains: a Regulator of Chromosome Condensation (RCC1) and a Broad complex, Tramtrack, and Bric-a-brac (BTB) domain²¹⁷. Analysis of RCBTB1 network using the STRING database (Supplementary figure S2A-B) highlights two

major nodes of putative or experimentally verified interactions: Ran (Ras-related nuclear protein) and UBC (Ubiquitin C).

The group of interactions that is connected to Ran may involve the RCC1 domain of RCBTB1. Among the 21 members of the RCC1 superfamily²²³, the best characterized is RCC1 itself. It is a nuclear guanine-nucleotide exchange factor for Ran²²⁵ and so, it has a critical role in mitotic spindle formation followed by nuclear envelope reassembly^{225,226}. The intracellular localization of RCBTB1 is still debated^{78,217} but no Nuclear Localization Signal (NLS) has been identified in *RCBTB1* sequence. According to our results, RCBTB1 plays a role in the mitotic phase and its localization in LMS cell lines is cytoplasmic. Hence, we put forward that the RCC1 domain of RCBTB1 could play a similar role to that of RCC1 protein during mitosis, when nuclear envelope breaks up. Nevertheless, this presumed interaction between RCBTB1 and the proteins of the Ran network is based on the fact that putative homologs interact together in other species and so, this remains to be experimentally verified in human cells.

The second node of RCBTB1 interacting network is UBC, a gene that encodes polyubiquitin precursor, encompassing several cullins. Cullin-RING E3 ligases (CRLs) are responsible for ubiquitination of about 20% of proteins degraded through the ubiquitin-proteasome system²²⁷. Not surprisingly, they are frequently dysregulated in malignancies such as cancer^{228,229}. *Plafker et al.* reported that in HEK293T cell line, RCBTB1 interacts with CUL3 and the class III ubiquitin-conjugating enzyme (E2) UbcM2, which is the 100% identical murine homolog of human UBE2E3²³⁰. CUL3 recruits substrates for ubiquitination through polypeptide adaptors containing a BTB

domain²³⁰; these complexes have been reported as tumor suppressors since they regulate degradation of substrates involved in oxidative stress and cell cycle. Consequently, *RCBTB1* loss could result in deleterious accumulation of oncoproteins and in the dysregulation of cellular processes such as mitosis and microtubule dynamics, in turn leading to resistance towards anti-mitotic drugs.

Furthermore, IntAct database indicates that human *RCBTB1* also interacts with *GABARAP*, *GABARAPL1*, *GABARAPL2*, *MAP1LC3A*, and *MAP1LC3B*. These partners are all involved in autophagy: LC3s are involved in elongation of the phagophore membrane, and the *GABARAP* subfamily is necessary for autophagosome maturation and interacts with microtubules. The impact of autophagy on mitosis is still debated²³¹. Nevertheless, dysregulation of autophagy genes leads to prolonged mitosis or abnormal chromosome congression and outer kinetochore assembly, reinforcing the idea that autophagy is active during mitosis²³¹. Interestingly, it has been demonstrated that the *CUL3-KBTBD6/KBTBD7* ubiquitin ligase complex cooperates with *GABARAP* for targeted degradation of specific substrates(37). So, the sensitization effect towards mitotic catastrophe-inducing agents could be mediated through cooperation between CRL3 complexes, LC3/*GABARAP* protein family and the BTB domain-containing adaptor protein *RCBTB1*. These assumptions are reinforced by studying the MFS in a cohort of 339 sarcomas with complex genomic profiles based on the expression of each of these genes (obtained by gene expression microarrays): interestingly, a lower expression of *GABARAP*, *GABARAPL1* or *GABARAPL2* is significantly associated with metastatic evolution for sarcoma patients (Figure S2C).

In conclusion, our data demonstrate that *RCBTB1* is commonly altered in sarcomas with complex genetic profiles, especially in LMS, and that this deletion is associated with a higher risk for developing metastasis. Furthermore, *RCBTB1* expression was found to be a prognostic marker for metastatic evolution only for patients receiving chemotherapy, by retrospective analysis of MFS in a cohort of sarcoma patients. *In vitro* experiments show that RCBTB1 supports mitotic catastrophe induced by depolymerizing agents advocating anti-mitotic drugs - that target microtubule dynamics - such as docetaxel, paclitaxel and eribulin for patients with high expression of *RCBTB1* in the neoplastic tissue.

Table :

	Cohort #1 (106 sarcomas)	Cohort #2 (204 sarcomas)
Median age at diagnosis [95% CI]	63 [59 - 66]	63 [61 - 67]
Median follow-up in months [95% CI]	20.40 [13.80 - 27.07]	29.04 [22.68 - 35.28]
Histotypes (N, %)		
• Leiomyosarcomas	• 33 (31.13)	• 60 (29.41)
• UPS	• 38 (35.85)	• 82 (40.20)
• Myxofibrosarcoma	• 23 (21.70)	• 0
• Pleomorphic liposarcomas	• 7 (6.60)	• 0
• Dedifferentiated liposarcomas	• 0	• 62 (30.39)
• Others	• 5 (4.72)	• 0
Gender (N,%)		
• Male	• 55 (51.89)	• 103 (50.49)
• Female	• 51 (48.11)	• 101 (49.51)
Metastasis (N,%)	32 (30.19)	83 (40.69)
Local recurrence (N, %)	23 (21.70)	49 (24.02)
FNCLCC grades (N, %)		
• Grade 1	• 8 (7.55)	• 12 (5.88)
• Grade 2	• 25 (23.58)	• 69 (33.82)
• Grade 3	• 68 (64.15)	• 103 (50.49)
• Unknown	• 5 (4.72)	• 20 (9.80)
Tumor size (cm) [95% CI]	9 [8 - 10] (n=103)	10 [10 - 11] (n=192)
Location (%)		
External Trunk	• 88 (83.02)	• 151 (74.02)
• Trunk wall	• 25 (23.59)	• 31 (15.20)
• Head and neck	• 1 (0.94)	• 2 (0.98)
• Extremities	• 62 (58.49)	• 118 (57.84)
Internal Trunk	• 15 (14.15)	• 53 (25.98)
Unknown	• 3 (2.83)	• 0 (0)

Table 1: Cohorts #1 and #2 patient characteristics

Description of the patient cohorts studied to identify the biological pathways altered in 90% of sarcomas with complex genomic profiles (cohort #1) and to analyze metastasis-free survival (cohort #2). UPS means Undifferentiated Pleomorphic Sarcoma.

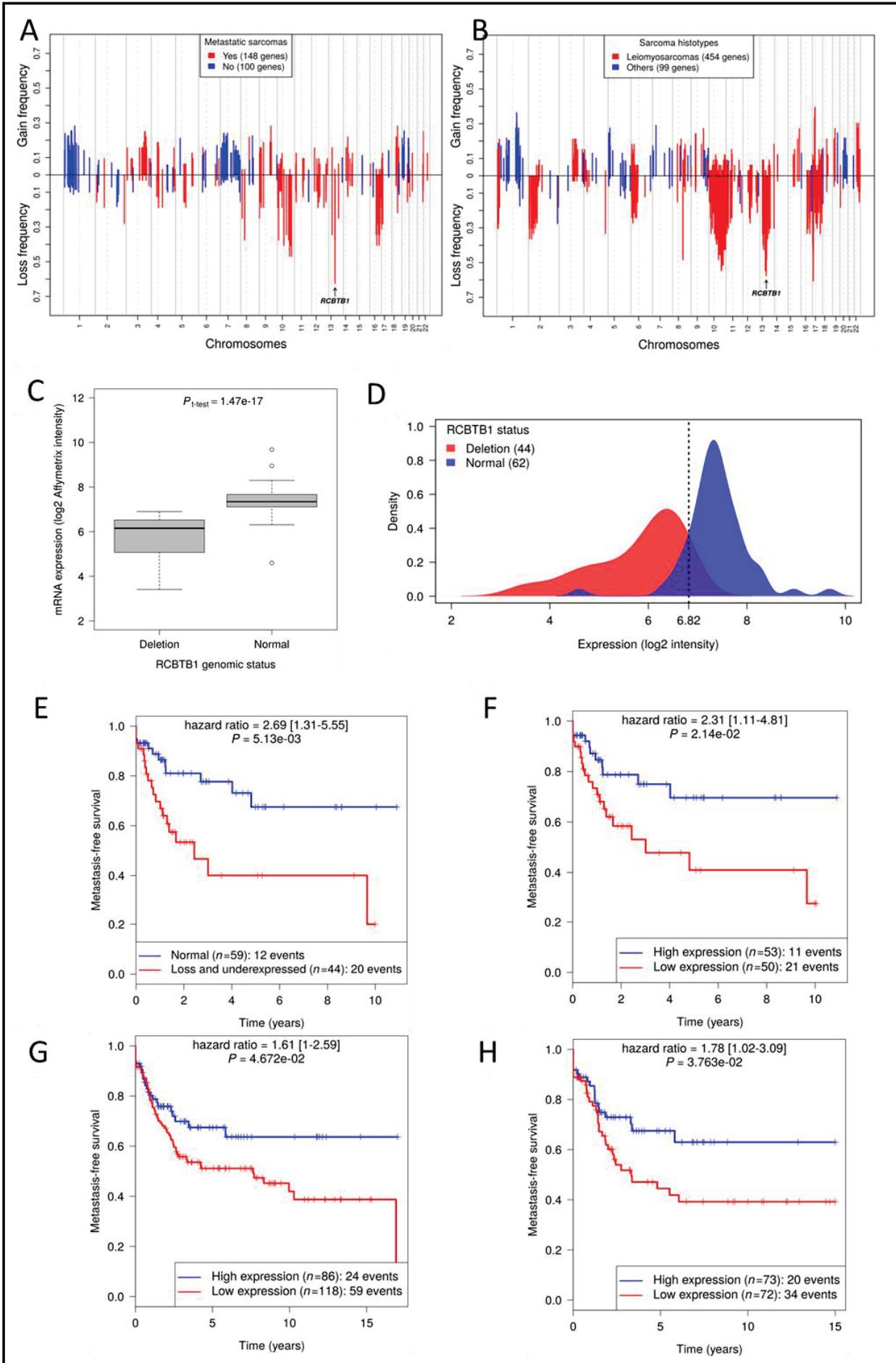


Figure 1: Association of *RCBTB1* deletion with leiomyosarcoma and metastatic evolution

(A-B) Each bar represents a gene on its chromosomal location and its alteration frequency: positive for a gain or negative for a deletion. **(A)** 151 altered genes are significantly correlated with metastatic evolution (in red) whereas 100 other altered genes are significantly associated with no metastatic evolution (in blue). **(B)** 462 altered genes are associated with leiomyosarcoma (LMS) histotype (in red) and 99 altered genes are associated with undifferentiated sarcomas (US) (in blue). The arrows indicate the most frequently deleted gene associated with both LMS histotype and metastatic evolution: *RCBTB1*.

(C-D) Analysis of *RCBTB1* expression depending on its genomic status (normal or deleted), represented **(C)** on a boxplot showing that *RCBTB1* deletion is significantly associated with a decrease of its expression ($P=1.47e-17$) and **(D)** on a density curve comparing signal intensities between the population of tumors with a normal copy number of *RCBTB1* and tumors harboring *RCBTB1* deletion. So, we defined the level (6.82) below which *RCBTB1* is under expressed.

(E-F) Kaplan-Meier curves of Metastasis-Free Survival (MFS) in 106 sarcoma patients divided into high- and low-risk groups according to **(E)** their *RCBTB1* genomic status (normal or deleted) or **(F)** their expression level of *RCBTB1* (high or low, cut-off = mean) determined by microarrays.

The prognostic value of *RCBTB1* expression on metastasis-free evolution was also investigated **(G)** by gene expression microarrays in an independent cohort of 204

sarcomas with complex genomic profiles and **(H)** by RNA-sequencing of 145 sarcomas.

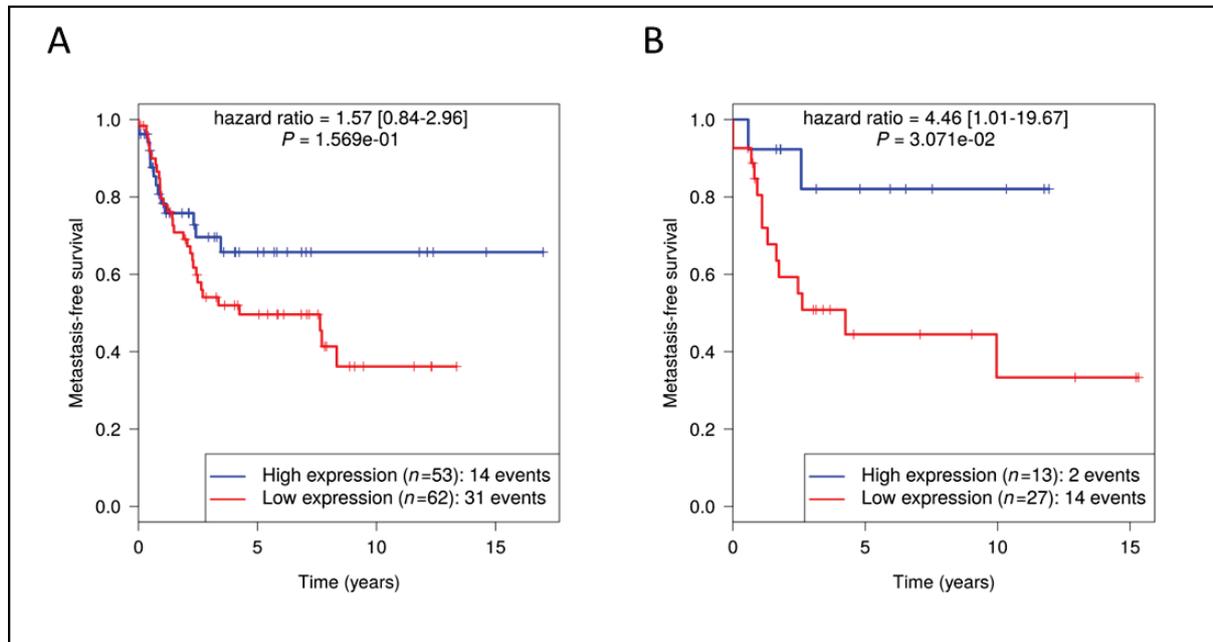


Figure 2: *RCBTB1* is significantly associated with metastasis-free survival for patients who received chemotherapy

Kaplan-Meier curves of Metastasis-Free Survival (MFS) for sarcoma patients divided into high- and low-risk groups according to their *RCBTB1* expression level determined by microarrays. Cohorts correspond to patients that did not **(A)** or did **(B)** receive chemotherapy.

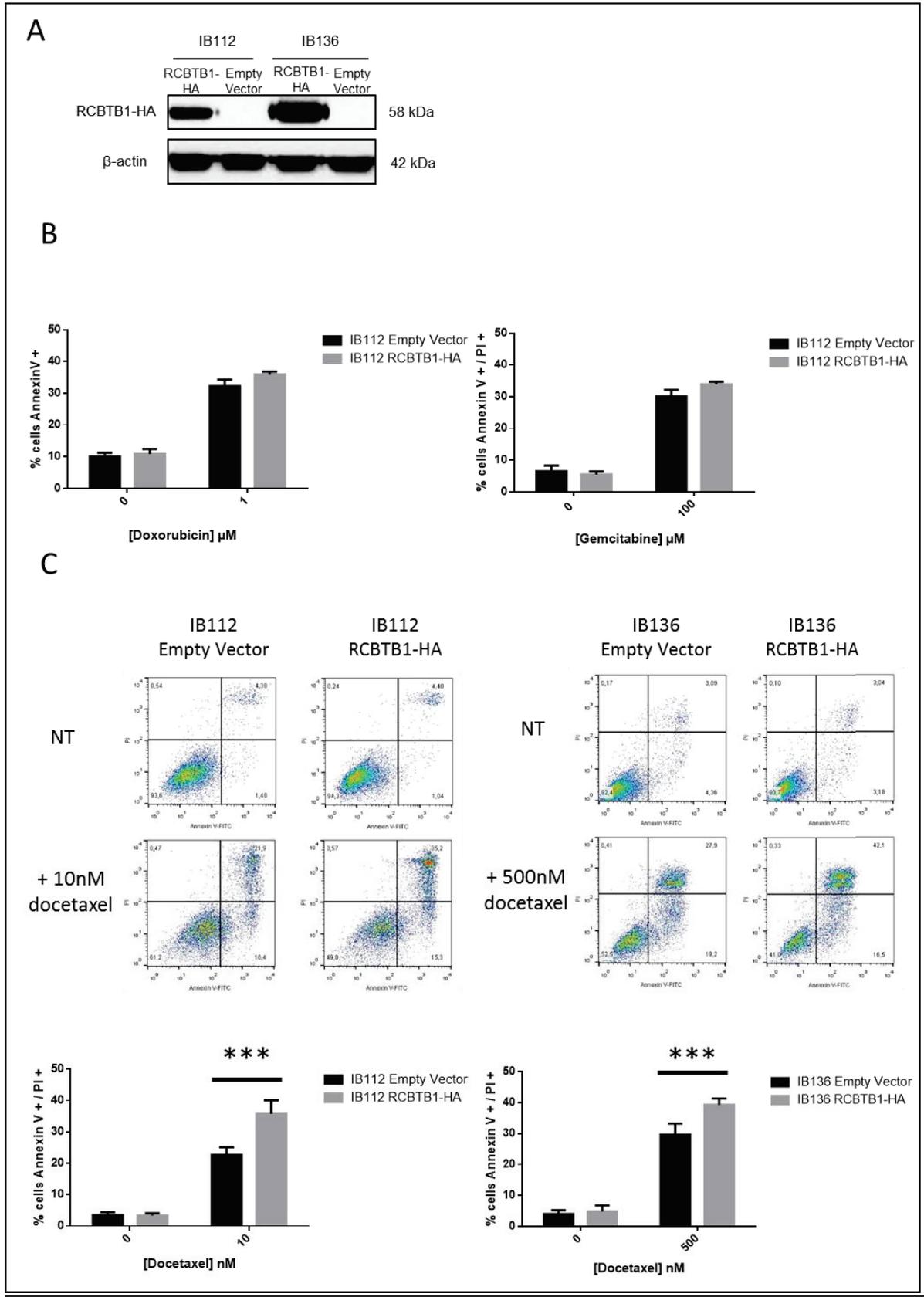


Figure 3: RCBTB1 overexpression increases *in vitro* docetaxel-induced apoptosis in two LMS cell lines.

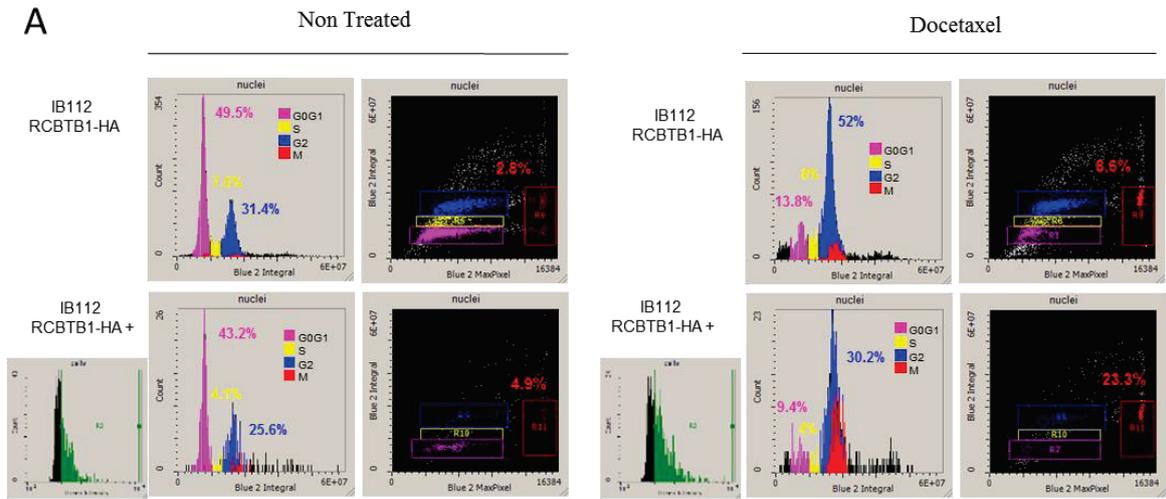
(A) Validation of RCBTB1-HA overexpression by western blotting after lentiviral transduction in two LMS cell lines. Control cell lines express an empty vector.

(B) Apoptosis induced by 72 h treatment with Doxorubicin and Gemcitabine was measured by dual Annexin V-FITC / PI staining (except for doxorubicin, for which simple Annexin V-FITC staining was used): 10 000 events were analyzed by flow cytometry for each of three replicates in three independent experiments.

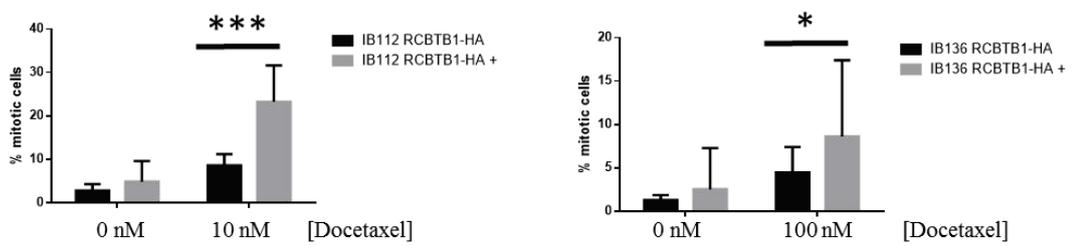
(C) Apoptosis induced by 72 h treatment with Docetaxel was measured by dual Annexin V-FITC / PI staining and analyzed by flow cytometry. Data shown summarize the results of three independent experiments with three replicates (unpaired t-test). Upper quadrants are representative of the results obtained for all experiments.

***: $P < 0.001$.

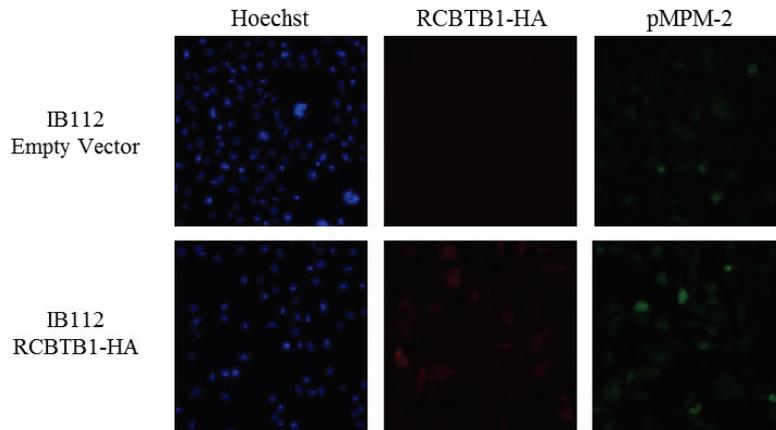
A



B



C



D

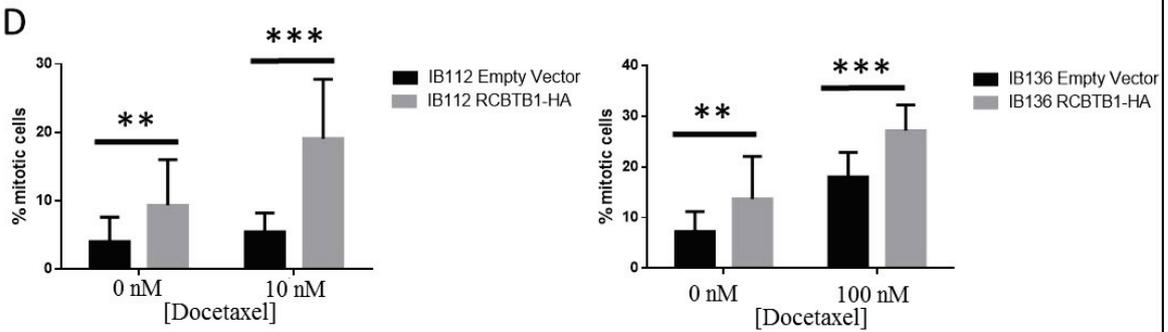
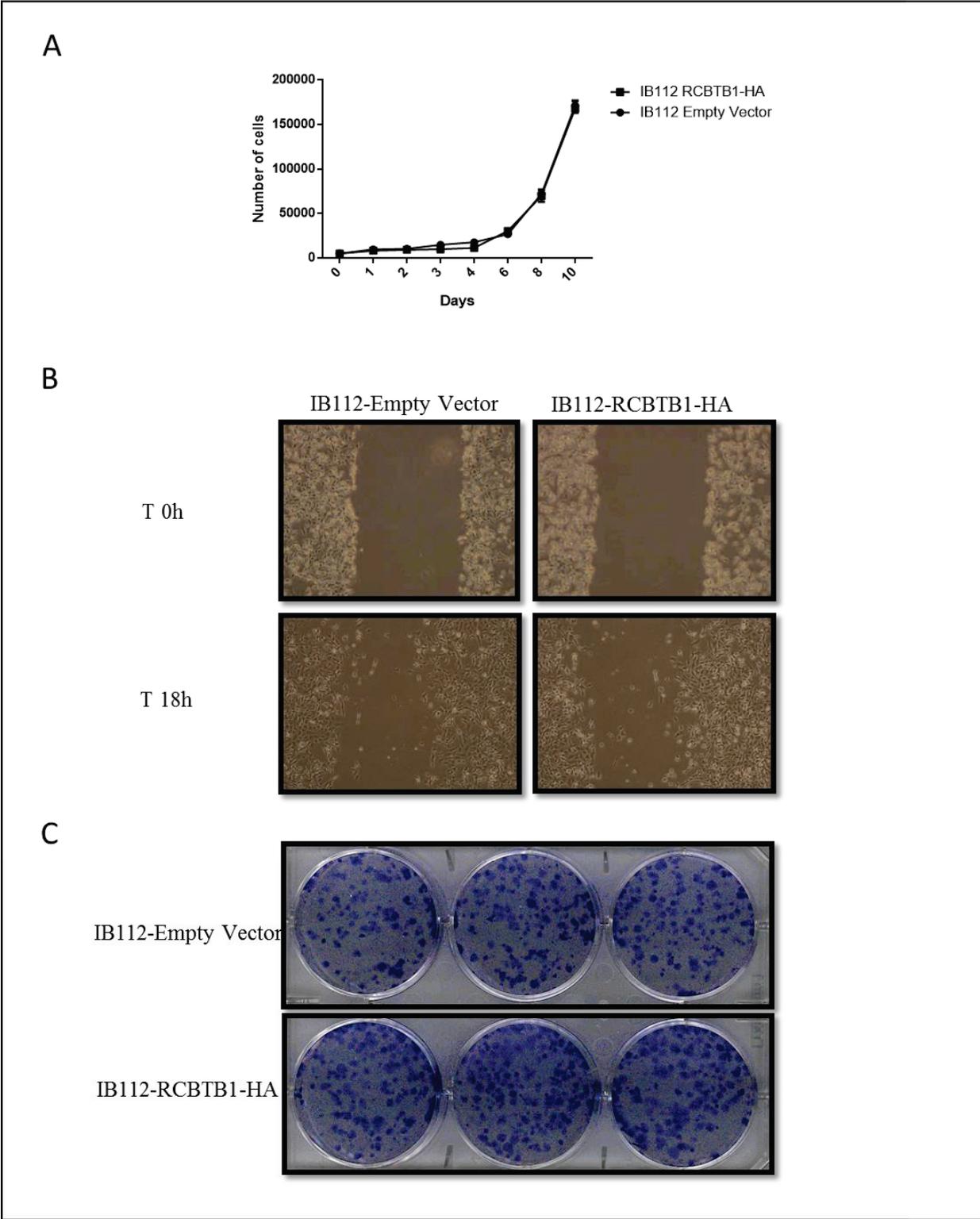


Figure 4: *RCBTB1* overexpression increases the percentage of mitotic cells in two LMS cell lines.

RCBTB1 was overexpressed by lentiviral transduction in two LMS cell lines derived from patient tumors (IB112 and IB136). Cells were treated with docetaxel (at 10 nM and 100 nM for IB112 and IB136, respectively) for 24 h and then were fixed for fluorescent staining.

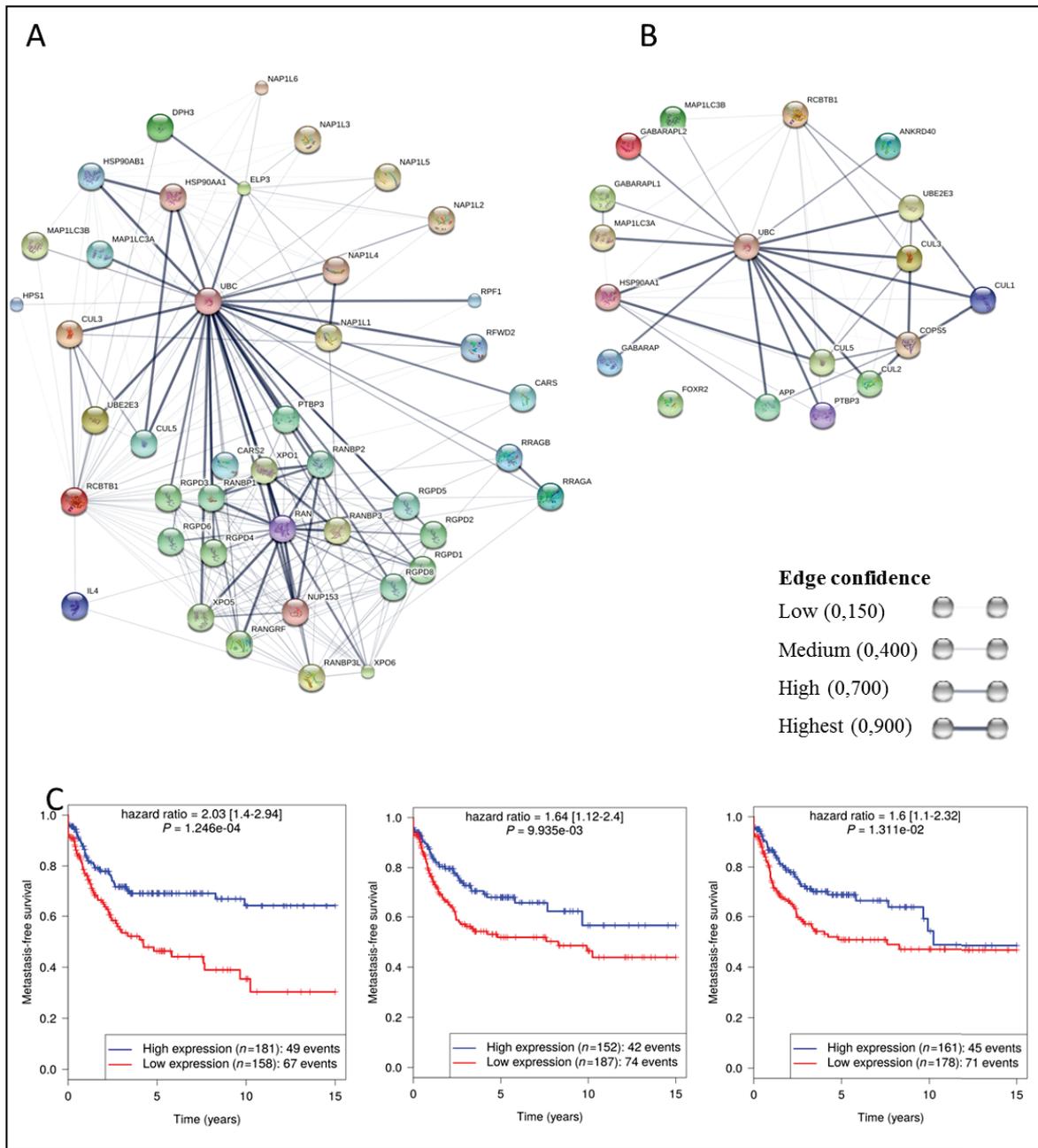
(A-B) On one hand, nuclei were stained with Hoechst and on the other, *RCBTB1*-HA was labeled with a green-fluorescent antibody. Chromatin condensation was evaluated with iCys technology. **(A)** Events (count) correspond to the number of nuclei identified by Hoechst staining, and max pixel feature allowed identification of cell cycle phases in control conditions and after docetaxel treatment. Lower panels show similar analyses on the cell population that highly expresses *RCBTB1* (strong green labeling). **(B)** Histograms summarize results obtained with iCys technology for IB112 and IB136 cell lines that are representative of two independent experiments with 15 replicates for each condition. P value was obtained by a paired t-test. *: $P < 0.05$ and ***: $P < 0.001$.

(C-D) Mitosis-specific marker pMPM-2 (green) and *RCBTB1*-HA (red) dual staining was analyzed by epifluorescence microscopy. **(C)** These images are representative of the results obtained with both LMS cell lines. **(D)** Histograms below sum up results that are representative of three independent experiments. A minimum of 700 cells in mitosis was analyzed to determine the percent of mitotic cells. P value was obtained by Mann-Whitney test. **: $P < 0.01$ and ***: $P < 0.001$.



Supplementary figure 1: *RCBTB1* overexpression does not alter phenotypes of aggressiveness in IB112 LMS cell line

RCBTBI was overexpressed by lentiviral transduction in the IB112 LMS cell line. **(A)** Cells were plated and counted by flow cytometry during 10 days with 6 replicates for each time point. **(B)** Cell migration was assessed with wound-healing assays. Photos were taken at 0h and 18h after scratching the cell layer. **(C)** Clonogenicity was measured with crystal violet staining. Data shown are representative of the results obtained in three independent experiments.



Supplementary figure S2: RCBTB1 interacting partners and their relevance in terms of metastasis-free survival

(A, B) RCBTB1 network obtained with STRING tool. Network nodes represent proteins and edges are protein-protein associations. Edge thickness corresponds to the confidence level of this association. Associations were provided by (A) STRING database and comprise interactions observed between homologs in other species or by (B) IntAct database that provided only human interacting partners.

(C) Kaplan-Meier curves of metastasis-free survival for sarcoma patients divided into high- and low-risk groups according to their expression level (cut-off = mean) of *GABARAP*, *GABARAPL1* or *GABARAPL2*.

Gene expression was determined by gene expression microarrays on 339 sarcomas with complex genomic profiles.

	Normal	Deletion	
Metastatic	37,5% (12)	62,5% (20)	Chi2=6.44e-03
Non metastatic	66,2% (47)	33,8% (24)	
LMS	42,4% (14)	57,6% (19)	Chi2=3.64e-02
Others	64,3% (45)	35,7% (25)	

Supplementary Table 1: Proportion of tumors harboring *RCBTB1* deletion according to clinical annotations (metastatic or not metastatic; LMS or other histotypes) and the statistical significance of the association between *RCBTB1* deletion and metastasis or LMS histotype.

ID	Biological Pathways of GO database :	P-value
GO:0006996	organelle organization	1.26e-88
GO:0007049	cell cycle	8.87e-81
GO:0016070	RNA metabolic process	3.52e-80
GO:0044267	cellular protein metabolic process	6.96e-78
GO:0007399	nervous system development	1.21e-77
GO:0090304	nucleic acid metabolic process	2.05e-77
GO:0030154	cell differentiation	1.29e-76
GO:0006950	response to stress	1.50e-76
GO:0048869	cellular developmental process	1.55e-76
GO:0010467	gene expression	3.13e-76
GO:0048731	system development	7.40e-75
GO:0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	5.63e-73
GO:0022008	neurogenesis	1.75e-72
GO:0009605	response to external stimulus	2.09e-72
GO:0007275	multicellular organismal development	6.04e-72
GO:0019538	protein metabolic process	8.07e-72
GO:0048699	generation of neurons	8.60e-72
GO:0048856	anatomical structure development	1.20e-70
GO:0071842	cellular component organization at cellular level	3.35e-70
GO:0016043	cellular component organization	3.35e-70
GO:0030182	neuron differentiation	7.44e-70
GO:0071841	cellular component organization or biogenesis at cellular level	1.31e-69
GO:0071840	cellular component organization or biogenesis	1.31e-69
GO:0042221	response to chemical stimulus	2.71e-69
GO:0034641	cellular nitrogen compound metabolic process	4.16e-66
GO:0032502	developmental process	1.18e-65
GO:0006807	nitrogen compound metabolic process	4.30e-63
GO:0007165	signal transduction	1.13e-60
GO:0023052	signaling	2.52e-57
GO:0051716	cellular response to stimulus	2.00e-52
GO:0032501	multicellular organismal process	6.63e-51
GO:0044260	cellular macromolecule metabolic process	1.92e-47
GO:0050896	response to stimulus	7.55e-41
GO:0043170	macromolecule metabolic process	2.50e-40
GO:0044237	cellular metabolic process	1.79e-30
GO:0044238	primary metabolic process	2.15e-29
GO:0050794	regulation of cellular process	2.97e-27
GO:0050789	regulation of biological process	4.02e-23
GO:0065007	biological regulation	1.43e-19
GO:0008152	metabolic process	1.32e-17
GO:0009987	cellular process	0.04

Supplementary Table 2: List of the 41 biological pathways commonly altered in 100% of sarcomas using Gene Ontology database.

➤ Impact de *RCBTB1* sur l'instabilité chromosomique

Afin de déterminer si l'inhibition du gène *RCBTB1* pourrait jouer un impact sur la stabilité génomique, j'ai utilisé un modèle présentant un profil génomique simple, sans remaniement complexe, et possédant une origine commune aux sarcomes. Mon choix s'est alors porté sur une lignée d'origine mésenchymateuse de fibroblastes transformée RST. L'ensemble du "Matériel et Méthodes" de cette partie est présenté en Annexe (Page 178).

Validation de l'inhibition du gène *RCBTB1* dans les modèles cellulaires

Afin de valider les modèles cellulaires d'inhibition du gène *RCBTB1*, son expression a été étudiée par RT-qPCR. Les données d'expression en fonction des cinq différents shARN utilisés sont présentées dans la Figure 28. D'après ces résultats, nous avons sélectionné le modèle cellulaire présentant l'inhibition la plus importante : la lignée exprimant le sh444 qui inhibe l'expression du gène *RCBTB1* à 79%.

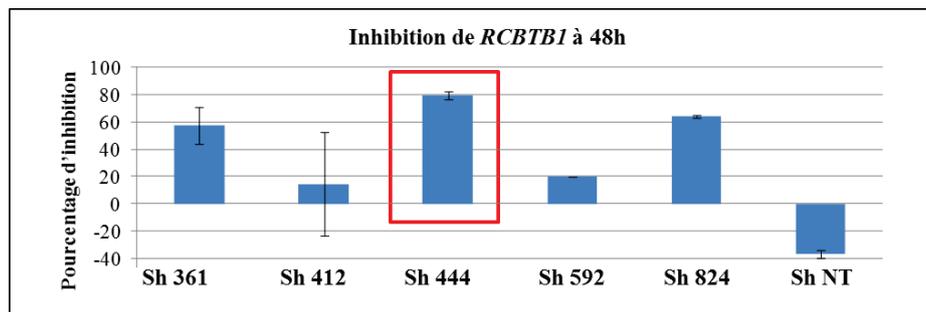


Figure 28 : Graphique présentant les valeurs d'inhibition pour *RCBTB1* en fonction des différents shARN après 48h post-induction à la doxycycline

Afin de déterminer si l'inhibition de *RCBTB1* joue un rôle dans les remaniements chromosomiques observés dans les sarcomes à génétique complexe, l'effet de l'inhibition de ce gène sur les cassures double-brin et sur les remaniements chromosomiques déséquilibrés a été étudié.

Effet sur les cassures double-brin de l'ADN

Pour observer les cassures double-brin de l'ADN, un marquage de la phosphorylation de l'histone γ H2AX (P- γ H2AX) a été réalisé. L'histone γ H2AX est phosphorylée au niveau des cassures double-brin de l'ADN. Nous avons visualisé le

marquage de l'histone phosphorylée P- γ H2AX par immunofluorescence et quantifié sa fluorescence par cytométrie en flux. Ces expériences ont été réalisées après 17 passages sous doxycycline pour la lignée induite et 17 passages sans induction pour la lignée contrôle.

Les résultats d'immunofluorescence permettent d'observer que toutes les lignées, avec ou sans induction de l'expression des shARN, présentent un marquage de l'histone P- γ H2AX. De la même manière, en cytométrie en flux, nous pouvons observer que toutes les lignées présentent un marquage P- γ H2AX. De plus, les expériences de cytométrie en flux permettent de distinguer deux populations cellulaires dans toutes les lignées induites ou non induites : des cellules présentant un marquage plus faible et des cellules présentant un marquage plus fort (Figure 29). Par exemple, pour la lignée contrôle shNT non induite, 47% des cellules présentent un marquage faible et 53% des cellules présentent un fort marquage. Afin de confirmer que le marquage plus fort représentait bien une augmentation de foyers γ H2AX et donc de cassures double-brin d'ADN, la lignée shNT a été traitée avec de l'étoposide (inhibiteur de la topoisomérase II humaine augmentant les complexes de clivage et donc le nombre de cassures double-brin). Nous pouvons observer sur la Figure 29 que suite au traitement à l'étoposide, toutes les cellules de la lignée présentent un marquage fort. Ainsi, il semble que les lignées de fibroblastes présentent un niveau basal important de cassures double-brin. Nous avons donc voulu savoir si l'inhibition du gène *RCBTB1* augmentait la proportion de cellules fortement marquées et induisait donc l'apparition de nouvelles cassures d'ADN.

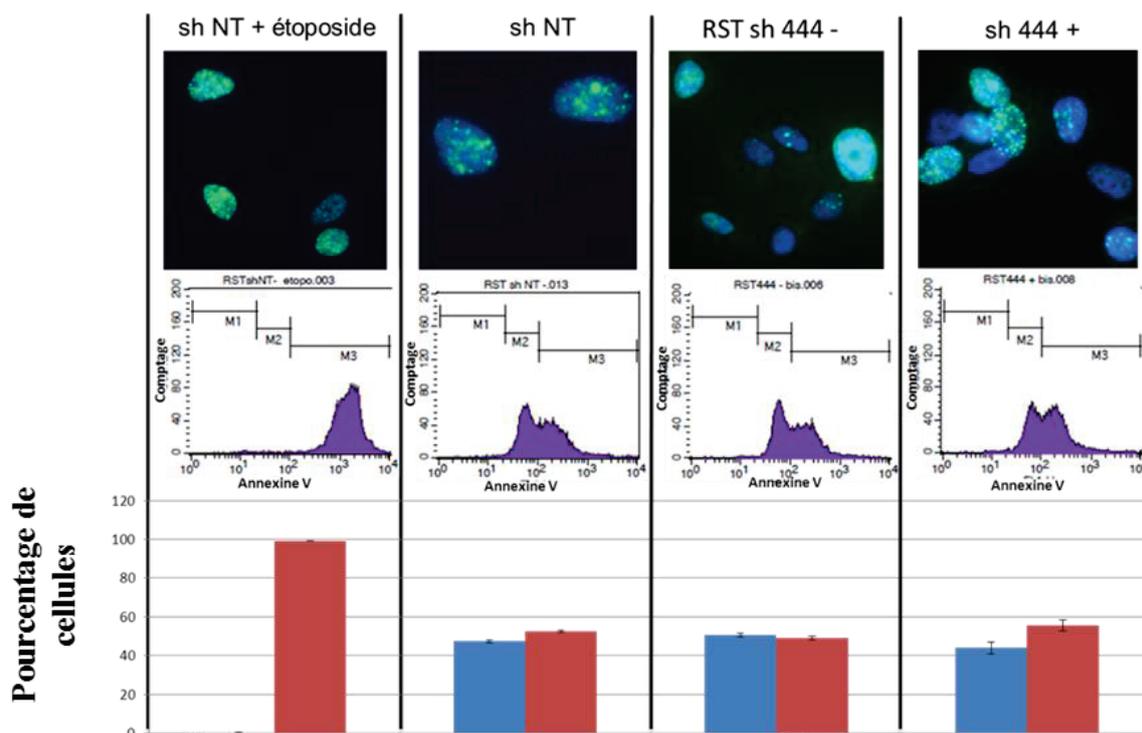


Figure 29 : Visualisation des foyers γ H2AX par immunofluorescence et quantification par FACS dans les fibroblastes dans lesquels *RCBTB1* est inhibé ou non

La lignée exprimant le sh444 présente, lorsqu'il est non induit et induit respectivement, 49% et 56% de cellules fortement marquées. Ces différences observées ne sont pas statistiquement significatives (test de Fisher) alors que l'inhibition du gène *RCBTB1* est de 79% au moment de l'expérience. L'inhibition du gène *RCBTB1* n'induit pas de cassures double-brin de l'ADN dans cette lignée.

Effet sur la complexité génomique

Afin de détecter de nouvelles anomalies génomiques déséquilibrées, l'ADNg des lignées exprimant le sh444 a été analysé par CGH-array suite à 17 passages sous induction avec la doxycycline. En comparant les profils des lignées induites par rapport aux lignées non induites, nous constatons qu'il n'y a pas de différence entre les profils (Figure 30). De plus, ils sont identiques au profil de la lignée exprimant le shNT. L'inhibition de *RCBTB1* n'induit pas l'apparition de nouveaux remaniements déséquilibrés dans les lignées de fibroblastes après 17 passages avec inhibition de *RCBTB1*.

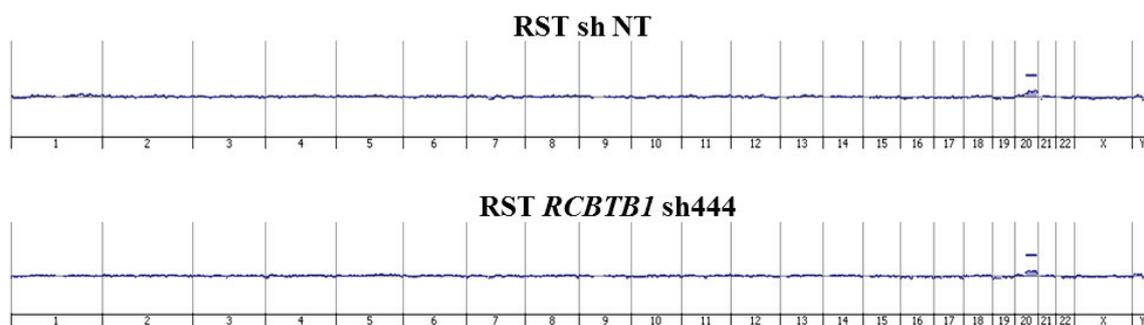


Figure 30 : Profils génomiques des lignées de fibroblastes obtenus suite à 17 passages sous induction à la doxycycline

Il apparaît alors que l'inhibition de *RCBTB1* dans la lignée de fibroblastes RST n'a pas d'incidence sur la stabilité génomique et infirme donc l'hypothèse selon laquelle la délétion de *RCBTB1* induirait des cassures de l'ADN et augmenterait ainsi l'évolution métastatique des sarcomes à génétique complexe.

Validation *in vivo* de la sensibilisation des tumeurs au docétaxel par *RCBTB1*

Plus récemment, j'ai également testé la sensibilisation des cellules surexprimant *RCBTB1* *in vivo*. Pour cela, j'ai réalisé des xénogreffes des lignées IB136-Vide et IB136-*RCBTB1* sur des souris NSG (10 souris par condition). J'ai ainsi pu constater que les tumeurs formées par les lignées surexprimant *RCBTB1* présentaient une plus forte croissance en volume en comparaison avec la lignée contrôle (Figure 31). Cette différence de croissance significative ($P < 0.001$) avant traitement est en accord avec l'augmentation du pourcentage de cellules en mitose observée précédemment *in vitro*.

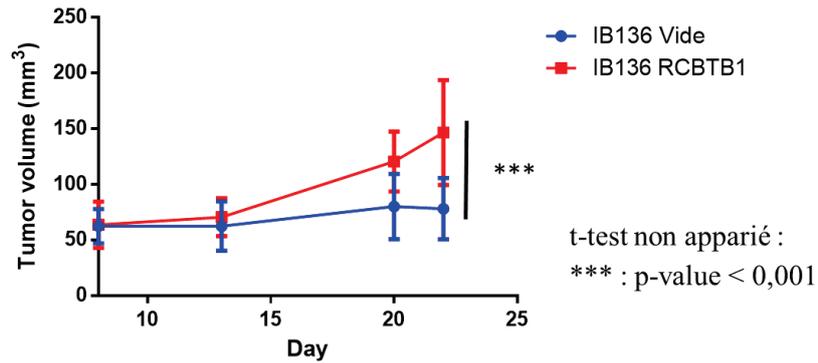


Figure 31 : Courbe de croissance des tumeurs en fonction de l'expression de RCBTB1

Lorsque les tumeurs ont atteint un volume de 150 mm³, les souris ont été traitées soit avec 100µl de DMSO, soit avec 100µl de docétaxel à 1mg/mL ou 2mg/mL. Trente jours plus tard, les souris ont été sacrifiées puis les tumeurs résiduelles ont été récupérées et pesées afin de comparer l'effet des différents traitements sur la masse tumorale. Pour l'analyse des résultats, les poids des tumeurs ont été normalisés par rapport à leur groupe contrôle respectif (traité avec le DMSO) (Figure 32).

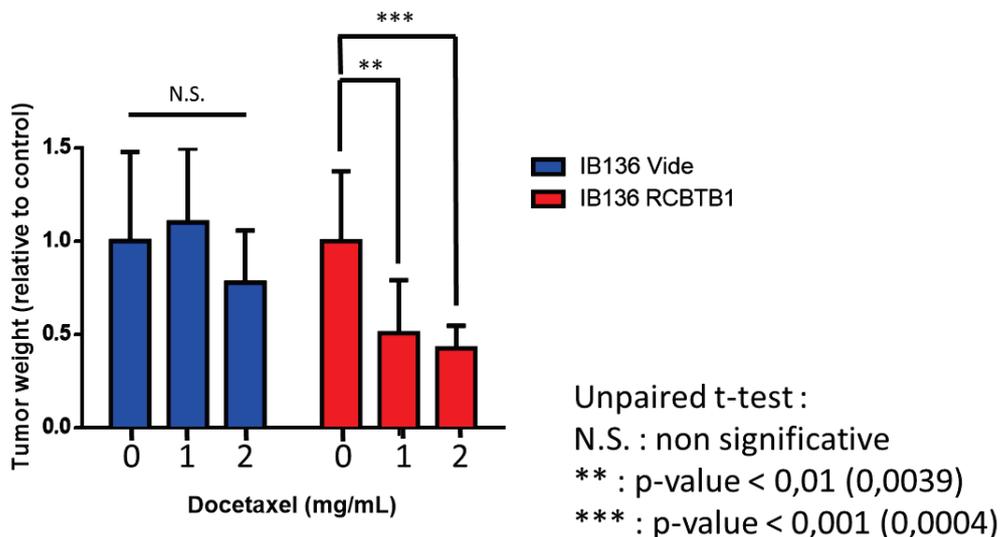


Figure 32 : Poids des tumeurs récupérées 30 jours après le début du traitement

t-test non apparié :
 N.S. : non significative
 ** : p-value < 0,01 (0,0039)
 *** : p-value < 0,001 (0,0004)

Nous pouvons voir que dans le groupe de souris greffées avec IB136-*vide*, le traitement au docétaxel n'a eu aucun effet significatif par rapport au groupe contrôle traité au DMSO. En revanche, dans les groupes IB136-RCBTB1 ayant été traités avec le docétaxel, nous observons une réduction significative du poids des tumeurs par rapport au groupe contrôle, en moyenne 49% et 58% de réduction respectivement pour les groupes traités avec 1mg/mL et 2mg/mL de docétaxel. Cette expérience *in vivo* confirme donc les résultats obtenus *in vitro* présentés précédemment montrant que la sur-expression de *RCBTB1* sensibilise les lignées de LMS au docétaxel.

En conclusion, lors de ma thèse j'ai pu montrer que la délétion de *RCBTB1* associée à une perte d'expression était associée à l'évolution métastatique. Contrairement à ce qui est décrit dans la littérature, je n'ai pas montré d'effet *in vitro* de RCBTB1 sur la stabilité du génome ni sur l'acquisition de phénotypes d'agressivité tels que la prolifération, la migration et l'invasion cellulaires. Cependant, étant donné que la sur-expression de *RCBTB1* augmente le pourcentage de cellules en mitose *in vitro* et donc la croissance tumorale *in vivo*, elle induit ainsi une sensibilisation au docétaxel qui agit lors de la division cellulaire dans les lignées cellulaires de LMS.

I- Rôle de la délétion 5' du gène *DMD* dans l'évolution métastatique

A) Contexte

Comme indiqué dans le 2^e chapitre, nous avons observé une délétion récurrente du gène *DMD* dans 17% des 321 sarcomes (comprenant des sarcomes à génétique complexe, des GISTs et des synovialosarcomes) étudiés. De plus, la délétion du gène *DMD* est significativement associée à l'évolution métastatique.

L'objectif de ma thèse était alors de comprendre le rôle de la délétion du gène *DMD* sur l'évolution métastatique. Observant que la délétion ciblée du gène *DMD* affecte systématiquement l'isoforme Dp427 et jamais l'isoforme Dp71, j'ai alors posé l'hypothèse que Dp427 pourrait jouer un rôle de gène suppresseur de tumeur et / ou Dp71 jouerait un rôle d'oncogène à cause de la délétion ciblée de Dp427 qui pourrait augmenter l'expression de Dp71 en libérant les complexes de transcription de Dp427.

Au vu de la complexité de ce gène, nous avons tout d'abord étudié l'expression de l'ensemble des isoformes par RNA-seq sur une cohorte de 145 sarcomes, afin d'identifier les isoformes pouvant jouer un rôle dans l'oncogenèse des sarcomes. Puis, dans un premier temps, j'ai cherché à comprendre le rôle de l'isoforme Dp427 dans des lignées de LMS car il s'agit de l'isoforme systématiquement ciblée dans toutes les délétions du gène *DMD*. Dans un second temps, j'ai étudié le rôle de Dp71 qui, contrairement à Dp427, est épargnée par la délétion ciblée du gène *DMD*.

L'ensemble du "Matériel et Méthodes" de cette partie est présenté en Annexe (Page181).

B) Résultats

1) Recherches de mécanismes d'altération alternatifs

Dans un premier temps, j'ai souhaité déterminer si *DMD* pouvait être inactivé par d'autres mécanismes que la délétion. Pour cela, grâce aux données de RNA-seq,

j'ai recherché les variations présentes dans le gène *DMD* de 145 tumeurs. Trente deux variations ont été identifiées, 19 étant déjà référencées dans la base dbSNP. Sur les 13 variations restantes, seulement 3 sont des variations faux-sens. J'ai cherché à les valider par séquençage de l'ADN de type Sanger et n'en ai validé aucune. Nous pouvons donc conclure que l'inactivation du gène *DMD* est exclusivement réalisée par un mécanisme de délétion et non de mutation.

2) Etude de l'expression du gène *DMD*

➤ Dans les tumeurs

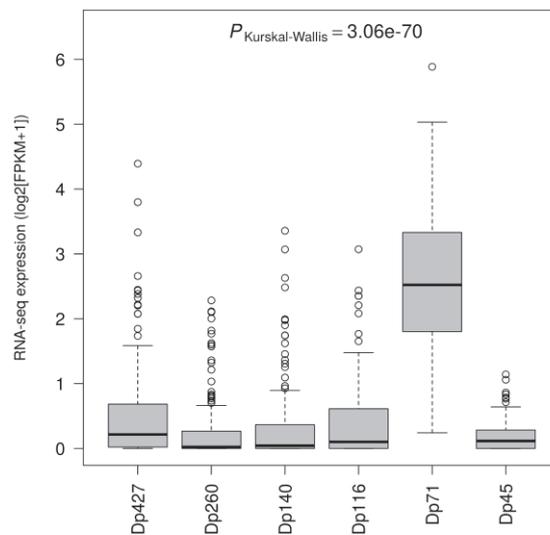


Figure 33 : Expression par RNA-seq des isoformes du gène *DMD* dans 145 sarcomes

Dans la série de 145 RNA-seq, nous voyons très nettement que seule l'isoforme Dp71 est exprimée. Cette observation est concordante avec les délétions observées ne ciblant jamais Dp71.

➤ Dans les lignées

Pour pouvoir établir des modèles d'étude de l'expression de Dp427 et Dp71, nous avons dans un premier temps étudié l'expression de ces deux isoformes dans les lignées suivantes :

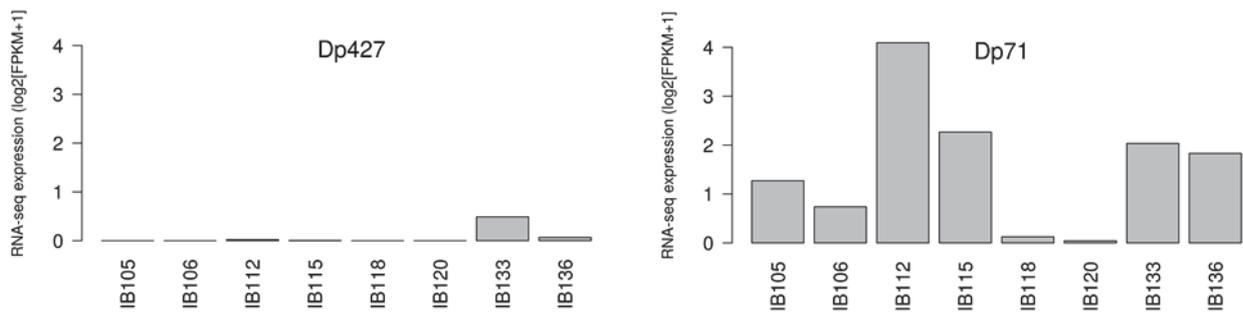


Figure 34 : Expression des isoformes Dp427 et Dp71 dans des lignées de sarcomes

Par RNA-seq, nous pouvons observer qu'une seule lignée exprime Dp427, il s'agit de la lignée de LMS IB133. Quant à Dp71, elle semble être exprimée dans toutes les lignées à l'exception de IB118 et IB120 au niveau transcriptomique.

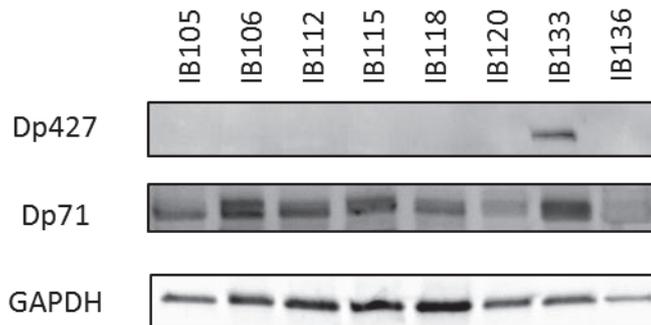


Figure 35 : Expression de Dp427 et Dp71 dans les lignées par Western blot

J'ai donc validé ces résultats par Western blot. Au niveau protéique, seule IB133 exprime l'isoforme Dp427, alors que toutes les lignées expriment Dp71.

3) Rôle suppresseur de tumeur de Dp427

Etant donné la taille de cette isoforme, il est très difficile de la sur-exprimer *in vitro* dans nos modèles. J'ai alors adopté comme stratégie d'inhiber ce gène par la technologie CRISPr/Cas9. Pour cela, j'ai sélectionné l'unique lignée à ma disposition l'exprimant : IB133 qui est un LMS. Etant donné que je n'avais qu'une seule lignée à ma disposition, j'ai réalisé 5 infections de manière indépendante sur la lignée IB133 avec le système CRISPr/Cas9 et le système contrôle, afin d'observer la variabilité biologique due à l'infection.

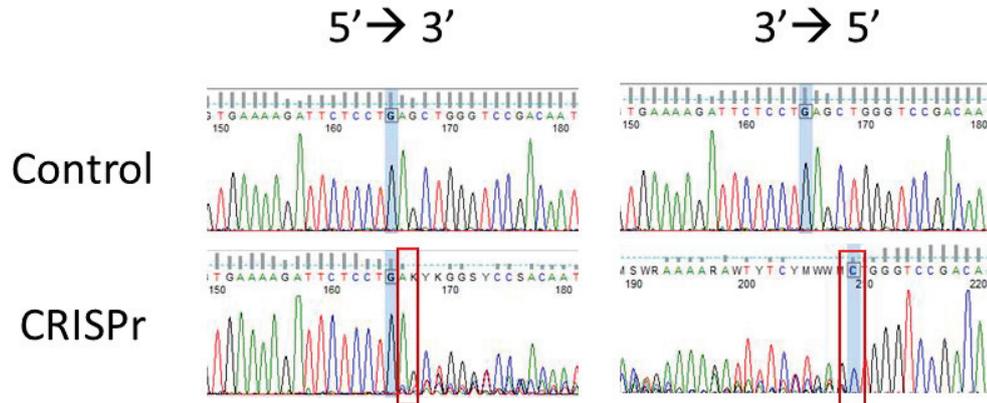


Figure 36 : Séquençage Sanger de la lignée IB133 après infection

Après avoir infecté la lignée IB133 avec les particules lentivirales contenant soit le système CRISPr/Cas9 dirigé contre Dp427 soit le système contrôle, j'ai vérifié le type de mutation induite. Sur la Figure 36, nous pouvons observer que dans la lignée infectée avec le système CRISPr/Cas9 complet, il y a une délétion de 2 nucléotides, introduisant donc un décalage du cadre de lecture. Ce séquençage a été réalisé sur l'ensemble des 5 lignées infectées, et dans tous les cas, on constate une délétion de 2 nucléotides.

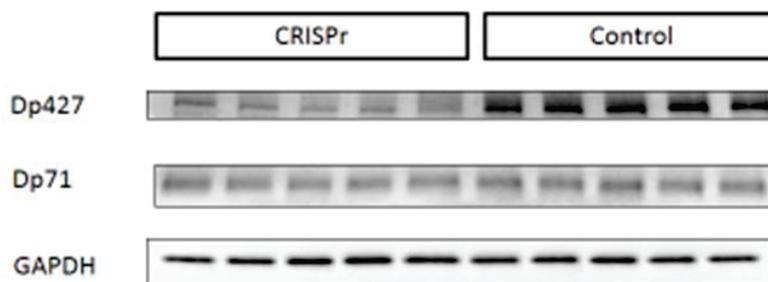


Figure 37 : Validation de l'inhibition de Dp427 par CRISPr

Par Western blot, j'ai validé que cette mutation induisait bien une diminution partielle d'au moins 50% de l'expression de Dp427 dans les 5 lignées. Par ailleurs, nous pouvons également noter qu'il n'y a pas d'augmentation significative de l'expression de Dp71 dans les lignées infectées avec le CRISPr/Cas9 ciblant Dp427.

Après cette validation des modèles cellulaires présentant une inhibition de Dp427, j'ai cherché à étudier si cette diminution de l'expression de Dp427 pouvait avoir une répercussion sur des caractéristiques d'agressivité *in vitro*.

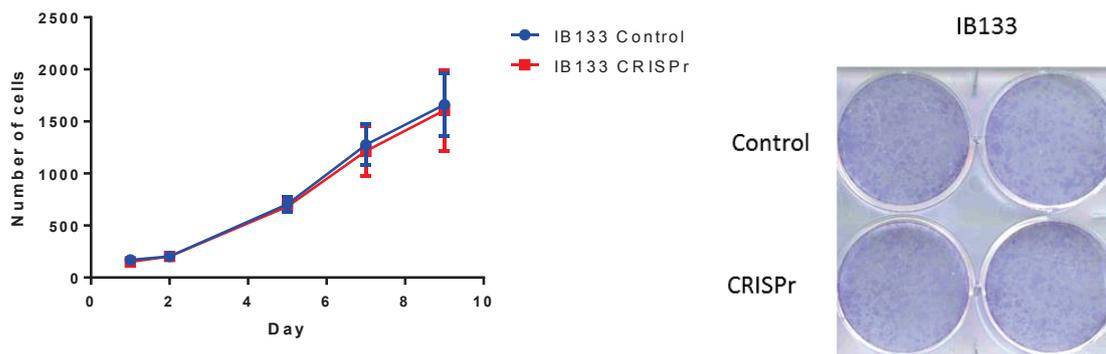


Figure 38: Tests de prolifération et de clonogénicité

Sur la Figure 38, j'ai représenté les courbes de prolifération moyennes de l'ensemble des 5 lignées CRISPr (en rouge) et des 5 lignées contrôles (en bleu). Il apparaît que l'inhibition de Dp427 ne semble pas avoir d'impact sur la prolifération des cellules de la lignée IB133. Il en est de même pour la clonogénicité où l'inhibition de Dp427 ne semble pas influencer sur la capacité des cellules à former des clones.

4) Rôle oncogène de Dp71 ?

Pour déterminer si Dp71 pouvait avoir un rôle oncogène, j'ai adopté deux stratégies. J'ai tout d'abord étudié l'effet de l'inhibition de Dp71 dans des lignées de sarcomes, puis l'effet de la sur-expression de Dp71 dans une lignée de myoblastes.

Concernant l'inhibition de Dp71, nous avons sélectionné 2 lignées de LMS (IB112 et IB136) et également une lignée commerciale de synoviosarcome (SW982, sans transcrit de fusion SSX indentifié), en raison de leurs types histologiques, dans lesquels nous observons fréquemment des délétions de *DMD*, mais également du fait qu'elles n'expriment pas Dp427. Ainsi nous pouvons utiliser des shARN ciblant la séquence de Dp71 (commune avec celle de Dp427) pour étudier spécifiquement l'inhibition de Dp71 en l'absence de Dp427. Les shARN utilisés ont été clonés dans un vecteur Tet-pLKO sous le contrôle d'un promoteur inductible avec la doxycycline.

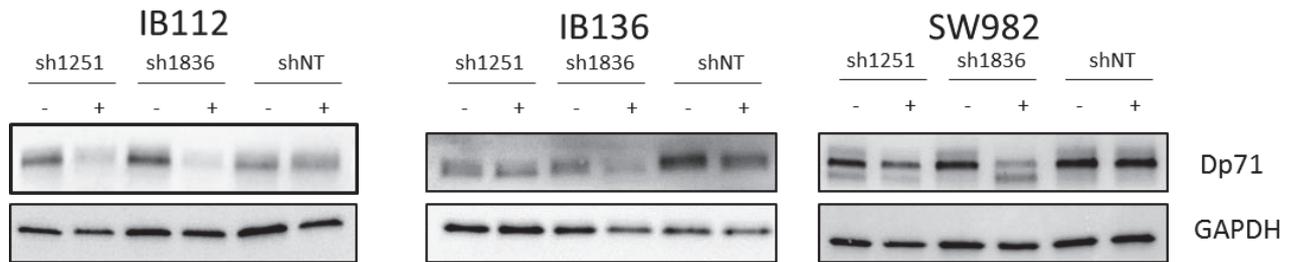


Figure 39 : Expression de Dp71 avec les shARN ciblant Dp71 dans les 3 lignées étudiées

Sur la Figure 39, nous pouvons observer que l'ajout de doxycycline (noté +) induit bien une inhibition de Dp71 dans les lignées exprimant les sh1251 et sh1836 (ciblant Dp71) contrairement à celles exprimant un shNT (ne ciblant aucun ARN humain). Notons néanmoins que le sh1251 dans la lignée IB136 semble avoir un faible effet inhibiteur sur l'expression de Dp71 lorsqu'il est induit par la doxycycline.

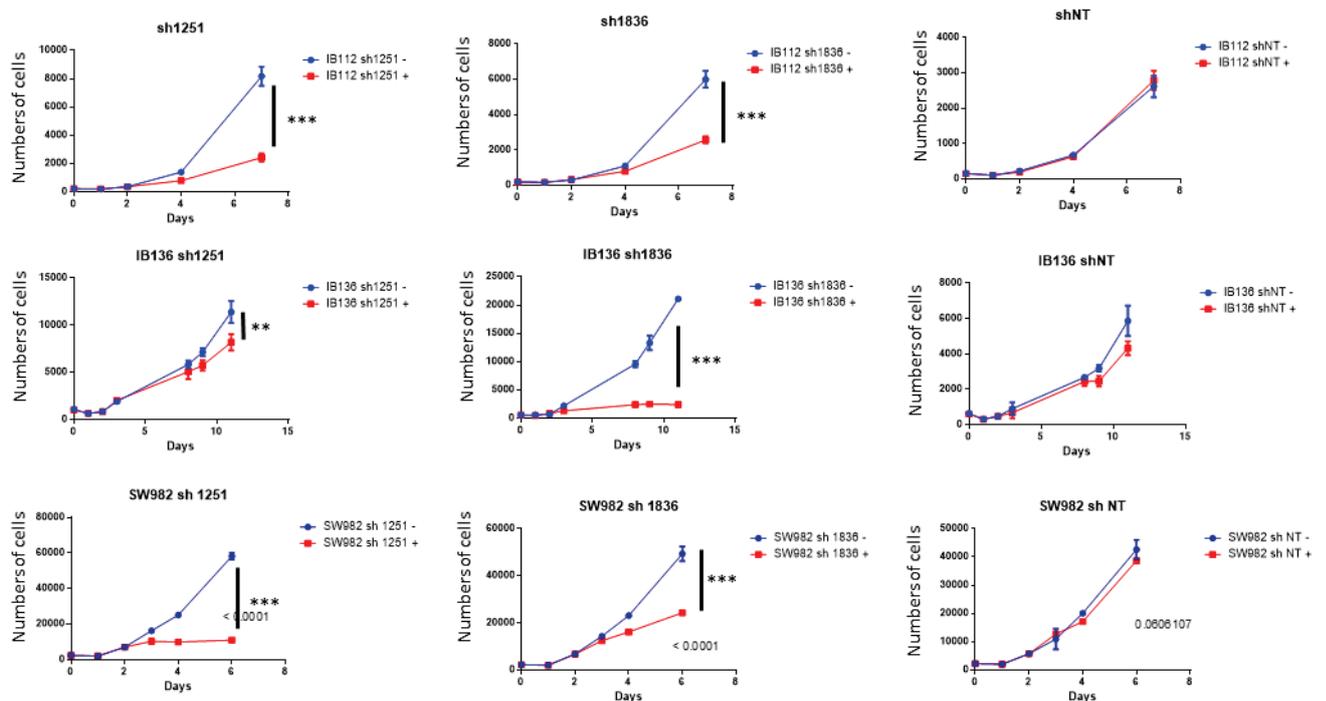


Figure 40: Courbes de prolifération en fonction de l'inhibition de Dp71

Sur la Figure 40 est représenté une expérience représentative des trois tests réalisés pour chaque lignée. Nous pouvons observer que pour l'ensemble des lignées, il n'y a pas de différence significative de prolifération entre les lignées n'exprimant pas le shNT (en bleu) par rapport aux lignées l'exprimant (en rouge). Concernant la lignée IB112, l'induction de l'expression des sh1251 et sh1836 entraîne une

diminution significative (p-value <0.0001) de la prolifération cellulaire. Pour la lignée IB136, nous pouvons noter également une diminution significative de la prolifération lorsque les sh1251 et sh1836 sont induits, néanmoins l'effet sur la prolifération est plus faible dans le cas de la lignée exprimant le sh1251. Cela est probablement dû à la faible inhibition exercée par le sh1251 dans la lignée IB136, comme observé sur le Western blot (Figure 39).

Concernant la lignée commerciale SW982, nous observons que les cellules exprimant le sh1251 présentent un arrêt de la prolifération, contrairement à la lignée contrôle dont la croissance n'est pas altérée (p-value <0.0001). Lorsque la lignée exprime le sh1836, nous observons une diminution significative de 50% de la prolifération cellulaire (p<0.0001). Pour l'ensemble des lignées, il apparaît alors que l'inhibition de Dp71 provoque une diminution voire dans certains cas l'arrêt de la prolifération des cellules.

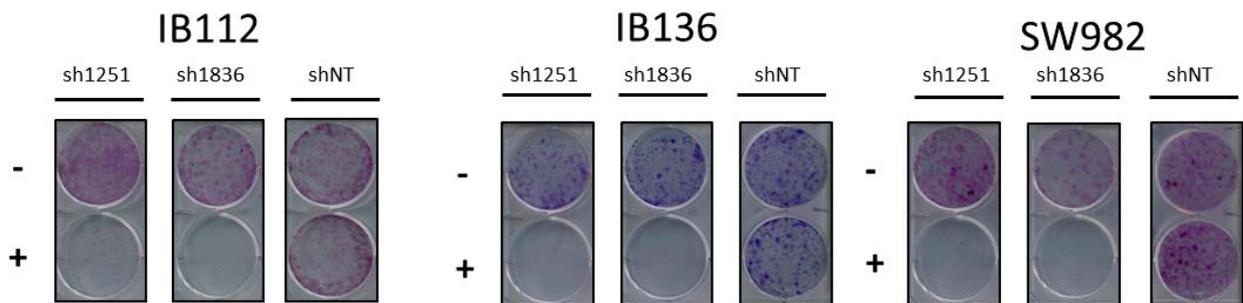


Figure 41 : Test de clonogénicité

En parallèle des expériences de prolifération cellulaire, j'ai réalisé des tests de clonogénicité (Figure 41) pour l'ensemble de ces lignées. Nous pouvons observer que pour les trois lignées, l'induction des sh1251 et sh1836 (notée +) inhibant Dp71 induit une diminution de la taille voire une absence de clone par rapport au contrôle non induit. Ceci peut être expliqué par l'incapacité des cellules à proliférer lorsque Dp71 est inhibée.

La diminution de la prolifération cellulaire suite à l'inhibition de Dp71 peut être due soit à un blocage des cellules dans une des phases du cycle cellulaire, soit à une augmentation de l'apoptose des cellules. N'observant pas de cellules mortes ou de débris cellulaires au microscope au cours des expériences précédentes, j'ai privilégié la

première hypothèse et donc effectué des études du cycle cellulaire par un marquage avec l'iodure de propidium et une analyse par cytométrie en flux pour valider cette hypothèse.

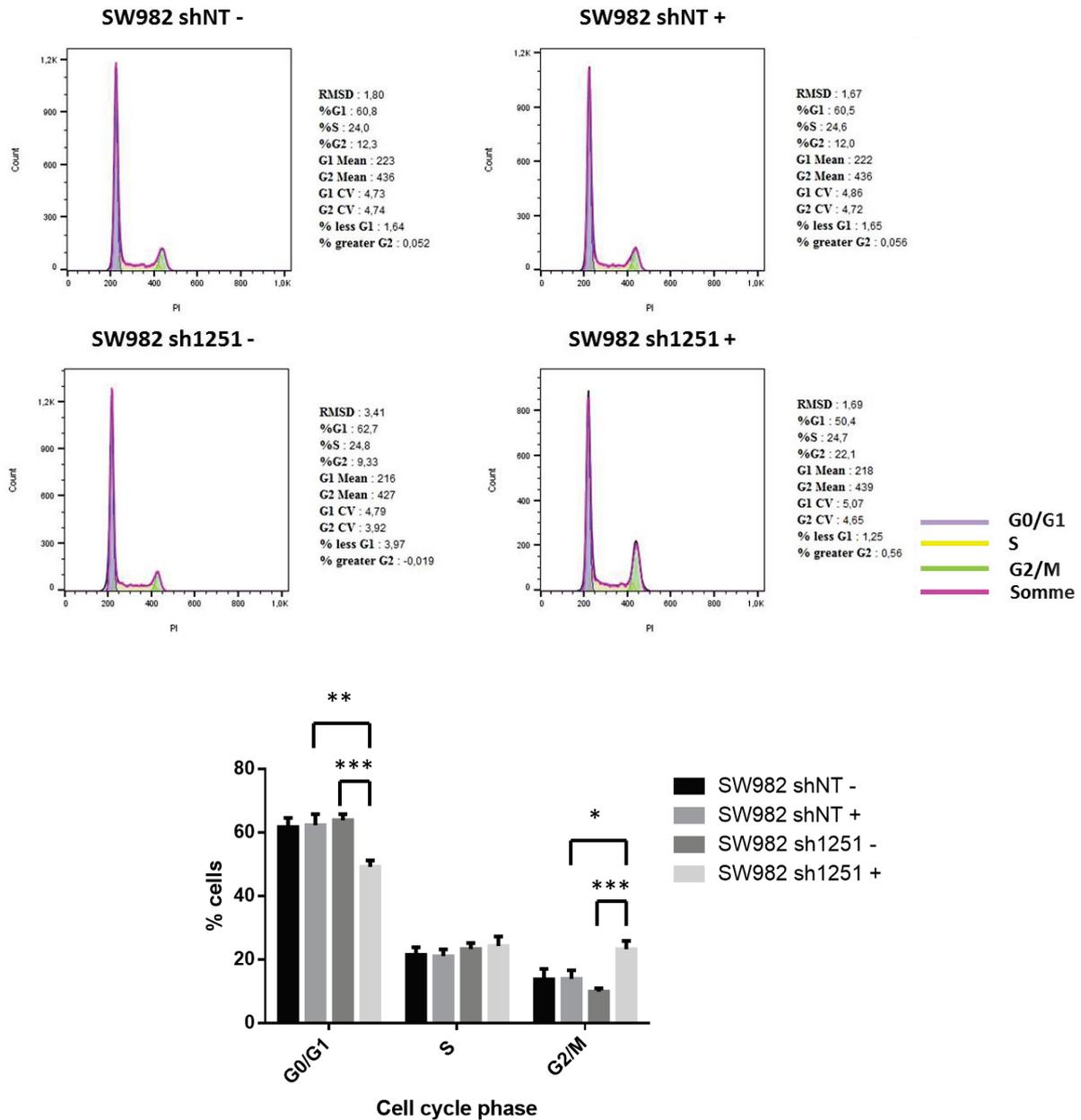


Figure 42 : Etude du cycle cellulaire par cytométrie en flux

Le marquage à l'iodure de propidium a montré que l'inhibition de Dp71 dans la lignée SW982 (Figure 42) provoque une accumulation des cellules en phases G2/M. Dans la lignée SW982 sh1251, le pourcentage moyen de cellules en phase G2/M est de

10% en condition contrôle et augmente à 23% lorsque l'expression du sh1251 est induite. Ainsi, la diminution de la prolifération est due à un blocage des cellules en phase G2/M au détriment de la phase G0/G1. Cette expérience a été réalisée pour l'ensemble des shARN de la lignée SW982 et le résultat a été confirmé dans la lignée exprimant le sh1836. Pour les lignées IB112 et IB136, les expériences doivent être confirmées.

La diminution de la prolifération due à l'inhibition de Dp71, qui provoque un blocage des cellules en phases G2/M du cycle cellulaire, semble renforcer l'hypothèse selon laquelle Dp71 pourrait jouer un rôle oncogène. Pour renforcer cette hypothèse, j'ai alors étudié l'effet de la sur-expression de cette isoforme dans une lignée de myoblastes « Myo D6 » fournie par le Dr Bénédicte CHAZAUD.

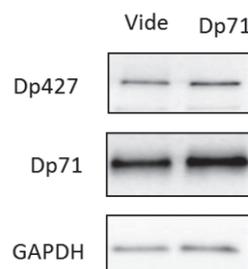


Figure 43: Western blot de validation de la sur-expression de Dp71

Après avoir infecté la lignée de myoblaste Myo D6 avec soit le vecteur contrôle, soit le vecteur contenant le l'ADNc de Dp71, j'ai validé sa sur-expression par Western blot (Figure 43).



Figure 44 : Tests de prolifération et de clonogénicité en fonction de la sur-expression de Dp71

Sur la

Figure 44, nous pouvons observer que la sur-expression de Dp71 ne semble pas avoir d'impact sur la prolifération ou sur la capacité des myoblastes à former des clones.

Synthèse des résultats

Ainsi, lors de ma thèse, j'ai pu montrer que l'altération de *Dp427* dans les sarcomes se produit principalement par un mécanisme de délétion et non de mutation. L'analyse RNA-seq des tumeurs a révélé que seule l'isoforme Dp71 est fortement exprimée, ce qui concorde avec nos observations au niveau génomique sur le fait que Dp71 est toujours épargnée par la délétion affectant *DMD*. L'étude d'une lignée de LMS dans laquelle j'ai inhibé spécifiquement Dp427 grâce au système CRISPr/Cas9 montre que son inactivation ne modifie ni la prolifération cellulaire, ni la clonogénicité de la lignée. Etant donné que la délétion de Dp427 n'a pas d'effet sur l'agressivité des cellules, j'ai testé l'hypothèse selon laquelle Dp71 aurait un rôle oncogénique. Ainsi, l'inhibition de son expression par shARN dans deux lignées de LMS et une lignée de synoviosarcome révèle que Dp71 est nécessaire à la croissance cellulaire et à la formation de clones, car son inhibition entraîne un blocage du cycle cellulaire en phase G2/M. J'ai alors voulu tester le rôle oncogénique de Dp71 dans une lignée de myoblaste. La surexpression de Dp71 n'a pas eu d'impact sur la prolifération et la clonogénicité de ces cellules, toutefois la forte expression basale de cette protéine peut expliquer l'absence d'effet notable. En effet, les cellules musculaires immatures expriment Dp71, qui est indispensable pour leur renouvellement, et ne perdent son expression au profit de celle de Dp427 qu'au cours de la différenciation myogénique. On peut donc émettre l'hypothèse que la délétion de Dp427 dans des cellules musculaires associée à une forte expression de Dp71 pourrait induire une dédifférenciation lors de la transformation oncogénique.

Discussion
et
Perspectives

Lors de ma thèse, j'ai montré d'une part que la sur-expression de *RCBTB1* augmente le pourcentage de cellules mitotiques, ceci ayant pour conséquence d'une part d'augmenter la croissance tumorale *in vivo* et d'autre part de sensibiliser les cellules à l'apoptose induite par le docétaxel. Je discuterai alors dans un premier temps de *RCBTB1* en tant que nouveau marqueur biologique. Puis, dans un second temps, je relierai mes résultats obtenus avec les partenaires de *RCBTB1* identifiés dans la bibliographie afin d'émettre des hypothèses quant à son rôle au niveau cellulaire, notamment son effet sur la mitose. Je finirai sur l'étude de l'altération de *RCBTB1* dans d'autres modèles tumoraux afin d'élargir l'intérêt de gène dans la biologie des cancers.

La deuxième partie de mes travaux de recherche sur le gène *DMD* a montré que l'isoforme Dp71 est essentielle à la progression tumorale. Etant donné que je n'ai pas constaté d'effet phénotypique lors de l'inhibition par délétion de Dp427, je mettrai alors en relation les dernières découvertes sur les fonctions des différentes isoformes du gène *DMD* avec mes résultats pour proposer de nouvelles perspectives d'étude de ce gène, visant à comprendre le rôle de la délétion 5' de *DMD* dans l'oncogenèse ou la progression tumorale des tumeurs. Notamment, je développerai l'effet de la perte de Dp427 sur l'instabilité génomique observé dans des modèles murins.

Dans la dernière partie de ma discussion, j'ai souhaité aborder deux problématiques que j'ai rencontrées lors de ma thèse. La première est : comment obtenir des informations précises et biologiquement pertinentes à partir de données informatiques volumineuses ? J'essaierai de donner quelques perspectives pour améliorer ces problèmes récurrents à l'ensemble des équipes de recherche qui utilisent les nouvelles techniques de séquençage à haut débit. Le deuxième aspect est une réflexion sur la pertinence de rechercher encore à l'heure actuelle des gènes « driver » de l'oncogenèse ou de la progression tumorale.

I- *RCBTB1* : un nouveau marqueur biologique ?

L'objectif initial de ce projet était d'identifier des gènes « driver » de l'oncogenèse ou de la progression tumorale des sarcomes à génétique complexe. Pour cela, notre équipe a élaboré une stratégie originale partant de l'ensemble des altérations génomiques présentes dans les tumeurs ayant des répercussions au niveau transcriptomique, les reliant aux voies biologiques dans lesquelles les gènes altérés interviennent, puis identifiant les voies altérées dans 100% des sarcomes. Grâce à cette méthodologie, un gène dont la délétion est associée à la métastase et au sous-type histologique des LMS a été identifié : *RCBTB1*. Je me suis donc ensuite attaché à élucider le rôle de *RCBTB1* dans ces tumeurs.

1) *RCBTB1* : un biomarqueur de la réponse au traitement ?

J'ai ainsi pu constater que *RCBTB1* n'avait de valeur pronostique de l'évolution métastatique que pour les patients dont le traitement incluait de la chimiothérapie. Il s'est avéré que la surexpression de *RCBTB1* sensibilise les lignées de LMS au docétaxel, un agent anti-mitotique qui perturbe la dynamique des microtubules lors de la mitose. Cette sensibilisation est due à une augmentation de la proportion de cellules en mitose *in vitro*. Bien que les tests de prolifération n'aient pas mis en évidence de différence de croissance entre les lignées surexprimant *RCBTB1* et les lignées contrôles dans les conditions que j'ai utilisées, l'effet de *RCBTB1* sur la mitose est confirmé *in vivo* par la croissance beaucoup plus rapide des tumeurs formées par la lignée IB136-*RCBTB1*. J'ai ainsi également confirmé *in vivo* que l'expression de *RCBTB1* augmentait l'efficacité du docétaxel sur la croissance tumorale. Étant donné que la plupart des molécules cytotoxiques utilisées en clinique ciblent les cellules en prolifération, comme la gemcitabine et la doxorubicine, il n'est pas exclu que *RCBTB1* augmente leur efficacité *in vivo*. Et ce, malgré le fait que mes conditions d'expérimentation ne m'aient pas permis d'observer *in vitro* un effet de *RCBTB1* sur la réponse apoptotique induite par ces deux drogues. Si cela se confirmait, ce pourrait être expliqué par le fait qu'*in vitro*, nous n'observons pas de différence de prolifération. Cependant, étant donné nos résultats avec le docétaxel et ceux de Lee *et al* (2013) avec le nocodazole²²⁰, il est plus vraisemblable que l'effet de *RCBTB1* soit

plus spécifique des drogues ciblant la dynamique des microtubules lors de la mitose telles que le docétaxel, le paclitaxel ou l'éribuline. Mes recherches sur *RCBTB1* ouvrent donc de nouvelles perspectives quant à la sensibilisation des tumeurs et ainsi il serait intéressant de tester *in vivo* l'efficacité de ces chimiothérapies en fonction de l'expression ou non de *RCBTB1*. Nous pourrions également envisager, dans des cohortes plus importantes, de déterminer la valeur pronostique de *RCBTB1* en fonction des chimiothérapies administrées aux patients. Selon la spécificité de la valeur pronostique de *RCBTB1* vis-à-vis des chimiothérapies, c'est-à-dire, si *RCBTB1* est pronostique à une, plusieurs ou toutes les chimiothérapies, la valeur d'expression de *RCBTB1* pourrait être utilisée comme un nouveau biomarqueur de la réponse à certaines chimiothérapies. Ceci permettrait d'améliorer la prise en charge thérapeutique des patients afin d'augmenter leur chance d'éviter la progression métastatique de leur maladie, facteur qui à l'heure actuelle reste la plus fréquente cause de décès des patients.

2) RCBTB1 : un nouveau biomarqueur... avec quels partenaires ?

La valeur pronostique de *RCBTB1* semble donc reposer notamment sur son rôle dans la réponse au traitement. Bien que les mécanismes moléculaires de son mode d'action demeurent inconnus, il se pourrait que son rôle soit en lien avec la Culline 3 (*CUL3*).

En effet, cette dernière est décrite comme régulant la mitose, l'apoptose ou encore la réponse au stress dans le cadre d'interactions avec d'autres partenaires possédant un domaine BTB⁹⁷. *CUL3* est considérée comme ayant un rôle suppresseur de tumeur dans certains cancers comme les cancers du sein²³³ et du poumon²³⁴. Toutefois, son rôle vis-à-vis du développement tumoral est ambivalent et dépend en réalité de la nature de ses partenaires. Ainsi, l'un de ses partenaires les plus connus est la protéine Keap1 qui permet la reconnaissance puis la dégradation du facteur de transcription Nrf2, qui active l'expression de gènes impliqués dans la réponse au stress oxydatif²³⁵. Une forte expression de Nrf2 induisant une chimiorésistance des tumeurs²³⁶, il apparaît attractif d'activer sa dégradation par la voie *CUL3*-Keap1. Cependant, il est également connu que le stress oxydatif promeut la carcinogenèse, ce

qui implique donc ici qu'activer la dégradation de Nrf2 favoriserait l'oncogenèse. Ainsi, le rôle du complexe CUL3-Keap1 dépend fortement du contexte cellulaire et son utilisation thérapeutique est distincte selon que l'on souhaite prévenir ou traiter le cancer. Un autre exemple est fourni par le complexe formé par CUL3 et SPOP, dont l'effet sur la progression tumorale dépend du substrat ciblé pour la dégradation²³⁵. En effet, ce complexe ubiquitinye d'une part les récepteurs aux androgènes (AR) responsables de la progression tumorale et de la résistance au traitement dans le cancer de la prostate²³⁷, ayant ainsi un effet suppresseur de tumeur, et d'autre part induit la dégradation de PTEN et BRMS1, favorisant ainsi la progression tumorale respectivement dans les cancers du rein et du sein²³⁵. Il se peut donc que la perte de *RCBTB1* participe à l'oncogenèse des sarcomes du fait de l'accumulation de substrats oncogènes normalement adressés à la dégradation par les complexes qu'il forme avec CUL3, et que ses effets sur la mitose et peut-être les voies de réponse au stress impliquent la reconnaissance de différents substrats.

De plus, son effet spécifique vis-à-vis du docétaxel suggère une action liée à la mitose et la dynamique des microtubules. Or, comme présenté dans l'introduction et l'article en cours de révision, *RCBTB1* interagit avec GABARAP, GABARAPL1, GABARAPL2, MAP1LC3A et MAP1LC3B. Ces partenaires sont tous impliqués dans le processus autophagique et interagissent avec les microtubules. L'impact de l'autophagie sur la mitose fait toujours l'objet de débats²³¹. En effet, le fait que l'autophagie persiste au cours de la mitose est sujet à controverses. Néanmoins, il est indubitable que l'autophagie est nécessaire pour fournir l'énergie et les nutriments qui seront consommés au cours de la mitose. De plus, la dérégulation des gènes de l'autophagie aboutit à une mitose prolongée ou à une anomalie dans le rassemblement des chromosomes ou l'assemblage des kinétochores, ce qui renforce l'idée selon laquelle l'autophagie est active durant la mitose²³¹. Etant donné la nature des partenaires d'interaction de *RCBTB1*, il est possible que son effet sur l'augmentation du pourcentage de cellules mitotiques dans les lignées de LMS puisse être expliqué par une régulation du processus autophagique.

De façon intéressante, le complexe ubiquitine ligase CUL3-KBTBD6/KBTBD7 coopère avec GABARAP pour la dégradation ciblée de substrats spécifiques²³². Ainsi, le rôle hypothétique de RCBTB1 dans l'autophagie pourrait avoir lieu grâce à une coopération entre les complexes recrutés par CUL3, les protéines de la famille LC3/GABARAP et la protéine adaptatrice RCBTB1, aboutissant à la dégradation spécifique de substrats induite par l'autophagie. Toutefois, d'autres hypothèses pourraient également être testées. En effet, nous avons montré qu'une faible expression de GABARAP est associée à l'évolution métastatique. Or le rôle suppresseur de tumeur de GABARAP a déjà été démontré : l'injection de cellules de mélanomes exprimant GABARAP dans des souris *GABARAP*^{-/-} réduit la croissance tumorale²³⁸, certainement de façon directe par son implication dans l'autophagie et indirecte *via* la régulation de récepteurs tels qu'EGFR^{239,240}. En effet, la famille GABARAP est aussi impliquée dans le trafic vésiculaire et l'internalisation de récepteurs membranaires²⁴¹. Ainsi, RCBTB1 et ses partenaires d'interaction pourraient réguler les voies de signalisation par des processus tels que l'autophagie ou encore la régulation de la localisation des récepteurs membranaires et la signalisation qui en découle.

Une autre piste de recherche intéressante est celle concernant RCBTB2. En effet, RCBTB1 et RCBTB2 sont capables de s'hétérodimériser et pourraient ainsi participer à la reconnaissance de substrats par les complexes recrutés par la CUL3. Le gène *RCBTB2* étant situé entre les gènes *RBI* et *RCBTB1*, il est également fréquemment délété dans les sarcomes à génétique complexe. L'étude de l'existence d'une coopération fonctionnelle entre ces deux paralogues permettra peut-être d'identifier de nouveaux mécanismes impliquant RCBTB1 dans la biologie des sarcomes.

3) *RCBTB1* : un gène d'intérêt dans d'autres cancers ?

De manière plus générale, le locus 13q14 sur lequel *RCBTB1* est situé est fréquemment perdu dans plusieurs cancers comme indiqué en introduction. D'un point de vue biologique, si l'on considère le lien entre RCBTB1 et CUL3, aucun argument ne montre la spécificité de leur fonction uniquement dans les LMS. Au contraire,

RCBTB1 participerait à des mécanismes biologiques plus globaux, et donc pourrait servir de marqueur de réponse à la chimiothérapie dans d'autres cancers, comme notamment les leucémies où sa perte est également associée à un mauvais pronostic⁷⁷. Sur le site www.cbioportal.org, recensant les données de séquençage à haut débit du TCGA, nous observons des délétions de *RCBTB1* dans 10 à 23% des cas de cancer de la prostate. Grâce aux données de l'ICGC, nous observons que *RCBTB1* est altéré dans 36.25% des mélanomes, 35.52% des cancers de la peau et dans 28.79% des cancers du poumon. Ainsi, l'altération de *RCBTB1* ne se limite pas aux sarcomes à génétique complexe et renforce donc ainsi l'intérêt de l'étude de ce gène dans le domaine de l'oncologie.

4) Perspectives

Concernant ce projet, les perspectives immédiates sont de vérifier dans un premier temps que l'inhibition de *RCBTB1* par shARN provoque une diminution du pourcentage de cellules mitotiques et éventuellement une diminution de la prolifération cellulaire. Pour cela, j'utiliserai les lignées de fibroblastes que j'ai déjà établies pour étudier l'instabilité chromosomique. Dans un second temps je vérifierai que l'inhibition de *RCBTB1* provoque une résistance au docétaxel *in vitro*. Je testerai également son effet sur la mort cellulaire induite par la doxorubicine et la gemcitabine. En effet, dans la mesure où je n'observe pas de différence de prolifération *in vitro* dans les lignées de LMS lorsque *RCBTB1* est sur-exprimé, il est possible que je n'observe pas d'effet de la doxorubicine ni de la gemcitabine *in vitro* si son inhibition n'altère pas la prolifération cellulaire. Néanmoins, au vu de l'augmentation de la croissance tumorale *in vivo* lorsque *RCBTB1* est surexprimé, et comme il est couramment admis en clinique que les tumeurs proliférant plus rapidement sont plus sensibles aux chimiothérapies, il est raisonnable de poser l'hypothèse que RCBTB1 sensibilise à d'autres chimiothérapies. Tester cette hypothèse *in vivo* et sur les modèles d'inhibition *in vitro* permettra de déterminer la spécificité ou non de la sensibilisation à l'apoptose par RCBTB1 vis-à-vis du docétaxel.

II- *DMD* : un gène complexe associé à la progression tumorale des sarcomes ?

Notre équipe a identifié par CGH-array un gène localisé sur le chromosome X qui est fréquemment perdu dans une cohorte de 317 sarcomes des tissus mous (STM). Il s'agit du gène *DMD*, perdu dans 16.7% des tumeurs. Ce pourcentage de délétion est particulièrement élevé en comparaison avec d'autres études rapportant une délétion dans 5.5% des lignées cellulaires de mélanomes¹⁹⁰ et une fréquence médiane de 3.4% dans 8052 tumeurs de différents types histologiques analysées par Luce *et al.* (2016)¹⁹³. Cependant, Wang *et al.*, (2014) ont observé une délétion dans 25/40 tumeurs myogéniques de haut grade, soit 63% des cas¹⁸⁹. Ces différences peuvent probablement s'expliquer d'une part, par la nature des différents types histologiques étudiés, puisque les délétions de *DMD* semblent concerner principalement les tumeurs d'origine musculaire. D'autre part, nous avons pu constater grâce à l'étude de notre cohorte que la délétion de *DMD* était associée à l'évolution métastatique. Or, Wang *et al.* s'étant intéressés aux tumeurs de haut-grade uniquement, cela peut être une des raisons pour lesquelles leur pourcentage de délétion dans les tumeurs myogéniques est bien supérieur à celui que nous observons dans notre cohorte (comprenant des GISTs, LMS et SSX).

1) Dp427 : suppresseur de tumeur – Dp71 : oncogène ?

Nous avons observé trois types de délétion dans notre cohorte de tumeurs : délétions du chromosome X, du bras court du chromosome X ou dans 61.1% des cas des délétions ciblant la partie 5' du gène *DMD*. Le gène *DMD* codant pour plusieurs isoformes, j'ai ensuite cherché à déterminer le rôle des isoformes Dp427 (toujours affectée par la délétion) et Dp71 (épargnée par les délétions ciblées du gène *DMD*). L'isoforme Dp427 code pour la dystrophine, protéine dont l'altération est connue pour être responsable des dystrophies musculaires de Duchenne et de Becker. Outre son implication dans les myopathies, des études rapportent un rôle de gène suppresseur de tumeur dans le mélanome¹⁹⁰ et les sarcomes myogéniques¹⁸⁹. Concernant Dp71, son rôle dans les cancers semble plus controversé, car la littérature peu exhaustive à ce sujet fait état de résultats contradictoires selon les modèles de cancer étudiés^{194,195}.

Pour déterminer le rôle de Dp427, nous avons étudié l'effet de son inhibition dans la seule lignée cellulaire l'exprimant que nous possédons : la lignée de LMS

IB133. L'inhibition par délétion n'ayant eu aucun effet sur les phénotypes étudiés, la perte de Dp427 ne serait pas essentielle pour l'acquisition des phénotypes d'agressivité tumorale. Or, la délétion de la région 5' de *DMD* étant associée à l'évolution métastatique, nous avons posé l'hypothèse que ce ne serait pas la perte de fonction de Dp427 qui serait impliquée dans l'évolution métastatique, mais que sa délétion permettrait à Dp71 d'assurer ses fonctions oncogéniques, soit par libération des sites de liaison des facteurs de transcription de Dp71 afin d'augmenter son expression, soit par un arrêt de la compétition entre les deux isoformes Dp427 et Dp71 pour la formation de complexes DAPC.

J'ai alors étudié l'effet de l'inhibition de Dp71 dans trois lignées n'exprimant que cette isoforme. J'ai pu démontrer que Dp71 est essentielle à la prolifération cellulaire et à la formation de clones car son inhibition provoque un blocage des cellules en phase G2/M. Cette observation est cohérente avec l'étude de Villarreal-Silva *et al*, en 2011 qui ont montré la localisation de Dp71 aux pôles du fuseau mitotique lors de la division cellulaire¹⁶⁴. Ils ont postulé que Dp71 se lie à la lamine B1 et aux β -dystroglycanes, permettant ainsi de stabiliser le complexe nécessaire à la division cellulaire^{164,165}. J'ai ainsi mis en évidence que dans les STM, Dp71 est essentielle à la division cellulaire et donc au développement des tumeurs. Ceci apporte donc une explication biologique au fait que cette isoforme soit majoritairement exprimée par les tumeurs et que les délétions ciblant *DMD* la préservent systématiquement. Dp71 étant exprimée dans toutes les tumeurs, il apparaît alors intéressant de développer des inhibiteurs de cette protéine à des fins thérapeutiques.

Cependant, l'ensemble de mes données ne me permet pas de conclure quant à un rôle oncogène de Dp71 car je n'observe pas d'augmentation de l'agressivité des myoblastes sur-exprimant Dp71. Pour cela, il serait nécessaire de montrer une augmentation de son expression dans le tissu tumoral par rapport au tissu sain, et/ou que sa sur-expression provoquerait une augmentation de la prolifération cellulaire et amènerait les cellules vers un processus de dédifférenciation. Dans nos lignées de myoblastes, il apparaît que Dp71 est fortement exprimée de manière basale, ceci pouvant expliquer le fait que la sur-expression dans ces lignées n'ait aucun effet.

2) Origine de la délétion 5' de *DMD* ?

Il a été montré que la perte de Dp427 provoque le développement spontané de rhabdomyosarcomes dans les modèles murins de dystrophie musculaire de Duchenne. La délétion de la région 5' de *DMD* pourrait donc être un événement précoce de l'oncogenèse des sarcomes. *DMD* est l'un des plus longs gènes humains (2.2Mb) et il est localisé au niveau d'un site fragile. Les sites fragiles sont des régions de profonde instabilité génomique, hautement susceptibles de subir des recombinaisons homologues dans les cellules normales²⁴² et reconnues comme impliquées dans les phases initiales de l'oncogenèse²⁴³. En effet, les réarrangements chromosomiques qui en découlent aboutissent souvent à l'amplification d'oncogènes et/ou l'inactivation de gènes suppresseurs de tumeur. Cette instabilité génomique au niveau des longs gènes s'explique notamment par les phénomènes de collision entre les complexes de réplication et de transcription. En effet, Helmrich *et al.*, (2011) ont rapporté que la transcription de ces longs gènes commençait lors de la phase G2/M du cycle cellulaire et se poursuivait jusqu'à la phase S du cycle suivant²⁴⁴. Cette observation a de nombreuses implications. Tout d'abord, la présence de pré-ARNm sur les chromosomes mitotiques peut interférer avec la condensation appropriée de la chromatine et donc, ils sont localisés au niveau de zones de fragilité génomique. De plus, cela signifie que l'arrêt de la transcription ou de la progression de la fourche de réplication mènent à la création d'hybrides ARN/ADN stables qui forment des structures en tige-boucle, responsables des cassures d'ADN double-brin et donc de l'instabilité génomique. Dans cette même étude, Helmrich et collaborateurs ont montré que la fréquence des cassures au niveau des sites fragiles n'était pas corrélée à la longueur du gène mais à son niveau de transcription²⁴⁴. Ceci confirme que la délétion de *DMD* est particulièrement pertinente dans les sarcomes myogéniques et plus largement dans les sarcomes dérivés de cellules exprimant *DMD*.

Cette observation m'amène étayer la thèse d'une origine myogénique des synoviosarcomes, dans lesquels nous observons une délétion de *DMD* dans 19 tumeurs sur les 88 étudiées. Cette hypothèse quant à l'origine musculaire des synoviosarcomes a déjà été avancée par Haldar *et al.*, en 2007 : lorsque la protéine

de fusion SYT—SSX2 est exprimée dans des myoblastes de souris, il y a développement de tumeurs ressemblant histologiquement aux synovialosarcomes humains²⁴⁵. Si cette protéine de fusion est exprimée plus tôt lors de la différenciation musculaire, cela provoque la mort de l'embryon, tandis que si elle est exprimée plus tard comme dans les myocytes ou dans les myofibres, cela provoque une myopathie. Une origine musculaire, partiellement engagée dans le processus de différenciation, des synovialosarcomes est donc tout à fait envisageable. Ainsi, il apparaîtrait que la délétion de *DMD* s'observe également dans des cellules d'origine musculaire peu ou pas différenciées, laissant ainsi penser que son expression n'est pas restreinte aux cellules musculaires squelettiques différenciées.

3) Conséquences de la délétion 5' de *DMD*

Concernant la délétion de Dp427 qui est associée à l'évolution métastatique dans notre cohorte, nous pouvons formuler deux hypothèses. Outre le rôle éventuel de Dp71 suite à la délétion, la perte de Dp427 pourrait être impliquée dans l'acquisition par les cellules cancéreuses de capacités migratoires et/ou invasives augmentées. Cette hypothèse est confortée par les travaux de Körner *et al* (2007) dans le mélanome¹⁹⁰ et de Wang *et al* (2014) dans les sarcomes myogéniques¹⁸⁹, qui ont montré que la ré-expression de Dp427 dans des lignées cancéreuses réduisait la migration et l'invasion cellulaires. Deuxièmement, il a été montré que la perte de *dmd* induisait une instabilité génomique dans les cellules musculaires des souris *mdx*, qui sont des modèles murins pour la myopathie de Duchenne. Précédemment, notre équipe a montré que l'instabilité génomique dans les sarcomes était significativement liée à l'évolution métastatique. Nous pouvons alors émettre l'hypothèse que la délétion de *DMD* pourrait causer une instabilité génomique, qui aurait alors pour conséquence d'augmenter la probabilité d'évolution métastatique pour ces tumeurs.

Les dystrophies musculaires présentent des caractéristiques autres que la dégénérescence musculaire, telles que la prolifération des tissus adipeux et fibreux. Ceci suggère que la perte de *DMD* n'affecte pas seulement les cellules musculaires squelettiques mais également le tissu mésenchymateux environnant et que les sarcomes portant une délétion de *DMD* pourraient provenir d'une cellule souche

multipotente mésenchymateuse ou dérivée d'une cellule musculaire. Ce point est renforcé par les travaux de Schmidt *et al.* (2011), qui rapportent que 39% des souris dystrophiques *mdx* (mutées pour *dmd*) développent spontanément avec l'âge des sarcomes mixtes comprenant des composantes de type rhabdo-, fibro- et liposarcomes, tous caractérisés par une instabilité génomique et des mutations récurrentes (*Trp53*, *Nf1*, *Cdkn2a...*)¹⁸⁷, que l'on retrouve dans les tumeurs humaines. Par ailleurs, ils ont détecté des mutations de type cancer et de l'aneuploïdie dans le muscle dystrophique, associées à des réponses de dommage à l'ADN prononcées dans les échantillons murins et humains non tumoraux. Ceci apporte la preuve que l'instabilité génomique et les réarrangements chromosomiques ont lieu avant la transformation oncogénique.

Enfin, ces résultats concernant le développement spontané de sarcomes mixtes ont également été obtenus avec des souris présentant des mutations affectant d'autres gènes impliqués dans les dystrophies (*Dysf*, *Capn3* et *Large*), suggérant ainsi qu'il existe une voie commune impliquant tous ces acteurs et préservant l'intégrité du génome. Le fait que les cas de sarcomes chez des patients atteints de myopathie de Duchenne soient extrêmement rares peut être expliqué par les différences entre la biologie des souris et celle des humains. En effet, la perte de la dystrophine est fatale chez les humains tandis qu'elle n'entraîne qu'une myopathie moyenne chez la souris. Enfin, les souris sont plus susceptibles de développer des tumeurs mésenchymateuses tandis que les humains sont plus prompts aux carcinomes et nécessitent plus d'événements mutationnels que les souris pour développer des tumeurs²⁴⁶.

Malgré ces différences, les observations faites chez les souris myopathes offrent de nouvelles perspectives de recherche pour la compréhension de l'oncogénèse des sarcomes. Puisque le gène *DMD* appartient à un site fragile, il est plus susceptible aux altérations génétiques et sa délétion pourrait provoquer une instabilité génomique, permettant l'amplification d'oncogènes et l'inactivation de gènes suppresseurs de tumeur. En conclusion, la délétion de *DMD* serait un événement « driver » de l'oncogénèse des sarcomes, favorisant le développement de tumeurs agressives aux profils génomiques remaniés, ces deux aspects étant pronostiques de l'évolution métastatique... Mais comment ?

Nos connaissances actuelles sur la dystrophine ne permettent pas encore de répondre à cette question. Cependant, de nouveaux rôles longtemps insoupçonnés sont en train d'émerger quant à Dp427 car jusqu'à récemment, l'essentiel des travaux étaient menés sur du muscle différencié. Des myoblastes isolés de souris *mdx* et de patients atteints de DMD présentent une prolifération et une différenciation altérées, associées à un métabolisme énergétique anormal et une mort autophagique, tous ces phénotypes perdurant dans des conditions de culture cellulaire en laboratoire²⁴². Ceci suggère donc que la perte de la dystrophine n'impacte pas uniquement le muscle différencié et que ces phénotypes sont indépendants du micro-environnement. Comment expliquer ces observations dans la mesure où les myoblastes n'expriment pas Dp427 ? Une des explications pourra éventuellement être fournie par l'étude approfondie du gène *DMD*. En effet, ce locus code également pour plusieurs longs ARN non codants (lncRNA, *long non-coding RNA*) dont la fonction n'est pas encore élucidée mais dont l'expression est réduite dans les myoblastes isolés de patients DMD²⁴⁷. Il ne s'agit cependant pas de la seule explication possible. En effet, le dogme selon lequel l'isoforme Dp427 n'est exprimée que dans le muscle différencié dans lequel elle ne participe qu'à sa structure est récemment tombé : Dumont *et al.* (2015) ont ainsi démontré que la dystrophine était exprimée dans les cellules satellites activées, dans lesquelles elle régule la polarité cellulaire lors de la division asymétrique²⁴³. Cette division asymétrique permet la régénération du muscle en orientant une cellule fille vers la différenciation musculaire et l'autre vers l'auto-renouvellement afin de maintenir le pool de cellules satellites. L'expression transitoire de la dystrophine permet la localisation adéquate des marqueurs de polarité Mark2 et Pard3 et sa perte implique donc une dégénérescence musculaire d'une part due à la fragilité des myofibres et d'autre part due à une régénération altérée du fait de la dysfonction des cellules satellites. Cette étude apporte la preuve formelle que Dp427 intervient dans les cellules souches du tissu musculaire. S'ajoute à cela le fait établi que les cellules satellites régulent également la prolifération des fibroblastes qui engendrent le tissu conjonctif. Or l'injection d'iPSCs (induced-Pluripotent Stem Cells) exprimant la dystrophine dans des blastocystes mutants *mdx* corrige la masse grasseuse des souris²⁴⁸. L'expression de la dystrophine dans les cellules souches du

tissu musculaire ouvre donc la voie à de nouveaux axes de recherche non seulement sur des fonctions de Dp427 autres que structurales mais également sur son effet sur le tissu conjonctif voisin, étroitement lié au muscle.

4) Perspectives

Afin de déterminer définitivement si la délétion 5' du gène *DMD* ciblant l'isoforme Dp427 joue un rôle dans l'initiation de l'oncogenèse ou de la progression tumorale, il faut tester dans des lignées exprimant Dp427 si son inhibition induit une instabilité chromosomique comme cela a déjà été décrit chez la souris. Pour cela, il faudra créer des modèles de myoblastes KO pour Dp427 et réaliser un marquage immunofluorescent de phospho- γ H2AX afin d'observer par microscopie une éventuelle augmentation du nombre de cassures double-brin. Cette analyse pourra être complétée par CGH-array qui permettra de déterminer précisément les altérations engendrées par la perte d'expression de Dp427. Enfin, nous pourrons répéter cette expérience dans d'autres lignées de myoblastes et éventuellement d'autres types cellulaires exprimant Dp427 afin d'observer si les altérations sont récurrentes pour un même type cellulaire et pour des types cellulaires différents.

De plus, il serait également intéressant de tester si l'inhibition de Dp427 augmente la migration ou l'invasion. En effet, Wang *et al.* ont montré que la surexpression de *miniDMD* provoque une inhibition de la migration et de l'invasion¹⁸⁹. Etant donné que ce modèle n'est pas le reflet de l'expression totale de Dp427, il serait intéressant de valider si son inhibition a un effet inverse.

III- Discussion générale

1) « Dig Data » : Comment filtrer les informations ?

Lors de ma thèse, je me suis attaché à comprendre l'impact de la délétion de deux gènes fréquemment altérés dans les sarcomes à génétique complexe. Au vu de la complexité de leur génome, ces tumeurs n'ont pour l'instant pas révélé les mécanismes à l'origine de leur oncogenèse. Grâce à des techniques d'étude pan-génomique

(comme la CGH-array) et pan-transcriptomique (comme les puces d'expression), nous avons pu identifier parmi les milliers de gènes altérés dans chaque tumeur deux gènes candidats avec un rôle important dans ces tumeurs. Cela montre que notre approche consistant à filtrer les gènes altérés par l'étude des voies biologiques a été pertinente. En effet, en raison de sa proximité avec *RBI*, le gène *RCBTBI* a été ignoré par les biologistes, qui considéraient que la délétion 13q14 ciblait uniquement le gène *RBI* impliqué dans le cycle cellulaire. Or notre analyse a mis en évidence de manière objective *RCBTBI* en plus de *RBI* (qui, dans notre cohorte, n'est pas significativement associé à la métastase).

A l'heure actuelle, des techniques d'étude du génome et du transcriptome bien plus exhaustives que les puces à ADN ou ARN ont fait leur arrivée dans nos laboratoires et nous apportent donc des quantités importantes de données à traiter. En plus des gènes gagnés ou perdus vont se rajouter les gènes mutés, méthylés... D'où la nécessité de filtrer les altérations observées avec une cohérence biologique. C'est ce que nous proposons grâce à notre analyse bio-informatique.

De manière plus générale, la rencontre entre la biologie et les mégadonnées est très récente. La plupart des experts s'accordent à dater cette rencontre au nouveau millénaire, lorsque le génome humain a été séquencé pour la première fois. Depuis, les données accessibles sur le site du NCBI dépassent le million de Giga octets. Ainsi, il est évident qu'un chercheur ne pourra à lui seul analyser ces données, d'une part par le manque de compétences et d'autre part par le manque de moyens physiques. De plus, ce flot de données ne va cesser d'augmenter avec le développement de projets de séquençage de nombreuses tumeurs comme le programme international ICGC et américain TCGA. Il est alors normal de voir le développement intense de la bio-informatique dont l'objectif est de faciliter l'accès des mégadonnées aux biologistes. Ainsi, la question qui se pose dans le domaine de la recherche biologique est : comment va évoluer l'analyse de ces données ? En arrivera-t-on à l'utilisation d'ordinateurs utilisant l'intelligence artificielle pour traiter ces données ? De même, afin de comparer les données entre les différentes équipes de recherche, il faudra alors

normaliser et homogénéiser les données. Ce sont ces objectifs qui devront être atteints afin d'exploiter au maximum les données que nous générons.

Ainsi, de la façon dont j'observe cette (r)évolution, je constate que la plupart des projets se concentrent sur l'acquisition des données aux volumes considérables et dans un second temps n'exploitent qu'une faible partie de ces informations. Ceci est dû au faible coût de ces techniques mais provoque par la suite une augmentation du coût dans le stockage et l'analyse de ces données. Nous pourrions alors envisager de redimensionner ces techniques d'acquisition afin d'apporter directement les informations finales, que nous scientifiques utilisons. Par exemple, envisager une technique pour l'analyse des tumeurs, qui donnerait de manière simple uniquement les nucléotides mutés par rapport à l'ADN contrôle du patient. Nous pourrions envisager une approche similaire à la technique DHPLC (Denaturing High-Performance Liquid Chromatography), permettant d'identifier les fragments d'ADN présentant une mutation grâce à la formation d'hétéroduplexes des brins mutés avec les brins sauvages. Par élution sur une colonne, les fragments d'ADN mutés peuvent être récupérés. En envisageant cette technique à haut débit, nous pourrions pour un échantillon, ne récupérer que les fragments d'ADN mutés qu'il suffirait de séquencer. Ainsi, après séquençage, nous aurions beaucoup moins de données à traiter et à filtrer, ce qui réduirait considérablement les coûts et le temps d'analyse.

2) Recherche de gène « driver » : Est-ce réellement pertinent ?

Notre analyse bio-informatique s'est révélée pertinente pour identifier *RCBTBI* en tant que gène supresseur de tumeur. Néanmoins, cette dénomination de gène supresseur de tumeur est à mon sens de plus en plus difficilement applicable. En effet, sa perte étant impliquée dans la progression tumorale, nous sommes tentés de le classer dans la catégorie de gène supresseur de tumeur. Néanmoins, mes expériences *in vitro* et *in vivo*, montrant que sa sur-expression provoque une augmentation du pourcentage de cellules en mitose ainsi qu'une augmentation de la croissance tumorale, suggèrent un rôle plutôt oncogène de *RCBTBI*. L'autre exemple est celui de *DMD*, que nous serions tentées de le classer comme *RCBTBI* en gène supresseur de tumeur de par la bibliographie et sa fréquente délétion. Néanmoins, nous voyons

nettement grâce à mes expériences sur Dp71 que nous ne pouvons pas classer le gène en entier en tant que gène supresseur de tumeur. Ainsi, il convient d'être prudent lors de l'observation d'altérations de gènes, puisqu'en fonction du contexte cellulaire, un gène pourra avoir des fonctions différentes.

Au vu de la grande diversité de ces tumeurs, il est certain qu'aucun gène n'expliquera à lui seul la biologie de ces sarcomes. Il est maintenant admis par la communauté scientifique qu'il n'existe pas un seul gène « driver » de l'oncogenèse d'une tumeur, mais plutôt une somme de verrous que la cellule doit débloquent avant de devenir tumorale, comme notamment *RBI* ou *TP53*. Mais l'altération de ces deux gènes ne suffit pas à induire des tumeurs avec un génome aussi complexe que ceux que nous observons. De plus, Davoli *et al.* en 2013, ont avancé l'hypothèse que l'aneuploïdie jouerait un rôle tout aussi important que la perte de gènes supresseurs de tumeur ou l'activation d'oncogènes en provoquant une haplo-insuffisance de certains gènes, augmentant ainsi la probabilité pour une cellule de devenir cancéreuse²¹⁹. Ainsi l'accumulation de l'haplo-insuffisance de plusieurs gènes et de l'activation d'oncogènes pourrait être à l'origine de l'oncogenèse d'une cellule plutôt que l'altération d'un seul et unique gène. Chaque tumeur étant unique d'un point de vue génétique, il semble impensable d'identifier un gène responsable de l'instabilité génomique dans ces tumeurs. Il faut alors résonner par voies biologiques altérées. En poussant le raisonnement que nous avons eu lors de l'établissement de notre analyse bio-informatique, nous pourrions déterminer quelles sont les voies biologiques spécifiquement altérées en fonction des annotations cliniques. Par exemple, quelles sont les voies biologiques altérées dans l'ensemble des sarcomes métastatiques ? En effet, pour comprendre la biologie de ces tumeurs, nous avons tendance à séparer les tumeurs en fonction de leurs sous-types histologiques, alors que nous voyons bien des ressemblances au niveau de la génétique de ces tumeurs. Pour exemple, les sarcomes indifférenciés et les leiomyosarcomes présentent des altérations communes. Il semble donc que les mécanismes donnant lieu à l'instabilité génomique dans ces sarcomes à génétique complexe soient communs et que la différence histologique soit plutôt due au programme de différenciation que la cellule a subi. Il pourrait alors être intéressant

d'étudier de manière simultanée les sarcomes présentant une génétique similaire plutôt qu'un même histotype.

En conclusion, l'histoire de l'étude moléculaire des sarcomes nous a montré qu'une meilleure compréhension de leur génétique permettait de mieux les diagnostiquer et de mieux les prendre en charge. Dans ce vaste groupe de sarcomes, il reste encore 50% de tumeurs avec une génétique complexe dans lesquelles aucun mécanisme moléculaire à l'origine de leur instabilité n'a été découvert. Néanmoins, mon travail de thèse a permis d'identifier un gène potentiellement impliqué dans l'initiation de la cellule tumorale, *DMD*, pour lequel le jeu d'expression entre les différentes isoformes lors de la différenciation pourrait conduire, suite à une délétion de ce gène, à une instabilité chromosomique et à une dérégulation de la différenciation de la cellule aboutissant aux tumeurs à remaniements complexes que l'on peut observer en clinique. Enfin, j'ai également pu caractériser un gène impliqué dans la progression tumorale, *RCBTBI*, dont l'expression augmente d'une part la croissance tumorale *in vivo* mais qui d'autre part est un potentiel biomarqueur de la réponse à la chimiothérapie notamment dans les LMS, montrant ainsi que certains acteurs aux rôles pro-oncogènes peuvent en réalité représenter des talons d'Achille que nous pouvons exploiter pour le traitement de ces tumeurs.

Références

1. WHO classification of tumours of soft tissue and bone - NLM Catalog - NCBI. Available at: <https://www.ncbi.nlm.nih.gov/nlmcatalog/101622555>. (Accessed: 8th January 2017)
2. Clark, M. A., Fisher, C., Judson, I. & Thomas, J. M. Soft-tissue sarcomas in adults. *N. Engl. J. Med.* **353**, 701–711 (2005).
3. Stiller, C. A. & Parkin, D. M. International variations in the incidence of childhood soft-tissue sarcomas. *Paediatr. Perinat. Epidemiol.* **8**, 107–119 (1994).
4. Chibon, F. *et al.* Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat. Med.* **16**, 781–787 (2010).
5. O’Sullivan, B. *et al.* Preoperative versus postoperative radiotherapy in soft-tissue sarcoma of the limbs: a randomised trial. *Lancet Lond. Engl.* **359**, 2235–2241 (2002).
6. ESMO/European Sarcoma Network Working Group. Soft tissue and visceral sarcomas: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **25 Suppl 3**, iii102-112 (2014).
7. Woll, P. J. *et al.* Adjuvant chemotherapy with doxorubicin, ifosfamide, and lenograstim for resected soft-tissue sarcoma (EORTC 62931): a multicentre randomised controlled trial. *Lancet Oncol.* **13**, 1045–1054 (2012).
8. Frustaci, S. *et al.* Adjuvant chemotherapy for adult soft tissue sarcomas of the extremities and girdles: results of the Italian randomized cooperative trial. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **19**, 1238–1247 (2001).
9. van der Graaf, W. T. A. *et al.* Pazopanib for metastatic soft-tissue sarcoma (PALETTE): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Lond. Engl.* **379**, 1879–1886 (2012).
10. Kollár, A. *et al.* Pazopanib in advanced vascular sarcomas: an EORTC Soft Tissue and Bone Sarcoma Group (STBSG) retrospective analysis. *Acta Oncol. Stockh. Swed.* **56**, 88–92 (2017).

11. Casali, P. G. *et al.* Gastrointestinal stromal tumours: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **20 Suppl 4**, 64–67 (2009).
12. Dufresne, A. *et al.* [Molecular biology of sarcoma and therapeutic choices]. *Bull. Cancer (Paris)* **102**, 6–16 (2015).
13. Coggon, D., Ntani, G., Harris, E. C., Jayakody, N. & Palmer, K. T. SOFT TISSUE SARCOMA, NON-HODGKIN LYMPHOMA AND CHRONIC LYMPHOCYTIC LEUKAEMIA IN WORKERS EXPOSED TO PHENOXY HERBICIDES: EXTENDED FOLLOW-UP OF A UK COHORT. *Occup. Environ. Med.* **72**, 435–441 (2015).
14. Burningham, Z., Hashibe, M., Spector, L. & Schiffman, J. D. The Epidemiology of Sarcoma. *Clin. Sarcoma Res.* **2**, 14 (2012).
15. Rubino, C. *et al.* Radiation dose and risk of soft tissue and bone sarcoma after breast cancer treatment. *Breast Cancer Res. Treat.* **89**, 277–288 (2005).
16. Kedes, D. H. *et al.* The seroepidemiology of human herpesvirus 8 (Kaposi's sarcoma-associated herpesvirus): distribution of infection in KS risk groups and evidence for sexual transmission. *Nat. Med.* **2**, 918–924 (1996).
17. Renwick, N. *et al.* Seroconversion for human herpesvirus 8 during HIV infection is highly predictive of Kaposi's sarcoma. *AIDS Lond. Engl.* **12**, 2481–2488 (1998).
18. Malkin, D. Li-Fraumeni Syndrome and p53 in 2015: Celebrating their Silver Anniversary. *Clin. Investig. Med. Med. Clin. Exp.* **39**, E37-47 (2016).
19. Capponcelli, S. *et al.* Evaluation of the molecular mechanisms involved in the gain of function of a Li-Fraumeni TP53 mutation. *Hum. Mutat.* **26**, 94–103 (2005).
20. Pérot, G. *et al.* Constant p53 pathway inactivation in a large series of soft tissue sarcomas with complex genetics. *Am. J. Pathol.* **177**, 2080–2090 (2010).
21. Chibon, F. *et al.* The RB1 gene is the target of chromosome 13 deletions in malignant fibrous histiocytoma. *Cancer Res.* **60**, 6339–6345 (2000).

22. Versteeg, I. *et al.* Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature* **394**, 203–206 (1998).
23. Chibon, F. *et al.* A subgroup of malignant fibrous histiocytomas is associated with genetic changes similar to those of well-differentiated liposarcomas. *Cancer Genet. Cytogenet.* **139**, 24–29 (2002).
24. Cassier, P. A. *et al.* Outcome of patients with platelet-derived growth factor receptor alpha-mutated gastrointestinal stromal tumors in the tyrosine kinase inhibitor era. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **18**, 4458–4464 (2012).
25. Biegel, J. A. *et al.* Germ-line and acquired mutations of INI1 in atypical teratoid and rhabdoid tumors. *Cancer Res.* **59**, 74–79 (1999).
26. Biegel, J. A. *et al.* Alterations of the hSNF5/INI1 gene in central nervous system atypical teratoid/rhabdoid tumors and renal and extrarenal rhabdoid tumors. *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.* **8**, 3461–3467 (2002).
27. Jackson, E. M. *et al.* High-density single nucleotide polymorphism array analysis in patients with germline deletions of 22q11.2 and malignant rhabdoid tumor. *Hum. Genet.* **122**, 117–127 (2007).
28. Savla, J. *et al.* Mutations of the hSNF5/INI1 gene in renal rhabdoid tumors with second primary brain tumors. *J. Natl. Cancer Inst.* **92**, 648–650 (2000).
29. Betz, B. L., Strobeck, M. W., Reisman, D. N., Knudsen, E. S. & Weissman, B. E. Re-expression of hSNF5/INI1/BAF47 in pediatric tumor cells leads to G1 arrest associated with induction of p16ink4a and activation of RB. *Oncogene* **21**, 5193–5203 (2002).
30. Chai, J., Charboneau, A. L., Betz, B. L. & Weissman, B. E. Loss of the hSNF5 gene concomitantly inactivates p21CIP/WAF1 and p16INK4a activity associated with replicative senescence in A204 rhabdoid tumor cells. *Cancer Res.* **65**, 10192–10198 (2005).

31. Dômont, J. *et al.* High frequency of beta-catenin heterozygous mutations in extra-abdominal fibromatosis: a potential molecular tool for disease management. *Br. J. Cancer* **102**, 1032–1036 (2010).
32. Escobar, C., Munker, R., Thomas, J. O., Li, B. D. & Burton, G. V. Update on desmoid tumors. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **23**, 562–569 (2012).
33. Kohler, E. M., Chandra, S. H. V., Behrens, J. & Schneikert, J. Beta-catenin degradation mediated by the CID domain of APC provides a model for the selection of APC mutations in colorectal, desmoid and duodenal tumours. *Hum. Mol. Genet.* **18**, 213–226 (2009).
34. Waller, A., Findeis, S. & Lee, M. J. Familial Adenomatous Polyposis. *J. Pediatr. Genet.* **5**, 78–83 (2016).
35. Chromosomal translocations in Ewing's sarcoma. *N. Engl. J. Med.* **309**, 496–498 (1983).
36. Delattre, O. *et al.* The Ewing family of tumors--a subgroup of small-round-cell tumors defined by specific chimeric transcripts. *N. Engl. J. Med.* **331**, 294–299 (1994).
37. Chibon, F. & Aurias, A. Biologie moléculaire des sarcomes des tissus mous. *Oncologie* **9**, 88–96 (2007).
38. O'Brien, K. P. *et al.* Various regions within the alpha-helical domain of the COL1A1 gene are fused to the second exon of the PDGFB gene in dermatofibrosarcomas and giant-cell fibroblastomas. *Genes. Chromosomes Cancer* **23**, 187–193 (1998).
39. Andrae, J., Gallini, R. & Betsholtz, C. Role of platelet-derived growth factors in physiology and medicine. *Genes Dev.* **22**, 1276–1312 (2008).
40. Heldin, C. H. & Westermark, B. Mechanism of action and in vivo role of platelet-derived growth factor. *Physiol. Rev.* **79**, 1283–1316 (1999).
41. Greco, A. *et al.* Transforming activity of the chimeric sequence formed by the fusion of collagen gene COL1A1 and the platelet derived growth factor b-chain gene in dermatofibrosarcoma protuberans. *Oncogene* **17**, 1313–1319 (1998).

42. Riggi, N., Cironi, L., Suvà, M.-L. & Stamenkovic, I. Sarcomas: genetics, signalling, and cellular origins. Part 1: The fellowship of TET. *J. Pathol.* **213**, 4–20 (2007).
43. Coindre, J.-M. *et al.* Most malignant fibrous histiocytomas developed in the retroperitoneum are dedifferentiated liposarcomas: a review of 25 cases initially diagnosed as malignant fibrous histiocytoma. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **16**, 256–262 (2003).
44. Guillou, L. & Aurias, A. Soft tissue sarcomas with complex genomic profiles. *Virchows Arch.* **456**, 201–217 (2010).
45. Dei Tos, A. P. Classification of pleomorphic sarcomas: where are we now? *Histopathology* **48**, 51–62 (2006).
46. Sandberg, A. A. Updates on the cytogenetics and molecular genetics of bone and soft tissue tumors: leiomyosarcoma. *Cancer Genet. Cytogenet.* **161**, 1–19 (2005).
47. Mandahl, N. *et al.* Comparative cytogenetic study of spindle cell and pleomorphic leiomyosarcomas of soft tissues: a report from the CHAMP Study Group. *Cancer Genet. Cytogenet.* **116**, 66–73 (2000).
48. Yang, J. *et al.* Genetic aberrations in soft tissue leiomyosarcoma. *Cancer Lett.* **275**, 1–8 (2009).
49. Wang, R. *et al.* Loss of 13q14-q21 and gain of 5p14-pter in the progression of leiomyosarcoma. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **16**, 778–785 (2003).
50. Hu, J. *et al.* Loss of DNA copy number of 10q is associated with aggressive behavior of leiomyosarcomas: a comparative genomic hybridization study. *Cancer Genet. Cytogenet.* **161**, 20–27 (2005).
51. Hernando, E. *et al.* The AKT-mTOR pathway plays a critical role in the development of leiomyosarcomas. *Nat. Med.* **13**, 748–753 (2007).
52. Miettinen, M. & Fetsch, J. F. Evaluation of biological potential of smooth muscle tumours. *Histopathology* **48**, 97–105 (2006).
53. Lin, C.-N. *et al.* Prognostic factors of myxofibrosarcomas: implications of margin status, tumor necrosis, and mitotic rate on survival. *J. Surg. Oncol.* **93**, 294–303 (2006).

54. Mentzel, T. *et al.* Myxofibrosarcoma. Clinicopathologic analysis of 75 cases with emphasis on the low-grade variant. *Am. J. Surg. Pathol.* **20**, 391–405 (1996).
55. Mertens, F. *et al.* Cytogenetic analysis of 46 pleomorphic soft tissue sarcomas and correlation with morphologic and clinical features: a report of the CHAMP Study Group. *Chromosomes and Morphology. Genes. Chromosomes Cancer* **22**, 16–25 (1998).
56. Nakayama, R. *et al.* Gene expression analysis of soft tissue sarcomas: characterization and reclassification of malignant fibrous histiocytoma. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **20**, 749–759 (2007).
57. Fletcher, C. D. Pleomorphic malignant fibrous histiocytoma: fact or fiction? A critical reappraisal based on 159 tumors diagnosed as pleomorphic sarcoma. *Am. J. Surg. Pathol.* **16**, 213–228 (1992).
58. Fletcher, C. D. M. The evolving classification of soft tissue tumours: an update based on the new WHO classification. *Histopathology* **48**, 3–12 (2006).
59. Mairal, A. *et al.* Loss of chromosome 13 is the most frequent genomic imbalance in malignant fibrous histiocytomas. A comparative genomic hybridization analysis of a series of 30 cases. *Cancer Genet. Cytogenet.* **111**, 134–138 (1999).
60. Carneiro, A. *et al.* Indistinguishable genomic profiles and shared prognostic markers in undifferentiated pleomorphic sarcoma and leiomyosarcoma: different sides of a single coin? *Lab. Investig. J. Tech. Methods Pathol.* **89**, 668–675 (2009).
61. Parham, D. M. & Ellison, D. A. Rhabdomyosarcomas in adults and children: an update. *Arch. Pathol. Lab. Med.* **130**, 1454–1465 (2006).
62. Furlong, M. A., Mentzel, T. & Fanburg-Smith, J. C. Pleomorphic rhabdomyosarcoma in adults: a clinicopathologic study of 38 cases with emphasis on morphologic variants and recent skeletal muscle-specific markers. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **14**, 595–603 (2001).
63. Li, G. *et al.* Cytogenetic and real-time quantitative reverse-transcriptase polymerase chain reaction analyses in pleomorphic rhabdomyosarcoma. *Cancer Genet. Cytogenet.* **192**, 1–9 (2009).

64. Little, D. J. *et al.* Adult rhabdomyosarcoma: outcome following multimodality treatment. *Cancer* **95**, 377–388 (2002).
65. Gebhard, S. *et al.* Pleomorphic liposarcoma: clinicopathologic, immunohistochemical, and follow-up analysis of 63 cases: a study from the French Federation of Cancer Centers Sarcoma Group. *Am. J. Surg. Pathol.* **26**, 601–616 (2002).
66. Miettinen, M. & Enzinger, F. M. Epithelioid variant of pleomorphic liposarcoma: a study of 12 cases of a distinctive variant of high-grade liposarcoma. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **12**, 722–728 (1999).
67. Hornick, J. L. *et al.* Pleomorphic liposarcoma: clinicopathologic analysis of 57 cases. *Am. J. Surg. Pathol.* **28**, 1257–1267 (2004).
68. Fritz, B. *et al.* Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. *Cancer Res.* **62**, 2993–2998 (2002).
69. Rieker, R. J. *et al.* Distinct chromosomal imbalances in pleomorphic and in high-grade dedifferentiated liposarcomas. *Int. J. Cancer* **99**, 68–73 (2002).
70. Taylor, B. S. *et al.* Functional copy-number alterations in cancer. *PLoS One* **3**, e3179 (2008).
71. Carroll, S. L. & Stonecypher, M. S. Tumor suppressor mutations and growth factor signaling in the pathogenesis of NF1-associated peripheral nerve sheath tumors. I. The role of tumor suppressor mutations. *J. Neuropathol. Exp. Neurol.* **63**, 1115–1123 (2004).
72. Mertens, F. *et al.* Cytogenetic characterization of peripheral nerve sheath tumours: a report of the CHAMP study group. *J. Pathol.* **190**, 31–38 (2000).
73. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
74. Solomou, E. E. *et al.* 13q Deletion in Chronic Lymphocytic Leukemia: Characterization of E4.5, a Novel Chromosome Condensation Regulator-like Guanine Nucleotide Exchange Factor. *Leuk. Lymphoma* **44**, 1579–1585 (2003).

75. Guo, D.-F. *et al.* A novel angiotensin II type 1 receptor-associated protein induces cellular hypertrophy in rat vascular smooth muscle and renal proximal tubular cells. *J. Biol. Chem.* **279**, 21109–21120 (2004).
76. Wu, J.-H. *et al.* Haploinsufficiency of RCBTB1 is associated with Coats disease and familial exudative vitreoretinopathy. *Hum. Mol. Genet.* **25**, 1637–1647 (2016).
77. Zhou, X. & Munger, K. Clld7, A Candidate Tumor Suppressor on Chromosome 13q14, Regulates Pathways of DNA Damage/Repair and Apoptosis. *Cancer Res.* **70**, 9434–9443 (2010).
78. Mabuchi, H. *et al.* Cloning and characterization of CLLD6, CLLD7, and CLLD8, novel candidate genes for leukemogenesis at chromosome 13q14, a region commonly deleted in B-cell chronic lymphocytic leukemia. *Cancer Res.* **61**, 2870–2877 (2001).
79. Legartova, S. *et al.* Expression of RAN, ZHX-2, and CHC1L genes in multiple myeloma patients and in myeloma cell lines treated with HDAC and Dnmts inhibitors. *Neoplasma* **57**, 482–487 (2010).
80. Latil, A. *et al.* CHC1-L, a candidate gene for prostate carcinogenesis at 13q14. 2, is frequently affected by loss of heterozygosity and underexpressed in human prostate cancer. *Int. J. Cancer* **99**, 689–696 (2002).
81. Pontén, F., Jirstrom, K. & Uhlen, M. The Human Protein Atlas—a tool for pathology. *J. Pathol.* **216**, 387–393 (2008).
82. Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **54**, 1.30.1-1.30.33 (2016).
83. Hadjebi, O., Casas-Terradellas, E., Garcia-Gonzalo, F. R. & Rosa, J. L. The RCC1 superfamily: from genes, to function, to disease. *Biochim. Biophys. Acta* **1783**, 1467–1479 (2008).
84. Zhang, L., Villa, N. Y. & McFadden, G. Interplay between poxviruses and the cellular ubiquitin/ubiquitin-like pathways. *FEBS Lett.* **583**, 607–614 (2009).
85. Dhanoa, B. S., Cogliati, T., Satish, A. G., Bruford, E. A. & Friedman, J. S. Update on the Kelch-like (KLHL) gene family. *Hum. Genomics* **7**, 13 (2013).

86. Plafker, K. S., Singer, J. D. & Plafker, S. M. The Ubiquitin Conjugating Enzyme, UbcM2, Engages in Novel Interactions with Components of Cullin-3 Based E3 Ligases[†]. *Biochemistry (Mosc.)* **48**, 3527–3537 (2009).
87. Chaharbakhshi, E. & Jemc, J. C. Broad-complex, tramtrack, and bric-à-brac (BTB) proteins: Critical regulators of development. *Genes. N. Y. N 2000* **54**, 505–518 (2016).
88. Bannasch, D. *et al.* LIFEdb: a database for functional genomics experiments integrating information from external sources, and serving as a sample tracking system. *Nucleic Acids Res.* **32**, D505-508 (2004).
89. Mehrle, A. *et al.* The LIFEdb database in 2006. *Nucleic Acids Res.* **34**, D415-418 (2006).
90. Guo, D.-F. *et al.* A Novel Angiotensin II Type 1 Receptor-associated Protein Induces Cellular Hypertrophy in Rat Vascular Smooth Muscle and Renal Proximal Tubular Cells. *J. Biol. Chem.* **279**, 21109–21120 (2004).
91. Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865–877 (2007).
92. Bamba, C., Bobinnec, Y., Fukuda, M. & Nishida, E. The GTPase Ran regulates chromosome positioning and nuclear envelope assembly in vivo. *Curr. Biol. CB* **12**, 503–507 (2002).
93. Askjaer, P., Galy, V., Hannak, E. & Mattaj, I. W. Ran GTPase cycle and importins alpha and beta are essential for spindle formation and nuclear envelope assembly in living *Caenorhabditis elegans* embryos. *Mol. Biol. Cell* **13**, 4355–4370 (2002).
94. Melchior, F. Ran GTPase cycle: oOne mechanism -- two functions. *Curr. Biol. CB* **11**, R257-260 (2001).
95. Geyer, R., Wee, S., Anderson, S., Yates, J. & Wolf, D. A. BTB/POZ domain proteins are putative substrate adaptors for cullin 3 ubiquitin ligases. *Mol. Cell* **12**, 783–790 (2003).
96. Xu, L. *et al.* BTB proteins are substrate-specific adaptors in an SCF-like modular ubiquitin ligase containing CUL-3. *Nature* **425**, 316–321 (2003).

97. Pintard, L., Willems, A. & Peter, M. Cullin-based ubiquitin ligases: Cul3–BTB complexes join the family. *EMBO J.* **23**, 1681–1687 (2004).
98. Zhao, Y. & Sun, Y. Cullin-RING Ligases as attractive anti-cancer targets. *Curr. Pharm. Des.* **19**, 3215–3225 (2013).
99. Sarikas, A., Hartmann, T. & Pan, Z.-Q. The cullin protein family. *Genome Biol.* **12**, 220 (2011).
100. Lee, J. & Zhou, P. Cullins and cancer. *Genes Cancer* **1**, 690–699 (2010).
101. Haglund, K. & Dikic, I. Ubiquitylation and cell signaling. *EMBO J.* **24**, 3353–3359 (2005).
102. Grabbe, C., Husnjak, K. & Dikic, I. The spatial and temporal organization of ubiquitin networks. *Nat. Rev. Mol. Cell Biol.* **12**, 295 (2011).
103. Sumara, I. *et al.* A Cul3-based E3 ligase removes Aurora B from mitotic chromosomes, regulating mitotic progression and completion of cytokinesis in human cells. *Dev. Cell* **12**, 887–900 (2007).
104. Genau, H. M. *et al.* CUL3-KBTBD6/KBTBD7 ubiquitin ligase cooperates with GABARAP proteins to spatially restrict TIAM1-RAC1 signaling. *Mol. Cell* **57**, 995–1010 (2015).
105. Genau, H. M. & Behrends, C. GABARAP proteins as scaffolds in localized TIAM1-RAC1 signaling. *Mol. Cell. Oncol.* **3**, e1027440 (2016).
106. Hermjakob, H. *et al.* IntAct: an open source molecular interaction database. *Nucleic Acids Res.* **32**, D452–D455 (2004).
107. Ripollés, L. *et al.* Genetic abnormalities and clinical outcome in chronic lymphocytic leukemia. *Cancer Genet. Cytogenet.* **171**, 57–64 (2006).
108. Cotter, F. E. & Auer, R. L. Genetic alteration associated with chronic lymphocytic leukemia. *Cytogenet. Genome Res.* **118**, 310–319 (2007).
109. Wada, M. *et al.* Frequent chromosome arm 13q deletion in aggressive non-Hodgkin's lymphoma. *Leukemia* **13**, 792–798 (1999).
110. Wada, M. *et al.* Delineation of the frequently deleted region on chromosome arm 13q in B-cell non-Hodgkin's lymphoma. *Int. J. Hematol.* **71**, 159–166 (2000).

111. Liu, Y. *et al.* 13q deletions in lymphoid malignancies. *Blood* **86**, 1911–1915 (1995).
112. Kwong, F. M., Wong, P. S. & Lung, M. L. Genetic alterations detected on chromosomes 13 and 14 in Chinese non-small cell lung carcinomas. *Cancer Lett.* **192**, 189–198 (2003).
113. Wistuba, I. I. *et al.* Sequential molecular abnormalities are involved in the multistage development of squamous cell lung carcinoma. *Oncogene* **18**, 643–650 (1999).
114. Latil, A. *et al.* Extensive analysis of the 13q14 region in human prostate tumors: DNA analysis and quantitative expression of genes lying in the interval of deletion: 13q14 and Prostate Cancer. *The Prostate* **57**, 39–50 (2003).
115. Dong, J. T., Boyd, J. C. & Frierson, H. F. Loss of heterozygosity at 13q14 and 13q21 in high grade, high stage prostate cancer. *The Prostate* **49**, 166–171 (2001).
116. Pack, S. D. *et al.* Molecular cytogenetic fingerprinting of esophageal squamous cell carcinoma by comparative genomic hybridization reveals a consistent pattern of chromosomal alterations. *Genes. Chromosomes Cancer* **25**, 160–168 (1999).
117. Hu, N. *et al.* Identification of novel regions of allelic loss from a genomewide scan of esophageal squamous-cell carcinoma in a high-risk Chinese population. *Genes. Chromosomes Cancer* **27**, 217–228 (2000).
118. Gupta, V. K., Schmidt, A. P., Pashia, M. E., Sunwoo, J. B. & Scholnick, S. B. Multiple regions of deletion on chromosome arm 13q in head-and-neck squamous-cell carcinoma. *Int. J. Cancer* **84**, 453–457 (1999).
119. Muller, F. L. *et al.* Passenger deletions generate therapeutic vulnerabilities in cancer. *Nature* **488**, 337–342 (2012).
120. Lee, D. H. *et al.* Overexpression of miR-26a-2 in human liposarcoma is correlated with poor patient survival. *Oncogenesis* **2**, e47 (2013).
121. Báez-Vega, P. M. *et al.* Targeting miR-21-3p inhibits proliferation and invasion of ovarian cancer cells. *Oncotarget* **7**, 36321–36337 (2016).

122. Genschik, P., Sumara, I. & Lechner, E. The emerging family of CULLIN3-RING ubiquitin ligases (CRL3s): cellular functions and disease implications. *EMBO J.* **32**, 2307–2320 (2013).
123. Singh, K. D. & Karnik, S. S. Angiotensin Receptors: Structure, Function, Signaling and Clinical Applications. *J. Cell Signal.* **1**, (2016).
124. Coppieters, F. *et al.* Isolated and Syndromic Retinal Dystrophy Caused by Biallelic Mutations in RCBTB1, a Gene Implicated in Ubiquitination. *Am. J. Hum. Genet.* **99**, 470–480 (2016).
125. Mirza, S., Plafker, K. S., Aston, C. & Plafker, S. M. Expression and distribution of the class III ubiquitin-conjugating enzymes in the retina. *Mol. Vis.* **16**, 2425–2437 (2010).
126. Plafker, K. S., Farjo, K. M., Wiechmann, A. F. & Plafker, S. M. The Human Ubiquitin Conjugating Enzyme, UBE2E3, Is Required for Proliferation of Retinal Pigment Epithelial Cells. *Invest. Ophthalmol. Vis. Sci.* **49**, 5611 (2008).
127. Verellen-Dumoulin, C. *et al.* Expression of an X-linked muscular dystrophy in a female due to translocation involving Xp21 and non-random inactivation of the normal X chromosome. *Hum. Genet.* **67**, 115–119 (1984).
128. Francke, U. *et al.* Minor Xp21 chromosome deletion in a male associated with expression of Duchenne muscular dystrophy, chronic granulomatous disease, retinitis pigmentosa, and McLeod syndrome. *Am. J. Hum. Genet.* **37**, 250–267 (1985).
129. Hoffman, E. P., Brown Jr., R. H. & Kunkel, L. M. Dystrophin: The protein product of the duchenne muscular dystrophy locus. *Cell* **51**, 919–928 (1987).
130. Nishio, H. *et al.* Identification of a novel first exon in the human dystrophin gene and of a new promoter located more than 500 kb upstream of the nearest known promoter. *J. Clin. Invest.* **94**, 1037–1042 (1994).
131. Hammonds, R. G. Protein sequence of DMD gene is related to actin-binding domain of alpha-actinin. *Cell* **51**, 1 (1987).
132. Bonet-Kerrache, A., Fabbriozio, E. & Mornet, D. N-terminal domain of dystrophin. *FEBS Lett.* **355**, 49–53 (1994).

133. Davison, M. D. & Critchley, D. R. alpha-Actinins and the DMD protein contain spectrin-like repeats. *Cell* **52**, 159–160 (1988).
134. Koenig, M. & Kunkel, L. M. Detailed analysis of the repeat domain of dystrophin reveals four potential hinge segments that may confer flexibility. *J. Biol. Chem.* **265**, 4560–4566 (1990).
135. Constantin, B. Dystrophin complex functions as a scaffold for signalling proteins. *Biochim. Biophys. Acta* **1838**, 635–642 (2014).
136. Pillers, D. A. Dystrophin and the retina. *Mol. Genet. Metab.* **68**, 304–309 (1999).
137. Tokarz, S. A. *et al.* Redefinition of dystrophin isoform distribution in mouse tissue by RT-PCR implies role in nonmuscle manifestations of duchenne muscular dystrophy. *Mol. Genet. Metab.* **65**, 272–281 (1998).
138. Lidov, H. G., Selig, S. & Kunkel, L. M. Dp140: a novel 140 kDa CNS transcript from the dystrophin locus. *Hum. Mol. Genet.* **4**, 329–335 (1995).
139. Chamova, T. *et al.* Association Between Loss of Dp140 and Cognitive Impairment in Duchenne and Becker Dystrophies. *Balk. J. Med. Genet. BJMG* **16**, 21–30 (2013).
140. Byers, T. J., Lidov, H. G. & Kunkel, L. M. An alternative dystrophin transcript specific to peripheral nerve. *Nat. Genet.* **4**, 77–81 (1993).
141. Matsumura, K., Yamada, H., Shimizu, T. & Campbell, K. P. Differential expression of dystrophin, utrophin and dystrophin-associated proteins in peripheral nerve. *FEBS Lett.* **334**, 281–285 (1993).
142. Hnia, K. *et al.* Effect of beta-dystroglycan processing on utrophin/Dp116 anchorage in normal and mdx mouse Schwann cell membrane. *Neuroscience* **141**, 607–620 (2006).
143. Lederfein, D. *et al.* A 71-kilodalton protein is a major product of the Duchenne muscular dystrophy gene in brain and other nonmuscle tissues. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 5346–5350 (1992).
144. Howard, P. L., Klamut, H. J. & Ray, P. N. Identification of a novel actin binding site within the Dp71 dystrophin isoform. *FEBS Lett.* **441**, 337–341 (1998).

145. Tozawa, T. *et al.* The shortest isoform of dystrophin (Dp40) interacts with a group of presynaptic proteins to form a presumptive novel complex in the mouse brain. *Mol. Neurobiol.* **45**, 287–297 (2012).
146. Fujimoto, T., Itoh, K., Yaoi, T. & Fushiki, S. Somatodendritic and excitatory postsynaptic distribution of neuron-type dystrophin isoform, Dp40, in hippocampal neurons. *Biochem. Biophys. Res. Commun.* **452**, 79–84 (2014).
147. Gao, Q. Q. & McNally, E. M. The Dystrophin Complex: Structure, Function, and Implications for Therapy. *Compr. Physiol.* **5**, 1223–1239 (2015).
148. Ibraghimov-Beskrovnaya, O. *et al.* Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **355**, 696–702 (1992).
149. Imamura, M. & Ozawa, E. Differential expression of dystrophin isoforms and utrophin during dibutyryl-cAMP-induced morphological differentiation of rat brain astrocytes. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 6139–6144 (1998).
150. Cisneros, B. *et al.* Expression of dystrophin Dp71 during PC12 cell differentiation. *Neurosci. Lett.* **213**, 107–110 (1996).
151. Tadayoni, R., Rendon, A., Soria-Jasso, L. E. & Cisneros, B. Dystrophin Dp71: the smallest but multifunctional product of the Duchenne muscular dystrophy gene. *Mol. Neurobiol.* **45**, 43–60 (2012).
152. de León, M. B. *et al.* Dystrophin Dp71 expression is down-regulated during myogenesis: role of Sp1 and Sp3 on the Dp71 promoter activity. *J. Biol. Chem.* **280**, 5290–5299 (2005).
153. Morales-Lázaro, S. L. *et al.* Induction of dystrophin Dp71 expression during neuronal differentiation: opposite roles of Sp1 and AP2alpha in Dp71 promoter activity. *J. Neurochem.* **112**, 474–485 (2010).
154. Howard, P. L. *et al.* Dystrophin isoforms DP71 and DP427 have distinct roles in myogenic cells. *Muscle Nerve* **22**, 16–27 (1999).

155. Leibovitz, S. *et al.* Exogenous Dp71 is a dominant negative competitor of dystrophin in skeletal muscle. *Neuromuscul. Disord. NMD* **12**, 836–844 (2002).
156. Cerecedo, D. *et al.* Role of dystrophins and utrophins in platelet adhesion process. *Br. J. Haematol.* **134**, 83–91 (2006).
157. Cerecedo, D., Cisneros, B., Suárez-Sánchez, R., Hernández-González, E. & Galván, I. beta-Dystroglycan modulates the interplay between actin and microtubules in human-adhered platelets. *Br. J. Haematol.* **141**, 517–528 (2008).
158. Austin, R. C. *et al.* Identification of Dp71 isoforms in the platelet membrane cytoskeleton. Potential role in thrombin-mediated platelet adhesion. *J. Biol. Chem.* **277**, 47106–47113 (2002).
159. Enríquez-Aragón, J. A. *et al.* Dystrophin Dp71 in PC12 cell adhesion. *Neuroreport* **16**, 235–238 (2005).
160. Cerna, J. *et al.* Dystrophin Dp71f associates with the beta1-integrin adhesion complex to modulate PC12 cell adhesion. *J. Mol. Biol.* **362**, 954–965 (2006).
161. Hance, J. E., Fu, S. Y., Watkins, S. C., Beggs, A. H. & Michalak, M. alpha-actinin-2 is a new component of the dystrophin-glycoprotein complex. *Arch. Biochem. Biophys.* **365**, 216–222 (1999).
162. Hynes, R. O. Targeted mutations in cell adhesion genes: what have we learned from them? *Dev. Biol.* **180**, 402–412 (1996).
163. Jagadha, V. & Becker, L. E. Brain morphology in Duchenne muscular dystrophy: a Golgi study. *Pediatr. Neurol.* **4**, 87–92 (1988).
164. Villarreal-Silva, M., Suárez-Sánchez, R., Rodríguez-Muñoz, R., Mornet, D. & Cisneros, B. Dystrophin Dp71 is critical for stability of the DAPs in the nucleus of PC12 cells. *Neurochem. Res.* **35**, 366–373 (2010).
165. Villarreal-Silva, M., Centeno-Cruz, F., Suárez-Sánchez, R., Garrido, E. & Cisneros, B. Knockdown of dystrophin Dp71 impairs PC12 cells cycle: localization in the spindle and cytokinesis structures implies a role for Dp71 in cell division. *PLoS One* **6**, e23504 (2011).

166. Martin, L. T., Glass, M., Dosunmu, E. & Martin, P. T. Altered expression of natively glycosylated alpha dystroglycan in pediatric solid tumors. *Hum. Pathol.* **38**, 1657–1668 (2007).
167. Higginson, J. R., Thompson, O. & Winder, S. J. Targeting of dystroglycan to the cleavage furrow and midbody in cytokinesis. *Int. J. Biochem. Cell Biol.* **40**, 892–900 (2008).
168. Fuentes-Mera, L. *et al.* Characterization of a novel Dp71 dystrophin-associated protein complex (DAPC) present in the nucleus of HeLa cells: members of the nuclear DAPC associate with the nuclear matrix. *Exp. Cell Res.* **312**, 3023–3035 (2006).
169. González-Ramírez, R., Morales-Lázaro, S. L., Tapia-Ramírez, V., Mornet, D. & Cisneros, B. Nuclear and nuclear envelope localization of dystrophin Dp71 and dystrophin-associated proteins (DAPs) in the C2C12 muscle cells: DAPs nuclear localization is modulated during myogenesis. *J. Cell. Biochem.* **105**, 735–745 (2008).
170. Rodríguez-Muñoz, R. *et al.* Neuronal differentiation modulates the dystrophin Dp71d binding to the nuclear matrix. *Biochem. Biophys. Res. Commun.* **375**, 303–307 (2008).
171. González, E. *et al.* Alternative splicing regulates the nuclear or cytoplasmic localization of dystrophin Dp71. *FEBS Lett.* **482**, 209–214 (2000).
172. Calderilla-Barbosa, L., Ortega, A. & Cisneros, B. Phosphorylation of dystrophin Dp71d by Ca²⁺/calmodulin-dependent protein kinase II modulates the Dp71d nuclear localization in PC12 cells. *J. Neurochem.* **98**, 713–722 (2006).
173. Lara-Chacón, B. *et al.* Characterization of an Importin alpha/beta-recognized nuclear localization signal in beta-dystroglycan. *J. Cell. Biochem.* **110**, 706–717 (2010).
174. Flanigan, K. M. Duchenne and Becker muscular dystrophies. *Neurol. Clin.* **32**, 671–688, viii (2014).
175. Monaco, A. P., Bertelson, C. J., Liechti-Gallati, S., Moser, H. & Kunkel, L. M. An explanation for the phenotypic differences between patients bearing partial deletions of the DMD locus. *Genomics* **2**, 90–95 (1988).

176. Flanigan, K. M. *et al.* DMD Trp3X nonsense mutation associated with a founder effect in North American families with mild Becker muscular dystrophy. *Neuromuscul. Disord. NMD* **19**, 743–748 (2009).
177. Abbadi, N. *et al.* Additional case of female monozygotic twins discordant for the clinical manifestations of Duchenne muscular dystrophy due to opposite X-chromosome inactivation. *Am. J. Med. Genet.* **52**, 198–206 (1994).
178. Cotton, S., Voudouris, N. J. & Greenwood, K. M. Intelligence and Duchenne muscular dystrophy: full-scale, verbal, and performance intelligence quotients. *Dev. Med. Child Neurol.* **43**, 497–501 (2001).
179. Daoud, F. *et al.* Role of mental retardation-associated dystrophin-gene product Dp71 in excitatory synapse organization, synaptic plasticity and behavioral functions. *PLoS One* **4**, e6574 (2008).
180. Sarig, R. *et al.* Targeted inactivation of Dp71, the major non-muscle product of the DMD gene: differential activity of the Dp71 promoter during development. *Hum. Mol. Genet.* **8**, 1–10 (1999).
181. Nico, B. *et al.* Altered blood-brain barrier development in dystrophic MDX mice. *Neuroscience* **125**, 921–935 (2004).
182. Dalloz, C. *et al.* Targeted inactivation of dystrophin gene product Dp71: phenotypic impact in mouse retina. *Hum. Mol. Genet.* **12**, 1543–1554 (2003).
183. Cibis, G. W., Fitzgerald, K. M., Harris, D. J., Rothberg, P. G. & Rupani, M. The effects of dystrophin gene mutations on the ERG in mice and humans. *Invest. Ophthalmol. Vis. Sci.* **34**, 3646–3652 (1993).
184. Sigesmund, D. A. *et al.* Characterization of the ocular phenotype of Duchenne and Becker muscular dystrophy. *Ophthalmology* **101**, 856–865 (1994).

185. Rossbach, H. C., Lacson, A., Grana, N. H. & Barbosa, J. L. Duchenne muscular dystrophy and concomitant metastatic alveolar rhabdomyosarcoma. *J. Pediatr. Hematol. Oncol.* **21**, 528–530 (1999).
186. Chamberlain, J. S., Metzger, J., Reyes, M., Townsend, D. & Faulkner, J. A. Dystrophin-deficient mdx mice display a reduced life span and are susceptible to spontaneous rhabdomyosarcoma. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* **21**, 2195–2204 (2007).
187. Schmidt, W. M. *et al.* DNA damage, somatic aneuploidy, and malignant sarcoma susceptibility in muscular dystrophies. *PLoS Genet.* **7**, e1002042 (2011).
188. Hosur, V. *et al.* Dystrophin and dysferlin double mutant mice: a novel model for rhabdomyosarcoma. *Cancer Genet.* **205**, 232–241 (2012).
189. Wang, Y. *et al.* Dystrophin is a tumor suppressor in human cancers with myogenic programs. *Nat. Genet.* **46**, 601–606 (2014).
190. Körner, H. *et al.* Digital karyotyping reveals frequent inactivation of the dystrophin/DMD gene in malignant melanoma. *Cell Cycle Georget. Tex* **6**, 189–198 (2007).
191. McAvoy, S., Ganapathiraju, S., Perez, D. S., James, C. D. & Smith, D. I. DMD and IL1RAPL1: two large adjacent genes localized within a common fragile site (FRAXC) have reduced expression in cultured brain tumors. *Cytogenet. Genome Res.* **119**, 196–203 (2007).
192. Georgakilas, A. G. *et al.* Are common fragile sites merely structural domains or highly organized ‘functional’ units susceptible to oncogenic stress? *Cell. Mol. Life Sci. CMLS* **71**, 4519–4544 (2014).
193. Luce, L. N., Abbate, M., Cotignola, J. & Giliberto, F. Non-myogenic tumors display altered expression of dystrophin (DMD) and a high frequency of genetic alterations. *Oncotarget* (2016). doi:10.18632/oncotarget.10426
194. Tan, S. *et al.* Decreased Dp71 expression is associated with gastric adenocarcinoma prognosis. *Oncotarget* **7**, 53702–53711 (2016).

195. Tan, S. *et al.* Knocking down Dp71 expression in A549 cells reduces its malignancy in vivo and in vitro. *Cancer Invest.* **34**, 16–25 (2016).
196. Guillou, L. & Aurias, A. Soft tissue sarcomas with complex genomic profiles. *Virchows Arch.* **456**, 201–217 (2010).
197. Pérot, G. *et al.* Smooth muscle differentiation identifies two classes of poorly differentiated pleomorphic sarcomas with distinct outcome. *Mod. Pathol.* **27**, 840–850 (2014).
198. Ratan, R. & Patel, S. R. Chemotherapy for soft tissue sarcoma. *Cancer* **122**, 2952–2960 (2016).
199. van der Graaf, W. T. A. *et al.* Pazopanib for metastatic soft-tissue sarcoma (PALETTE): a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Lond. Engl.* **379**, 1879–1886 (2012).
200. Brodowicz, T. *et al.* Study protocol of REGOSARC trial: activity and safety of regorafenib in advanced soft tissue sarcoma: a multinational, randomized, placebo-controlled, phase II trial. *BMC Cancer* **15**, 127 (2015).
201. Demetri, G. D. *et al.* Efficacy and Safety of Trabectedin or Dacarbazine for Metastatic Liposarcoma or Leiomyosarcoma After Failure of Conventional Chemotherapy: Results of a Phase III Randomized Multicenter Clinical Trial. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **34**, 786–793 (2016).
202. Chibon, F. *et al.* Validated prediction of clinical outcome in sarcomas and multiple types of cancer on the basis of a gene expression signature related to genome complexity. *Nat. Med.* **16**, 781–787 (2010).
203. Stable Instability of Sarcoma Cell Lines Genome Despite Intra-Tumoral Heterogeneity: A Genomic and Transcriptomic Study of Sarcoma Cell Lines. Available at:
<http://austinpublishinggroup.com/genetics-genomic-research/fulltext/ajggr-v2-id1014.php>.
(Accessed: 11th October 2016)
204. Gibault, L. *et al.* New insights in sarcoma oncogenesis: a comprehensive analysis of a large series of 160 soft tissue sarcomas with complex genomics. *J. Pathol.* **223**, 64–71 (2011).

205. Lesluyes, T. *et al.* RNA sequencing validation of the Complexity INdex in SARComas prognostic signature. *Eur. J. Cancer Oxf. Engl.* **1990** **57**, 104–111 (2016).
206. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25**, 2078–2079 (2009).
207. Li, H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinforma. Oxf. Engl.* **27**, 2987–2993 (2011).
208. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).
209. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
210. 1000 Genomes Project Consortium *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
211. Landrum, M. J. *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res.* **44**, D862-868 (2016).
212. Forbes, S. A. *et al.* COSMIC: exploring the world’s knowledge of somatic mutations in human cancer. *Nucleic Acids Res.* **43**, D805-811 (2015).
213. Henriksen, M., Miller, B., Newmark, J., Al-Kofahi, Y. & Holden, E. Laser scanning cytometry and its applications: a pioneering technology in the field of quantitative imaging cytometry. *Methods Cell Biol.* **102**, 161–205 (2011).
214. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
215. Trojani, M. *et al.* Soft-tissue sarcomas of adults; study of pathological prognostic variables and definition of a histopathological grading system. *Int. J. Cancer* **33**, 37–42 (1984).

216. Campbell, A. E., Hsiung, C. C.-S. & Blobel, G. A. Comparative analysis of mitosis-specific antibodies for bulk purification of mitotic populations by fluorescence-activated cell sorting. *BioTechniques* **56**, 90–91, 93–94 (2014).
217. Zhou, X. & Munger, K. Clld7, A Candidate Tumor Suppressor on Chromosome 13q14, Regulates Pathways of DNA Damage/Repair and Apoptosis. *Cancer Res.* **70**, 9434–9443 (2010).
218. Dommering, C. J. *et al.* RB1 mutation spectrum in a comprehensive nationwide cohort of retinoblastoma patients. *J. Med. Genet.* **51**, 366–374 (2014).
219. Davoli, T. *et al.* Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948–962 (2013).
220. Lee, D. H. *et al.* Overexpression of miR-26a-2 in human liposarcoma is correlated with poor patient survival. *Oncogenesis* **2**, e47 (2013).
221. Bhardwaj, A. *et al.* CXCL12/CXCR4 signaling counteracts docetaxel-induced microtubule stabilization via p21-activated kinase 4-dependent activation of LIM domain kinase 1. *Oncotarget* **5**, 11490–11500 (2014).
222. Pautier, P. *et al.* Randomized multicenter and stratified phase II study of gemcitabine alone versus gemcitabine and docetaxel in patients with metastatic or relapsed leiomyosarcomas: a Federation Nationale des Centres de Lutte Contre le Cancer (FNCLCC) French Sarcoma Group Study (TAXOGEM study). *The Oncologist* **17**, 1213–1220 (2012).
223. Hadjebi, O., Casas-Terradellas, E., Garcia-Gonzalo, F. R. & Rosa, J. L. The RCC1 superfamily: from genes, to function, to disease. *Biochim. Biophys. Acta* **1783**, 1467–1479 (2008).
224. Bos, J. L., Rehmann, H. & Wittinghofer, A. GEFs and GAPs: critical elements in the control of small G proteins. *Cell* **129**, 865–877 (2007).
225. Bamba, C., Bobinnec, Y., Fukuda, M. & Nishida, E. The GTPase Ran regulates chromosome positioning and nuclear envelope assembly in vivo. *Curr. Biol. CB* **12**, 503–507 (2002).

226. Askjaer, P., Galy, V., Hannak, E. & Mattaj, I. W. Ran GTPase cycle and importins alpha and beta are essential for spindle formation and nuclear envelope assembly in living *Caenorhabditis elegans* embryos. *Mol. Biol. Cell* **13**, 4355–4370 (2002).
227. Zhao, Y. & Sun, Y. Cullin-RING Ligases as Attractive Anti-cancer Targets. *Curr. Pharm. Des.* **19**, 3215–3225 (2013).
228. Lee, J. & Zhou, P. Cullins and Cancer. *Genes Cancer* **1**, 690–699 (2010).
229. Zhao, Y. & Sun, Y. Cullin-RING Ligases as Attractive Anti-cancer Targets. *Curr. Pharm. Des.* **19**, 3215–3225 (2013).
230. Plafker, K. S., Singer, J. D. & Plafker, S. M. The Ubiquitin Conjugating Enzyme, UbcM2, Engages in Novel Interactions with Components of Cullin-3 Based E3 Ligases[†]. *Biochemistry (Mosc.)* **48**, 3527–3537 (2009).
231. Li, Z. & Zhang, X. Autophagy in mitotic animal cells. *Sci. Bull.* **61**, 105–107 (2016).
232. Genau, H. M. *et al.* CUL3-KBTBD6/KBTBD7 ubiquitin ligase cooperates with GABARAP proteins to spatially restrict TIAM1-RAC1 signaling. *Mol. Cell* **57**, 995–1010 (2015).
233. Gao, K. *et al.* Tumor suppressor SPOP mediates the proteasomal degradation of progesterone receptors (PRs) in breast cancer cells. *Am. J. Cancer Res.* **5**, 3210–3220 (2015).
234. Zhang, C. *et al.* Cullin3-KLHL25 ubiquitin ligase targets ACLY for degradation to inhibit lipid synthesis and tumor progression. *Genes Dev.* **30**, 1956–1970 (2016).
235. Chen, H.-Y. & Chen, R.-H. Cullin 3 Ubiquitin Ligases in Cancer Biology: Functions and Therapeutic Implications. *Front. Oncol.* **6**, 113 (2016).
236. Wang, X.-J. *et al.* Nrf2 enhances resistance of cancer cells to chemotherapeutic drugs, the dark side of Nrf2. *Carcinogenesis* **29**, 1235–1243 (2008).
237. Chen, C. D. *et al.* Molecular determinants of resistance to antiandrogen therapy. *Nat. Med.* **10**, 33–39 (2004).

238. Salah, F. S. *et al.* Tumor suppression in mice lacking GABARAP, an Atg8/LC3 family member implicated in autophagy, is associated with alterations in cytokine secretion and cell death. *Cell Death Dis.* **7**, e2205 (2016).
239. Keulers, T. G. *et al.* GABARAPL1 is required for increased EGFR membrane expression during hypoxia. *Radiother. Oncol. J. Eur. Soc. Ther. Radiol. Oncol.* **116**, 417–422 (2015).
240. Wells, A. *et al.* Ligand-induced transformation by a noninternalizing epidermal growth factor receptor. *Science* **247**, 962–964 (1990).
241. Schaaf, M. B. E., Keulers, T. G., Vooijs, M. A. & Rouschop, K. M. A. LC3/GABARAP family proteins: autophagy-(un)related functions. *FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol.* (2016). doi:10.1096/fj.201600698R
242. Hirsch, B. Sister chromatid exchanges are preferentially induced at expressed and nonexpressed common fragile sites. *Hum. Genet.* **87**, 302–306 (1991).
243. Gorgoulis, V. G. *et al.* Activation of the DNA damage checkpoint and genomic instability in human precancerous lesions. *Nature* **434**, 907–913 (2005).
244. Helmrich, A., Ballarino, M. & Tora, L. Collisions between replication and transcription complexes cause common fragile site instability at the longest human genes. *Mol. Cell* **44**, 966–977 (2011).
245. Haldar, M., Hancock, J. D., Coffin, C. M., Lessnick, S. L. & Capecchi, M. R. A Conditional Mouse Model of Synovial Sarcoma: Insights into a Myogenic Origin. *Cancer Cell* **11**, 375–388 (2007).
246. Rangarajan, A., Hong, S. J., Gifford, A. & Weinberg, R. A. Species- and cell type-specific requirements for cellular transformation. *Cancer Cell* **6**, 171–183 (2004).
247. Górecki, D. C. Dystrophin: The dead calm of a dogma. *Rare Dis. Austin Tex* **4**, e1153777 (2016).
248. Beck, A. J. *et al.* Differential requirement for utrophin in the induced pluripotent stem cell correction of muscle versus fat in muscular dystrophy mice. *PLoS One* **6**, e20065 (2011).

Annexes

I- Matériel et Méthode du projet RCBTB1

Culture Cellulaire

La lignée de fibroblastes utilisée lors de ce projet provient de la lignée IMR90 qui a été transformée avec les oncoprotéines : E6 (dégrade p53), E7 (inactive pRB), SmallT (altère la voie mTOR), Ras mutée (active les MAP kinases) et Tert (maintient la longueur des télomères). Cette lignée est cultivée dans un milieu DMEM Glutamax™ (GIBCO®).

Création de modèles cellulaires sous-exprimant le gène *RCBTB1*

Afin d'étudier l'inhibition de l'expression du gène *RCBTB1* dans une lignée de fibroblastes (RST), 5 séquences codant des shARN dirigés contre *RCBTB1* ont été clonées dans le plasmide TET-pLKO-puro inducible selon les recommandations du fournisseur (Addgene) (Figure 45). Leur séquence a été vérifiée par séquençage. Une séquence codant un shARN NT ne ciblant aucun gène humain a également été clonée afin de servir de contrôle négatif.

de 0.22 μ m. Les cellules à infecter sont alors incubées toute la nuit dans un milieu DMEM + 10% SVF sans résidus de tétracycline + 1‰ polybrène + 1mL de surnageant contenant les particules lentivirales. Le lendemain, le milieu contenant les particules virales est remplacé par du milieu DMEM + 10% SVF sans résidus de tétracycline + 1% PS.

Visualisation des cassures de l'ADN par marquage de l'histone γ H2AX

Le marquage a été réalisé après 10 et 15 passages cellulaires après induction à la doxycycline (ajout de doxycycline au milieu toutes les 48h). Les cellules sontensemencées à 250 000 cellules par puits en duplicats. Le lendemain, de l'étoposide (40 μ M), un inhibiteur de la topoisomérase II humaine induisant des cassures doubles brins sur l'ADN, est ajouté dans les puits de contrôle positif 2h avant le début de l'expérience. Suite à deux lavages en PBS1X, les cellules reprises dans du PBS1X sont fixées à l'aide d'éthanol absolu froid ajouté à raison de 3 volumes au goutte-à-goutte sur un vortex. Les cellules sont ensuite centrifugées à 1300 rpm pendant 5 min avant d'être reprises dans 200 μ l d'une solution contenant du PBS1x, du Tween 20 0.05%, de la BSA (*bovine albumin serum*) 5% et l'anticorps primaire anti- γ H2AX dilué au 400^{ème}. Les cellules sont alors incubées 2h à 4°C. Ensuite elles sont lavées deux fois à l'aide de PBS1X avant d'être reprises dans la même solution contenant cette fois-ci l'anticorps secondaire : anti-mouse FITC dilué au 100^{ème}. Les cellules sont incubées 1h à 4°C à l'abri de la lumière. Enfin, elles sont lavées deux fois à l'aide de PBS 1X et reprises dans 500 μ l de PBS 1X puis conservées dans la glace à l'abri de la lumière avant d'être analysées au FACS (BD FACSCalibur flow cytometer, BD Biosciences).

Immunofluorescence: visualisation de foyers γ H2AX

En parallèle de l'analyse au FACS, des lames Lab-Tek (Nunc) de 8 puits sontensemencées avec 10 000 cellules par puits après 15 passages sous doxycycline. Comme précédemment, les puits contrôles positifs sont traités avec de l'étoposide (10 μ M) 2h avant l'expérience. Entre chaque étape trois lavages en PBS1X sont réalisés. Ainsi les cellules sont tout d'abord fixées à l'aide de paraformaldéhyde 3,7% pendant 20 minutes avant d'être perméabilisées avec du triton 0,5% pendant 10 min à température ambiante. Ensuite, les cellules sont incubées en présence d'anticorps

primaire anti- γ H2AX dilué au 200^e dans du PBS1X + 10% de SVF pendant 1h à température ambiante. Elles sont ensuite incubées en présence d'anticorps secondaire Anti-mouse FITC dilué au 400^e dans une solution de PBS 1X – 10% SVF pendant 45 min à température ambiante. Enfin, du milieu de montage (Dapi Vectashield) et une lamelle sont ajoutés sur les cellules. Les lames sont alors observées au microscope à fluorescence Nikon Eclipse 80i. Les photos sont prises avec le logiciel Genikon.

Hybridation génomique comparative sur puce à ADN (CGH-array)

Les profils génomiques des lignées RST exprimant les shARN et cultivées après 17 passages sous doxycycline ont été analysés par CGH-array sur des lames 8 array de 60 000 sondes (Agilent) selon les recommandations du fournisseur. Ces expériences ont été réalisées par Pauline LAGARDE.

In Vivo

Les souris NSG ont reçu une injection sur le flanc droit de $2.5 \cdot 10^6$ cellules IB136-Vide (Contrôle) ou IB136-RCBTB1. Lorsque les tumeurs ont atteint un volume de 150 mm³, les souris sont traitées par 2 injections espacées de 4 jours de 100 μ l en intra-péritonéal de :

- DMSO : contrôle
- Docetaxel à 1mg/mL
- Docetaxel 2mg/mL

Trente jours après 1^{ère} injection, les souris sont sacrifiées, et les tumeurs sont récupérées et pesées.

II- Matériel et Méthode du projet *DMD*

Produit chimique

La doxycycline (D9891, Sigma-Aldrich) est reçue sous forme de poudre lyophilisée et est reconstituée avec de l'eau stérile à une concentration de 10mg/mL. La doxycycline est utilisée à une concentration finale de 1mg/mL pour induire l'expression des shARN.

Lignées cellulaires

Les lignées cellulaires ont été établies comme décrit précédemment (6). Leur authentification est réalisée par aCGH et comparaison avec la tumeur d'origine. Les lignées cellulaires SW982 (synoviosarcome), IB112, IB133 et IB136 (léiomyosarcomes) sont cultivées dans un milieu RPMI 1640 + GlutaMAX I (Life Technologies, Inc.) complété par 10% sérum de veau fœtal (SVF) et 1% pénicilline-streptomycine (Life Technologies). La lignée HEK-293T est cultivée en DMEM + GlutaMAX I (Life Technologies) complété par 10% de SVF et 1% pénicilline-streptomycine (Life Technologies). Les cellules sont cultivées à 37°C en atmosphère humide et à 5% CO₂.

Les lignées SW982, IB112 et IB136 ont été infectées avec un vecteur lentiviral (Tet-pLKO-Puro, Figure 45) contenant les séquences des shARN sous le contrôle d'un promoteur inductible par la Tétracycline (*Tet-on*). La lignée contrôle a été établie par transduction lentivirale d'un vecteur contenant la séquence d'un shARN ne ciblant aucun ARN humain (shNT, « *Non-targeting* »). La lignée IB133 a subi une infection lentivirale afin d'exprimer le système CRISPR/Cas9 comprenant l'ARN-guide ciblant spécifiquement *Dp427* (CRISPR/CAS gRNA-DMD, Hs0000003709, Sigma Aldrich). La lignée contrôle a été établie par infection lentivirale d'un vecteur codant pour un ARN-guide ne ciblant aucune séquence du génome humain (CRISPR/CAS gRNA-NEGATIVE CONTROL 1, CRISPR06, Sigma Aldrich). Les particules virales pseudotypées avec la protéine G du virus de la stomatite vésiculaire (VSV-G) ont été produites par co-transfection de cellules HEK-293T avec les vecteurs présentés précédemment et les plasmides d'encapsulation psPAX2 et pVSVg. Les lignées cellulaires sont incubées sur la nuit avec les surnageants contenant les particules lentivirales en présence de 8 µg/ml de polybrene (Sigma H9268). Les cellules exprimant stablement les vecteurs sont sélectionnées à l'aide de puromycine (2µg/mL, Sigma P9620) dans le milieu de culture. L'efficacité de l'inhibition de l'expression des dystrophines induite par la doxycycline dans le cas des lignées exprimant les shARN

d'une part, et l'efficacité du CRISPr exprimé par IB133 d'autre part, sont ensuite vérifiées par western blot.

Séquencage ADN de type Sanger

La PCR est réalisée dans un volume final de 20 µl contenant 2 µl de Tampon (10X), 1.5 µl de MgCl₂ (25mM), 1.6µl de dNTP (2.5 mM chacun), 0.1 µl de Taq Gold (0.5 unité), 1µl d'amorces (10µM pour chaque amorce : sens et antisens), 12.8 µl d'eau et 1µl d'ADNg (50ng/µl). Le produit de PCR est contrôlé sur un gel d'agarose à 2 % en tampon TAE (Tris, Acétate, EDTA) complété par 0.035 % de Bromure d'éthidium. Les produits de PCR sont ensuite purifiés à l'aide du kit ExoSAP-it PCR cleanup Kit™ (GE Healthcare) selon les recommandations du fournisseur. La réaction de séquence est réalisée à l'aide du kit Big Dye V1.1 puis elle est purifiée avec le kit Big Dye XTerminator Purification (Applied) selon les recommandations du fournisseur. Enfin les échantillons sont séquencés à l'aide du séquenceur 3130xl Genetic Analyzer (Applied) et analysés à l'aide du logiciel SeqScape v2.5 (Applied).

Western blot

Pour l'extraction des protéines, les cellules sont lavées deux fois avec du PBS 1X puis lysées 30min à 4°C avec un tampon d'extraction RIPA (R0278, Sigma) complété avec un cocktail d'inhibiteurs de phosphatases et protéases (11697498001, Roche). Les lysats sont ensuite récoltés puis centrifugés pendant 10min à 13500rpm à 4°C. Le surnageant est récupéré puis le dosage des protéines est réalisé selon les recommandations du fournisseur (DC protein assay kit, Biorad). Pour chaque échantillon, 40µg de protéines sont déposés dans des gels de SDS-PAGE. Après électro-transfert à sec des protéines sur une membrane de PVDF, la membrane est saturée avec une solution de PBS-Tween 01% + lait 5% puis incubée sur la nuit à 4°C avec l'anticorps primaire : l'anticorps 7A10 dilué au 1/200e (sc-47760, Santa Cruz Biotechnology) pour détecter Dp71, l'anticorps Dys1 dilué au 1/50e (NCL-DYS1, Leica Biosystems) pour détecter Dp427 et l'anticorps spécifique de la GAPDH (1/2000e, sc-166574, Santa Cruz Biotechnology) qui sert de contrôle de charge. Après lavages, les membranes sont hybridées avec l'anticorps secondaire couplé à la

peroxydase (Amersham) pendant 1h à température ambiante puis les protéines d'intérêt sont révélées par chimioluminescence à l'aide d'un kit de réactifs ECL (Amersham ECL Select detection reagent kit, Sigma) selon les recommandations du fournisseur. Le signal est détecté à l'aide du système Fusion Fx7 (Thermo Fisher Scientific).

Etude de l'expression par la technologie Taqman

Les ADNc sont synthétisés à partir de 1µg d'ARN à l'aide du kit « GeneAmp RNA PCR core Kit » (Applied Biosystem, Courtaboeuf, France). L'analyse par PCR quantitative a été réalisée avec les réactifs « TaqMan Assays-on-demand Gene expression reagents » (Applied Biosystem), « qPCR Mastermix Plus without UNG » (Eurogentec, Belgium) et système de PCR en temps réel fourni par Applied Biosystems. Les sondes utilisées sont les suivantes : HsXXXX_m1 pour DMD et Hs99999902_m1 pour RPLP0, qui sert de gène de référence pour la normalisation des résultats.

Test de prolifération cellulaire

Les cellules sontensemencées en plaques 96 puits à 1000 cellules/puits. Le milieu de culture est renouvelé aux jours 4, 7 et 9. Tous les deux ou trois jours, les cellules sont lavées, décollées à l'aide d'une solution de trypsine puis re-suspendues dans un volume final de 200µL avec une solution de PBS 1X. Le nombre de cellules est évalué par cytométrie en flux (FACS Calibur, BD Biosciences) à partir de leurs caractéristiques morphologiques (dispersion en FSC/SSC). L'acquisition des données est réalisée grâce au logiciel CellQuest Pro (BD FACS Systems, Sunnyvale, CA) puis l'analyse est faite avec les logiciels FlowJo (Tree Star, Celeza GmbH) et GraphPad (La Jolla, CA).

Test de clonogénicité

La clonogénicité est évaluée par la capacité des cellules à former des clones (au moins 50 cellules) à partir d'une seule cellule dans des conditions d'adhérence. Dans ce but, 2000 cellules sontensemencées en triplicats dans des plaques de 6 puits et

cultivées 12 jours. Les colonies formées sont ensuite fixées avec une solution d'éthanol à 70% puis colorées 5min avec une solution de Crystal Violet 2.3% (Sigma-Aldrich).

Analyse du cycle cellulaire

Pour l'analyse du cycle cellulaire, 75 000 cellules sont ensemencées en triplicats dans des plaques 6 puits. Lorsque les cellules atteignent 75% de confluence, elles sont décollées à l'aide d'une solution de trypsine puis lavées deux fois avec une solution de PBS 1X. Après la dernière centrifugation, le culot de cellules est repris dans 500µL de PBS 1X puis 1,5mL d'éthanol absolu est ajouté au goutte à goutte au vortex. Les cellules sont ensuite incubées à 4°C sur la nuit. Pour le marquage avec l'iodure de propidium (PI), la suspension cellulaire subit deux lavages avec une solution de PBS 1X puis le culot est repris avec la solution de marquage FxCycle™ PI/RNase Staining Solution (F10797, Thermofisher Scientific) selon les recommandations du fournisseur. Après 15min d'incubation, au moins 10 000 cellules par condition sont analysées par cytométrie en flux (FACS Calibur, BD Biosciences). L'acquisition des données est réalisée grâce au logiciel CellQuest Pro (BD FACS Systems, Sunnyvale, CA) puis l'analyse des résultats est faite avec le logiciel FlowJo (Tree Star, Celeza GmbH), qui permet l'identification des phases du cycle à l'aide du modèle de Watson et le logiciel GraphPad (La Jolla, CA).

Analyse statistique

La survie sans métastase (MFS, Metastasis-Free Survival) est définie comme l'intervalle de temps entre le diagnostic et le moment de la rechute métastatique ou le dernier suivi. Les taux de survie sont estimés à l'aide d'une analyse Kaplan-Meier et comparés avec un test log-rank. Les statistiques descriptives sont utilisées pour montrer la distribution des variables dans la population. Descriptive statistics were used to show the distribution of variables in the population. Une différence est considérée comme significative lorsque $P < 0.05$. Toutes analyses statistiques ont été réalisées avec le logiciel R (R version 2.14.1).