



Profils d'expression des microARN dans les sarcomes : des données brutes aux applications cliniques

Daniel Pissaloux

► To cite this version:

Daniel Pissaloux. Profils d'expression des microARN dans les sarcomes : des données brutes aux applications cliniques. Médecine humaine et pathologie. Université Claude Bernard - Lyon I, 2012. Français. NNT : 2012LYO10323 . tel-00997015

HAL Id: tel-00997015

<https://theses.hal.science/tel-00997015>

Submitted on 27 May 2014

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 : 323 – 2012

Année 2012

THÈSE DE L'UNIVERSITÉ DE LYON

Délivrée par

L'UNIVERSITÉ CLAUDE BERNARD LYON 1

ECOLE DOCTORALE
BIOLOGIE MOLÉCULAIRE INTÉGRATIVE ET CELLULAIRE

DIPLÔME DE DOCTORAT

(arrêté du 7 août 2006)

Discipline : cancérologie

soutenue publiquement le 18 décembre 2012 par

M. Daniel PISSALOUX

TITRE :

**Profils d'expression des microARN dans les sarcomes :
des données brutes aux applications cliniques**

Directeur de Thèse : **Pr Jean-Yves BLAY**
Codirecteur de Thèse : **Dr Laurent ALBERTI**

JURY :	M. Bernard ASSELAIN	Rapporteur
	M. Jean-Michel COINDRE	Rapporteur
	M. Michel RIVOIRE	Président
	M. Laurent ALBERTI	Examineur
	M. Jean-Yves BLAY	Examineur
	M. Sébastien SALAS	Examineur

REMERCIEMENTS

A l'issue de ces années de thèse, voici venu le moment de remercier toutes les personnes qui m'ont entouré et aidé à mener à bien ce projet. La liste est longue, mais je voudrais m'attacher à n'oublier personne.

Mes premiers remerciements vont tout naturellement à mon directeur de thèse, le Professeur Jean-Yves Blay, pour m'avoir tout d'abord encouragé à entreprendre ce travail de thèse, mais également pour son écoute et ses conseils avisés pendant toutes ces années. Je ne le remercierai jamais assez pour le dynamisme et la motivation qu'il a su me transmettre.

Je remercie vivement mon codirecteur de thèse, le Docteur Laurent Alberti, pour m'avoir accueilli au sein de son laboratoire, avant même que je ne commence ce travail de thèse. Merci d'avoir cru en moi et de m'avoir donné l'opportunité de travailler dans un grand laboratoire de recherche. Merci également de m'avoir laissé une grande liberté d'organisation, et d'avoir excusé mes nombreuses absences lorsque mon travail de routine prenait le dessus.

Je tiens également à adresser mes remerciements aux membres de mon jury de thèse : le Docteur Bernard Asselain et le Professeur Jean-Michel Coindre, qui m'ont fait l'honneur d'accepter de juger ce travail en tant que rapporteurs, le Docteur Sébastien Salas, qui a accepté d'être examinateur de cette thèse, ainsi que le Professeur Michel Rivoire, qui a accepté de présider ce jury.

Je remercie le Professeur Alain Puisieux de m'avoir accueilli au sein de l'unité mixte de recherche INSERM 1052 - CNRS 5286, et le Docteur Christophe Caux, au sein de l'équipe Ciblage de la Tumeur et de son Environnement Immunitaire.

Ce travail de thèse ayant été réalisé en tant que salarié du Centre Léon Bérard, je tiens à remercier le Professeur Sylvie Négrier, Directeur Général du Centre Léon Bérard, et Monsieur Hervé Valérian-Bessac, Directeur des Ressources Humaines, qui m'ont autorisé à effectuer ce travail de thèse en parallèle à mes obligations d'Ingénieur Biologiste.

Je souhaite également remercier les Docteurs Dominique Ranchère-Vince et Anne-Valérie Decouvlaere, pour les responsabilités qui m'ont été confiées au sein du département de Biopathologie, et pour m'avoir soutenu tout au long de ma thèse.

Merci beaucoup à Jean-Phillipe Michot, mon partenaire de tous les jours, et notamment de mes longues soirées au Centre, avec qui je prends toujours beaucoup de plaisir à refaire le monde. J'espère avoir pu t'aider à remplir tes objectifs !

Je remercie tout particulièrement l'ensemble des membres des équipes sarcomes avec qui j'ai eu l'occasion de travailler : merci à Martine, Adeline, Anthony, Héloïse, Angélique, Hiba, Séverine, Thomas et Jennifer pour leur aide précieuse, et tous les bons moments que j'ai pu passer en leur compagnie.

Merci à toutes les personnes du Centre et d'ailleurs, qui ont croisé, ou qui croisent régulièrement ma route, et avec qui j'apprécie beaucoup de travailler : merci à Frédéric et Pauline, à Florence, Frédérique et Rita, et à Benoît pour leurs innombrables conseils en cytogénétique moléculaire, merci à Aurélie pour sa gentillesse, sa bonne humeur et ses cafés-cookies, merci à Françoise, la reine du multitâche, merci à Didier, le maître des vecteurs, et merci à Erik pour son accueil chaleureux à Rotterdam.

Merci également à tous les collègues avec lesquels je travaille quotidiennement dans la joie et la bonne humeur ! Marlène, Géraldine (x2), Sandrine, Aurélie, Elodie, Amandine, Cyrille, Arnaud, Christine, Isabelle, Marie-Eve, Anne-Isabelle, Kalsoum, Fatiha, Dalhia, Brigitte, Claire, Séverine, Thierry et toute la bande de la Biopathologie...

Un grand merci aux différents organismes (Conticanet, EUROSARC, LYRIC et les Amis de l'Université de Lyon) et à leurs administrateurs (Marta, Simon et Marina) pour leurs soutiens financiers m'ayant permis de réaliser ces travaux.

Je tiens à remercier ma famille et ma belle-famille pour leur soutien. Un grand merci à mon père, François, qui m'a toujours encouragé dans cette voie, ainsi que pour ses conseils judicieux. Un grand merci également à ma belle-mère, Béatrice, qui m'a permis de finir ce travail de thèse dans les meilleures conditions.

Merci à mes deux grands amours ! Sacha, mon fils, merci pour ta joie de vivre, indestructible et contagieuse, et merci d'avoir accepté mes nombreuses absences sans m'en tenir rigueur. Enfin, merci à la personne qui m'a toujours soutenu et avec qui j'ai déjà parcouru un bon bout de chemin dans la vie, Vanessa. Merci infiniment pour tout ton amour, ta compréhension et ta confiance.

SOMMAIRE

RÉSUMÉ	1
LISTE DES PRINCIPALES ABRÉVIATIONS.....	2
PRÉAMBULE	5
INTRODUCTION BIBLIOGRAPHIQUE	9
I. PRESENTATION DU RHABDOMYOSARCOME ET DE L'OSTEOSARCOME .	10
1. Les rhabdomyosarcomes	10
1.1. Description et épidémiologie	10
1.2. Diagnostic et classification.....	11
- Présentation clinique et diagnostique	11
- Classification.....	12
1.3. Caractéristiques biologiques des ERMS	15
- Anatomopathologie	15
- Caractéristiques génétiques et moléculaires	16
1.4. Caractéristiques biologiques des ARMS.....	17
- Anatomopathologie	17
- Caractéristiques génétiques et moléculaires	18
1.5. Événements génétiques secondaires présents dans les RMS	21
- Événements favorisant la survie cellulaire.....	22
- Événements favorisant la prolifération cellulaire	22
- Autres événements contribuant à la tumorigenèse des RMS.....	23
1.6. Le diagnostic différentiel ARMS <i>versus</i> ERMS	24
1.7. Facteurs pronostiques.....	27
1.8. Traitement des RMS	28
2. Les ostéosarcomes	30
2.1. Description et épidémiologie	30
2.2. Diagnostic et classification.....	31
- Présentation clinique	31
- Anatomopathologie	32
- Caractéristiques génétiques et moléculaires	34
2.3. Traitement	35
- Traitement conventionnel.....	35
- Chimiothérapie.....	36
- Nouvelles stratégies thérapeutiques	38
2.4. Risques et effets secondaires du traitement à l'ifosfamide	39
2.5. Evolution et facteurs pronostiques.....	40
- Récidives et métastases	40
- Facteurs pronostiques cliniques.....	40
- Facteurs pronostiques moléculaires	41

II.	LES MICROARN	43
1.	Définition et mode d'action	43
1.1.	Définition et découverte des microARN	43
1.2.	Biogenèse des microARN	46
1.3.	Principe de l'interaction entre les microARN et les ARN messagers.....	49
1.4.	Mécanismes post-transcriptionnels régulés par les microARN	51
	- Régulation de la traduction	52
	- Régulation de la stabilité des ARN messagers	53
2.	MicroARN et cancer	53
2.1.	Dérégulation de l'expression des microARN dans les cancers	53
2.2.	Potentiel clinique des profils d'expression des microARN	57
	- Diagnostic et classification.....	57
	- Indicateurs pronostiques	58
	- MicroARN circulants : biomarqueurs non invasifs	58
	- Indicateurs prédictifs d'une réponse au traitement.....	59
2.3.	Utilisation des microARN en tant que stratégie thérapeutique.....	60
III.	OUTILS APPLIQUES A L'ETABLISSEMENT DES PROFILS D'EXPRESSION DES MICROARN	64
1.	Technologies pour l'étude des profils d'expression des microARN	64
1.1.	Puces à oligonucléotides	64
	- Définition et historique	64
	- Principe de fonctionnement des puces microARN	66
1.2.	PCR quantitative à moyen débit	69
	- Définition et historique de la PCR quantitative.....	69
	- Principe de la PCR quantitative appliquée aux microARN.....	71
2.	Méthodes de normalisation des données d'expression	73
2.1.	Le besoin de normalisation.....	73
	- Principes généraux.....	73
	- Détection du type de biais	75
2.2.	Méthodes de normalisation globales	78
	- Médiane ou moyenne	78
	- Méthode du delta-delta Ct et choix de l'endogène	80
2.3.	Méthodes de normalisation locales	81
	- Normalisation LOESS.....	82
	- Normalisation par les quantiles	83
3.	Outils statistiques pour l'analyse des profils d'expression des microARN ..	85
3.1.	Filtrage des échantillons et des microARN	85
3.2.	Méthodes d'analyse non supervisées.....	87
	- Classification hiérarchique	88
	- Analyse en composantes principales	89
3.3.	Tests statistiques appliqués aux analyses supervisées : T-test et ANOVA	91

RÉSULTATS EXPÉRIMENTAUX	93
INTRODUCTION AU TRAVAIL EXPÉRIMENTAL	94
PROBLÉMATIQUE	94
OUTILS	96
I. MISE AU POINT D'UNE MÉTHODOLOGIE D'ANALYSE DES RÉSULTATS ISSUS DE TLDA	104
1. Des Cts bruts à l'interprétation biologique : importance du traitement des données primaires	104
Article I en soumission : <i>Defining the optimal analysis methodology for miRNA profiling using TaqMan® Low Density Arrays (TLDA) with frozen and formalin-fixed paraffin-embedded samples.</i>	106
2. Validation par comparaison avec les puces microARN	140
2.1. Présentation de la plateforme de puce à microARN	140
2.2. Analyse des résultats des profils d'expression obtenus sur puce	141
2.3. Conclusion	146
II. PRÉDICTION D'UNE RÉPONSE THÉRAPEUTIQUE GRÂCE AUX PROFILS MICROARN	147
Article II : <i>micro-RNA profiles in osteosarcoma as a predictive tool for ifosfamide response.</i>	149
III. APPORTS DES PROFILS MICROARN A LA CLASSIFICATION DES RHABDOMYOSARCOMES	161
1. Difficultés liées à la détection des transcrits rares : essai de développement d'une technologie de puce à ADNc	161
Prix 2010 des Amis de l'Université de Lyon, doté par BIOMNIS, dans la catégorie Nanotechnologie et Génétique Moléculaire, Pharmacogénomique, pour le sujet de thèse : Développement d'une puce à ADN de diagnostic des sarcomes à translocation	162
2. Classification des rhabdomyosarcomes grâce aux profils microARN	185
Article III : <i>MiRNA profiling: how to bypass the current difficulties in the diagnosis and treatment of sarcomas</i>	186
3. Caractérisation génomique des rhabdomyosarcomes à transcrits rares	200
3.1. Recherche de cas rares grâce aux profils d'expression des microARN	200
3.2. Recherche des transcrits rares par technique FISH et PCR	202
3.3. Caractérisation génomique par technique CGH sur puce	203
- Etude de cas : RMS embryonnaire à transcrit rare	204
- Etude de cas : RMS alvéolaire à transcrit rare	206
PUBLICATION ANNEXE	209
Article IV en soumission : <i>Autocrine role for Gas6, Tyro3 and Axl in leiomyosarcomas</i>	210

DISCUSSION ET PERSPECTIVES	238
I. IMPORTANCE DE LA DÉTECTION DES BIAIS BIOLOGIQUES ET DES BIAIS D'ANALYSE	239
1. Biais biologiques	239
2. Biais de manipulations et d'analyse	244
II. LIMITES DES PROFILS D'EXPRESSION DES MICROARN.....	254
1. Aléas technologiques et évolutions	254
2. Limites biologiques	261
2.1. Limites de la classification actuelle des rhabdomyosarcomes	261
2.2. Complexité génomique des ostéosarcomes	263
III. VOIES PEU EXPLORÉES ET PERSPECTIVES.....	267
1. Profils d'expression des microARN et voies métaboliques.....	267
2. Développement d'outils cliniques de routine	270
3. Poursuite générale des projets en cours.....	272
RÉFÉRENCES BIBLIOGRAPHIQUES	275

RESUME en français

Les sarcomes sont des tumeurs malignes des tissus conjonctifs, représentant moins de 1% des tumeurs malignes de l'adulte, mais près de 8% de l'ensemble des cancers pédiatriques. En raison de leur rareté, de leur grande variété histologique et de leur potentiel évolutif hétérogène, les sarcomes sont des pathologies difficiles à traiter, tant sur le plan diagnostique, pronostique que thérapeutique. Ces dernières années, l'avènement de techniques d'analyse pangénomiques par biologie moléculaire a permis d'améliorer la prise en charge clinique des sarcomes, mais les microARN sont des biomarqueurs émergents encore peu utilisés.

Au cours de ce travail de thèse, nous avons choisi d'étudier la valeur des profils d'expression des microARN dans les rhabdomyosarcomes et les ostéosarcomes. Les données brutes des profils d'expression ont été obtenues à l'aide d'une technologie à moyen débit basée sur des réactions de PCR quantitative. Nous avons tout d'abord développé une méthodologie d'analyse permettant d'obtenir des données d'expression précises, reproductibles et à forte valeur ajoutée, à partir de matériel biologique hétérogène. Dans un second temps, nous avons montré que les profils d'expression des microARN permettent d'améliorer la prise en charge clinique des deux types de sarcomes étudiés : il est possible d'affiner la classification nosologique des rhabdomyosarcomes, et de prédire la réponse des ostéosarcomes à la chimiothérapie néo-adjuvante. La recherche de nouvelles applications cliniques liées aux profils d'expression des microARN doit donc être poursuivie, et peut désormais l'être grâce à l'outil robuste que nous avons développé au cours de cette thèse.

TITRE en anglais	Expression profiles of microRNAs in sarcomas: from raw data to clinical applications
------------------	--

RESUME en anglais

Sarcomas are malignant soft tissue tumors, accounting for 1% of adult tumors and 8% of all pediatric malignancies. Sarcomas are rare, and display a variety of histological subtypes and clinical characteristics. Therefore, everyday management is difficult in terms of diagnosis, prognosis and treatment. Recently, the development of pangenomic molecular techniques improved the clinical management of sarcomas, but the use of microRNAs as biomarkers is still being investigated.

In the present work, we studied the value of microRNA expression profiles in rhabdomyosarcomas and osteosarcomas. Raw data of expression profiles were obtained using a medium throughput technology based on quantitative PCR. We first developed an analysis methodology to gain accurate, reproducible and relevant expression data, starting from heterogeneous samples. Furthermore, we showed that microRNA expression profiles can improve the clinical management of both sarcoma entities: they are helpful to upgrade the fine nosological classification of rhabdomyosarcomas, and they are able to predict the response of osteosarcomas to neoadjuvant chemotherapy. Searching for new clinical applications to microRNA expression profiles must be pursued.

DISCIPLINE	Cancérologie
------------	--------------

MOTS-CLES	MicroARN, cancer, sarcomes, profils d'expression, rhabdomyosarcomes, ostéosarcomes, PCR quantitative, normalisation
-----------	---

INTITULE ET ADRESSE DE L'U.F.R. OU DU LABORATOIRE

Equipe Ciblage de la Tumeur et de son Environnement Immunitaire
Centre de Recherche en Cancérologie de Lyon (CRCL), UMR INSERM 1052 - CNRS 5286
Centre Léon Bérard, 28 rue Laënnec, 69373 Lyon cedex 08

LISTE DES PRINCIPALES ABRÉVIATIONS

ABC	Famille des transporteurs ABC (ABCB1 à ABCB6)
ABCB1	<i>ATP-binding cassette, sub-family B (MDR/TAP), member 1</i>
aCGH	<i>Array comparative genomic hybridization</i>
ACP	Analyse en composantes principales
ADICAP	Association pour le développement de l'informatique en cytologie et anatomie pathologique
ADN	Acide désoxyribonucléique
ADNc	Acide désoxyribonucléique complément
AGO	Famille des protéines argonautes (AGO1 à AGO4)
AM	Représentation graphique de type AM
ANM	Académie Nationale de Médecine
ANOVA	<i>Analysis of variance</i>
AP-2 beta	<i>Activating enhancer binding protein 2 beta</i>
ARMS	<i>Alveolar rhabdomyosarcoma</i>
ARN	Acide ribonucléique
ARNi	Acide ribonucléique interférence
ARNm	Acide ribonucléique messenger
ATCC	<i>American Type Culture Collection</i>
BCL2	<i>B-cell CLL/lymphoma 2</i>
CCND1	<i>Cyclin D1</i>
CDK	<i>Cyclin dependent kinase</i>
CDK4	<i>Cyclin dependent kinase 4</i>
CDK6	<i>Cyclin dependent kinase 6</i>
CDKN1A	<i>Cyclin dependent kinase inhibitor 1A (P21)</i>
CDKN2A	<i>Cyclin dependent kinase inhibitor 2A (P16, P14)</i>
CGH	<i>Comparative genomic hybridization</i>
CINSARC	<i>Complexity index in sarcomas</i>
CIT	Carte d'Identité des Tumeurs®
CONTICANET	<i>Connective tissue cancer network</i>
Ct	<i>Cycle threshold</i>
$\Delta\Delta Ct$	Méthode du delta delta Ct
DDIT3	<i>DNA-damage-inducible transcript 3</i>
DGCR8	<i>DiGeorge syndrome critical region gene 8</i>
E	Efficacité (de la PCRq)
E2F1	<i>E2F transcription factor 1</i>
EDTA	Acide éthylène diamine tétraacétique
EGFR	<i>Epidermal growth factor receptor</i>
EJC	<i>European Journal of Cancer</i>
ERBB2	<i>v-erb-b2 erythroblastic leukemia viral oncogene homolog 2</i>
ERMS	<i>Embryonal rhabdomyosarcoma</i>

FFPE	<i>Formalin fixed, paraffin embedded</i>
FGFR1	<i>Fibroblast growth factor receptor 1</i>
FISH	<i>Hybridation in situ fluorescente</i>
FOXO	<i>Famille des gènes FOXO (FOXO1, FOXO4)</i>
FOXO1	<i>Forkhead box O1 (FKHR)</i>
FOXO4	<i>Forkhead box O4 (MLLT7)</i>
GIST	<i>Gastro-intestinal stromal tumor</i>
GO	<i>Gene ontology</i>
HES	<i>Héματοxyline, éosine, safran</i>
HRAS	<i>v-Ha-ras Harvey rat sarcoma viral oncogene homolog</i>
ICR	<i>International Classification of Rhabdomyosarcoma</i>
IGF-1R	<i>Insulin-like growth factor 1 receptor</i>
IGF2	<i>Insulin-like growth factor 2</i>
IHC	<i>Immunohistochimie</i>
IFO	<i>Ifosfamide</i>
IL1	<i>Interleukin 1</i>
IRM	<i>Imagerie par résonance magnétique</i>
IRSG	<i>Intergroup Rhabdomyosarcoma Study Group</i>
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i>
KIT	<i>v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog</i>
KRAS	<i>v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog</i>
LMS	<i>Léiomyosarcome</i>
LNA	<i>Locked nucleic acid</i>
LOESS	<i>Locally estimated scatterplot smoothing</i>
LOH	<i>Loss of heterozygosity</i>
LOWESS	<i>Locally weighted scatterplot smoothing</i>
MAPK	<i>Famille des gènes Mitogen-activated protein kinase</i>
MAP3K5	<i>Mitogen-activated protein kinase kinase kinase 5</i>
MDM2	<i>p53 E3 ubiquitin protein ligase homolog</i>
MDM4	<i>Mdm4 p53 binding protein homolog</i>
MDR	<i>multi-drug resistance</i>
MINDACT	<i>Microarray In Node negative Disease may Avoid ChemoTherapy</i>
MRE	<i>MicroRNA response element</i>
MYC	<i>v-myc myelocytomatosis viral oncogene homolog</i>
MYCN	<i>v-myc myelocytomatosis viral related oncogene (N-myc)</i>
nARMS	<i>Translocation-negative alveolar rhabdomyosarcoma</i>
NCOA	<i>Famille des gènes NCOA (NCOA1, NCOA2)</i>
NCOA1	<i>Nuclear receptor coactivator 1</i>
NCOA2	<i>Nuclear receptor coactivator 2</i>
NF1	<i>Neurofibromin 1</i>
NOS	<i>Not otherwise specified</i>
NRAS	<i>neuroblastoma RAS viral (v-ras) oncogene homolog</i>
OMS	<i>Organisation Mondiale de la Santé</i>

OS	Ostéosarcome
pARMS	<i>Translocation-positive alveolar rhabdomyosarcoma</i>
PAX	Famille des gènes PAX (PAX3, PAX7)
PAX3	<i>Paired box 3</i>
PAX3-FOXO1	Transcrit de fusion constitué des gènes PAX3 et FOXO1
PAX3-FOXO4	Transcrit de fusion constitué des gènes PAX3 et FOXO4
PAX3-NCOA1	Transcrit de fusion constitué des gènes PAX3 et NCOA1
PAX3-NCOA2	Transcrit de fusion constitué des gènes PAX3 et NCOA2
PAX7	<i>Paired box 7</i>
PAX7-FOXO1	Transcrit de fusion constitué des gènes PAX7 et FOXO1
PCR	<i>Polymerase chain reaction</i>
PCRq	<i>Quantitative polymerase chain reaction</i>
PTEN	<i>Phosphatase and tensin homolog</i>
RAS	Famille des gènes RAS (HRAS, KRAS, NRAS)
RB1	<i>Retinoblastoma 1</i>
RISC	<i>RNA-induced silencing complex</i>
RMS	Rhabdomyosarcome
RQ	Quantité relative
RT	<i>Reverse transcription</i>
RT-PCR	<i>Reverse transcription - Polymerase chain reaction</i>
SAI	Sans autres indications
SIOP	Société Internationale d'Oncologie Pédiatrique
siARN	<i>Small interfering RNA</i>
SNP	<i>Single nucleotide polymorphism</i>
SQSTM1	<i>Sequestosome 1</i>
SS18	<i>Synovial sarcoma translocation, chromosome 18</i>
TFAP2B	<i>Transcription factor AP-2 beta</i>
TGFβ	<i>Transforming growth factor, beta 1 (TGFB1)</i>
TLDA	<i>TaqMan® Low Density Array</i>
TMA	<i>Tissue micro array</i>
TNF	<i>Tumor necrosis factor</i>
TNFRSF11A	<i>Tumor necrosis factor receptor superfamily, member 11a, NFKB activator (RANK)</i>
TP53	<i>Tumor protein P53</i>
TRBP	<i>TAR (HIV-1) RNA binding protein 2</i>
ULS	<i>Universal linkage system</i>
UNA	<i>Unlocked nucleic acid</i>
UTR	<i>Untranslated region</i>
VEGF	<i>Vascular endothelial growth factor</i>
VEGFR	<i>Vascular endothelial growth factor receptors</i>
Wnt	Famille des gènes <i>wingless-type MMTV integration site</i>
XPO5	<i>Exportin 5</i>

PRÉAMBULE

Les premières descriptions de cancer ont été rapportées par les égyptiens (approximativement 1500 ans avant J.C.), et la toute première tentative de classification de ces maladies a été effectuée par Hippocrate, environ 400 avant J.C. (Weiss, 2000). La prise en charge des cancers a évolué très lentement au cours des siècles, et c'est seulement à partir des années 1950 que les avancées technologiques en biologie cellulaire et moléculaire ont permis une accélération des connaissances. Les cancers sont aujourd'hui en cours de caractérisation génomique exhaustive, grâce à des technologies fines de biologie moléculaire permettant de traiter l'ensemble des informations biologiques des cellules tumorales en une seule expérience. Grâce à ces nouvelles technologies, les concepts classiques de classification sont bouleversés, et nous n'avons jamais été aussi proches de pouvoir réaliser une analyse exhaustive du génome tumoral à la faveur de l'identification de la carte d'identité des tumeurs (CIT, telle que souhaitée par le programme français du même nom initié en 2000 et financé par la Ligue contre le cancer).

Les techniques permettant l'analyse de l'ensemble du génome, du transcriptome et du protéome ont débouché sur de nouvelles classifications nosologiques, et sur de multiples applications au diagnostic, au pronostic et au traitement des tumeurs grâce à la découverte concomitante de nombreux biomarqueurs d'intérêt. En particulier, l'étude du transcriptome tumoral par l'établissement des profils d'expression des ARN messagers (ARNm) a énormément fait progresser la recherche fondamentale sur le cancer, mais ces profils d'expression sont en cours d'évaluation dans des essais cliniques et sont encore peu utilisés pour la prise en charge clinique de routine des patients (Tiwari, 2012). On peut toutefois citer deux signatures d'expression ARNm prometteuses actuellement en cours d'évaluation : la signature d'Amsterdam, via l'étude MINDACT (*Microarray In Node negative Disease may Avoid ChemoTherapy*), qui permettrait de mieux traiter les patients porteurs d'un cancer du sein en situation adjuvante (van de Vijver et al., 2002), et la signature CINSARC (*Complexity index in sarcomas*), permettant de prédire le risque métastatique dans les sarcomes (Chibon et al., 2010). De manière générale,

on observe aujourd'hui une démocratisation de l'utilisation de ces techniques d'établissement des profils d'expression. Cela entraînera très certainement leur utilisation en routine dans un futur proche.

Les profils d'expression des microARN, contrairement à ceux des ARNm, sont encore peu étudiés, en partie en raison de la découverte récente de ces entités, mais aussi parce que les technologies d'analyse ne sont pas encore stabilisées. De nouveaux gènes de microARN sont découverts quotidiennement : les bases de données actuelles ne sont donc pas exhaustives en termes de gènes décrits, et ne le sont pas également en termes de cibles ARNm potentielles. Agissant à un tout autre niveau dans la cellule tumorale, l'étude des profils d'expression des microARN est un terrain d'exploration qui pourrait apporter des réponses là où d'autres analyses ciblant l'ADN, l'ARN ou les protéines n'ont pas encore pu apporter de réponses. A terme, l'intégration de ces différents niveaux d'analyse pourrait permettre une caractérisation optimale de chaque type tumoral.

Dans ce contexte, notre travail se propose de définir les modalités optimales d'analyse des profils d'expression de ces microARN, afin de les explorer dans deux types de tumeurs malignes des tissus conjonctifs, le rhabdomyosarcome et l'ostéosarcome. Nous avons choisi de travailler sur ces deux types de sarcomes en raison d'un fort recrutement local lié au panel pluridisciplinaire d'experts présents au Centre Léon Bérard, mais également parce que ce sont des tumeurs difficiles à prendre en charge sur le plan clinique. De par leur rareté et leur diversité, ces sarcomes constituent de bons modèles d'étude, pour lesquels il est possible d'apporter une réelle plus-value dans le cadre de la prise en charge clinique des patients. Les profils d'expression des microARN ont été établis au moyen d'une technologie dite à moyen débit basée sur la PCR quantitative (PCRq), à partir d'échantillons tumoraux congelés ou inclus en paraffine.

Ce travail de thèse a tout d'abord consisté à montrer qu'il était possible, à partir des données brutes délivrées par les appareils, d'obtenir des données biologiquement

significatives, représentatives des niveaux d'expression des microARN, et utilisables par des outils d'analyse puissants adaptés à des problématiques de démantèlement et de découverte de biomarqueurs. Pour cela, nous avons développé une méthode de normalisation et d'analyse des données en nous inspirant des méthodes traditionnellement utilisées pour des technologies de type puces d'expression, en prenant soin de réduire au maximum les biais biologiques et techniques. Dans un second temps, nous avons appliqué cette méthode d'analyse aux profils d'expression des microARN obtenus à partir de tumeurs de type rhabdomyosarcome ou ostéosarcome, et évalué la capacité des profils d'expression des microARN à résoudre des problématiques propres à chaque entité. Ainsi, nous avons montré qu'il était possible d'affiner la classification des rhabdomyosarcomes grâce aux profils d'expression des microARN, et que ces profils permettaient également, chez les ostéosarcomes, de prédire la réponse à la chimiothérapie.

Dans l'introduction bibliographique, nous aborderons la biologie des rhabdomyosarcomes et des ostéosarcomes, en insistant sur les problématiques cliniques pour lesquelles l'étude des profils d'expression des microARN nous a permis d'apporter des réponses. Nous présenterons ensuite la nature, le rôle et les fonctions physiologiques des microARN, ainsi que leurs implications dans les processus tumoraux. Enfin, nous aborderons les technologies couramment utilisées pour l'établissement des profils d'expression des microARN, et les principales méthodes de normalisation et d'analyse statistique qui peuvent leur être appliquées.

INTRODUCTION BIBLIOGRAPHIQUE

I. PRESENTATION DU RHABDOMYOSARCOME ET DE L'OSTEOSARCOME

1. Les rhabdomyosarcomes

1.1. Description et épidémiologie

Les rhabdomyosarcomes (RMS) sont des tumeurs malignes des tissus mous qui se développent à partir de cellules mésenchymateuses de la lignée musculaire squelettique (Parham, 2001). Le nom rhabdomyosarcome dérive des mots grecs *rhabdos*, signifiant baguette, en référence aux fibres musculaires, et *mus*, signifiant muscle. Un RMS fut décrit pour la première fois par Weber en 1854, à propos d'une masse de la langue présentant des cellules musculaires striées à différents stades de différenciation, chez un homme de 21 ans (Weber, 1854). Un siècle plus tard, Stout donna une définition histologique plus précise de la morphologie particulière des rhabdomyoblastes (Stout, 1946).

Les RMS sont les sarcomes des tissus mous les plus courants chez l'enfant : ils représentent environ la moitié des sarcomes des tissus mous pédiatriques (Pastore et al., 2006; Howlader et al., 2012). Ce sont cependant des tumeurs rares, représentant seulement environ 4,5 % de la totalité des cancers pédiatriques (Leaphart and Rodeberg, 2007; Huh and Skapek, 2010). L'incidence annuelle des RMS est d'environ 5 cas par million d'enfants de moins de 15 ans, aux Etats-Unis comme en France (Sommelet et al., 2009; Howlader et al., 2012).

L'étiologie et les facteurs de risques spécifiques sont mal connus. Quelques études suggèrent une incidence plus importante des RMS chez les individus exposés à des radiations *in utero* (Grufferman et al., 2009), chez les enfants ayant reçu des antibiotiques peu après la naissance (Hartley et al., 1988) et chez ceux dont les parents consommaient des drogues pendant la grossesse (Grufferman et al., 1993).

Le développement des RMS est associé à un certain nombre de syndromes familiaux rares liés au cancer tels que la neurofibromatose de type 1 (Matsui et al.,

1993), caractérisée par des mutations germinales du gène NF1 (Goldberg and Collins, 1991), le syndrome de Li-Fraumeni (Li and Fraumeni, Jr., 1969), caractérisé par des mutations germinales du gène TP53 (Diller et al., 1995), le syndrome de Beckwith-Wiedemann (Smith et al., 2001), lié à une dérégulation de l'expression du gène IGF2 (Reik et al., 1994), ainsi que le syndrome de Costello (Gripp et al., 2002), caractérisé par des mutations germinales du gène HRAS (Aoki et al., 2005).

1.2. Diagnostic et classification

- Présentation clinique et diagnostique

Bien que les RMS puissent apparaître dans toutes les parties du corps (Wexler et al., 2006), les localisations préférentielles se situent au niveau de la tête et du cou, de l'appareil urogénital et des membres. Trois études de l'*Intergroup Rhabdomyosarcoma Study Group* (IRSG) ont montré que 35 % des tumeurs se développaient au niveau de la tête (principalement dans l'orbite et en localisation paraméningée) ou sur la nuque, environ 24 % au niveau de l'appareil urogénital (vessie et prostate, paratesticulaire, vagin et utérus), 19 % sur les membres et 22 % dans d'autres sites (Maurer et al., 1988; Maurer et al., 1993; Crist et al., 1995).

Les RMS appartiennent au groupe des tumeurs à petites cellules rondes basophiles. L'analyse histopathologique s'attache à identifier les caractéristiques de la lignée musculaire squelettique en utilisant des techniques variées. L'immunohistochimie (IHC) est impérativement utilisée pour orienter le diagnostic. On s'attache en particulier à identifier la présence de myogénine, un marqueur spécifique du muscle squelettique indispensable à l'authentification d'un RMS. On utilise la myogénine en association avec la desmine, cette dernière étant presque toujours positive (Molenaar et al., 1985; Kumar et al., 2000; Wexler et al., 2006).

- Classification

La classification des RMS en différents types et sous-types a été la source de débats considérables, car il existe de nombreux variants morphologiques. Le tableau 1 résume les principales évolutions de cette classification. Le premier système de classification standard a été élaboré par Horn et Enterline en 1958, sur une série de 39 cas (Horn and Enterline, 1958). Essentiellement basé sur la morphologie, il reconnaissait l'existence de 4 grands groupes de RMS : les rhabdomyosarcomes alvéolaires (ARMS), les rhabdomyosarcomes embryonnaires (ERMS), les ERMS de type botryoïdes et les rhabdomyosarcomes pléomorphes.

De nombreuses autres classifications ont ensuite été proposées, dont la classification cyto-histologique de Palmer en 1978, celle de la Société Internationale d'Oncologie Pédiatrique (SIOP) en 1989 (Caillaud et al., 1989), et celle du *National Cancer Institute* (NCI) en 1990 (Parham, 2001). C'est en 1995 que sera élaborée pour les RMS pédiatriques l'*International Classification of Rhabdomyosarcoma* (ICR), qui sera largement adoptée par la communauté médicale (Newton, Jr. et al., 1995). L'ICR reconnaît alors trois grandes entités : les ARMS, de mauvais pronostique, les ERMS, de pronostique intermédiaire (dont les sous-types botryoïde et à cellules fusiformes, de pronostique plus favorable), et les RMS indifférenciés ou sans autres indications (SAI), de pronostique défavorable.

En 2002, l'organisation mondiale de la santé (OMS) publie une classification générale des tumeurs des tissus mous et des os, intégrant les dernières données de génétique (Fletcher et al., 2002). Une révision de cette classification est actuellement à l'étude, avec une publication prévue pour 2013. Dans cette dernière classification, les formes anaplasiques ne représenteraient plus des variants à part entière : ce critère sera cependant retrouvé implicitement dans les RMS pléomorphes de l'adulte, et la notion d'anaplasie sera précisée, si présente dans les RMS pédiatriques. Cette

nouvelle classification devrait voir apparaître trois nouvelles entités : le RMS à cellules fusiformes de l'adulte, le RMS sclérosant et le RMS épithélioïde (Nascimento and Fletcher, 2005; Jo et al., 2011; Robinson et al., 2012). Les RMS SAI (ou NOS pour *not otherwise specified*) devraient également être redéfinis avec de nouveaux critères (non représentativité de l'échantillon / zone grise).

De manière générale, les ERMS et les ARMS sont considérés comme les deux sous-types majeurs de RMS. Les ERMS représentent plus de la moitié des nouveaux cas de RMS chez les patients les plus jeunes, tandis que les ARMS sont plus communément retrouvés chez les adolescents, et représentent 20 à 30 % des cas (Merlino and Helman, 1999; Wexler et al., 2006).

Date	1958	1989	1995		2002	2013 (orientations)	
Nom	Horn-Enterline	SIOP	Newton (ICR)		Fletcher (OMS)	Fletcher (OMS)	
Base	-Morphologie	- Morphologie - Cytologie	- Morphologie - Cytologie - Pronostique		- Morphologie - Cytologie - Génétique	- Morphologie et Cytologie - Génétique - Âge du patient	
Classification	ARMS	ARMS (dont les formes mixtes si > 50 % d'histologie alvéolaire)	MP	ARMS (dont les formes mixtes quelle que soit l'étendue de la zone alvéolaire)	ARMS - conventionnel - forme solide - mixte ARMS/ERMS	ARMS	
	ERMS	ERMS - dense, bien / peu différencié - SAI - lâche, non botryoïde	PI	ERMS	ERMS - à cellules fusiformes - botryoïde - anaplasique	P	ERMS conventionnel
			BP	ERMS à cellules fusiformes		P	ERMS à cellules fusiformes
	ERMS botryoïde	ERMS lâche, botryoïde		ERMS botryoïde			ERMS botryoïde
	-	Embryonal sarcoma	-		-	-	-
	RMS pléomorphe	RMS pléomorphe	Absents (classification pédiatrique)		RMS pléomorphe	A	RMS pléomorphe
	-	-	MP	RMS indifférencié / SAI	-		RMS SAI (si non représentativité de l'échantillon / zone grise)
	-	-	-		-		RMS sclérosant
	-	-	-		-		RMS à cellules fusiformes
	-	-	-		-		RMS épithélioïde

Tableau 1 : principales évolutions de la classification des RMS, de 1958 à nos jours.

(BP = bon pronostic, PI = pronostic intermédiaire, MP = mauvais pronostic, A = formes présentes préférentiellement chez les adultes, P = formes préférentiellement pédiatriques).

1.3. Caractéristiques biologiques des ERMS

- Anatomopathologie

Histologiquement, les ERMS sont caractérisés par un polymorphisme cellulaire et la présence de zones myxoïdes. La population cellulaire est constituée de cellules rondes ou fusiformes à noyaux hyperchromatiques et/ou de grandes cellules tumorales à cytoplasme éosinophile abondant. Les cellules tumorales peuvent montrer tous les stades de différenciation musculaire, de cellules mésenchymateuses indifférenciées à des myoblastes allongés, des myotubes et des myofibres complètement différenciées (figure 1) (Parham, 2001). Le variant à cellules fusiformes (principalement retrouvé au niveau des régions paratesticulaire et de la tête et du cou) montre des cellules allongées avec une architecture fasciculée, et une morphologie qui peut être pseudomusculaire lisse. Le variant botryoïde se présente comme une masse polypoïde de cellules tumorales denses se développant principalement sous la surface des épithéliums du vagin ou de la vessie. Il existe également des variants plus indifférenciés, qui sont alors difficiles à distinguer des ARMS de forme solide.

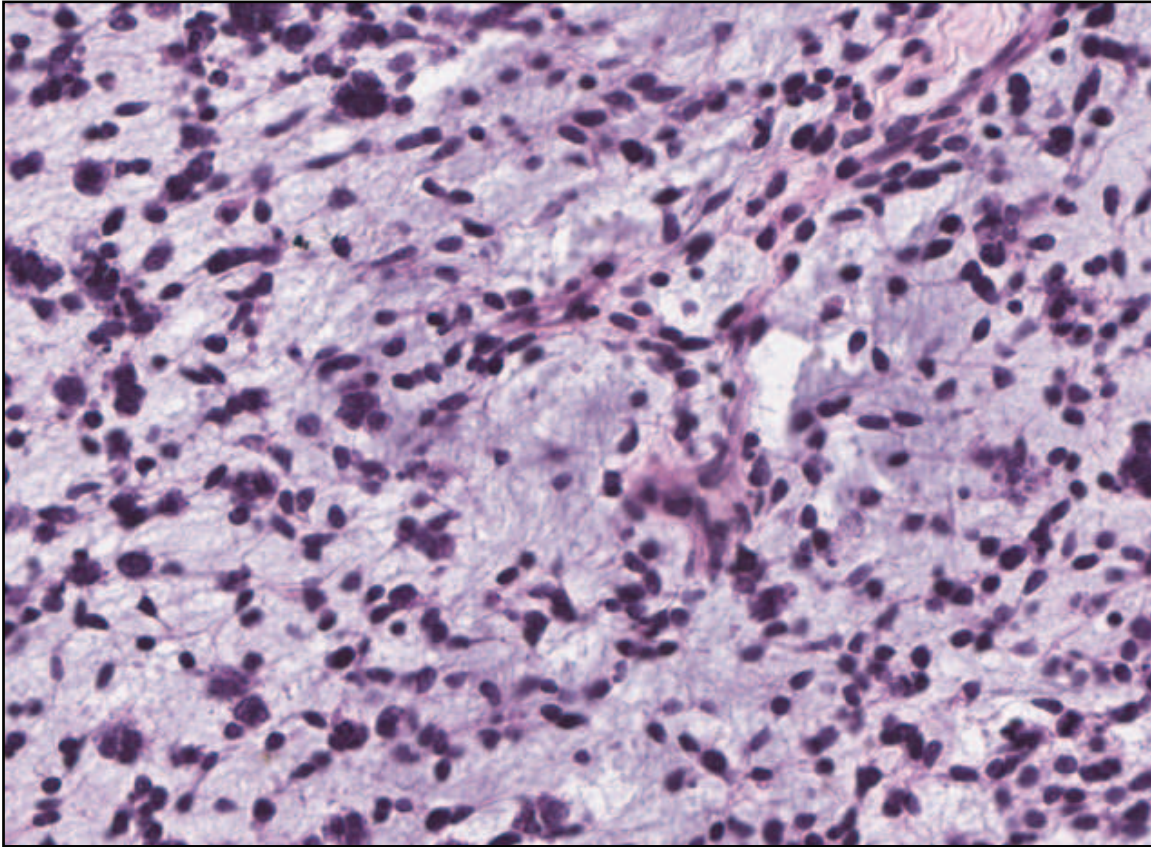


Figure 1 : coupe en paraffine colorée hématoxyline-éosine-safran (HES) d'un ERMS.

Les cellules montrent différents stades de différenciation sur un fond myxoïde. Les cellules peu différenciées sont généralement peu chromatiques ou en fuseau. Les cellules différenciées sont larges et possèdent un cytoplasme éosinophile (image microscopique capturée par le Dr. Dominique Ranchère-Vince, Département de Biopathologie, Centre Léon Bérard, Lyon).

- Caractéristiques génétiques et moléculaires

Le sous-type des ERMS est caractérisé par une perte d'hétérozygotie (LOH) sur le bras court du chromosome 11 (11p15) (Scrabble et al., 1989). Des publications précédentes ont montré que la LOH complète ou partielle de cette région pouvait entraîner la formation de cancers variés tels que les cancers du sein, de l'ovaire, de la vessie, du poumon, des testicules ainsi que les tumeurs de Wilms (Winqvist et al., 1993). Le locus 11p15 contient le gène *insulin-like growth factor 2* (IGF2), qui chez les mammifères est soumis à l'empreinte maternelle. L'IGF2 est un facteur de croissance crucial qui se lie à l'*insulin-like growth factor 1 receptor* (IGF-1R) (Anderson et al.,

1999a). La voie de signalisation passant par l'IGF-1R est associée à la tumorigenèse et à l'inhibition de l'apoptose (Werner and Le Roith D., 1997). La LOH avec disomie uniparentale paternelle conduit à une surexpression de l'IGF2, phénomène démontré comme étant important lors de la survenue d'un ERMS (Zhan et al., 1994). Il est intéressant de noter que la perte de l'empreinte du gène IGF2 a déjà été retrouvée dans des tumeurs primitives de type ARMS, produisant également une surexpression de l'IGF2 (Anderson et al., 1999a).

Les nombreuses études en cytogénétique conventionnelle et moléculaire n'ont pas identifié d'autre anomalie chromosomique récurrente et spécifique. Il est cependant fréquemment observé de nombreuses anomalies numériques de type gain, avec un gain particulièrement élevé du chromosome 8 (Anderson et al., 1999b; Williamson et al., 2010). Il est décrit, de manière exceptionnelle, des translocations ou des points de cassure impliquant le gène PAX3, mais les éventuels gènes partenaires n'ont pas encore été identifiés (Hayashi et al., 1988; Ho et al., 2004; Hosoi et al., 2009).

1.4. Caractéristiques biologiques des ARMS

- Anatomopathologie

Les ARMS sont généralement caractérisés par une cytologie particulière : les cellules sont monomorphes, les noyaux possèdent des contours plus ou moins irréguliers, avec une chromatine hétérogène. Il existe deux principaux types d'architecture : les ARMS classiques sont caractérisés par des agrégats denses de petites cellules rondes indifférenciées qui fixent intensément l'hématoxyline. Ces agrégats de cellules sont typiquement séparés par des septa fibreux, formant des structures ressemblant à des alvéoles pulmonaires (figure 2). Les ARMS peuvent également se présenter sans cette caractéristique, avec des septa fibreux peu visibles : on parle alors de variant alvéolaire de forme solide (Tsokos et al., 1992).

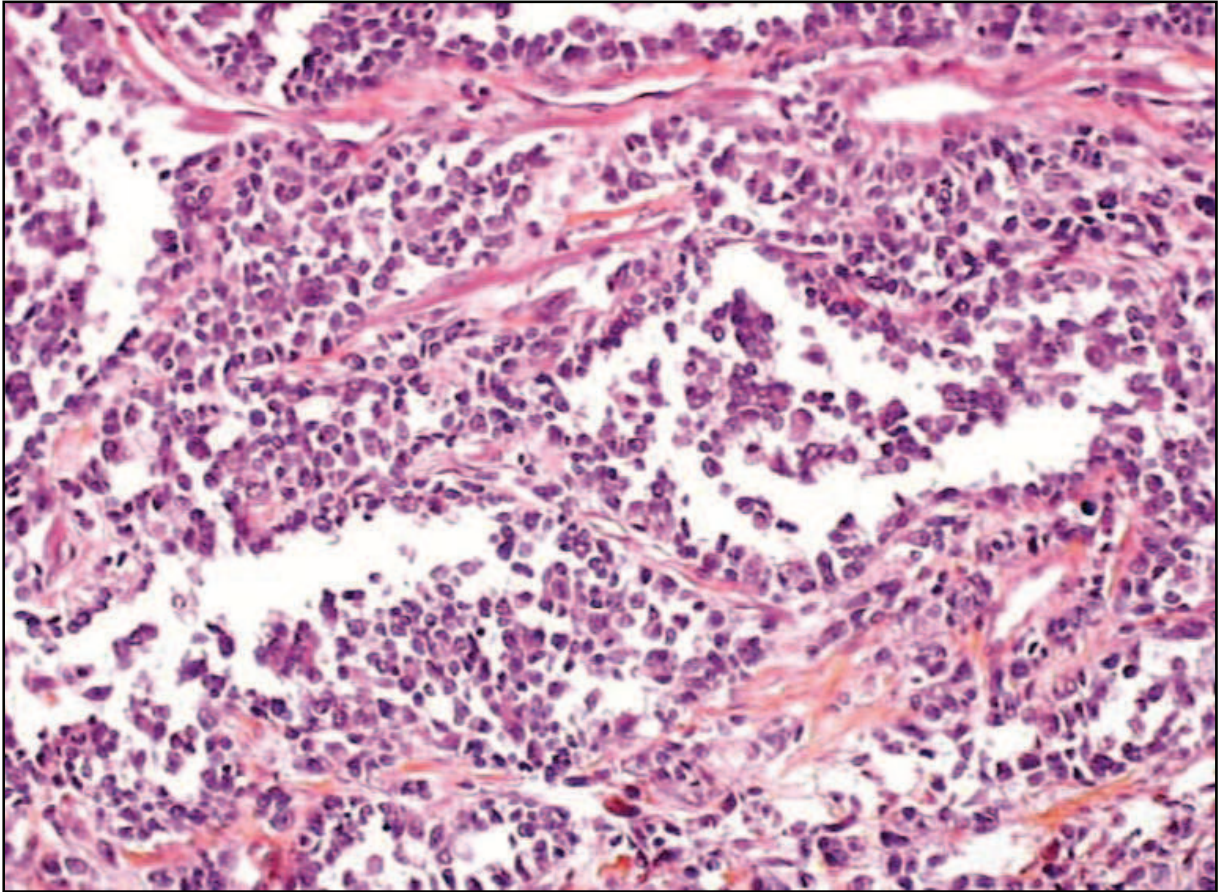


Figure 2 : coupe en paraffine colorée HES d'un ARMS.

L'image montre de petites cellules rondes avec un marquage bleu intense des noyaux. Les cellules sont organisées en groupes séparés par des septa fibreux (image microscopique capturée par le Dr. Dominique Ranchère-Vince, Département de Biopathologie, Centre Léon Bérard, Lyon).

- Caractéristiques génétiques et moléculaires

La signature moléculaire qui caractérise le plus souvent les ARMS implique la translocation chromosomique $t(2;13)(q35;q14)$ (Turc-Carel et al., 1986). Le transcrit de fusion résultant, liant les gènes *paired box 3* et *forkhead box O1* (PAX3-FOXO1) a été identifié pour la première fois en 1993 (Barr et al., 1993). Le second variant le plus fréquent est issu de la translocation $t(1;13)(p36;q14)$, et génère le transcrit de fusion PAX7-FOXO1 (Biegel et al., 1991; Davis et al., 1994b). Plus rarement, quelques cas d'ARMS ont été décrits avec des translocations alternatives de type $t(X;2)(q13;q35)$, $t(2;2)(p23;q35)$ et $t(2;8)(q35;q13)$, qui génèrent respectivement les transcrits de fusion

PAX3-FOXO4, PAX3-NCOA1 et PAX3-NCOA2 (Barr et al., 2002; Wachtel et al., 2004; Sumegi et al., 2010). Il existe également un unique cas d'ARMS de variant solide pour lequel il est décrit un transcrit de fusion amplifié de type FOXO1-FGFR1 (Liu et al., 2011).

La majorité des ARMS exprime donc la protéine de fusion PAX3-FOXO1 ou PAX7-FOXO1. Les gènes PAX3 et PAX7 sont tous deux des membres de la famille des gènes *paired box*. PAX7 est l'homologue le plus proche du gène PAX3. Ces deux gènes partagent une organisation génétique similaire, ainsi des fonctions similaires, même si leurs profils d'expression sont distincts (Barr et al., 1999; Barr, 2001; Relaix et al., 2006). Dans les deux cas, les points de cassure se produisent au niveau du septième intron des gènes PAX, et dans le premier grand intron du gène FOXO1. Il en résulte des protéines chimériques contenant le domaine N-terminal de liaison à l'ADN du gène PAX3 ou PAX7 (domaines *paired box* et *homeobox*) fusionné au domaine C-terminal de transactivation du gène FOXO1 (Barr, 2001) (figure 3). Ce principe de création d'une protéine chimérique possédant les capacités spécifiques des deux partenaires impliqués reste similaire dans le cas des transcrits rares.

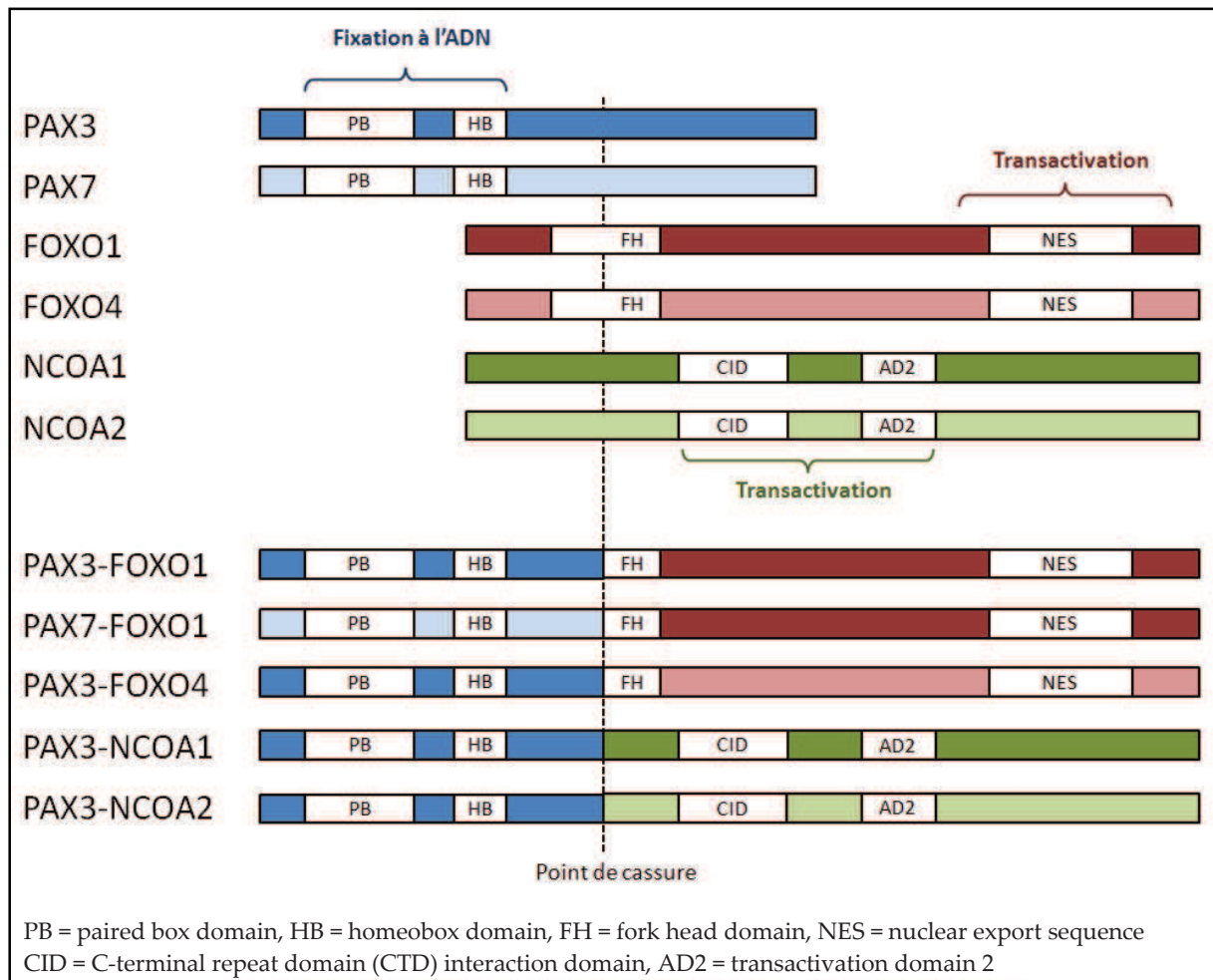


Figure 3 : représentation schématique des principales translocations retrouvées dans les ARMS.

Les protéines chimériques résultantes contiennent les domaines de liaison à l'ADN des gènes PAX fusionnés avec les domaines de transactivation des gènes FOXO ou NCOA. Le résultat est la production de facteurs de transcription aberrants, qui possèdent les capacités de liaison à l'ADN des gènes PAX, et les activités potentielles de transactivation transcriptionnelle des gènes FOXO ou NCOA.

L'année de sa découverte, le transcrit de fusion PAX3-FOXO1 a été cloné par deux groupes indépendants (Galili et al., 1993; Shapiro et al., 1993). Depuis son clonage, de nombreuses études ont été menées pour comprendre le rôle PAX3-FOXO1 dans la « rhabdomyosarcomagenèse ». Les premiers travaux réalisés dans des fibroblastes de poulet ou dans des lignées cellulaires myogéniques ont montré que PAX3-FOXO1 induisait une croissance indépendante de l'ancrage et inhibait la différenciation myogénique (Epstein et al., 1995; Scheidler et al., 1996). Plus tard,

d'autres études se sont penchées sur les multiples aspects de la transformation cellulaire opérée par PAX3-FOXO1, montrant que les effets monogéniques de son expression influençaient de manière aberrante la prolifération, la survie et la différenciation cellulaire (Hollenbach et al., 1999; Margue et al., 2000; Anderson et al., 2001; Wang et al., 2005; Miller and Hollenbach, 2007; Roeb et al., 2007; Roeb et al., 2008).

De manière générale, une meilleure compréhension de la fonction des protéines chimériques permettra d'élucider le mécanisme moléculaire aboutissant à la formation des ARMS, et permettra d'identifier des cibles thérapeutiques potentielles. Des études ont toutefois montré que l'expression seule de PAX3-FOXO1 n'était pas suffisante pour induire la formation d'un ARMS (Anderson et al., 2001; Lagutina et al., 2002; Keller et al., 2004). Il est probable qu'une combinaison de voies métaboliques altérées affectant la prolifération et la survie cellulaires ainsi qu'une expression de PAX3-FOXO1 soient nécessaires pour le développement d'un ARMS (Diller et al., 1995).

1.5. Événements génétiques secondaires présents dans les RMS

Un résumé des événements génétiques primaires et additionnels qui participent à la tumorigenèse des RMS est présenté dans la figure 4. Il est intéressant de remarquer que les gènes décrits ci-dessous sont souvent connus pour être les moteurs des différents syndromes familiaux de prédisposition aux RMS. Il semblerait que pour les ARMS comme pour les ERMS, plusieurs événements génétiques séquentiels soient indispensables à l'émergence du cancer ; ces événements sont complémentaires, agissant d'une part sur la prolifération et la survie cellulaires, et d'autre part sur le blocage de la différenciation cellulaire.

- Événements favorisant la survie cellulaire

Un des facteurs clés contribuant au développement des RMS est le gène TP53. Il s'agit d'un facteur de transcription qui, en réponse à un stress cellulaire et/ou une activation oncogénique, peut induire de nombreux signaux anti-prolifératifs et apoptotiques (Sherr, 2004). Le gène MDM2 est un régulateur négatif crucial de l'activité de TP53 : l'activité ligase E3 du gène MDM2 permet l'ubiquitination du domaine de transactivation de TP53, ce qui induit sa dégradation par le protéasome (Kruse and Gu, 2009; Wade et al., 2010). L'inactivation de la voie TP53 se produit dans de très nombreux cancers, dont les sarcomes (Kusafuka et al., 1997; Yoo et al., 1997), avec une fréquence variable selon les types de tumeurs (Soussi, 2011). Plusieurs études ont décrit des mutations de TP53 et une amplification du gène MDM2 dans des lignées cellulaires et des tumeurs primitives de RMS (Keleti et al., 1996; Taylor et al., 2000; Takahashi et al., 2004). Ces altérations suggèrent donc un rôle important de la voie TP53 dans la biogenèse des RMS.

- Événements favorisant la prolifération cellulaire

La voie du gène suppresseur de tumeur *retinoblastoma 1* (RB1) est une autre voie métabolique qui pourrait contribuer au développement des RMS. Le blocage de cette voie par la dérégulation et l'inactivation des inhibiteurs des *cyclin dependent kinase* (CDK) CDKN2A et CDKN1A est souvent impliqué dans la formation de tumeurs (Sherr, 1996). En particulier, CDKN2A peut inhiber l'activité des *cyclin dependent kinases* 4 et 6 (CDK4 / CDK6), ce qui entraîne une hypophosphorylation et une activation de la protéine RB1 qui, à son tour, empêche le facteur de transcription E2F1 d'être activé, ce qui bloque les cellules en phase G1 (Sherr and Weber, 2000). La perte de CDKN2A a été décrite dans des lignées cellulaires et dans des tumeurs primitives de type ARMS et ERMS, ce qui veut dire que cet inhibiteur de CDK pourrait être impliqué dans le développement des RMS pédiatriques (Iolascon et al., 1996; Obana et al., 2003).

- Autres événements contribuant à la tumorigenèse des RMS

En plus de l'inactivation des voies TP53 / RB1, plusieurs autres gènes ont été décrits comme étant impliqués dans le développement des RMS.

L'oncogène MYCN a été très étudié dans les neuroblastomes : il joue un rôle important dans la prolifération cellulaire et la différenciation (Westermarck et al., 2011). De nombreuses études ont montré une amplification du gène MYCN dans des tumeurs primitives, ainsi que dans des lignées cellulaires d'ARMS (Dias et al., 1990; Driman et al., 1994; Pandita et al., 1999). De manière intéressante, il semblerait que comme dans le neuroblastome, l'amplification du gène MYCN soit un marqueur de mauvais pronostic dans les ARMS (Hachitanda et al., 1998).

Des mutations dans les gènes de la famille RAS ont été retrouvées dans de nombreux cancers (Bos, 1989). La famille RAS est constituée des gènes HRAS, KRAS et NRAS. L'activation de la voie RAS est retrouvée de façon prédominante dans les ERMS, sous la forme de mutations activatrices pouvant toucher les trois membres de cette famille (Shukla et al., 2012). Des expérimentations *in vitro* et *in vivo* ont démontré que l'activation de la voie RAS était impliquée dans la tumorigenèse des ERMS et des ARMS, par blocage de la différenciation (Linardic and Counter, 2008; Ren et al., 2008). Toujours dans les ERMS, une étude récente a démontré que la voie RAS pouvait être indirectement activée, en l'absence de mutations des gènes RAS, par une perte de fonction du gène NF1 via des délétions homozygotes (Paulson et al., 2011). Le gène NF1 étant inactivé, la voie RAS est activée puisque dans des conditions physiologiques normales, NF1 augmente l'activité GTPase de RAS, qui se retrouve alors inactive (Le and Parada, 2007).

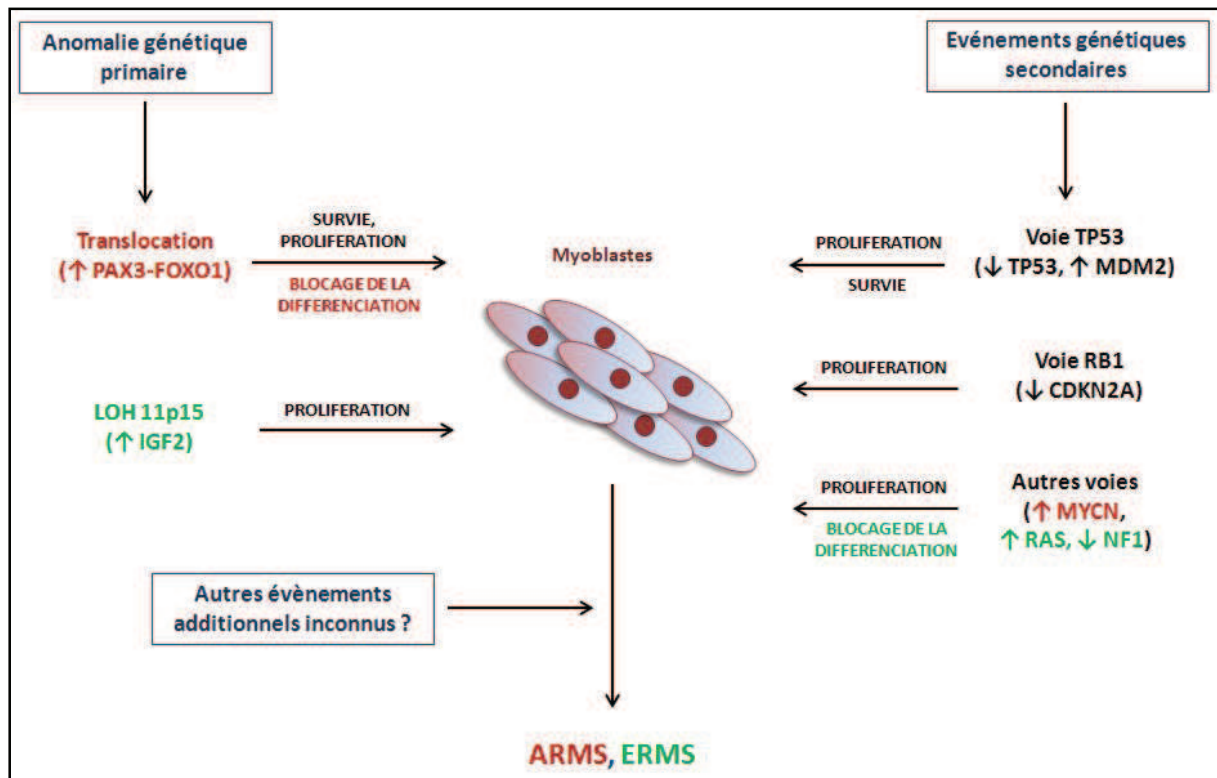


Figure 4 : résumé des événements génétiques qui se produiraient durant la myogenèse et qui pourraient mener à la formation des RMS.

Dans ce modèle, les effets induits par la présence d'un transcrite de fusion (ARMS) ou d'une LOH (ERMS), sont complétés par des événements secondaires. Les anomalies génétiques ciblant préférentiellement les ARMS sont représentées en rouge, celles ciblant les ERMS en vert.

1.6. Le diagnostic différentiel ARMS *versus* ERMS

Ces dernières années, des efforts importants ont été faits afin de distinguer de façon fiable et reproductible les ERMS et les ARMS, ces derniers étant des tumeurs beaucoup plus agressives. Alors que, dans le passé, la distinction était plutôt faite sur des critères d'architectures (par exemple, la présence de structures alvéolaires), il est aujourd'hui reconnu que les caractéristiques cytologiques sont plus fiables. En particulier, les cellules des ARMS possèdent une morphologie plus uniforme, sont beaucoup plus rondes et ont de plus gros noyaux que ceux des ERMS.

Cependant, les ERMS peuvent varier d'une présentation lâche, myxoïde et peu cellulaire à une présentation plus dense, à petites cellules rondes très indifférenciées (figure 5). Ils rappellent alors certains ARMS, en particulier les variants solides pour

lesquels les septa fibreux ont disparu. Cette similarité d'apparence peut être accentuée lors de l'analyse de petites biopsies non représentatives de la totalité de la lésion. Sans étude complémentaire en IHC ou en biologie moléculaire, ces types de lésions sont souvent impossibles à distinguer entre elles, et parfois même vis-à-vis d'autres lésions à petites cellules rondes. L'IHC peut aider à orienter le diagnostic vers l'un des deux sous-types de RMS, puisque la positivité d'AP-2 beta semble corrélée à la présence d'un ARMS à transcrit courant (marqueur actuellement à l'essai), alors que l'EGFR est souvent positif chez les ERMS. En diagnostic de routine, l'identification des translocations spécifiques des ARMS via les techniques de PCR et d'hybridation *in situ* fluorescente (FISH) permet de distinguer les ARMS des ERMS, dans ces cas où la morphologie seule n'est d'aucune aide (Mehra et al., 2008).

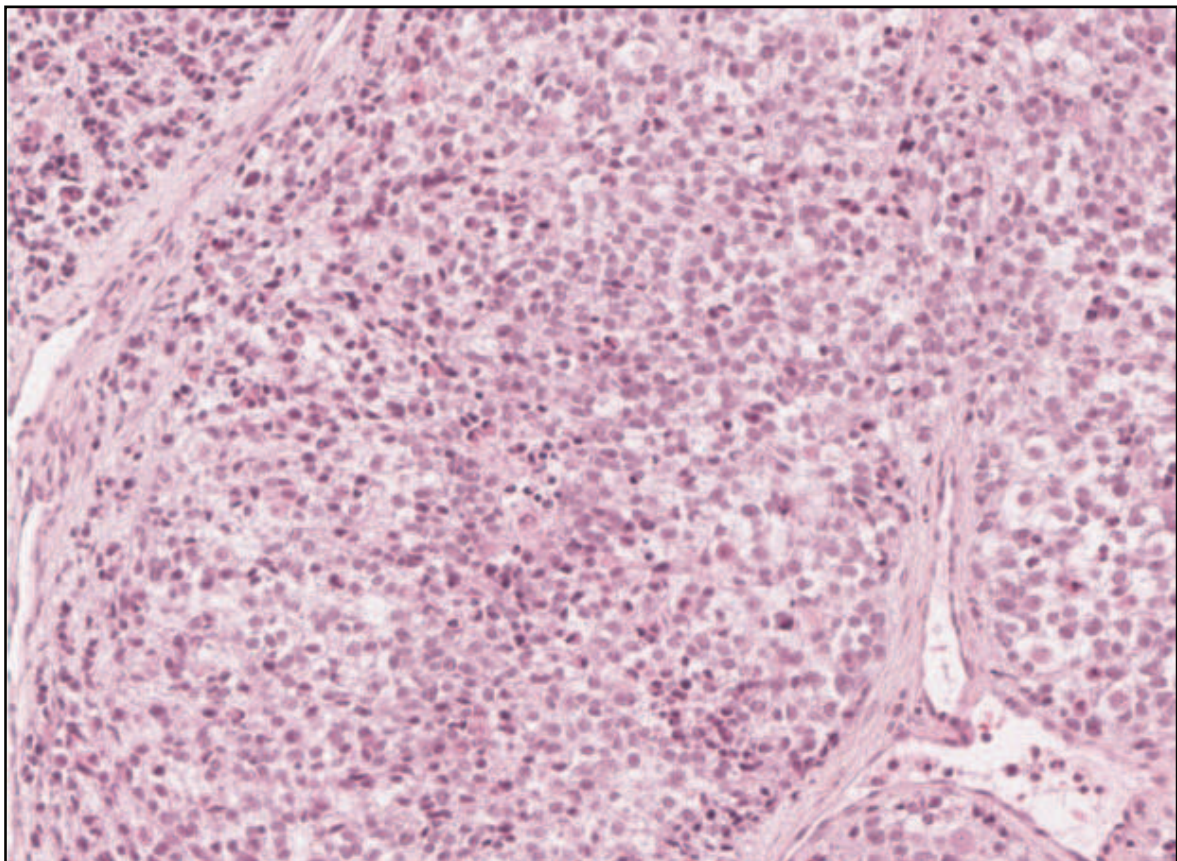


Figure 5 : coupe colorée HES d'un RMS indifférencié de la zone grise, difficile à classer.

(Image microscopique capturée par le Dr. Dominique Ranchère-Vince, Département de Biopathologie, Centre Léon Bérard, Lyon).

Le tableau 2 résume les principales caractéristiques permettant aux médecins anatomopathologistes, aux médecins cytogénéticiens et aux biologistes moléculaires de différencier ces deux entités. En ce qui concerne les RMS à petites cellules rondes très indifférenciées, il subsiste cependant une zone grise avec de nombreuses incertitudes, car certaines caractéristiques moléculaires sont difficilement détectables par des techniques de routine sur des prélèvements inclus en paraffine, tels que les transcrits rares décrits dans les ARMS ou les pertes d'hétérozygoties décrites dans les ERMS.

Caractéristiques	ERMS	ARMS
Age au diagnostic	- généralement < 5 ans - les variants à cellules fusiformes plutôt chez les adolescents	- tous les âges - plus de risque d'être un ARMS si se développe chez des patients de 5 ans et plus
Sites préférentiels	- tête et cou - abdomen - appareil urogénital - scrotum	- extrémités des membres - tronc - tête et base de crâne
Anatomopathologie	IHC : desmine +, myogénine + variable, EGFR souvent +	IHC : desmine +, myogénine ++, AP-2 beta + (transcrits courants, en cours de développement)
Cytogénétique	- perte d'hétérozygotie du bras court du chromosome 11 - anomalies de l'empreinte parentale - gains des chromosomes 2, 7, 8, 11, 12, 13q21, et 20 - pertes des zones 1p35-36.3, 6, 9q22, 14q21-32, et 17 - diploïdie ou hyperploïdie - profils d'expression hétérogènes sur puces d'expression	- translocations et transcrits de fusions : t(2;13), PAX3-FOXO1 t(1;13), PAX7-FOXO1 t(2;2), PAX3-NCOA1 t(2;8), PAX3-NCOA2 t(X;2), PAX3-FOXO4 - amplifications géniques - diploïdie ou tétraploïdie - profils d'expression homogènes sur puces d'expression (pour les tumeurs à transcrits de fusion)

Tableau 2 : tableau comparatif des principales caractéristiques biologiques des ARMS et ERMS.

1.7. Facteurs pronostiques

Les RMS présentent généralement un risque élevé de développer des métastases, l'issue du variant alvéolaire étant la plus défavorable, avec une forte propension à disséminer précocement et de manière large, une mauvaise réponse à la chimiothérapie et des rechutes rapides (Tsokos et al., 1992). Le taux de survie à 5 ans des patients atteints de RMS s'est considérablement amélioré ces dernières années, puisqu'il est passé de 55% lors des premières études de l'IRSG (Maurer et al., 1988), à une survie globale d'environ 70 à 75%, d'après les études récentes de l'IRSG et de la Société Internationale d'Oncologie Pédiatrique (SIOP) (Stevens et al., 2005a; Meza et al., 2006). La stratification en groupes de risques a été la clé de l'amélioration de la prise en charge clinique des RMS. Ainsi, les risques sont calculés en fonction de paramètres compliqués tels que l'histologie, la présence ou non de métastases, l'âge du patient, le stade, le site et la taille de la tumeur, ainsi qu'en fonction des marges chirurgicales et du bilan d'extension.

Parmi les ARMS, il existe également des phénotypes cliniques distincts associés à la nature du transcrit de fusion (Kelly et al., 1997). Il a été démontré qu'en situation métastatique, le transcrit de type PAX3-FOXO1 était de bien plus mauvais pronostic que le type PAX7-FOXO1 (Sorensen et al., 2002). En ce qui concerne les patients porteurs d'un ARMS locorégional, l'impact de la nature du transcrit de fusion sur le devenir des patients a été récemment remis en cause (Stegmaier et al., 2011), mais une autre étude ultérieure a finalement confirmé la valeur de mauvais pronostic du transcrit de fusion PAX3-FOXO1 (Missiaglia et al., 2012).

Les deux sites les plus communs pour l'émergence des métastases de RMS sont les poumons (40 à 50 % des cas) et la moëlle osseuse (20 à 30 % des cas) ; les sites les moins courants sont les os (10 %) et les ganglions lymphatiques (jusqu'à 20%, en fonction du site de la tumeur primitive) (Lawrence, Jr. et al., 1977; Ruymann et al.,

1984; Koscielniak et al., 1992). Plus rarement, des métastases peuvent être trouvées dans les organes viscéraux (cerveau ou foie), généralement chez les patients en phase terminale (Shimada et al., 1987; Wexler et al., 2006).

1.8. Traitement des RMS

Le traitement des RMS dépend de plusieurs facteurs : le stade de la tumeur, sa localisation et son sous-type histopathologique. Les ERMS et les ARMS répondent bien à la chimiothérapie, mais les rechutes sont plus fréquentes pour les sous-types alvéolaires. Les formes alvéolaires étant plus agressives que les autres, leur traitement est beaucoup plus lourd et nécessite le recours à des radiothérapies et à des chimiothérapies plus intensives. Des traitements à base de cycles alternés de chimiothérapies combinées (généralement vincristine, adriamycine et cyclophosphamide) sont utilisés dans tous les RMS, alors que la radiothérapie est principalement utilisée pour les stades II à IV des ARMS et les stades IV des ERMS.

Grâce à ces approches thérapeutiques modernes, l'issue des ARMS s'est améliorée de façon constante sur plusieurs décennies. Mais ces traitements intensifs peuvent entraîner des séquelles plus ou moins graves à long terme, surtout chez les jeunes enfants (Stevens, 2005; Mansky et al., 2007). Il est donc crucial de les administrer uniquement aux enfants présentant des formes de RMS très agressives.

La classification moléculaire est donc d'un apport essentiel pour orienter la prise en charge des patients porteurs de RMS, puisqu'elle permet d'affiner le diagnostic lorsque les critères classiques fondés sur la morphologie et l'IHC atteignent leurs limites. Afin d'apporter une aide aux médecins anatomopathologistes dans leur prise de décision, nous avons entrepris de développer deux approches différentes permettant de démembrer les RMS. La première approche a consisté à concevoir une puce à ADN ayant pour but la détection de la totalité des transcrits de fusion décrits dans les ARMS. Dans un second temps, nous avons choisi d'utiliser une technologie de PCRq à moyen débit pour établir les profils d'expression des microARN, afin d'évaluer leur potentiel pouvoir discriminant entre ARMS et ERMS. Cette même technologie permettant d'analyser les profils d'expression des microARN a également été appliquée à une autre pathologie, l'ostéosarcome, mais cette fois-ci dans un contexte thérapeutique.

2. Les ostéosarcomes

2.1. Description et épidémiologie

L'ostéosarcome (OS) est une tumeur maligne rare des tissus conjonctifs qui est caractérisée par la production d'une matrice ostéoïde par les cellules tumorales. Le terme ostéosarcome a été introduit dès 1805 par le chirurgien français Alexis Boyer (Peltier, 1993). Malgré le fait que ce cancer ait été très étudié depuis le début du XX^{ème} siècle, il est toujours caractérisé par un taux de mortalité élevé, et sa génétique est encore mal comprise.

L'incidence des OS est faible, de 2 à 4,8 cas par an et par million de personnes. Il représente moins de 0,2 % de l'ensemble des cancers (Whelan et al., 2010), mais il s'agit du cancer des os le plus courant et du troisième cancer le plus courant parmi les adolescents (Kansara and Thomas, 2007). Deux tiers des patients porteurs d'un OS sont âgés entre 10 et 25 ans (Bullough, 2004), période où la croissance osseuse est maximale. Dans un tiers des cas, les OS surviennent chez des individus de plus de 40 ans, mais ces événements surviennent généralement dans un contexte génétique particulier (Schimke et al., 1974).

Des facteurs génétiques et environnementaux jouant un rôle dans la biogenèse des OS ont été identifiés, et plusieurs syndromes familiaux sont associés à l'OS. L'exposition à des radiations ou à la radiothérapie augmente les risques de développer un OS à long terme, sur une période de plusieurs dizaines d'années (Amendola et al., 1989). Les agents de contraste utilisés au diagnostic et les agents alkylants utilisés lors des traitements pourraient également augmenter le risque de développement d'OS (Ottaviani and Jaffe, 2009).

Les patients possédant des remaniements désordonnés du tissu osseux, tels que les patients atteints de la maladie de Paget, ont un risque accru de développement d'OS. Dans ces syndromes, les patients développent généralement des OS après 40

ans. La maladie de Paget a été découverte en 1877 par Sir James Paget, et a ensuite été associée à des anomalies délétères du gène TNFRSF11A (RANK) dans la voie de signalisation du TNF/IL1, ou des gènes SQSTM1 et MAPK8, qui sont des acteurs de la voie de signalisation de TNFRSF11A (Kansara and Thomas, 2007). Les formes héréditaires de rétinoblastomes sont également fortement associées aux OS (Hansen and Cavenee, 1987). Ce syndrome est causé par une mutation germinale du gène suppresseur de tumeur RB1 (Friend et al., 1986), qui conduit à des développements de tumeurs malignes de la rétine chez des patients très jeunes, généralement avant 5 ans. Ces mêmes patients ont 2000 fois plus de risques de développer un OS du crâne en territoire irradié, et 500 fois plus de risques de développer un OS des extrémités (Wong et al., 1997). Les patients porteurs du syndrome de Li-Fraumeni (cf. introduction, biologie des rhabdomyosarcomes) ont également plus de risques de développer un OS. A noter qu'il existe encore d'autres syndromes moins courants associés à la production d'OS (Calvert et al., 2012).

2.2. Diagnostic et classification

- Présentation clinique

Les patients porteurs d'un OS présentent le plus souvent une douleur liée à une masse au niveau d'un os ou d'une articulation (Arndt and Crist, 1999). La douleur est plus forte lors d'une activité physique, et persiste la nuit. Des fractures pathologiques surviennent dans 5% des cas. Ces symptômes apparaissent généralement plusieurs mois avant le diagnostic. Ces lésions, qui peuvent être ostéolytiques et/ou sclérosantes, peuvent être détectées par radiographie (Rosenberg et al., 1995). La scintigraphie et l'imagerie par résonance magnétique (IRM) permettent d'évaluer l'extension de la tumeur et de planifier l'approche biopsique ou chirurgicale (Gibbs et al., 2002). Ces techniques sont également fondamentales pour la recherche de

métastases. Avant traitement, le diagnostic d'OS doit être confirmé au niveau histologique par une biopsie, pour le distinguer d'autres tumeurs osseuses.

L'OS peut se développer sur tous les os, mais il affecte principalement les os longs au niveau du cartilage de croissance : le fémur distal, le tibia proximal et l'humérus (Bullough, 2004).

- Anatomopathologie

Histologiquement, la lésion apparaît comme une prolifération d'ostéoblastes plus ou moins bien différenciés, produisant une matrice osseuse d'architecture anarchique, plus ou moins abondante. Cette matrice osseuse peut être isolée, ou associée à une production de matrice cartilagineuse ou fibreuse (Schajowicz et al., 1995).

Selon la classification de l'OMS, les OS sont séparés en deux groupes, les OS centraux (intra-médullaires) et les OS de surface (Fletcher et al., 2002). Plus de 75 % des cas sont des OS centraux. Ils sont situés dans le canal médullaire des os et sont caractérisés par plusieurs sous-types histologiques. Le type histologique le plus commun est l'OS conventionnel de haut grade. Parmi ceux là, le sous-type appelé OS ostéoblastique représente près de la moitié des cas, suivi des formes fibroblastiques et chondroblastiques (figure 6), représentant chacun environ un quart des cas (Cleton-Jansen et al., 2005). Les cas restants sont des variants rares incluant par exemple l'OS télangiectasique, l'OS riche en cellules géantes, l'OS ressemblant à l'ostéoblastome et l'OS ressemblant au chondroblastome. Des sous-types histologiques différents sont fréquemment retrouvés au sein d'une même tumeur, rendant une classification précise difficile.

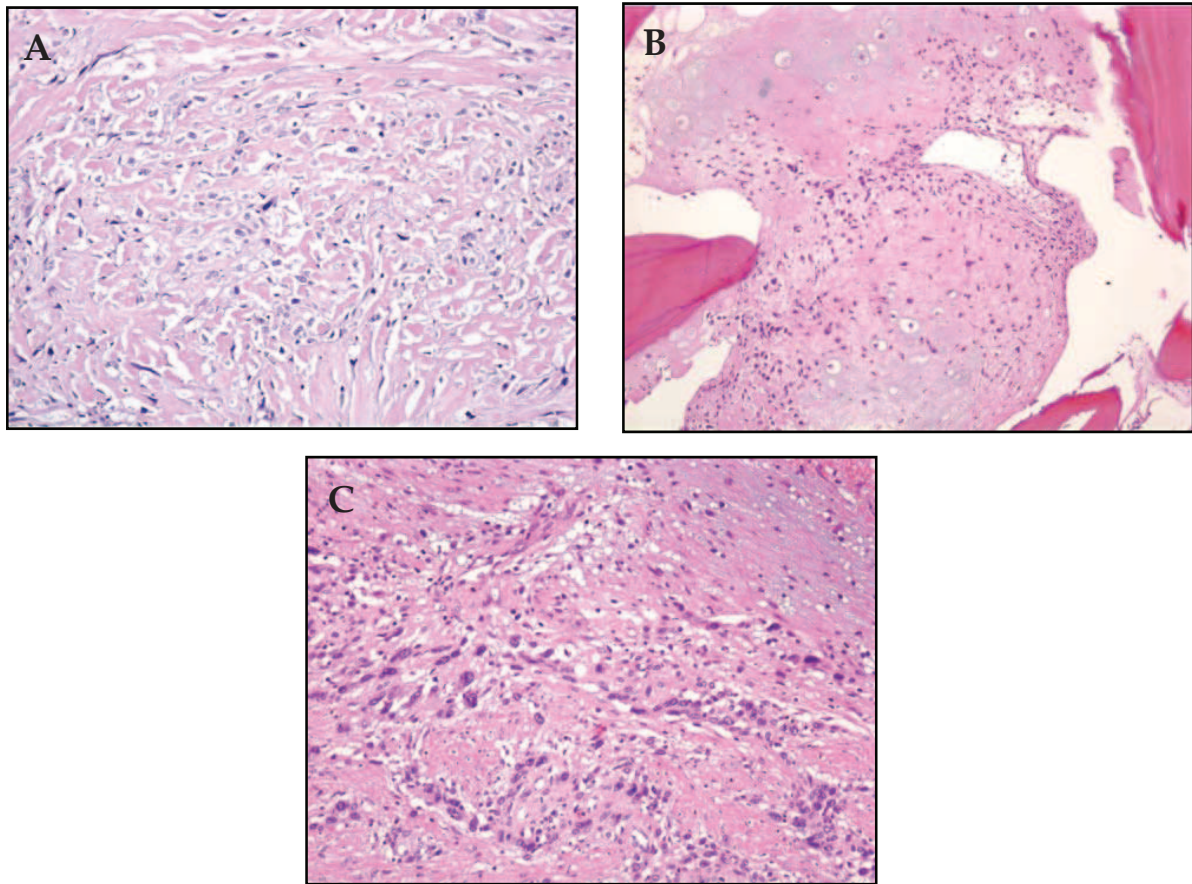


Figure 6 : coupes en paraffine colorées HES des trois formes les plus courantes d'OS.

(A) L'OS ostéoblastique, représentant environ 50 % des OS médullaires, est riche en ostéogénèse. (B) L'OS chondroblastique, représentant environ 25 % des OS médullaires. (C) L'OS fibroblastique, pauvre en matrice ostéoïde, représente environ 25 % des OS médullaires (photos microscopiques reproduites à partir de l'*Atlas of Genetics and Cytogenetics in Oncology and Haematology*, <http://atlasgeneticsoncology.org/Tumors/ConvOsteoID5344.html>).

Les lésions de surface sont moins souvent observées et comprennent deux principaux sous-types : l'OS parostéal, de bas grade de malignité, et l'OS périosté, de grade intermédiaire de malignité. A la différence des OS centraux retrouvés principalement chez les adolescents, les OS parostéaux surviennent chez les jeunes adultes (avec un pic à environ 30 ans), et sont plus communs chez les femmes (Bullough, 2004). Selon la classification de l'OMS, l'OS périosté serait plus fréquent chez les adolescents.

- Caractéristiques génétiques et moléculaires

Les ostéosarcomes ne présentent pas d'anomalie chromosomique récurrente. A la différence du rhabdomyosarcome et de nombreux autres sarcomes à génomique simple, aucune translocation n'a été identifiée. Des caryotypes anormaux avec aneuploïdies et des anomalies génétiques complexes ont été répertoriés, mais sans recoupements clairs entre les différents cas (Guillou and Aurias, 2010). Il a été suggéré qu'une déficience dans les mécanismes de stabilité chromosomique ou de réparation de l'ADN pourrait être un événement précurseur de la pathogenèse des OS, menant ensuite à ces anomalies désordonnées (Khanna, 2008).

Comme chez les RMS, les voies RB1 et TP53 seraient impliquées dans l'oncogenèse des OS (figure 7). La dérégulation de l'expression du gène RB1 a des effets prolifératifs. La perte d'expression du gène RB1 a été fréquemment décrite dans des cas sporadiques d'OS (Kansara and Thomas, 2007), par LOH dans 40 à 70 % des cas (Feugeas et al., 1996; Heinsohn et al., 2007), mais également par des réarrangements chromosomiques ou des mutations ponctuelles (Miller et al., 1996b). D'autres gènes de la voie RB1 sont également fréquemment dérégulés, tels que CDK4 et CDKN2A (Miller et al., 1996a; Nielsen et al., 1998; Yoshida et al., 2010).

La perte de fonction du gène TP53 a été décrite dans 40 à 60 % des OS de haut grade (Tsuchiya et al., 2000). Elle peut résulter d'une délétion (70 à 80 % des cas), de mutations ponctuelles (20 à 30 % des cas), parfois de réarrangements structuraux (Hoogerwerf et al., 1994). La voie du gène TP53 peut être dérégulée via d'autres gènes. Le gène MDM2, un inhibiteur de TP53, est amplifié dans les OS de manière indépendante des altérations de TP53 (Ladanyi et al., 1993; Duhamel et al., 2012).

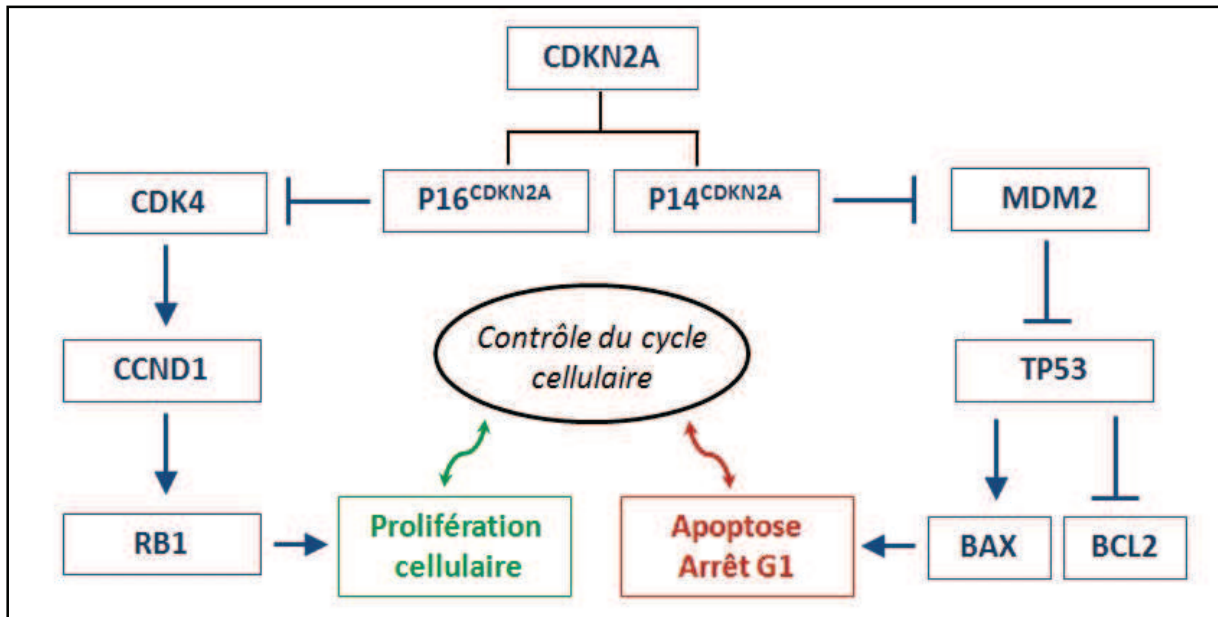


Figure 7 : schéma des voies métaboliques TP53 et RB1.

Les voies RB1 et TP53 sont reliées par le gène CDKN2A, qui code pour deux protéines différentes P16 et P14 (p16^{CDKN2A} et p14^{CDKN2A}). Ces deux protéines sont impliquées dans des cascades de régulation permettant un contrôle précis du cycle cellulaire (figure adaptée d'après (Castello et al., 2010)).

2.3. Traitement

L'ostéosarcome de haut grade et de grade intermédiaire est actuellement traité par chimiothérapie néoadjuvante et adjuvante. A des doses standard utilisées en clinique, la radiothérapie ne permet pas de stériliser ces tumeurs : elle n'est donc généralement pas prescrite en tant que traitement (Picci, 2007), même si des études récentes ont démontré de bons résultats lorsqu'elle est utilisée en combinaison avec une chimiothérapie (Mahajan et al., 2008).

- Traitement conventionnel

Avant l'introduction de la chimiothérapie dans les années 70, les OS étaient traités uniquement par chirurgie. Le taux de survie global était alors seulement de 10 à 20 % (Eilber and Rosen, 1989). Malgré l'arrivée de la chimiothérapie, l'ablation complète de la tumeur est primordiale, car seulement 10 % des patients sont guéris

par la seule chimiothérapie (Jaffe et al., 2002). Grâce aux progrès de la chirurgie orthopédique, l'amputation du membre complet peut être évitée dans la plupart des cas (Kim et al., 2010a) : des procédures de sauvetage des membres sont désormais utilisées dans 90 % des cas et n'affectent pas la survie des patients quand on pratique des marges larges. Comme les OS affectent principalement les jeunes enfants et les adolescents, ces procédures chirurgicales limitent le poids de la chirurgie à long terme et améliore grandement la qualité de vie des patients (Punzalan and Hyden, 2009). Pour les formes d'OS de bas grade, le traitement reste exclusivement chirurgical.

- Chimiothérapie

L'introduction de la chimiothérapie pré- et postopératoire comme principal traitement des OS a grandement amélioré la survie globale des patients (Philip et al., 1999). Elle était au maximum de 20 % lorsque la chirurgie était la seule prise en charge, elle est désormais d'environ 60 à 70 % (Greene, 2002). Depuis, aucun changement significatif relatif à cette survie n'a été observé. La chimiothérapie préopératoire cible les micro-métastases indétectables présentes chez la plupart des patients (Bruland et al., 2005). La chimiothérapie est également utile pour réduire la masse tumorale, facilitant sa résection et l'utilisation de chirurgies moins invasives (Gibbs et al., 2002; Kim et al., 2010a).

Les agents chimiothérapeutiques couramment utilisés pour le traitement des OS sont la doxorubicine (Cortes et al., 1974; Pratt and Shanks, 1974), le cisplatine (Baum et al., 1981; Gasparini et al., 1985), le méthotrexate (Jaffe et al., 1974; Pratt et al., 1980), et l'ifosfamide (Marti et al., 1985; Miser et al., 1987; Saleh et al., 1990). L'ifosfamide (IFO) est fréquemment utilisé en combinaison avec l'étoposide, un inhibiteur de topoisomérase II (Ferguson and Goorin, 2001).

L'IFO est un agent alkylant qui interagit directement avec l'ADN pour former des liaisons covalentes par l'intermédiaire de ses radicaux alcoyles (figure 8). Ceci aboutit à la formation de ponts alcoyles intrabrin et interbrins, et secondairement à une inhibition de la transcription et de la réplication de l'ADN aboutissant à la mort cellulaire par apoptose (Furlanut and Franceschi, 2003). Son action est cycle-dépendante, elle respecte les cellules en phase G0.

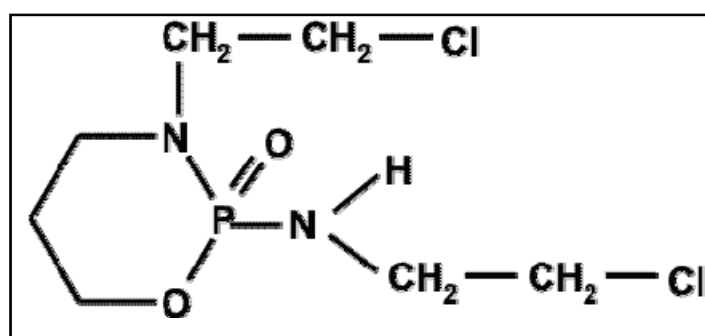


Figure 8 : formule chimique de l'ifosfamide.

Formule complète : *N*-3-bis(2-chloroéthyl)-1,3,2-oxazaphosphinan-2-amide-2-oxyde.

La réponse d'un patient à la chimiothérapie préopératoire est évaluée par un médecin anatomopathologiste, sur la pièce opératoire après résection. Moins de 30 % des patients sont considérés comme de bons répondeurs à la chimiothérapie (Provisor et al., 1997). Une bonne réponse correspond à la présence de plus de 90 % de cellules tumorales nécrotiques (Salzer-Kuntschik et al., 1983; Winkler et al., 1988). La survie à 5 ans des patients qui présentent une bonne réponse à la chimiothérapie est d'environ 75 à 80 %, à comparer aux taux de 45 à 55 % pour les mauvais répondeurs (Bielack et al., 2002).

- Nouvelles stratégies thérapeutiques

De nombreux OS étant résistants aux agents cytotoxiques classiquement utilisés lors des traitements chimiothérapiques, de nouvelles stratégies thérapeutiques ont été élaborées. Le développement de traitements antiangiogéniques est certainement l'approche la plus prometteuse. Il a été montré que l'angiogenèse était un facteur essentiel pour la croissance tumorale et l'émergence de métastases, et que la densité vasculaire tumorale pouvait être un indicateur de la malignité de la tumeur (Folkman, 1990; Gasparini and Harris, 1995). L'angiogenèse est donc rapidement devenue une cible pour les traitements anticancéreux (Sedlacek, 2001). Les OS sont des tumeurs très vascularisées, et l'angiogenèse serait impliquée dans la dissémination métastatique, notamment après résection de la tumeur primitive (Tsunemi et al., 2003; Kaya et al., 2004). De nombreuses études *in vitro* et *in vivo* en phases précliniques ont prouvé l'efficacité de la stratégie antiangiogénique dans les OS, via l'inhibition de la voie du *vascular endothelial growth factor* (VEGF) (Ando et al., 2012). Les médicaments développés peuvent être des anticorps monoclonaux humanisés dirigés contre le VEGF ou ses récepteurs VEGFR, mais également des inhibiteurs des tyrosines kinases associées au récepteur VEGF (Halimi et al., 2008). L'endostatine, une molécule naturelle dérivée du collagène, est également un inhibiteur de l'angiogenèse à large spectre, interférant avec de nombreux acteurs de la voie du VEGF (Folkman, 2006). Cet inhibiteur endogène a donné des résultats prometteurs sur des modèles *in vivo* d'ostéosarcomes chez le rat (Dutour et al., 2005).

De nombreuses études cliniques sont actuellement à l'essai, car les modalités optimales d'administration des thérapies antiangiogéniques vis-à-vis des chimiothérapies sont encore inconnues (Stempak et al., 2006). En effet, les nombreux vaisseaux sanguins irriguant la tumeur permettent également d'améliorer la distribution de la chimiothérapie aux cellules cancéreuses. Combattre ces voies d'approvisionnement pourrait donc se révéler être une action à double tranchant (Clark et al., 2008).

Les approches de thérapies géniques sont également à l'étude (Witlox et al., 2007). Elles regroupent des stratégies aussi différentes que l'utilisation d'un gène suicide, l'immuno-potentialisation et la compensation des mutations des gènes TP53 et RB1. Cependant, ces stratégies sont aujourd'hui encore limitées par des problèmes techniques liés à la toxicité des vecteurs viraux et à la difficulté de les utiliser de manière ciblée.

2.4. Risques et effets secondaires du traitement à l'ifosfamide

Comme la plupart des chimiothérapies, l'IFO possède une toxicité immédiate, provoquant vomissements, alopecies et neutropénies courtes (Ferrari et al., 2009). De plus, l'IFO possède une neurotoxicité dose-limitante, appelée encéphalopathie à ifosfamide, d'origine encore incertaine. L'incidence de cet effet secondaire, qui peut être fatal, est de 10 à 20 % après un traitement intraveineux (Pelgrims et al., 2000). Il existerait des facteurs de risque, tels que les antécédents d'atteintes neurologiques (Cain and Bender, 1995). L'IFO peut également être associée à des complications de types arythmies et arrêts cardiaques, ainsi que des cystites hémorragiques (Klastersky, 2003). En raison de la toxicité vésicale de l'IFO, il est recommandé de lui associer systématiquement un médicament uroprotecteur (Voute et al., 1996).

L'utilisation fréquente de l'étoposide en association avec l'IFO peut aussi induire des effets secondaires graves. Les doses cumulées élevées d'étoposide peuvent accroître le risque de cancers secondaires, notamment le risque de leucémie myéloïde aiguë secondaire (Shen et al., 2009). De plus, même s'il n'existe pas encore d'études ayant démontré un lien direct entre cancers secondaires et traitement à l'IFO dans les ostéosarcomes, en tant qu'agent alkylant, il pourrait également être à l'origine de cancers secondaires de type syndrome myélodysplasique (Pedersen-Bjergaard et al., 2008).

2.5. Evolution et facteurs pronostiques

L'évolution spontanée de l'OS est toujours fatale, vers une agression locale et une dissémination hématogène métastatique précoce.

- Récidives et métastases

Les OS présentent fréquemment des récurrences ou des métastases. Les métastases peuvent être précoces et présentes au moment du diagnostic, ou tardives, apparaissant des années après le traitement. Le site métastatique le plus répandu est le poumon, puisqu'il concentre environ 80 % des localisations métastatiques, mais des métastases osseuses ou dans les tissus mous sont également rapportées (Kaste et al., 1999). Environ 20 % des patients présentent des métastases distantes détectables au moment du diagnostic (Cleton-Jansen et al., 2005). Les patients en évolution métastatique ont un pronostic plus mauvais et leur survie à 5 ans est beaucoup plus faible, puisqu'elle chute à 10 % quand des métastases sont présentes au diagnostic (en comparaison aux 55 % de la maladie localisée, et aux 90 % des OS de bas grade) (Bielack et al., 2002).

- Facteurs pronostiques cliniques

A l'heure actuelle, il existe peu de facteurs pronostiques utilisés en routine pour la prise en charge des OS. La réponse du patient à la chimiothérapie a été le facteur pronostique le plus utilisé, avec la présence ou non de métastases (Philip et al., 1999). Les patients présentant une mauvaise réponse à la chimiothérapie ou des métastases disposent de peu de choix quant à un traitement additionnel. Ainsi, ces patients seraient les premiers bénéficiaires de l'identification de nouvelles cibles thérapeutiques.

Les différents sous-types histologiques montrent des réponses différentes à la chimiothérapie, et peuvent être utilisés en tant que marqueurs pronostiques, même si leur poids est limité et soumis à de grandes variabilités. Par exemple, le sous-type chondroblastique répond classiquement moins bien à la chimiothérapie (Cleton-Jansen et al., 2005). Ceci est peut-être dû à la faible cellularité de cette tumeur et à la présence d'une grande quantité de matrice chondroïde, qui pourrait gêner la distribution du médicament aux cellules cancéreuses.

Une chirurgie incomplète est un facteur pronostique péjoratif et la classification des marges chirurgicales peut être utilisée pour prédire le risque de rechute. Des marges chirurgicales larges sont recommandées pour limiter les risques de récurrence (Picci et al., 1994).

- Facteurs pronostiques moléculaires

L'expression de la P-glycoprotéine est certainement le marqueur pronostique le plus prometteur. Cette protéine est codée par le gène ABCB1, un membre de la famille des transporteurs ABC et de la sous-famille *multi-drug resistance* (MDR). Cette molécule permet aux cellules d'expulser la chimiothérapie, ce qui pourrait expliquer son association avec un mauvais pronostic dans 94 cas d'ostéosarcomes (Serra et al., 2006). Les auteurs suggèrent son utilisation pour stratifier le risque pronostique, sous réserve de validation sur une cohorte plus large.

De nombreux autres biomarqueurs sont rapportés dans la littérature comme ayant une valeur de mauvais pronostic, mais la plupart de ces études nécessiterait une validation sur de plus grandes cohortes (Clark et al., 2008). Une surexpression du gène ERBB2 serait corrélée à un mauvais pronostic (Zhou et al., 2003). Au niveau génomique, plusieurs études ont montré que la délétion ou la LOH du locus 13q14 contenant le gène RB1 était un facteur de mauvais pronostic (Ozaki et al., 2002; Patino-Garcia et al., 2003).

Comme nous l'avons vu, à l'heure actuelle, aucun marqueur moléculaire à valeur pronostique n'a été confirmé comme étant un prédicteur fort de la réponse à la chimiothérapie, ou du risque de développer une maladie métastatique. Il existe un fort besoin pour un tel biomarqueur, qui permettrait une attribution plus adaptée des traitements aux patients, tout en minimisant les risques d'effets secondaires liés aux traitements chimiothérapiques. Dans ce contexte, nous avons analysé les profils d'expression des microARN dans des OS provenant de patients bons et mauvais répondeurs par rapport au traitement à l'IFO (cf. article II) (Gougelet et al., 2011b). La réponse à la chimiothérapie néoadjuvante est actuellement évaluée à partir du pourcentage de cellules tumorales nécrotiques sur la pièce d'exérèse. S'il était possible de trouver une signature d'expression prédictive de la réponse à la chimiothérapie avant chirurgie, cela permettrait également d'orienter le geste chirurgical.

II. LES MICROARN

1. Définition et mode d'action

1.1. Définition et découverte des microARN

Les microARN sont des ARN simple brin d'une longueur de 20 à 24 nucléotides dans leur forme mature. Ils sont transcrits à partir du génome, mais leurs séquences ne sont pas traduites en protéines. Ils régulent l'expression de protéines cibles en s'appariant de manière spécifique à leurs ARN messagers correspondants, inhibant ainsi leur traduction ou entraînant leur dégradation.

Les microARN font partie de la grande famille des petits ARN, molécules régulatrices possédant des fonctions très variées chez les eucaryotes. Les petits ARN sont tous dépendants d'enzymes de type RNase III pour leur biogénèse, et s'associent avec des protéines de la famille Argonaute pour agir lors des différents mécanismes de régulation négative. Il existe deux grandes classes de petits ARN : les microARN, dérivés de précurseurs endogènes possédant une structure en tige-boucle, et les *small interfering* ARN (siARN), générés à partir de longues molécules d'ARN double brin, pouvant être endogènes ou exogènes. Ces deux classes participent au phénomène commun d'ARN interférence (ARNi).

Le phénomène d'ARNi a été observé pour la première fois chez les végétaux en 1990 (Napoli et al., 1990), mais il n'a réellement été compris et expliqué que quelques années plus tard, grâce à des travaux réalisés sur le nématode *Caenorhabditis elegans* (Fire et al., 1998). Les résultats de ces travaux de recherche ont montré que l'expression de gènes pouvait être régulée de manière spécifique et efficace par l'injection d'ARN double brin de séquence homologue au gène visé. Parallèlement à la découverte de ce mécanisme d'ARNi, d'autres travaux, toujours sur *Caenorhabditis elegans*, ont permis de mettre en évidence l'existence de petits ARN interférents endogènes : les microARN (Lee et al., 1993; Wightman et al., 1993). Ainsi, lin-4 est le

premier microARN décrit. Cet ARN fonctionnel semble être responsable de la répression traductionnelle de l'ARNm du gène *lin-14*, en s'appariant de façon imparfaite à des sites multiples situés dans sa partie 3'UTR. En 2000, une équipe découvre le microARN *let-7*, responsable de l'inhibition de l'expression de la protéine *lin-41* (Reinhart et al., 2000). *Lin-4* et *let-7* sont tous deux impliqués dans l'embryogenèse de *Caenorhabditis elegans* (Olsen and Ambros, 1999). Contrairement à *lin-4*, des homologues de *let-7* sont rapidement découverts chez les mammifères, et une structure conservée de type tige-boucle est prédite pour le transcrit précurseur (Pasquinelli et al., 2000). Le terme microARN est adopté en 2001, lorsque plusieurs groupes de recherche publient simultanément des études montrant l'existence de nombreux autres gènes de microARN conservés entre le ver et l'homme (Lee and Ambros, 2001; Lagos-Quintana et al., 2001; Lau et al., 2001).

Depuis ces études, en un peu plus d'une dizaine d'années, on observe une explosion du nombre de publications liées aux microARN, et bon nombre d'entre elles décrivent leurs contributions à l'oncogenèse (figure 9). On compte aujourd'hui plus de 2000 microARN matures identifiés, séquencés et annotés dans la version 19 de la base de données internationale des microARN nommée miRBase (Griffiths-Jones, 2004; Kozomara and Griffiths-Jones, 2011).

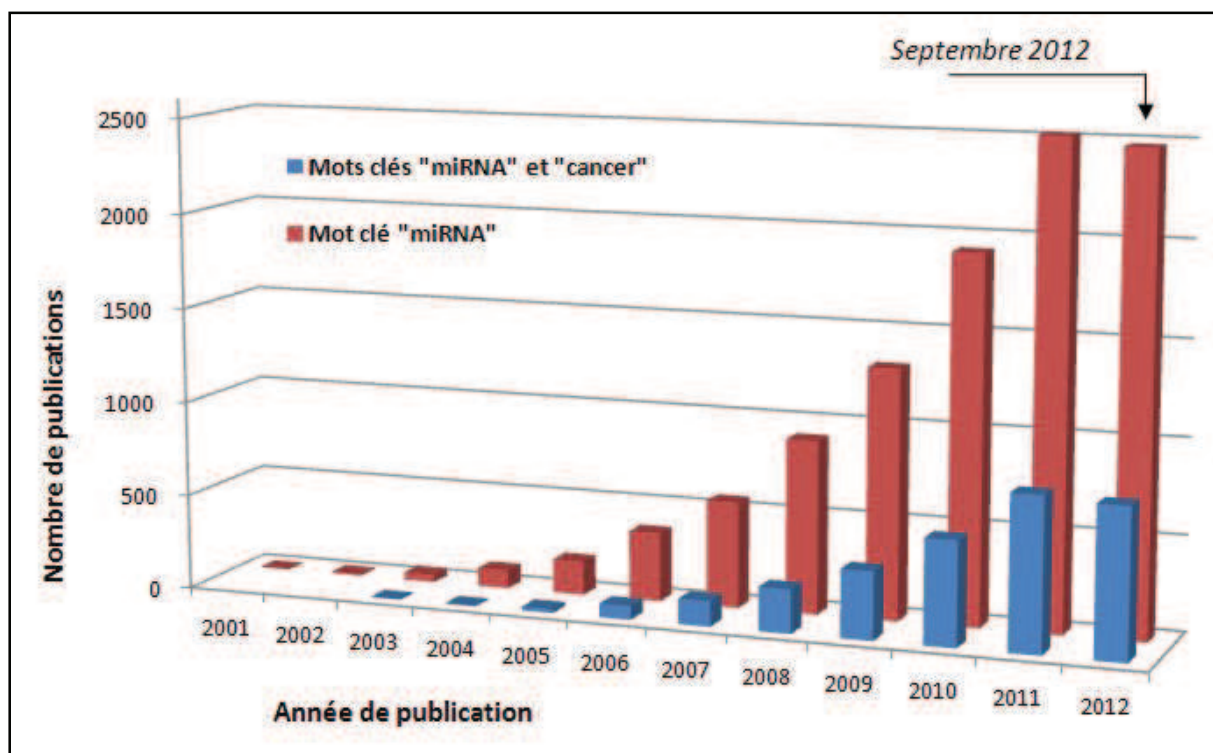


Figure 9 : évolution du nombre de publications annuelles traitant des microARN et de leurs implications dans le cancer, jusqu'en septembre 2012.

Ces histogrammes ont été générés à partir des résultats de requêtes effectuées sur le site de l'US National Library of Medicine National Institutes of Health (<http://www.ncbi.nlm.nih.gov/pubmed>).

Au sein du génome, les microARN sont codés par des unités transcriptionnelles pouvant faire partie de régions déjà codantes pour des protéines. Ainsi, environ la moitié des gènes de microARN humains seraient présents dans les introns d'autres gènes classiques, ou dans les introns de longs ARN non codants déjà connus (Rodriguez et al., 2004). L'autre moitié des gènes de microARN sont dits « intergéniques », parce qu'ils ne sont pas situés dans des unités transcriptionnelles déjà connues. Chez les mammifères, environ un tiers des gènes de microARN sont dits polycistroniques, parce qu'ils sont regroupés sur le génome et partagent un progéniteur pri-microARN commun (Lee et al., 2004) (voir paragraphe suivant).

Même si on peut penser que la grande majorité des microARN matures sont aujourd'hui identifiés, la compréhension de leur organisation génomique et de leurs mécanismes de transcription est encore à approfondir.

1.2. Biogenèse des microARN

La majorité des microARN introniques seraient co-transcrits avec leur gène hôte. Cette hypothèse a été confirmée par la fréquente corrélation entre l'expression des microARN introniques et celle de leur gène hôte (Baskerville and Bartel, 2005; Morlando et al., 2008). Cependant, environ 30 % des microARN introniques peuvent être transcrits à partir d'un promoteur indépendant du gène hôte (Corcoran et al., 2009; Monteys et al., 2010). Les microARN subissent ensuite une série d'étapes de maturation et de migration avant d'acquérir leur forme mature et active (figure 10).

Les gènes codant pour les microARN sont principalement transcrits par l'ARN polymérase II, qui produit une molécule d'ARN immature qui se replie en une structure en tige-boucle d'une centaine de nucléotides, appelée pri-microARN ou pri-pré-microARN (Lee et al., 2002). Dans le noyau, cette molécule est ensuite prise en charge par un complexe multiprotéique appelé *microprocessor*, contenant une RNase de type III, la protéine Drosha (Lee et al., 2003), et une protéine de liaison à l'ARN double brin appelée *DiGeorge syndrome critical region gene 8* (DGCR8) (Gregory et al., 2004; Han et al., 2004). Ce complexe va cliver le pri-microARN pour générer des structures tiges-boucles plus petites d'environ 70 nucléotides, qui constituent les pré-microARN. A noter que certains microARN situés dans de petits introns, appelés « mirtrons », donnent directement des pré-microARN par épissage, mécanisme qui ne nécessite pas l'intervention du *microprocessor* (Ruby et al., 2007; Schamberger et al., 2012).

Les pré-microARN générés dans le noyau sont ensuite exportés vers le cytoplasme de manière active, grâce à l'exportine-5 (XPO5) (Lund et al., 2004). Une fois dans le cytoplasme, les pré-microARN vont à nouveau être clivés par une autre RNase de type III, la protéine Dicer, assistée de la protéine TAR *RNA binding* (TRBP) (Hutvagner et al., 2001; Filipowicz et al., 2008). Ce clivage génère un duplex microARN, qui correspond au pré-microARN débarrassé de sa boucle (Zhang et al.,

2004). La protéine Dicer est aussi bien impliquée dans la voie de biogenèse de microARN que dans celle des siARN.

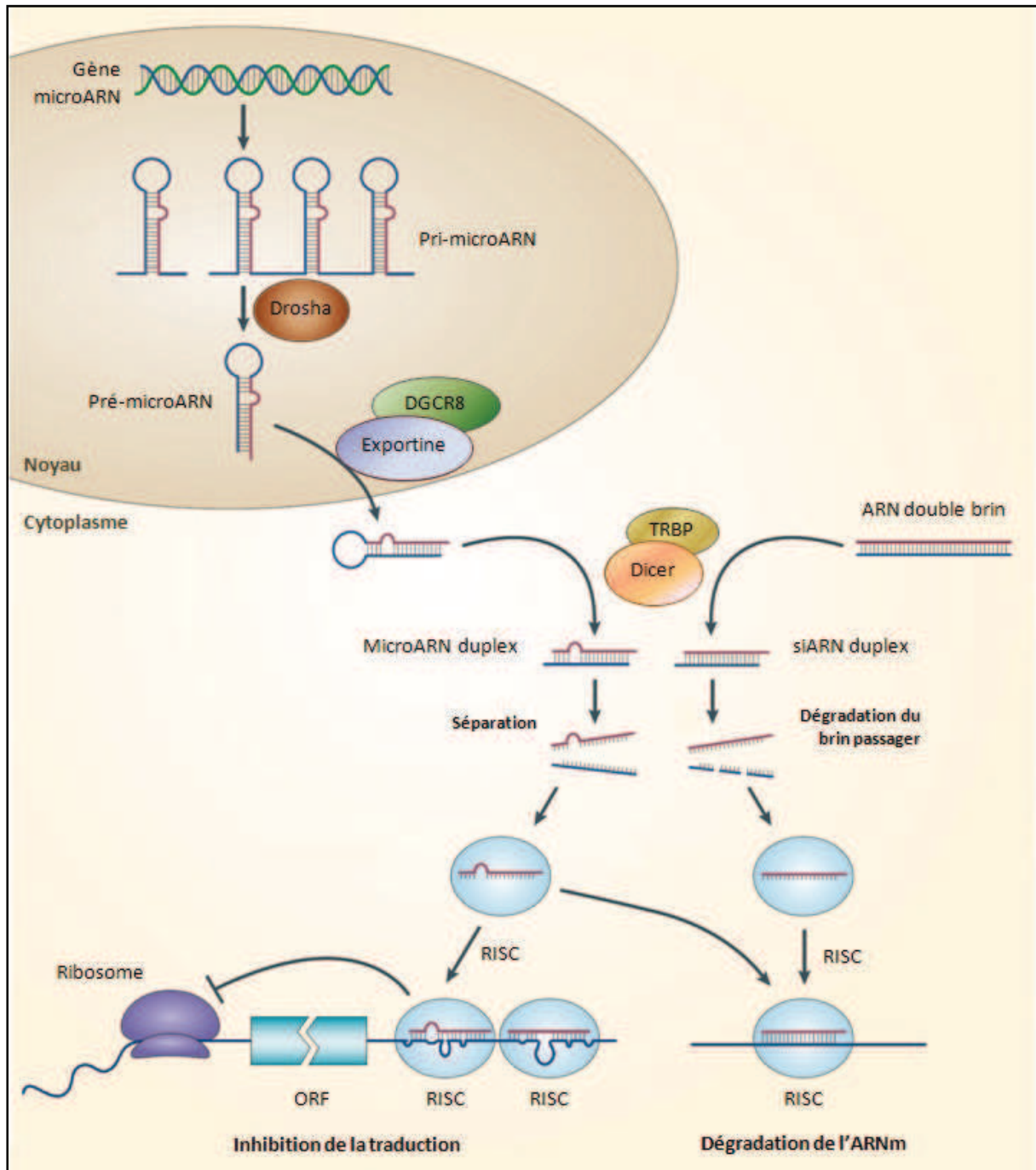


Figure 10 : voie de biogenèse des microARN chez les animaux.

Principales étapes de la voie biogenèse de microARN, et similarités avec la voie de biogenèse des siARN (figure adaptée de (He and Hannon, 2004; de Franca et al., 2010)).

Le duplex microARN est alors pris en charge par le complexe multiprotéique *RNA-induced silencing complex* (RISC) (Hutvagner and Zamore, 2002). L'un des deux brins du duplex microARN est appelé le brin « guide », et correspond au microARN mature. Ce brin est incorporé dans le complexe RISC, alors que le brin complémentaire, appelé brin « passager » ou encore miRNA*, est exclu par le complexe et éliminé. Certains microARN matures proviennent indifféremment de l'un ou l'autre des deux brins du duplex. Ces microARN matures provenant du même duplex microARN sont distingués des autres au niveau de leur nomenclature : ainsi, on retrouve les microARN notés miR-x-3p, et ceux notés miR-x-5p (respectivement pour brin en 3' et brin en 5'). Cette nomenclature s'explique par le fait qu'en conditions normales, le brin majoritairement incorporé dans le complexe RISC est le brin le plus faiblement apparié à son extrémité 5', pour des raisons de stabilité thermodynamique (Schwarz et al., 2003). La nomenclature miRNA* pour un microARN mature est utilisée lorsque que le brin traditionnellement passager est produit de manière non négligeable (Okamura et al., 2008). Les microARN matures de type miRNA* qui sont exprimés sont également actifs et induisent une répression ciblée des ARNm (Yang et al., 2011).

Le complexe RISC est caractérisé par la présence d'une protéine effectrice de la famille Argonaute. Chez les mammifères, il existe quatre protéines argonautes (AGO1 à AGO4). Toutes les protéines AGO sont capables de se lier aux petits duplex d'ARN, mais seule la protéine AGO2 possède une activité catalytique de clivage des ARNm cibles lorsqu'elle est associée à un microARN ou un siARN (Meister et al., 2004). Les complexes RISC associés à des microARN sont appelés des *miRNA-associated ribonucleoprotein complex* (miRNP) (Mourelatos et al., 2002).

Contrairement aux duplex siARN, les duplex microARN présentent des mésappariements empêchant le clivage du brin « passager » lorsque le brin « guide » est incorporé dans le complexe RISC. En effet, il a été montré qu'un mésappariement central favorisait l'incorporation du duplex microARN au sein des complexes RISC AGO1 à AGO4, et qu'un mésappariement dans les régions excentrées favorisait la

séparation des brins du duplex. En revanche, les duplex siARN ne sont maturés que par les complexes RISC contenant la protéine AGO2, grâce à son activité catalytique capable de cliver le brin passager (Yoda et al., 2010).

1.3. Principe de l'interaction entre les microARN et les ARN messagers

Les microARN sont des régulateurs spécifiques de leurs ARNm cibles grâce à un appariement du microARN mature contenu dans le complexe RISC avec une séquence partiellement complémentaire de l'ARNm cible. Cette séquence appelée MRE pour *microRNA response element* est généralement située dans la partie 3'UTR de l'ARNm cible.

La complémentarité entre le microARN et l'ARNm est imparfaite et trois principales règles ont été établies pour expliquer la répression plus ou moins forte exercée par un microARN sur sa cible (figure 11) :

- Le critère le plus important est un appariement parfait entre les nucléotides 2 et 8 du microARN. Cette séquence est appelée la région noyau. Tout mésappariement dans cette région diminue fortement l'efficacité de la répression (Doench and Sharp, 2004). En revanche, la présence d'un nucléotide A en position 1 du microARN, ou la présence d'un nucléotide A ou U en position 9 du microARN augmente l'efficacité de la répression (Lewis et al., 2005).
- Il est généralement retrouvé un mésappariement dans la région centrale du duplex microARN-ARNm. Dans ce cas, le renflement causé par ce mésappariement permet de stopper l'activité endonucléolytique d'AGO2 et de prévenir la dégradation de l'ARNm. Plus rarement, l'appariement dans cette région centrale est parfait et induit le clivage de l'ARNm (Shin et al., 2010).

➤ Enfin, la qualité de l'appariement dans la deuxième moitié du microARN du côté 3' peut être soumise à variation. Même si un bon niveau d'appariement participe à stabiliser le duplex microARN-ARNm, des mésappariements sont tolérés dans cette région, à condition que l'appariement dans la région noyau soit optimal (Brennecke et al., 2005; Grimson et al., 2007).

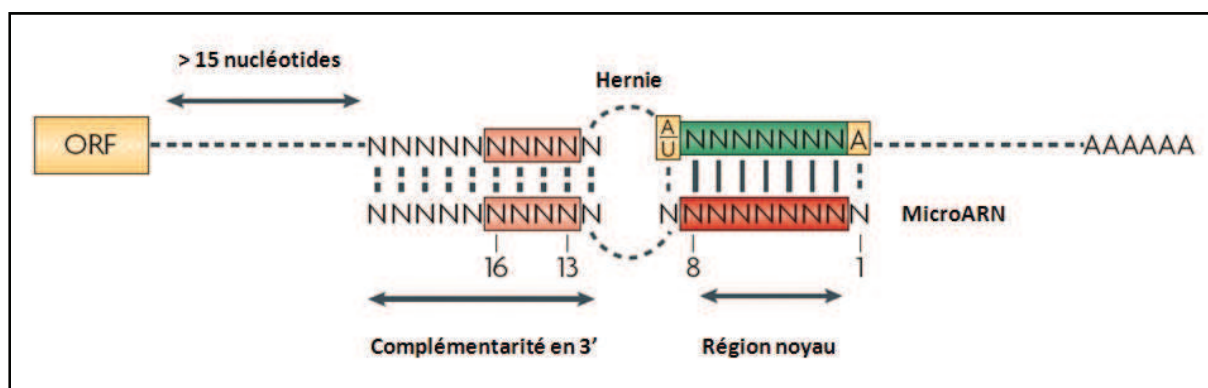


Figure 11 : principes de l'interaction microARN-ARNm.

ORF = *open reading frame* (figure adaptée de (Filipowicz et al., 2008)).

Sur la base de ces règles d'appariement, il existe de nombreux logiciels prédisant les cibles ARNm potentielles des microARN, dont les plus utilisés sont PicTar, TargetScan et miRanda (Lewis et al., 2003; John et al., 2004; Krek et al., 2005). Ces logiciels utilisant chacun des algorithmes différents, il est conseillé d'en utiliser plusieurs et de croiser les résultats obtenus pour obtenir des listes de cibles plus robustes (Witkos et al., 2011). Ainsi, la région 3'UTR d'un ARNm donné est souvent la cible de plusieurs microARN différents, et un microARN unique serait capable de réguler des centaines de cibles, ce qui fait de la voie microARN un réseau de régulations extrêmement complexes (Selbach et al., 2008). Dans tous les cas, seule une validation expérimentale *in vivo* peut confirmer qu'un ARNm est bien la cible d'un microARN.

1.4. Mécanismes post-transcriptionnels régulés par les microARN

La répression d'un ARNm via un complexe miRNP peut s'effectuer de quatre manières différentes : par inhibition de l'initiation ou de l'élongation de la traduction, par protéolyse du peptide en formation, ou par induction de la dégradation de l'ARNm cible (figure 12).

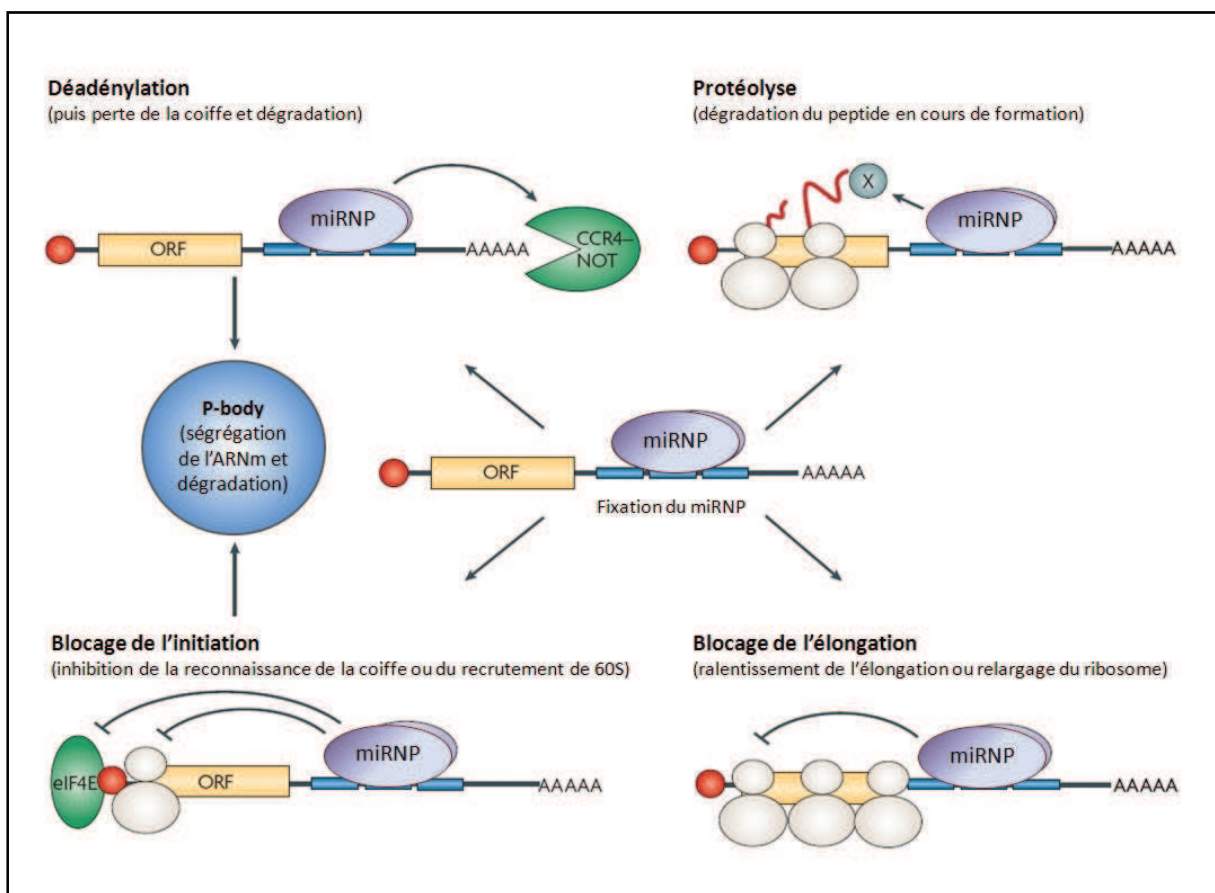


Figure 12 : modes d'action du complexe miRNP.

Le complexe miRNP possède quatre modes d'action aboutissant à la répression de l'expression de l'ARNm cible. ORF = *open reading frame* (figure adaptée d'après (Filipowicz et al., 2008)).

- Régulation de la traduction

La traduction d'un ARNm peut être inhibée par le complexe miRNP au moment de son initiation. Dans ce cas, deux principaux mécanismes sont décrits :

- L'inhibition de la reconnaissance de la coiffe est un mécanisme par lequel les protéines argonautes se fixeraient sur la coiffe par affinité, en compétition avec la protéine eIF4E, qui permet en condition normale d'initier la traduction (Standart and Jackson, 2007; Kiriakidou et al., 2007).
- L'inhibition du recrutement de la sous-unité 60S du ribosome permettrait également de bloquer l'initiation de la traduction. Ce mécanisme implique la protéine eIF6, dont la fonction première est de bloquer l'association des sous-unités ribosomales en se liant avec la sous-unité 60S. Or, il a été montré que le complexe miRNP s'associe avec cette protéine eIF6, ainsi qu'avec les protéines de la sous-unité 60S. La formation de ce complexe multi-protéique empêche donc la formation du ribosome, ce qui bloque l'initiation de la traduction (Chendrimada et al., 2007).

L'inhibition de la traduction peut également être obtenue après son initiation, par détachement des ribosomes pendant l'élongation. Il a été montré que la fixation des microARN dans la région 3'UTR des ARNm interfère à distance avec l'élongation. Le ribosome sera relargué avant même que la synthèse de la protéine soit achevée (Petersen et al., 2006).

Enfin, toujours après l'initiation, d'autres études ont montré qu'il existerait un mécanisme de dégradation des protéines en cours de formation, bien qu'il n'ait pas encore été prouvé que le complexe miRNP soit capable de recruter des enzymes de type protéase (Nottrott et al., 2006).

- Régulation de la stabilité des ARN messagers

De nombreuses études ont montré que la répression des ARNm messenger cibles pouvait également être directement liée à leur déadénylation et à leur dégradation. Les ARNm déadénylés sont en effet plus instables et vulnérables aux exonucléases (Wu et al., 2006; Eulalio et al., 2009). Cette déstabilisation fait intervenir les protéines argonautes, qui colocalisent avec de nombreuses autres enzymes de dégradation dans des compartiments cellulaires spécialisés appelés *processing-bodies* (P-bodies). Les P-bodies sont donc des lieux de séquestration et de dégradation des ARNm ciblés par les complexes miRNP (Eulalio et al., 2007a; Eulalio et al., 2007b).

2. MicroARN et cancer

Les microARN sont impliqués dans la quasi-totalité des processus biologiques. Ils jouent un rôle essentiel dans le cancer, pouvant fonctionner comme gènes suppresseurs de tumeur ou comme oncogènes (appelés alors oncomirs) (Zhang et al., 2007).

2.1. Dérégulation de l'expression des microARN dans les cancers

Les microARN régulent les voies métaboliques impliquées dans la cancérogenèse en ciblant de nombreux oncogènes et de nombreux gènes suppresseurs de tumeur. La perte d'activité ou l'augmentation d'activité d'un microARN peut donc être responsable ou participer à la croissance tumorale (figure 13) (Kong et al., 2012).

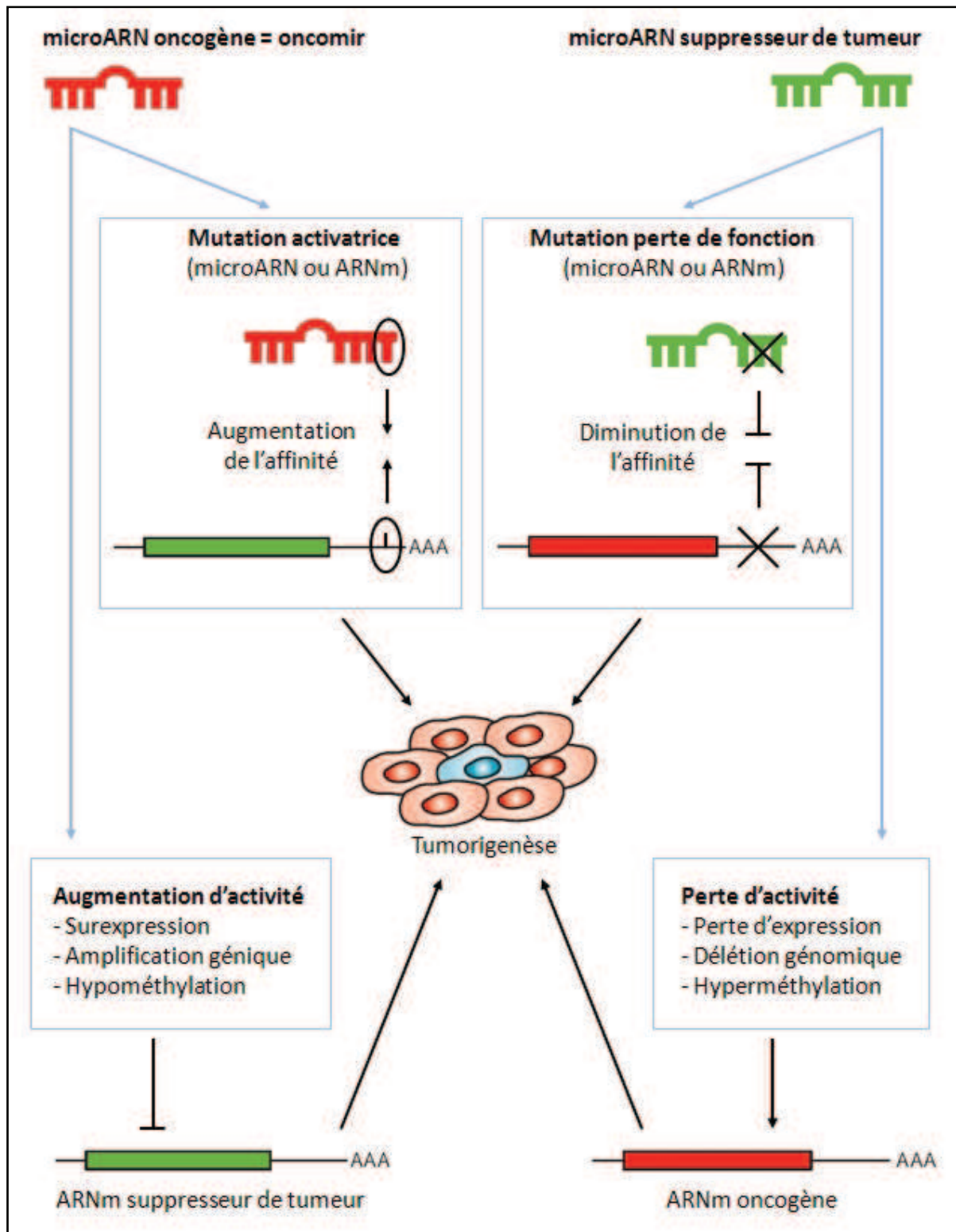


Figure 13 : les différents niveaux de régulation de la tumorigenèse par les microARN.

L'augmentation d'activité des oncomirs ou la perte d'activité des microARN suppresseurs de tumeur conduit à la tumorigenèse. La présence de mutations qui augmentent l'affinité oncomir $\leftrightarrow$ gène suppresseur de tumeur, ou qui diminuent l'affinité microARN suppresseur de tumeur $\leftrightarrow$ oncogène, favorisent également la tumorigenèse (figure adaptée d'après (Kong et al., 2012)).

La première étude rapportant le rôle de plusieurs microARN dans un cancer a montré que le cluster de gènes contenant les microARN miR-15 et miR-16 était perdu chez la plupart des patients ayant une leucémie lymphoïde chronique (Calin et al., 2002). D'autres expériences ont ensuite montré que miR-15 et miR-16 agissaient en tant que suppresseurs de tumeur en ciblant l'oncogène BCL2 (Cimmino et al., 2005). D'après une étude publiée en 2004 par le même groupe de travail, environ 50 % des gènes de microARN seraient localisés sur des régions génomiques associées au cancer ou sur des sites fragiles, suggérant que les microARN pourraient jouer un rôle important dans la pathogenèse de nombreux cancers (Calin et al., 2004). Dans le cancer du poumon, il a été montré que la famille de microARN suppresseurs de tumeur let-7 est située dans une région génomique fréquemment perdue par délétion (Boyerinas et al., 2010).

A l'inverse, certains microARN sont surexprimés dans la majorité des cancers. Le microARN miR-21 est un oncomir bien représentatif, puisqu'il est surexprimé dans le cancer du sein, le cancer colorectal, le cancer du poumon, le cancer du pancréas, le glioblastome, le neuroblastome, la leucémie et le lymphome (Chan et al., 2005; Iorio et al., 2005; Si et al., 2007; Schetter et al., 2008). La surexpression de miR-21 réprime plusieurs gènes cibles suppresseurs de tumeur tel que PTEN, et active des voies métaboliques impliquées dans la prolifération, la survie et le blocage de la différenciation (Gabriely et al., 2008; Medina et al., 2010; Darido et al., 2011).

Les cascades de régulation sont parfois plus complexes, puisque certains oncogènes régulent l'expression de microARN qui, à leur tour, vont réguler des oncogènes ou des gènes suppresseurs de tumeur. C'est par exemple le cas de l'oncogène MYC dans les lymphomes à cellules B, qui induit une surexpression de l'oncomir miR-19, qui à son tour va réprimer le gène suppresseur de tumeur PTEN (O'Donnell et al., 2005; Olive et al., 2009). En parallèle, le gène MYC inhibe l'expression de la famille des microARN let-7, dont les cibles sont les gènes de la

famille RAS (Johnson et al., 2005). Dans les RMS, il est intéressant de noter que certaines études analysant l'expression des microARN se sont intéressées aux mêmes voies métaboliques. En effet, ces études se sont focalisées sur les ARMS de type PAX7-FOXO1, présentant une amplification de la région 13q31-32 et du cluster miR-17-92, contenant en particulier l'oncomir miR-19 (Williamson et al., 2007; Reichek et al., 2011).

Des mutations au niveau des sites de reconnaissance des microARN, dans les parties 3'UTR d'oncogènes, ont également été décrites et seraient impliquées dans la tumorigenèse. Ainsi, le polymorphisme d'un seul nucléotide dans la partie 3' UTR du gène KRAS accroît significativement le risque de développer un cancer du poumon non à petites cellules (Chin et al., 2008).

En plus de ces phénomènes de régulation complexes, il semblerait que dans certains types tumoraux, le niveau d'expression d'un grand nombre de microARN soit diminué par des phénomènes touchant directement les acteurs de la biogenèse des microARN. De nombreuses études portant sur des pathologies différentes ont montré des déficits d'expression des protéines Dicer et Drosha (Merritt et al., 2008; Dedes et al., 2011; Torres et al., 2011). Ce phénomène permettrait d'expliquer la diminution généralisée du niveau d'expression des microARN rapportée dans certains types tumoraux (Ozen et al., 2008; Lin et al., 2010).

2.2. Potentiel clinique des profils d'expression des microARN

- Diagnostic et classification

De la même façon que cela a été montré pour des tissus normaux d'origines variées, les tumeurs possèdent des signatures microARN spécifiques, avec des profils d'expression spécifiques du type tumoral (Lu et al., 2005; Volinia et al., 2006). Ces observations ont également été reproduites dans les sarcomes : dans une étude portant sur 27 tumeurs variées (GIST, RMS, synovialosarcome, liposarcome et léiomyosarcome), des signatures d'expression des microARN ont été identifiées (Subramanian et al., 2008). Ces signatures sont clairement distinctes selon le type tumoral, indiquant que les microARN sélectionnés sont certainement impliqués dans la tumorigenèse et peuvent être utilisés à des fins diagnostiques. Au sein de notre équipe, nous avons pu confirmer ces résultats et affiner la classification des RMS (Gougelet et al., 2011a). Nous avons montré que parmi les ARMS, il était possible de différencier les ARMS avec transcrit de type PAX3-FOXO1, des ARMS possédant un transcrit de type PAX7-FOXO1. Nous avons également prouvé que les ARMS sans transcrit, bien que représentant un groupe hétérogène, ne se comportaient ni comme les ARMS à transcrits courants, ni comme les ERMS, suggérant la présence de probables anomalies génomiques différentes. Toujours dans les sarcomes, l'équipe de Sarver a établi une base de données en ligne des niveaux d'expression des microARN dans les sarcomes, nommée S-MED, et regroupant plus de 300 échantillons provenant de 25 types de sarcomes différents, incluant des OS et des RMS (Sarver et al., 2010).

En étudiant des tumeurs primitives variées et leurs métastases associées, il a également été montré que l'étude des profils d'expression d'un groupe de microARN pouvait permettre d'identifier le tissu d'origine d'une tumeur d'origine inconnue (Fire et al., 1998; Rosenfeld et al., 2008).

- Indicateurs pronostiques

Des niveaux d'expression anormaux de microARN peuvent avoir une valeur pronostique dans différentes pathologies (Cho, 2012). Par exemple, dans le carcinome à cellules squameuses du poumon, on a découvert qu'un niveau d'expression élevé de miR-31 était corrélé à une mauvaise survie (Tan et al., 2011). Dans le carcinome nasopharyngé, une signature d'expression de cinq microARN (miR-142-3p, miR-29c, miR-26a, miR-30e et miR-93) est significativement associée à la survie sans progression des patients (Liu et al., 2012). Les patients atteints d'épendymomes présentent également une signature d'expression de trois microARN (let-7d, miR-367 et miR-596) fortement corrélée à la survie globale, ainsi qu'un marqueur indépendant de rechute, le microARN miR-203 (Costa et al., 2011).

Les profils d'expression des microARN peuvent également être utilisés en tant que marqueurs prédictifs de l'évolution d'un cancer, permettant en particulier de prédire l'invasion et la dissémination métastatique. Dans le cancer colorectal, une signature d'expression de cinq microARN (let-7a, miR-21, miR-135a, miR-206 et miR-335) a été identifiée, permettant de détecter la présence de métastases (Vickers et al., 2012). En ce qui concerne le cancer du sein, une baisse du niveau d'expression du microARN let-7g est un marqueur de mauvais pronostic, traduisant un envahissement ganglionnaire et un faible taux de survie (Qian et al., 2011).

- MicroARN circulants : biomarqueurs non invasifs

Récemment, la découverte de la présence de microARN dans plusieurs types de fluides corporels, chez des patients sains comme chez des patients présentant des cancers, a ouvert la voie à un nouveau champ d'application (Mitchell et al., 2008; Park et al., 2009; Kosaka et al., 2010; Michael et al., 2010; Weber et al., 2010). Les microARN sont exportés activement, dans des exosomes ou simplement liés à des protéines Argonautes (Valadi et al., 2007; Arroyo et al., 2011).

Les microARN circulants sont des biomarqueurs potentiels très intéressants pour une évaluation non invasive du diagnostic et du pronostic des cancers (cf. paragraphe 2.3) (Cho, 2011). Par exemple, une faible expression de let-7b et miR-223 dans le plasma de patients porteurs d'un cancer du poumon non à petites cellules serait un marqueur de mauvais pronostic (Heegaard et al., 2012). A partir du sérum de 74 patients porteurs de cancers du poumon de types variés, il a également été montré que trois microARN (miR-155, miR-197 et miR-182) pouvaient être utilisés comme biomarqueurs diagnostiques pour la détection précoce du cancer du poumon (Zheng et al., 2011). En effet, ils étaient significativement surexprimés chez les patients porteurs d'un cancer du poumon, quelque soit son type, par rapport aux patients sains. Chez les patients porteurs d'un syndrome myélodysplasique, les niveaux d'expression dans le plasma des microARN let-7a et miR-16 étaient prédictifs de la survie globale et de la survie sans progression (Zuo et al., 2011). Dans les RMS, il a initialement été montré qu'une expression faible du microARN miR-206 était corrélée à une mauvaise survie globale dans les ERMS et les ARMS sans transcrits (Missiaglia et al., 2010). Le microARN miR-206 a ensuite été recherché avec succès dans le plasma ; son niveau d'expression chez les patients porteurs d'un RMS étant significativement différent de celui des patients sains ou de patients porteurs de tumeurs différentes, ce microARN peut donc être utilisé en tant que marqueur diagnostique et pronostique (Miyachi et al., 2010).

- Indicateurs prédictifs d'une réponse au traitement

Les microARN sont de plus en plus utilisés comme biomarqueurs permettant de prédire la réponse à un traitement anticancer. Ainsi, dans le cancer du poumon non à petites cellules, la signature d'expression des trois microARN miR-92a-2*, miR-147 et miR-574-5p est significativement associée à la chimiorésistance (Ranade et al., 2010). Dans le cancer pancréatique après résection chirurgicale, il a été montré que les patients traités à la gemcitabine, et pour lesquels l'expression de miR-142-5p était élevée, avaient des temps de survie significativement plus longs (Ohuchida et al.,

2011). Dans le cancer de l'ovaire, les tumeurs des patientes traitées au paclitaxel et ne présentant pas de réponse complète, possèdent des niveaux d'expression de miR-200c plus bas que ceux obtenus dans le cas d'une réponse complète (Leskela et al., 2011). Au sein de notre équipe, nous avons montré que les profils d'expression des microARN permettaient de prédire la réponse à l'ifosfamide chez les patients atteints d'ostéosarcomes, puisqu'une signature d'expression de cinq microARN permettait de distinguer les patients bons répondeurs des patients mauvais répondeurs (cf. article II) (Gougelet et al., 2011b).

2.3. Utilisation des microARN en tant que stratégie thérapeutique

A partir du principe général d'ARNi propre aux microARN, et grâce à la progression rapide des connaissances sur les voies microARN dérégulées dans les cancers de manière spécifique, il semble aujourd'hui envisageable de définir de nouvelles stratégies thérapeutiques inspirées des microARN. Il existe deux stratégies possibles : l'inhibition des oncomirs surexprimés grâce à l'utilisation de molécules anti-microARN (également appelées antagomir), et l'administration de microARN artificiels (appelés alors *small interfering* ARN ou siARN) en remplacement des microARN suppresseurs de tumeur dont l'activité est en défaut, ou spécifiquement conçus *de novo* pour cibler un oncogène particulier (Kong et al., 2012) (figure 14).

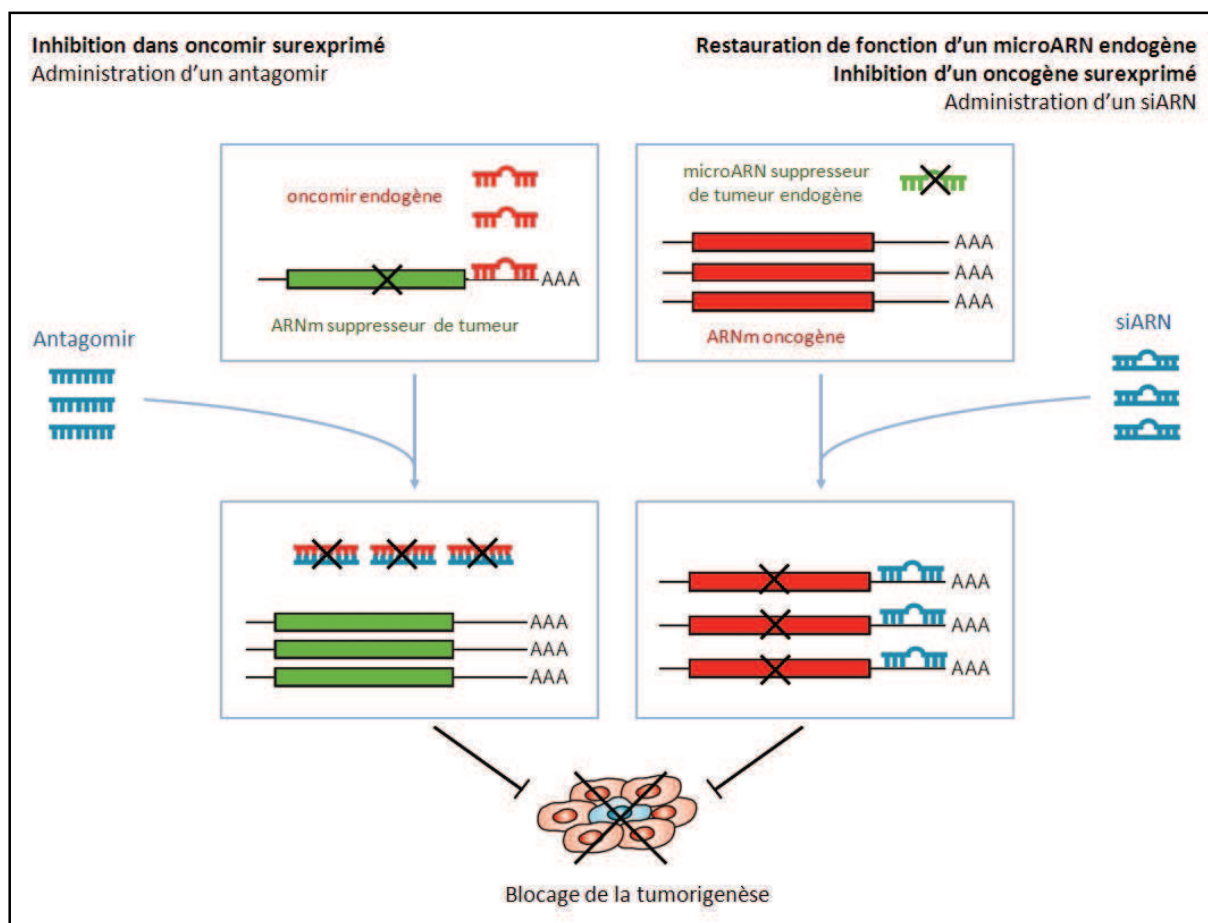


Figure 14 : principales stratégies pour les traitements basés sur l'ARNi.

Lors de la stratégie d'inhibition d'un oncomir (à gauche), des molécules simple brin appelées antagomir vont se lier de façon complémentaire au microARN ciblé, empêchant la fixation de ce dernier sur l'ARNm. Lors de la stratégie de complémentation (à droite), les microARN suppresseurs de tumeur qui sont inhibés ou perdus au cours de la tumorigenèse sont réintroduits sous la forme de siARN de même fonction, réprimant leur oncogène cible.

En 2004, seulement six ans après la découverte du phénomène d'ARNi (Fire et al., 1998), le premier médicament basé sur un siARN était développé et entré en phase I d'un essai clinique traitant des patients atteints de dégénérescence maculaire liée à l'âge (Whelan, 2005). Aujourd'hui, de nombreux essais cliniques sont en cours dans diverses pathologies, et on en dénombre environ une dizaine appliqués au cancer (Miele et al., 2012). Cependant, plusieurs essais cliniques précoces ont mis à jour des problèmes de sécurité et des difficultés d'administration pour ces médicaments basés sur des siARN (Castanotto and Rossi, 2009; Pecot et al., 2011).

Plusieurs obstacles techniques doivent d'abord être surmontés afin de pouvoir développer et démocratiser les stratégies thérapeutiques impliquant l'injection de siARN. Tout d'abord, les petits ARN ont une stabilité faible *in vivo*. Des études ont montré que des substituts de microARN introduits dans le système veineux de souris étaient éliminés de l'organisme en trente minutes (Trang et al., 2011). Pour qu'un traitement soit efficace, la durée de demi-vie des microARN artificiels *in vivo* doit être améliorée. Cela peut passer par une augmentation de la stabilité de ces siARN, ou par l'élaboration de systèmes de protection. La stabilité des siARN a été améliorée par incorporation de nucléotides de type *locked nucleic acid* (LNA) lors de leur synthèse. Cette technique est par ailleurs très répandue pour la fabrication des puces d'expression dédiées aux microARN, car elle permet d'augmenter la spécificité de l'hybridation et d'obtenir des températures d'hybridation normalisées (cf. introduction, partie III). Les nucléotides de type LNA possèdent un pont supplémentaire entre un atome d'oxygène et un atome de carbone, ce qui augmente la thermostabilité des duplex LNA-ARN, leur spécificité vis-à-vis de la cible, et leur confère une résistance accrue aux exonucléases et aux endonucléases. Ces siARN modifiés sont donc plus stables *in vitro* et *in vivo* (Vester and Wengel, 2004). D'autres pistes, encore à l'étude, consistent à substituer des groupes d'atomes dans des positions variées. Par opposition aux LNA, les nucléotides ainsi modifiés sont appelés *unlocked nucleic acid* (UNA), et confèrent aux siARN des propriétés de résistance aux nucléases (Jensen et al., 2008). Enfin, les siARN, stabilisés ou non, peuvent être encapsulés dans des structures (par exemple des vecteurs lipidiques ou des vecteurs viraux) faisant alors office de coques de protection, ou être transportés par des nanoparticules de conception beaucoup plus évoluée (Haque et al., 2012; Ali et al., 2012).

Le deuxième point important à améliorer afin de pouvoir démocratiser l'utilisation des siARN en tant que médicaments, est de pouvoir cibler avec précision le relargage des siARN aux seuls tissus tumoraux, pour minimiser d'éventuels effets

secondaires indésirables. A cette fin, la meilleure approche reste l'ajout d'anticorps ou de ligands tissus-spécifiques sur les structures utilisées pour le transport des siARN. Dans le futur, les nanoparticules contenant des siARN prendront très probablement une place prépondérante dans le traitement des cancers par ARNi, car elles semblent capables de résoudre de nombreux problèmes techniques, tels que la stabilisation du siARN, le passage des barrières biologiques, la reconnaissance du tissu cancéreux par rapport aux tissus sains environnants et le relargage des siARN transportés (Miele et al., 2012). En effet, il a été montré, au cours d'un essai clinique de phase I incluant des patients atteints de mélanomes réfractaires, que des nanoparticules contenant des siARN dirigés contre le produit du gène RRM2 étaient capables de réprimer spécifiquement ce gène, avec une action ciblée au tissu tumoral (Davis et al., 2010).

Toutes les études décrites ci-dessus permettent de restaurer l'activité d'un microARN défectif ou de réprimer un oncogène. Les stratégies basées sur l'inhibition d'un oncomir sont moins répandues. On peut toutefois citer deux études prometteuses appliquées au cancer, réalisées *in vivo* dans des modèles de souris. Dans le neuroblastome, en étudiant des tumeurs résistantes à la thérapie ou avec amplification du gène MYCN, il a été montré qu'un antagomir dirigé contre le microARN miR-17-5p permettait de réprimer ce microARN et d'inactiver la voie MYCN, induisant un arrêt du cycle cellulaire et une activation de l'apoptose (Fontana et al., 2008). Dans le cancer du sein, un antagomir dirigé contre le microARN miR-10b permet de réduire significativement la formation de métastases dans le poumon (De and Naldini, 2010).

Il est donc possible d'utiliser les profils d'expression des microARN en tant que biomarqueurs pour des applications diagnostiques, pronostiques et thérapeutiques. Les deux principales plateformes appliquées à la quantification des niveaux d'expression des microARN sont détaillées dans la partie suivante.

III. OUTILS APPLIQUES A L'ETABLISSEMENT DES PROFILS D'EXPRESSION DES MICROARN

Presque toutes les études publiées traitant des microARN dans les cancers sont fondées sur des études comparatives de profils d'expression des microARN entre cellules cancéreuses et cellules normales, ou entre cellules cancéreuses provenant de différents types ou sous-types de cancers.

Ainsi, dans cette partie, nous ne traiterons pas des techniques permettant d'étudier les profils d'expression de microARN uniques, telles que le *northern blot* ou la PCR simplex, qui ne sont quasiment plus utilisées actuellement par la communauté scientifique. Les deux plateformes technologiques les plus utilisées aujourd'hui sont les puces à oligonucléotides de type LNA, et la RT-PCRq moyen débit réalisée dans des plaques microfluidiques appelées *Taqman Low Density Array* (TLDA), toutes deux conçues spécifiquement pour les microARN.

1. Technologies pour l'étude des profils d'expression des microARN

1.1. Puces à oligonucléotides

- Définition et historique

La conception des puces à ADN a été guidée par la volonté de développer un outil permettant de quantifier de manière fiable le niveau d'expression d'un grand nombre de gènes simultanément. Dans le domaine des microARN, le développement des puces à oligonucléotides a permis de lever les principales limitations de la technique de *northern blot*, à savoir l'efficacité inégale des sondes d'hybridation, et la difficulté de pouvoir détecter plusieurs microARN simultanément (Thomson et al., 2004).

Le premier concept de puce à ADN date de 1987 (Bellis and Casellas, 1997). Des efforts ont été faits pour miniaturiser le principe du *reverse dot blot*, en faisant appel à diverses compétences telles que la robotisation, l'acquisition et l'analyse d'image

(Lennon and Lehrach, 1991). Le terme « puce à ADN » fait désormais référence aux lames de verre sur lesquelles sont ordonnés des oligonucléotides synthétisés *in situ*, mais elles sont passées par de nombreux stades « primitifs » avant amélioration et optimisation (Lagoda and Regad, 2000). Initialement, les sondes des puces à ADN de type *macroarray* étaient fabriquées à partir de clones bactériens et déposées sur des membranes de nylon. La technique a ensuite été miniaturisée avec les *microarrays*, des puces basées sur le dépôt de produits de PCR sur lame de verre. Enfin, la première puce à ADN de type oligonucléotide à être commercialisée a été la puce « *GeneChip HIV PRT* », fabriquée par la société Affymetrix (Santa Clara, CA, USA) (Vahey et al., 1999). Il existe deux méthodes principales de synthèse des oligonucléotides *in situ*, sur lame de verre. La première est basée sur un procédé de photolithographie (Fodor et al., 1993; Pease et al., 1994) ; la deuxième repose sur le dépôt de très petites sphères de fluide à la manière des imprimantes à jet d'encre (Hughes et al., 2001).

Les puces à ADN possèdent de nombreuses applications. L'application la plus répandue est certainement la mesure du niveau d'expression des gènes constituant le transcriptome, par hybridations des ARN totaux. Cependant, les puces à ADN peuvent également être utilisées pour la recherche de mutations génomiques (Ahrendt et al., 1999; Hacia, 1999), pour la détection des *single-nucleotide polymorphisms* (SNP) (Gunderson et al., 2005), pour l'étude des variations du nombre de copies des gènes (technique de *comparative genomic hybridization* ou CGH) ou encore des pertes d'hétérozygoties (LOH) (Carter, 2007).

Par ailleurs, de nombreux laboratoires ont rapidement détourné la technologie des puces à ADN dédiées à l'analyse du transcriptome, et l'ont adapté afin de pouvoir l'appliquer aux microARN (Thomson et al., 2004; Barad et al., 2004; Liang et al., 2005).

- Principe de fonctionnement des puces microARN

Toutes les puces à ADN reposent sur le même principe de fonctionnement. Les milliers d'oligonucléotides présents sur une petite surface de verre, appelés sondes, sont constitués de séquences complémentaires des acides nucléiques que l'on souhaite analyser. Chaque groupe de sondes ou *spot* est séparé des autres de quelques micromètres, et sa séquence représentative ainsi que ses coordonnées sont parfaitement connues (Ramsay, 1998). Ce micro-dispositif est ensuite mis en contact avec les cibles à analyser, c'est-à-dire les acides nucléiques dont les quantités sont à évaluer, lors de l'étape d'hybridation. Au préalable, les cibles ont été marquées par incorporation ou accrochage d'un fluorochrome, grâce à des techniques chimiques ou enzymatiques. Après hybridation et lavages des cibles non accrochées, la mesure des intensités de chaque hybride appelé *spot* est quantifiée par un scanner à fluorescence. L'intensité de chaque *spot* est directement corrélée à la quantité de cibles initialement présentes dans l'échantillon.

Les puces à ADN classiquement utilisées pour l'étude des profils d'expression des ARNm ont dû être adaptées pour pouvoir être utilisées pour l'étude des profils d'expression des microARN. Les microARN matures étant des entités d'environ 22 nucléotides, il a fallu réduire la taille des sondes de capture sans diminuer la spécificité de l'hybridation. Pour cela, les sondes des puces à ADN conçues pour la détection du niveau d'expression des microARN respectent plusieurs critères (figure 15). Les sondes sont généralement constituées de deux parties : une séquence d'accrochage et une séquence de capture. La séquence d'accrochage est souvent constituée de poly(dT) ou de poly(dA), avec une amine terminale modifiée. Cette séquence d'accrochage est fixée sur une lame de verre traitée, ce qui permet de rendre la séquence de capture la plus accessible possible pour sa cible. La séquence de capture est donc complémentaire du microARN cible. Cette séquence étant très

courte, il n'est pas possible de la choisir en fonction d'une température d'hybridation souhaitée (T_m), comme c'est le cas pour les sondes ciblant les ARNm sur les puces classiques. Pour résoudre ce problème, des sondes constituées d'oligonucléotides incorporant des nucléotides de type LNA ont été développées pour ce type de puces (Castoldi et al., 2006). Les sondes LNA ont la particularité de posséder des T_m plus élevés, puisque la stabilité des hétéroduplexes est augmentée. De plus, en jouant sur le nombre de nucléotides de type LNA incorporés dans une même sonde, il est possible d'obtenir un T_m normalisé.

Comme pour les puces d'expression classiques, les cibles microARN peuvent être marquées en fluorescence par deux techniques majeures : un marquage direct, dans lequel la molécule de microARN s'associe directement à un fluorochrome par réaction chimique, et un marquage indirect, où on marque les produits de RT ou les produits d'une RT-PCR pendant la réaction enzymatique *in vitro*, par incorporation de nucléotides fluorescents.

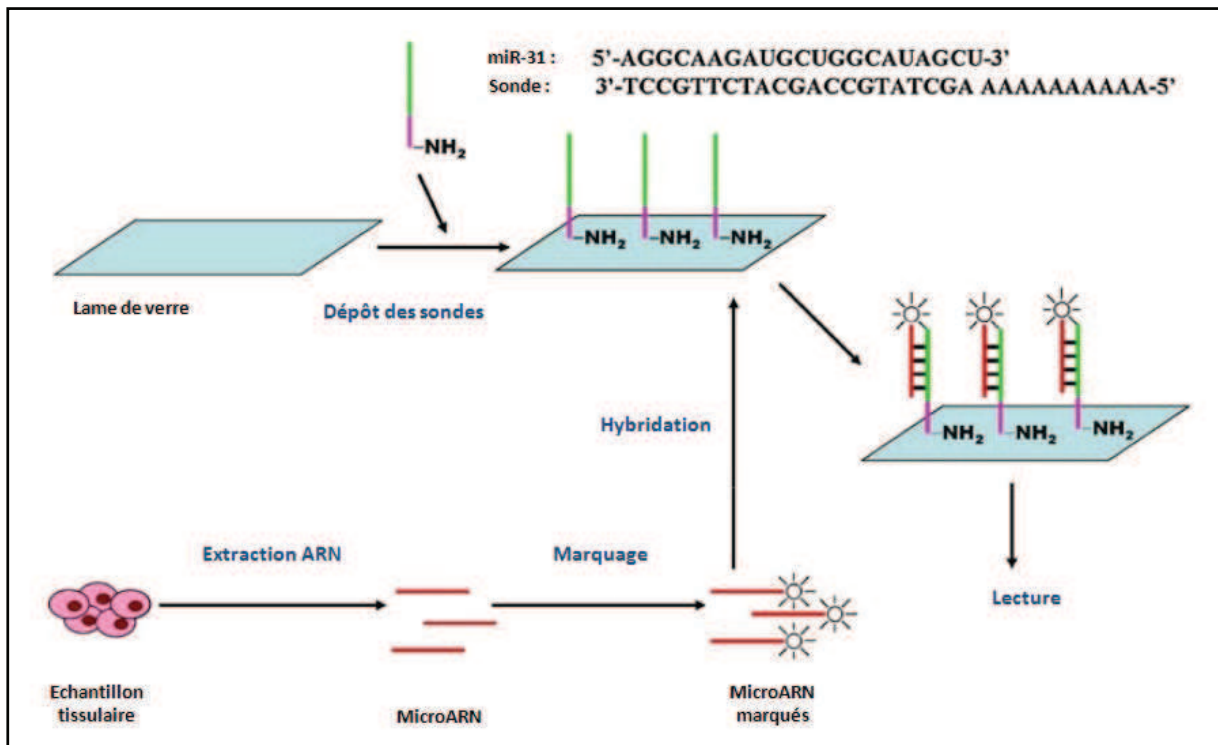


Figure 15 : principe des puces à ADN adaptées à la quantification de l'expression des microARN.

Les sondes déposées sur la lame de verre sont constituées d'une séquence d'attache (en violet) et d'une séquence de capture (en vert). Après extraction des ARN totaux, les microARN sont marqués en fluorescence par des techniques chimiques (directes) ou enzymatiques (indirectes) (figure adaptée d'après (Li and Ruan, 2009)).

Les résultats bruts sont donnés en intensité de fluorescence, dépendant du niveau d'expression des microARN mais également d'un grand nombre de paramètres expérimentaux. Ces résultats bruts ne sont donc pas directement analysables, et nécessitent une étape cruciale de normalisation.

1.2. PCR quantitative à moyen débit

- Définition et historique de la PCR quantitative

La PCRq est une amélioration de la technique bien connue de *polymerase chain reaction* (PCR) (Mullis et al., 1986). Jusqu'il y a peu, la PCRq était rarement utilisée comme outil de criblage à grande échelle, contrairement aux puces à ADN. Toutefois, sa robustesse, sa sensibilité élevée ainsi que sa grande gamme dynamique en font un outil de prédilection pour quantifier de petites variations d'expression : c'est pourquoi elle est souvent utilisée pour valider les résultats obtenus sur puces à ADN (Provenzano and Mocellin, 2007).

La PCR quantitative repose sur la quantification en temps réel de l'amplification enzymatique d'un fragment ADN, ou d'ADN complément (ADNc) obtenu à partir d'un ARN après une étape de transcription inverse (RT). Cette amplification étant basée sur une répétition d'étapes cycliques, la quantité d'ADN est évaluée au cours du temps, à chaque cycle, par quantification d'un niveau de fluorescence directement corrélé à la quantité de cible générée par la réaction d'amplification (Nolan et al., 2006).

Il existe de nombreuses méthodes pour obtenir ce signal fluorescent. Cependant, seules deux méthodes ont été appliquées avec succès pour la quantification des microARN (figure 16). La première méthode consiste à suivre en temps réel l'apparition de l'ADN cible double brin, par l'incorporation d'un marqueur non spécifique de type SYBR® Green I (Shi and Chiang, 2005; Varkonyi-Gasic et al., 2007). Ce marqueur a le désavantage de s'incorporer dans tous les fragments d'ADN double brin générés, y compris d'éventuels amplicons non spécifiques tels que les dimères d'amorces (Zipper et al., 2004). On lui préfère généralement la chimie TaqMan, basée sur un système de sonde hydrolysable couplant un fluorochrome et un *quencher*, qui absorbe la lumière émise par le fluorochrome (Kuimelis et al., 1997). Le fluorochrome et le *quencher* étant dissociés lors des étapes d'élongation de la PCRq, le signal fluorescent apparaît et peut être

quantifié. Cette méthode, bien que nécessitant plus de travail de conception, permet de s'assurer que la fluorescence mesurée est bien celle de l'amplicon recherché, grâce à la spécificité de la sonde.

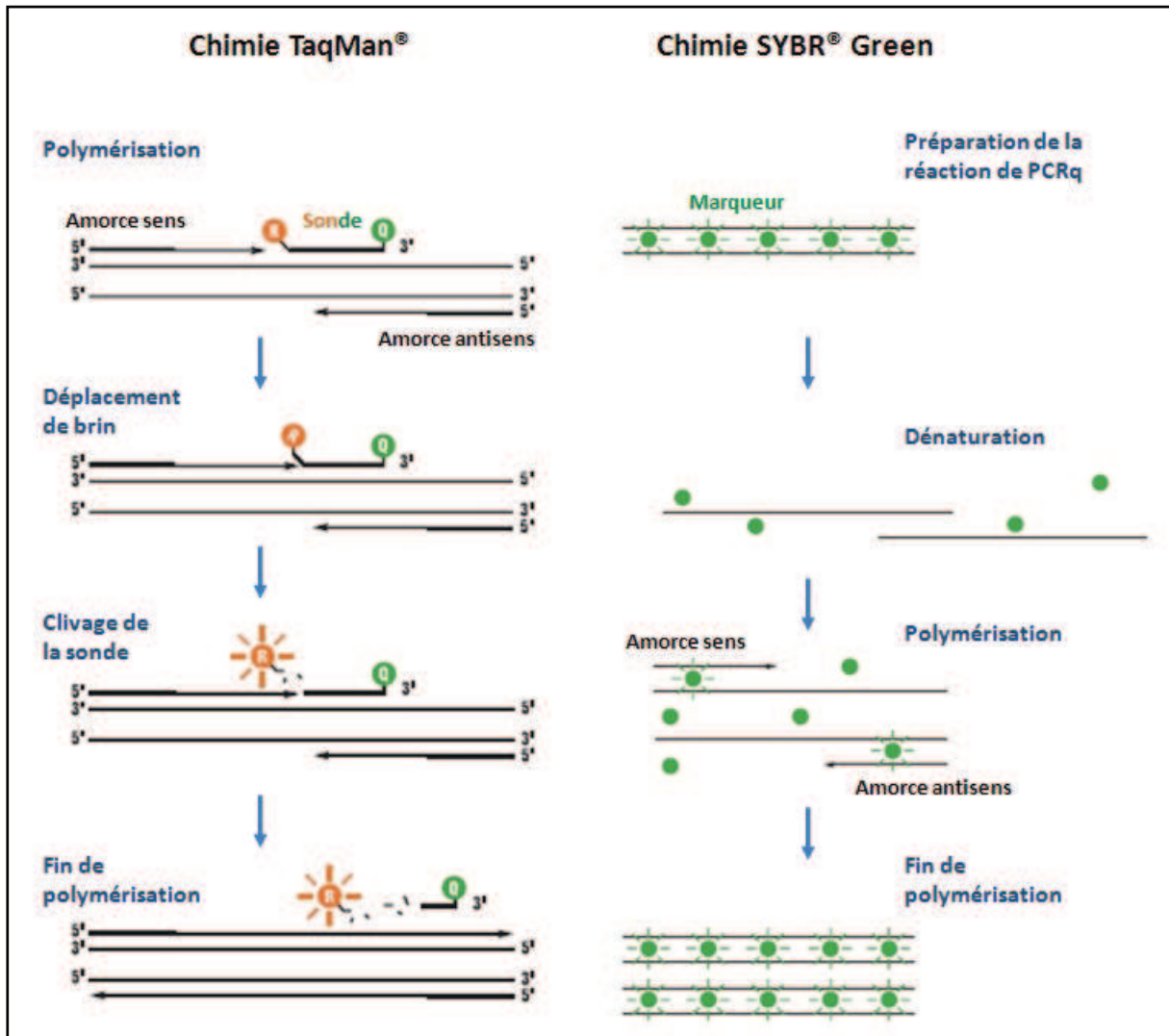


Figure 16 : principe de fonctionnement des chimies TaqMan et SYBR Green.

La sonde utilisée par la chimie TaqMan (à gauche) est hydrolysée à chaque cycle d'élongation par l'ADN polymérase, et le marqueur rapporteur émet alors de la fluorescence. La chimie SYBR Green (à droite) utilise un marqueur qui a la capacité de se lier seulement aux molécules d'ADN double-brin. Lors de sa fixation sur les amplicons en fin de chaque étape d'élongation, ce marqueur émet de la fluorescence (figure adaptée d'après le site Internet de la société *Life Technologies*, <http://fr-fr.invitrogen.com/site/fr/fr/home/Products-and-Services/Applications/PCR/real-time-pcr/qpcr-education/what-can-you-do-with-qpcr/introduction-to-gene-expression.html>).

Les résultats bruts sont donnés en cycles seuil ou Ct, correspondant au moment à partir duquel l'intensité de fluorescence dépasse un seuil très supérieur au bruit de fond, alors que l'amplification est encore en phase exponentielle. Comme les résultats bruts issus des puces à ADN, ces Ct bruts sont influencés par un grand nombre de paramètres expérimentaux et nécessitent d'être normalisés.

- Principe de la PCR quantitative appliquée aux microARN

Il a tout d'abord été montré que la PCRq pouvait être utilisée pour quantifier avec précision les profils d'expression des microARN, dans des lignées cellulaires de cancers humains (Schmittgen et al., 2004). A l'époque limitée à l'analyse de quelques microARN, cette technique a été rapidement améliorée pour l'adapter à l'analyse de plusieurs centaines de microARN par échantillon (Jiang et al., 2005). Cependant, ces premières approches se limitaient à quantifier les précurseurs des microARN. Cette limitation a rapidement été levée par l'utilisation d'amorces de RT possédant une structure de type tige-boucle (Chen et al., 2005). Ces amorces sont conçues de telle sorte qu'elles possèdent une petite partie simple brin complémentaire de l'extrémité 3' du microARN cible, et une partie double brin formant également une boucle qui contient une séquence universelle utilisable par la suite comme site de liaison d'une amorce de PCR (figure 17). Ces amorces de RT sont difficiles à concevoir, mais leur structure réduit fortement leur hybridation sur les formes immatures des microARN. Par cette méthode, la quantification de quelques microARN matures était alors fiable et précise, mais il était fastidieux de réaliser des études pangénomiques, sur l'ensemble des microARN.

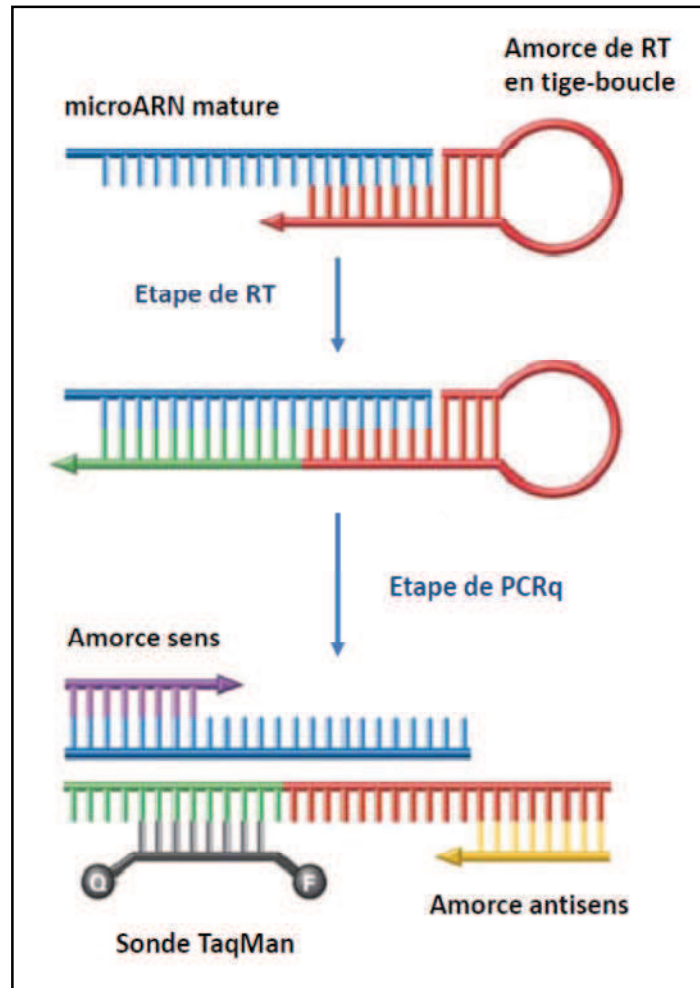


Figure 17 : principe d'une réaction de RT-PCRq adaptée aux microARN.

L'étape de RT s'effectue à l'aide d'une amorce en tige-boucle contenant une séquence universelle au sein de la boucle. L'étape de PCRq est réalisée à l'aide d'une amorce sens et d'une sonde spécifiques du microARN ciblé, et d'une amorce antisens universelle. Q = *quencher*, F = *fluorochrome* (figure adaptée d'après le site Internet de la société *Life Technologies*, <http://www.invitrogen.com/1/1/2144-taqman-microrna-cells-ct-kit-manual.html>).

Peu de temps après ces développements, une équipe de chercheurs a mis en œuvre une approche de multiplexage permettant de quantifier l'expression de 220 microARN à partir d'une cellule unique (Tang et al., 2006). En parallèle, la société *Applied Biosystems* (Foster City, CA) a développé et a commercialisé une plaque PCRq microfluidique, appelée *Taqman Low Density Array* (TLDA), basée sur le format 384 puits et personnalisable à façon pour quantifier l'expression de gènes. En 2007, la rencontre de ces deux techniques a donné naissance aux premières plaques TLDA

pour microARN (Resnick et al., 2009). Dans ce format, les plaques TLDA sont utilisées pour quantifier l'expression de nombreux microARN en une seule expérience. Chaque puits individuel contient les amorces PCR et la sonde TaqMan spécifiques d'un microARN unique, sous forme déshydratée. Après l'extraction de l'ARN total d'un échantillon à analyser, la réaction de RT est réalisée grâce à des amorces tige-boucle massivement multiplexées. Les ADNc générés sont injectés dans chacun des puits par centrifugation, grâce à un système ingénieux de micro-canaux, qui sont ensuite scellés avant réalisation de la PCRq proprement dite (voir le principe détaillé dans la partie Résultats – Introduction au travail expérimental – Outils). Aujourd'hui, un ensemble de deux plaques TLDA permet de quantifier les niveaux d'expression de 754 microARN matures, avec une conception basée sur la version 14 de la miRBase.

2. Méthodes de normalisation des données d'expression

2.1. Le besoin de normalisation

- Principes généraux

Les données brutes issues des puces à ADN ou de TLDA sont dites « bruitées ». En effet, une part non négligeable de chaque mesure est soumise à l'influence de paramètres expérimentaux qui peuvent biaiser l'analyse des résultats, si on ne les corrige pas (Schuchhardt et al., 2000; Udvardi et al., 2008). De manière générale, les erreurs de mesures peuvent être classées en deux grands groupes (Gentleman and Huber, 2007) : les erreurs systématiques et les erreurs stochastiques.

Les erreurs systématiques sont dues à des biais se produisant sur l'ensemble des données, avec une amplitude différente d'une expérience à l'autre, et pouvant parfois être variables au sein d'une même expérience, selon certaines propriétés

physiques ou chimiques liées à la nature de la cible. Dans le cas des expériences de puces microARN, ces biais sont la somme de différences liées principalement à la nature de l'échantillon analysé, à ses différents modes de conservation possibles avant analyse, au rendement de l'extraction des ARN totaux, à l'efficacité de la technique de marquage, de l'hybridation et du scan de puces. Dans le cas des expériences de TLDA, après extraction des ARN totaux, les biais peuvent dépendre de l'efficacité des réactions de RT et de PCR, ainsi des mesures de fluorescence. Ces biais sont toutefois détectables et mesurables car ils sont appliqués à l'ensemble des données : il est donc possible de les corriger.

Les erreurs stochastiques sont des erreurs non reproductibles, non prévisibles et ne touchant qu'une partie des cibles analysées. Ces erreurs de mesure peuvent être par exemple liées à la variabilité inter-individus, ou à la dégradation partielle de certains microARN. Dans le cadre d'expériences d'hybridation sur puces à ADN, ces erreurs peuvent être liées à des anomalies locales de l'hybridation : présence d'une poussière gênant la prise de mesure de la fluorescence, ou problème lors de la fabrication de la puce. Dans le cadre des expériences basées sur les TLDA, ces erreurs peuvent être liées à des problèmes lors de la fabrication des plaques TLDA, au moment du dépôt des sondes et des amorces déshydratées dans chaque puits, ou à une fuite dans l'un des puits réactionnels lors de la réaction de PCRq. La seule façon d'éviter ces biais est de répéter les expériences portant sur un même échantillon un certain nombre de fois, et d'analyser les coefficients de variations des valeurs mesurées. Cependant, les techniques de puces et de TLDA étant coûteuses, et les échantillons analysés étant souvent uniques et précieux, ces répétitions ne sont généralement pas réalisées en cancérologie.

Les données brutes d'expression des microARN doivent impérativement être décortiquées finement avant toute analyse ou interprétation, à l'aide d'outils permettant de visualiser la nature des biais engendrés par les erreurs précédemment décrites. Ainsi, en fonction de la nature des biais retrouvés, il sera possible d'appliquer une méthode de normalisation optimale à l'ensemble des données. La plupart de ces méthodes d'analyse des biais, ainsi que les méthodes de normalisation, ont été développées pour les données issues de puces à ADN. Dans les paragraphes suivants, pour présenter les grands principes de ces approches, nous prendrons des exemples issus des plateformes de puces d'expression, même si nous avons prouvé que ces approches sont également applicables aux données issues de plaques TLDA (cf. article I, en soumission au journal *RNA*).

- Détection du type de biais

Afin d'effectuer un contrôle qualité initial des données brutes et de déterminer quelle méthode de normalisation est la plus adaptée à un set de données, il est indispensable de réaliser des représentations graphiques informatives. Ces représentations peuvent également être utilisées, par la suite, pour visualiser les résultats des différentes étapes de normalisation, et vérifier que les modifications apportées sont adéquates.

Les données d'expressions des microARN obtenues à partir de puces à ADN sont généralement visualisées grâce à un graphique de corrélation, c'est-à-dire à un nuage de points en logarithme base 2 des intensités d'un échantillon contre celles d'un autre échantillon. La transformation en valeurs logarithmiques permet d'étirer le nuage de points, et de mieux visualiser l'aspect général de ce dernier. Ainsi, les faibles valeurs sont mises en valeur, alors que la contribution des valeurs extrêmes diminue (figure 18).

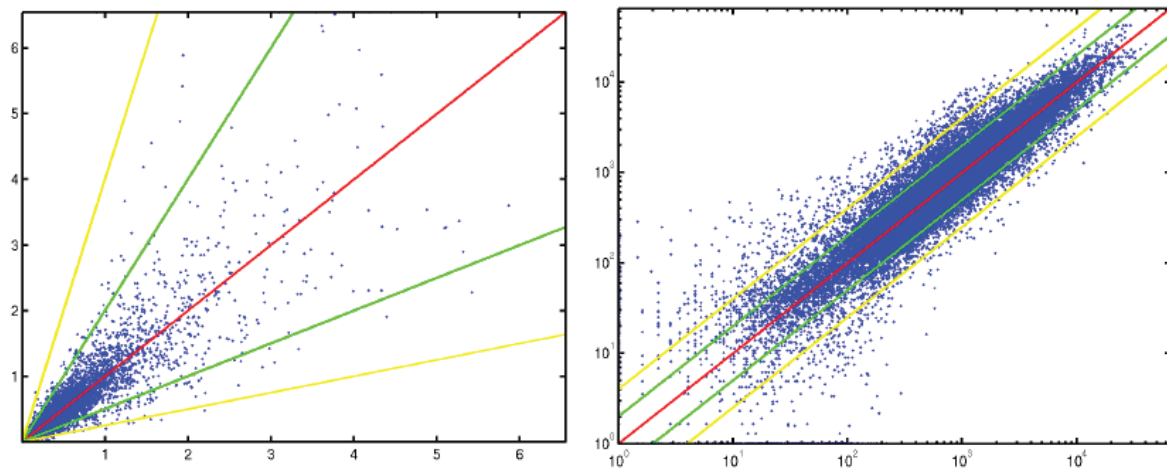


Figure 18 : exemples de graphiques de corrélation démontrant l'importance de la transformation des valeurs brutes d'intensité en valeurs logarithmiques.

A gauche, nuage de points représentant les valeurs d'intensité brutes. A droite, ces mêmes données en logarithme base 2 (figure adaptée d'après (Beißbarth, 2005)).

Il est cependant préférable de réaliser un nuage de points du ratio des intensités contre la valeur moyenne des intensités. Cette représentation appelée *plot AM* se calcule de la manière suivante (Dudoit S. et al., 2002) :

$$M = \log_2 (\text{Intensité_échantillon_1} / \text{Intensité_échantillon_2})$$

$$A = \log_2 (\sqrt{(\text{Intensité_échantillon_1} \times \text{Intensité_échantillon_2})})$$

Cette représentation correspond en fait à une rotation horaire de 45° du système de coordonnées précédemment décrit. Elle permet de mettre plus facilement en évidence la présence d'éventuels artefacts, ou un biais local, dépendant des intensités (figure 19).

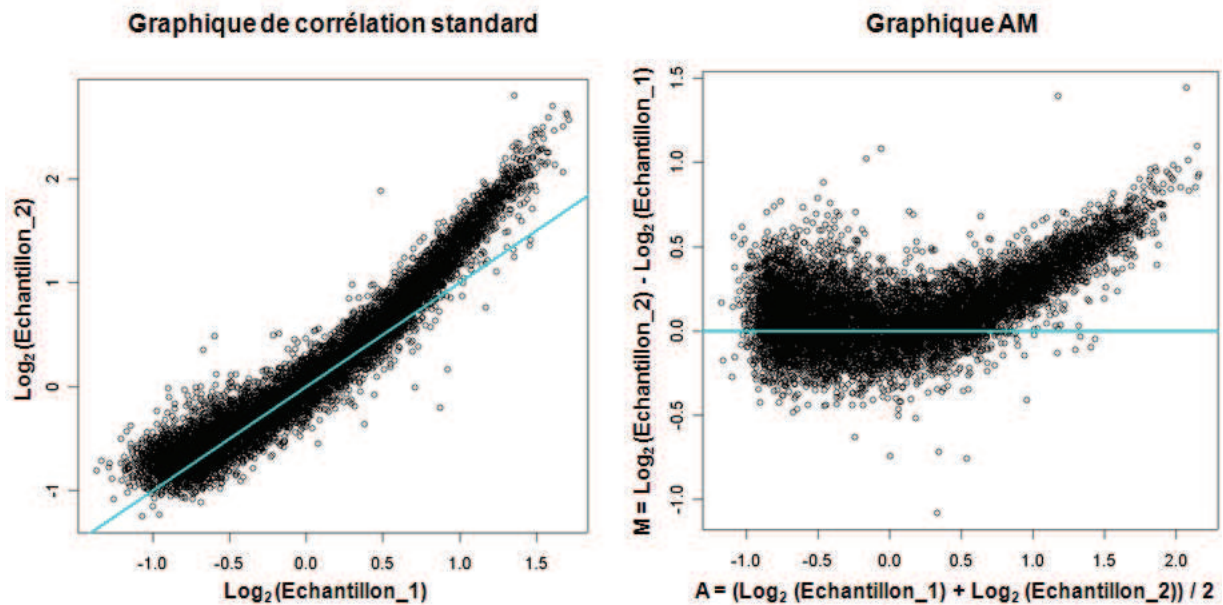


Figure 19 : exemples de graphiques démontrant l'utilité de la représentation AM.

A gauche, représentation graphique classique ou *plot* de corrélation, représentant les valeurs d'intensité brutes. A droite, graphique AM correspondant à une rotation horaire de 45° par rapport au graphique précédent. Le nuage de points possédant une forme de virgule, il existe un biais dépendant des intensités (figure adaptée d'après (Nettleton, 2009)).

On parle de biais global lorsque le nuage de points n'est pas centré sur la droite des ratios nuls, mais que sa forme reste droite et horizontale. Dans d'autres cas, le nuage de points peut suivre une diagonale, ou même posséder une forme courbe. Cela signifie que le biais est fonction des intensités ou dépendant de A, on parle alors de biais local. En d'autres termes, le biais est plus ou moins fort selon qu'il touche une valeur de faible intensité ou une valeur de forte intensité (Park et al., 2003).

2.2. Méthodes de normalisation globales

Certaines méthodes de normalisations sont basées sur l'utilisation d'un ou plusieurs gènes microARN dont l'expression est constante, et qui peuvent être considérés comme des gènes microARN de ménage. Cependant, lorsqu'on utilise des techniques dites à moyen ou haut-débit, générant un très grand nombre de données, il est souvent plus efficace d'utiliser l'ensemble des microARN ciblés à des fins de normalisation (Deo et al., 2011). En effet, dans le cas d'analyses pangénomiques (par opposition à des analyses orientées, ciblées sur certaines voies métaboliques), on s'attend toujours à trouver une symétrie dans les niveaux d'expression entre les microARN sur ou sous-exprimés.

Les méthodes de normalisation globales partent du postulat que la valeur moyenne des Ct d'une plaque TLDA doit théoriquement être égale à celle d'une autre plaque TLDA, ou que la valeur moyenne des intensités d'une puce à ADN doit être égale à celle d'une autre puce à ADN.

- Médiane ou moyenne

La méthode de normalisation par la médiane ou la moyenne est traditionnellement appliquée aux puces à ADN lorsque le biais observé sur les données brutes est global. Nous avons également montré que cette méthode pouvait être appliquée aux données brutes issues de plaques TLDA (cf. article I, en soumission au journal *RNA*).

Un facteur de normalisation k peut être déterminé de la manière suivante :

$k = \text{moyenne (Intensité_échantillon_1)} / \text{moyenne (Intensité_échantillon_2)}$, ou

$k = \text{médiane (Intensité_échantillon_1)} / \text{médiane (Intensité_échantillon_2)}$

On préfère le plus souvent utiliser la médiane, car cette fonction est moins influencée par les valeurs extrêmes que la moyenne.

Sur une représentation AM, cette normalisation a pour effet de ramener le nuage de points sur la ligne $M = 0$, c'est-à-dire de le centrer sur les ratios nuls (figure 20). Il s'agit d'une translation, la forme du nuage de points n'est pas affectée. Ce type de normalisation globale ne peut donc pas corriger un biais local, puisque le même facteur de correction est appliqué à l'ensemble des données, quelles que soient leurs intensités moyennes.

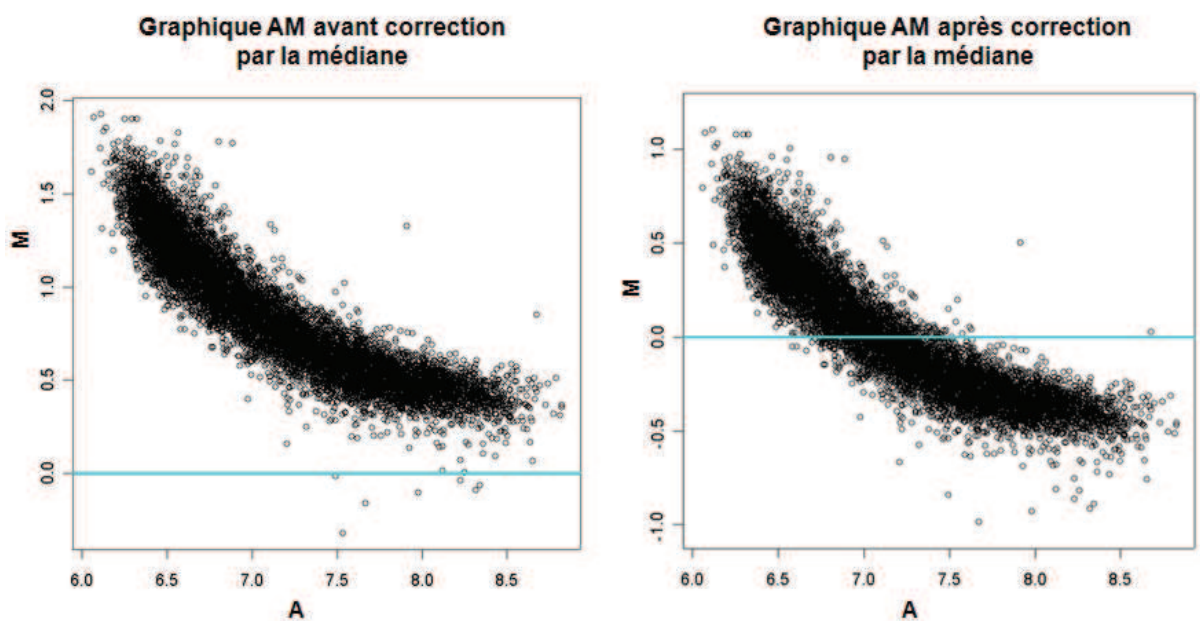


Figure 20 : exemples de graphiques AM montrant les effets d'une méthode de normalisation globale par la médiane.

A gauche, nuage de points avant normalisation par la médiane. A droite, nuage de points après normalisation par la médiane. Les données sont bien centrées sur la droite $M=0$, mais le biais dépendant des intensités est toujours présent puisque le nuage reste déformé (figure adaptée d'après (Nettleton, 2009)).

- Méthode du delta-delta Ct et choix de l'endogène

La méthode du delta-delta Ct ($\Delta\Delta Ct$) est une méthode de quantification relative, développée pour normaliser différentes expériences de PCRq (Livak and Schmittgen, 2001). Même si le terme n'est généralement pas employé, il s'agit d'une méthode de normalisation globale. Elle permet de comparer les différences de niveaux d'expression d'un gène entre plusieurs échantillons ou plusieurs conditions expérimentales. Les variations sont exprimées en rapport de niveau d'expression (*fold-change* ou *relative quantity*, *RQ*) entre les échantillons tests et un échantillon de référence, appelé calibrateur.

Le principe de cette méthode de normalisation repose entièrement sur l'utilisation de gènes de références, appelés également gènes de ménage endogènes, qui sont indispensables au fonctionnement de leur cellule d'origine, et donc censés être exprimés de manière très stable. Un gène de ménage parfait ne devrait pas voir son niveau d'expression varier entre deux expériences de PCRq. Dans la pratique, ce n'est jamais le cas, et on en déduit que la variation observée traduit les erreurs systématiques liées aux manipulations. La correction de ces erreurs est appliquée à l'ensemble des gènes cibles grâce au calcul suivant :

- pour chaque échantillon testé et pour l'échantillon calibrateur,

$$\Delta Ct (\text{gène cible}) = Ct (\text{gène cible}) - Ct (\text{gène de référence})$$

- puis, pour chaque gène cible,

$$\Delta\Delta Ct = \Delta Ct (\text{échantillon testé}) - \Delta Ct (\text{échantillon calibrateur})$$

- enfin, si les efficacités des réactions de PCRq pour chaque gène sont de 100 %, c'est-à-dire s'il y a bien un doublement de la quantité d'ADN amplifié à chaque cycle, alors le rapport des niveaux d'expression est donné par la formule suivante :

$$RQ = 2^{-\Delta\Delta Ct}$$

Ainsi, un RQ de 2 signifie que la cible est deux fois plus exprimée dans l'échantillon test que dans l'échantillon calibrateur, et un RQ de -0,5 signifie que la cible est deux fois moins exprimée. Si les résultats de PCRq doivent ensuite être injectés dans des logiciels d'analyse statistique de type clustering, il est préférable de travailler en $\log_2(\text{RQ})$. En effet, on obtient par cette transformation des valeurs symétriques autour de zéro : une cible deux fois plus exprimée donne un $\log_2(\text{RQ}) = 1$ et une cible deux fois moins exprimée un $\log_2(\text{RQ}) = -1$.

Le choix du gène de référence à utiliser pour le calcul du ΔCt est l'objet de nombreux débats. D'après plusieurs études, le gène de référence universel unique n'existe pas, et il convient plutôt de déterminer le (ou les) gène(s) dont l'expression est la plus stable dans un contexte donné (Meyer et al., 2010). A cette fin, de nombreux outils ont été développés, mais seulement deux approches ressortent du lot, GeNorm et Normfinder (Vandesompele et al., 2002; Andersen et al., 2004). Ces deux approches ne sont bien sûr utiles que si on dispose d'un nombre de gènes cibles important, et sont donc particulièrement adaptées aux données de PCRq générées par les plaques TLDA. Il est ainsi possible d'obtenir un classement des gènes dont l'expression est la plus stable entre les différents échantillons testés. La normalisation s'effectue ensuite sur un gène de référence virtuel correspondant à la moyenne géométrique des gènes stables sélectionnés.

2.3. Méthodes de normalisation locales

La plupart des méthodes de normalisation locales ont été développées pour les puces à ADN. Elles avaient pour but de corriger les biais intensités-dépendants fréquemment observés. Historiquement, les données générées par des expériences de PCRq étant peu nombreuses (techniques à bas débit), ces approches de normalisation ne pouvaient pas être appliquées, car elles nécessitent un grand nombre de données

pour être valides. Avec l'avènement de la plateforme TLDA, ces méthodes ont été progressivement testées et parfois transférées pour une utilisation sur des données de PCRq (cf. article I, en soumission au journal *RNA*) (Deo et al., 2011; Meyer et al., 2012).

- Normalisation LOESS

L'acronyme LOESS signifie *locally estimated scatterplot smoothing*. Il s'agit d'une méthode de normalisation dérivée de la méthode originale LOWESS (Cleveland WS., 1979; Cleveland WS. and Devlin SJ., 1988). Elle est très employée pour la correction des biais survenant lors d'expériences d'hybridation sur puce (Yang et al., 2002).

Dans de nombreux cas, les biais entre expériences semblent être dépendants des intensités, comme le révèlent par exemple les nuages de points des figures 19 et 20. Dans ces cas précis, une méthode de normalisation en fonction des intensités ou dépendante de A est préférable aux méthodes globales. La fonction LOESS permet un lissage des formes des nuages de points à partir d'une courbe de régression locale. La normalisation proprement dite est effectuée en ramenant la courbe LOESS sur la droite des ratios nuls (figure 21). Une des forces de la fonction LOESS est sa capacité à pouvoir travailler sur des sets de données de tailles très variables. En effet, il est possible de lui définir un paramètre de lissage, qui permet à la fonction de ne pas être affectée par la présence de quelques points aberrants, ou qui lui permet de s'ajuster au nuage de points de manière un peu moins affirmée lorsque les données sont peu nombreuses.

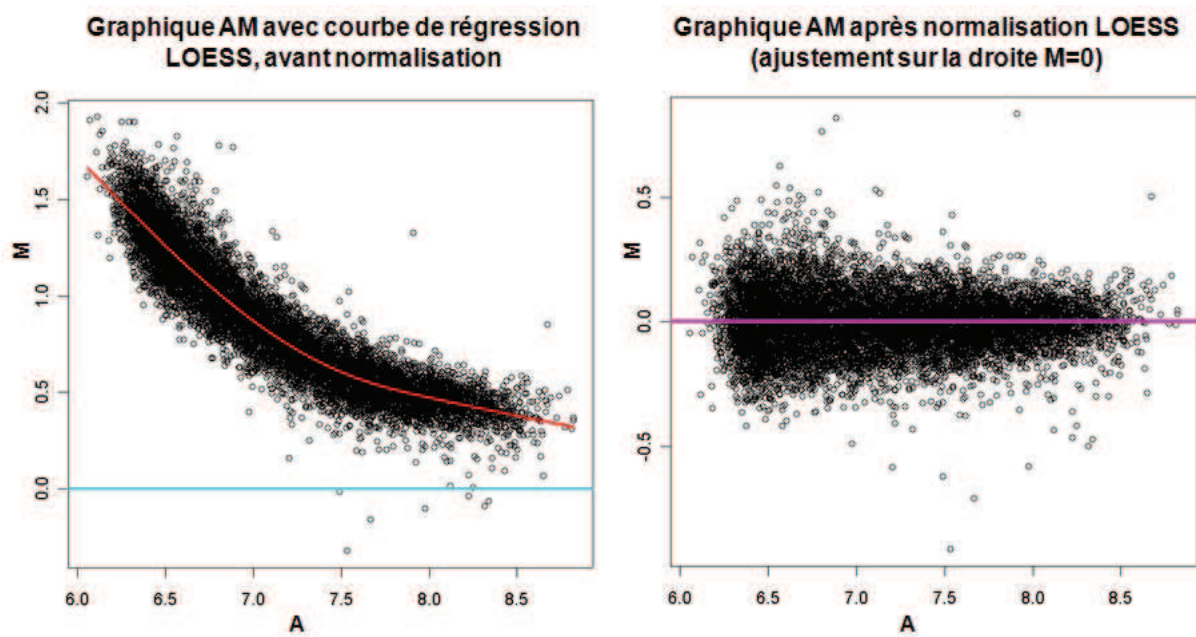


Figure 21 : exemples de graphiques AM montrant les effets d'une normalisation locale par la méthode LOESS.

A gauche, nuage de points avant normalisation par la méthode LOESS. A droite, nuage de points après normalisation par la méthode LOESS. Les données sont centrées sur la droite $M=0$, et la forme du nuage devient linéaire (figure adaptée d'après (Nettleton, 2009)).

- Normalisation par les quantiles

La normalisation par les quantiles est une procédure qui rend les distributions de plusieurs sets de données strictement identiques (Bolstad et al., 2003; Rao et al., 2008). Le principe de cette normalisation est simple à comprendre à partir d'un set de données de petite taille :

- Dans un premier temps, les données brutes sont organisées sous la forme d'une matrice où les échantillons sont disposés en colonnes, et les gènes cibles en lignes :

		Echantillon 1	Echantillon 2	Echantillon 3
Gène A	(I)	6	13	8
Gène B	(II)	12	10	6
Gène C	(III)	1	3	11

- Dans un deuxième temps, les données de chaque échantillon sont triées par ordre croissant, puis la moyenne de chaque ligne est calculée (pour chaque échantillon, l'ordre initial des valeurs est indiqué entre parenthèses) :

	Echantillon 1		Echantillon 2		Echantillon 3		Moyenne
Gène A	1	(III)	3	(III)	6	(II)	3
Gène B	6	(I)	10	(II)	8	(I)	8
Gène C	12	(II)	13	(I)	11	(III)	12

- Ensuite, les valeurs de chaque ligne sont remplacées par la moyenne :

	Echantillon 1		Echantillon 2		Echantillon 3		Moyenne
Gène A	3	(III)	3	(III)	3	(II)	3
Gène B	8	(I)	8	(II)	8	(I)	8
Gène C	12	(II)	12	(I)	12	(III)	12

- Enfin, les valeurs sont remises dans l'ordre initial pour chaque échantillon :

	Echantillon 1		Echantillon 2		Echantillon 3	
Gène A	8	(I)	12	(I)	8	(I)
Gène B	12	(II)	8	(II)	3	(II)
Gène C	3	(III)	3	(III)	12	(III)

A la fin du traitement, chaque échantillon contient donc le même jeu de données, mais ces données sont classées dans un ordre différent. Cette normalisation se révèle être particulièrement puissante sur les grands volumes de données : elle est donc très adaptée à la normalisation des expériences d'hybridation sur puces d'expression pangénomiques (figure 22).

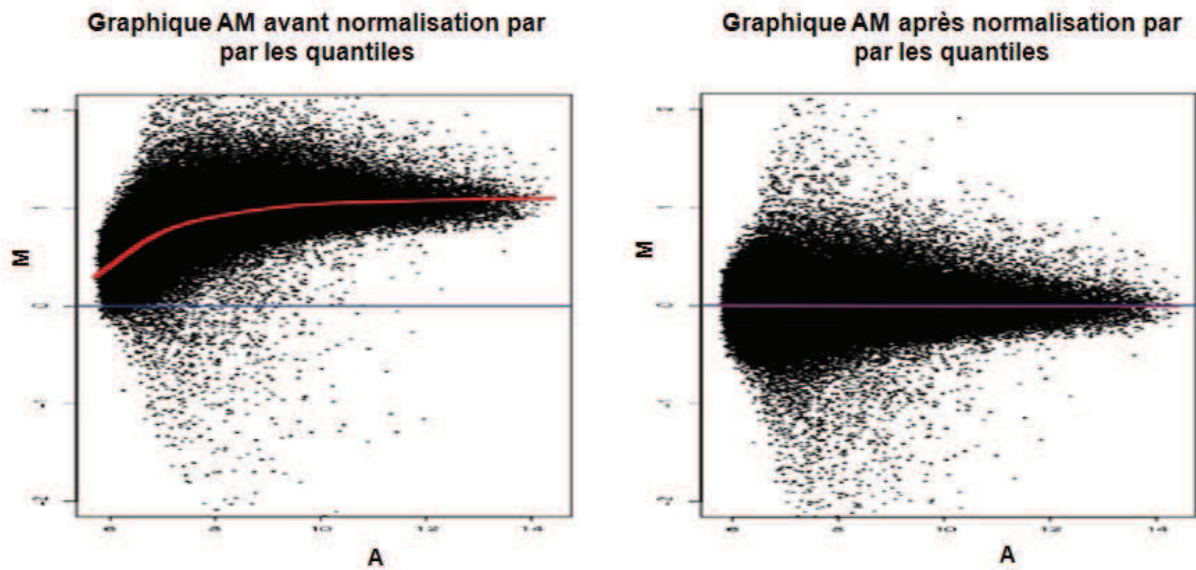


Figure 22 : exemples de graphiques AM montrant les effets d'une normalisation locale par la méthode des quantiles.

A gauche, nuage de points avant normalisation, à droite, nuage de points après normalisation par les quantiles. Sur le nuage de points avant normalisation, une courbe de régression de type LOESS est également visible à titre informatif, elle ne participe pas au processus de normalisation. Le nuage étant centré sur zéro et linéaire, les biais locaux sont corrigés (figure adaptée d'après (Bolstad et al., 2004)).

3. Outils statistiques pour l'analyse des profils d'expression des microARN

3.1. Filtrage des échantillons et des microARN

L'ensemble des méthodes de filtrages des échantillons ou des microARN sont des moyens d'obtenir des données finales de grande qualité et les plus informatives possibles. Une partie de ces méthodes statistiques sert également d'outils de contrôle qualité.

Lors du démarrage d'une étude de profils d'expression, des outils statistiques permettent tout d'abord d'éliminer les échantillons anormaux. Il s'agit principalement des graphiques de type boîte à moustaches (*box plot*) et des diagrammes en couleur (*heat maps*) basés sur les coefficients de corrélation de Pearson (Benjamini, 1988; Altman, 1991). Ainsi, les échantillons présentant une

distribution anormale de leurs données d'expression, et les échantillons se comportant de manière très différente des autres sont éliminés du processus d'analyse (figure 23).

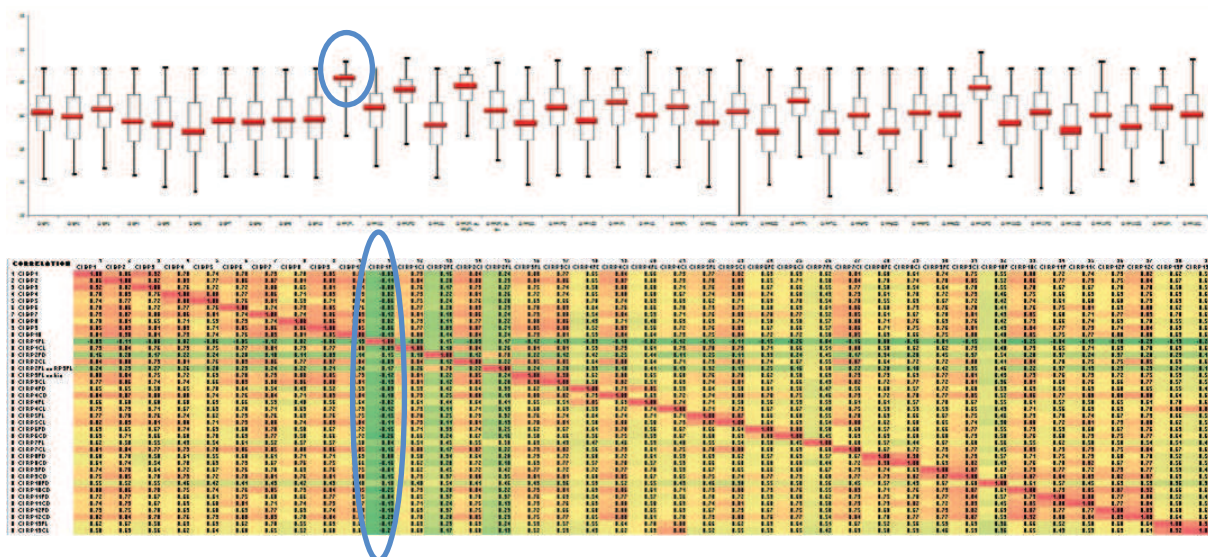


Figure 23 : boîtes à moustache (en haut) et diagramme de Pearson (en bas) réalisés à partir du même set de données brutes issues de plaques TLDA, sous le logiciel Microsoft Excel.

Ces représentations permettent de mettre en évidence les échantillons possédant une distribution anormale ou au comportement anormal. Par exemple, l'échantillon entouré en bleu possède une distribution très tassée, et des Ct moyens très tardifs. De plus, il n'est pas du tout corrélé avec les autres échantillons (analyses non publiées, réalisées au sein de notre laboratoire).

D'autres types de contrôle qualité, spécifiques à chaque plateforme mais ne reposant pas sur des outils statistiques, peuvent être utilisés afin de repérer d'éventuels échantillons ou gènes anormaux. Par exemple, pour les puces à ADN, un contrôle visuel de l'image scannée permet de visualiser du bruit de fond parasite ou une poussière, alors que pour les plaques TLDA, on peut visualiser la quantité de puits indéterminés sur un plan de plaque schématique, ou encore vérifier l'absence de fuites par l'analyse des données brutes de fluorescence.

Enfin, il existe d'autres méthodes appliquées au filtrage de gènes en fonction de leurs valeurs d'expression. On élimine en général les gènes faiblement exprimés dans tous les échantillons. Pour les puces à ADN, il s'agit des gènes dont les intensités brutes de fluorescence sont toujours proches du bruit de fond. Pour les plaques TLDA, il s'agit des gènes dont les Ct sont toujours indéterminés ou au-dessus d'un certain seuil. Il est important de supprimer ces gènes, car les valeurs faibles ne peuvent pas être considérées comme fiables et risquent de se retrouver différemment exprimées en fin d'analyse. Il est également courant d'éliminer les gènes présentant de faibles variations après normalisation. En effet, ils ne sont utilisables que pour la recherche de gènes de ménage. Ils n'apportent aucune autre information, puisque le but de toute analyse est de trouver des gènes différemment exprimés. En revanche, ils risquent de perturber les analyses statistiques ultérieures non supervisées, en fonction de l'algorithme choisi.

Dans tous les cas, ces filtres ont pour but de ne travailler que sur les données les plus robustes pour le reste du processus d'analyse.

3.2. Méthodes d'analyse non supervisées

L'objectif des méthodes d'analyses non supervisées est d'explorer les données « à l'aveugle » pour découvrir des associations intéressantes entre échantillons ou entre microARN. Ainsi, dans un clustering ou une analyse en composantes principales (ACP), les échantillons ou les microARN ayant des profils d'expression similaires sont regroupés entre eux.

- Classification hiérarchique

La classification hiérarchique est une méthode de clustering itérative, pouvant être réalisée sur les échantillons et sur les microARN (Eisen et al., 1998). A partir d'une matrice contenant les valeurs d'expression normalisées, où les échantillons sont disposés en colonnes et les gènes cibles en lignes, la classification hiérarchique est réalisée par la répétition des trois étapes suivantes (exemple de classification des échantillons) :

- 1/ l'algorithme calcule tout d'abord les distances entre tous les échantillons deux à deux, et groupe ensemble les deux échantillons les plus proches
- 2/ l'algorithme calcule ensuite la distance entre tous les échantillons restants deux à deux, le groupe précédemment formé étant considéré comme un échantillon unique, puis regroupe entre eux les deux échantillons ou groupe les plus proches
- 3/ l'algorithme répète encore la deuxième étape jusqu'à ce que tous les échantillons soient regroupés en un cluster unique.

Sans entrer dans les détails, il est nécessaire de préciser deux paramètres importants avant de lancer l'algorithme : la méthode de calcul de la distance entre les échantillons, et la méthode de calcul de l'agrégation entre les clusters. Une fois que l'algorithme a regroupé entre eux les échantillons ou les gènes qui possèdent des profils d'expression similaires, les résultats sont présentés sous la forme d'une représentation graphique matricielle, sur laquelle les valeurs d'expression sont remplacées par un code couleur. Par convention, les gènes sur-exprimés sont souvent représentés en rouge, et les gènes sous-exprimés en vert. Les associations en cluster sont généralement représentées en regard des lignes ou des colonnes, sous la forme d'arbres hiérarchisés (figure 24).

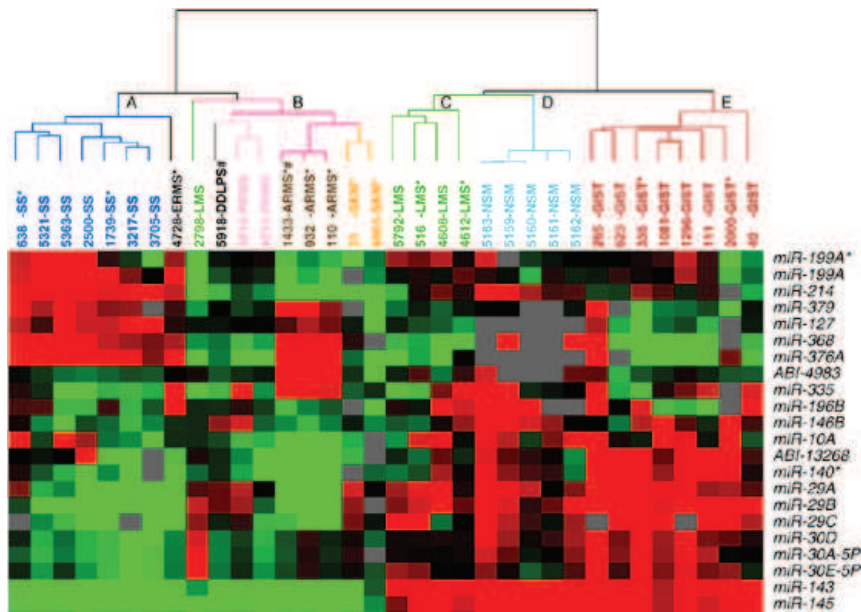


Figure 24 : exemple de clustering hiérarchique non supervisé des profils d'expression des microARN dans les sarcomes, réalisé à partir de données de puces d'expression.

Les échantillons sont disposés en colonne, et les gènes microARN en ligne. La classification hiérarchique est effectuée sur les échantillons (figure adaptée d'après (Subramanian et al., 2008)).

- Analyse en composantes principales

L'analyse en composantes principales (ACP) est un outil statistique puissant pour l'étude des profils d'expression (Li, 2010). Il s'agit d'une méthode de réduction de dimension qui, dans sa forme la plus informative, permet de représenter les informations biologiques les plus discriminantes entre échantillons ou entre microARN dans un espace en trois dimensions.

Il est possible de comprendre cette transformation de manière intuitive, sans entrer dans le détail des calculs mathématiques complexes sous-jacents. Si on imagine une étude portant sur la caractérisation des niveaux d'expression de trois microARN dans divers échantillons, il serait possible de représenter chaque échantillon dans l'espace, en lui attribuant pour coordonnées x , y et z les valeurs des niveaux d'expression des trois microARN respectifs. Plus les valeurs de x , y et z seront similaires entre deux échantillons, plus ces échantillons seront représentés par des points proches dans cet espace tridimensionnel. Les puces et les TLDA générant

des données d'expression de plusieurs centaines de microARN, il n'est donc pas possible de représenter ces données dans l'espace, qui nécessiterait autant de dimensions que de microARN étudiés. L'ACP va donc chercher à donner la meilleure visualisation possible de notre nuage de points, en s'appuyant sur les critères permettant de maximiser la dispersion des points dans le sous-espace à trois dimensions choisi.

Les logiciels informatiques implémentant l'ACP permettent généralement de faire tourner la représentation obtenue dans les trois axes (figure 25). La visualisation du nuage de points selon plusieurs perspectives permet généralement de faciliter l'observation de regroupements entre échantillons ou entre microARN.

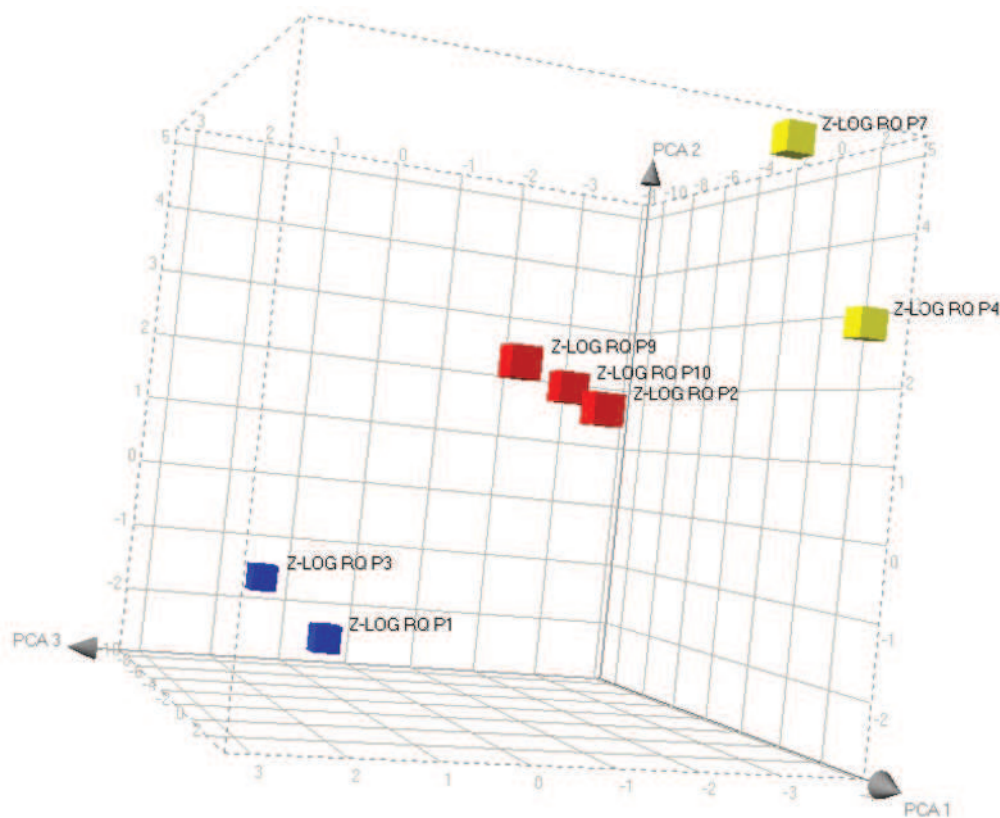


Figure 25 : exemple d'ACP non supervisée des profils d'expression des microARN dans des sarcomes, réalisée à partir de données issues de plaques TLDA.

Les deux échantillons bleus sont des liposarcomes différenciés, les trois échantillons rouges représentent deux sarcomes inclassés et un myxofibrosarcome, et les échantillons jaunes sont deux léiomyosarcomes (analyses non publiées, réalisées au sein de notre laboratoire).

3.3. Tests statistiques appliqués aux analyses supervisées : T-test et ANOVA

Les analyses non supervisées sont idéales pour résoudre les problématiques de classification, ou lorsque peu de données biologiques ou cliniques sont connues sur les échantillons. Elles permettent souvent de découvrir les grandes tendances, et de mettre en évidence des regroupements inattendus.

Cependant, il est parfois nécessaire d'aller plus loin pour comprendre certains phénomènes biologiques. Il peut être informatif de connaître la liste complète des gènes dérégulés entre deux ou plusieurs conditions très particulières relatives aux échantillons. Grâce aux analyses supervisées, il est alors possible de visualiser de manière plus claire les voies métaboliques qui sont impliquées dans les phénomènes étudiés. Ces informations étant souvent noyées dans la masse des données pangénomiques, seules les analyses supervisées permettent de les mettre en avant.

Les analyses supervisées sont également indispensables pour la recherche de biomarqueurs d'intérêt (Orr and Scherf, 2002). Ainsi, en obtenant une liste restreinte comportant les gènes les plus différentiellement exprimés dans une problématique donnée, il est ensuite possible d'utiliser ces quelques gènes pour des applications cliniques de routine, avec des techniques de dépistage beaucoup moins coûteuses que les techniques pangénomiques d'établissement des profils d'expression.

Concrètement, le test de student, ou T-test, est utilisé pour la recherche des gènes différentiellement exprimés entre deux groupes d'échantillons, alors que l'analyse ANOVA (pour *analysis of variance*), basée sur la même logique, est utilisée dès que le nombre de groupes est supérieur à deux (Churchill, 2004; Ayroles and Gibson, 2006). Une des conditions requises pour pouvoir appliquer ces tests est que les groupes soient équilibrés, c'est-à-dire qu'il y ait à peu près le même nombre d'échantillons dans chaque groupe.

Lors de l'analyse de deux groupes d'échantillons, le T-test va rendre compte de deux règles principales. Le meilleur gène candidat sera celui qui sera le plus

différentiellement exprimé possible entre les deux groupes, mais dans le même temps, l'expression de ce gène devra être la plus stable possible au sein de chaque groupe. Le T-test et l'ANOVA génèrent tous deux des *p-values* attribuées à chacun des gènes. Plus cette valeur est basse, et plus le gène correspondant répond aux deux règles citées ci-dessus. La *p-value* peut être vue comme la probabilité d'obtenir un faux positif, c'est-à-dire la probabilité de considérer que le gène est différentiellement exprimé entre les deux groupes, alors qu'il ne l'est pas dans la réalité. On fixe généralement arbitrairement un seuil à 5 % ($p = 0,05$), pour ne retenir que les gènes situés en-dessous de ce seuil. Par ailleurs, il existe des méthodes statistiques permettant de corriger ce seuil (par exemple, les techniques de Bonferroni ou du *False Discovery Rate*), ou des méthodes de calcul de la puissance d'un test *a posteriori*, qui permettent d'éviter d'éventuelles erreurs liées à la fixation d'un seuil arbitraire (Reiner et al., 2003; Streiner and Norman, 2011).

A la suite d'un T-test ou d'une ANOVA, il est possible de ne conserver que les gènes dont la *p-value* est inférieure au seuil fixé, et de relancer des algorithmes de classification. On parlera alors de clustering supervisé, ou encore d'ACP supervisée (Frades and Matthiesen, 2010).

Comme nous l'avons vu précédemment, les principales méthodes de normalisation et d'analyse statistique des grands volumes de données ont été initialement développées pour les puces à ADN. Cependant, les données générées par la plateforme TLDA dite à moyen débit étant relativement nombreuses, il serait intéressant de voir si elles sont applicables à des données de type PCRq. Durant la première partie de ce travail de thèse, nous avons développé une méthodologie d'analyse des TLDA largement inspirée de celle des puces à ADN, en utilisant les mêmes outils, mais avec quelques adaptations liées à la nature des données, dont les valeurs sont exprimées en Ct.

RÉSULTATS EXPÉRIMENTAUX

INTRODUCTION AU TRAVAIL EXPÉRIMENTAL

PROBLÉMATIQUE

Ce travail de thèse a été réalisé en deux temps, la première partie du projet pouvant être considérée comme une phase de faisabilité et d'optimisation d'une méthodologie d'analyse sur une plateforme technique innovante. Nous avons étudié la possibilité d'obtenir des profils d'expression des microARN de manière fiable et reproductible, à partir d'une technologie de PCRq à moyen débit, réalisée sur des plaques microfluidiques de type TLDA. Les tests de faisabilité ont été réalisés à partir d'échantillons variés de sarcomes, conservés en congélation, ou fixés au formol et inclus en paraffine (principalement GIST, rhabdomyosarcomes et ostéosarcomes). La première étude basée sur la technologie TLDA ayant été publiée en 2005 (Abruzzo et al., 2005), et les premières plaques TLDA dédiées au microARN ayant été commercialisées en 2007, l'analyse d'expériences de ce type était encore balbutiante lors du démarrage de ce travail de thèse. En effet, la plupart des utilisateurs restaient alors cantonnés à des analyses conventionnelles de PCRq, et ne tiraient pas partie de la relative abondance de données générées par cette technologie. Il n'existait pas non plus de solution logicielle dédiée au contrôle qualité et à l'analyse de ces grands volumes de données de type PCRq. Nous avons donc dû développer une nouvelle méthodologie d'analyse, inspirée à la fois des outils d'analyse classique des expériences de PCRq, et des outils développés initialement pour les puces à ADN. Nous avons fait en sorte que cette méthodologie soit applicable à l'analyse de cohortes homogènes ou hétérogènes, intégrant tous types d'échantillons, qu'ils soient congelés ou fixés au formol et inclus en paraffine (FFPE).

Dans un second temps, nous avons essayé d'appliquer cet outil à la résolution de problématiques liées à la prise en charge des patients porteurs de tumeurs solides de type rhabdomyosarcome ou ostéosarcome. Nous avons choisi de nous focaliser sur ces deux types de sarcomes en raison d'un fort recrutement au Centre Léon Bérard, mais aussi pour répondre à la demande de notre panel d'experts médicaux, à la recherche de nouvelles approches qui pourraient permettre de mieux orienter leurs prises de décision. Au sein des sarcomes pédiatriques, les RMS sont des tumeurs pour lesquelles le résultat du typage morphologique conditionne de manière très forte la prise en charge thérapeutique. Or, il existe au sein de cette entité des tumeurs très difficiles à classer sur le plan morphologique, que les médecins anatomopathologistes positionneraient plutôt dans une zone grise entre ERMS et ARMS. De plus, les fragments analysables étant généralement de petite taille (biopsies et microbiopsies), cela ajoute une difficulté supplémentaire liée à la non représentativité des prélèvements. Les formes alvéolaires étant de bien plus mauvais pronostic, le développement d'un nouvel outil leur permettant de démembrer ces entités difficiles serait d'une grande aide pour la prise en charge des patients concernés. A travers ce travail de thèse, nous avons donc décidé d'évaluer la capacité des profils d'expression des microARN à apporter une réponse à cette problématique.

En ce qui concerne la prise en charge clinique des OS, la réponse à la chimiothérapie néoadjuvante est considérée comme le facteur pronostique le plus important. Cependant, actuellement, cette information ne peut être déterminée qu'après l'acte chirurgical, par évaluation de la nécrose tumorale. S'il était possible de prédire la réponse à la chimiothérapie avant chirurgie, cela permettrait de la planifier au mieux, et il serait également envisageable d'adapter ou de personnaliser les traitements. Nous avons donc choisi d'évaluer également la capacité des profils d'expression des microARN à prédire la réponse des patients atteints d'OS.

OUTILS

Modèles animaux et lignées cellulaires :

La plupart de nos études ont été réalisées à partir de tumeurs congelées ou FFPE, que ce soit pour le développement de notre méthodologie d'analyse des données de TLDA, ou pour les études plus appliquées, dans les OS et les RMS. Cependant, nous avons ponctuellement utilisé les lignées cellulaires humaines suivantes :

➤ Lignée humaine de RMS SJCRH30 :

Encore appelée RMS13 (ATCC# CRL-2061), cette lignée cellulaire a été isolée à partir d'une métastase osseuse chez un enfant porteur d'un ARMS. Nous avons au préalable vérifié la présence d'un transcrit de fusion de type PAX3-FOXO1, par technique de RT-PCRq.

Au cours de ce travail de thèse, cette lignée a été utilisée pour la mise au point de notre puce prototype de diagnostic des sarcomes à translocation (cf. résultats partie III, prix 2010 des Amis de l'Université de Lyon). Secondairement incluse en paraffine, nous avons également utilisé cette lignée pour le développement de notre méthodologie d'analyse des données issues de TLDA, en particulier lors de la comparaison des échantillons congelés et FFPE en miroir (cf. article I, en soumission au journal *RNA*).

➤ Lignées humaines d'OS MNNG/HOS Cl #5 et Saos-2 :

Les lignées cellulaires MNNG/HOS Cl #5 (ATCC# CRL-1547) et Saos-2 (ATCC# HTB-85) sont toutes deux des lignées d'OS utilisées au cours de notre étude sur les profils d'expression des microARN appliqués à la prédiction de la réponse à l'IFO (cf. article II) (Gougelet et al., 2011b). La lignée MNNG/HOS Cl #5 peut être considérée comme une lignée chimiorésistante, alors que la lignée Saos-2 est chimiosensible.

➤ Lignée humaine de chondrosarcome SW1353 :

La lignée SW1353 (ATCC# HTB-94) est une lignée de chondrosarcome que nous avons utilisée pour démontrer que les signatures de réponses à la chimiothérapie sont différentes en fonction du type tumoral et du type de médicament (cf. article III) (Gougelet et al., 2011a).

➤ Lignées humaines de lymphomes Daudi et Namalwa :

Les lignées Daudi (ATCC# CCL-213) et Namalwa (ATCC# CRL-1432) sont toutes deux des lignées de lymphomes de Burkitt, utilisées pour la même raison que la lignée de chondrosarcome (cf. article III) (Gougelet et al., 2011a).

➤ Lignées humaines de léiomyosarcomes SK-LMS-1 et CNIO-AA :

La lignée SK-LMS-1 (ATCC# HTB-88) est dérivée d'un léiomyosarcome utérin, alors que la lignée CNIO-AA, gracieusement fournie par la société Pharmamar (Tres Cantos, Espagne) est dérivée d'un léiomyosarcome des tissus-mous. Ces deux lignées ont été utilisées pour les tests fonctionnels dans le cadre de notre étude portant sur l'analyse des profils d'expression protéiques des léiomyosarcomes (cf. article IV, en soumission au journal *EJC*).

➤ Modèle d'OS de rat :

Il s'agit d'un modèle orthotopique et métastatique d'OS de rat (Allouche et al., 1980; Dutour et al., 2007). Les tumeurs générées sont de type ostéoblastique, dans des rats immunocompétents de souche Sprague-Dawley (Laboratoires Charles River, Wilmington, MA, USA). Nous avons utilisé ce modèle pour la génération d'OS dans le cadre de notre étude sur les profils d'expression des microARN appliqués à la prédiction de la réponse à l'IFO (cf. article II) (Gougelet et al., 2011b)(figure 26).

La tumeur est implantée en plaçant un fragment tumoral hyperprolifératif sur la diaphyse tibiale, après abrasion périostée. Quatorze jours après transplantation, les rats subissent un ^{18}F FDG PET Scan. Ils sont ensuite assignés de manière aléatoire à

un groupe contrôle, traité à l'aide d'une solution saline, ou à un groupe traité par injection sous-cutanée d'IFO, deux fois à sept jours d'intervalle. Un dernier PET Scan est réalisé une semaine après la deuxième injection, et 24h avant le sacrifice des rats.

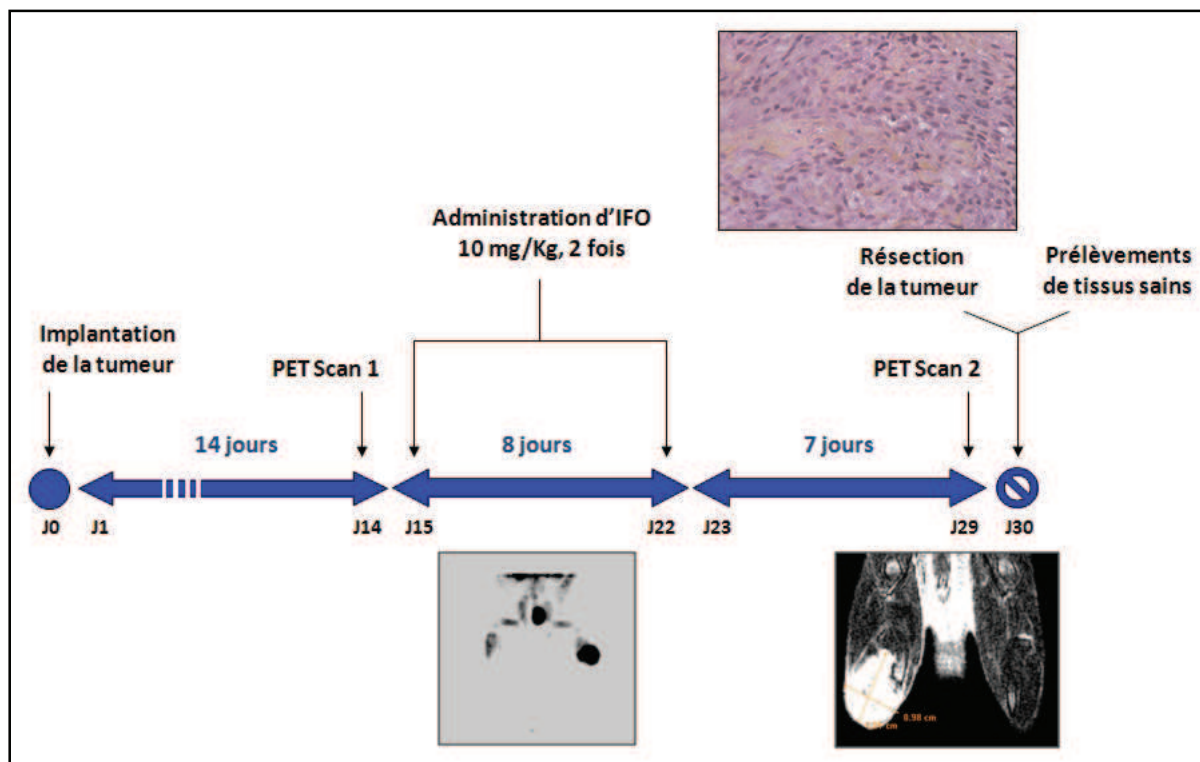


Figure 26 : organisation des manipulations sur le modèle d'OS de rat.

Plateforme TLDA pour microARN :

Toutes les expériences de TLDA ont été réalisées sur un appareil PCRq de type Applied Biosystems 7900HT. Selon l'origine des prélèvements, nous avons utilisé des TLDA pour l'homme ou pour les rongeurs. Dans la version 2 des TLDA, l'ensemble des microARN est détecté à l'aide de deux plaques A et B. Pour l'homme, 377 microARN sont détectés par la plaque A, et 290 microARN par la plaque B, pour un total de 667 microARN. Pour les rongeurs, 226 microARN de rat sont détectés par la plaque A, et 77 par la plaque B, pour un total de 303 microARN. La plateforme TLDA simplifie considérablement les étapes de transcription inverse et de chargement des plaques, par rapport à des PCRq uniques (figure 27).

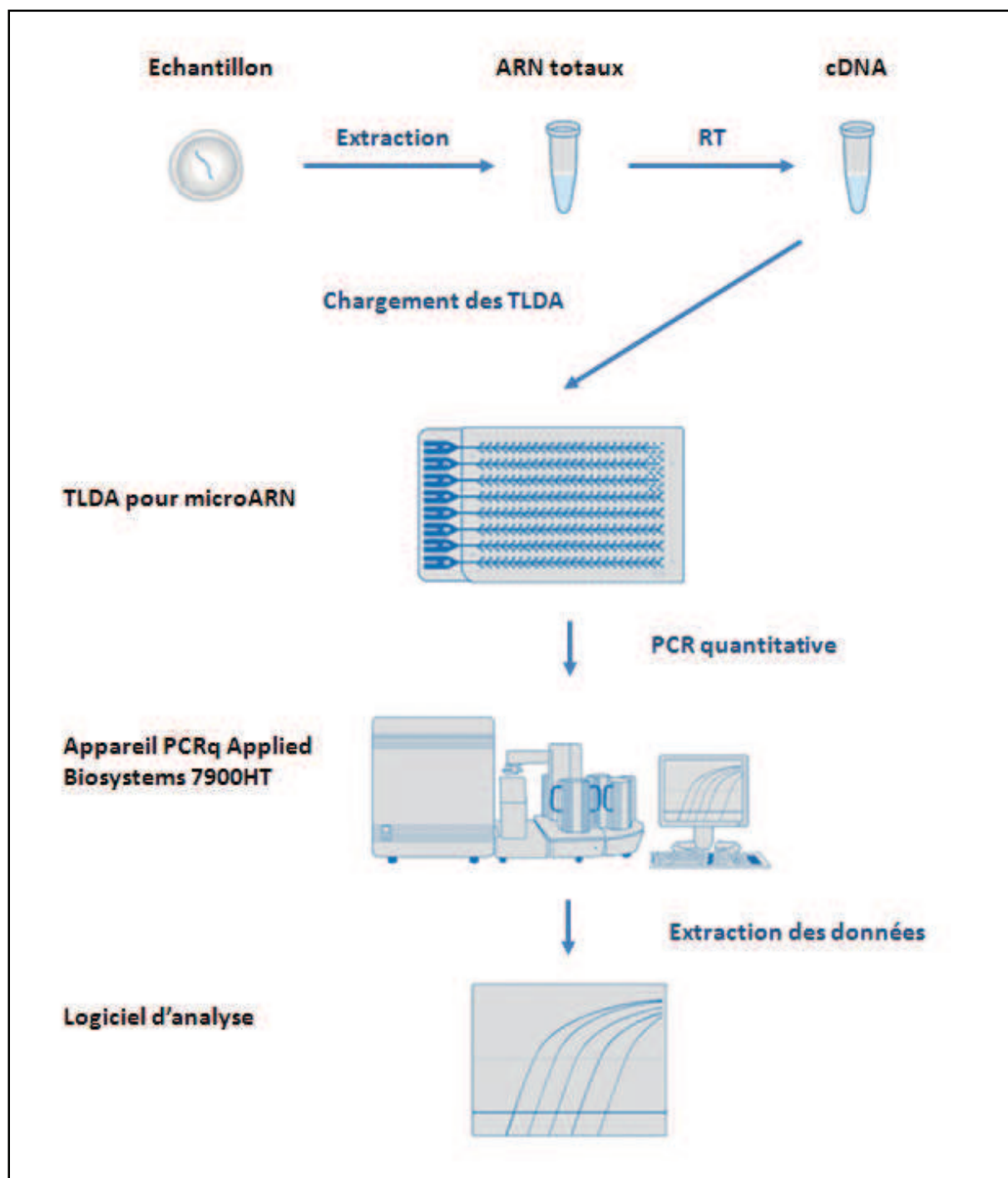


Figure 27 : processus résumant une expérience réalisée sur plaque TLDA.

Après extraction des ARN totaux, les microARN sont transformés en ADNc lors d'une réaction de transcription inverse, à l'aide d'amorces de type tige-boucle massivement multiplexées. Les ADNc sont ensuite introduits dans les huit puits de chargement de la plaque TLDA, qui contient déjà les amorces et les sondes de PCRq, à l'état déshydraté, dans chacun des 384 puits. Les puits sont remplis par centrifugation, grâce à un réseau de micro-canaux. Une fois que la plaque est remplie et scellée, il est possible de réaliser la réaction proprement dite de PCRq.

Techniques de FISH pour la détection des translocations rares dans les ARMS :

Par technique FISH, la détection de la présence d'un transcrit rare s'effectue en deux temps. Dans un premier temps, on vérifie l'absence d'un transcrit de fusion courant, à l'aide de la sonde FOXO1 de type break-apart (figure 28). L'absence de réarrangement du gène FOXO1 élimine la présence des translocations générant les transcrits de fusion courant de type PAX3-FOXO1 et PAX7-FOXO1.

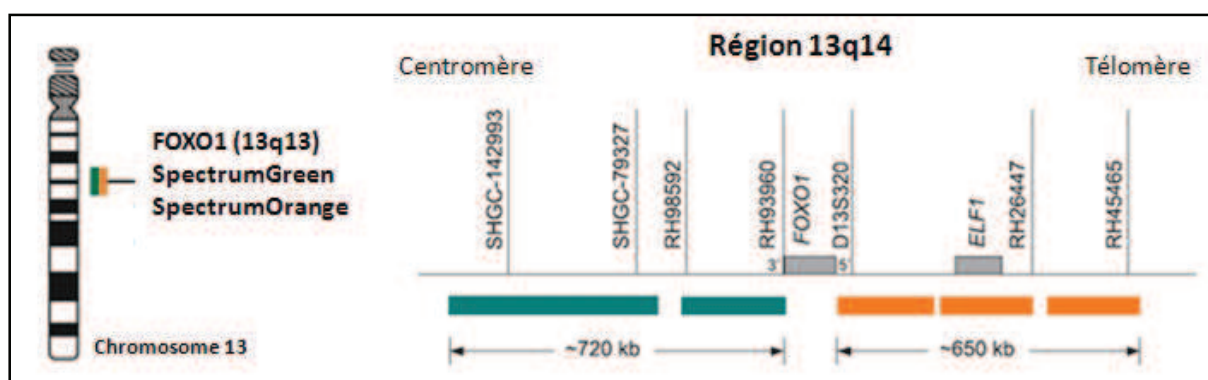


Figure 28 : idéogramme présentant la sonde FOXO1 de type *break-apart*.

A gauche, position de la sonde sur le bras long du chromosome 13. A droite, présentation de la disposition des sondes de part et d'autre du point de cassure à tester, au niveau du gène FOXO1. Dans un noyau normal, on observe deux signaux de fusion (signaux orange et vert très proches). Dans un noyau avec réarrangement, on observe un signal de fusion accompagnée d'un signal vert et d'un signal rouge séparés (sonde Vysis LSI FOXO1 *dual color, break apart rearrangement probe*, # 05J48-006, d'après le manuel d'utilisation de la sonde fourni par Abbott Molecular).

Dans un deuxième temps, on effectue la recherche d'un réarrangement du gène PAX3, à l'aide de la sonde PAX3 de type break-apart (figure 29). La présence d'un réarrangement du gène PAX3, en l'absence de réarrangement du gène FOXO1, signe la présence d'une translocation rare générant un transcrit de fusion de type PAX3-NCOA1, PAX3-NCOA2 ou PAX3-FOXO4 (figure 30).

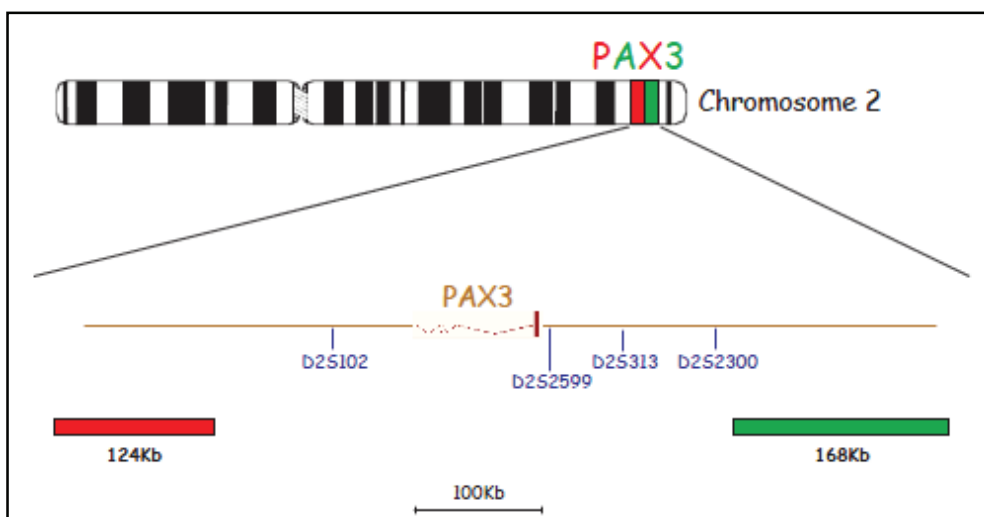


Figure 29 : idéogramme présentant la sonde PAX3 de type *break-apart*.

En haut, position de la sonde sur le bras long du chromosome 2. En bas, présentation de la disposition des sondes de part et d'autre du point de cassure à tester, au niveau du gène PAX3. Dans un noyau normal, on observe deux signaux de fusion (signaux orange et vert très proches). Dans un noyau avec réarrangement, on observe un signal de fusion accompagnée d'un signal vert et d'un signal rouge séparés (sonde PAX3 *break-apart*, #LPS 012, d'après le manuel d'utilisation de la sonde fourni par Cytocell).

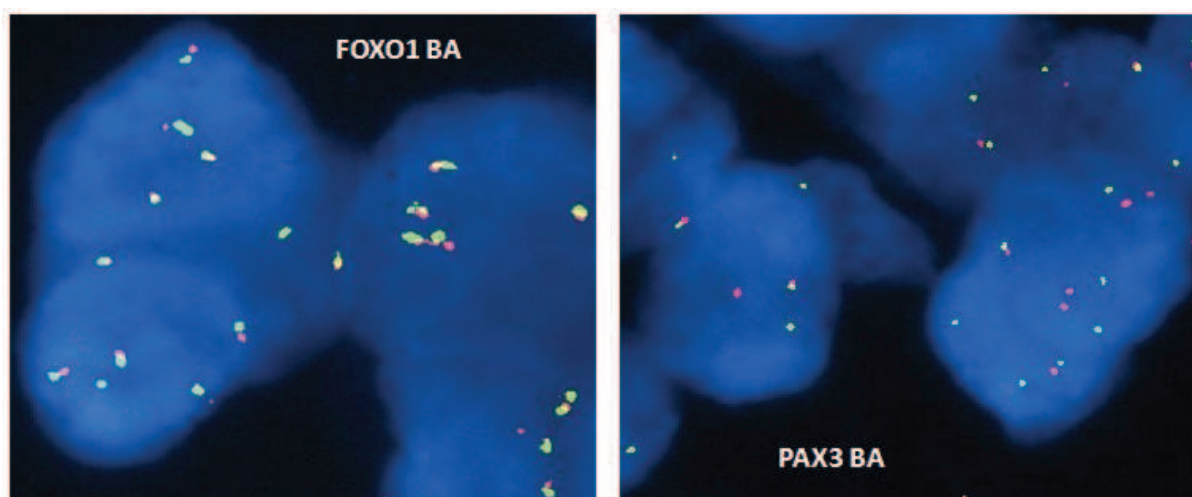


Figure 30 : photos de techniques de FISH prises au microscope à fluorescence sur un cas d'ARMS à transcrit rare.

A gauche, résultat de l'hybridation à l'aide de la sonde FOXO1 de type *break-apart*. A droite, résultat de l'hybridation à l'aide de la sonde PAX3 de type *break-apart* (photos réalisées au grossissement x63).

Plateforme CGH utilisée pour l'établissement des profils génomiques :

Pour l'établissement des profils génomiques des cas de RMS de diagnostic difficile, la plateforme de puces CGH Agilent a été utilisée (figure 31). Brièvement, l'ADN génomique de la tumeur est extrait à partir de grattages d'un bloc de paraffine. Cet ADN tumoral est marqué à l'aide d'un fluorochrome rouge, puis co-hybridé avec un ADN témoin marqué à l'aide d'un fluorochrome vert sur une puce à ADN possédant quatre zones contenant chacune 180 000 sondes réparties sur l'ensemble du génome humain (il est possible d'analyser quatre tumeurs à la fois). Par compétition entre l'ADN tumoral et l'ADN témoin sur chacun des spots de la puce, il est alors possible de déterminer les pertes, les gains et les amplifications présents dans le génome tumoral.

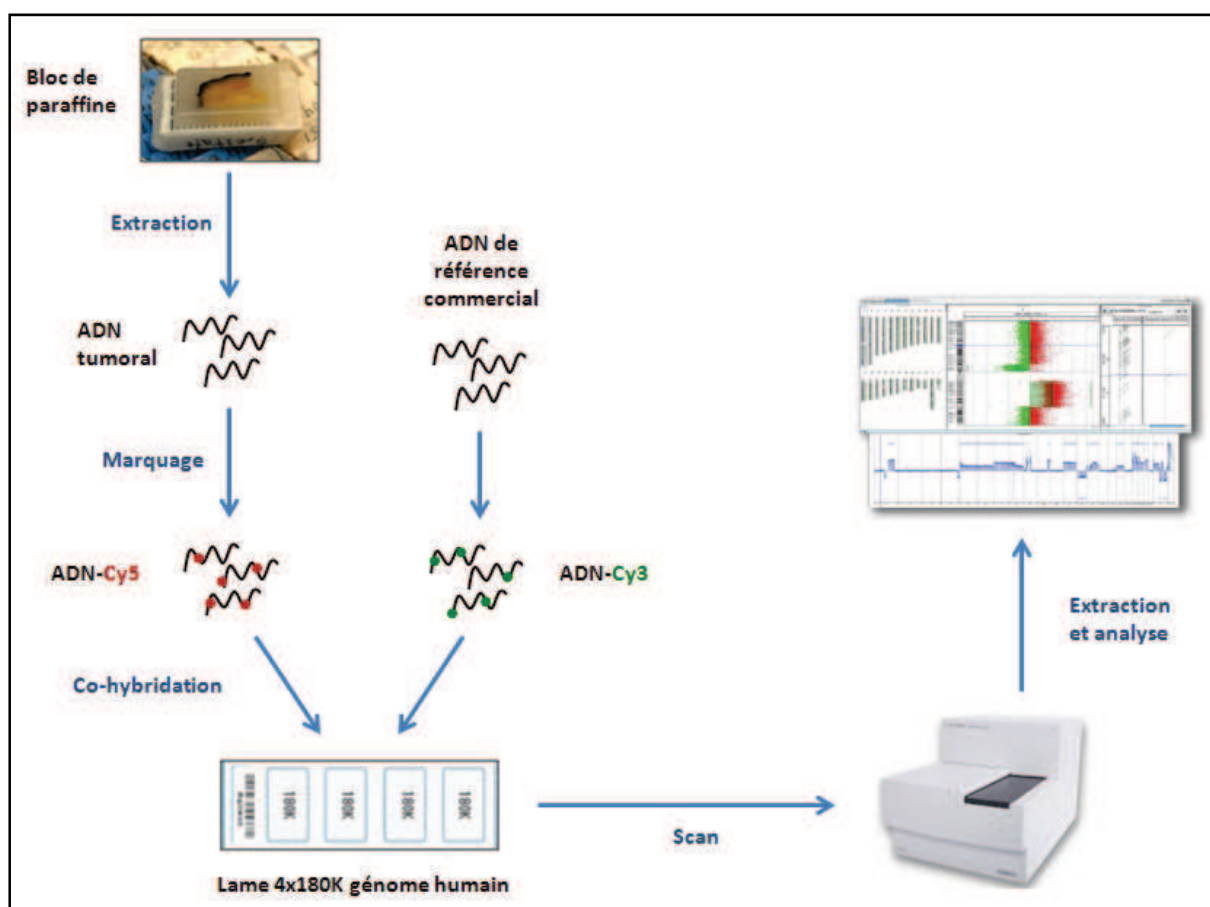


Figure 31 : déroulement d'une expérience de CGH réalisée sur puces pangénomiques Agilent.

Logiciels de programmation de d'analyse :

Les manipulations de données ont été réalisées principalement sous trois logiciels. Le logiciel Perl a été utilisé pour la manipulation des fichiers en sortie de l'appareil PCRq. Ainsi, nous avons développé un script Perl permettant de concaténer les données brutes issues de différentes plaques TLDA, par des manipulations de lignes de texte (figure 32). L'ensemble des données nécessaires à l'analyse d'une cohorte était alors disponible sous un seul fichier importable dans un logiciel d'analyse.

```
Ce programme permet de creer un fichier .txt avec tabulations
contenant les Cts bruts des plaques TLDA pour import sous Excel.

Entrez le nom de l'etude : Clustering Léios
Entrez les numeros absolus des plaques a analyser (sous la forme "num1+num2") : 12+26+30
Vous avez specifie les numeros de 3 plaques TLDA.
```

Figure 32 : récupération des données brutes via le script Perl.

Invite de commande du micrologiciel permettant de créer une cohorte d'analyse à partir des fichiers bruts donnés par l'appareil PCRq (script Perl développé en interne).

Les fondamentaux de la méthodologie d'analyse des données brutes issues de plaques TLDA ont été développés dans le logiciel Microsoft Excel, à l'aide de formules et de macros. Enfin, l'intégralité des méthodes de traitement et de normalisation des données a ensuite été regroupée sous un script R, qui nous a permis de pouvoir appliquer des méthodes de normalisation locales beaucoup plus poussées, et d'éviter de jongler continuellement d'un outil informatique à l'autre.

Enfin, les analyses statistiques ont été réalisées tour à tour à l'aide des logiciels R, Tiger MeV et TIBCO Spotfire DecisionSite for Functional Genomics Analysis.

I. MISE AU POINT D'UNE MÉTHODOLOGIE D'ANALYSE DES RÉSULTATS ISSUS DE TLDA

1. Des Cts bruts à l'interprétation biologique : importance du traitement des données primaires

Pour mener à bien ce travail de thèse, nous avons choisi d'étudier les profils d'expression des microARN à l'aide d'une technologie basée sur la PCRq, appelée TLDA. Ces plaques microfluidiques correspondent en fait à une technologie de PCRq dite à « moyen débit », et sont de plus en plus utilisées pour l'établissement des profils d'expression des microARN. Comme pour l'obtention des profils d'expression des ARNm sur puces à hybridation de type *microarray*, les données provenant de plaques TLDA nécessitent des méthodologies d'analyse efficaces, pour produire des résultats robustes et à forte valeur ajoutée. Une étape cruciale du traitement des données est la normalisation : elle a pour but de rendre comparables entre elles des données provenant de plusieurs plaques TLDA, réalisées à des moments différents et avec des échantillons variés. La normalisation s'attache à réduire les erreurs de mesures et les variations techniques qui se produisent lors des différentes étapes expérimentales, de l'exérèse de la tumeur à l'analyse de la fluorescence dans les puits des plaques TLDA. Ces erreurs, appelées également biais, peuvent être purement techniques, mais également biologiques, dépendant alors de la nature des échantillons testés.

Il existe aujourd'hui un grand nombre de méthodes de normalisation et d'analyse, parfois radicalement différentes. Il convient de les choisir avec discernement, car une méthode inadaptée peut à son tour introduire des biais entraînant des variations n'ayant aucune signification biologique. L'analyse et l'interprétation des résultats s'en trouveraient alors faussées. Le choix de la méthode de normalisation est donc un point fondamental. Il permet d'obtenir des données de

départ les plus « propres » possibles, et représentant la réalité biologique de manière juste et précise.

Le but de cette étude était de développer une méthodologie d'analyse reproductible, à partir des données brutes générées par les plaques TLDA, et qui soit indifféremment applicable à des échantillons congelés ou FFPE. Pour cela, nous avons réalisé les profils d'expression des microARN de 117 échantillons tumoraux congelés ou FFPE. Cette grande cohorte était plus précisément composée de 104 échantillons FFPE de type GIST, RMS et OS, ainsi que de 13 sarcomes de types variés pour lesquels nous disposons à la fois d'un fragment congelé et d'un fragment FFPE correspondant. La totalité de ces échantillons tumoraux a été utilisée pour déterminer la meilleure méthode de normalisation de données d'expression des microARN. Les performances des différentes méthodes de normalisation testées ont été évaluées sur la base de leur capacité à classer correctement les différents types de sarcomes lors d'un clustering non supervisé.

Nous avons tout d'abord montré qu'il était possible de normaliser les données brutes d'expression des microARN générées à partir de plaques TLDA à l'aide de la méthode standard du delta-delta CT ($\Delta\Delta Ct$), à la condition que les cohortes analysées soient constituées d'échantillons tumoraux homogènes, c'est-à-dire exclusivement constituées d'échantillons congelés ou FFPE. Dans ce cas, l'utilisation de la médiane en tant qu'endogène ou gène de référence virtuel nous a permis d'obtenir les meilleurs résultats. Quelques modifications mineures de la méthode du $\Delta\Delta Ct$ ont également été introduites, afin de pouvoir analyser ultérieurement les données normalisées à l'aide de logiciels initialement développés pour l'analyse de puces d'expression et la découverte de biomarqueurs d'intérêt.

Dans un second temps, nous avons montré qu'il était également possible d'analyser des cohortes constituées d'échantillons hétérogènes (congelés et FFPE), en appliquant une méthode de normalisation locale supplémentaire avant la

normalisation globale par la méthode du $\Delta\Delta\text{Ct}$. En effet, en travaillant sur des cohortes hétérogènes, nous avons découvert la présence d'un biais local lié à la fixation de l'échantillon, c'est-à-dire un biais dépendant de la valeur des Cts. Deux méthodes de normalisation locales ont été testées, et seule la normalisation de type LOESS a permis de corriger de manière efficace ce biais local.

Tout au long du processus de normalisation et d'analyse des données, de l'extraction des données brutes à l'obtention des valeurs d'expression en quantité relatives (RQ), nous avons veillé à ce que les données ne subissent aucune modification inappropriée, en surveillant chacune des étapes à l'aide d'outils de contrôle qualité adaptés.

Le développement de cette méthodologie d'analyse nous a permis par la même occasion d'obtenir les profils d'expression des microARN pour de nombreuses tumeurs (GIST, RMS, OS, liposarcome, léiomyosarcome, chondrosarcome, synoviosarcome, histiocytome fibreux malin, tumeur maligne des gaines nerveuses périphériques et sarcome inclassé) et en particulier pour des grandes cohortes de tumeurs de type RMS et OS. Nous avons également montré que les profils d'expression obtenus étaient directement exploitables pour des analyses statistiques non supervisées et supervisées de type clustering. Par la suite, nous avons utilisé ces données d'expression pour affiner la classification nosologique des RMS, et nous avons également étudié leur pouvoir prédictif de la réponse à la chimiothérapie néo-adjuvante dans les OS.

Document à suivre :

Article I en soumission : *Defining the optimal analysis methodology for miRNA profiling using TaqMan® Low Density Arrays (TLDA) with frozen and formalin-fixed paraffin-embedded samples.*

Title:

Defining the optimal analysis methodology for miRNA profiling using TaqMan® Low Density Arrays (TLDA) with frozen and formalin-fixed paraffin-embedded samples.

Authors:

Daniel Pissaloux ^{*1,2,3}, Héloïse Philippon ^{*2}, Anthony Besse ², Adeline Duc ^{1,2}, Anne-Valérie Decouvelaere ¹, Sébastien Salas ⁵, Jean-Yves Blay ^{2,3,4} and Laurent Alberti ^{1,2,3}

Author affiliations:

*These authors contributed equally to this work, ¹ Département de Biopathologie, Centre Léon Bérard, Lyon, France, ² UMR INSERM 1052 - CNRS 5286 équipe Ciblage de la tumeur et de son environnement immunitaire, Centre de Recherche en Cancérologie de Lyon, Centre Léon Bérard, Lyon, France, ³ Conticanet (FP6-06188), www.conticanet.eu, ⁴ EORTC, Brussels Belgium, ⁵ CHU La Timone, Marseille, France.

Short Title:

An optimized methodology for analyzing miRNA TLDA data.

Keywords:

miRNA; normalization; TLDA; cancer; FFPE; clustering

Abstract:

MicroRNAs (miRNAs) are small non-coding RNAs regulating gene expression in a complex network and with coordinate actions. They are involved in many cell processes and play a crucial role in cancer development. High throughput miRNA expression profiling is the best way to identify single miRNA biomarkers in cancer cell lines and clinical samples, and to conduct classification studies providing new insights.

Global miRNA expression profiling using TaqMan® Low Density Array technology (TLDA, Life Technologies) was conducted in 117 tumor samples, mainly composed of gastrointestinal stromal tumors (GISTs), rhabdomyosarcomas and osteosarcomas. Frozen and formalin-fixed, paraffin embedded (FFPE) specimens were used to identify the most appropriate method for the accurate normalization of miRNA expression data. The performances of the different normalization methods were compared for their ability to correctly classify these distinct tumor entities.

When working on cohorts exclusively composed of frozen or FFPE samples, applying the standard delta-delta Ct ($\Delta\Delta Ct$) method alone was sufficient, and optimal results were obtained choosing the median as a virtual endogenous housekeeping gene. For mixed frozen and FFPE cohorts, a prior LOESS normalization was the most efficient way to correct the local bias induced by the fixation.

The methodology developed here enables a reliable determination of miRNA expression profiles. In addition, this methodology works equally on frozen and FFPE archived material, which is a relevant advantage in the field of rare pathologies. Obtaining statistically significant data is no longer an issue for miRNA discovery.

Introduction:

MicroRNAs (miRNAs) are small non-coding RNAs that repress gene expression through sequence-specific base pairing with their target mRNAs (Doench and Sharp 2004). There are currently 2042 entries for *Homo sapiens* mature miRNAs in the 19th release of the miRBase sequence database, but new entities are still being identified (Griffiths-Jones 2004; Griffiths-Jones et al. 2006; Griffiths-Jones et al. 2008). MiRNAs are involved in many processes such as control of cell proliferation, differentiation and apoptosis (He and Hannon 2004). MiRNAs are deregulated in many tumor types, and can either act as oncogenes or tumor suppressors (Tie and Fan 2011; Zhang et al. 2007). Each miRNA has approximately one hundred targets (Brennecke et al. 2005). MiRNA profiling combined with a solid target database represents an attractive strategy to identify novel nosological entities, activated pathways in cancer cells, and new targets for research, diagnosis or clinical use (Betel et al. 2008; Friedman et al. 2009; Krek et al. 2005).

MiRNA expression profiles are usually assessed using microarray hybridization and high-throughput quantitative real-time PCR (qPCR) (Jing et al. 2010). MiRNA profiling using microarray hybridization allows the screening of all known miRNAs at once, but their limitation is that probes cannot distinguish mature miRNAs from miRNA precursors, and accurate microarray profiling requires their separation by polyacrylamide gel electrophoresis (Shingara et al. 2005). Results obtained through microarray hybridization are expressed in relative intensities and need to be validated by qPCR assays (Morey et al. 2006).

Conversely, qPCR-based miRNA expression profiling relies on the 384 wells format, with a simplified sample preparation and loading process. Therefore, the number of simultaneous miRNAs that can be studied in a single experiment is lower than the microarray-based technique. QPCR experiments are often presented as fold changes over a control sample.

Because of their specific technologies, these two platforms provide non comparable results, and cross-platform analysis can be challenging (Wang et al. 2011). Moreover, there is still no consensus on the optimal normalization method to be used for miRNA profiling studies (Meyer et al. 2010).

MiRNAs are assumed to be less prone to degradation than other longer nucleic acid materials. It has already been shown that qPCR is suitable for miRNA quantification on single assays (Siebolts et al. 2009), and that FFPE sample could give interpretable results using the TLDA technology (Hui et al. 2009; Sotolongo et al. 2010). When FFPE samples are compared to

frozen samples, a bias varying along the Ct values, i.e. a local bias, is visible on correlation plots (Goswami et al. 2010; Steg et al. 2006). However, no correction method has been successfully applied so far to improve the reliability of results. The ability to use FFPE tumor samples would open up large tumor collections from clinical trials and from other tumor bank collection centers (Gnanapragasam 2010). Their analysis might enable researchers to identify new targets that are linked to disease progression or response to cancer therapy (Oulas et al. 2009; Trang et al. 2008).

As any emerging technology, there are few studies evaluating the key parameters influencing TLDA results and showing how to analyze optimally these datasets (Gillis et al. 2007). Scientists often use the standard $\Delta\Delta\text{Ct}$ method with no consensus concerning the choice of the endogenous housekeeping gene(s), and limiting themselves to fold change analyses (Flamant et al. 2010; Mor et al. 2011; Packer et al. 2008; Tatro et al. 2010). Commercial software solutions have emerged to enable TLDA analysis (Sorby et al. 2010; Zhao et al. 2012). These solutions implement several analysis tools related to qPCR normalization. However, selecting the parameters within the software is at the discretion of the user, and inappropriate selections can lead to non consistent results (Chugh and Dittmer 2012). Moreover, normalization algorithms are applied globally, over the whole dataset, and local biases can remain uncorrected with these approaches (Steg et al. 2006).

We propose in the present work a standardized approach for the analysis of TLDA for miRNA raw data. QPCR will be processed using the $\Delta\Delta\text{Ct}$ method, with a carefully selected optimal endogenous housekeeping gene. We will also investigate here the best significant way to normalize TLDA data derived from a heterogeneous set of samples, composed of FFPE and frozen samples, using tools originally designed for microarray data.

Results:

The main objective was to establish a standard and reproducible analysis methodology to obtain reliable miRNA expression profiles, starting from TLDA raw data. The aim is the production of data with high quality, directly available for statistical analyses such as unsupervised or supervised clusterings. This methodology should be applicable to frozen, FFPE or mixed frozen/FFPE cohorts. The entire cohort is composed of a variety of sarcoma entities (Table 1). The evaluation of the normalization methods is based on their ability to accurately classify the different samples.

➤ **Obtaining the cleanest dataset: raw data extraction, filtering and quality control**

The method used to collect the raw data was found to be a potential source of bias. Using an automatic threshold, Ct values obtained for a given GIST sample processed alone differed from the Ct values obtained for the same sample extracted as a part of different groups of five distinct GIST samples (Figure 1A). In this example, Ct values of three representative miRNAs of the GIST sample #F1728 are presented. The first miRNA shows no variation, the second one a limited variation and the third one the maximum variation observed. The maximum variation was of 7.72 cycles for the miR-628-5p. Among expressed miRNAs, 10 % of Ct values showed no variation, but the variation of Ct values was greater than 2 cycles for 18% of Ct Values (Figure 1B). A plate by plate extraction was found to be the best way to retrieve raw data, which were then not affected by the constitution of the analysis group. Subsequently, each sample can be easily included in various cohorts. However, in this case, a manual threshold must be chosen, since the use of an automatic threshold on a single plate produced an unusual stepwise distribution of Ct values, due to algorithm approximations (Figure 1C). The threshold for Ct determination was set up as close as possible to the base of the exponential phase. Following the visualization of amplification curves for more than a hundred of TLDA arrays, a manual threshold of 0.25 was chosen. All Ct values were extracted with this threshold and an automatic baseline.

For each TLDA, quality controls were performed on raw data (see Materials and Methods). Box plots, the following of the expression level of the MammU6 internal control, and heat maps based on Pearson's correlation coefficients were sufficient to detect samples with an abnormal data distribution or with unusual overall late Ct values (Figure 2A and 2B). Inconsistent samples were removed to avoid any biases during the normalization procedures.

Among all tumor samples, 2 out of 104 (2%) FFPE samples and 4 out of 13 frozen (32%), overall 5% of all samples, did not pass these criteria and were removed from the analysis process.

Commercial analysis software products often allow users to allocate an arbitrary value to undetermined Cts or Cts above a chosen cut-off before any calculations (Xia M. et al. 2010). Using this approach, a miRNA whose expression was undetermined for several samples may end up with different expression levels after applying the $\Delta\Delta\text{Ct}$ method described thereafter (Table 2). Thereby, undetermined Ct should not be changed into arbitrary Ct values at this step. At the end of the analysis procedure, a miRNA which is non-expressed across several samples would show artefactual variations of no biological significance; this could influence final results, especially if there are a high number of undetermined or late Cts in samples. The arbitrary value for undetermined or late Cts must be given after the normalization by the endogenous housekeeping gene(s).

Overall, these results show that the extraction of raw data is a critical preanalytic step for frozen and FFPE samples. The proposed method prevents the emergence of biases.

➤ **Choosing the optimal endogenous housekeeping gene(s) for the $\Delta\Delta\text{Ct}$ method**

Prior to the processing of Ct values with the global $\Delta\Delta\text{Ct}$ normalization method, the absence of any potential local bias on TLDA data had to be verified on MA plot (see Materials and Methods). Raw Ct datasets coming from all 102 FFPE samples were plotted against each other (see example Figure 3A). No local biases were detected, as shown by scatter plots being always straight and horizontal. The $\Delta\Delta\text{Ct}$ method can therefore be used on FFPE-only cohorts as the only and sufficient normalization method.

The next step was to optimize the $\Delta\Delta\text{Ct}$ normalization method. To that aim, the 102 FFPE samples composed of 25 rhabdomyosarcomas, 21 osteosarcomas and 56 GISTs were used. Different sizes of training sets within the cohort of 102 samples were tested (random samplings, see Materials and Methods). The main issue concerning the $\Delta\Delta\text{Ct}$ method is the need to determine the optimal endogenous housekeeping gene(s). Four different methods for the choice of the endogenous housekeeping gene(s) were evaluated on these different training sets: 1/ the normalization factor MammU6, 2/ normalization using the geometric mean of the

four most stable miRNAs identified by geNorm or Normfinder, and 3/ normalization by the sample medians.

As a performance indicator, the ability of these four methods to accurately classify samples according to histological diagnosis was investigated, using unsupervised clusterings. To compare the results of the classification algorithm, confusion rates and purity indexes were calculated (see Material and Methods). Overall, the curves generated showed that the purity index is always better in large cohorts, while the confusion rate is more stable. We describe hereunder the results obtained with the four methods (Figure 4A and 4B).

1/ Single gene control method: MammU6 endogenous control

The MammU6 endogenous control produced the worst purity index and confusion rate when compared to the three other methods. Results are particularly inaccurate for small cohorts, as seen with the purity index. Conversely, the confusion rate increases together with the size of the cohort. These opposite results emphasize the fact that the purity index values are different whether misclassified samples are all located in the same cluster, or in several different clusters. In our study, purity index was reduced because of a large amount of errors occurring in the same cluster (e.g. GIST cluster, data not shown). This index is therefore not representative of the overall error rate. Whatever the size of the cohort, the 3 other approaches always outperformed the MammU6 method, probably indicating that the expression of MammU6 was varying across samples.

2/ GeNorm and Normfinder methods

GeNorm and Normfinder algorithms have to be run again for each training sets, because the most stable miRNAs can change depending on the size and the composition of the cohort. Using both methods and studying the same subsets of samples within two different cohort sizes with two random drawings, the selected miRNAs varied substantially (Table 3). The two methods never retrieved the exact same group of stable miRNAs, and similar miRNAs were always differently ranked. Even for different drawings within a training set, the retrieved miRNA normalizers were found different. Interestingly, using the geNorm method, the expression of MammU6 was found to be stable in some training sets, as shown by its presence in the table.

The analysis of purity indexes and confusion rates shows that the Normfinder method performs better, especially for small cohorts. When using cohorts containing more than 50

samples, the efficacy of the Normfinder method is very similar to that of the median described hereunder. However, the disadvantage of the Normfinder method is that it requires a prior knowledge of the constitution of the groups to determine the best miRNA normalizers, which is an unknown factor in most discovery studies.

3/ Median method

With this approach, the median is calculated for each sample and is considered as a virtual endogenous housekeeping gene. The median curves of the purity index and the confusion rate can be considered more stable, regardless of the size of the cohort, yet improving slowly as the size of the cohort increases. Together with Normfinder, the median method performed best with both indicators, regardless of the cohort size. Due to the advantage of the median method for studies where there is no prior knowledge about the constitution of groups, this method was selected as the best compromise.

➤ Normalizing heterogeneous analyses cohorts containing FFPE and frozen samples

We previously showed that it is possible to analyze homogenous FFPE cohorts with the standard $\Delta\Delta\text{Ct}$ method. However, the feasibility of comparing FFPE and frozen samples together in the same cohort had to be evaluated. To that aim, we used a cohort composed of 13 tumors and one rhabdomyosarcoma cell line having both frozen and their FFPE equivalent counterpart (see Materials and Methods). TLDA for miRNAs were run for all samples, and 9 out of these 13 paired tumor samples passed all quality control criteria. The SJCRH30 rhabdomyosarcoma cell line was added to this study for a total of 10 paired samples.

The presence of any potential biases was then assessed. For all paired samples, the two raw data sets were visualized on a MA plot. The example of a leiomyosarcoma sample is shown on Figure 3B. Scatter plot followed an oblique line with a slight curvature, a plot characteristic of the presence of a local bias, varying along the Ct values. A local normalization must be applied to correct this bias prior to the $\Delta\Delta\text{Ct}$ method.

To find out the best method to correct this local bias, the 10 paired samples (9 paired frozen and FFPE tumor sample, and paired samples of an RMS cell line) were analyzed using 3 different approaches : without any local normalization, with the quantile method (Rao et al.

2008), and with the LOESS method (Cleveland WS. 1979; Cleveland WS. and Devlin SJ. 1988). All samples were analyzed using our previously established $\Delta\Delta C_t$ method, with medians as endogenous housekeeping genes. All data modifications due to normalization steps were controlled by box plots and MA plots (Figure 5).

After the quantile normalization of all paired samples, box plots showed exactly the same distribution, with a shared median (Figure 5B1). However, the local biases were only partially corrected using this method, as shown in the MA plot. Moreover, the distribution of the data shows an increased overall spread of dots (Figure 5B2). After LOESS normalization, box plot showed that frozen and FFPE samples were globally corrected by pairs, with pairs having the same median value but a different distribution (Figure 5C1). Using an MA plot, samples correlations were correctly adjusted along the C_t values, so the scatter plot is straight and horizontal around the zero line (Figure 5C2). All samples will share the same median and will be comparable for further analysis after the application of the $\Delta\Delta C_t$ method.

The accuracy of these two local normalization methods were then assessed measuring their ability to pair corresponding frozen and FFPE samples when performing an unsupervised clustering on $\log_2(RQ)$ values (Figure 5A3, B3 and C3). With no local normalization, 8 samples out of 10 were correctly paired: tumors #4 and #12 were split in different groups (Figure 5A3). As expected, samples correctly paired in the absence of local correction exhibited a flat distribution on MA plots (data not shown). The quantile normalization method was able to classify accurately only 7 samples (Figure 5B3). The LOESS normalization performed the best with a perfect recognition of all paired samples (Figure 5C3). Importantly, entities with similar histologies were clustered together, such as liposarcoma samples (WDLPS, DDLPS and US tumors) and rhabdomyosarcoma samples (ARMS cell line and ERMS tumor). In our study, the rhabdomyosarcoma cell line is not separated from fresh tumor samples, in contrast with what was reported in previous molecular profiling studies (Ertel et al. 2006; Houshdaran et al. 2010; Kao et al. 2009; Virtanen et al. 2002).

The LOESS normalization is therefore an appropriate method for the analysis of TLDA data when there is a need for a local bias correction between FFPE and frozen samples.

Discussion:

In the present work, we present a novel analysis methodology for miRNA profiling using TLDA, in the third largest cohort ever reported for miRNA expression profiling on the TLDA platform (104 FFPE samples) (Nakata et al. 2011; Schultz et al. 2012). The objectives of the present work was to delineate a method to enable a reliable, reproducible, evaluation of miRNA expression profile among a variety of tumor samples, with the aim of a potential clinical use.

TLDA data are generally analyzed as usual qPCR results, using the $\Delta\Delta C_t$ method as the sole normalization method, but with no consensus about the optimal endogenous housekeeping gene(s). In the present report, we observe that the use of gross and heterogeneous methods may significantly affect the reproducibility of the results across laboratories using TLDA platforms (Wang et al. 2011; Wang et al. 2012). The management of the quality of the data was found to be essential at all steps of the procedure, with the aim to prevent or correct any possible technical or biological bias. The adapted $\Delta\Delta C_t$ method proposed in the present manuscript, using the median as a virtual endogenous housekeeping gene, is applicable to homogeneous cohort of frozen samples or FFPE samples and lead to optimal results in our dataset. For heterogeneous cohorts, composed of FFPE and frozen samples, a local bias was identified, which was found correctible using an additional LOESS normalization.

Usually, for the validity of quantification using the comparative $\Delta\Delta C_t$ method, the efficiency of the target amplification and the efficiency of the active reference amplification must approximately be equal. The median being a virtual endogenous housekeeping gene, it has no quantifiable efficiency, but averages the efficiencies across the whole sample set. In the present work, we show that this median method was found to contribute to very limited and non detectable quantification errors and outperformed all other methods evaluating real endogenous housekeeping gene(s).

The results presented here show in particular that MammU6 cannot be always considered as an endogenous control. MammU6 frequently stands out as a differentially expressed target in several statistical analyses, as shown also in other studies (Peltier and Latham 2008; Tatro et al. 2010). In general, using a single control reference has been reported to be suboptimal in several studies (Huggett et al. 2005; Lee et al. 2002).

Using geNorm and Normfinder, the most stable miRNAs were found to differ significantly in the various training subsets tested. These algorithms have therefore to be run again every time

a change in the composition of samples occurs, in particular with the addition of a novel sample. Conversely, the median is sample-related only, and needs to be calculated once and for all. Choosing the median as a virtual endogenous housekeeping gene is therefore saving many efforts and calculation times (Deo et al. 2011). This easy way of normalizing TLDA data with the $\Delta\Delta C_t$ method is of particular interest for discovery studies in rare pathologies, where a given stable miRNA found in a small cohort might not be representative.

In cancer research, high-throughput platforms are often used to find a small subset of deregulated targets of interest. In order to be cost-effective, the expression level of these targets can then be measured in single assays for diagnostic, prognostic or therapeutic applications. Using the median method recommended here as a virtual endogenous housekeeping gene can no longer be used in single qPCR assays. At the present time, the only solution is to select one or several real endogenous control miRNAs using large cohort studies, constituted with adequate samples. However, the cost of TLDA, which could limit the present approach, is significantly lowering, making this issue considerably less relevant for the near future.

In order to compare FFPE and frozen samples, we directed our efforts on local normalization methods given the presence of local biases on the C_t values. It has already been shown that non-linear normalization methods perform the best for microarray data analysis (Bolstad et al. 2003), and two commonly used methods are used in this context: the LOESS normalization and the quantile normalization methods. In the present work, the LOESS normalization method was found to be superior to the quantile normalization method for TLDA data correction. The LOESS function actually fits a smoothing curve to a dataset. This represents an advantage for miRNA profiling, as the family of known miRNAs represents a relatively small dataset: an adjustable smooth parameter in the LOESS function allows users to work with a few hundreds of targets. Besides, LOESS was also proven to be efficient with miRNA microarrays, containing much fewer spots than whole-genome gene expression microarrays (Hua et al. 2008). Conversely, the quantile normalization was originally designed for the accurate correction of expression values of thousands of targets. It is a more “aggressive” procedure than LOESS, because it redistributes the data in one data set to match the distribution of another data set. Quantile normalization is therefore less appropriate in small datasets.

In this work, the use of paired frozen and FFPE samples allowed us to use a simple LOESS normalization method. When working on various non-related FFPE and frozen sample, the cyclic LOESS has to be used instead as there is no matched reference sample (Yang et al. 2002). This generalization of the global LOESS method normalizes values for a set of samples by working in a pairwise manner. Alternatively, it should be possible to use more recent algorithms based on LOESS smooths (Ballman et al. 2004).

We used a cohort composed of frozen and FFPE samples, but a lot of other non tested sampling conditions could also require this LOESS normalization. Ct values could be locally disturbed by therapeutic treatments of patients, sample decalcifications, and fixation methods used before the embedding of samples in paraffin. Therefore, an assessment of biases using MA plot should always be performed when samples with different treatment and storage conditions are analyzed. In the present study, the complexity was limited by using samples sharing the same formaldehyde conservation method, even if they were processed in several different hospitals.

Data treatment methods were evaluated according to their performances during clustering, based on the assumption that tumors were initially optimally classified by pathologists using morphological and immunophenotypical criteria, the institution being a reference center for sarcoma pathology in France. Evaluating the performance of a normalization method on a well-defined real data set and clustering analyses has already been proven to be particularly efficient with biological data (Kim S.Y. et al. 2006). A superior cluster robustness and accuracy can be considered as the best biological proof of an unbiased normalization. Using this strategy, we validated a novel analysis methodology for miRNA profiling using TLDA: when there is no local bias in the raw data, the $\Delta\Delta\text{Ct}$ method using the median as the endogenous housekeeping gene is the best approach to obtain accurate results. If a local variation is detectable in the raw data, a prior LOESS normalization enables a good correction of this particular bias. This analysis methodology is therefore virtually applicable to cohorts composed of homogenous or non homogenous biological material. This is of particular relevance in rare tumors, for retrospective or prospective cohorts gathering samples from various origins, with different preparations as well as storage.

This novel analysis methodology was successfully applied within our group to improve the management of patients with sarcoma. MiRNA profiles generated by this analysis

methodology were used to identify a miRNA signature that can predict therapeutic responses in osteosarcomas (Gougelet et al. 2011b). The set of discriminating miRNAs is consistent with other published studies, targeting signaling pathways involved in bone resorption and osteogenesis, such as the TGF β and the Wnt pathways, and other various pathways implicated in immune response (Kingsley et al. 2007; Mintz et al. 2005; Salas et al. 2009). It was also shown that miRNA profiles obtained through this methodology have a diagnostic value, and may guide nosological classification when standard molecular diagnostic criteria are not enough differentiating. MiRNA profiles provide further information, especially for rhabdomyosarcoma fine classification (Gougelet et al. 2011a). The use of clustering analysis for qPCR data is a great advantage over simple fold-change analysis. The sharing of miRNA expression databases resulting from such studies, on well characterized samples, will open to novel, refined understanding of rare pathologies (Sarver et al. 2010; Subramanian et al. 2008). Acting at a different level, miRNA expression profiles can therefore be used in the field of cancer research to explore another layer of complexity on top of other high throughput technologies like CGH, mRNA expression microarrays or protein arrays (Subramanian and Kartha 2012).

Finally, another application could be the quantification of miRNA expressions in a non invasive way for the patient, starting from human serum samples (Zampetaki et al. 2012). Recent advances have been made to detect small quantities of miRNAs with an alternative technology (Chapin and Doyle 2011).

Overall, the work described enables to define a robust methodology to explore miRNA expression profiles in FFPE samples, as well as in tumor samples with different conservation techniques. This method is reliable, and particularly relevant to compare expression profiles across tumor samples in rare tumors, where retrospective collection of samples stored with different conditions is a major limitation to investigate expression profile.

Materials and Methods:

➤ Tumor samples selection

Frozen and FFPE tumor samples were obtained from the department of Biopathology at the Centre Léon Bérard in Lyon, the La Timone hospital in Marseille and the Ambroise Paré hospital in Boulogne. A total of 117 tumors were collected. The cohort was composed of three major groups of FFPE sarcoma samples: 21 osteosarcomas, 26 rhabdomyosarcomas and 57 gastrointestinal stromal tumors (GISTs). In addition, 13 tumors having both frozen and FFPE samples were randomly picked out among a variety of sarcoma subtypes, including sarcomas with simple genomics such as rhabdomyosarcoma, synovial sarcoma, liposarcoma, myxoid liposarcoma, and sarcomas with complex genomics such as chondrosarcoma, leiomyosarcoma, osteosarcoma, and unclassified sarcoma (Table 1). A rhabdomyosarcoma cell line was also selected to be frozen and formalin fixed, paraffin embedded (SJCRH30 cell line, ATCC# CRL-2061).

➤ RNA extraction

Frozen tumor fragments conserved in liquid nitrogen were cryo-grounded. Total RNA was extracted from tissue powders using a single phenol/chloroform extraction protocol with Trizol®, according to the manufacturer's instructions (Invitrogen, Carlsbad, CA). RNA pellets were dissolved in 20 µl of RNase-free water and stored at -80°C.

FFPE samples were cut into approximately 30 x 4 µm sections and placed in a microcentrifuge tube. Deparaffinization was performed using xylene followed by ethanol and buffer TE washes. After centrifugation and removal of the buffer, tissues were placed in a tissue lysis buffer with proteinase K (Qiagen) and rotated at 65 °C for 24 to 72 h as needed for full digestion. Subsequent total RNA extraction and purification was carried out with a double phenol/chloroform extraction separated by a DNase step. RNA pellets were resuspended in 20 µl of RNase-free water and stored at -80°C.

No enrichment for small RNAs was made, as recommended per the manufacturer.

➤ Multiplex reverse transcription

All RNA samples were first diluted using RNase-free water, in order to input 500 ng of total RNA into the Megaplex RT reactions. Briefly, each sample of total RNA is reverse transcribed using two Megaplex pools of RT primers (A and B pools corresponding to A and B cards), and the Taqman® MicroRNA Reverse Transcription kit. The RT master mix is prepared as per the manufacturer's instructions. For each tube, 3 µl of diluted RNA solution is

added to 4.5 µl of the RT reaction mix. Thermal cycler conditions are set as follows: 40 cycles with 2 min at 16°C, 1 min at 42 °C and 1 sec at 50 °C, followed by a 5 min step at 85°C. The resulting cDNA samples were used immediately for subsequent TLDA analyses or stored at -20°C.

➤ **Quantitative real-time PCR using TLDA**

All qPCR experiments were run with the TLDA Human MicroRNA Panel v2.0, which allowed the analysis of 365 and 290 mature human miRNAs, respectively on the A and B cards. No pre-amplification step was added. 6 µl of RT products are added to 444 µl of nuclease-free water and 450 µl of the TaqMan Universal PCR Master Mix (Life Technologies). Final volumes are introduced through the height loading ports of the TLDA plate. By centrifugation and thanks to micro-channels, mixtures reach miniature wells that are preloaded with dehydrated specific primers and probes. The card is then sealed and introduced in the thermal cycler upon the software-locked conditions recommended by the manufacturer for miRNA amplifications.

➤ **qPCR data extraction**

For each miRNA, the threshold cycle (Ct) was calculated plate by plate thanks to the ABI 7900 Sequence Detection System Software v2.4 and the RQ manager v1.2.1 (Life Technologies). Along with the Ct values, flags based on fluorescence raw data were also retrieved in a tabulated text file (Woo DC. and Subramaniam Y. 2008). Containing valuable information, some of these flags had to be interpreted immediately. Wells flagged with "Empty Well" or "Bad Passive Reference" were considered as missing data, whether they were undetermined or not; in our experiments, it could happen when the TLDA plate was not correctly sealed, inducing leaks of reagents during the amplification process. On the contrary, the "Non-Amplified Well" flags really signified non-expressed targets. Ignoring these flags would induce impaired results, especially if the well is producing an absurd Ct.

In practice, late Ct values of qPCR assays are discarded in order not to take into account hazardous values that would induce analysis bias. According to previously published data using the TLDA technology for miRNA expression profiling, a cut-off of 37 was applied to exclude these late Ct values (Lao K. et al. 2008; Witwer K.W. et al. 2011). All Ct above these cut-offs were considered as undetermined for the rest of the analyses.

➤ qPCR data quality control

The initial quality of raw data was verified for each sample, checking the valid expression of internal controls (small nucleolar RNAs MammU6, RNU44 and RNU48), using box plots and heat maps based on Pearson's correlation coefficients (Altman 1991; Benjamini 1988).

Another quality control, the MA plot, was performed on raw data and after every normalization procedure (Dudoit S. et al. 2002). The MA plot is basically a scaled, 45 degree rotation of the standard correlation plot. On the Y axis M represents the Ct values ratio, and on the X axis A represents the average Ct values of every dot in the plot. MA plots are particularly useful to visualize local biases, i.e. biases varying along the Ct values.

➤ qPCR data global normalization

All data manipulations were done using the R software v2.14.2 (Ihaka R. and Gentleman R. 1996). Samples were normalized in a global way using the standard $\Delta\Delta Ct$ method, with four different approaches for the choice of the endogenous housekeeping gene(s). The first approach consisted in using the Life Technologies internal control gene MammU6, which is repeated four times on each TLDA card and stands for the U6 small nuclear 1 RNA (Accession Number NR_004394). The second and third approaches were the use of the geometric mean of the four most stable miRNAs found by geNorm (Vandesompele et al. 2002) or Normfinder (Andersen et al. 2004). The fourth approach consisted in using of the median, calculated for each sample. When comparing geNorm and Normfinder methods, a subset consisting of four most stable miRNAs was used. This particular set of miRNAs was chosen on the basis of published experience (Vandesompele et al. 2002).

The standard $\Delta\Delta Ct$ method was used to determine the relative quantities (RQ) as follows:

$$\Delta Ct = Ct - [Ct \text{ of the chosen endogenous housekeeping gene(s)}]$$

$$\Delta\Delta Ct = \Delta Ct - [\text{Mean of } \Delta Ct \text{ across samples for each miRNA}]$$

$$RQ = 2^{-(\Delta\Delta Ct)}$$

In this formula, the Ct of the chosen endogenous housekeeping gene(s) is determined by the four approaches we decided to test. This Ct is therefore either real or virtual. For the $\Delta\Delta Ct$ calculation, it was more relevant for the upcoming statistical analyses to use the mean of all ΔCt obtained across samples for each miRNA, instead of using the ΔCt of a reference sample. The fold change being obsolete when there is a need for classification, the calculated RQ values had to be adapted. To obtain symmetrical data around zero, RQ values were

transformed thanks to the logarithmic function. Log2(RQ) values could directly be used in any statistical analysis software dedicated to microarray, for clustering purposes.

➤ **qPCR data local normalization**

In order to prove the utility of a local correction and to find out the best suited method for TLDA data when working with mixed frozen and FFPE samples, paired samples were analyzed using 3 different approaches : without any local normalization, with the quantile or with the LOESS normalization methods. LOESS corrects systematic deviations in the MA plot by carrying out a locally weighted linear regression as a function of the X axis and subtracting the calculated best-fit average ratio from the experimentally observed ratio for each data point (Cleveland WS. 1979; Cleveland WS. and Devlin SJ. 1988). The LOESS function was empirically set with a smooth parameter of 0.8. The quantile normalization is a procedure that redistributes the data in one dataset to match the distribution of another dataset. This is made possible by sorting each sample distribution, and setting each sample value to the mean of the other sample values of the same rank (Bolstad et al. 2003; Rao et al. 2008). Both of these local normalization methods were applied prior to the $\Delta\Delta C_t$ method when needed.

➤ **Statistical analysis**

After data normalization, some quality markers were used to improve subsequent statistical analysis. In our study, miRNAs that were never expressed in our cohorts and miRNAs whose expression was always weak were tagged. Thus, it was possible to work on the most reliable data at first, with the possibility to broaden the analysis to lower quality data if needed. Logarithmic values of normalized RQ data were directly input into the Spotfire® DecisionSite for Functional Genomics analysis software v9.1.1 (TIBCO Software Inc., Somerville, MA) and the R software v2.14.2. We performed unsupervised or supervised hierarchical clustering to classify samples, in order to prove the accuracy or not of the various data treatments. These hierarchical clusterings were done using Euclidean distance and Ward's linkage.

The best way of choosing the endogenous housekeeping gene(s) stood out comparing the ability of all methods to separate three groups of different sarcoma entities in an unsupervised clustering test. Among the 102 samples used for these tests, five subsets of different sizes (20, 31, 51, 61 or 71 tumor samples, respectively around 20%, 30%, 50%, 60% et 70% of the size of the initial cohort) were randomly picked out by drawing without replacement. For each

cohort size, 50 random sampling were analyzed. On every subset, we performed unsupervised clusterings corresponding to the four tested endogenes.

Two cluster validation indicators were calculated to evaluate the accuracy of clusterings in terms of correct classification: the confusion rate (Kohavi R. and Provost F. 1998) and the purity index (Rendon E. et al. 2011). The confusion rate directly represents the percentage of samples that are misclassified. The purity index corresponds to the weighted sum of the number of misclassified tumors in each cluster. These calculations show how different a clustering is from the expected one. In this way, when comparing different normalization approaches, the lowest the confusion rate and the purity index are, the better the method has to be considered.

ACKNOWLEDGMENTS:

This work is supported by the Conticanet Network of Excellence (EU-FP6-06188), the EUROSARC network of clinical trials (EU-FP7-278742), and the LYRIC Integrated Cancer Research Site (INCa_4664). We are very grateful to Corinne Bouvier, M.D., from the La Timone hospital for providing the osteosarcoma tumor samples, Jean-François Emile, M.D., from the Ambroise Paré hospital for providing GIST tumor samples, and to Anne-Valérie Decouvelaere, M.D., from the Centre Léon Bérard hospital in Lyon for providing rhabdomyosarcoma tumor samples and paired frozen and FFPE tumor samples.

REFERENCES:

- Altman, D. G. 1991. *Practical Statistics for Medical Research*. 285-288.
- Andersen, C. L., Jensen, J. L., and Orntoft, T. F. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64: 5245-5250.
- Ballman, K. V., Grill, D. E., Oberg, A. L., and Therneau, T. M. 2004. Faster cyclic loess: normalizing RNA arrays via linear models. *Bioinformatics.* 20: 2778-2786.
- Benjamini, Y. 1988. Opening the box of a boxplot. *The American Statistician* 42: 257-262.
- Betel, D., Wilson, M., Gabow, A., Marks, D. S., and Sander, C. 2008. The microRNA.org resource: targets and expression. *Nucleic Acids Res.* 36: D149-D153.
- Bolstad, B. M., Irizarry, R. A., Astrand, M., and Speed, T. P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* 19: 185-193.
- Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. 2005. Principles of microRNA-target recognition. *PLoS.Biol.* 3: e85-
- Chugh, P. and Dittmer, D. P. 2012. Potential pitfalls in microRNA profiling. *Wiley.Interdiscip.Rev.RNA.* 3: 601-616.
- Cleveland WS. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74: 829-836.
- Cleveland WS. and Devlin SJ. 1988. Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* 83: 596-610.
- Deo, A., Carlsson, J., and Lindlof, A. 2011. How to choose a normalization strategy for mirna quantitative real-time (qpcr) arrays. *J.Bioinform.Comput.Biol.* 9: 795-812.
- Doench, J. G. and Sharp, P. A. 2004. Specificity of microRNA target selection in translational repression. *Genes Dev.* 18: 504-511.
- Dudoit S., Yang YH., Callow MJ., and Speed TP. 2002. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica* 12: 111-139.
- Ertel, A., Verghese, A., Byers, S. W., Ochs, M., and Tozeren, A. 2006. Pathway-specific differences between tumor cell lines and normal and tumor tissue cells. *Mol.Cancer* 5: 55-

- Flamant, S., Ritchie, W., Guilhot, J., Holst, J., Bonnet, M. L., Chomel, J. C., Guilhot, F., Turhan, A. G., and Rasko, J. E. 2010. Micro-RNA response to imatinib mesylate in patients with chronic myeloid leukemia. *Haematologica* 95: 1325-1333.
- Friedman, R. C., Farh, K. K., Burge, C. B., and Bartel, D. P. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 19: 92-105.
- Gillis, A. J., Stoop, H. J., Hersmus, R., Oosterhuis, J. W., Sun, Y., Chen, C., Guenther, S., Sherlock, J., Veltman, I., Baeten, J. et al. 2007. High-throughput microRNAome analysis in human germ cell tumours. *J.Pathol.* 213: 319-328.
- Gnanapragasam, V. J. 2010. Unlocking the molecular archive: the emerging use of formalin-fixed paraffin-embedded tissue for biomarker research in urological cancer. *BJU.Int.* 105: 274-278.
- Goswami, R. S., Waldron, L., Machado, J., Cervigne, N. K., Xu, W., Reis, P. P., Bailey, D. J., Jurisica, I., Crump, M. R., and Kamel-Reid, S. 2010. Optimization and analysis of a quantitative real-time PCR-based technique to determine microRNA expression in formalin-fixed paraffin-embedded samples. *BMC.Biotechnol.* 10: 47-
- Gougelet, A., Perez, J., Pissaloux, D., Besse, A., Duc, A., Decouvelaere, A. V., Ranchere-Vince, D., Blay, J. Y., and Alberti, L. 2011a. miRNA Profiling: How to Bypass the Current Difficulties in the Diagnosis and Treatment of Sarcomas. *Sarcoma.* 2011: 460650-
- Gougelet, A., Pissaloux, D., Besse, A., Perez, J., Duc, A., Dutour, A., Blay, J. Y., and Alberti, L. 2011b. Micro-RNA profiles in osteosarcoma as a predictive tool for ifosfamide response. *Int.J.Cancer* 129: 680-690.
- Griffiths-Jones, S. 2004. The microRNA Registry. *Nucleic Acids Res.* 32: D109-D111.
- Griffiths-Jones, S., Grocock, R. J., van, Dongen S., Bateman, A., and Enright, A. J. 2006. miRBase: microRNA sequences, targets and gene nomenclature. *Nucleic Acids Res.* 34: D140-D144.
- Griffiths-Jones, S., Saini, H. K., van, Dongen S., and Enright, A. J. 2008. miRBase: tools for microRNA genomics. *Nucleic Acids Res.* 36: D154-D158.
- He, L. and Hannon, G. J. 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat.Rev.Genet.* 5: 522-531.
- Houshdaran, S., Hawley, S., Palmer, C., Campan, M., Olsen, M. N., Ventura, A. P., Knudsen, B. S., Drescher, C. W., Urban, N. D., Brown, P. O. et al. 2010. DNA methylation profiles of ovarian epithelial carcinoma tumors and cell lines. *PLoS.One.* 5: e9359-

- Hua, Y. J., Tu, K., Tang, Z. Y., Li, Y. X., and Xiao, H. S. 2008. Comparison of normalization methods with microRNA microarray. *Genomics* 92: 122-128.
- Huggett, J., Dheda, K., Bustin, S., and Zumla, A. 2005. Real-time RT-PCR normalisation; strategies and considerations. *Genes Immun.* 6: 279-284.
- Hui, A. B., Shi, W., Boutros, P. C., Miller, N., Pintilie, M., Fyles, T., McCready, D., Wong, D., Gerster, K., Waldron, L. et al. 2009. Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. *Lab Invest* 89: 597-606.
- Ihaka R. and Gentleman R. 1996. R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* 5: 299-314.
- Jing, H., Song, Q. X., and Zhou, G. H. 2010. [Advances in approaches for the quantitative detection of microRNAs]. *Yi.Chuan* 32: 31-40.
- Kao, J., Salari, K., Bocanegra, M., Choi, Y. L., Girard, L., Gandhi, J., Kwei, K. A., Hernandez-Boussard, T., Wang, P., Gazdar, A. F. et al. 2009. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS.One.* 4: e6146-
- Kim S.Y., Lee J.W., and Bae J.S. 2006. Effect of data normalization on fuzzy clustering of DNA microarray data. *BMC Bioinformatics* 7:
- Kingsley, L. A., Fournier, P. G., Chirgwin, J. M., and Guise, T. A. 2007. Molecular biology of bone metastasis. *Mol.Cancer Ther.* 6: 2609-2617.
- Kohavi R. and Provost F. 1998. Glossary of Terms: Special Issue on Applications of Machine Learning and the Knowledge Discovery Process. *Machine Learning* 30: 271-274.
- Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da, Piedade, I, Gunsalus, K. C., Stoffel, M. et al. 2005. Combinatorial microRNA target predictions. *Nat.Genet.* 37: 495-500.
- Lao K., Xu N.L., Yeung V., Chen C., Livak K.J., and Straus N.A. 2008. Multiplexing RT-PCR for the detection of multiple miRNA species in small samples. *Biochemical and Biophysical Research Communications* 343: 85-89.
- Lee, P. D., Sladek, R., Greenwood, C. M., and Hudson, T. J. 2002. Control genes and variability: absence of ubiquitous reference transcripts in diverse mammalian expression studies. *Genome Res.* 12: 292-297.
- Meyer, S. U., Pfaffl, M. W., and Ulbrich, S. E. 2010. Normalization strategies for microRNA profiling experiments: a 'normal' way to a hidden layer of complexity? *Biotechnol.Lett.* 32: 1777-1788.

- Mintz, M. B., Sowers, R., Brown, K. M., Hilmer, S. C., Mazza, B., Huvos, A. G., Meyers, P. A., Lafleur, B., McDonough, W. S., Henry, M. M. et al. 2005. An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Res.* 65: 1748-1754.
- Mor, E., Cabilly, Y., Goldshmit, Y., Zalts, H., Modai, S., Edry, L., Elroy-Stein, O., and Shomron, N. 2011. Species-specific microRNA roles elucidated following astrocyte activation. *Nucleic Acids Res.* 39: 3710-3723.
- Morey, J. S., Ryan, J. C., and Van Dolah, F. M. 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol.Proced.Online.* 8: 175-193.
- Nakata, K., Ohuchida, K., Mizumoto, K., Kayashima, T., Ikenaga, N., Sakai, H., Lin, C., Fujita, H., Otsuka, T., Aishima, S. et al. 2011. MicroRNA-10b is overexpressed in pancreatic cancer, promotes its invasiveness, and correlates with a poor prognosis. *Surgery* 150: 916-922.
- Oulas, A., Reczko, M., and Poirazi, P. 2009. MicroRNAs and cancer-the search begins! *IEEE Trans.Inf.Technol.Biomed.* 13: 67-77.
- Packer, A. N., Xing, Y., Harper, S. Q., Jones, L., and Davidson, B. L. 2008. The bifunctional microRNA miR-9/miR-9* regulates REST and CoREST and is downregulated in Huntington's disease. *J.Neurosci.* 28: 14341-14346.
- Peltier, H. J. and Latham, G. J. 2008. Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA.* 14: 844-852.
- Rao, Y., Lee, Y., Jarjoura, D., Ruppert, A. S., Liu, C. G., Hsu, J. C., and Hagan, J. P. 2008. A comparison of normalization techniques for microRNA microarray data. *Stat.Appl.Genet.Mol.Biol.* 7: Article22-
- Rendon E., Abundez I.M., Gutierrez C., Diaz Sagal S., Arizmendi A., Quiroz E.M., and Arzate H.E. 2011. A Comparison of internal and external cluster validation indexes. *Proceedings of the 2011 American conference on applied mathematics and the 5th WSEAS international conference on Computer engineering and applications* 158-163.
- Salas, S., Jezequel, P., Champion, L., Deville, J. L., Chibon, F., Bartoli, C., Gentet, J. C., Charbonnel, C., Gouraud, W., Voutsinos-Porche, B. et al. 2009. Molecular characterization of the response to chemotherapy in conventional osteosarcomas: predictive value of HSD17B10 and IFITM2. *Int.J.Cancer* 125: 851-860.
- Sarver, A. L., Phalak, R., Thayanithy, V., and Subramanian, S. 2010. S-MED: sarcoma microRNA expression database. *Lab Invest* 90: 753-761.

- Schultz, N. A., Andersen, K. K., Roslind, A., Willenbrock, H., Wojdemann, M., and Johansen, J. S. 2012. Prognostic MicroRNAs in Cancer Tissue from Patients Operated for Pancreatic Cancer-Five MicroRNAs in a Prognostic Index. *World J.Surg.*
- Shingara, J., Keiger, K., Shelton, J., Laosinchai-Wolf, W., Powers, P., Conrad, R., Brown, D., and Labourier, E. 2005. An optimized isolation and labeling platform for accurate microRNA expression profiling. *RNA*. 11: 1461-1470.
- Siebolts, U., Varnholt, H., Drebber, U., Dienes, H. P., Wickenhauser, C., and Odenthal, M. 2009. Tissues from routine pathology archives are suitable for microRNA analyses by quantitative PCR. *J.Clin.Pathol.* 62: 84-88.
- Sorby, L. A., Andersen, S. N., Bukholm, I. R., and Jacobsen, M. B. 2010. Evaluation of suitable reference genes for normalization of real-time reverse transcription PCR analysis in colon cancer. *J.Exp.Clin.Cancer Res.* 29: 144-
- Sotolongo, B., Asaoka, T., Island, E., Carreno, M., Delacruz, V., Cova, D., Russo, C., Tryphonopoulos, P., Moon, J., Weppeler, D. et al. 2010. Gene expression profiling of MicroRNAs in small-bowel transplantation paraffin-embedded mucosal biopsy tissue. *Transplant.Proc.* 42: 62-65.
- Steg, A., Wang, W., Blanquicett, C., Grunda, J. M., Eltoum, I. A., Wang, K., Buchsbaum, D. J., Vickers, S. M., Russo, S., Diasio, R. B. et al. 2006. Multiple gene expression analyses in paraffin-embedded tissues by TaqMan low-density array: Application to hedgehog and Wnt pathway analysis in ovarian endometrioid adenocarcinoma. *J.Mol.Diagn.* 8: 76-83.
- Subramanian, S. and Kartha, R. V. 2012. MicroRNA-mediated gene regulations in human sarcomas. *Cell Mol.Life Sci.*
- Subramanian, S., Lui, W. O., Lee, C. H., Espinosa, I., Nielsen, T. O., Heinrich, M. C., Corless, C. L., Fire, A. Z., and van de Rijn, M. 2008. MicroRNA expression signature of human sarcomas. *Oncogene* 27: 2015-2026.
- Tatro, E. T., Scott, E. R., Nguyen, T. B., Salaria, S., Banerjee, S., Moore, D. J., Masliah, E., Achim, C. L., and Everall, I. P. 2010. Evidence for Alteration of Gene Regulatory Networks through MicroRNAs of the HIV-infected brain: novel analysis of retrospective cases. *PLoS.One.* 5: e10337-
- Tie, J. and Fan, D. 2011. Big roles of microRNAs in tumorigenesis and tumor development. *Histol.Histopathol.* 26: 1353-1361.
- Trang, P., Weidhaas, J. B., and Slack, F. J. 2008. MicroRNAs as potential cancer therapeutics. *Oncogene* 27 Suppl 2: S52-S57.

- Vandesompele, J., De, Preter K., Pattyn, F., Poppe, B., Van, Roy N., De, Paepe A., and Speleman, F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3: RESEARCH0034-
- Virtanen, C., Ishikawa, Y., Honjoh, D., Kimura, M., Shimane, M., Miyoshi, T., Nomura, H., and Jones, M. H. 2002. Integrated classification of lung tumors and cell lines by expression profiling. *Proc.Natl.Acad.Sci.U.S.A* 99: 12357-12362.
- Wang, B., Howel, P., Bruheim, S., Ju, J., Owen, L. B., Fodstad, O., and Xi, Y. 2011. Systematic evaluation of three microRNA profiling platforms: microarray, beads array, and quantitative real-time PCR array. *PLoS.One.* 6: e17167-
- Wang, B., Zhang, S. G., Wang, X. F., Tan, M., and Xi, Y. 2012. Testing for differentially-expressed microRNAs with errors-in-variables nonparametric regression. *PLoS.One.* 7: e37537-
- Witwer K.W., Sarbanes S.L., Liu J., and Clements J.E. 2011. A plasma microRNA signature of acute lentiviral infection: biomarkers of central nervous system disease. *AIDS* 25: 2057-2067.
- Woo DC. and Subramaniam Y. 2008. Automated quality control method and system for genetic analysis. United States Patent
- Xia M., Sherlock J., Hegerich P., You X., Lee K., Walworth C., and Spier E. 2010. DataAssist™ - Data Analysis Software for Taqman® Real-Time PCR Data. *Proceedings of the International MultiConference of Engineers and Computer Scientists* 2010
- Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J., and Speed, T. P. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 30: e15-
- Zhang, B., Pan, X., Cobb, G. P., and Anderson, T. A. 2007. microRNAs as oncogenes and tumor suppressors. *Dev.Biol.* 302: 1-12.
- Zhao, L., Morgan, T. E., Mao, Z., Lin, S., Cadenas, E., Finch, C. E., Pike, C. J., Mack, W. J., and Brinton, R. D. 2012. Continuous versus cyclic progesterone exposure differentially regulates hippocampal gene expression and functional profiles. *PLoS.One.* 7: e31267

Type of information	Screening cohort	Remaining samples after QC
Number of tumors	117	111
Histological subtypes		
Gastrointestinal stromal tumor (GIST)	57	56
Osteosarcoma (OS)	22	21
Rhabdomyosarcoma (ERMS and ARMS)	27	26
Unclassified sarcoma (US)	2	1
Dedifferentiated liposarcoma (DDLPS)	1	1
Well differentiated liposarcoma (WDLPS)	1	1
Myxoid liposarcoma (MLPS)	1	1
Chondrosarcoma (CS)	1	0
Malignant fibrous histiocyoma (MFH)	2	1
Synovialosarcoma (SS)	1	1
Malignant peripheral nerve sheath tumors (MPNST)	1	1
Leiomyosarcoma (LMS)	1	1
Conservation method		
FFPE samples (GIST + OS + RMS)	104	102
Paired samples (Frozen / FFPE)	13	9

Table 1. Summary of tumors characteristics: number, histological subtypes and conservation method. A total of 111 tumors were included in the analyses, consisting of three major subtypes: gastrointestinal stromal tumor (GIST), osteosarcoma (OS) and rhabdomyosarcoma (RMS).

A

miRNA name	F1728 alone	F1728 (cohort #1)	F1728 (cohort #2)
hsa-miR-18b	33.96	33.96	33.96
hsa-miR-17	26.96	26.57	25.96
hsa-miR-628-5p	33.98	26.26	31.84

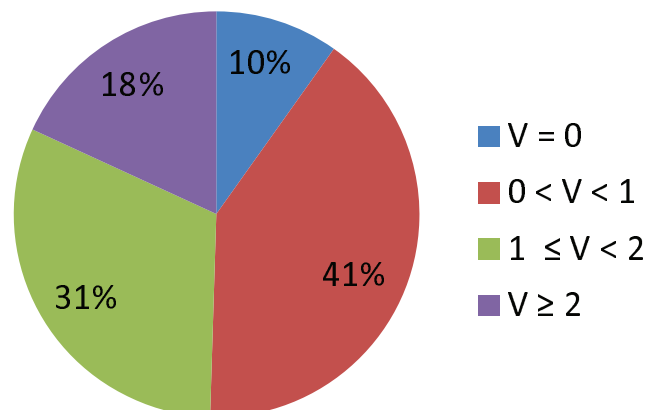
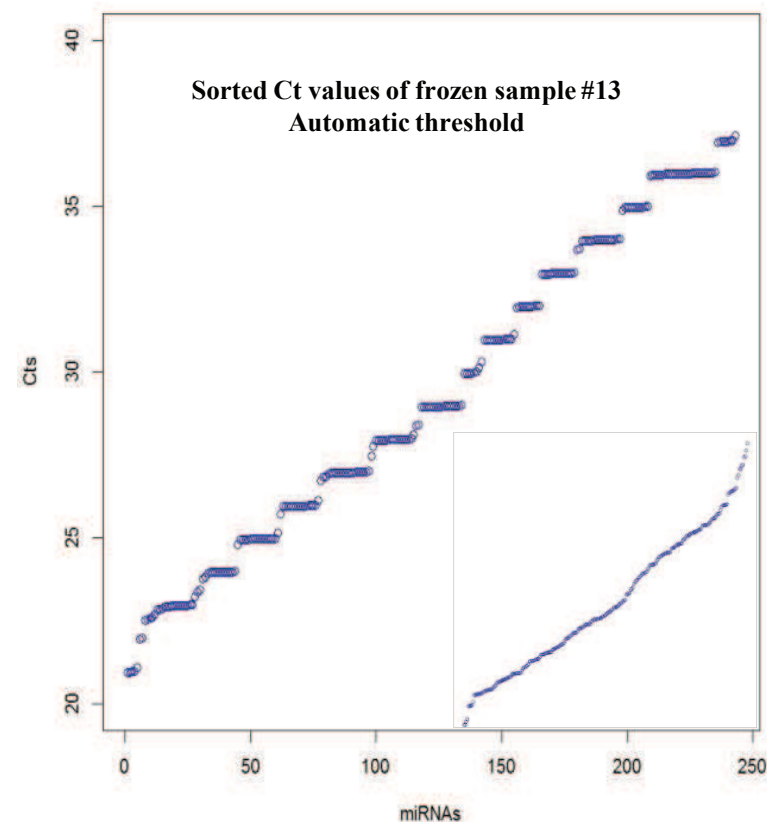
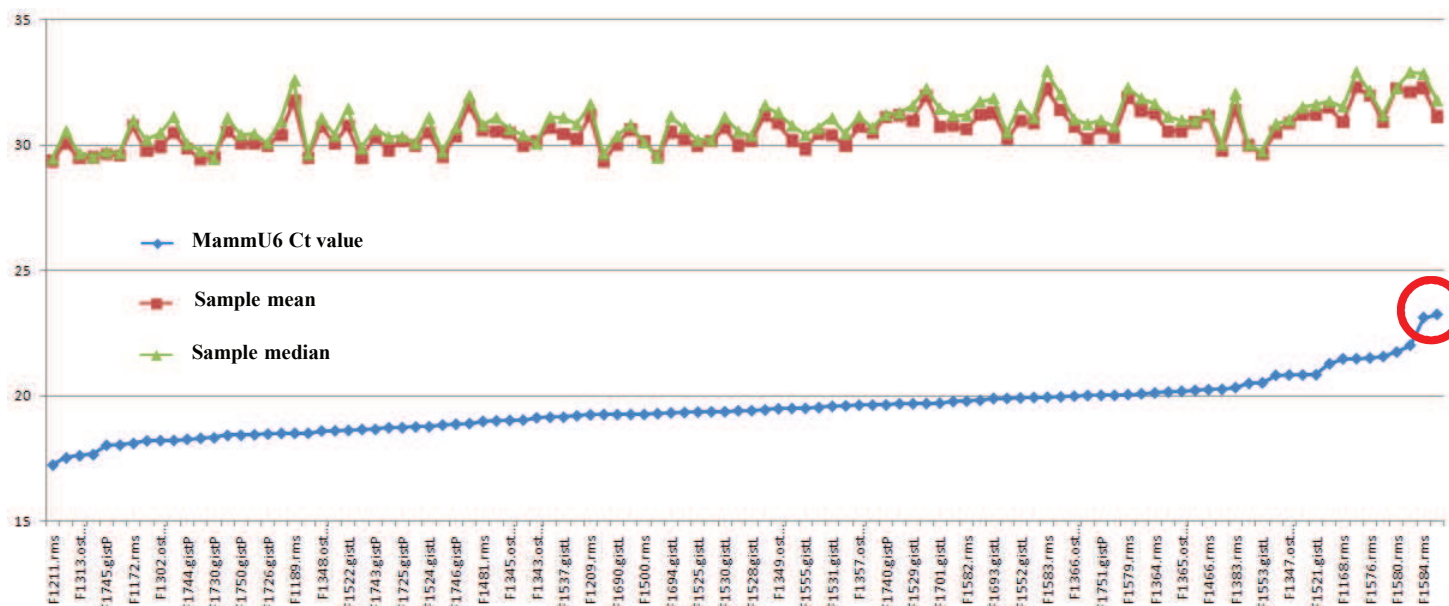
B**C**

Figure 1. Extraction of TLDA raw data from the software. (A) Example of raw Ct values obtained for one GIST sample extracted with an automatic threshold calculation, alone or as a part of two different cohorts of five samples. (B) Histogram representing the major variations (V) of the Ct Values observed in the F1728 GIST sample. Values are expressed in percentage over the total number of expressed miRNA. (C) Distribution curve of the Ct values obtained for the frozen sample #13, used here as an example; the TLDA was extracted alone, using an automatic threshold. An example of an usual distribution curve is shown in the insert at the bottom right corner of the figure.

A



B

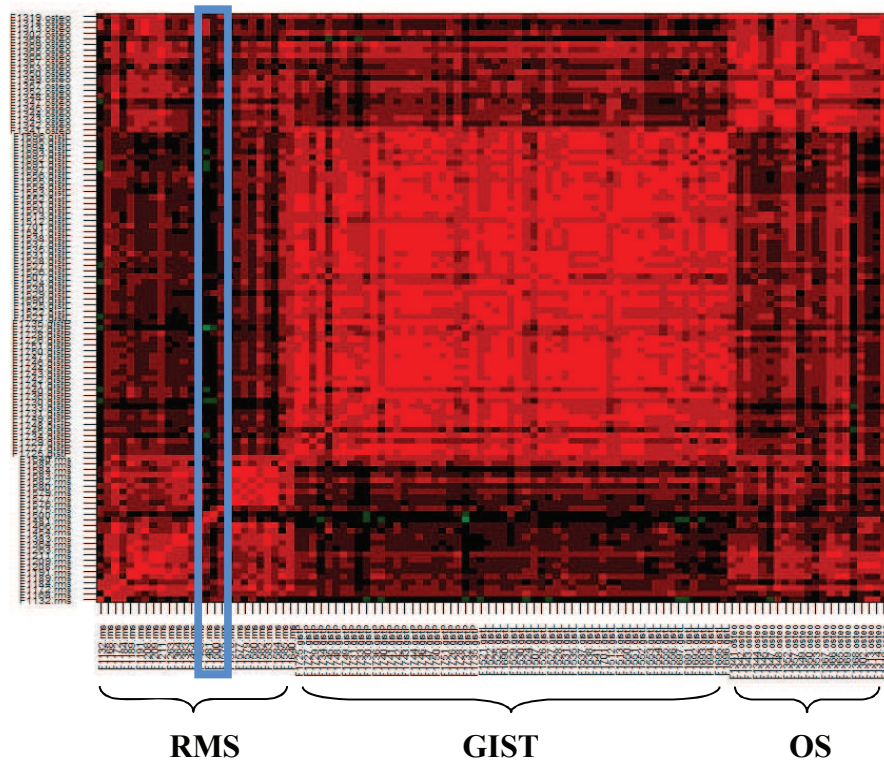


Figure 2. Quality controls of TLDA raw data performed on the 104 FFPE samples. (A) Graphics representing the expression of the mean, the median and the MammU6 internal control. Two samples were removed because of a late MammU6 expression (highlighted in red) (B) Heat map based on Pearson's correlation coefficients representing the correlation of the all samples with each other. The same two inconsistent samples stood out and were removed from the analysis (highlighted in blue).

Calculations	Frozen sample #4	Frozen sample #3	Frozen sample #5	Frozen sample #6
Raw Ct hsa-miR-133a	31,14	Undet	Undet	Undet
Undetermined Ct for hsa-miR-133a are replaced by an arbitrary value	31,14	40	40	40
Δ Ct hsa-miR-133a	2,70	12,06	12,82	14,46
$\Delta\Delta$ Ct hsa-miR-133a	-0,51	8,84	9,60	11,24

Table 2. Management of undetermined Ct values. This example shows the evolution of Ct values during the $\Delta\Delta$ Ct procedure, and the effect of changing undetermined Ct values prior to the Δ Ct step. The three Ct values that were undetermined at the beginning of the procedure end up with three different $\Delta\Delta$ Ct values with obviously no biological significance

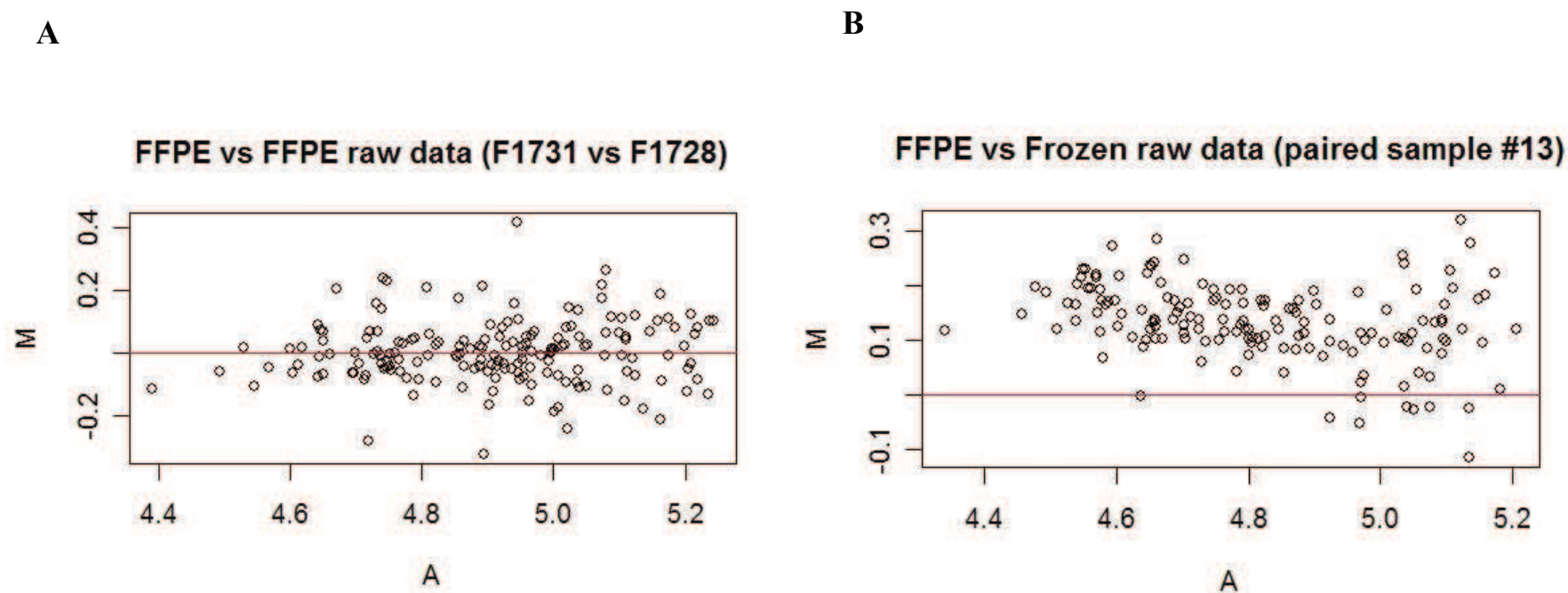


Figure 3. Assessment of the nature of the bias using MA plots. (A) Example of a MA plot of two distinct FFPE samples (F1731 versus F1728). The scatter plot is straight and horizontal, reflecting the absence of any local bias. (B) Example of a MA plot of a paired frozen and FFPE sample (Leiomyosarcoma, frozen sample #13 versus FFPE sample #13). The scatter plot follows an oblique line with a slight curvature, reflecting the presence of a local bias.

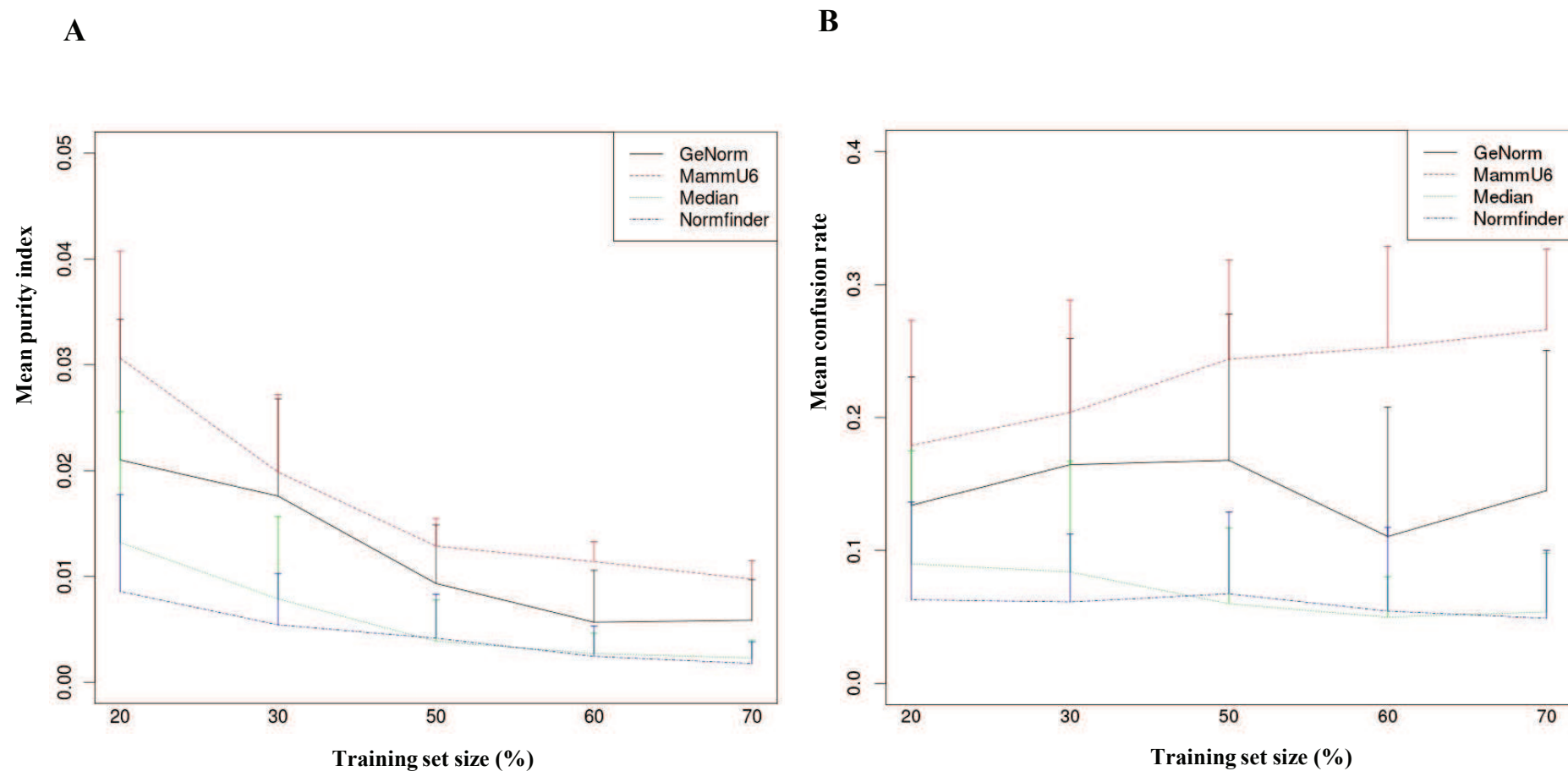
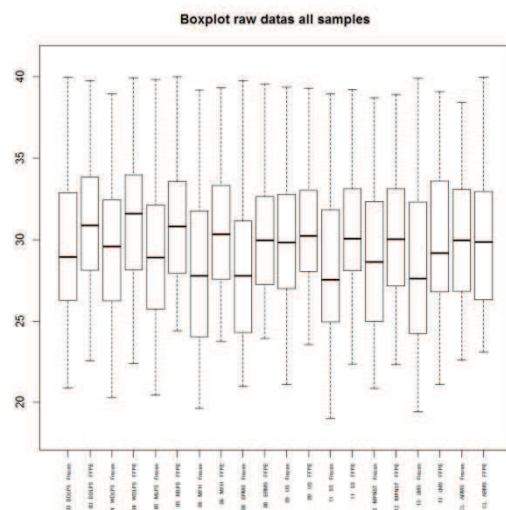


Figure 4. Evolution of two performance indicators, according to the size of the training set. (A) Evolution of the purity index. Overall, the purity index is always better in large cohorts. The median and the Normfinder methods give the best results. MammU6 obtains the worst results. (B) Evolution of the confusion rate.

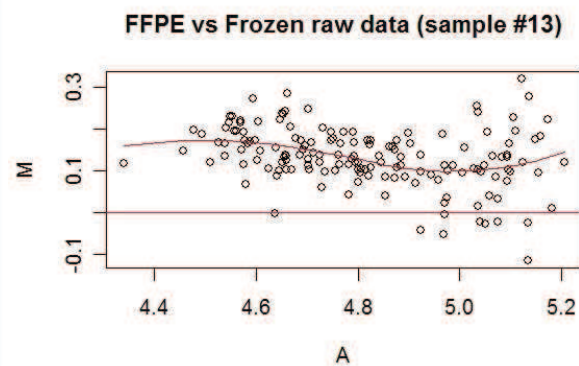
Method & rank	20% panel, first random sampling	20% panel, second random sampling	60% panel, first random sampling	60% panel, second random sampling
geNorm				
miR 1	hsa-miR-125a-5p	hsa-let-7g	hsa-miR-191	MammU6.1
miR 2	hsa-miR-191	hsa-miR-26b	hsa-miR-331-3p	MammU6.3
miR 3	hsa-miR-331-3p	hsa-miR-103	hsa-miR-324-3p	hsa-miR-191
miR 4	hsa-miR-422a	hsa-miR-331-3p	hsa-miR-28-3p	hsa-miR-331-3p
Normfinder				
miR 1	hsa-miR-191	hsa-miR-331-3p	hsa-miR-331-3p	hsa-miR-324-3p
miR 2	hsa-miR-28-5p	hsa-miR-103	hsa-miR-324-3p	hsa-miR-331-3p
miR 3	hsa-miR-125a-5p	hsa-miR-324-3p	hsa-miR-191	hsa-miR-125a-5p
miR 4	hsa-miR-331-3p	hsa-let-7g	hsa-miR-342-3p	hsa-miR-191

Table 3. Analysis of geNorm and Normfinder results in four distinct training sets. The four most stable miRNA retrieved with both methods are presented for each training set. This table shows an example of two random sampling on 2 different sizes of training set. The four most stable miRNA are different, and differently ranked.

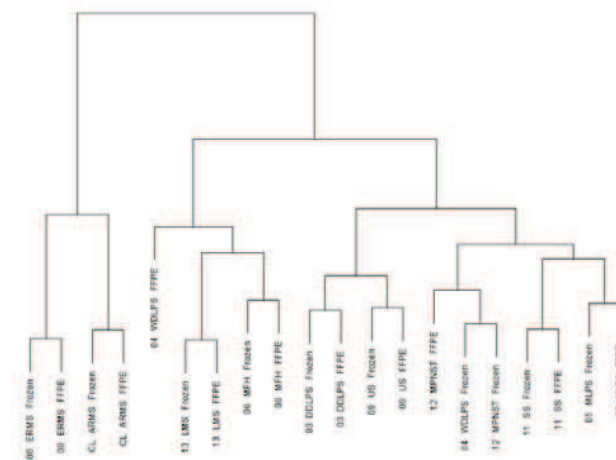
A1



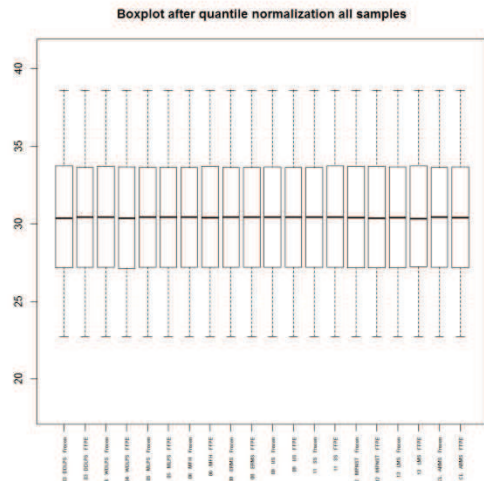
A2



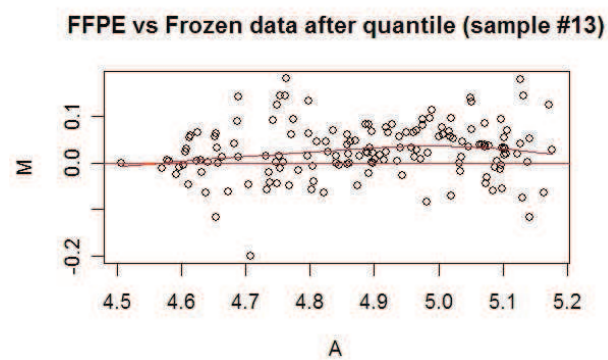
A3



B1



B2



B3

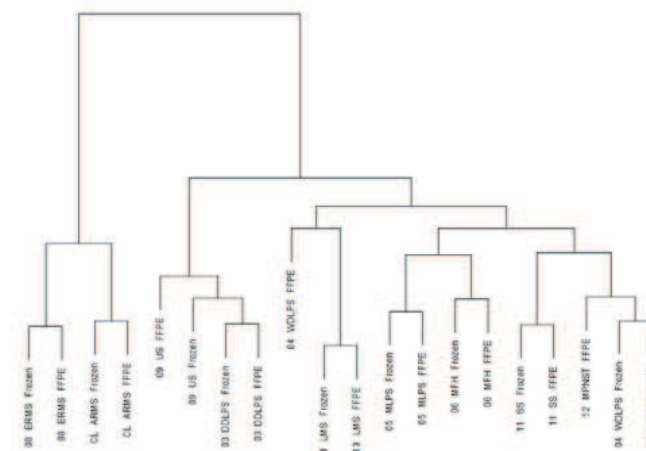
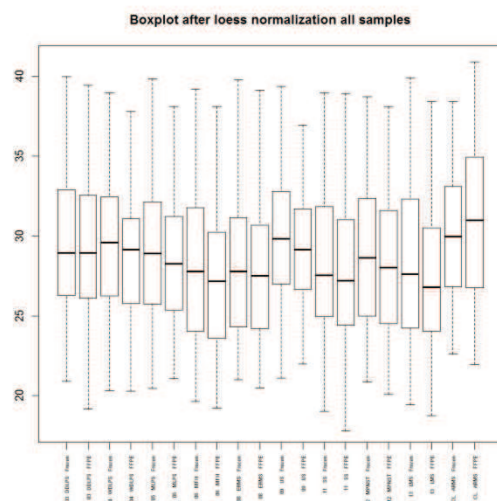
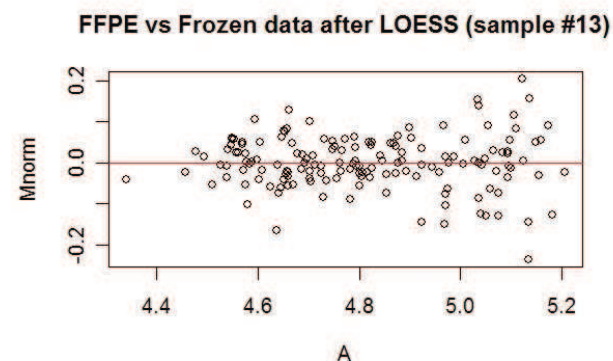


Figure 5 (1 of 2). Results of the local normalization procedures for the 9 fresh tumor samples and the rhabdomyosarcoma cell line.

C1



C2



C3

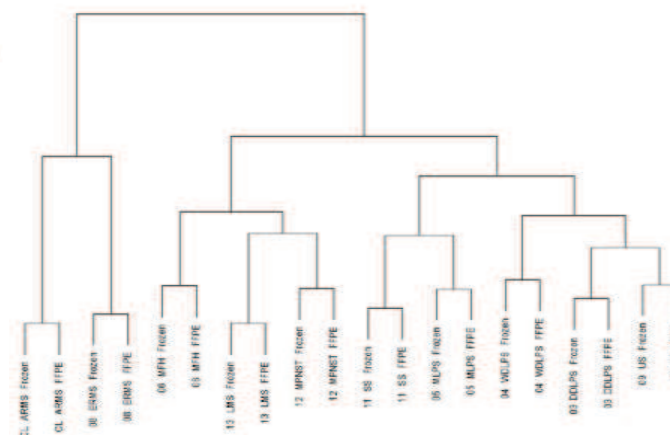


Figure 5 (2 of 2). Results of the local normalization procedures for the 9 fresh tumor samples and the rhabdomyosarcoma cell line. (A1) Boxplot of raw data paired samples. (B1) Boxplot of data after the quantile normalization. (C1) Boxplot of data after the LOESS normalization. (A2) MA plot of the raw data of sample #13, used here as an example. (B2) MA plot of sample #13 after quantile normalization. (C2) MA plot of sample #13 after the LOESS normalization. (A3) Unsupervised clustering with no local normalization. 8 paired samples are well-classified. (B3) Unsupervised clustering after the quantile normalization method. 7 paired samples are well-classified. (C3) Unsupervised clustering after the LOESS normalization method. All paired samples are well-classified.

2. Validation par comparaison avec les puces microARN

2.1. Présentation de la plateforme de puce à microARN

Afin de réaliser les profils d'expression des microARN dans le cadre de ce travail de thèse, nous avons initialement dû faire un choix concernant la plateforme technologique à utiliser. La plateforme TLDA s'est rapidement imposée à nous, non seulement parce que nous disposions d'un appareil PCRq compatible, mais également parce que nous étions certains de la faisabilité de la PCRq sur matériel FFPE, en raison de l'expertise diagnostique de notre équipe dans la détection des transcrits de fusion présents dans les sarcomes, à partir de matériel FFPE.

Par l'intermédiaire du projet Européen CONTICANET, nous avons eu l'opportunité de monter une collaboration de quelques mois avec l'*Erasmus Medical Center* de Rotterdam (Erasmus MC, Pays-Bas), qui possède une plateforme dédiée aux puces d'expression de type microARN (figure 33).

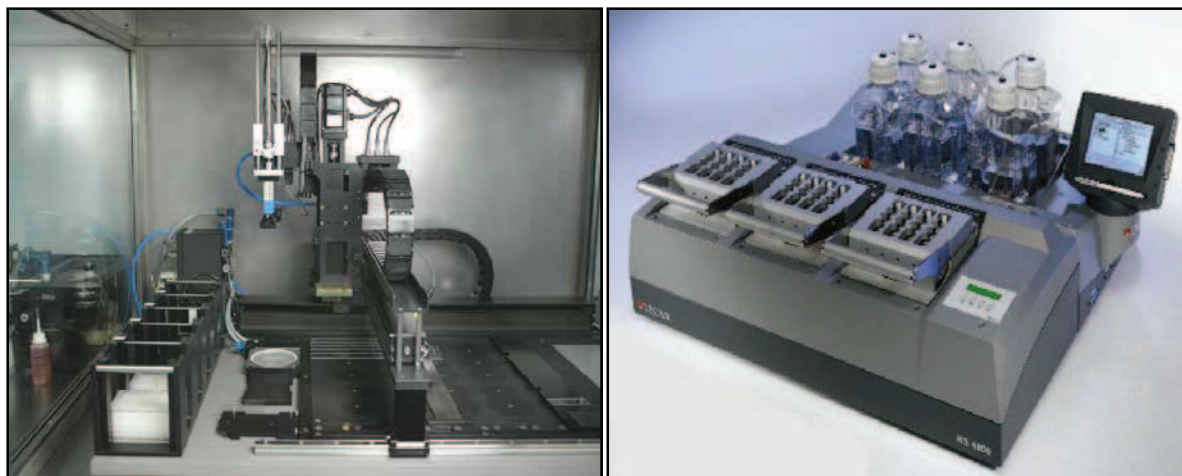


Figure 33 : présentation de la plateforme de puce présente à l'*Erasmus MC*.

A gauche, robot utilisé pour la fabrication des puces maisons. A droite, station d'hybridation.

Cette plateforme repose sur des puces fabriquées en pièce blanche dans les locaux de l'Erasmus MC, sur lesquelles sont déposées des sondes à températures d'hybridation normalisées, grâce à l'inclusion de bases nucléiques de type LNA développées par la société Exiqon (Danemark). La puce proprement dite contenait 718 sondes de capture déposées en doublons, dont le design était basé sur la version 10 de la miRBase. Par rapport à la technologie TLDA, la technologie de type puce permet donc en une seule expérience de réaliser les profils d'expression d'un nombre beaucoup plus élevé de microARN, mais a l'inconvénient de ne pas pouvoir distinguer les microARN matures des microARN immatures. Il existe en effet de nombreux phénomènes de régulation des protéines impliquées dans la maturation des microARN (Obernosterer et al., 2006), et les données obtenues pourraient donc ne pas être exactement représentatives du niveau d'expression des microARN réellement actifs dans les cellules tumorales.

Après extraction des ARN totaux, les cibles contenant les microARN sont marquées par un marquage chimique direct de type ULS, puis hybridées sur la puce microARN toute une nuit.

2.2. Analyse des résultats des profils d'expression obtenus sur puce

Sur les 13 couples d'échantillons FFPE et congelés sélectionnés pour le développement de notre méthodologie d'analyse des résultats issus de TLDA, seuls 10 couples d'échantillons ont été utilisables, en raison de problèmes techniques rencontrés lors de l'extraction. Ainsi, 10 couples d'échantillons ont pu bénéficier de l'établissement des profils d'expression des microARN sur cette plateforme de puce, parallèlement aux profils d'expression obtenus sur plateforme TLDA. Les mêmes solutions d'ARN totaux, extraites dans notre laboratoire, ont pu être utilisées sur les deux plateformes.

Nous avons également ajouté à notre cohorte 19 échantillons tumoraux provenant de l'Erasmus MC, et contenant 7 léiomyosarcomes (dont 4 congelés et 3 FFPE), 4 liposarcomes (dont 3 congelés et 1 FFPE), et 8 échantillons de tissus sains (dont 5 congelés et 3 FFPE).

Les résultats obtenus par hybridation sur puces ont été normalisés par la méthode des quantiles, d'après les recommandations des équipes bioinformatiques de l'Erasmus MC. Quatre groupes ont été constitués, selon la provenance des échantillons (CLB ou Erasmus MC), et selon le mode de conservation (congelé ou FFPE). La normalisation par les quantiles a été effectuée en deux passages, au sein de chaque groupe puis sur la totalité des échantillons. Les modifications apportées par la normalisation par les quantiles ont été appréciées à l'aide de graphiques de type boîtes à moustaches (figure 34).

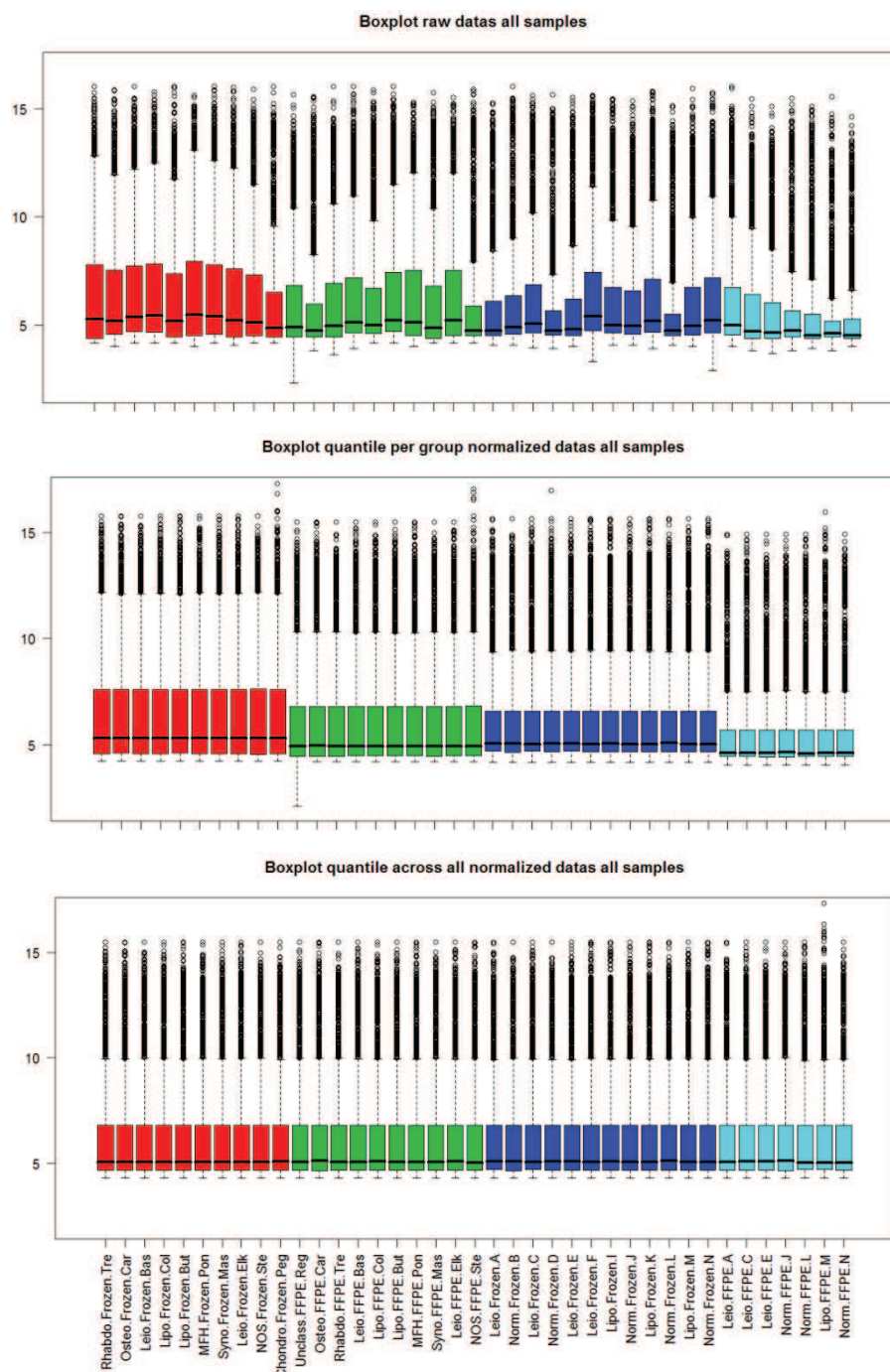


Figure 34 : boîtes à moustaches montrant le suivi des résultats de la normalisation par les quantiles. En haut, données brutes avant normalisation. Les couleurs représentent les quatre groupes. Les dix échantillons congelés en provenance du CLB sont représentés en rouge, les dix échantillons FFPE correspondant sont en vert, les douze échantillons congelés de l'Erasmus MC sont représentés en bleu foncé, et les sept échantillons FFPE de l'Erasmus MC sont représentés en bleu turquoise. Au milieu, données normalisées par les quantiles en intra-groupe. En bas, données normalisées par les quantiles en inter-groupe.

Deux types d'analyse ont ensuite été réalisées. Dans un premier temps, nous avons analysé les dix couples d'échantillons congelés et FFPE provenant du Centre Léon Bérard en clustering non supervisé, pour voir s'il était possible d'apparier les vingt échantillons deux par deux. Malheureusement, en clustering non supervisé, deux grands groupes sont formés, correspondant d'un côté aux échantillons congelés, et de l'autre aux échantillons fixés (figure 35).

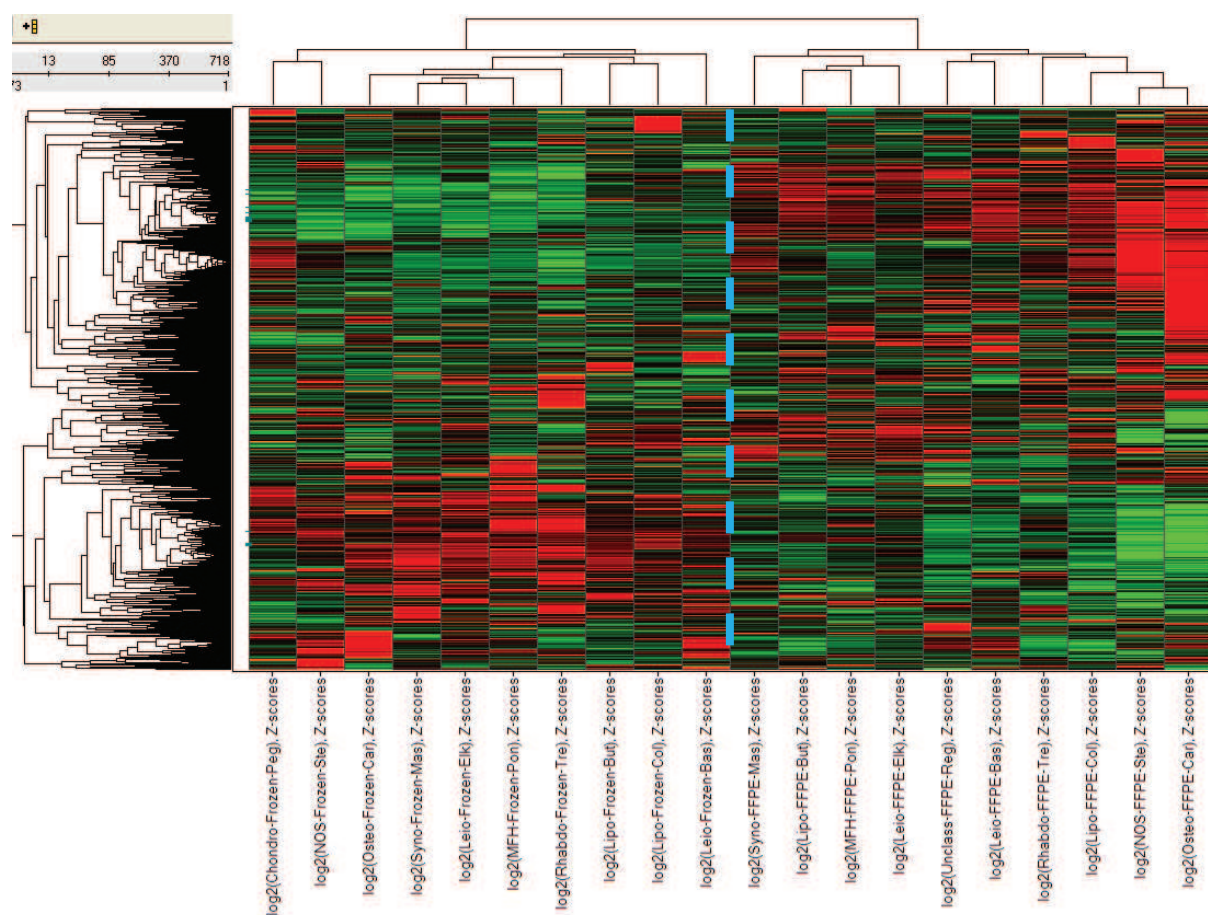


Figure 35 : clustering hiérarchique non supervisé des dix couples d'échantillons représentant des sarcomes variés, en provenance du CLB.

Le clustering non supervisé sépare les échantillons en deux groupes distincts, les échantillons congelés et les échantillons FFPE. Aucun couple d'échantillons n'est apparié.

Partant du principe que ces mauvais résultats venaient peut-être du fait que la plateforme de puce était inadaptée pour une utilisation avec du matériel FFPE, dans une deuxième analyse, nous avons créé une cohorte de 11 échantillons congelés, composée de 4 échantillons provenant du CLB (2 léiomyosarcomes et 2 liposarcomes), et de 7 échantillons provenant de l'Erasmus MC (4 léiomyosarcomes et 3 liposarcomes). Après normalisation, nous avons essayé de voir si les léiomyosarcomes et les liposarcomes se regroupaient entre eux en clustering non supervisé. A notre grande surprise, le clustering non supervisé a séparé les tumeurs en deux groupes non liés à la nature des échantillons, mais en fonction de leur origine hospitalière. De plus, au sein de chaque groupe, les mêmes entités ne se regroupaient pas (figure 36).

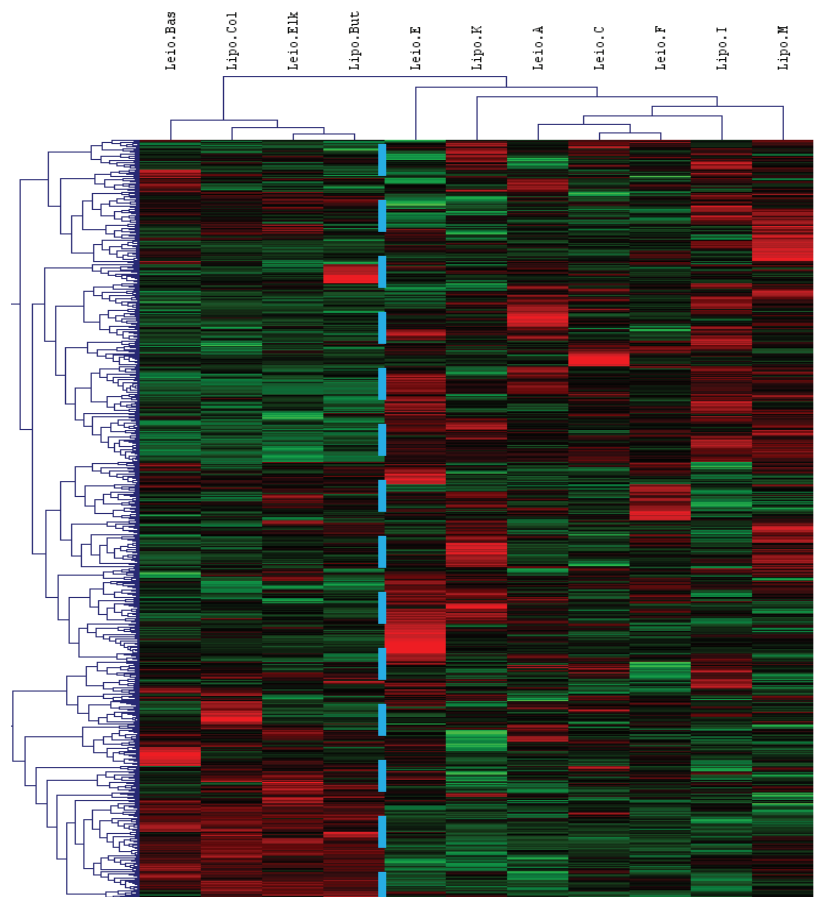


Figure 36 : clustering hiérarchique non supervisé des 11 échantillons congelés.

Ce clustering non supervisé est uniquement constitué de léiomyosarcomes et d'ostéosarcomes, qui ne se regroupent pas entre eux.

2.3. Conclusion

D'après notre expérience, et bien qu'elle soit limitée à un petit nombre d'échantillons, la plateforme de puce à ADN ne nous a pas permis d'obtenir des résultats satisfaisants pour l'établissement des profils d'expression des microARN, en particulier à partir de matériel FFPE.

Cette plateforme semble beaucoup plus sensible aux variations expérimentales, puisque lors de l'analyse d'échantillons exclusivement congelés, une séparation naturelle s'est effectuée lors du clustering non supervisé entre les échantillons en provenance de Lyon et les échantillons en provenance de Rotterdam.

Les technologies en jeu étant très différentes, plusieurs autres raisons pourraient expliquer ces résultats inattendus. Comme nous l'avons vu lors de la normalisation des données brutes issues de prélèvements FFPE passés en plaques TLDA, la normalisation par les quantiles n'est peut être pas non plus adaptée aux données générées par les puces microARN. En effet, même si le nombre de cibles miRNA est beaucoup plus élevé qu'en technologie TLDA, cela reste encore un set de données de très petite taille, et la normalisation par les quantiles est peut-être trop « agressive ».

La suite de nos travaux portant sur les applications cliniques des profils d'expression des microARN dans les sarcomes a été intégralement réalisée sur plateforme TLDA, grâce à la méthodologie d'analyse que nous avons développée précédemment.

II. PRÉDICTION D'UNE RÉPONSE THÉRAPEUTIQUE GRÂCE AUX PROFILS MICROARN

Comme nous l'avons vu dans l'introduction bibliographique, les OS sont des tumeurs agressives présentant des anomalies génomiques complexes. La complexité de ces tumeurs et leur grande hétérogénéité font que leur biologie est encore mal connue, et qu'aucun marqueur universel n'existe pour le développement de stratégies thérapeutiques ciblées. Avec la chirurgie, la chimiothérapie reste le principal traitement applicable aux OS.

Depuis les années 1990, la survie médiane n'a pas connu d'améliorations significatives, malgré la généralisation de nombreux agents chimiothérapiques, utilisés actuellement en combinaisons (Ferrari and Palmerini, 2007). Cette phase de plateau s'explique en partie par le fait qu'à peine 30 % des patients peuvent être considérés comme de bons répondeurs (Provisor et al., 1997). De nombreux efforts sont faits pour essayer de mieux comprendre pourquoi certains patients ne répondent pas ou peu à la chimiothérapie, dans l'espoir de pouvoir un jour lever cette principale limitation (Chou and Gorlick, 2006).

Actuellement, hormis la présence de métastases au diagnostic, la réponse à la chimiothérapie néoadjuvante est le marqueur pronostique le plus fort (Davis et al., 1994a). En effet, la survie globale à cinq ans des bons répondeurs est de presque 80 %, alors que celle des mauvais répondeurs n'est que d'environ 60 % (Bacci et al., 2005). Cependant, la réponse à la chimiothérapie néoadjuvante ne peut être déterminée qu'à la suite de l'acte chirurgical, par une évaluation du pourcentage de nécrose tumorale. Il serait très utile de pouvoir prédire la réponse tumorale avant l'exérèse complète, et de pouvoir la monitorer à partir de prélèvements microbiopsiques ou mieux, à partir de marqueurs sanguins. Cette information permettrait de planifier des marges de résection optimales, et permettrait également une personnalisation du traitement (Cheon et al., 2009).

Dans cette optique, nous avons entrepris d'analyser les profils d'expression des microARN dans 27 OS conventionnels de haut grade, issus de patients traités au Centre Léon Bérard et à La Timone (Marseille), exclusivement FFPE. Classiquement, avant inclusion, les prélèvements osseux subissent des traitements de décalcification ayant pour but de ramollir les tissus calcifiés, et de permettre la réalisation de coupes histologiques indispensables aux médecins anatomopathologistes. Par rapport aux analyses précédentes (cf. article I, en soumission au journal *RNA*), nous avons donc ajouté une difficulté technique supplémentaire qui nous a permis de tester la robustesse de notre méthodologie d'analyse appliquée à des cohortes hétérogènes, puisque la décalcification dégrade fortement les acides nucléiques. En parallèle, nous avons également pu analyser 12 tumeurs fraîches issues d'un modèle d'OS de rat bien caractérisé, les rats étant traités ou non à l'IFO (Dutour et al., 2007), ainsi que 2 lignées tumorales humaines d'OS.

A partir du modèle d'OS de rat, nous avons tout d'abord montré que les tumeurs possédaient des profils d'expression spécifiques par rapport à d'autres tissus sains (os, poumon et muscle). Une analyse supervisée entre les tumeurs traitées et les tumeurs non traitées nous a ensuite permis de montrer que les profils d'expression des OS traités à l'IFO se classaient automatiquement en deux groupes distincts, corrélés à la bonne ou mauvaise réponse des rats à la chimiothérapie. Sur la base de la signature des microARN permettant de distinguer les bons des mauvais répondeurs, les tumeurs non traitées ont été regroupées soit avec les bons répondeurs, soit avec les mauvais répondeurs, révélant alors tout le potentiel prédictif de cette signature.

A partir des OS issus de patients traités au Centre Léon Bérard, grâce à une analyse supervisée entre les bons et les mauvais répondeurs, nous avons identifié une signature d'expression de cinq microARN significativement dérégulés (miR-92a, miR99b, miR-132, miR-193 et miR-422a), dont trois étaient communs à ceux trouvés chez le rat (miR-92a, miR99b et miR-193). Cette signature à cinq gènes permet de

séparer avec précision plus de 85 % des tumeurs en fonction de leur réponse à la chimiothérapie.

L'analyse des deux lignées humaines d'OS nous a permis de démontrer, sur la base de la signature d'expression précédemment trouvée, que l'une d'entre elle était sensible à la chimiothérapie, et l'autre résistante. Ainsi, ces deux lignées constituent de bons modèles pour déchiffrer les mécanismes sous-jacents à la dérégulation des microARN. Ces lignées pourraient nous permettre de comprendre pourquoi certaines tumeurs sont chimiorésistantes, avec pour but ultime la levée de cette résistance par une approche thérapeutique ciblée.

Nous avons donc démontré qu'il était possible de prédire la réponse à la chimiothérapie des patients porteurs d'OS grâce à l'analyse des niveaux d'expression de cinq microARN. Cette signature d'expression pourrait être transférable en routine, où elle serait idéalement recherchée à partir de prélèvements microbiopsiques ou sanguins. Elle permettrait alors une meilleure prise en charge clinique des patients, sur le plan chirurgical comme thérapeutique.

Document à suivre :

Article II : *micro-RNA profiles in osteosarcoma as a predictive tool for ifosfamide response.*

Micro-RNA profiles in osteosarcoma as a predictive tool for ifosfamide response

Angélique Gougelet^{1,3}, Daniel Pissaloux², Anthony Besse¹, Jennifer Perez^{1,3}, Adeline Duc^{1,3}, Aurélie Dutour^{1,3}, Jean-Yves Blay^{1,3,4} and Laurent Alberti^{1,3}

¹Unité INSERM U590 équipe Cytokines et Cancer, Centre Léon Bérard, 69373 Lyon cedex 08, France

²Service d'Anatomie et Cytologie Pathologiques, Centre Léon Bérard, 69373 Lyon cedex 08, France

³Conticanet (FP6-06188)

⁴EORTC, Brussels, Belgium

Micro-RNAs (miRNA) are currently used as cancer biomarkers for hematological cancers and solid tumors. Osteosarcoma is the first primary malignant bone tumor, characterized by a complex genetic and resistance to conventional treatments. For this latter property, the median survival has not been improved since 1990 despite preoperative administration of chemotherapeutic agents. The prediction of tumor response before chemotherapy treatment would constitute a major progress for this pathology. We assessed in this study if miRNA profiling could surpass the current limitations for osteosarcoma diagnosis. We measured the miRNA expression in different osteosarcoma samples: (i) 27 osteosarcoma paraffin-embedded tumors from patients, (ii) human osteosarcoma cell lines, and (iii) tumors from a syngeneic rat osteosarcoma model, recapitulating human osteosarcoma. miRNA profiles were determined using microfluidic cards performing high-throughput TaqMan[®]-based PCR assays, called TaqMan[®] Low Density Arrays. Osteosarcoma of rat and human origins showed a miRNA signature, which could discriminate good from bad responders. In particular, we identified five discriminating miRNAs (miR-92a, miR-99b, miR-132, miR-193a-5p and miR-422a) in patient tumors, which could be easily transferable to diagnosis. These discriminating miRNAs, as well as those identified in rat, targeted the TGF β , the Wnt and the MAP kinase pathways. These results indicate that our platform constitutes a potent diagnostic tool to predict tumor sensitivity to a drug in attempt to better adapt treatment to tumor biological specificities and also to identify new potential therapeutic strategies.

Micro-RNAs (miRNAs) are 22 nucleotide non-coding small RNA molecules, involved in the regulation of growth, development and differentiation.¹ According to the miRBase

Sequence Database release 15, 940 miRNA sequences have been identified to date in *Homo sapiens* and 326 in *Rattus norvegicus*. miRNAs are synthesized from intronic regions, processed by specific complexes of proteins containing Drosha and Dicer and integrated in RISC complex following maturation.^{2,3} Their matching with complementary sequences in messenger RNA (mRNA) results in translation inhibition and accelerated mRNA degradation. miRNAs could also act through a mechanism based on chromatin silencing.⁴

miRNA expression levels are characteristic for one tissue to regulate tissue-specific target genes during development.⁵ Recent studies have suggested miRNA implication in skeletal tissue development, that is, miR-29 for osteoblast phenotype attainment⁶ or miR-223 for osteoclast differentiation.⁷ miRNAs also regulate oncogenesis, being both oncogene and tumor suppressor.⁸ Their expression is thus modified in various cancers as compared to normal tissues,^{9,10} resulting in a tumor miRNA signature, which could be useful for their classification in line with their tissue origin and molecular alterations^{9,11,12} and their diagnosis.^{11,13}

This study focused on miRNA expression in osteosarcoma, the most frequent primary malignant bone tumor, whose major localizations arise in the femur or the tibia. High grade tumors are highly metastatic, particularly in lung sites, when adjuvant chemotherapy is not administered.¹⁴

Key words: osteosarcoma, miRNA profiling, diagnosis, chemotherapy, targeted therapy

Abbreviations: Ct: threshold cycle; FFPE: formalin-fixed paraffin embedded; ifos: ifosfamide; KEGG: Kyoto Encyclopedia of Genes and Genomes; mus: muscle; NT: non treated; PCA: principal component analysis; RQ: relative quantity; RT-qPCR: real-time quantitative PCR; TLDA: TaqMan Low Density Array; Tum: tumor; Tum ori: original tumor; 15-47: human osteosarcoma cells CRL-15-47.

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

Grant sponsors: The Ligue Nationale contre le Cancer (grants from the Ain department), The Institut National du Cancer, The Conticanet Consortium

DOI: 10.1002/ijc.25715

History: Received 9 Jul 2010; Accepted 24 Sep 2010; Online 14 Oct 2010

Correspondence to: Dr. Angélique Gougelet, Unité INSERM U590 Cytokines et Cancer, Centre Léon Bérard, 28 rue Laennec, 69373 Lyon cedex 08, France, Tel.: +33478782964, Fax: +33478782720, E-mail: angeliquegougelet@gmail.com

Since 1990, the median survival has not been further improved through preoperative administration of chemotherapeutic agents (doxorubicin, cyclophosphamide, cisplatin or methotrexate), currently used in combination.¹⁵ Unfortunately, many cases of resistance or relapse occur after treatment, and only few therapeutic options are possible and generally noncurative.¹⁶ The discovery of new targets involved in osteosarcoma tumorigenesis could hence be of great benefit to offer targeted therapies. miRNAs are also promising diagnosis biomarkers with their tissue specificities and since they are detected in cancer patient sera.^{17,18} A non-invasive diagnostic tool based on miRNAs may be useful to adapt chemotherapy protocols to tumor biological specificities.

Here, we studied miRNA expression in different osteosarcoma samples: (i) tumors from a syngeneic rat osteosarcoma model, recapitulating human osteosarcoma, (ii) osteosarcoma tumors from patients, and (iii) human osteosarcoma cell lines derived from patient tumors. miRNA profiles were determined using microfluidic cards performing high-throughput TaqMan[®]-based PCR assays, called TaqMan[®] Low Density Arrays (TLDA). Interestingly, a miRNA signature was identified in osteosarcoma of rat and human origins; this profile was diagnostic and predictive of tumor response to treatment. These results indicate that miRNA profiling could constitute a potent diagnostic tool to predict tumor sensitivity to a drug in attempt to better adapt treatment to tumor biological specificities and finally to improve the patient initial care.

Experimental Procedures

Rat osteosarcoma model

Procedures for animal care were performed according to institutional and national guidelines. Animals were anesthetized throughout all surgical and imaging procedures with isoflurane/oxygen (2.5%/2.5%, v/v) (Minerve, Esternay, France). The transplantable orthotopic and metastatic rat osteosarcoma model has been previously described.^{19,20} This model mimics its human counterpart in terms of aggressiveness, metastatic spreading and chemoresistance phenotype.^{20,21} All the tumors obtained were classified as osteoblastic following histological analyses. Briefly, small tumor fragments (100 mm³) taken from a hyperproliferative osteogenic tumor area (Tum ori), were grafted on 3-weeks old immunocompetents Sprague-Dawley rats (Charles River Laboratories, Wilmington, MA). Using a lateral approach, a tumor fragment was placed contiguous to tibial diaphysis after periosteal abrasion; then, the cutaneous and muscular wounds were sutured. Fourteen days after tumor transplantation, animals underwent a first ¹⁸F-FDG PET Scan and were randomly assigned to a control group treated with saline solution or a treated group exposed to a subcutaneous dose of 10 mg/kg ifosfamide (ifos, Baxter, Deerfield, IL), 7 days apart (at days 15 and 22 after tumor transplantation). A second ¹⁸F-FDG PET Scan was performed 7 days after the second ifos administration. Animals were sacrificed 1 week after the

end of the treatment. Tumor and normal tissue fragments (muscle, bone and lung) were collected for RNA extractions. A second (S2) and third (S3) series of untreated tumors were independently obtained

Human osteosarcoma tumors

Twenty-seven patients treated for high grade conventional osteosarcoma in the Centre Léon Bérard or the La Timone hospital were included in this study. Histological diagnosis was made on biopsies before treatment for all patients. Informed consent was obtained from all patients. All tumors were formalin-fixed paraffin-embedded tumors (FFPE).

Osteosarcoma cell lines

Two human osteosarcoma cell lines were obtained from ATCC (Manassas, VA): MNNG/HOS Cl #5 [R-1059-D] (reference CRL-15-47) and Saos-2 (HTB-85). Cells were grown in DMEM (Gibco, Carlsbad, CA), supplemented with 10% decomplemented fetal calf serum (Lonza, Basel, Switzerland), 10 mL penicillin streptomycin (10 U/mL/10 µg/mL, Gibco, Carlsbad, CA) and 5 mL L-glutamin (200 mM; Gibco) at 37°C humidified atmosphere containing 5% CO₂. Cells were exposed to 50 µM ifosfamide (ifos) (Baxter) for 24 hr.

RNA extraction and quantitative real-time PCR

Tissue and tumor fragments were cryo-ground. Total RNA was extracted from tissue powders and cell pellets using a single phenol/chloroform extraction protocol with TRIzol[®], according to the manufacturer's instructions (Invitrogen, Carlsbad, CA). Five hundred nanograms of total RNA were subjected to the microfluidic PCR technology performed by Applied Biosystems (Foster City, CA). In brief, RNA was reversed transcribed, using multiplexed specific looped miRNA primers from the TaqMan[®] MicroRNA Reverse Transcription kit. The second step consists in a real-time quantitative PCR (RT-qPCR) on TLDA: RT products are introduced through micro-channels into miniature wells that are preloaded with deshydrated specific primers and probes. Recently, Applied released the second version of TLDA, which consists of two cards A and B. Rodent card A allows the analysis of 226 rat-specific miRNAs, while card B analyses 77 miRNAs. For human, analyses were performed for 377 miRNAs on card A and 290 on card B, respectively.

PCR data normalization

For each miRNA, the threshold cycle (Ct) was calculated by the ABI 7900 Sequence Detection System software (plate by plate manual Ct analysis with a threshold at 0.25 and automatic baseline). All further data manipulations were done using R scripts. A cut-off of 32 was applied to discard the late Ct values. Around 60% of miRNAs passed the filtering criteria and were used for further analysis. For each TLDA, quality controls were performed on the raw data by checking internal controls and using box plot and scatter plot diagrams. Samples with any kind of problems were discarded so

they would not introduce bias during the following normalization procedures. We tested different methods of normalization since the recommended “pseudo” normalization factor mammU6 plotted in each card was not stably expressed in our different samples. Normalization with the two most stable miRNAs identified by GeNorm, miR-744 and miR-335-5p, was not optimal too. Finally, a global normalization by the median was chosen for its reliability over experiments. Tissues included in a given analysis were treated altogether, the normalization procedure being applied separately for the two types of card, A and B. Distribution of normalized data was checked with box plots and correlation plots. The following formula was used to correct Ct values of every card:

$$[\text{Normalized Ct}] = \text{Ct} \times [\text{mean of medians}] / [\text{median of the card}]$$

Through this approach, the new median value shared by all samples can be considered as a sort of perfect “virtual housekeeping gene.” Therefore the standard $\Delta\Delta\text{Ct}$ method can be used to determine the relative quantities (RQ) as follows:

$$\Delta\text{Ct} = [\text{Normalized Ct}] - [\text{New shared median}]$$

For the $\Delta\Delta\text{Ct}$ calculation, it was more relevant for the statistical analyses to use the mean of all ΔCt obtained across samples for each miRNA, instead of using the ΔCt of a reference sample:

$$\begin{aligned} \Delta\Delta\text{Ct} &= \Delta\text{Ct} \\ &- [\text{Mean of } \Delta\text{Ct} \text{ across samples for each miRNA}] \\ \text{RQ} &= 2^{\Delta\Delta\text{Ct}} \end{aligned}$$

miRNA target predictions

We compiled 4 databases to determine miRNA targets: TargetScan 5.1, MiRanda, PICTAR and the miRbase databases. These databases search the presence of conserved 8 mer and 7 mer sites on the 3'UTR parts of messenger RNA that match the seed region of each miRNA. It also predicts the efficacy of targeting for each matching site. We created our own database, which regrouped each miRNA with the geneID of all their protein targets, for rat and human. We only conserved couples miRNA/geneID present in two databases at least.

miRNA-regulated cell signaling pathways predictions

We used the “G-language microarray” web application, which allows the mapping of molecular dataset onto “Kyoto Encyclopedia of Genes and Genomes” (KEGG) pathway maps.²² We first input miRNA-targeted proteins of interest and the sum of RQ values for all miRNAs that regulate these proteins, contained between 1 and 50; the software then generates KEGG data to create FLASH graphics of cell signaling pathways in which proteins are involved. The color intensity

of a highlighted protein varies with the strength of its regulation by miRNAs.

Proliferation assay

Cells were plated in 96 well plates at 5,000 cells/well and exposed to 50 μM ifosfamide or not (NT). Cell growth was measured 24, 48 and 72 hr later with 20 μL Cell Titer Glo[®] luminescent reagent (Promega, Madison, WI) for 10 min. Luminescence was recorded using a Microbeta reader (PerkinElmer, Fremont, CA).

Statistical analysis

Normalized RQ data were directly input into the TIBCO Spotfire[®] DecisionSite for Functional Genomics analysis software. We performed unsupervised hierarchical clustering to classify samples by groups. The selection of miRNAs useful to predict tumor response to treatment was statistically realized using ANOVA tests with p -values of 0.03 at least. Results were verified through supervised hierarchical clustering.

Data from miRNA lists of interest were then used as variables in a two dimensional principal component analysis (PCA) performed with R 2.9.0 package to demonstrate their capabilities to distinguish between sensitive and resistant tumors. PCA supplies a simplified two-dimensional picture to our multivariate dataset of miRNA RQ values. By mathematical combination of values according to their strength, two principal components are created that represent as much as possible the variability of the data. Thus, tumors possesses two new coordinates in a two-dimensional space. According to their localization in this space, tumors form groups and can be predicted sensitive or resistant.

Results

miRNA signature of osteosarcoma tumors in rat model

miRNA expression profiles were performed on a series of 34 samples [of tumor (Tum), lung, muscle (mus) and bone origins] collected from an orthotopic osteosarcoma rat model, randomly treated or not with 10 mg/kg ifosfamide.^{19,20} As mentioned previously, this model closely mimics its human counterpart, regarding aggressiveness, chemoresistance and the onset of lung metastases. Tumors as normal tissue fragments collected at the end of the treatment were submitted to RNA extraction and analyzed using the TLDA platform, using two 384-microfluidic cards A and B designed to measure miRNA levels. We did not used card B plotted with recently discovered and poorly described miRNAs, since only one miRNA plotted on this card was significantly modulated by treatment in our preliminary experiments (data not shown). About 230 miRNAs among the 381 represented in the card A (60%) met the filtering criteria. Ct values of each miRNA were normalized to median Ct of all miRNAs plotted on all cards, the best normalization method as described by D. Pissaloux in our laboratory (in press).

A three-dimensional scatter plot analysis showed that three different samples of one tissue, of tumoral or normal

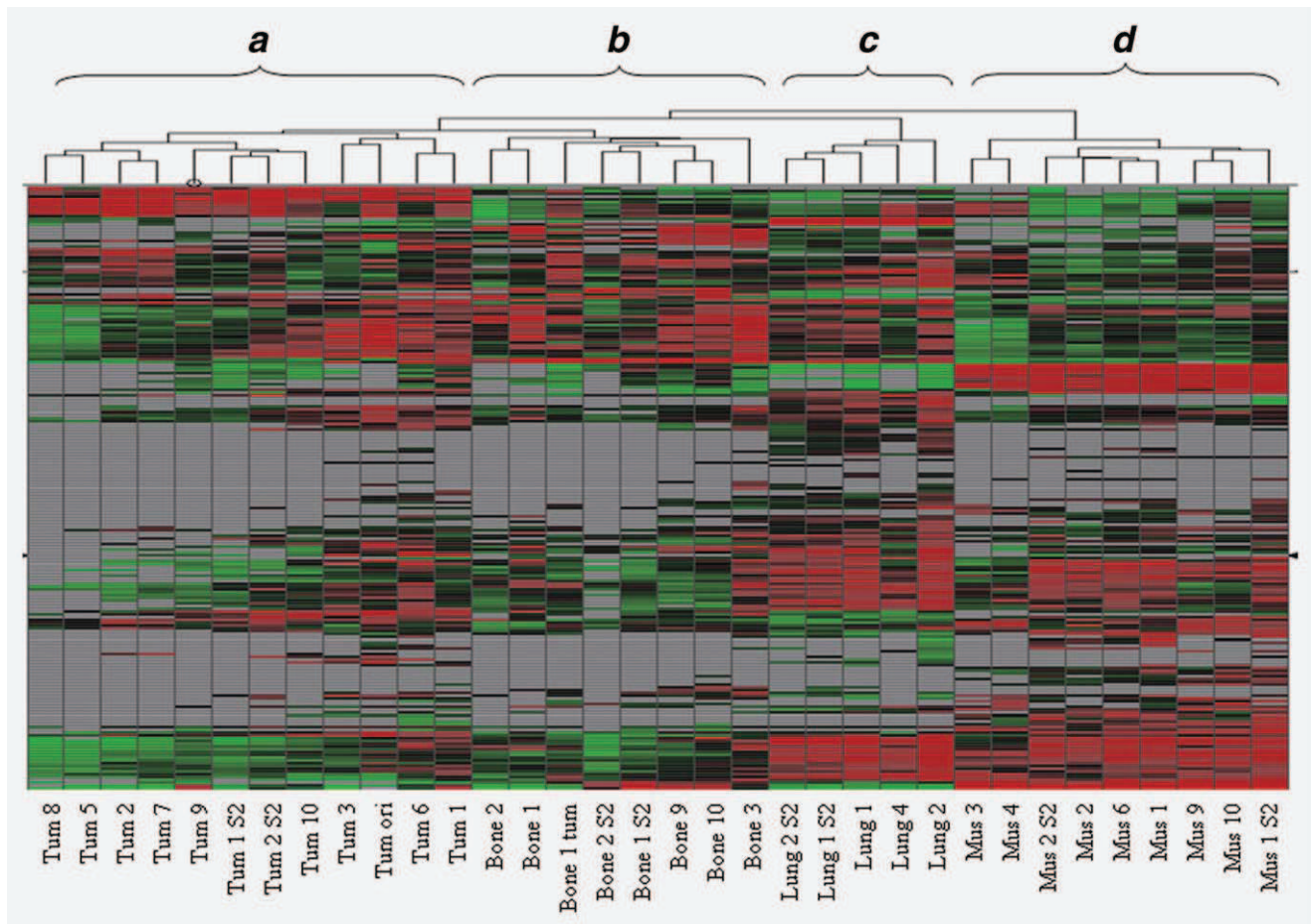


Figure 1. Heat map clustering of miRNA expression in osteosarcoma tumors (12 samples) as compared to bone (8), lung (5) and muscle (9). Each row represents the relative levels of expression for each miRNA and each column shows the expression levels for each sample. The red or green color indicates relatively high or low expression, respectively, while grey squares indicate absent data points. The four main groups of the dendrogram are gathered by tissue: *a*: tumor; *b*: bone; *c*: lung; *d*: muscle.

origin, expressed similar miRNA patterns (Supporting Information Fig. S1a). In addition, an unsupervised clustering based on Spearman's correlation revealed that all samples from one tissue clustered together (Supporting Information Fig. S1b). Finally, an unsupervised hierarchical clustering based on euclidian distance confirmed that all tissues were grouped according to their origins (Fig. 1). Osteosarcoma tumors presented a miRNA signature distinct from those of muscle and lung and also from those of bones. These profiles were highly conserved since other samples from independent experiments perfectly clustered with the previous ones (noted S2 for series 2 in Fig. 1).

Tumor miRNA profiles were correlated to their response to ifosfamide treatment

Thereafter, we focused on miRNA profiles of the tumor samples obtained at the end of treatment. We performed a supervised hierarchical clustering following an ANOVA ($p < 0.03$) between two groups, treated or not with ifos. Rat

treatment was detailed in Supporting Information Table S1. The clustering highlighted that tumors clustered in three groups: (i) a group with tumors 2, 5, 7 and 8, (ii) a group containing Tum 3, 9 and 10, and (iii) a group composed of Tum 1 and 6 (Fig. 2a). This subsequent clustering was found in perfect correlation with the response observed *in vivo* during the experiment: untreated tumors were tumors 3, 9 and 10; tumors 1 and 6 were treated and sensitive to treatment (tumor volume <50% volume of control group, necrotic and hypoxic according ^{18}F -FDG PET Scan and immunohistochemistry analyses), while Tum 2, 5, 7 and 8 were resistant tumors (Supporting Information Table S1). Then, we performed a similar supervised hierarchical clustering ($p < 0.03$) with normal tissues and observed that muscle as well as bone samples were randomly distributed, regardless of treatment (Fig. 2b). Therefore, tumor miRNA profiles were associated to response to treatment and not a bias related to the exposure of the animal to ifosfamide.

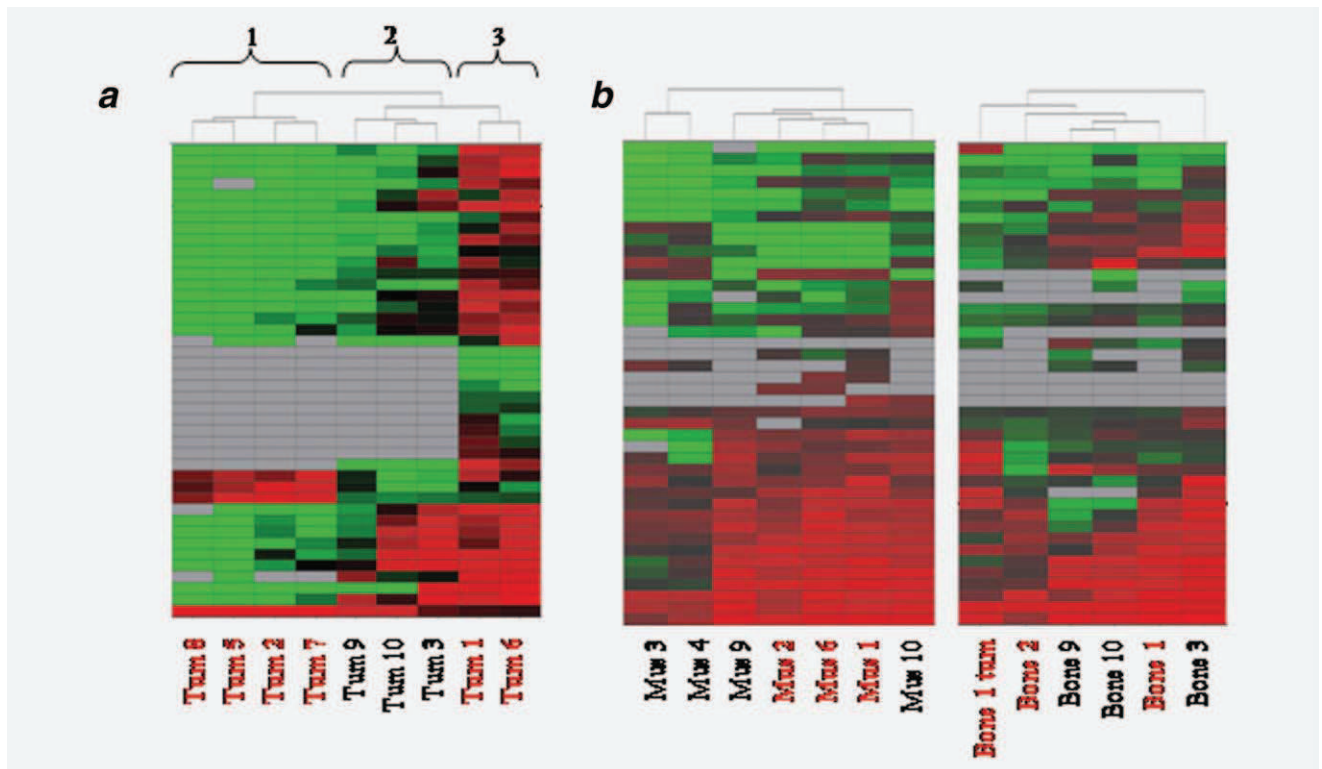


Figure 2. miRNA expression profiles of osteosarcoma tumors were consistent with their response to treatment. This hierarchical clustering only conserved miRNA differently expressed in tumors according to their response to treatment (untreated, good response, no/low response) following an ANOVA ($p < 0.03$). *a*: The three main groups of the dendrogram are gathered by conditions of treatment: 1: tumors resistant to ifosfamide, 2: untreated tumors, 3: tumors sensitive to ifosfamide. *b*: muscle and bone profiles. Samples from ifosfamide treated animals were indicated in bold red.

The predictions by miRNA profiling were consistent with tumor response *in vivo*

Since the TLDA platform was reliable for the identification of tumor miRNA profiles correlated to response to treatment, we investigated whether this technology allowed us to predict the hypothetical response of untreated tumors to ifosfamide. To this aim, we compiled relative quantity values for miRNAs allowing us to discriminate sensitive and resistant tumors in Figure 2. We statistically selected a panel of 10 miRNAs shared by rat and human and stably expressed in all the tumors in the three experiments. Then, we performed a statistical analysis based on Principal Component Analysis, which studies the variability between a set of variables. In brief, this mathematical procedure assigned for each contributory miRNA in treated tumors a new coordinate system, composed of two coordinates in our case (Fig. 3a). We could notice that all miRNAs presented a negative contributory factor according to their vector representation in left quadrant in Figure 3a. This corroborated the previous observation in Figure 2 that discriminating miRNAs were overexpressed in sensitive tumors. In addition, these selected miRNAs were representative of tumor response since their variance in the new 2D-system represented more than 97% variance of initial values (component 1 + component 2). Finally, some

miRNAs were particularly involved in tumor response, that is, miR-142-3p, miR-30c, miR-19b and miR-93, since their graphic representation showed a longer vector and a larger angle with the axis "component 1" (Fig. 3a). After this analysis, we adjusted RQ values of each miRNA among the 10 selected for each untreated tumors by the new coefficients obtained previously. Thus, two new coordinates were assigned to each untreated tumor following the sum up of all miRNAs. The couples of coordinates were useful to realize a 2D PCA diagram in attempt to predict tumor response to treatment (Fig. 3b). These computations suggested that tumor 3 may respond to ifosfamide, since it appeared in left window near Tum 1 and 6, while tumor 9 might be resistant. These predictions were consistent with the miRNA profiles obtained in Figure 2a: Tum 9 clustered with refractory tumors 2, 5, 7 and 8 in the dendrogram, while Tum 3 clustered with Tum 1 and 6. Tumor 10 appeared in left panel in Figure 3b. Although its position was in favor of a good response, it appeared quite far from axis "component 2". In the next steps, we considered this type of tumor as undetermined on the basis of statistical analysis and by cross-checking with the clustering.

Altogether, these different analyses showed the reliability and the reproducibility of this technology resulting in relevant

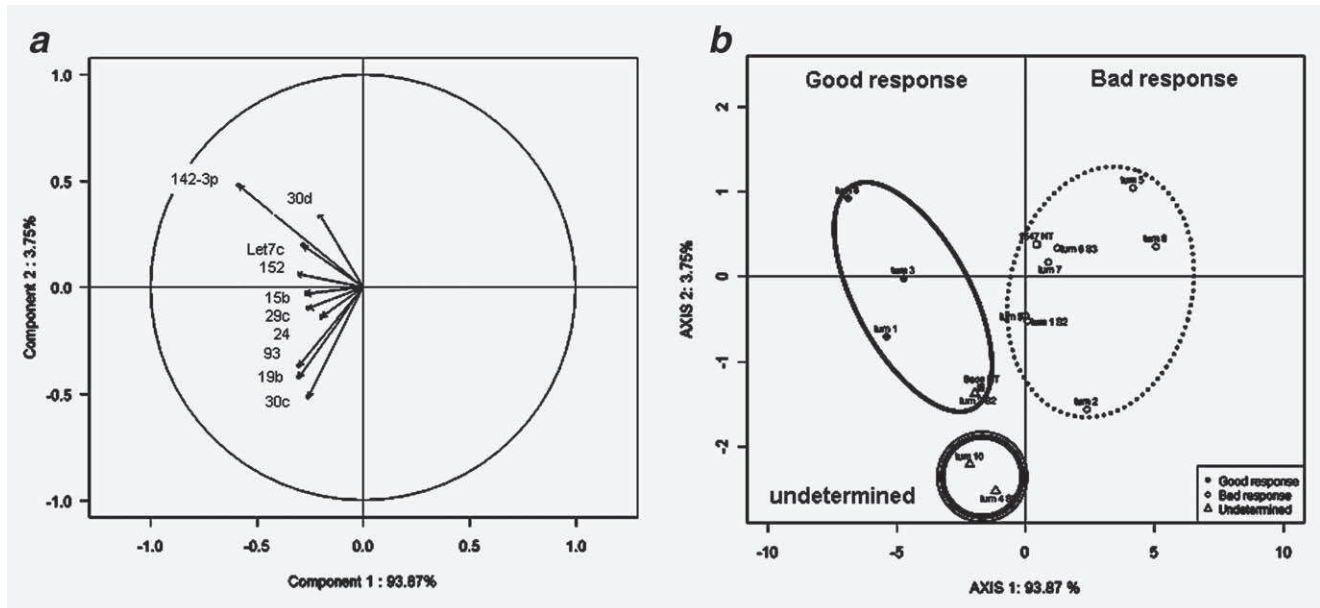


Figure 3. Predictive response of untreated tumors were consistent with hierarchical clustering. *a*: Principal component analysis diagram of the 10 miRNAs, which signed tumor response to treatment and were stably expressed in all tumor samples. The arrow represents the contribution of each miRNA expression level to tumor response as described in Material and Methods. *b*: PCA of untreated tumors as a tool to determine the potential tumor response to treatment. RQ values of the ten selected miRNAs for each tumor were corrected by the coefficients determined in A, as described in Material and Methods. Tumors and cells represented in left panel, (●) and (■), respectively, were predicted sensitive to treatment, while those in right panel were considered as resistant, (○) and (□). Tumors with very low value for component 2 were considered undetermined (Δ).

biological miRNA profiles. miRNA profiling could constitute a tool of choice to predict tumor susceptibility to treatment. In particular, these data strongly supported that tumor miRNA profiles were correlated to tumor response, even if these tumors presented a similar genetic background and were injected in syngeneic rats.

miRNA signature in patient osteosarcoma tumors

Thereafter, we performed a similar analysis with 27 osteosarcoma FFPE tumors obtained from patients. As previously suggested by Liu *et al.*, FFPE tissues are suitable for miRNA expression analysis.²³ We performed a supervised hierarchical clustering with a *p* value < 0.07 with good responders versus poor responders. As previously for rat model, we only used the TLDA card A for human analysis. Five miRNAs discriminated good responders from bad ones: miR-92a, miR-99b, miR-193a-5p and miR-422a were overexpressed in good responders contrary to miR-132, which was reduced (Fig. 4 and Supporting Information Fig. S3). Interestingly, three of them (miR-193, miR-99b and miR-92a) were also deregulated in the rat model according to tumor sensitivity to ifosfamide. On the basis of these five miRNAs, we realized a similar PCA analysis than previously, which successfully classed the good responders in a distinct group from the poor responders (Fig. 5). The selected miRNAs were representative of tumor response since their variance represented 80% inertia (component 1 + component 2). Three tumors were consid-

ered as undetermined as previously mentioned for the *in vivo* model (Fig. 3b).

Relationship between miRNA profiles of human osteosarcoma cell lines and cell response to ifosfamide

Finally, we investigated if miRNA signature identified in human osteosarcoma tumors was also representative of two human osteosarcoma cell lines, Saos-2 and CRL-15-47 (15-47), which could constitute an interesting *in vitro* model. We first tested the response of these cells to ifosfamide by a proliferation assay based on ATP measurement. As shown in Figure 6a, 15-47 cells are resistant to 50 μ M ifosfamide, while Saos-2 growth only decreased after 72 hr treatment. Thus, these two cell lines could mimic the different response observed *in vivo* (Supporting Information Table S1). We investigated if these two cell lines presented different miRNA profiles when they were exposed to 50 μ M ifosfamide or untreated. We chose an exposure of 24 hr, since apoptosis program might be engaged at this time in sensitive Saos-2 cells but no mortality was observed, to obtain enough cells for RNA extraction.

Saos-2 and 15-47 cells effectively expressed different miRNA patterns at the basal level and after exposure to ifosfamide (Fig. 6b). 80% filtered miRNAs were over-expressed in the sensitive Saos-2 cells, as it was observed for sensitive tumors in Figure 2a. A supervised hierarchical clustering with a *p*-value < 0.03 highlighted 61 miRNAs significantly

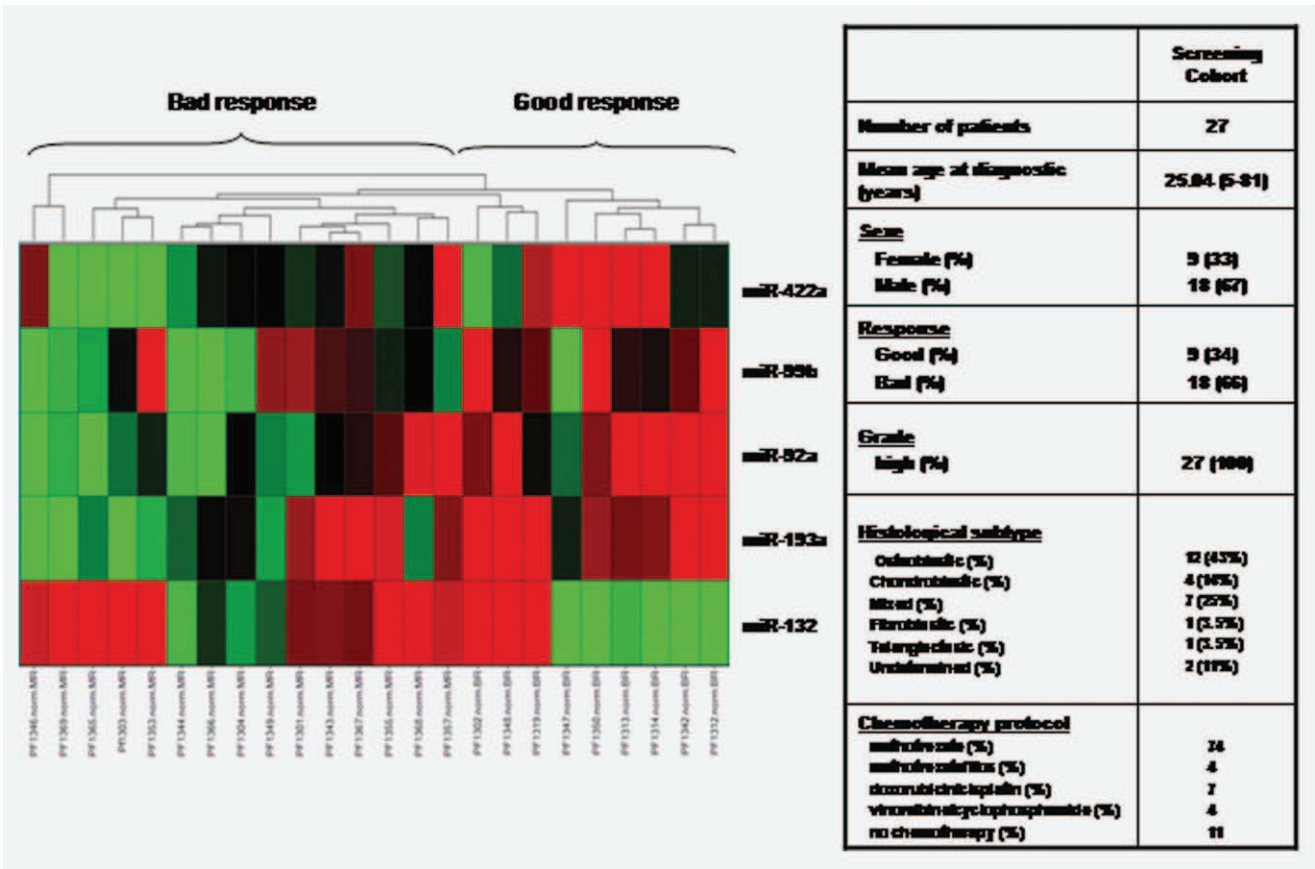


Figure 4. miRNA expression profiles of osteosarcoma tumors from patients were consistent with their response to treatment. This supervised hierarchical clustering only conserved miRNA differently expressed in tumors according to their response to treatment (good response, no/low response) following an ANOVA ($p < 0.07$). The table summarizes the characteristics of tumors obtained from patients: mean age, sex, response, histological subtypes and type of neoadjuvant chemotherapy administered.

and differently expressed between the two cell lines. Of great interest, we could notice that miR-99b, miR-132 and miR-193a-5p identified in human tumors, were also significantly deregulated in these cells according to their response to ifosfamide. A panel of six miRNAs identified as essential for tumor response in rat (miR-15b, miR-29c, miR-24, miR-93, miR-19b and miR-30c) were also significantly affected by treatment of these cell models (Fig. 3a). The PCA analysis indicated that Saos-2 cells were predicted as sensitive to ifosfamide, contrary to 15–47 cells (Fig. 3b), in agreement with the results obtained in the proliferation assay in Fig. 6b.

In conclusion, these two cell lines could constitute a promising model, mimicking osteosarcoma tumors from patients, to decipher the importance of the miRNAs identified as discriminating for drug response.

miRNA targets were involved in tumor cell cycle, invasion, osteosarcoma tumorigenesis and bone physiology

Thereafter, we searched through an *in silico* approach the predicted targets of the five miRNAs shown as differently expressed in tumors according to their response to ifosfamide (Fig. 4). We created a database, as described in Material and

Methods, which predicted 1,349 potential targets for these miRNAs. Since selected miRNAs were, in most cases, increased in good responders, their protein targets might be decreased. It was thus not surprising that these miRNAs mainly target regulators of cell proliferation like the MAP kinase pathway, of invasion (ADAM, MMP) and of bone resorption (SMAD, BMP, NFIB).

Then, we concatenated RQ values for each miRNA with the geneID of their protein targets and inserted these data in the G-language micro-array web application. This application connects miRNA targets according to their involvement in similar KEGG pathways. Supporting Information Figure S2 showed that two essential pathways were inhibited by the discriminating miRNAs: (i) the TGF β pathway, stimulating bone and tumor cell proliferation,²⁴ and associating with higher osteosarcoma grade²⁵; (ii) the Wnt pathway, shown to control osteoblast and osteoclast differentiation²⁶ and osteosarcoma invasiveness.²⁷ These two pathways were similarly targeted by the discriminating miRNA identified in osteosarcoma tumors (Supporting Information Fig. S2 left panel) as well as in rat model in response to ifosfamide (Supporting Information Fig. S2 right panel). In consequence, it showed

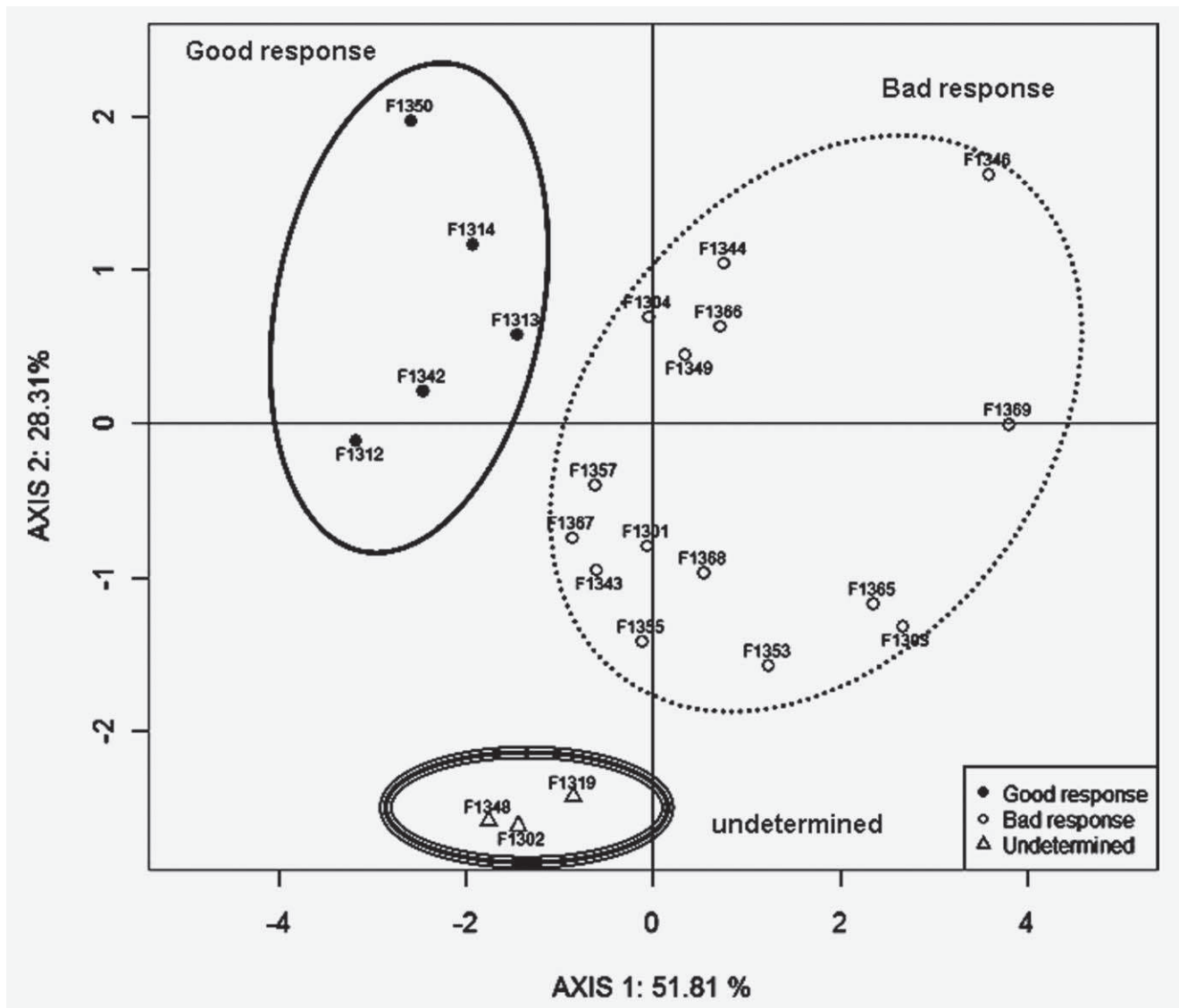


Figure 5. Predictive response of osteosarcoma tumors was consistent with hierarchical clustering. This diagram represents a PCA of the five miRNAs, which signed tumor response to treatment and were stably expressed in all tumor samples. The arrow represents the contribution of each miRNA expression level to tumor response as described in Material and Methods. Tumors represented in left group were predicted sensitive to treatment (●), while those in right panel were considered as resistant (○). Tumors with very low value for component 2 were considered undetermined (△).

that even if the miRNA profiles were not fully similar in our three models, the signaling pathways they targeted were identical. In future, these few miRNAs could constitute biomarkers of choice for osteosarcoma diagnosis; some of them, or the proteins they repress, could also constitute promising therapeutic targets as an alternative for chemotherapeutic treatment.

Discussion

miRNA signatures are observed for many types of cancers, that is, sarcoma,¹² breast and prostate cancers.^{11,28} These signatures constitute potent diagnosis and prognosis tools for chronic lymphocytic leukaemia,¹³ colon adenocarcinoma²⁹ or

lung cancers.³⁰ It was the first study showing that osteosarcoma also expressed miRNA patterns, which allow us to discriminate tumor response to chemotherapeutic agents (Fig. 4) and to predict it with our algorithms (Fig. 5). This miRNA signature was observed in osteosarcoma from patients (Fig. 4) but also in the two models that we possess, the rat osteosarcoma model (Fig. 2) and the osteosarcoma cell lines (Fig. 6).

Tumor miRNA profiles could constitute a potent predictive tool since supervised hierarchical clustering was perfectly correlated to tumor response predictions calculated by a PCA for human osteosarcoma tumors (Fig. 5) as well as rat osteosarcoma model (Fig. 3). Thus, it seems that miRNA profiling could surpass the current limitations for osteosarcoma

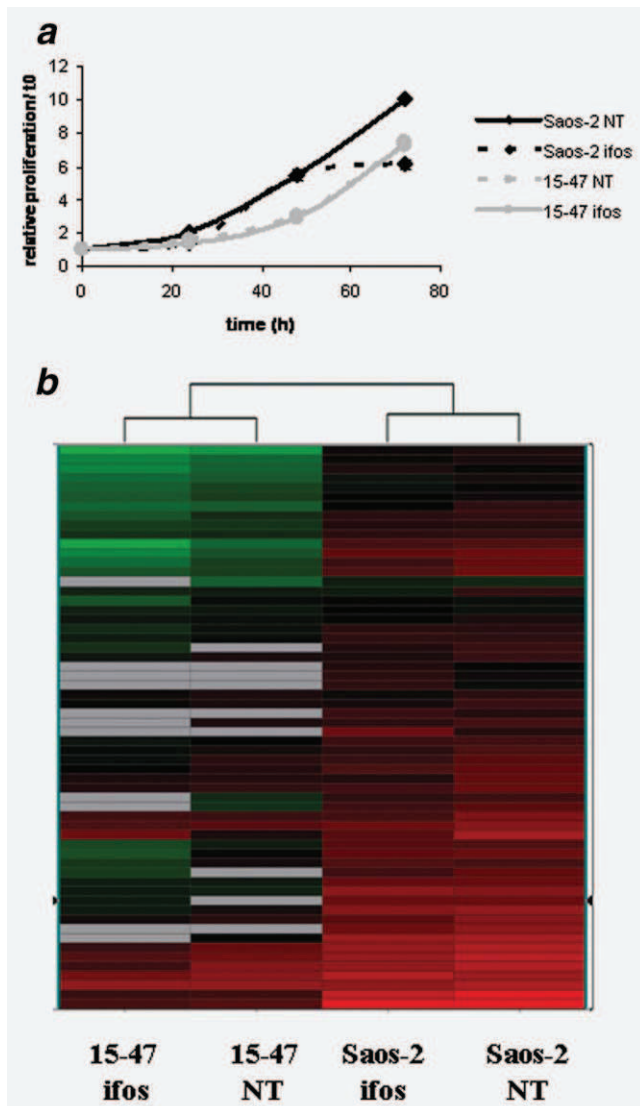


Figure 6. Relationship between osteosarcoma cell response to ifosfamide and their miRNA profiles. Saos-2 and 15-47 osteosarcoma cells were exposed to 50 μ M ifosfamide (ifos) or not (NT). A: Cell growth was measured by the Cell Titer GloLuminescent assay as described in materials and methods 24, 48 and 72 hr after drug exposure. Results were represented as the mean of relative proliferation normalized to time 0 of two independent experiments realized in duplicate. B: hierarchical clustering of miRNA profiles in osteosarcoma cells from cards A and B. This clustering was realized after an ANOVA (resistant, sensitive) ($p < 0.03$).

diagnosis since molecular abnormality detection appeared not fully adapted for this pathology. Indeed, osteosarcoma is characterized by complex chromosomal abnormalities in 70% of cases instead of specific chromosomal translocation or molecular defects, as it is observed for other sarcomas.³¹ In fact, osteosarcoma is the consequence of various sequential events leading to a complex karyotype³² and tumors with a similar genetic background could differently respond to treatment. We confirmed this concept through CGH array performed

with the rat tumors previously used for miRNA profiling. The majority of chromosomal aberrations observed in CGH array was common to untreated tumors and treated tumors, regardless of their response to treatment (data not shown). No miRNA were found in these regions, suggesting that the different miRNA profiles were closely linked to tumor response to treatment and not due to upstream chromosomal rearrangements.

miRNA signatures differed for all models but some miRNAs were common: miR-132, miR-99b and miR-193 for rat and human osteosarcoma, miR-132, miR-92a and miR-99b for osteosarcoma cell lines and patient tumors (Figs. 3a and 4). Interestingly, miR-92a belongs to miR-17-92 cluster which signs pancreas, stomach, prostate and colon cancers.^{10,28,33} For our interest, this anti-angiogenic factor seems to be upregulated in activated osteoblasts^{34,35}; miR-132 is implicated in osteoarthritis and rheumatoid arthritis.³⁶ miR-193 is poorly described but might be involved in caspase activation.³⁷ Of great interest, circulating miR-92a and miR-99b could be detected in the serum of ovarian cancer patient¹⁸ and colon cancer patient for miR-92a.³⁸ In fact, a non-invasive diagnosis tool based on miRNA profiling from patient serum could be a major advance for osteosarcoma, requiring a biopsy for its diagnosis, which could result in a secondary amputation. These past 2 years, numbers of studies identified circulating miRNAs as potential clinical biomarkers for ovarian cancer¹⁸ or breast cancer patients³⁹ as compared to healthy donors. We have recently succeeded in the detection of let-7a and miR-142 in the serum of sarcoma patients with the TaqMan[®] technology.

In addition, even if miRNA profiles were not fully similar between all models, we showed that all miRNA signatures potentially targeted the same signaling pathways, in particular the calcium signaling, the MAP kinase, the TGF β and the Wnt pathways (Supporting Information Fig. S2). Of great interest, these pathways regulate bone resorption and osteogenesis⁴⁰ and are deregulated in chemotherapy resistant pediatric osteosarcoma.⁴¹ Besides a clinical trial has been opened in Korea to test if the level of blood TGF β could predict osteosarcoma response to chemotherapy. On the basis of our data, we could envisage a therapy targeting the MAP kinase, the TGF β and the Wnt pathways for which some specific inhibitors are currently tested in clinical trials: the MAP Kinase Kinase MEK inhibitors AKD6244 or GSK20212, the anti-TGF β GC1008, or the anti-Dkk1 BHQ880.

Interestingly, the KEGG pathway analysis we performed with the miRNAs discriminating poor and good responders showed that various pathways implicated in immune response were altered, especially the Fc γ mediated phagocytosis and the natural killer mediated cytotoxicity. This was in accordance with the SSH analysis published by S. Salas with the same osteosarcoma tumors, suggesting that IFITM2, IFITM3, HLA-DRA, HLA-DQA1, HLA-DQB1 and LY6E were differentially expressed in poor responder tumors.⁴² This could explain that the response of fragments collected

from the same tumor varied in the rat model (Supporting Information Table S1). Indeed, all tumors in our study were grafted from the same initial osteogenic tumor, called "Tum ori." As shown in Figure 1, Tum ori clustered with the sensitive tumors, while tumors derived from this latter could be sensitive or resistant. This was not due to different genetic aberrations, since CGH arrays were very similar for all tumors. The first experiments we performed consisting to graft tumor fragments in the bone cavity revealed that miRNA profile was not altered in this intraosteal model as compared to the paraosteal model. This excludes that tumor response depends on the site of tumor implantation. Although we could not rule out that the different zones of a tumor could be more or less osteogenic and aggressive, we could hypothesize that tumor environment (microvessel density or immune cell infiltration) was a major determinant of tumor sensitivity to treatment. This was corroborated by a number of recent studies exploring the impact of immune cells on osteosarcoma progression, that is, natural killers,⁴³ tumor-infiltrating lymphocytes⁴⁴ or killer dendritic cells⁴⁵ and also by the use of immunotherapy for osteosarcoma treatment.⁴⁶

To resume, our study highlighted promising therapeutic targets, a panel of miRNAs whose expression level might determine osteosarcoma susceptibility to treatment but also the proteins they target. In future, we would like to confirm the implication of these miRNAs in treatment response *in vitro* through the use of miRNA mimics or inversely of Locked Nucleic Acid against these miRNAs. As mentioned in this study, we possess an interesting *in vitro* osteosarcoma model on which we could test the miRNA functionality in the presence of different drugs. The identification of a panel of miRNAs specifically involved in the response to one drug

could highlight new targeted therapy strategies to encounter resistance to treatment, we could test on these cell lines and also *in vivo*.

To conclude, we possess a reliable TLDA platform for osteosarcoma diagnosis. This platform is practical and easily transferable to another laboratory. This technology was highly reproducible since analyses with RNA samples obtained in an independent manner perfectly clustered altogether (Fig. 1). This analysis revealed a distributional similarity (Supporting Information Fig. S1a) and a strong correlation between two samples from a similar origin (Supporting Information Fig. S1b). It generates relevant biological data (Figs. 2, 4, and 6), in accordance with predictions obtained with our algorithms (Figs. 3 and 5) for the rat model and the patient osteosarcoma tumors. We also showed the reliability of miRNA measurement by TLDA on FFPE samples, as it was previously shown by Daniel Pissaloux in our laboratory (in press) and others.^{23,47,48} Even if these data might be confirmed with others samples to obtain a bigger cohort of patient tumors, we envisage to realize a prospective study with patient tumors. The great advantage of this platform is its need for a low level of total RNA (500 ng), which is highly required for osteosarcoma diagnosis from micro-biopsies of very small volumes. This allows us detecting circulating miRNAs in patient serum.

Altogether, these promising results open up the way to a new diagnosis tool, which could decrease the intensity of pre-operative chemotherapy cures and thus improve patient survival.

Acknowledgements

We are very grateful to Corinne Bouvier and Sébastien Salas from the La Timone Hospital for providing osteosarcoma tumor samples.

References

- He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet* 2004;5:522–31.
- Hutvagner G, Zamore PD. A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 2002;297:2056–60.
- Lee Y, Jeon K, Lee JT, Kim S, Kim VN. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J* 2002;21:4663–70.
- Doench JG, Sharp PA. Specificity of microRNA target selection in translational repression. *Genes Dev* 2004;18:504–11.
- Zhao Y, Srivastava D. A developmental view of microRNA function. *Trends Biochem Sci* 2007;32:189–97.
- Li Z, Hassan MQ, Jafferji M, Aqeilan RI, Garzon R, Croce CM, van Wijnen AJ, Stein JL, Stein GS, Lian JB. Biological functions of miR-29b contribute to positive regulation of osteoblast differentiation. *J Biol Chem* 2009;284:15676–84.
- Sugatani T, Hruska KA. MicroRNA-223 is a key factor in osteoclast differentiation. *J Cell Biochem* 2007;101:996–9.
- Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. *Nat Rev Cancer* 2006;6:259–69.
- Lu J, Getz G, Miska EA, varez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, et al. MicroRNA expression profiles classify human cancers. *Nature* 2005;435:834–8.
- Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 2006;103:2257–61.
- Israel A, Sharan R, Ruppin E, Galun E. Increased microRNA activity in human cancers. *PLoS One* 2009;4:e6045.
- Subramanian S, Lui WO, Lee CH, Espinosa I, Nielsen TO, Heinrich MC, Corless CL, Fire AZ, van de RM. MicroRNA expression signature of human sarcomas. *Oncogene* 2008;27:2015–26.
- Calin GA, Ferracin M, Cimmino A, Di LG, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med* 2005;353:1793–801.
- Wittig JC, Bickels J, Priebat D, Jelinek J, Kellar-Graney K, Shmookler B, Malawer MM. Osteosarcoma: a multidisciplinary approach to diagnosis and treatment. *Am Fam Physician* 2002;65:1123–32.
- Ferrari S, Palmerini E. Adjuvant and neoadjuvant combination chemotherapy for osteogenic sarcoma. *Curr Opin Oncol* 2007;19:341–6.
- Chou AJ, Gorlick R. Chemotherapy resistance in osteosarcoma: current

- challenges and future directions. *Expert Rev Anticancer Ther* 2006;6:1075–85.
17. Lodes MJ, Caraballo M, Suci D, Munro S, Kumar A, Anderson B. Detection of cancer with serum miRNAs on an oligonucleotide microarray. *PLoS One* 2009;4:e6229.
 18. Resnick KE, Alder H, Hagan JP, Richardson DL, Croce CM, Cohn DE. The detection of differentially expressed microRNAs from the serum of ovarian cancer patients using a novel real-time PCR platform. *Gynecol Oncol* 2009;112: 55–9.
 19. Dutour A, Monteil J, Paraf F, Charissoux JL, Kaletta C, Sauer B, Naujoks K, Rigaud M. Endostatin cDNA/cationic liposome complexes as a promising therapy to prevent lung metastases in osteosarcoma: study in a human-like rat orthotopic tumor. *Mol Ther* 2005;11:311–9.
 20. Dutour A, Leclers D, Monteil J, Paraf F, Charissoux JL, Rousseau R, Rigaud M. Non-invasive imaging correlates with histological and molecular characteristics of an osteosarcoma model: application for early detection and follow-up of MDR phenotype. *Anticancer Res* 2007;27:4171–8.
 21. Allouche M, Delbruck HG, Klein B, Masse R, Matar A, Morin M, Lafuma J, Jasmin C. Malignant bone tumours induced by a local injection of colloidal radioactive ¹⁴⁴Cerium in rats as a model for human osteosarcomas. *Int J Cancer* 1980;26: 777–82.
 22. Arakawa K, Kono N, Yamada Y, Mori H, Tomita M. KEGG-based pathway visualization tool for complex omics data. *In Silico Biol* 2005;5:419–23.
 23. Liu A, Tetzlaff MT, Vanbelle P, Elder D, Feldman M, Tobias JW, Sepulveda AR, Xu X. MicroRNA Expression Profiling Outperforms mRNA Expression Profiling in Formalin-fixed Paraffin-embedded Tissues. *Int J Clin Exp Pathol* 2009;2: 519–27.
 24. Linkhart TA, Mohan S, Baylink DJ. Growth factors for bone growth and repair: IGF, TGF beta and BMP. *Bone* 1996;19: 1S–12S.
 25. Franchi A, Arganini L, Baroni G, Calzolari A, Capanna R, Campanacci D, Caldora P, Masi L, Brandi ML, Zampi G. Expression of transforming growth factor beta isoforms in osteosarcoma variants: association of TGF beta 1 with high-grade osteosarcomas. *J Pathol* 1998;185:284–9.
 26. Glass DA, Bialek P, Ahn JD, Starbuck M, Patel MS, Clevers H, Taketo MM, Long F, McMahon AP, Lang RA, Karsenty G. Canonical Wnt signaling in differentiated osteoblasts controls osteoclast differentiation. *Dev Cell* 2005;8:751–64.
 27. Hoang BH, Kubo T, Healey JH, Yang R, Nathan SS, Kolb EA, Mazza B, Meyers PA, Gorlick R. Dickkopf 3 inhibits invasion and motility of Saos-2 osteosarcoma cells by modulating the Wnt-beta-catenin pathway. *Cancer Res* 2004;64:2734–9.
 28. Ambs S, Prueitt RL, Yi M, Hudson RS, Howe TM, Petrocca F, Wallace TA, Liu CG, Volinia S, Calin GA, Yfantis HG, Stephens RM, et al. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer Res* 2008;68: 6162–70.
 29. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, Yuen ST, Chan TL, Kwong DL, Au GK, Liu CG, Calin GA, et al. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* 2008;299:425–36.
 30. Yu SL, Chen HY, Chang GC, Chen CY, Chen HW, Singh S, Cheng CL, Yu CJ, Lee YC, Chen HS, Su TJ, Chiang CC, et al. MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell* 2008;13: 48–57.
 31. Fletcher JA, Gebhardt MC, Kozakewich HP. Cytogenetic aberrations in osteosarcomas. Nonrandom deletions, rings, and double-minute chromosomes. *Cancer Genet Cytogenet* 1994;77:81–8.
 32. Biegel JA, Womer RB, Emanuel BS. Complex karyotypes in a series of pediatric osteosarcomas. *Cancer Genet Cytogenet* 1989;38:89–100.
 33. Motoyama K, Inoue H, Takatsuno Y, Tanaka F, Mimori K, Uetake H, Sugihara K, Mori M. Over- and under-expressed microRNAs in human colorectal cancer. *Int J Oncol* 2009;34:1069–75.
 34. Bonauer A, Carmona G, Iwasaki M, Mione M, Koyanagi M, Fischer A, Burchfield J, Fox H, Doebele C, Ohtani K, Chavakis E, Potente M, et al. MicroRNA-92a controls angiogenesis and functional recovery of ischemic tissues in mice. *Science* 2009;324: 1710–3.
 35. Palmieri A, Pezzetti F, Brunelli G, Martinelli M, Lo ML, Scarano A, Degidi M, Piattelli A, Carinci F. Peptide-15 changes miRNA expression in osteoblast-like cells. *Implant Dent* 2008;17:100–8.
 36. Murata K, Yoshitomi H, Tanida S, Ishikawa M, Nishitani K, Ito H, Nakamura T. Plasma and synovial fluid microRNAs as potential biomarkers of rheumatoid arthritis and osteoarthritis. *Arthritis Res Ther* 2010;12:R86.
 37. Ovcharenko D, Kelnar K, Johnson C, Leng N, Brown D. Genome-scale microRNA and small interfering RNA screens identify small RNA modulators of TRAIL-induced apoptosis pathway. *Cancer Res* 2007;67: 10782–8.
 38. Huang Z, Huang D, Ni S, Peng Z, Sheng W, Du X. Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *Int J Cancer* 2010;127:118–26.
 39. Zhu W, Qin W, Atasoy U, Sauter ER. Circulating microRNAs in breast cancer and healthy subjects. *BMC Res Notes* 2009;2:89.
 40. Kingsley LA, Fournier PG, Chirgwin JM, Guise TA. Molecular biology of bone metastasis. *Mol Cancer Ther* 2007;6: 2609–17.
 41. Mintz MB, Sowers R, Brown KM, Hilmer SC, Mazza B, Huvo AG, Meyers PA, Lafleur B, McDonough WS, Henry MM, Ramsey KE, Antonescu CR, et al. An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Res* 2005;65:1748–54.
 42. Salas S, Jezequel P, Campion L, Deville JL, Chibon F, Bartoli C, Gentet JC, Charbonnel C, Gouraud W, Voutsinos-Porche B, Bouchet A, Duffaud F, et al. Molecular characterization of the response to chemotherapy in conventional osteosarcomas: predictive value of HSD17B10 and IFITM2. *Int J Cancer* 2009; 125:851–60.
 43. Luksch R, Perotti D, Cefalo G, Gambacorti PC, Massimino M, Spreafico F, Casanova M, Ferrari A, Terenziani M, Polastri D, Gambirasio F, Podda M, et al. Immunomodulation in a treatment program including pre- and post-operative interleukin-2 and chemotherapy for childhood osteosarcoma. *Tumori* 2003;89: 263–8.
 44. Theoleyre S, Mori K, Cherrier B, Passuti N, Gouin F, Redini F, Heymann D. Phenotypic and functional analysis of lymphocytes infiltrating osteolytic tumors: use as a possible therapeutic approach of osteosarcoma. *BMC Cancer* 2005;5:123.
 45. Chauvin C, Philippeau JM, Hemont C, Hubert FX, Wittrant Y, Lamoureux F, Trinite B, Heymann D, Redini F, Josien R. Killer dendritic cells link innate and adaptive immunity against established osteosarcoma in rats. *Cancer Res* 2008;68: 9433–40.
 46. Mori K, Redini F, Gouin F, Cherrier B, Heymann D. Osteosarcoma: current status of immunotherapy and future trends (Review). *Oncol Rep* 2006;15:693–700.
 47. Hui AB, Shi W, Boutros PC, Miller N, Pintilie M, Fyles T, McCready D, Wong D, Gerster K, Waldron L, Jurisica I, Penn LZ, et al. Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. *Lab Invest* 2009;89: 597–606.
 48. Leite KR, Canavez JM, Reis ST, Tomiyama AH, Piantino CB, Sanudo A, Camara-Lopes LH, Srougi M. miRNA analysis of prostate cancer by quantitative real time PCR: Comparison between formalin-fixed paraffin embedded and fresh-frozen tissue. *Urol Oncol* 2009.

III. APPORTS DES PROFILS MICROARN A LA CLASSIFICATION DES RHABDOMYOSARCOMES

1. Difficultés liées à la détection des transcrits rares : essai de développement d'une technologie de puce à ADNc

Les RMS peuvent être classés en deux groupes majoritaires de pronostics très différents, les ARMS et les ERMS. Les ARMS, de pronostic plus défavorable, sont caractérisés par la présence de transcrits de fusion induisant la production de protéines chimériques. A ce jour, 5 transcrits de fusion différents impliquant des partenaires variés ont été décrits, mais en laboratoire, seuls les deux réarrangements les plus courants sont recherchés en routine par les techniques classiques de biologie moléculaire : les transcrits de fusion PAX3-FOXO1 et PAX7-FOXO1 par technique PCRq, et les réarrangements du gène FOXO1 par technique FISH. Les limitations propres à ces deux techniques font que les transcrits de type PAX3-FOXO4, PAX3-NCOA1 et PAX3-NCOA2 ne sont pas détectées. On peut alors penser qu'à l'heure actuelle, une partie des ARMS de diagnostic anatomopathologique difficile, pour lesquels il n'est pas retrouvé de transcrits de fusion de type PAX3-FOXO1 et PAX7-FOXO1, ou de réarrangement du gène FOXO1, sont finalement classés dans le groupe des ERMS, alors qu'il pourrait en fait s'agir d'ARMS à transcrits de fusion rares qui nécessiteraient une prise en charge thérapeutique plus lourde.

Avant d'aborder la problématique du diagnostic différentiel entre ARMS et ERMS par le biais des profils d'expression des microARN, nous avons essayé de développer un autre outil diagnostique qui nous permettrait de détecter la totalité des transcrits de fusion des ARMS. Cet outil était basé sur la fabrication d'une puce à ADN maison contenant des sondes correspondant aux séquences de quatre des transcrits de fusion décrits pour les ARMS : PAX3-FOXO1, PAX7-FOXO1, PAX3-

FOXO4 et PAX3-NCOA1. Sur cette puce à ADN seraient hybridés les produits d'amplification d'une RT-PCR ancrée et multiplexe. Le principal avantage de cette technologie était sa capacité à détecter l'intégralité des séquences déposées sur la puce en une seule expérience. Cela lui permet donc d'être particulièrement économe en quantité de matériel biologique de départ. En raison de cette caractéristique, le nombre de séquences correspondant aux transcrits de fusion déposés sur la puce a été étendu pour permettre la détection de l'ensemble des transcrits de fusion des sarcomes pédiatriques à petites cellules, à savoir les rhabdomyosarcomes, les synovialosarcomes et les tumeurs d'Ewing. Ce projet, établi en collaboration avec la société Ipsogen basée à Marseille, a été réalisé au sein du réseau européen d'excellence CONTICANET, financé au titre du 7^{ème} programme-cadre de l'Union Européenne (FP7/2007-2013, agrément n°278742, code contrat : FP-018806).

Les premiers tests d'une puce prototype ont montré que cette puce était capable de détecter avec spécificité le transcrit de fusion de type PAX3-FOXO1 présent dans la lignée cellulaire de RMS utilisée.

Malgré des résultats prometteurs, ce projet a finalement été abandonné en raison d'un partenariat rendu difficile par la restructuration de la société Ipsogen, qui n'était plus en mesure de produire la puce à ADN. Il n'a pas été possible de trouver une solution alternative. L'intégralité du travail réalisé autour de ce projet de thèse a toutefois été récompensé par le prix des Amis de l'Université de Lyon, dans la catégorie Nanotechnologie et Génétique Moléculaire, Pharmacogénomique.

Document à suivre :

Prix 2010 des Amis de l'Université de Lyon, doté par BIOMNIS, dans la catégorie Nanotechnologie et Génétique Moléculaire, Pharmacogénomique, pour le sujet de thèse : Développement d'une puce à ADN de diagnostic des sarcomes à translocation.

Daniel PISSALOUX

Etudiant en thèse de Cancérologie
Université de Lyon - Ecole doctorale BMIC



À l'attention des membres du comité d'évaluation
de l'Association des Amis de l'Université de Lyon
pour le prix 2010 doté par le laboratoire Biomnis

CONFIDENTIEL

Développement d'une puce à ADN de diagnostic des sarcomes à translocation

Description détaillée du projet de recherche - Février 2010

Directeur de thèse : Jean-Yves Blay, PUPH, Ph. D.
Codirecteur de thèse : Laurent Alberti, Ph. D.

INSERM U590
Equipe Cytokines et Cancer / Département de Biopathologie
Centre Léon Bérard
28 rue Laënnec
69373 LYON Cedex 08



INTRODUCTION

La rareté et la complexité histologique des sarcomes font qu'ils nécessitent une prise en charge organisée et multidisciplinaire dans des centres spécialisés. L'immunohistochimie a contribué à clarifier la classification des sarcomes, qui est depuis peu complétée par des diagnostics moléculaires pour certains sarcomes dits à génétique simple. Nous nous proposons, dans le cadre de ce projet de thèse, de réaliser une puce à ADN de diagnostic de ces sarcomes, qui permettra la détection des translocations chromosomiques, et plus précisément des ARN messagers chimériques associés appelés transcrits de fusion.

Cette nouvelle technologie permettra de compléter, d'améliorer ou de remplacer les techniques actuelles utilisées pour le diagnostic en routine.

Ces travaux sont réalisés au Centre régional de lutte contre le cancer Léon Bérard (CLB), et plus précisément au sein du laboratoire Cytokines et Cancer de l'INSERM U590.

1. Contexte et motivations du projet de thèse

1.1. L'importance du diagnostic initial des sarcomes

Les sarcomes sont considérés comme des pathologies rares en cancérologie, puisqu'ils représentent environ 1 à 3% des tumeurs malignes de l'adulte. On enregistre 4500 à 5000 nouveaux cas par an en France¹. Ils constituent une entité très hétérogène tant sur le plan de la localisation que de l'histologie. Plus d'une cinquantaine de types majeurs de sarcomes ont été identifiés^{2, 3}. En incluant les sous-types, l'OMS a répertorié au total plus de 150 entités différentes. Les sarcomes sont des tumeurs curables dans 50% des cas environ. Cependant, la rareté et la complexité de ces pathologies réservent leur prise en charge à des équipes multidisciplinaires dans des centres spécialisés, et ce dès le diagnostic initial^{4, 5, 6}.

Dans une étude rétrospective réalisée en 2002 à partir de dossiers de patients traités entre 1999 et 2001, le CLB a évalué la prise en charge des patients atteints de sarcomes des tissus mous selon les recommandations nationales⁷. Cette étude a montré que la prise en charge des sarcomes suivait les recommandations de pratiques cliniques pour 30% des patients, notamment en ce qui concerne la prise en charge initiale de ces sarcomes (diagnostic et premier geste chirurgical)⁸. Ces observations sont semblables à celles recueillies dans d'autres études internationales réalisées dans des pays occidentaux. De plus, il a été montré que la mise en place de concertations multidisciplinaires augmentait le taux de conformité des pratiques par rapport aux recommandations nationales et était associée, dans cette étude rétrospective, à une amélioration de la survie sans rechute des patients. Il apparaît donc très clairement que le diagnostic initial est un facteur primordial pour le bien-être du patient.

1.2. Le diagnostic moléculaire des sarcomes et ses limitations

Depuis plus de 20 ans, un nombre croissant d'anomalies chromosomiques sont décrites dans les sarcomes. Pour certains types de tumeurs, ces anomalies sont suffisantes et spécifiques pour avoir une valeur diagnostique. A ce jour, la biologie moléculaire est une aide bien établie pour les pathologistes. Elle concerne une dizaine de types de sarcomes et permet de consolider un diagnostic obtenu sur des bases morphologiques et immunohistochimiques. Ces informations se révèlent tout particulièrement utiles dans le cas des micro-prélèvements, et dans le cas de prélèvements peu représentatifs ou peu différenciés^{9, 10}. La biologie moléculaire permet ainsi le diagnostic de 35% des sarcomes à anomalie génétique simple.

Ainsi, mon laboratoire d'accueil s'est spécialisé dans le diagnostic moléculaire en routine de sept grands types de sarcomes à génétique simple : le sarcome d'Ewing, le synoviosarcome, le rhabdomyosarcome, le liposarcome bien différencié ou dédifférencié, le liposarcome myxoïde, le sarcome à cellules claires et les tumeurs desmoplastiques à petites cellules rondes. Ces diagnostics sont basés sur la détection d'amplifications ou de translocations de gènes par PCR quantitative en temps réel^{11, 12, 13}. Cette technique est réalisée à partir de matériel biologique congelé ou d'échantillons fixés au formol et inclus en paraffine (FFPE), en dépit des désavantages que ce dernier type de fixation implique : présence d'inhibiteurs de PCR, dégradations et modifications chimiques des acides nucléiques¹⁴.

Cette technique, bien que performante, a ses limites. Aujourd'hui, à partir d'un échantillon inclus en paraffine, il faut utiliser pas moins de treize sets différents de RT-PCRq pour diagnostiquer un sarcome pédiatrique difficile, lorsqu'il existe un doute entre un synoviosarcome, un sarcome d'Ewing ou un rhabdomyosarcome (annexe 1). Malgré ce nombre important de tests, il est en plus possible de passer à côté de certains types rares. Cette technique de routine se révèle donc très consommatrice en matériel biologique. En particulier, pour les sarcomes dont la famille comprend de nombreux sous-types de translocations, il n'est pas rare de ne pas pouvoir rendre de résultat par manque de matériel biologique de départ. Les sarcomes pédiatriques sont particulièrement touchés par ce problème, car les chirurgiens essaient dans la mesure du possible d'effectuer des micro-prélèvements les moins invasifs possibles.

D'autres alternatives existent pour la détection de transcrits de fusion dans les sarcomes, avec leurs avantages et leurs inconvénients (tableau 1)¹⁵. Pour l'instant, aucune d'elles ne combine les avantages que promettrait une approche plus globale de type puce à ADN. En effet, un screening de toutes les entités connues en une seule expérience permet un rendu de résultat plus rapide, une consommation moindre en matériel biologique et une réduction des coûts financiers. De plus, grâce au principe d'hybridation de séquences nucléiques complémentaires, il est possible d'obtenir une précision à l'échelle des gènes et même des exons impliqués dans les translocations recherchées.

		Avantages	Inconvénients
Cytogénétique	Caryotype	- Détection des anomalies de nombre - Détection des translocations non connues	- Technique longue et difficile - Sensibilité insuffisante
	FISH	- Bonne sensibilité - Technique facile et rapide	- Détection des seules translocations connues - Coût élevé des sondes fluorescentes - un seul test (sonde) par expérience
	Multi-FISH	- Détection des translocations non connues	- Mise en évidence des chromosomes impliqués mais pas des gènes précis
Biologie Moléculaire	PCR simplex	- Technique rapide et sensible - Mise en évidence des gènes et des exons précis impliqués	- Translocations connues seulement - un seul test (amorce) par expérience - Coûteux en matériel biologique si plusieurs partenaires
	PCR multiplex	- Plusieurs tests par expérience - Economie en matériel biologique	- Précision inférieure à la PCR simplex - Très rapidement limité par des problèmes de réactions aspécifiques

Tableau 1 : méthodes actuelles les plus utilisées pour la détection des transcrits de fusion. Les techniques de cytogénétique et les techniques de biologie moléculaire présentent chacune des avantages et des inconvénients distincts détaillés dans ce tableau ci-dessus.

2. Objectifs scientifiques du projet de thèse

Au vu des divers inconvénients liés aux techniques de diagnostic moléculaire déjà existantes, mes travaux de thèse consistent à développer une approche plus globale de caractérisation moléculaire des sarcomes à translocations, basée sur une technologie de puce à ADN.

Plus précisément, je m'attacherai à développer et à valider une puce à ADN de diagnostic fiable, capable de fonctionner à partir de matériel génétique extrait aussi bien de tissus frais ou congelés, qu'à partir de tissus fixés en paraffine. Une seule expérience d'hybridation doit suffire à la détection de la totalité des transcrits de fusion répertoriés aujourd'hui dans les sarcomes à génétique simple. Cette nouvelle technique sera donc très économe en termes temps de manipulation et de matériel biologique.

La première étape de mon travail de thèse a été de dresser la liste complète des points de cassure précis des translocations déjà référencées dans la littérature. Ce travail nous a permis de définir précisément l'ensemble des séquences "sondes" à immobiliser sur la puce. Après avoir choisi et affiné notre stratégie d'amplification et de marquage des cibles, les amorces de RT et de PCR nécessaires à ces étapes ont été dessinées. Une puce prototype a alors été fabriquée sur un nombre restreint d'entités. Cette première puce nous a permis de vérifier la faisabilité de notre projet, et a tout d'abord été évaluée à l'aide de lignées cellulaires de sarcomes bien établies. Dans un second temps, nous envisageons de tester plus précisément sa sensibilité et sa spécificité à partir d'échantillons de patients déjà diagnostiqués et caractérisés au niveau moléculaire au Centre Léon Bérard. Cela nous permettra d'évaluer sa propension à être utilisé comme outil de diagnostic moléculaire de routine des sarcomes à translocations.

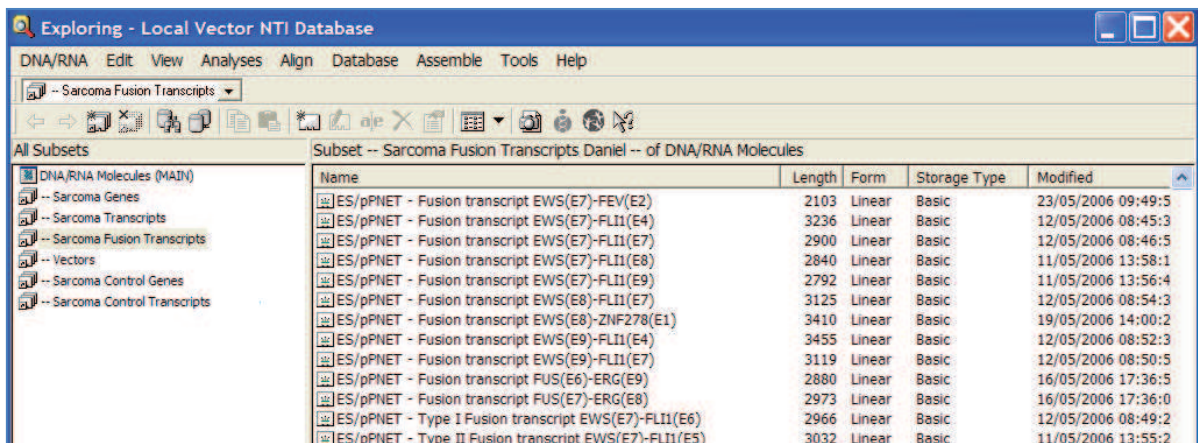
Enfin, nous élaborerons le cahier des charges d'une version pré-finale de la puce, en tenant compte de nos résultats et en incorporant l'ensemble des translocations initialement répertoriées et absentes sur la puce prototype. Cette nouvelle version de puce sera de nouveau testée de façon exhaustive ce qui nous permettra, après de derniers ajustements, d'obtenir la version finale de la puce à ADN de diagnostic des sarcomes à translocations. En cas de succès lors de son application en routine au Centre Léon Bérard, le Réseau d'Excellence Européen Conticanet contribuera à la diffusion de cette nouvelle technologie dans les principaux centres anticancéreux Européens.

RESULTATS OBTENUS LORS DE LA PREMIERE ANNEE DE THESE

1. Travaux préliminaires : partie pré-analytique

1.1. Création d'une base de données exhaustive des séquences d'intérêt

L'intégralité des séquences des transcrits de fusion résultant des translocations connues à ce jour a été récupérée à partir des banques nucléiques publiques et des résultats de séquençages des points de cassure indiqués dans la littérature. Ces séquences ont alors été entrées dans une base de données spécialisée dans la manipulation et l'analyse de séquences nucléiques, à l'aide du logiciel Vector NTI® (figure 1).



Name	Length	Form	Storage Type	Modified
ES/pPNET - Fusion transcript EWS(E7)-FEV(E2)	2103	Linear	Basic	23/05/2006 09:49:5
ES/pPNET - Fusion transcript EWS(E7)-FLI1(E4)	3236	Linear	Basic	12/05/2006 08:45:3
ES/pPNET - Fusion transcript EWS(E7)-FLI1(E7)	2900	Linear	Basic	12/05/2006 08:46:5
ES/pPNET - Fusion transcript EWS(E7)-FLI1(E8)	2840	Linear	Basic	11/05/2006 13:58:1
ES/pPNET - Fusion transcript EWS(E7)-FLI1(E9)	2792	Linear	Basic	11/05/2006 13:56:4
ES/pPNET - Fusion transcript EWS(E8)-FLI1(E7)	3125	Linear	Basic	12/05/2006 08:54:3
ES/pPNET - Fusion transcript EWS(E8)-ZNF278(E1)	3410	Linear	Basic	19/05/2006 14:00:2
ES/pPNET - Fusion transcript EWS(E9)-FLI1(E4)	3455	Linear	Basic	12/05/2006 08:52:3
ES/pPNET - Fusion transcript EWS(E9)-FLI1(E7)	3119	Linear	Basic	12/05/2006 08:50:5
ES/pPNET - Fusion transcript FUS(E6)-ERG(E9)	2880	Linear	Basic	16/05/2006 17:36:5
ES/pPNET - Fusion transcript FUS(E7)-ERG(E8)	2973	Linear	Basic	16/05/2006 17:36:0
ES/pPNET - Type I Fusion transcript EWS(E7)-FLI1(E6)	2966	Linear	Basic	12/05/2006 08:49:2
ES/pPNET - Type II Fusion transcript EWS(E7)-FLI1(E5)	3032	Linear	Basic	11/05/2006 13:55:2

Figure 1 : extrait du contenu de la base de données de séquences d'intérêt sous le logiciel Vector NTI® commercialisé par la société Invitrogen.

Dans cette base, les séquences nucléiques de chacun des transcrits de fusion peuvent être annotées de façon très détaillée. Ainsi, les points de cassure ont été indexés, ainsi que les délimitations de chacun des exons, les variants d'épissage connus, les séquences codantes et enfin les amorces de RT-PCRq utilisées actuellement dans l'équipe de diagnostic moléculaire des sarcomes par PCR au Centre Léon Bérard (figure 2).

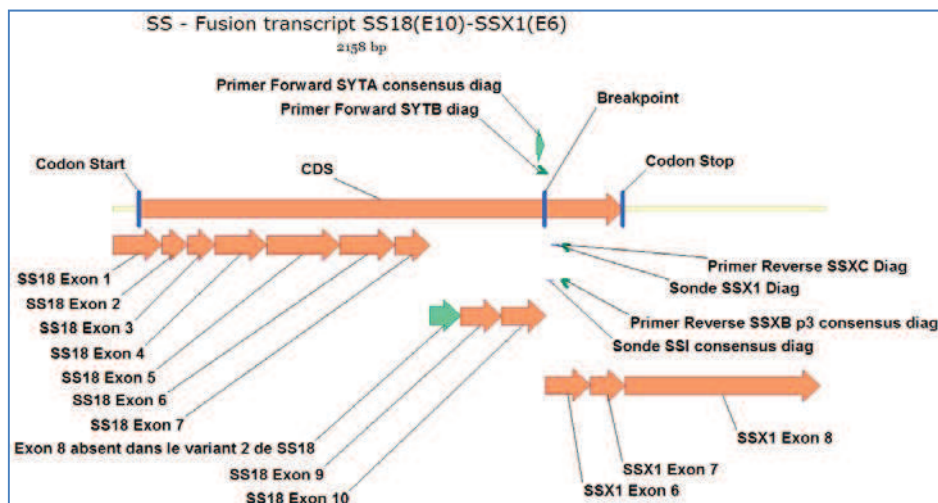


Figure 2 : représentation du transcript de fusion SYT(E10)-SSX1(E6) montrant les possibilités graphiques et les annotations contenues dans la base de données du logiciel Vector NTI®.

De plus, cette base de données contient les séquences ADN et ARN annotées de tous les gènes normaux impliqués dans les translocations présentes dans les sarcomes. A ce jour, cette base de données contient 14 familles de sarcomes représentant 159 transcrits de fusion différents. Elle contient également plusieurs centaines de séquences nucléiques correspondant aux gènes normaux non transloqués.

1.2. Création d'une banque des transcrits de fusion par clonage

Lors des futures étapes indispensables de tests de la puce à ADN de diagnostic, il se posera le problème de la disponibilité du matériel biologique de support. En effet, le matériel tumoral est précieux car difficile à obtenir sans justification diagnostique ou consentement éclairé écrit du patient concerné. Nous avons donc rapidement réfléchi à deux solutions alternatives : le clonage des transcrits de fusion déjà obtenus au Centre Léon Bérard via la technique de PCRq, et l'utilisation de lignées tumorales bien établies.

Dans cette optique, nous avons cloné avec succès cinq séquences de transcrits de fusion (SYT-SSX1, SYT-SSX2, PAX3-FKHR, PAX7-FKHR et EWS-FLI1), à partir d'ARN extraits d'échantillons tumoraux déjà diagnostiqués et positifs pour l'un des types de sarcomes recherchés par RT-PCRq. Ces transcrits de fusion ont été insérés dans des vecteurs pGEM®-T de chez Promega ou des vecteurs pCR®8/GW/TOPO® de chez Invitrogen (figure 3). Ces transcrits de fusion clonés seront ensuite utilisés lors des premiers tests de marquage et d'hybridation de la puce à ADN de diagnostic.

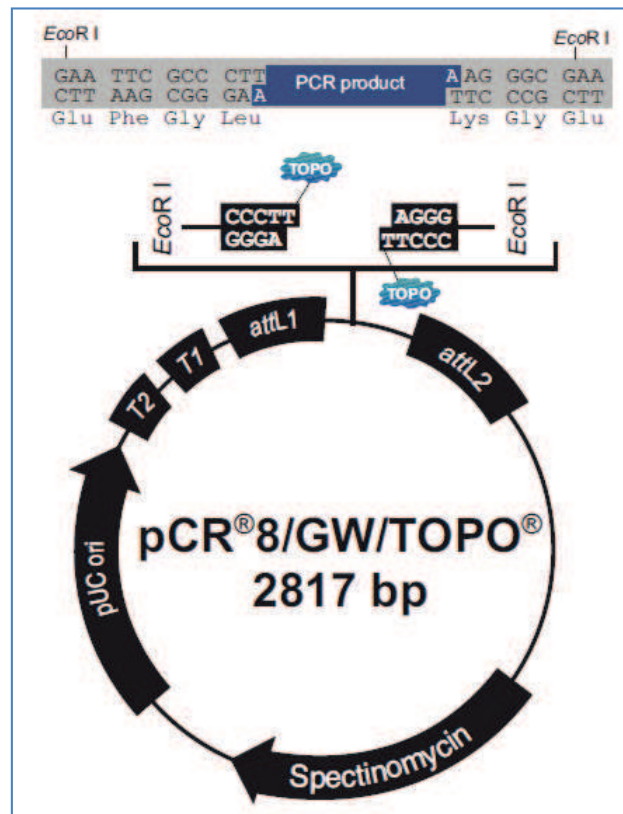


Figure 3 : schéma de l'un des vecteurs de clonage utilisé, le pCR®8/GW/TOPO® de chez Invitrogen. Un produit PCR (dans notre cas le transcrit de fusion d'intérêt) peut facilement être introduit dans ce vecteur au niveau des sites d'action de la topoisomérase. Ainsi, nous disposons de matériel immortalisé "inépuisable" pour nos phases de test ultérieures de la puce à ADN de diagnostic.

1.3. Principe et mode opératoire de la puce à ADN de diagnostic

Les réflexions sur le principe et le mode opératoire les plus adaptés pour une puce à ADN capable de détecter des transcrits de fusion ont été menées en collaboration avec la société Ipsogen (Marseille), qui disposait déjà d'une expertise reconnue dans ce domaine¹⁶. Ainsi, Ipsogen a suggéré un enrichissement des cibles avant leur détection par hybridation sur puce, selon une méthode précédemment publiée¹⁷. Les grandes lignes et limites de ce projet ont rapidement été posées. Il a alors été décidé, dans un premier temps, de réaliser une première puce de diagnostic centrée sur les sarcomes pédiatriques, à savoir les sarcomes d'Ewing, les synoviosarcomes et les rhabdomyosarcomes (liste des sous-types en annexe 2). Ces sarcomes ont l'énorme avantage de disposer tous trois d'un diagnostic via technique PCRq bien établi au Centre Léon Bérard. Il sera donc possible, lors des phases de test, de comparer les résultats obtenus entre notre technique innovante et la technique de référence. A l'issue de cette phase de faisabilité, la puce pourra être élargie à l'ensemble des sarcomes caractérisés par des translocations.

Afin de réduire au maximum les coûts de notre outil destiné à être utilisé en routine, il a été décidé que la puce serait constituée d'une membrane en nylon sur laquelle deux types de sondes seraient fixées (figure 4) :

- des sondes de recherche

Pour chacun des transcrits de fusion recherchés, deux sondes de recherche constituées d'oligonucléotides d'environ 50 mers, correspondant à des séquences localisées de part et d'autre du point de cassure, mais très proches de celui-ci (condition essentielle pour réussir à générer par PCR des cibles marquées à partir de matériel dégradé car fixé en paraffine).

- des sondes contrôles positifs

Les autres sondes présentes sur la puce (50 mers également) serviront de contrôles positifs, et correspondront à une séquence située en amont du point de cassure du gène partenaire situé du côté 3' du transcrit de fusion.

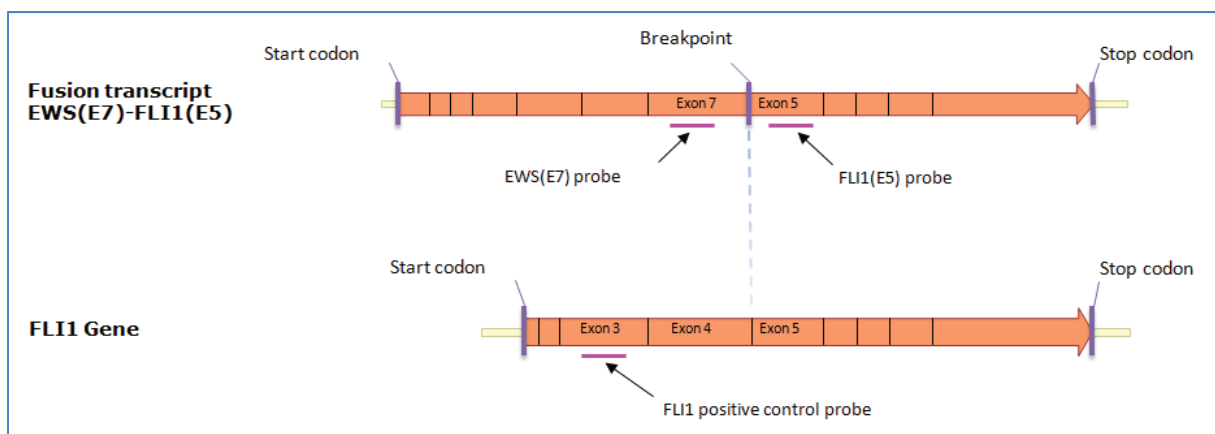


Figure 4 : schéma explicatif du choix des sondes, basé sur la détection du transcrit de fusion de type EWS(E7)-FLI1(E5) présent dans les sarcomes d'Ewing. Concernant cette entité, trois sondes seront créées : deux sondes de recherche EWS(E7) et FLI1(E5), de part et d'autre du point de cassure, et une sonde contrôle positif située en regard de l'exon 3 du gène FLI1 (l'exon le plus en amont des différents points de cassures connus pour le gène FLI1).

Le test diagnostique sera basé sur la détection après amplification des transcrits de fusion au moyen d'une transcription inverse de l'ARN du patient suivie d'une PCR ancrée. La technologie mise au point pour l'étape de traitement de l'échantillon garantira que tous les événements de réarrangement d'un gène donné puissent être détectés avec une bonne fiabilité, quelles que soient la longueur du transcrit et la position du point de fusion des deux gènes sur la séquence.

Brièvement, l'ARN total est extrait de l'échantillon congelé ou fixé en paraffine. Une transcription inverse (RT) multiplexe, basée sur l'utilisation d'amorces ancrées spécifiques des partenaires de la translocation du côté 3', est réalisée sur cet ARN total (figure 5). L'ADNc ainsi obtenu est amplifié par PCR multiplexe ancrée à l'aide d'amorces spécifiques du gène partenaire de la translocation du côté 5' (en amont du point de cassure) et d'amorces spécifiques de l'ancre utilisée lors de l'étape de RT. Au cours de cette *PCR*, le produit d'amplification est marqué par l'incorporation d'un nucléotide biotinylé (figure 6). Les amplicons complexes ainsi obtenus sont hybridés sur la puce.

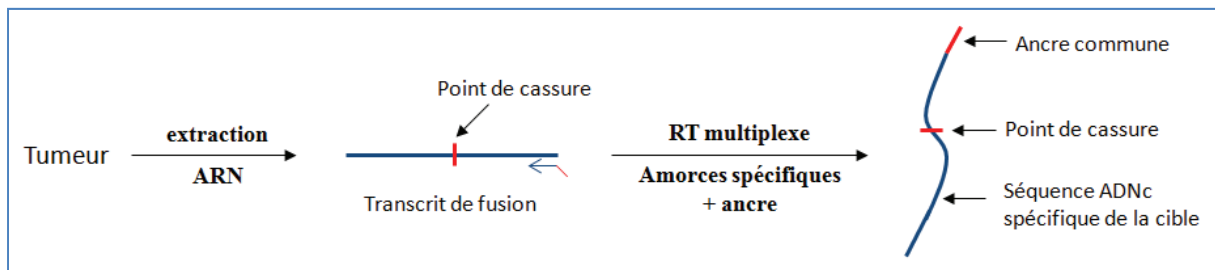


Figure 5 : principe de l'extraction et de la RT multiplexe à l'aide d'amorces possédant une ancre. L'ADNc produit contient la séquence de l'ancre suivie de celle du transcrit de fusion.

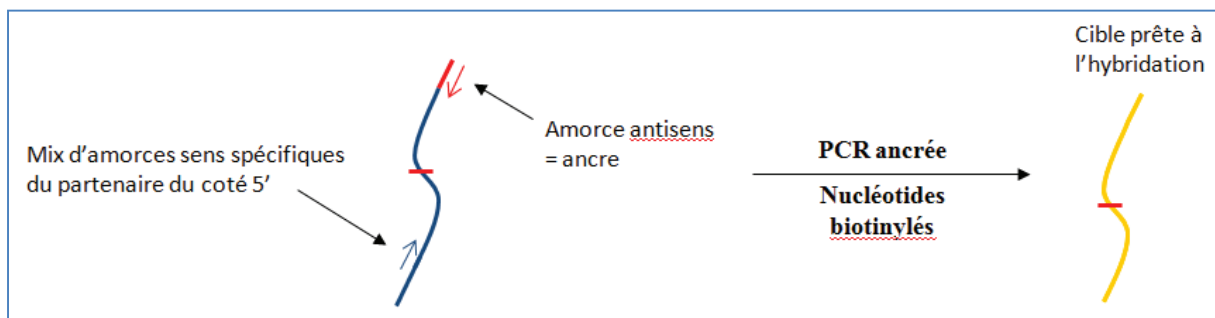


Figure 6 : principe de la PCR multiplexe ancrée permettant l'amplification et le marquage des cibles. L'amorce sens spécifique des partenaires situés du côté 5' de la translocation permet de n'amplifier que les amplicons des allèles transloqués. En effet, le mécanisme de translocation ne touchant qu'un seul allèle, cela permet de ne pas amplifier le produit de RT résultant de l'allèle normal.

La biopuce sera conçue de manière à permettre une lecture simple et directe du résultat. Une réaction colorimétrique montrera la présence ou non d'une translocation génique et, le cas échéant, indique quel est le gène partenaire impliqué et quelle est la position approximative du point de cassure sur ce dernier.

2. Analyse bioinformatique des séquences d'intérêt

2.1. Design des sondes oligonuléotides à déposer sur la puce

En se basant sur les séquences et les annotations entrées dans la base de données, nous avons procédé à une première étape globale d'analyse bioinformatique. Dans un premier temps, des sous-séquences de 400 nucléotides provenant de chacune des séquences d'intérêt ont été extraites (200 nucléotides en amont et 200 nucléotides en aval du point de cassure). C'est dans ces sous-séquences que les sondes oligonucléotides doivent être choisies. Pour les transcrits de fusion des trois sarcomes pédiatriques sur lesquels nous nous focaliserons sur la première puce prototype, cela représente vingt-quatre séquences uniques pour les partenaires du côté 3' de la translocation, et dix-huit séquences uniques pour les partenaires du côté 5' de la translocation. Chacune de ces séquences a été comparée par BLAST et alignements multiples contre toutes les autres séquences. Cette analyse a montré de fortes homologies entre certains partenaires différents : c'est en particulier le cas des partenaires 3' SSX1, SSX2 et SSX4 qui sont trois gènes phylogénétiquement très proches impliqués dans les synoviosarcomes, et c'est également le cas des partenaires 5' du rhabdomyosarcome, PAX3 et PAX7. Ces résultats sont à garder en mémoire, car ils pourront expliquer des phénomènes d'hybridations croisées non désirés sur la puce à ADN de diagnostic finale.

Après cette étape préliminaire, nous avons essayé de localiser les meilleures sondes oligonucléotides sur ces séquences. Pour chaque secteur concerné, correspondant généralement à un exon, une liste d'oligonucléotides candidats d'environ 50 mers est créée, par système de fenêtre coulissante. Un script permet alors de calculer les propriétés physiques de chaque sonde potentielle. Un fichier Excel renseigne la longueur exacte de ces oligonucléotides, leur position relative par rapport au point de cassure, le nombre de bases G et C, le pourcentage de GC, la présence de triplets C ou G et la température de fusion (T_m) de l'oligonucléotide (tableau 2). Cette dernière est calculée selon deux méthodes, la méthode des GC et la méthode thermodynamique (*Nearest Neighbour analysis*). Enfin, les oligonucléotides sont soumis à un BLAST contre les séquences humaines de référence du NCBI et les zones encadrant les points de cassure. Pour chacun des partenaires uniques, nous avons alors sélectionné une sonde finale entre 40 et 50 mers.

Oligo	Length	Pos	percentGC	nbG	nbC	GC_last5	Tm_NN	Tm_GC	hairpin	tripleC	tripleG
GTAGAAGTGTATCTCTTTC AGTTGAACGTCCTCAGAT GACTTTCGG	46	51	43,5	11	9	3	76,9	66,9	OK	0	0
GTAGAAGTGTATCTCTTTC AGTTGAACGTCCTCAGAT GACTTTCGGC	47	52	44,7	11	10	4	78,1	67,6	OK	0	0
GTAGAAGTGTATCTCTTTC AGTTGAACGTCCTCAGAT GACTTTCGGCAGG	50	55	46	13	10	4	80,3	68,8	OK	0	0
GAAGTGTATCTCTTTCAGT TGAACGTCCTCAGATGACT TTCGGCAGGC	48	56	47,9	12	11	4	81,8	69,1	OK	0	0

Tableau 2 : extrait du tableau de rendu des résultats du logiciel d'analyse montrant les critères thermodynamiques essentiels pour le choix des sondes.

L'ensemble des sondes choisies est vérifié une ultime fois en positionnant ces dernières sur les séquences des transcrits de fusion de notre base de donnée sous Vector NTI®. Nous pouvons ainsi être sûrs que la taille des amplicons (distance entre les deux sondes d'un transcrit de fusion) n'excède pas 150 pb, la taille moyenne des fragments d'ARN dégradés rencontrés dans les échantillons FFPE.

Il est à noter qu'aucune solution satisfaisante n'a été trouvée pour régler les problèmes d'homologie entre les transcrits de fusion impliquant les gènes SSX1, SSX2 ou SSX4. Il est fort probable que les trois sondes correspondantes se comportent de la même façon lors de l'hybridation et ne permettent pas de discriminer les sous-types. En ce qui concerne les transcrits de fusion impliquant les gènes PAX3 et PAX7, il serait possible de choisir des sondes en dehors des zones d'homologies, mais les amplicons auraient une taille de 200 pb.

Notre liste finale est composée de 52 sondes, dont 36 sondes de recherche de transcrits de fusion et 16 sondes contrôles positifs des étapes de RT et de PCR multiplexes. Ces sondes nous permettront de diagnostiquer les 4 sous-types de rhabdomyosarcomes, les 24 sous-types d'Ewing et les 7 sous-types de synoviosarcomes connus.

2.2. Design des amorces de RT et de PCR

Nous avons choisi d'élaborer une puce de diagnostic qui soit compatible avec le matériel fixé au formol et inclus en paraffine (FFPE). En effet, ce mode de conservation présente un avantage certain en termes de coûts de stockage puisque les échantillons sont alors stabilisés et peuvent être entreposés à température ambiante. Il permet également de faire circuler plus facilement et plus sûrement les échantillons entre les laboratoires et les différents services d'un ou plusieurs hôpitaux. En revanche, ces échantillons nécessitent un traitement beaucoup plus compliqué lorsqu'il s'agit de réaliser des études de biologie moléculaire, et les techniques classiques doivent très souvent être adaptées. En effet, les acides nucléiques (ADN et ARN) extraits sont très dégradés et les bases nucléiques subissent des modifications chimiques pouvant gêner les enzymes couramment utilisées en biologie moléculaire (figure 7).

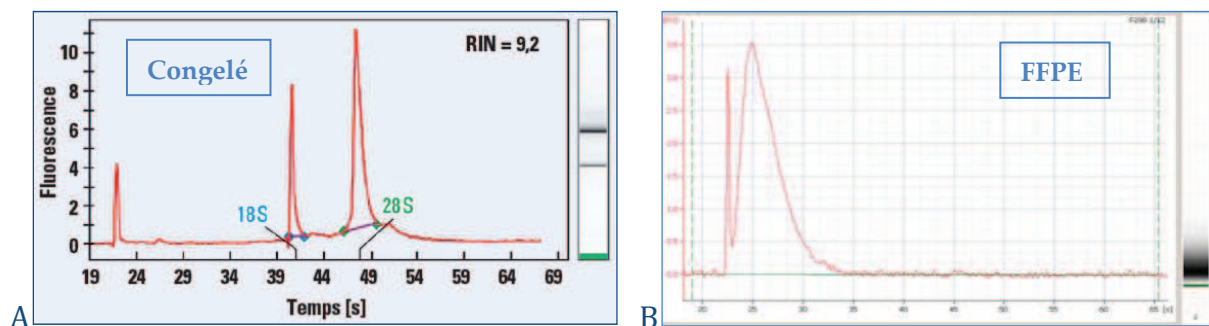


Figure 7 : exemples de profils de dégradation obtenus grâce au Bioanalyzer Agilent. Il s'agit d'électrophorèses capillaires des ARN extraits (A) à partir d'un tissu congelé et (B) à partir d'un tissu fixé au formol et inclus en paraffine (FFPE). Les ARN extraits du tissu FFPE sont dégradés, les fragments obtenus ont une taille moyenne d'une centaine de paires de base. Les pics caractéristiques des ARNs ribosomiques 18S et 28S disparaissent.

Mon laboratoire d'accueil fait partie des rares laboratoires spécialisés (il en existe seulement trois en France) qui réalisent des diagnostics moléculaires par PCRq sur les sarcomes inclus en paraffine. Les résultats obtenus sont excellents et le CLB est devenu un des centres de référence pour le traitement des échantillons FFPE. Le savoir-faire et l'expertise acquis dans ce domaine peuvent donc être mis à profit pour le design des amorces nécessaires aux étapes de RT et de PCR. Ainsi, par rapport à une PCR classique, la PCR à partir d'échantillons FFPE doit être robuste, et la taille d'un amplicon de PCR ne doit pas excéder 150 pb. Une stratégie de PCR de type ancrée a été retenue, pour sa capacité à réduire drastiquement la complexité du pool d'amorces, qui se doivent d'être multiplexées en correspondance à l'ensemble des transcrits de fusion recherchés¹⁸. Cela a pour effet direct de limiter au maximum les produits d'amplification non spécifiques gênants qui se produisent généralement au cours d'une PCR multiplexe. Ainsi, lors de notre processus d'amplification et de marquage des sondes, seules les amorces sens vont varier, les amorces antisens correspondant à l'ancre choisie (séquence artificielle absente du génome humain). Cette ancre aura été introduite dans l'ADNc servant de matrice à la réaction de PCR grâce à une amorce de RT chimérique, contenant une partie de la séquence cible du côté 3' et l'ancre artificielle du côté 5'.

Concernant le design des amorces de RT, la séquence de l'ancre nous a été fournie par la société Ipsogen. Pour assurer la spécificité de l'amorce de RT à sa cible, une séquence de dix nucléotides correspondant à l'extrémité de la sonde située sur le partenaire du côté 3' a été ajoutée à l'ancre. Différents logiciels de design d'amorces PCR ont été évalués. Le logiciel libre disponible en ligne Primer3 a finalement été retenu pour son adaptabilité face aux séquences difficiles, sur lesquelles nous avions peu de latitude de mouvement en raison des contraintes de taille des tissus FFPE¹⁹. Les résultats présentés à partir d'ici concerneront le rhabdomyosarcome, car il s'agit du sarcome de plus faible degré de complexité (seulement quatre sous-types connus) par rapport aux autres sarcomes qui seront présents sur la puce prototype (figure 8). Les résultats de design des deux autres sarcomes pédiatriques sont donnés en annexe (annexe 3).

Primer F	Sonde côté 5'	Breakpoint	Sonde côté 3'	Amorce RT	AMORCE ANTISENS = ANCRE
>PAX3(E7)-FKHR(E2)ANCRE2	Tgcctcgtcggggccctccaaccccatgaacccaccattggcaatggcctctcactcag	↑	aattcaattcgtcataatctgtcctcacagcaagttcattcgtgtgcagaatgaaggactgc	CCAGTACCGTGTAGTCTCTCTCGTG	
>PAX7(E7)-FKHR(E2)ANCRE2	Cttcatgaatccggcgccctccaacacatgaacccggctagcaacggcctgtctctcag	↑	aattcaattcgtcataatctgtcctcacagcaagttcattcgtgtgcagaatgaaggactgc	CCAGTACCGTGTAGTCTCTCTCGTG	
>PAX3(E7)-MLLT7(E2)ANCRE2	Tgcctcgtcggggccctccaaccccatgaacccaccattggcaatggcctctcactcag	↑	aactcgtatccgccaacactgctcctgcacagcaagtcatcaagggttcacaacgagccaccggcaaaagctcttgatgata	CCAGTACCGTGTAGTCTCTCTCGTG	
>PAX3(E6)-NCOA1(E13)ANCRE2	Gccatgccagccttaccacgtacacgtctcggagacctctaccagccacatctattccacaag	↑	gattacctgagctgggaattggaagcaattgataaccaattggacaaccaggaacagggc	CCAGTACCGTGTAGTCTCTCTCGTG	
>FKHR(E1)-FKHR(E2)	caagagcgtgccctatctcaaggataaggtgacagcaacagctcggcggtggaagaattcaattcgtcataatctgtcctcacagcaagttcattcgtgtgcagaatgaaggactgc	↑	CCAGTACCGTGTAGTCTCTCTCGTG		
>MLLT7(E1)-MLLT7(E2)	atgtctcgtactgtaccctcttcaaggacaagggtgacagcaacagctcagcagatggaagaactcgtccgccaacactgctcctgcacagcaagttcatcaagggttcacaacgagccaccggcaaaagctcttgatgata	↑	CCAGTACCGTGTAGTCTCTCTCGTG		
>NCOA1(E12)-NCOA1(E13)	catcaaatcggagatctaccagcttactcagtcgccactgccagaccacttcaggctaaatagattactcgaactggaattggaagcaattgataaccaattggacaaccaggaacagggc	↑	CCAGTACCGTGTAGTCTCTCTCGTG		

Figure 8 : résultats du design d'amorces de RT et de PCR obtenus pour le rhabdomyosarcome. Les quatre premières séquences correspondent aux quatre transcrits de fusion connus dans les rhabdomyosarcomes : PAX3-FKHR, PAX7-FKHR, PAX3 MLLT7 et PAX3-NCOA1. Les trois dernières séquences sont les gènes normaux des partenaires du côté 3' des transcrits de fusion servant de contrôles positifs.

Une fois la liste des différentes amorces de RT et de PCR obtenue au format texte tabulé, nous avons commandé ces oligonucléotides chez la société Eurogentec (tableau 3).

Nom des amorces	Séquence	Num	Taille
Primer R ANCRE2	CACGAGAGAGACTACACGGTACTGG	1	25
Primer F PAX3E7	ccctccaaccccatgaac	2	18
Primer F PAX7E7	gccctccaaccacatgaac	3	19
Primer F PAX3E6	tgccaacgtaccagctgtc	4	19
Primer F MLLT7E1	cctacttcaaggacaagggtga	5	22
Primer F FKHRE1	cttcaaggataagggtgacagc	6	22
Primer F NCOA1E12	cctgccagcttcacttcag	7	19
Primer RT FKHR+ANCRE2	CACGAGAGAGACTACACGGTACTGGccagttcctt	8	35
Primer RT MLLT7+ANCRE2	CACGAGAGAGACTACACGGTACTGGcatccaccaa	9	35
Primer RT NCOA1+ANCRE2	CACGAGAGAGACTACACGGTACTGGcgctgttcc	10	35

Tableau 3 : exemple de la liste d'amorces de RT et de PCR commandés chez Eurogentec pour le diagnostic du rhabdomyosarcome. Cette liste contient 7 amorces de PCR (6 amorces sens et 1 amorce antisens correspondant à l'ancre) et 3 amorces de RT (car il existe 3 partenaires différents du côté 3' des transcrits de fusion).

Ces amorces ont tout d'abord été testées et optimisées sur des culots congelés d'une lignée cellulaire de rhabdomyosarcome, caractérisée par RT-PCRq classique, et possédant une translocation de type PAX3-FKHR. Après avoir réalisé une RT multiplexe (mix de RT contenant les 3 amorces), nous avons réalisé les PCR des 6 couples d'amorces possibles ainsi qu'une PCR multiplexe contenant toutes les amorces à la fois (figure 9 et annexe 4).

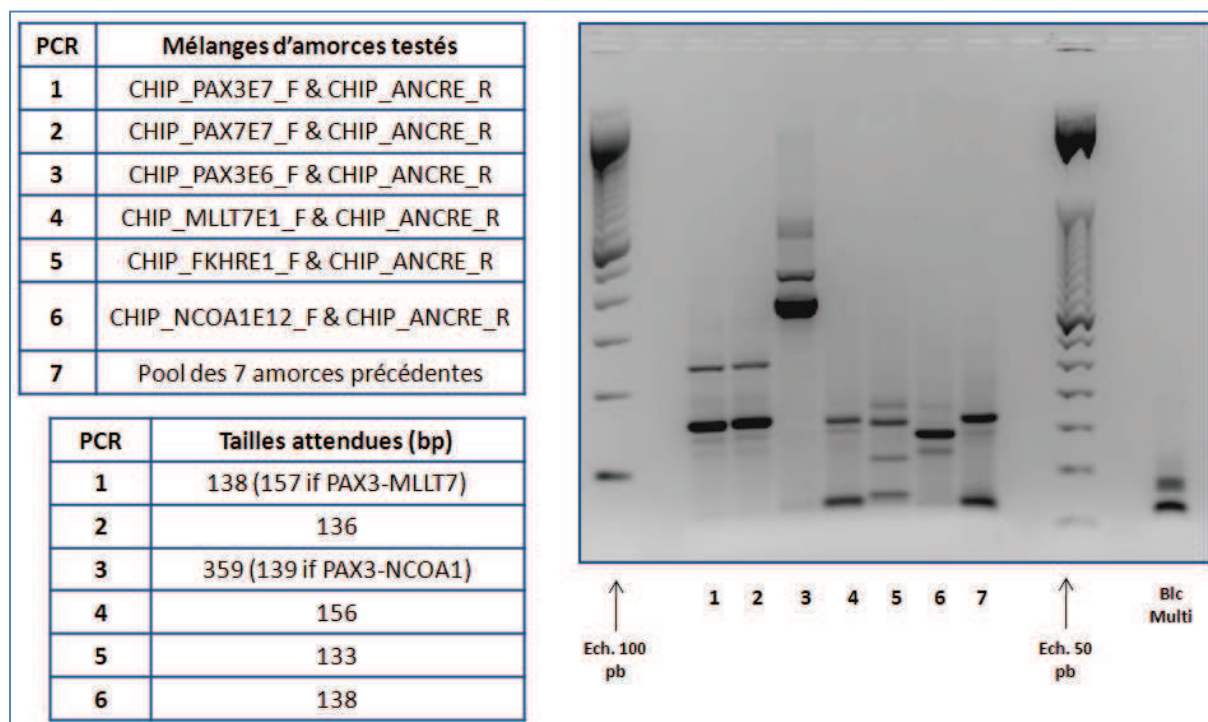


Figure 9 : résultats de migration sur gel des produits de la RT-PCR multiplexe ancree réalisée sur de l'ARN extrait à partir d'une lignée de rhabdomyosarcome congelée. Sachant que cette lignée est caractérisée par une translocation de type PAX3-FKHR, il est possible de déduire les tailles d'amplicons attendues. Dans chacun des puits d'électrophorèse, nous obtenons des bandes aux tailles attendues. Il existe cependant des bandes non spécifiques, en particulier au niveau du front de migration, qui correspondent probablement à des dimères d'amorces, puisqu'on les retrouve dans le puits contrôle sans matrice ADNc.

Les résultats de ce premier test sont très encourageants. En effet, pour chacun des six couples d'amorce testés, nous retrouvons bien des bandes aux tailles attendues, aux environs de 150 pb (couples n° 1, 2, 4, 5 et 6) et de 360 pb (couple n°3). Il existe toutefois des bandes aspécifiques dans de nombreux puits d'électrophorèse, mais on peut supposer que ces produits non spécifiques n'auront pas d'influence au moment de l'hybridation de l'ensemble de ces cibles sur la puce. De même, le résultat de la PCR contenant l'ensemble des sept amorces multiplexées est positif puisqu'on ne retrouve pas les bandes parasites des autres puits (hormis une bande entre 50 et 100 pb correspondant très probablement à des dimères d'amorce), comme si l'amplicon d'intérêt prenait le dessus sur les réactions parasites.

Nous avons ensuite essayé d'optimiser la réaction de PCR multiplexe en essayant de trouver la température de fusion idéale, c'est-à-dire celle favorisant le plus l'amplicon d'intérêt par rapport aux produits de PCR non spécifiques. Pour cela, nous avons réalisé, pour chacun des six couples, une PCR en gradient de température, avec des T_m variants de 54 à 65°C (figure 10). La température idéale a pu être estimée à environ 58°C.

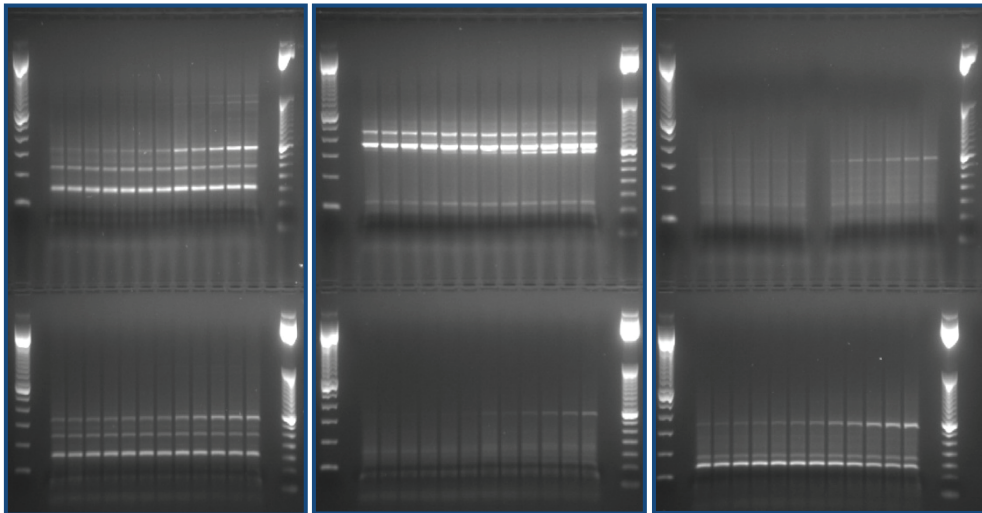


Figure 10 : résultats de la migration sur gel des PCR réalisées en gradient de température pour chacun des six couples d'amorce de la PCR multiplexe du rhabdomyosarcome. Une température optimale d'environ 58°C permettra de limiter au maximum l'apparition de produits aspécifiques.

Nous utiliserons désormais cette température de 58°C pour l'ensemble de nos autres tests. Après avoir testé notre design d'amorces sur un culot cellulaire congelé, nous avons souhaité l'appliquer à du matériel fixé au formol et inclus en paraffine. Pour cela, des culots cellulaires congelés ont été envoyés au service d'Anatomie et Cytopathologie du Centre Léon Bérard, pour procéder à leur fixation dans du formol puis à leur enrobage dans de la paraffine. L'ARN de ces cellules fixées a ensuite été extrait selon les protocoles classiques d'extraction utilisés pour le diagnostic de tumeurs en routine. Un test de RT-PCR multiplexe a été réitéré, par couples d'amorces et avec l'ensemble des amorces PCR (figure 11).

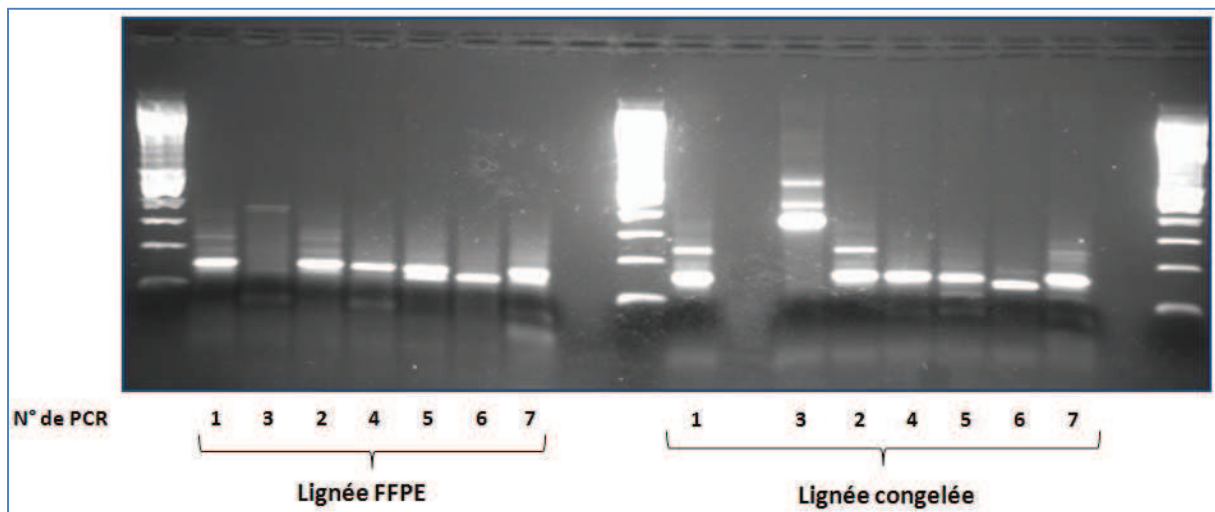


Figure 11 : résultats de la migration sur gel des RT-PCR multiplexes réalisées à un T_m de 58°C , à partir d'ARN extrait de culots cellulaires de la même lignée de rhabdomyosarcome, congelée ou fixée au formol puis incluse en paraffine. Il est à noter que le nouveau T_m permet de limiter l'apparition de bandes non spécifiques. De plus, les PCR réalisées à partir de matériel FFPE sont beaucoup plus "propres" que celles réalisées à partir de matériel congelé. Enfin, la bande à environ 360 pb de la PCR n°3 disparaît pour l'échantillon inclus en paraffine, ce qui est cohérent avec le fait que le matériel génétique d'un tissu FFPE est très dégradé (les tailles des acides nucléiques oscillent autour de 150 pb).

Le design et l'optimisation de la RT-PCR multiplexe du rhabdomyosarcome est une réussite, en particulier sur le matériel FFPE. Le travail effectué a été renouvelé pour le sarcome d'Ewing et le synovialosarcome. Par la suite, il suffira de rajouter des nucléotides biotinylés à la PCR multiplexe pour obtenir des cibles marquées.

3. Réalisation de la première puce prototype

3.1. Design de la puce à ADN prototype

Après obtention des premiers résultats positifs en ce qui concerne la partie de préparation des cibles (RT-PCR multiplexe de marquage des cibles), nous avons initié la production d'une première puce prototype. L'ensemble des sondes choisies précédemment a été synthétisé par la société Invitrogen. Ces sondes ont ensuite été déposées sur une membrane de nylon chez notre collaborateur Ipsogen, grâce à une séquence d'accroche qui permet d'augmenter l'accessibilité des sondes pour les cibles (figure 12).

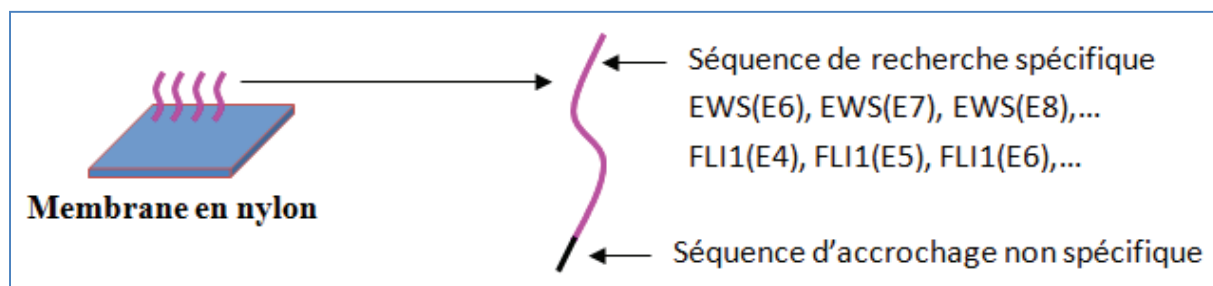


Figure 12 : principe d'accrochage des sondes sur la membrane de nylon.

Un schéma de positionnement des sondes a été décidé, et les sondes représentant les partenaires du côté 5' des translocations (les plus importantes) sont déposées en duplicats (figure 13).

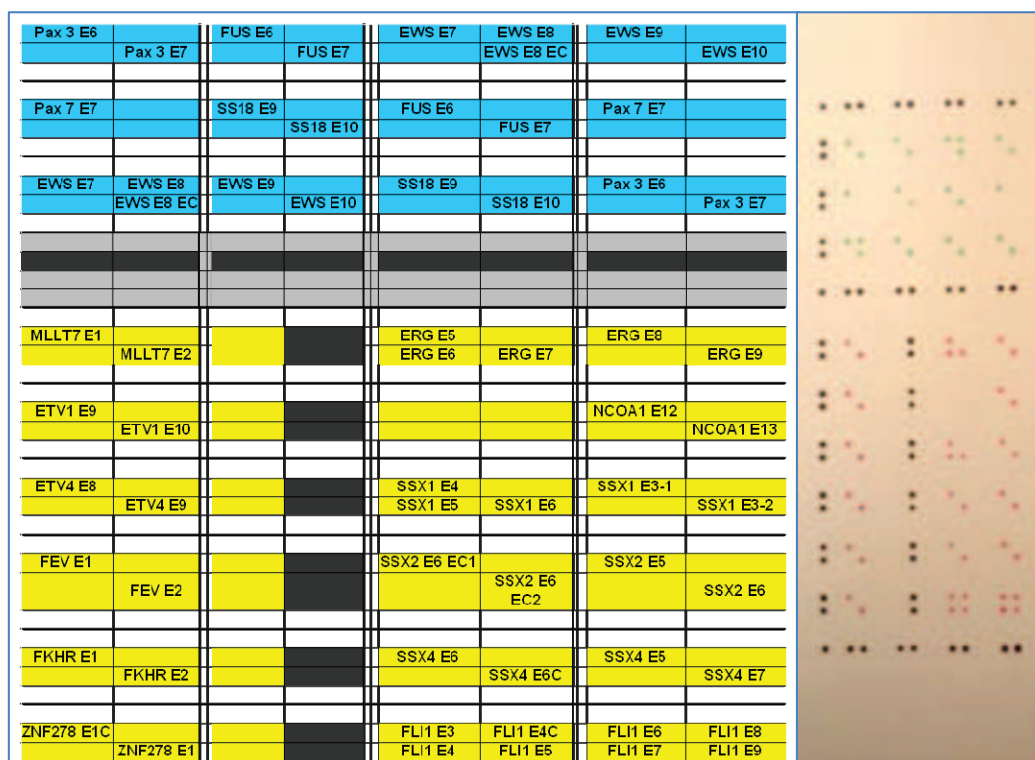


Figure 13 : schéma de positionnement des sondes sur la membrane (à gauche) et photo de la première puce prototype (à droite).

3.1.1. Premiers tests de la puce à ADN prototype

Une première expérience d'hybridation a été menée sur les premiers échantillons de la puce prototype, en marquant les cibles lors de l'étape de PCR ancrée multiplexe à l'aide de nucléotides marqués radioactivement au ^{32}P . Ce marquage radioactif est juste utilisé en test pour sa grande sensibilité, mais ne sera bien sûr pas retenu comme principe de marquage final, en raison des contraintes qu'il engendre et qu'on ne peut pas transposer dans un laboratoire de diagnostic en routine.

Pour ce premier test, nous avons encore une fois utilisé notre lignée de rhabdomyosarcome possédant la translocation PAX3-FKHR. Les produits de PCR radioactifs ont été hybridés sur la nuit sur la puce à ADN, puis révélés pendant 24h (figure 14).



Figure 14 : premier résultat d'hybridation sur puce à ADN prototype des produits d'amplification de la lignée de rhabdomyosarcome de type PAX3-FKHR. Les spots attendus comme positifs sont entourés sur le plan de la puce (à gauche). À droite, on peut voir le résultat de l'étape de révélation, sur laquelle on a superposé un masque permettant le repérage des sondes.

Les sondes contrôles MLLT7(E1), MLLT7(E2), FKHR(E1) et FKHR(E2) sortent clairement positives. Les sondes contrôles NCOA1 sortant très mal, elles seront donc abandonnées lors du design final de la puce, et les amorces correspondantes seront retirées du pool multiplexe utilisé pour la préparation des cibles. Les seules sondes s'allumant en duplicat et correspondant au partenaire du côté 5' de la translocation sont les sondes PAX3(E7). La puce à ADN nous montre donc la présence d'une translocation de type PAX3(E7)-gène3', le gène3' correspondant plutôt au gène FKHR qu'au gène MLLT7 au vu de l'intensité des signaux obtenus.

CONCLUSION ET PERSPECTIVES

Les travaux réalisés jusqu'à présent ont clairement prouvé la faisabilité de la détection sur puce à ADN de transcrits de fusion amplifiés et marqués par RT-PCR ancrée multiplexe. Ainsi, en une seule expérience, il est possible de détecter l'ensemble des translocations répertoriées dans les sarcomes complexes.

Les travaux de design d'amorces ont été achevés sur les sarcomes pédiatriques (rhabdomyosarcome, synoviosarcome et sarcome d'Ewing). Le test du rhabdomyosarcome a été couronné de succès. Il reste à tester les design d'amorces des deux sarcomes pédiatriques restants, à l'aide de lignées caractérisées si possible, ou à l'aide des transcrits de fusion clonés dans des vecteurs bactériens. Une fois les différents tests effectués, la puce prototype pourra être améliorée, et sa complexité sera réduite, en ne gardant que les meilleurs contrôles positifs pour chacun des types de sarcomes.

Ensuite, une fois la puce à ADN pédiatrique éprouvée, ce principe de détection sera élargi à l'ensemble des sarcomes à translocations répertoriés dans la base Vector NTI® qui est fréquemment maintenue à jour grâce à l'utilisation de flux RSS.

Enfin, nous pourrions lancer une étude plus large basée sur de vrais échantillons tumoraux. Ces tests seront simultanément menés sur puce et via notre approche diagnostique traditionnelle par RT-PCRq, pour comparer exhaustivement ces deux techniques en termes de sensibilité et de spécificité. Après validation des résultats obtenus et en cas de réussite, nous envisageons de transférer cette technologie en routine, en tant que nouveau test diagnostique moléculaire de référence au Centre Léon Bérard.

ANNEXES

1. Liste des sous-types pédiatriques diagnostiqués au Centre Léon Bérard par RT-PCRq

- > Sarcome d'Ewing /pPNET: 8 sous-types
 - ES/pPNET - Type I EWS (E7) - FLI1 (E6)
 - ES/pPNET - Type II EWS (E7) - FLI1 (E5)
 - ES/pPNET - EWS (E7) - FLI1 (E7)
 - ES/pPNET - EWS (E7) - FLI1 (E8)
 - ES/pPNET - EWS (E7) - FLI1 (E9)
 - ES/pPNET - EWS (E10) - FLI1 (E5)
 - ES/pPNET - EWS (E7) - ERG (E6)
 - ES/pPNET - EWS (E7) - ERG (E8)
- > Synovialosarcome: 3 sous-types
 - SS - SS18 (E10) - SSX1 (E6)
 - SS - SS18 (E10) - SSX2 (E6)
 - SS - SS18 (E10) - SSX4 (E6)
- > Rhabdomyosarcome : 2 sous-types
 - ARMS - PAX3 (E7) - FKHR (E2)
 - ARMS - PAX7 (E7) - FKHR (E2)

2. Liste des sous-types présents sur la puce à ADN prototype de diagnostic des sarcomes pédiatriques

- > Sarcome d'Ewing /pPNET: 24 sous-types
 - ES/pPNET - Type I EWS (E7) - FLI1 (E6)
 - ES/pPNET - Type II EWS (E7) - FLI1 (E5)
 - ES/pPNET - EWS (E7) - FLI1 (E4)
 - ES/pPNET - EWS (E7) - FLI1 (E7)
 - ES/pPNET - EWS (E7) - FLI1 (E8)
 - ES/pPNET - EWS (E7) - FLI1 (E9)
 - ES/pPNET - EWS (E7) - ERG (E6)
 - ES/pPNET - EWS (E7) - ERG (E7)
 - ES/pPNET - EWS (E7) - ERG (E8)
 - ES/pPNET - EWS (E7) - ERG (E9)
 - ES/pPNET - EWS (E7) - ETV1 (E10)
 - ES/pPNET - EWS (E7) - ETV4 (E9)
 - ES/pPNET - EWS (E7) - FEV (E2)
 - ES/pPNET - EWS (E8) - FLI1 (E7)
 - ES/pPNET - EWS (E8) - ZNF278 (E1)
 - ES/pPNET - EWS (E9) - FLI1 (E4)
 - ES/pPNET - EWS (E9) - FLI1 (E7)
 - ES/pPNET - EWS (E10) - FLI1 (E5)
 - ES/pPNET - EWS (E10) - FLI1 (E6)
 - ES/pPNET - EWS (E10) - FLI1 (E8)
 - ES/pPNET - EWS (E10) - ERG (E6)
 - ES/pPNET - EWS (E10) - FEV (E2)
 - ES/pPNET - FUS (E6) - ERG (E9)
 - ES/pPNET - FUS (E7) - ERG (E8)
- > Synovialosarcome: 7 sous-types
 - SS - SS18 (E9) - SSX1 (E4) variant
 - SS - SS18 (E10) - SSX1 (E6)
 - SS - SS18 (E10) - SSX2 (E6)
 - SS - SS18 (E10) - SSX2 (E6) variant
 - SS - SS18 (E10) - SSX2 (E6) variant 2
 - SS - SS18 (E10) - SSX4 (E6)
 - SS - SS18 (E10) - SSX4 (E7) variant
- > Rhabdomyosarcome : 4 sous-types
 - ARMS - PAX3 (E7) - FKHR (E2)
 - ARMS - PAX7 (E7) - FKHR (E2)
 - ARMS - PAX3 (E6) - NCOA1 (E13)
 - ARMS - PAX3 (E7) - MLLT7 (E2)

3. Liste des amorces de RT et de PCR dessinées pour le synoviosarcome et le sarcome d'Ewing

Amorces synoviosarcome	Séquences	Num	Taille
Primer R ANCRE2	CACGAGAGAGACTACACGGTACTGG	1	25
Primer F SS18E10	agcctggaccaccacagc	2	18
Primer F SS18E9	tatccaccccagcagcag	3	18
Primer F SSX1E3	ctcggagaaaatcagctatgtg	4	22
Primer F SSX24E5	aggttgaacgtcctcagatg	5	20
Primer RT SSX1E6+ANCRE2	CACGAGAGAGACTACACGGTACTGGGTGGGCCAGA	6	35
Primer RT SSX1E4+ANCRE2	CACGAGAGAGACTACACGGTACTGGCTGGAAGTCT	7	35
Primer RT SSX1E5+ANCRE2	CACGAGAGAGACTACACGGTACTGGCTTCGGGATG	8	35
Primer RT SSX24E6+ANCRE2	CACGAGAGAGACTACACGGTACTGGGGCACTTCCT	9	35
Primer RT SSX2E6V2+ANCRE2	CACGAGAGAGACTACACGGTACTGGGAAGTGAAGC	10	35
Primer RT SSX2E6V3+ANCRE2	CACGAGAGAGACTACACGGTACTGGCCTGCCGAAA	11	35
Primer RT SSX4E7+ANCRE2	CACGAGAGAGACTACACGGTACTGGCAGCTGCTTT	12	35

Amorces Ewing	Séquences	Num	Taille
Primer R ANCRE2	CACGAGAGAGACTACACGGTACTGG	1	25
CHIP_EWSE7_F	cagccaagctccaagtcaat	2	20
CHIP_EWSE8_F	gaggcatgagcagaggtg	3	18
CHIP_EWSE8EC_F	gttcctggagcttaggatgtgt	4	22
CHIP_EWSE9_F	aggaggacgcggtggaat	5	18
CHIP_EWSE10_F	ggacccatggatgaaggac	6	19
CHIP_FUSE6_F	gcagtgggtggctatgaacc	7	19
CHIP_FUSE7_F	gaggcagaggtggcatgg	8	18
CHIP_ERGE5_F	gcttttattttccaaatacttca	9	24
CHIP_ETV1E9_F	ggccccaggcagttttatg	10	19
CHIP_ETV4E8_F	gggtggtgatcaaacaggaac	11	21
CHIP_FLI1E3_F	caacatgaccaccaacgagag	12	21
CHIP_ZNFE1C_F	gtggctggacccctatcc	13	18
CHIP_ERGE6_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGCCGTGACCG	14	35
CHIP_ERGE7_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGGACGCTGGT	15	35
CHIP_ERGE8_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGGATTTGCAA	16	35
CHIP_ERGE9_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGCTGTCCGAC	17	35
CHIP_FEVE2_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGCCCCAGCTC	18	35
CHIP_FLI1E4_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGGCCCACTCC	19	35
CHIP_FLI1E5_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGTGAGGATTG	20	35
CHIP_FLI1E6_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGAATTCATGT	21	35
CHIP_FLI1E7_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGCCGTTGCTCT	22	35
CHIP_FLI1E8_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGGGTTGGCTAG	23	35
CHIP_FLI1E9_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGCTGTCCGAGA	24	35
CHIP_ETV1E10_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGCGTTGGTATG	25	35
CHIP_ETV4E9_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGCAGAGAAGCC	26	35
CHIP_ZNFE1_RT+ANCRE2	CACGAGAGAGACTACACGGTACTGGCCTTGGGCGG	27	35

4. Protocoles utilisés pour les RT et PCR multiplexes**-> RT multiplexe**

Volume final de RT : 40 µl

Quantité ARN : 3 µg

Mix	Volume (µl)
ARN (C = 1,461 ng/µl)	2,1
Oligos RT multiplexés (10 µM chacun)	2
dNTP (10 mM)	2
H2O qsp 24 µl	17,9
5 min à 65°C	
Tampon 5X	8
dTT	4
RNasine	2
2 min à 42°C	
SuperScript II	2
50 min à 42°C	
15 min à 70°C	

Oligos RT multiplexés

CHIP_A+FKHR_RT
 CHIP_A+MLLT7_RT
 CHIP_A+NCOA1_RT

Préparation du pool Multi RT Rhabdo

> Solutions mères à 100 µM
 > Solutions filles à 30 µM
 60 µl sol. mère + 140 µl H2O
 > Solution multi RT Rhabdo
 100 µl de chacune des solutions filles
 soit 300 µl à 10 µM chacuns

-> PCR ancré multiplexe

Volume final de PCR : 50 µl

Volume de RT : 5 µl

[primers] / réaction : 0,2 µM

Mix	Volume pour 1 réaction (µl)
ADNc	5
Tampon 10x	5
dNTP	2
MgCl2	1,5
Primer Mix (10 µM chacun)	1
Taq Polymérase	0,5
H2O qsp 50 µl	35

Programme thermocycleur PCR

Tubes = 0,2 ml

Vf = 50 µl

Température (°C)	Temps	Répétitions
94	2 min	1
94	1 min	40
60	1 min	
72	45 sec	
72	7 min	1
16	inf	1

BIBLIOGRAPHIE

1. Ménégoz F, Black RJ, Arveux P, Magne V, Ferlay J, Buémi A, Carli PM, Chapelain G, Faivre J, Gignoux M, Grosclaude P, Mace-Lesec'h J, Raverdy N, Schaffer P. Cancer incidence and mortality in France in 1975-95. *European Journal of Cancer Prevention* 1997; 6: 442-466.
2. Enzinger FM, Weiss SW. Soft Tissue Tumors. 3 ed. St Louis: Mosby-Year Book 1995.
3. Coindre JM, Ranchère D, Collin F. How to classify a soft tissue sarcoma in 2006. *Annales de Pathologie* 2006; 1: 98-109.
4. Wiklund T, Huuhtanen R, Blomqvist C, Tukiainen E, Virolainen M, Virkkunen P, Asko-Seljavaara S, Björkenheim JM, Elomaa I. The importance of a multidisciplinary group in the treatment of soft tissue sarcomas. *European Journal of Cancer* 1996; 32: 269-273.
5. Gustafson P, Dreinhofer KE, Rydholm A. Soft tissue sarcoma should be treated at a tumor center. A comparison of quality of surgery in 375 patients. *Acta orthopaedica Scandinavica* 1994; 65: 47-50.
6. Rydholm A. Improving the management of soft tissue sarcoma. Diagnosis and treatment should be given in specialist centres. *British Medical Journal* 1998; 317: 93-94.
7. Ouvrage collectif Fédération Nationale des Centres de Lutte contre le Cancer. Sarcomes des tissus mous et ostéosarcomes. Vol 1. Des "standards, options et recommandations". Paris: Arnette-Blackwell 1995.
8. Ray-Coquard I, Thiesse P, Ranchere-Vince D, Chauvin F, Bobin JY, Sunyach MP, Carret JP, Mongodin B, Marec-Berard P, Philip T, Blay JY. Conformity to clinical practice guidelines, multidisciplinary management and outcome of treatment for soft tissue sarcomas. *Annals of Oncology* 2004; 15: 307-315.
9. Mankin HJ, Mankin CJ, Simon MA. The hazards of the biopsy, revisited. Members of the Musculoskeletal Tumor Society. *The Journal of Bone and Joint Surgery. American volume* 1996; 78: 656-663.
10. Mankin HJ, Lange TA, Spanier SS. The hazards of biopsy in patients with malignant primary bone and soft-tissue tumors. *The Journal of Bone and Joint Surgery. American volume* 1982; 64: 1121-1127.
11. Peter M, Gilbert E, Delattre O. A multiplex real-time pcr assay for the detection of gene fusions observed in solid tumors. *Laboratory Investigation* 2001; 81: 905-912.
12. Guillou L, Coindre J, Gallagher G, Terrier P, Gebhard S, de Saint Aubain Somerhausen N, Michels J, Jundt G, Vince DR, Collin F, Trassard M, Le Doussal V, Benhattar J. Detection of the synovial sarcoma translocation t(X;18) (SYT;SSX) in paraffin-embedded tissues using reverse transcriptase-polymerase chain reaction: a reliable and powerful diagnostic tool for pathologists. A molecular analysis of 221 mesenchymal tumors fixed in different fixatives. *Human Pathology* 2001; 32: 105-112.
13. Coindre JM, Hostein I, Terrier P, Bouvier-Labit C, Collin F, Michels JJ, Trassard M, Marques B, Ranchere D, Guillou L. Diagnosis of clear cell sarcoma by real-time reverse transcriptase- polymerase chain reaction analysis of paraffin embedded tissues:

clinicopathologic and molecular analysis of 44 patients from the French sarcoma group. *Cancer* 2006; 107: 1055-1064.

14. Foss RD, Guha-Thakurta N, Conran RM, Gutman P. Effects of fixative and fixation time on the extraction and polymerase chain reaction amplification of RNA from paraffin-embedded tissue. Comparison of two housekeeping gene mRNA controls. *Diagnostic Molecular Pathology* 1994; 3: 148-155.

15. Fletcher JA. Cytogenetics and molecular biology of soft tissue tumors. *Monographs in Pathology* 1996; 38: 37-64.

16. Giusiano S, Formisano-Tréziny C, Benziane A, Maroc N, Picard C, Hermitte F, Taranger-Charpin C, Gabert J. Development of a biochip-based assay integrated in a global strategy for identification of fusion transcripts in acute myeloid leukemia: a work flow for acute myeloid leukemia diagnosis. *International Journal of Laboratory Hematology* 2009, sous presse.

17. Nasedkina T, Domer P, Zharinov V, Hoberg J, Lysov Y, Mirzabekov A. Identification of chromosomal translocations in leukemias by hybridization with oligonucleotide microarrays. *Haematologica* 2002; 87: 363-372.

18. Schutz T, Kairat A, Schroöder CH. Anchored oligo(dT) primed RT:PCR: identification and quantification of related transcripts with distinct 3'-ends. *Journal of Virological Methods* 2000; 86: 167-171.

19. Rutherford P, Churcher C, McCallum J. An interactive visualisation for selecting PCR primers. *ACM International Conference Proceeding Series* 2005; 109.

RESUME

En cancérologie, les sarcomes sont considérés comme de tumeurs rares et difficiles à prendre en charge. Malgré ces difficultés, de nombreuses anomalies chromosomiques sont décrites et permettent d'établir un diagnostic fiable à l'aide de techniques de biologie moléculaire éprouvées (FISH, PCRq...). Cependant, en raison de la grande complexité des sarcomes, ces techniques classiques sont laborieuses et nécessitent une grande quantité de matériel biologique de départ. Au cours de mes travaux de thèse, je me suis attaché à développer une méthode innovante de diagnostic moléculaire basée sur la technologie des puces à ADN. Cette puce de diagnostic, révolutionnaire par sa facilité d'utilisation et son exhaustivité, permet de rechercher en une seule expérience la totalité des sarcomes caractérisés par des translocations chromosomiques. Une puce prototype a d'ores et déjà été testée avec succès. A terme, nous espérons que cette puce à ADN sera adoptée comme technique de diagnostic moléculaire de routine au sein du Centre Léon Bérard, puis dans les principaux centres anticancéreux européens partenaires du Réseau d'Excellence Européen Conticanet. Elle permettra alors une meilleure prise en charge des patients atteints de sarcomes, et ce dès le diagnostic initial.

2. Classification des rhabdomyosarcomes grâce aux profils microARN

Fort de notre expérience réussie dans les ostéosarcomes, nous avons ensuite voulu savoir s'il était possible de mettre à profit les profils d'expression des microARN pour la résolution d'une autre problématique, celle de la classification fine des RMS. En particulier, notre but était d'apporter une aide pour le démembrement des RMS de la zone grise. En effet, ces entités possèdent des morphologies très indifférenciées, et il est parfois très difficile pour un médecin anatomopathologiste aguerri de distinguer un ARMS d'un ERMS. Le traitement des patients dépendant très fortement de ce diagnostic initial, il est important de pouvoir mieux cerner ces entités.

Alors que nous étions en train de démarrer notre étude, l'équipe du Dr Olivier Delattre a publié une étude analysant les profils d'expressions des ARNm, ainsi que les anomalies génomiques d'une cohorte de 210 RMS (Williamson et al., 2010). Ces RMS ont initialement été classés en trois groupes distincts : les pARMS, les nARMS et les ERMS, sur la base d'une relecture histopathologique et d'une recherche systématique des transcrits de fusion caractéristiques des ARMS par technique PCRq. En plus des transcrits de fusion courants décrits chez les ARMS (PAX3-FOXO1 et PAX7-FOXO1), le transcrit rare PAX3-NCOA1 a également été recherché de façon systématique. Cette étude a montré que globalement, les nARMS partageaient des caractéristiques moléculaires identiques aux ERMS. Plus que l'histologie, le facteur clé pour différencier les ARMS de ERMS en termes de biologie et de progression clinique serait donc la présence ou l'absence de gènes de fusion. Au cours de cette étude, sur la base des profils d'expression des ARNm, seul le cas de pARMS à transcrit rare de type PAX3-NCOA1 se comportait différemment des autres pARMS à transcrits courants.

Ces résultats sont en parfaite adéquation avec les résultats préliminaires que nous avons publiés dans la revue *Sarcoma* (cf. article III) (Gougelet et al., 2011a), puisque sur la base des profils d'expression des microARN, nous avons montré qu'il existait des signatures d'expression des microARN permettant de discriminer à la fois les pARMS des ERMS et des entités de la zone grise, mais également, au sein des pARMS, les entités possédant des transcrits de fusion type PAX3-FOXO1 de celles possédant des transcrits de fusion de type pAX7-FOXO1. Les profils d'expression des microARN sont donc potentiellement utilisables en tant qu'outil de diagnostic à forte valeur ajoutée.

L'article suivant publié dans *Sarcoma* présente également la poursuite de nos travaux entrepris dans les ostéosarcomes, à savoir l'utilisation des microARN pour répondre aux problématiques liées au traitement des sarcomes. En utilisant des lignées cellulaires humaines d'ostéosarcome, de chondrosarcome et de lymphome de Burkitt, nous montrons tout d'abord que les effets des différents agents chimiothérapeutiques sur les profils d'expression des microARN sont cellule-spécifiques et médicament-spécifiques. Nous montrons ensuite que pour les ostéosarcomes, les modifications observées dans les niveaux d'expression des microARN reflètent plutôt le phénotype tumoral que les caractéristiques génomiques intrinsèques, puisque ces modifications ne sont pas corrélées avec les profils génomiques obtenus par technique CGH.

Plus généralement, cet article peut être vu comme une synthèse des applications cliniques envisageables dans un futur proche, grâce aux profils d'expression des microARN.

Document à suivre :

Article III : *MiRNA profiling: how to bypass the current difficulties in the diagnosis and treatment of sarcomas.*

Research Article

miRNA Profiling: How to Bypass the Current Difficulties in the Diagnosis and Treatment of Sarcomas

Angélique Gougelet,^{1,2} Jennifer Perez,^{1,2} Daniel Pissaloux,³ Anthony Besse,¹ Adeline Duc,¹ Anne-Valérie Decouvelaere,³ Dominique Ranchere-Vince,³ Jean-Yves Blay,^{1,2,4} and Laurent Alberti^{1,2}

¹ Unité INSERM U590 équipe Cytokines et Cancer, Centre Léon Bérard, 28 rue Laennec, 69373 Lyon cedex 08, France

² Conticanet (FP6-06188), France

³ Service d'Anatomie et Cytologie Pathologiques, Centre Léon Bérard, 69373 Lyon cedex 08, France

⁴ EORTC, 83/11 avenue Mounierlaan, 1200 Brussels, Belgium

Correspondence should be addressed to Angélique Gougelet, angeliquegougelet@gmail.com

Received 26 August 2010; Revised 17 November 2010; Accepted 3 January 2011

Academic Editor: Stephen Lessnick

Copyright © 2011 Angélique Gougelet et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Sarcomas are divided into a group with specific alterations and a second presenting a complex karyotype, sometimes difficult to diagnose or with few therapeutic options available. We assessed if miRNA profiling by TaqMan low density arrays could predict the response of undifferentiated rhabdomyosarcoma (RMS) and osteosarcoma to treatment. We showed that miRNA signatures in response to a therapeutic agent (chemotherapy or the mTOR inhibitor RAD-001) were cell and drug specific on cell lines and a rat osteosarcoma model. This miRNA signature was related to cell or tumour sensitivity to this treatment and might be not due to chromosomal aberrations, as revealed by a CGH array analysis of rat tumours. Strikingly, miRNA profiling gave promising results for patient rhabdomyosarcoma, discriminating all types of RMS: (Pax+) or undifferentiated alveolar RMS as well as embryonal RMS. As highlighted by these results, miRNA profiling emerges as a potent molecular diagnostic tool for complex karyotype sarcomas.

1. Introduction

Sarcomas are rare malignant tumours arising in connective tissues like fat, muscle, bones, and cartilage. According to molecular cytogenetic alterations, sarcomas could be divided into two classes: (1) sarcomas with specific alterations (translocation, oncogenic mutation) including Ewing sarcoma, gastrointestinal stromal tumours, and alveolar rhabdomyosarcoma (2) sarcomas with complex karyotype like leiomyosarcoma, pleomorphic liposarcoma, or osteosarcoma. Osteosarcoma is the most frequent primary malignant bone tumours, characterized by its metastatic potent particularly in lung sites and its resistance to conventional treatments like chemotherapy and radiotherapy [1]. Even if the median survival of osteosarcoma patients has been improved through preoperative administration of chemotherapeutic agents, there are nowadays around 40%

poor-responder patients [2]. In fact, osteosarcoma tumours often resist or relapse to presurgical chemotherapeutic treatment, and only few therapeutic options are possible and generally noncurative [3]. A second intensive cure of chemotherapy is currently administered in this case. Thus, it seems essential to develop a diagnosis tool to predict tumour response to chemotherapy to avoid the administration of inefficient drugs. There is also a need for efficient therapeutic alternatives based on the discovery of new targets involved in osteosarcoma tumourigenesis.

Rhabdomyosarcoma (RMS) is one of the most common soft-tissue sarcoma. Three types of RMS are observed: alveolar RMS (20%), embryonal RMS (eRMS, 60%), and pleomorphic RMS (20%). 70% aRMS present a specific translocation of the transcription factor Pax3 at the 3' end of FOXO1, creating a potent transcription factor able to induce myogenesis and survival [4]. 10% aRMS present

a translocation of Pax7 with FOXO1 [5]. aRMS are of bad prognosis as compared to eRMS, particularly those with Pax3 fusion gene [6]. Thus, it appears primordial to obtain a diagnosis tool identifying precisely the RMS subtypes, and particularly discriminating Pax-aRMS from eRMS, difficult to separate according to patient survival characteristics, gene expression profiles, and CGH arrays [7].

Micro-RNAs (miRNAs) are promising diagnosis biomarkers with their tissue specificities and their involvement in oncogenic process [8]. miRNAs are non-coding small RNA molecules synthesized from intronic regions with a size range from 16 to 35 nucleotides. They are processed by specific complexes of proteins containing Drosha and Dicer to be matured and finally integrated in RISC complexes [9, 10]. Mature miRNAs match with complementary sequences in messenger RNAs resulting in translation inhibition and accelerated mRNA degradation [11]. miRNA expression levels are characteristic for one tissue to regulate gene expression during growth and development, as it was shown for skeletal tissue and muscle development [12–14]. Their expression is also deregulated in many cancers [15, 16], resulting in a tumour miRNA signature, which could be useful for their classification in line with their tissue origin and molecular alterations [17–19]. Thus, they currently constitute potent biomarkers for cancer diagnosis [18, 20] with their abilities to be detected in patient serum. A noninvasive diagnostic tool based on miRNAs for osteosarcoma could be very useful to adapt chemotherapy protocols to tumour biological specificities.

In this study, we performed the miRNA profiling of sarcoma cell lines, human or rat tumours, to assess if miRNAs could constitute potent biomarkers to surpass the current limitations for rhabdomyosarcoma diagnosis and osteosarcoma treatment. miRNA expression levels were determined using microfluidic cards performing high-throughput TaqMan Low Density Arrays (TLDA), a real-time quantitative PCR (RT-qPCR) assays based on TaqMan technology. We firstly studied the effects of different chemotherapeutic agents on osteosarcoma cell miRNA profiles; we observed that these miRNA signatures were cell specific and drug specific. A CGH array of osteosarcoma tumours obtained from a rat model revealed that this miRNA signature, conserved in rat and human cells, was independent of chromosomal rearrangements, suggesting that miRNA profiles were linked to tumour phenotypes rather than to their genetic background. Of great interest, a miRNA signature was identified in rhabdomyosarcoma tumours from patients in accordance with the molecular translocation Pax3 or Pax7. This signature was in fact a potent tool to discriminate alveolar RMS (Pax-) from embryonal RMS, indistinguishable by the molecular techniques currently used. In conclusion, miRNA profiling constitutes a promising technology as an alternative or a partner of usual molecular techniques to overcome the present difficulties in diagnosis and treatment of sarcomas.

2. Experimental Procedures

2.1. Human Rhabdomyosarcoma Tumours. Seventeen patients treated for rhabdomyosarcoma in the Centre Léon

Bérard were included in this study. Four frozen tumours and thirteen formalin-fixed paraffin-embedded tumours were obtained from biopsies realized at the diagnosis. Tumour diagnoses were realized by a referent anatomopathologist specialist for this pathology by immunohistochemistry, FISH, and qPCR.

2.2. Cancer Cell Lines. Five cancer cell lines were obtained from ATCC (Manassas, VA, USA): the two human osteosarcoma MNNG/HOS Cl #5 [R-1059-D] (reference CRL-15-47) and Saos-2 (HTB-85) cells, the chondrosarcoma cell line SW1353 (HTB-94) and the two Burkitt lymphoma Daudi (CCL-213) and Namalwa (CRL-1432) cells. Osteosarcoma and chondrosarcoma cells were grown in DMEM (Gibco, Carlsbad, CA, USA), supplemented with 10% decomplexed fetal calf serum (Lonza, Basel, Switzerland), 10 mL penicillin streptomycin (10 U/mL/10 µg/mL, Gibco, Carlsbad, CA, USA), and 5 mL L-glutamin (200 mM; Gibco) at 37°C humidified atmosphere containing 5% CO₂. Lymphoma cells were grown in RPMI (Gibco, Carlsbad, CA, USA). Cells were exposed to 100 nM RAD-001 (Novartis), 50 µM ifosfamide (ifos, Baxter) or 1 µM cisplatin (CDDP, TEVA) or 100 µM methotrexate (MTX, TEVA) for 24, 48, and 72 h.

2.3. Rat Osteosarcoma Model. Procedures for animal care were performed according to institutional and national guidelines. Animals were anesthetized throughout all surgical and imaging procedures with isoflurane/oxygen (2.5%/2.5%, v/v) (Minerve, Esternay, France). The transplantable orthotopic and metastatic rat osteosarcoma model has been previously described [21–23]. This model mimics its human counterpart in terms of aggressiveness, metastatic spreading and chemoresistance phenotype [21–23]. All the tumours obtained were classified as osteoblastic following histological analyses. Briefly, small tumour fragments (100 mm³) taken from a hyperproliferative osteogenic tumour area were grafted on 3-weeks old immunocompetent Sprague-Dawley rats (Charles River Laboratories, Wilmington, MA, USA). Using a lateral approach, a tumour fragment was placed contiguous to tibial diaphysis after periosteal abrasion; then, the cutaneous and muscular wounds were sutured. Fourteen days after tumour transplantation, animals underwent a first ¹⁸F-FDG PET Scan and were randomly assigned to a control group treated with saline solution or a treated group exposed to a subcutaneous dose of 10 mg/kg ifosfamide (ifos, Baxter, Deerfield, IL, USA), 7 days apart (at days 15 and 22 after tumour transplantation). A second ¹⁸F-FDG PET Scan was performed 7 days after the second ifos administration. Animals were sacrificed one week after the end of the treatment. Tumour and normal tissue fragments (muscle, bone, and lung) were collected for RNA extractions.

2.4. RNA Extraction and Quantitative Real-Time PCR. FFPE tumours were lysed for 24 h in ATL buffer (Qiagen, France) supplemented with proteinase K (Qiagen) at 60°C in rotative agitation after different washes with toluene, ethanol, and

tris/EDTA in this order. Total RNA was extracted from tumour or cell pellets using a single phenol/chloroform extraction protocol with Trizol, according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA). Five hundred nanograms of total RNA were subjected to the microfluidic PCR technology performed by Applied Biosystems (Foster City, CA, USA). In brief, RNA was reversed transcribed, using multiplexed specific looped miRNA primers from the Taqman MicroRNA Reverse Transcription kit. The second step consists in a real-time quantitative PCR on TLDA: RT products are introduced through microchannels into miniature wells that are preloaded with dehydrated specific primers and probes. Recently, Applied biosystems released the second version of TLDA, consisting of two cards A and B. Analyses were performed for 377 miRNAs on card A and 290 on card B.

2.5. PCR Data Normalization. For each miRNA, the threshold cycle (Ct) was calculated by the ABI 7900 Sequence Detection System software (plate by plate manual Ct analysis with a threshold at 0.25 and automatic baseline). All further data manipulations were done using R scripts. A cutoff of 32 was applied to discard the late Ct values, except for RMS analysis. Around 60% of miRNAs passed the filtering criteria and were used for further analysis. For each TLDA, quality controls were performed on the raw data by checking internal controls and using box plot and scatter plot diagrams. Samples with any kind of problems were discarded so they would not introduce bias during the following normalization procedures. We tested different methods of normalization since the recommended "pseudo" normalization factor mammU6 plotted in each card was not stably expressed in our different samples. Normalization with the two most stable miRNAs identified by GeNorm was not optimal too. Finally, a global normalization by the median was chosen for its reliability over experiments. Tissues included in a given analysis were treated altogether, the normalization procedure being applied separately for the two types of card, A and B. Distribution of normalized data was checked with box plots and correlation plots. The following formula was used to correct Ct values of every card:

$$(\text{Normalized Ct}) = \frac{\text{Ct} \times (\text{mean of medians})}{(\text{median of the card})}. \quad (1)$$

Through this approach, the new median value shared by all samples can be considered as a sort of perfect "virtual housekeeping gene". Therefore, the standard $\Delta\Delta\text{Ct}$ method can be used to determine the relative quantities (RQ) as follows:

$$\Delta\text{Ct} = (\text{Normalized Ct}) - (\text{New shared median}). \quad (2)$$

For the $\Delta\Delta\text{Ct}$ calculation, it was more relevant for the statistical analyses to use the mean of all ΔCt obtained across

samples for each miRNA, instead of using the ΔCt of a reference sample

$$\Delta\Delta\text{Ct}$$

$$= \Delta\text{Ct} - (\text{Mean of } \Delta\text{Ct} \text{ across samples for each miRNA}), \quad (3)$$

$$\text{RQ} = 2^{-(\Delta\Delta\text{Ct})} \quad (4)$$

2.6. miRNA Target Predictions. We compiled 4 databases to determine miRNA targets: TargetScan 5.1, MiRanda, PICTAR, and the miRbase databases. These databases search the presence of conserved 8mer and 7mer sites on the 3'UTR parts of messenger RNA that match the seed region of each miRNA. It also predicts the efficacy of targeting for each matching site. We created our own database which regrouped each miRNA with the geneID of all their protein targets, for rat and human. We only conserved couples miRNA/geneID present in two databases at least.

2.7. miRNA-Regulated Cell Signalling Pathways Predictions. We used the "G-language microarray" web application, which allows the mapping of molecular dataset onto "Kyoto Encyclopedia of Genes and Genomes" (KEGG) pathway maps [24]. We first input miRNA-targeted proteins of interest and the sum of RQ values for all miRNAs that regulate these proteins, contained between 1 and 50; the software then generates KEGG data to create FLASH graphics of cell signalling pathways in which proteins are involved. The colour intensity of a highlighted protein varies with the strength of its regulation by miRNAs.

2.8. Proliferation Assay. Cells were plated in 96 well plates at 5000 cells/well and exposed to 100 nM RAD-001, 50 μM ifosfamide, 100 μM methotrexate, or 1 μM cisplatin or not (NT). Cell growth was measured 24, 48, and 72 h later with 20 μL Cell Titer Glo luminescent reagent (Promega, Madison, WI, USA) for 10 min. Luminescence was recorded using a Microbeta reader (PerkinElmer, Fremont, CA, USA).

2.9. Western Blot. Pelleted cells were resuspended in lysis buffer (Tris 50 mM pH 7.4, NaCl 250 mM, EDTA 5 mM, NaF 50 mM, Triton X-100 0.1%, orthovanadate 1 μM) plus protease inhibitors for 30 min on ice. After a centrifugation at 14000 rpm for 10 min, supernatants were boiled for 5 min in Laemmli sample buffer (Biorad, Hercules, CA, USA). Analysis of protein content was performed on 4%–12% gradient gel. After electrophoretic separation, 30 μg proteins were electrotransferred on a polyvinylidene difluoride membrane (Immobilon P, Millipore corp., Bedford, MA, USA). The membrane was then blocked for 1 h at room temperature with blocking agent 0.2% in PBS/Tween 0.1%, probed overnight with a primary rabbit antibody against the protein of interest, and finally revealed with a secondary antirabbit antibody HRP conjugated (Upstate Biotechnology, Lake Placid, NY, USA) and ECL Advance system (GEhealthcare, Chicago, IL, USA). Primary antibody used was obtained from Cell Signaling (New England Biolabs, Beverly, MA,

USA) used at 1/1000. The β actin was used as a reference (Sigma).

2.10. CGH Array. Oligonucleotide-based microarray analysis was performed using a custom-designed, 244K-feature whole-rat genome microarray manufactured by Agilent Technologies (Santa Clara, CA). Genomic DNA labeling, array hybridization, and washing were performed as specified by the manufacturer (Agilent Technologies). Results of aberration calls consisting of three or more consecutive oligos were then displayed using custom oligonucleotide CGH analysis software (Genespring).

2.11. Statistical Analysis. Normalized RQ data were directly input into the TIBCO Spotfire DecisionSite for Functional Genomics analysis software. We performed unsupervised hierarchical clustering to classify samples by groups. The selection of miRNAs useful to predict tumour response to treatment was statistically realized using ANOVA tests with P values of .05 at least. Results were verified through supervised hierarchical clustering.

Data from miRNA lists of interest were then used as variables in a three-dimensional principal component analysis (PCA) performed with R 2.9.0 package to demonstrate their capabilities to distinguish types of tumours. PCA supplies a simplified three-dimensional picture to our multivariate dataset of miRNA RQ values. By mathematical combination of values according to their strength, three principal components are created that represent as much as possible the variability of the data. Thus, tumours possess three new coordinates in a three-dimensional space. According to their localization in this space, tumours form groups, and their subtypes can be predicted.

3. Results

3.1. miRNA Signatures of Osteosarcoma Cell Lines. In our recent study published in International Journal of Cancer, we showed that the two osteosarcoma Saos-2 and CRL-15-47 (15-47) cells mimic the biological response of human osteosarcoma and tumours obtained from a rat model. In fact, we identified in an osteosarcoma rat model a panel of 61 miRNAs discriminating tumours with a good response to ifosfamide from those with a bad response [25]. On the basis of this signature, we realized a principal component analysis allowing predicting tumour response. In this PCA diagram, we could notice that the Saos-2 cells were predicted as sensitive to ifosfamide contrary to 15-47 cells (Figure 3(b) [25]), according to the results obtained by a proliferation assay (Figure 6(a) [25] and Figure S1). This was confirmed by a PCA analysis realized with the miRNA signature identified in human tumours (Figure S2). We so considered that these two cell lines were an interesting model to study the importance of miRNAs in cell response to treatment and to identify new therapeutic strategies.

3.2. miRNA Signatures of Human Cancer Cell Lines. We firstly performed a preliminary miRNA profiling on different cell models to compare the miRNA profiles of osteosarcoma

cells used in our laboratory to perform *in vitro* experiments, Saos-2 and 15-47 cells, with the chondrosarcoma cells SW-1353 (chondro) and the Burkitt lymphoma Daudi and Namalwa cells. In a previous study, we identified 61 miRNAs involved in osteosarcoma cell response to treatment [25]. We only conserved these miRNAs to realize an unsupervised hierarchical clustering with the five cancer cell lines. As shown in Figure 1(a), this miRNA signature was representative of the two human osteosarcoma cell lines, since these two cells clustered together independently but closely to the chondrosarcoma cells. These three cell lines were classed in a distinct group from the two lymphoma cells Daudi and Namalwa. This confirmed that each cancer cell line presents a miRNA signature in accordance with their origin, as shown by others [15, 16].

3.3. miRNA Profiles in Response to Chemotherapeutic Agents Were Cell Specific. Then, we assessed if miRNA profiles were specifically modified in response to chemotherapy. We chose to expose osteosarcoma and lymphoma cells to ifosfamide, an alkylating chemotherapeutic agent currently used for paediatric osteosarcoma. A proliferation assay based on ATP measurement showed that the only Saos-2 cell line was moderately sensitive to 50 μ M ifosfamide after 48 h exposure (proliferation inhibition around 30%) (Figure S1). Based on this observation, we decided to expose these cells to 50 μ M ifosfamide for 24 h to realize miRNA profiling. On the basis of the panel of 61 miRNAs identified in our previous study [25], osteosarcoma cells were markedly different from lymphoma cells, confirming that miRNA profiles were cell specific as shown by the unsupervised hierarchical clustering in Figure 1(b). We could notice that Saos-2 cells present a unique miRNA signature in which the majority of miRNAs were overexpressed (in red in Figure 1(b)). A supervised hierarchical clustering realized following an ANOVA $P < .03$ between the Saos-2 sensitive cells versus the resistant cells revealed that they effectively clustered according to their sensitivity to ifos: Saos-2 in one hand, independently to 15-47 cells and both lymphoma cells (Figure 1(c)). We confirmed this observation with the other chemotherapeutic agent cisplatin. As previously, cells were classified according to their susceptibility to CDDP on the supervised hierarchical clustering in Figure S3A (ANOVA $P < .03$): the 15-47 and Namalwa cells, sensitive to CDDP based on the proliferation assay in Figure S3B, clustered together, independently to Daudi and Saos-2 cells refractory to this treatment.

3.4. Osteosarcoma Cell miRNA Profiles Were Specific of Each Chemotherapeutic Agent. Thus, since miRNA signatures of untreated as well as treated cells were cancer specific, we assessed if each chemotherapeutic drug induced a different miRNA profile in a same cell. As suggested previously for osteosarcoma cells, cisplatin and ifosfamide exposure resulted in quite different miRNA profiles. After a statistical analysis with an ANOVA $P < .03$, we only found two discriminating miRNAs common to both miRNA signatures induced by ifos and CDDP in the two cell lines (Figure S3). In this context, we test a third cytotoxic agent currently administered in osteosarcoma pathology, the methotrexate.

Sarcoma

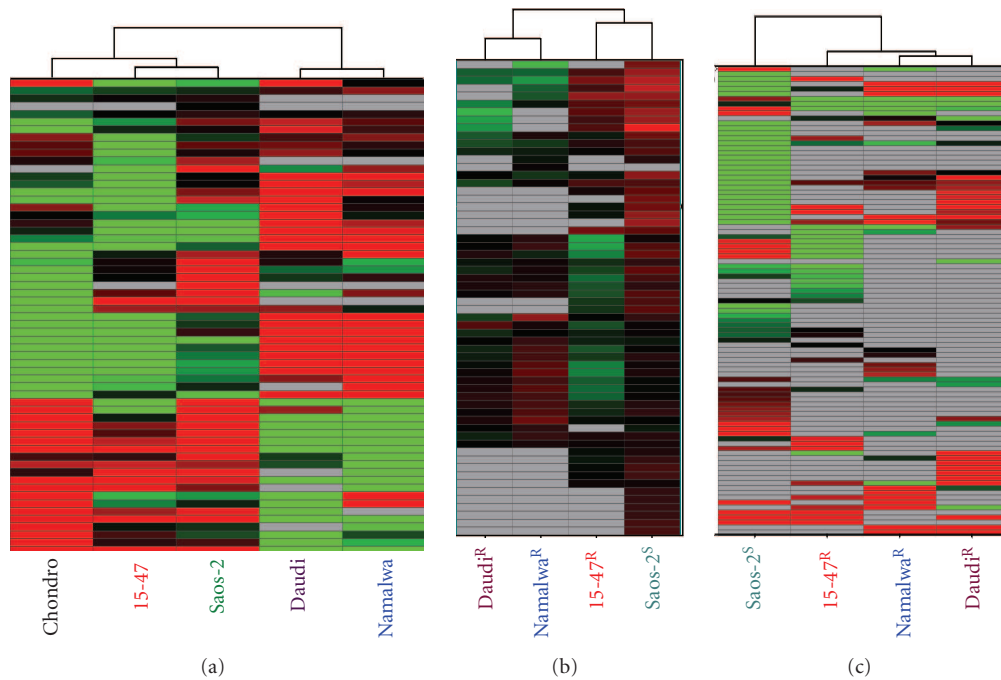


FIGURE 1: Cancer cell miRNA signatures were consistent with their tissue origin and with their sensitivity to ifosfamide. (a) This unsupervised hierarchical clustering only conserved the 61 miRNAs which discriminated osteosarcoma cells according to their response to treatment. Osteosarcoma cell lines clustered together near chondrosarcoma cells and independently to the lymphoma Daudi and Namalwa cells. Each row represents the relative levels of expression for each miRNA, and each column shows the expression levels for each sample. The red or green colour indicates relatively high or low expression, respectively, while grey squares indicate no expressed miRNA. (b) and (c) miRNA profiles after exposure to 50 μ M ifosfamide for 24 h. (b) This unsupervised hierarchical clustering only conserved the 61 miRNA differently expressed in osteosarcoma cells according to their sensitivity to ifos following an ANOVA ($P < .03$). (c) This supervised hierarchical clustering conserved miRNAs differently expressed in cells according to their response to treatment following an ANOVA ($P < .05$) after removing the miRNAs whose expression depends on the cell types. The red or green colour indicates relatively high or low expression, respectively, while grey squares indicate no expressed miRNA.

As shown in the unsupervised hierarchical clustering in Figure 2, only conserving the 61 miRNAs of interest for osteosarcoma response, as explained above, the miRNA signature in the two osteosarcoma cells Saos-2 and 15-47 strongly differed from those observed for ifosfamide and cisplatin. It is important to note that a majority of these miRNAs were overexpressed in both cell lines in response to MTX. This was relevant with their sensitivity to MTX as shown in the proliferation assay in Figure 2(b). In brief, it seems that discriminating miRNAs were generally overexpressed in the cells after exposure to a cytotoxic agent, to which they were sensitive, as it was also shown for ifosfamide in the Saos-2 cells (Figure 1(b)). This also confirmed that miRNAs predicting cell response to a treatment differed according to the drug.

On the basis of these preliminary *in vitro* results, we could suggest that miRNA profiles, due to their drug specificity, could be a potent tool to predict a cancer cell response to a treatment. Since osteosarcoma is currently resistant to conventional treatments, the prediction of its response to one agent could be a progress for this pathology.

3.5. Osteo- and Chondrosarcoma Cell Response to the mTOR Inhibitor RAD-001. As highlighted by these previous data,

we were able to classify and predict osteosarcoma cell response to chemotherapy. Our algorithms were not only interesting for chemotherapeutic agents but also promising to identify new targeted therapies to encounter osteosarcoma resistance. Thus, we tested a potent drug for skeletal sarcoma treatment, which inhibits the pro-oncogenic protein mTOR, called RAD-001 (Everolimus, Novartis). mTOR is often aberrantly activated in cancers and, in particular in chondrosarcoma [26] and osteosarcoma [27]. mTOR signalling has been described as implicated in tumour development, metastasis, and drug resistance [28, 29]; thus, mTOR targeting successfully inhibits tumour growth and renders them sensitive to conventional treatments [30, 31]. RAD-001, acting in a similar manner than rapamycin through the inhibition of mTORC1 complexes, is currently tested in various clinical trials for renal cell carcinoma (RECORD program), advanced papillary tumours (RAPTOR), metastatic neuroendocrine tumours (RAMSETE), or breast cancers (BOLERO).

Thus, we performed *in vitro* experiments on chondrosarcoma and osteosarcoma cells with 100 nM RAD-001. The Saos-2 and chondrosarcoma cell proliferation was reduced of 40% following exposure to RAD-001 during 72 h contrary to 15-47 cell growth (Figure 3(a)). In parallel, we

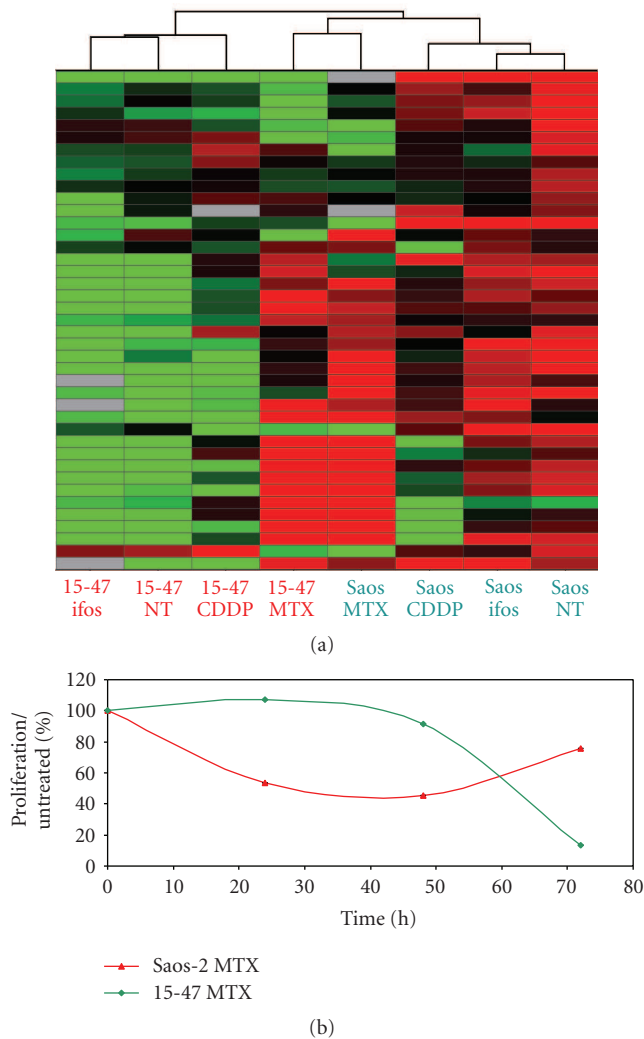


FIGURE 2: miRNA expression profiles osteosarcoma cells were specific for each chemotherapeutic agent. (a) This unsupervised hierarchical clustering conserved the miRNAs identified as discriminating for ifosfamide response after removing the miRNAs whose expression depends on the cell types. Each row represents the relative levels of expression for each miRNA, and each column shows the expression levels for each sample. The red or green colour indicates relatively high or low expression, respectively, while grey squares indicate no expressed miRNA. (b) Cell growth was measured by the Cell Titer GloLuminescent assay as described in Section 2.6 24, 48, and 72 h after exposure to 100 μ M methotrexate. Results were represented as the mean % of proliferation normalized to untreated cells of two independent experiments realized in duplicate.

realized Western blot with RAD-001 on chondrosarcoma and osteosarcoma cells concerning the major actors of the mTOR cell signalling pathways. This revealed that the mTOR pathway was inhibited by RAD-001 in chondrosarcoma cells contrary to 15-47 cells, in particular eIF4G and p70 S6 kinase whose phosphorylation level was decreased (Figure 3(b)).

Thus, we analysed if the miRNA signatures of these cells were different and could explain their differential response to RAD-001. We performed a supervised hierarchical clustering

between untreated Saos-2, chondrosarcoma and 15-47 cells following an ANOVA with $P < .05$. This clustering revealed that 16 miRNAs discriminated the chondrosarcoma and Saos-2 cells in one hand and the 15-47 cells in the other hand (Figure 4). Except miR-146, Saos-2 and chondrosarcoma overexpressed these contributory miRNAs.

Thereafter, as we have explained in our previous study on osteosarcoma [25], miRNA profiling constitutes a potent tool to identify miRNA-targeted cell signaling pathways through an *in silico* approach. In our case, we searched if these miRNAs shown as differently expressed in cells according to their response to RAD-001 potentially target the mTOR signalling pathway. We created a database, as described in Section 2.6 which determine the predicted targets for these miRNAs described in the miRbase. Then, we summed up the RQ values for each miRNA in the Saos-2 and chondrosarcoma cells sensitive to RAD-001 and concatenated with the geneID of their protein targets. We finally inserted these data in the G-language microarray web application, which connects miRNA targets according to their involvement in similar KEGG pathways, the mTOR pathway in this case. As shown in Figure 5, the mTOR pathway is targeted by these miRNAs and particularly its downstream proteins implicated in VEGF signaling and autophagy processes, in particular RICTOR, ATG1, and HIF-1 α . Thus, the Saos-2 and chondrosarcoma cells overexpressed miRNAs that potentially inhibit mTOR signalling. Inhibition of these miRNAs through the use of Locked Nucleic Acid (LNA) and qPCR measurement of RICTOR, ATG1, and HIF1 α could confirm this concept.

To resume, miRNAs constitute potent biomarkers to determine the susceptibility to a treatment and could be very useful to identify new therapeutic targets as an alternative of chemotherapy for chondrosarcoma and osteosarcoma often refractory to this treatment. In the next steps, we assessed if these observations were relevant *in vivo*, with a model of rat osteosarcoma and with patient samples.

3.6. Predictive miRNA Signature of a Rat Osteosarcoma Model Was Probably Not Related to DNA Aberrations. As described in other studies realized by members of our team [21, 22, 25], we possess a rat osteosarcoma model mimicking the human pathology concerning aggressiveness, chemoresistance and the apparition of lung metastases (see Section 2.6). The treatment of animals with ifosfamide results in two groups, the good versus the bad or moderate responders, in a proportion closer to that observed for patients. By miRNA profiling, we were able to distinguish tumours sensitive to ifosfamide from those refractory to this drug and above all to predict the response of untreated tumours with ten miRNAs through the use of statistical algorithms created in our lab [25]. Following these interesting data, we would like to confirm that this miRNA signature was specific of tumour response to treatment and not related to different tumour genetic backgrounds. We thus realized an analysis in CGH array with the same tumours used for miRNA profiling. We analysed two tumours of each type, untreated, treated with ifosfamide and good responder, or treated with ifosfamide and bad responder, as compared to the

Sarcoma

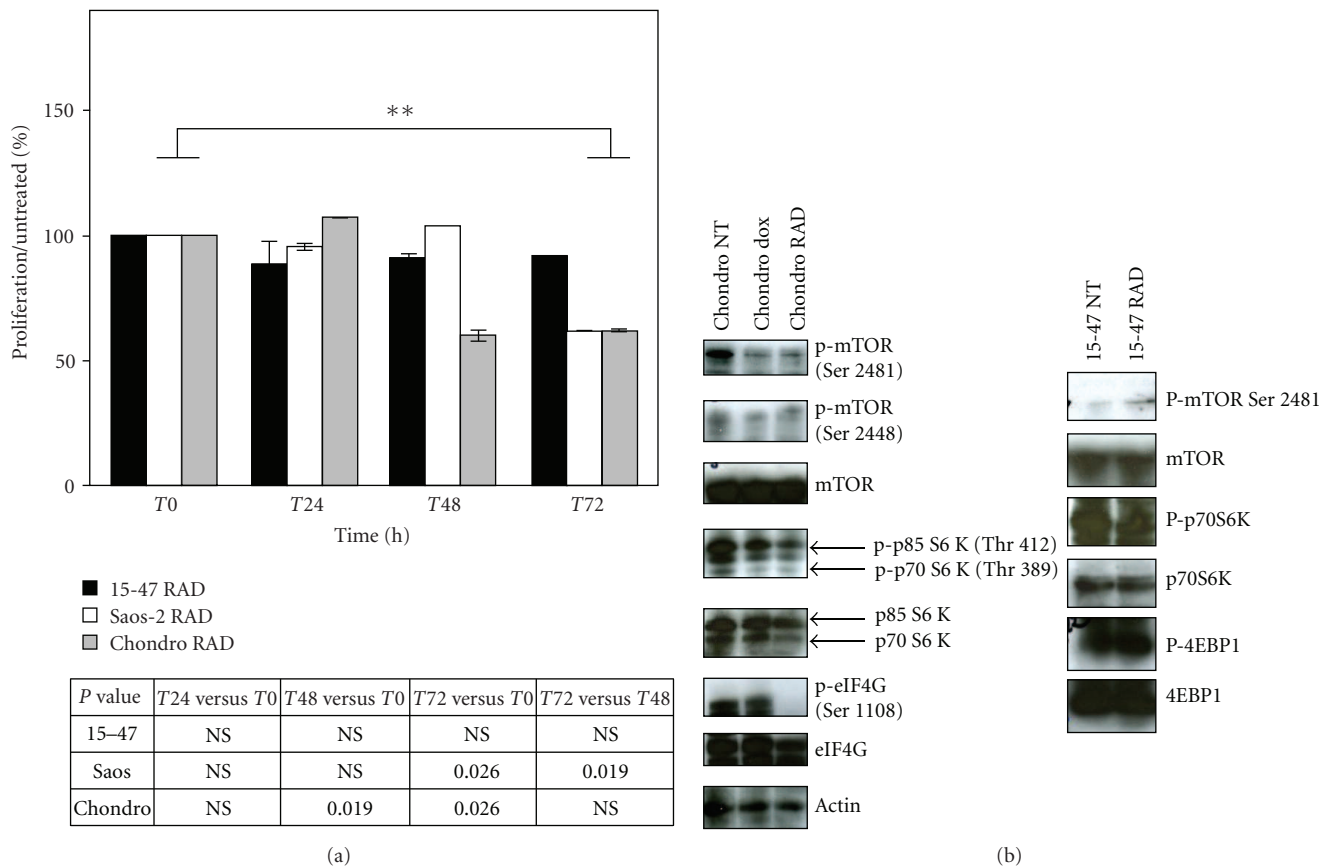


FIGURE 3: mTOR inhibition by RAD-001 in chondrosarcoma resulted in cell proliferation inhibition. (a) Cell growth was measured by the Cell Titer GloLuminescent assay as described in Section 2.6 24, 48, and 72 h after exposure to 100 nM RAD-001. Results were represented as the mean % of proliferation normalized to untreated cells of two independent experiments realized in duplicate. A Fisher test was realized, NS corresponds to “nonsignificant”. (b) Western blot analysis of RAD-001 effects on the mTOR pathway in the chondrosarcoma and 15-47 osteosarcoma cells exposed or not (NT) to doxorubicin (dox). 30 μ g protein extracts were analysed by Western blot with antibodies 1/1000 against actors of the mTOR pathway (phosphorylated form or not) (Cell signalling, Beverly MA).

same untreated bone sample, the reference tissue in CGH analysis. The majority of chromosomal aberrations observed in CGH array was common to untreated tumours and treated tumours, regardless of their response to treatment (Figure 6). The few different abnormalities were essentially linked to individual tumour biological specificities.

We compiled all abnormalities and verified in our “home-made” database if any miRNA, identified as discriminating of tumour response, was located in these DNA regions. Interestingly, this *in silico* analysis also revealed that neither miRNA nor gene were present in the few differential aberrations observed in these tumours, in particular in the chromosome 4 (Figure S4), suggesting that the different miRNA profiles were rather linked to tumour response to treatment and not due to upstream chromosomal rearrangements. Although we could not rule out that some transacting proteins could be deregulated consequently to these aberrations, this suggested that all tumours were homogeneous and that an increase in some miRNAs in sensitive tumours were due to their upregulation and not to a genetic amplification.

It seems that molecular diagnosis based on miRNA profiling highlights the tumour behaviour, that is, in response to a treatment, and thus a phenotype rather than a genotype contrary to CGH array. These two molecular techniques could be a couple of choice to improve the care of patients with pathologies currently hardly to diagnose.

3.7. Rhabdomyosarcoma miRNA Profiles Were Correlated to their Histological Subtypes. Finally, to corroborate the previous idea considering that miRNA profiling could be very helpful for uncertain diagnoses, we performed the miRNA profiling of rhabdomyosarcoma samples. In fact, we recently showed that miRNA profiling was reliable for osteosarcoma diagnosis on 29 formalin-fixed paraffin embedded (FFPE) biopsies of patients [25]. Based on the expression level of a panel of five miRNAs, we successfully separated good responders from bad responders to treatment. So, we assessed if our TLDA platform was also competitive for RMS diagnosis. We obtained seventeen tumours including alveolar RMS patients, (Pax3+) (3 patients) or Pax7+ (2), embryonal RMS patients (6) and negative fusion aRMS (6).

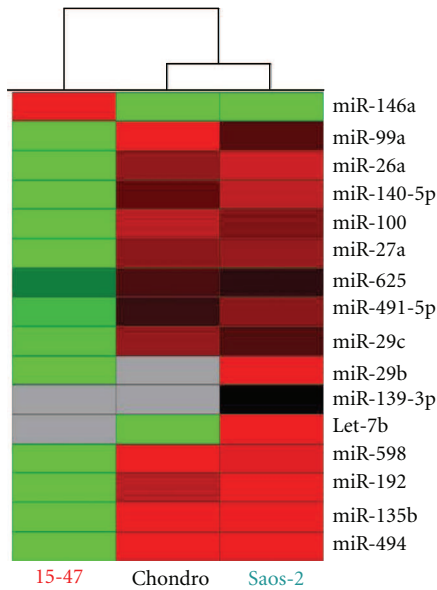


FIGURE 4: miRNA expression profiles of osteosarcoma and chondrosarcoma cells were consistent with their sensitivity to the mTOR inhibitor RAD-001. This hierarchical clustering only conserved miRNA differently expressed in tumours according to their response to treatment following an ANOVA ($P < .05$). Each row represents the relative levels of expression for each miRNA and each column shows the expression levels for each sample. The red or green colour indicates relatively high or low expression, respectively, while grey squares indicate no expressed miRNA.

All these tumours were diagnosed through the use of immunohistochemistry, FISH and qPCR, which were validated by a referent anatomopathologist (Table S5). A supervised hierarchical clustering on rhabdomyosarcoma tumours following an ANOVA with a P value $< .03$ between the four types of RMS, revealed that tumours clustered according to their molecular alterations Pax3/FOXO1, Pax7/FOXO1 or no translocation, on the basis of the expression level of 10 miRNAs (Figure 7(a)). (Pax+) tumours, particularly those (Pax3+) overexpressed all these miRNAs. Then, we performed a statistical analysis with these ten miRNAs based on Principal Component Analysis, a method which allows studying the variability between a set of variables. This consists of assigning a new system of three coordinates to each contributory miRNA by a mathematical procedure. Then, RQ values of each miRNA are adjusted for each tumour by the new coefficients obtained previously and summed up. Thus, a 3-dimension PCA diagram was realized with the three new coordinates for each tumour (Figure 7(b)). Through this mathematical representation, we could distinguish (Pax+) from fusion negative aRMS and eRMS. eRMS also constitutes an independent group with a high value of component 2 (represented in the y-axis on Figure 7(b)). The fusion negative aRMS constitute a separate group even if some samples were difficult to classify in accordance with their uncertain diagnosis. Even if the number of samples was low for each subset, a statistical

analysis showed a significant P value between (Pax3+) and (Pax7+) and between (Pax+) and (Pax-) tumours, 0.05 and 0.0005 respectively (Figure S6).

We showed that miRNA profiling was a potent tool to discriminate fusion negative aRMS from embryonal RMS. miRNAs could be useful biomarkers to improve the diagnosis of this type of RMS, since fusion negative aRMS are currently molecularly indistinguishable from eRMS [7].

4. Discussion

miRNA signatures are observed for many types of cancers, that is, sarcoma [19], breast and prostate cancers [18, 32]. These signatures constitute potent diagnosis and prognosis tools for chronic lymphocytic leukemia [33], colon adenocarcinoma [34], or lung cancers [35]. Here, we showed that osteosarcoma cell lines also expressed miRNA patterns different from those of chondrosarcoma and lymphoma cells (Figure 1(a)) and which allow us to discriminate cell response to chemotherapeutic treatment (Figures 1(b), 2, and S3). In addition, osteosarcoma miRNA signatures were cell and drug specific (Figure 2(a)). This drug specificity of osteosarcoma has also been observed by Song et al. with U2-OS osteosarcoma tumour xenografts, in which different miRNAs were deregulated in response to the chemotherapeutic agents doxorubicin, cisplatin, and ifosfamide; only 3 miRNAs were commonly found deregulated in response to all drugs [36]. With their specificity, miRNAs constitute promising biomarkers to anticipate the tumour response to a treatment of interest. As we have recently shown, through miRNA profiling, we were able to predict osteosarcoma tumour response to chemotherapy for rat tumours as well as for patient FFPE biopsies [25]. Here, we showed that miRNA profiles of osteosarcoma cells were in accordance with their response to the mTOR inhibitor, RAD-001 (Figures 3 and 4). The miRNAs deregulated in response to this drug in sensitive cells, effectively targeted the mTOR pathway, in particular the downstream proteins eIF4G and p70 S6 kinase (Figure 3(b)), and potentially RICTOR, ATG1 and HIF1a, which might be validated by qPCR analysis (Figure 5).

In brief, miRNAs appeared very useful for the identification of new exciting therapeutic approaches through the targeting of some miRNA protein targets or some miRNAs involved in tumour development themselves. In future, we would like to confirm the implication of these miRNAs in treatment response *in vitro* through the use of miRNA mimics or inversely of Locked Nucleic Acid (LNA) against these miRNAs. As mentioned in this study, we possess an interesting *in vitro* osteosarcoma model, on which we could test the miRNA functionality in the presence of the different drugs used in this work. Following the validation of miRNA involvement *in vitro*, we would also test these mimics or LNAs *in vivo* in the model of rat osteosarcoma. This approach has been successfully employed in rhabdomyosarcoma through the conditional expression of miR-206 in mice [37] and could become a potent therapeutic strategies [38].

In addition to the identification of new targets, miRNA also constitute an interesting alternative to the conventional

Sarcoma

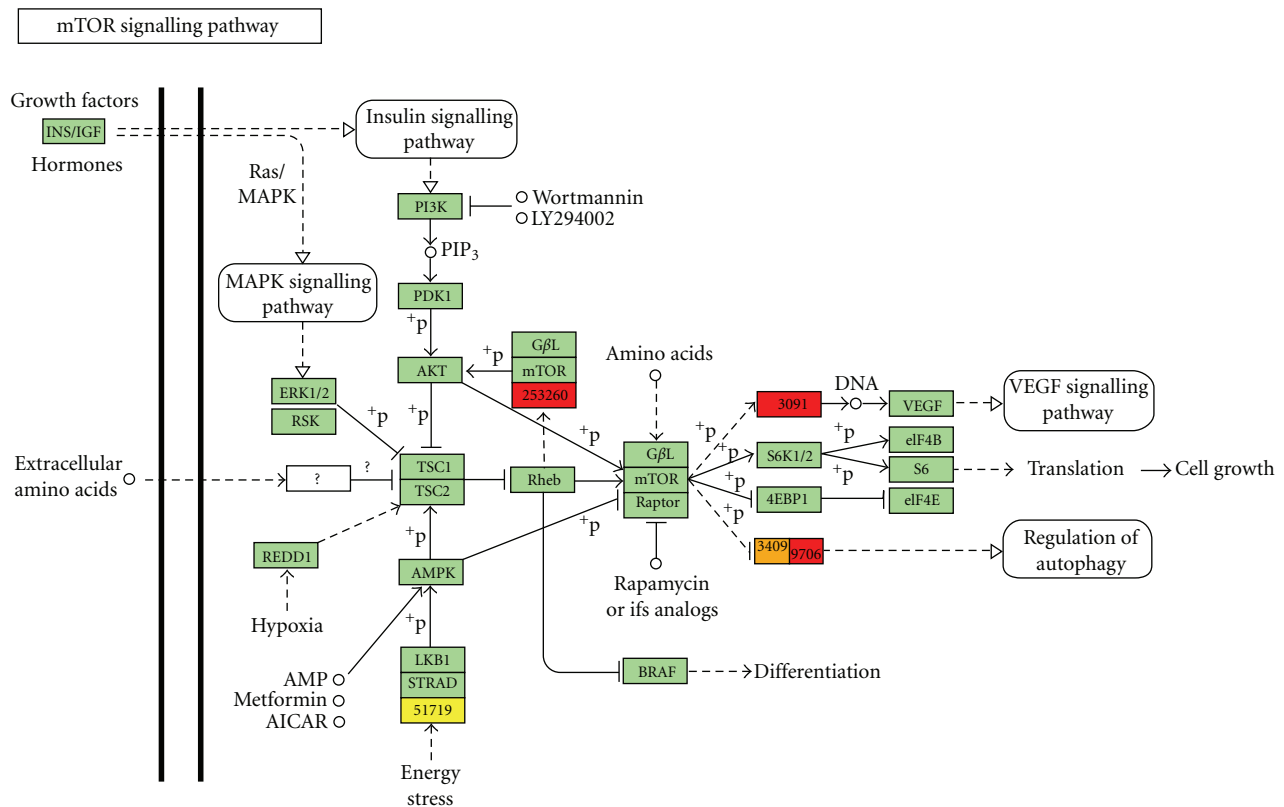


FIGURE 5: The discriminating miRNAs interfered with the mTOR pathway. Proteins in yellow, orange, and red colours represent targets of miRNAs; a yellow square represents a weak repression, while red represents the maximal repression; green squares were not targeted.

molecular technologies routinely used for cancer diagnosis. In fact, osteosarcoma present complex karyotypic alterations rendered them difficult to diagnose with current diagnostic methods, like CGH array [39]. With the rat osteosarcoma model, we confirmed that tumours presented numerous long chromosomal aberrations (Figure 6). These abnormalities were generally common to all tumours, regardless to their susceptibility to treatment and neither miRNAs of interest nor genes were located in these regions (Figure S4). Even if some proteins involved in the regulation of miRNA expression (trans-acting factors or epigenetic regulating factors) could be deregulated following these mutations, the miRNA profiles observed in rat tumours might be correlated to the effects of the cytotoxic drugs on the miRNA machinery and no to upstream DNA rearrangements. Even if miRNAs could be submitted to epigenetic regulation like methylation or acetylation, this only concerns 5% to 10% miRNAs, and we could consider that this process is minor for the miRNA signature of osteosarcoma tumours and cells based on 61miRNAs [40, 41]. Enthusiastically, our work is the first suggesting that miRNA signatures were not correlated to DNA amplifications, as it was observed for neuroblastoma [42] or in mixed lineage leukemia [43]. Although our cohort was not fully satisfying, it appeared that miRNA profiling could predict tumour response to treatment by reflecting tumour biological specificities and not genotypic characteristics. This work also highlights

miRNA measurement as an interesting partner to CGH array in the case of pathologies with unstable karyotypes. In the same way, Selvarajah et al. was the first to suggest a combination of CGH array and interphase FISH to better understanding osteosarcoma pathogenesis [44].

miRNA patterns were not only related with osteosarcoma phenotypic properties but also with rhabdomyosarcoma histological subtypes. By miRNA profiling, we were able to discriminate the different subtypes of rhabdomyosarcoma: Pax3+ or Pax7+ or fusion negative, classically difficult to diagnose by histological analysis (Figure 7(a)). This miRNA pattern was unique since all miRNAs identified as discriminating are no or weakly described in the literature. Very interestingly, on the basis of their miRNA profiles, our algorithms allow us to discriminate embryonal RMS from fusion negative aRMS (Figure 7(b)). It was in agreement with the work of Wachtel et al. identifying different expression profiles linked to aRMS (Pax+), fusion negative aRMS, and eRMS [45].

Altogether, it seems that miRNA measurement is advantageous for sarcoma with complex karyotype, since fusion negative RMS, similarly to osteosarcoma, are characterized by a complex karyotype linked to allelic imbalance, loss of heterozygosity and heterogeneous gene expression profiles. Although the molecular classification of fusion negative RMS is always controversial, our work corroborates the study of Davicioni et al. suggesting that Pax/FOXO1 dictates a specific

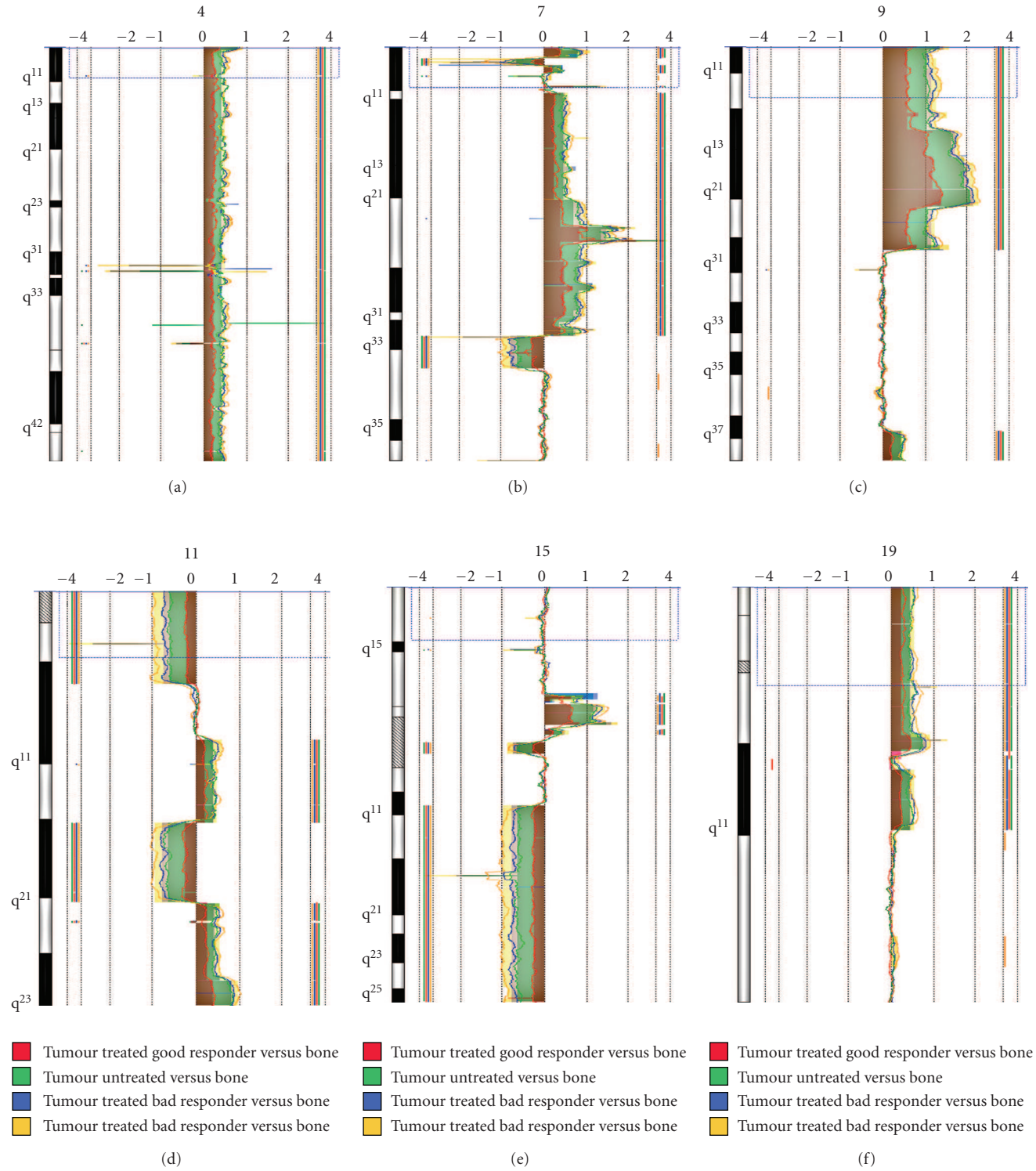


FIGURE 6: CGH analysis of six tumours from an osteosarcoma rat model. Here, we represent six chromosomes among the twenty + X chromosomes present in rat genome. All the analyses were performed with the same untreated bone sample as a reference.

expression signature in RMS by oligonucleotide microarray expression profiling [46]. Inversely, this differs from the recent work of Williamson suggesting that fusion negative alveolar rhabdomyosarcoma is difficult to distinguish from embryonal rhabdomyosarcoma concerning patient survival

characteristics, gene expression profiles, and CGH arrays [7]. In fact, our work and theirs were not totally contradictory, since they only focused on genomic analysis. As suggested for osteosarcoma, miRNA patterns reflect the phenotypic tumour properties rather than its genetic and could be

Sarcoma

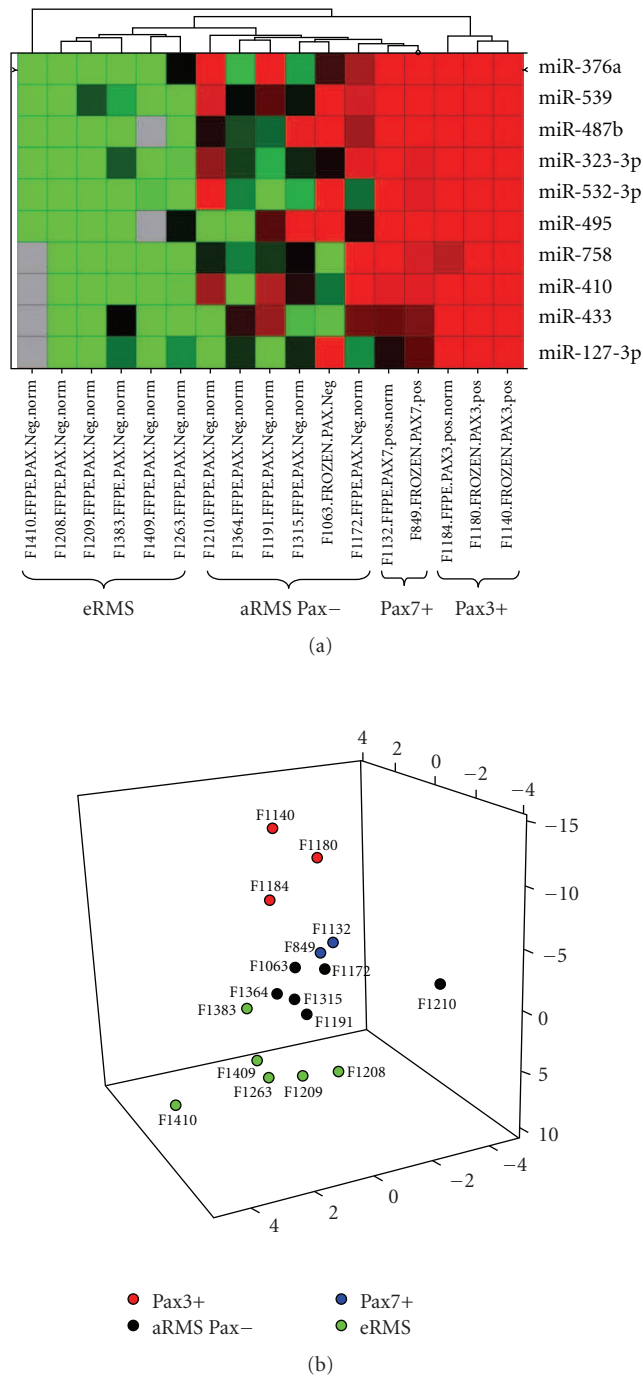


FIGURE 7: Rhabdomyosarcoma miRNA signatures were consistent with their molecular alterations. (a) This supervised hierarchical clustering only conserved miRNA differently expressed in the different subtypes of RMS following an ANOVA ($P < .03$). Each row represents the relative levels of expression for each miRNA, and each column shows the expression levels for each sample. The red or green colour indicates relatively high or low expression, respectively, while grey squares indicate no expressed miRNA. (b) Principal component analysis of RMS tumours as a tool to determine the potential tumour response to treatment. RQ values of the ten selected miRNAs for each tumour were corrected by the coefficients determined in (a), as described in Section 2.6.

a promising alternative for RMS diagnosis to surpass the current limitations of molecular analysis combined to traditional histopathology.

Thus, it seems that miRNA profiling could be very useful for osteosarcoma and rhabdomyosarcoma diagnosis. Here, we showed that on the basis of ten miRNAs, we were able to separate the different subtypes of RMS. We have previously suggested that a panel of five miRNAs was statistically sufficient to distinguish the potent response of osteosarcoma patients to treatment [25]. The TLDA technology presents numerous advantages including its need for few amount of total RNA and the possible analysis of FFPE samples, as it was previously shown by others [47–49]. This method is especially useful to detect circulating miRNAs in patient serum, an emerging field these two past years [50, 51]. A blood-based molecular diagnosis tool through miRNA profiling from patient serum could be a major advance for osteosarcoma, requiring a biopsy for its diagnosis, which could result in a secondary amputation.

Altogether, these promising results open up the way to a new diagnosis tool based on miRNA for osteosarcoma as well as rhabdomyosarcoma, which could improve patient survival in both cases through the prediction of patient response to chemotherapy and the precise identification of RMS subtypes, respectively.

Abbreviations

- CDDP: Cisplatin
- Chondro: Chondrosarcoma cells
- Ct: Threshold cycle
- Dox: Doxorubicin
- FFPE: Formalin-fixed paraffin embedded
- ifos: Ifosfamide
- KEGG: Kyoto encyclopedia of genes and genomes
- LNA: Locked nucleic acid
- NT: Non treated
- PCA: Principal component analysis
- RAD: RAD-001
- RMS: Rhabdomyosarcoma
- RQ: Relative quantity
- RT-qPCR: Real-time quantitative PCR
- TLDA: Taqman low density array
- 15-47: Human osteosarcoma cells CRL-15-47.

Acknowledgments

This work has been supported by the Ligue Nationale contre le Cancer (grants from the Ain department), the Institut National du Cancer, and the Conticanet Consortium. The authors acknowledge Imaxio, which realized the CGH analysis.

References

- [1] L. B. Rozeman, A. M. Cleton-Jansen, and P. C. W. Hogendoorn, "Pathology of primary malignant bone and cartilage tumours," *International Orthopaedics*, vol. 30, no. 6, pp. 437–444, 2006.

- [2] S. Ferrari and E. Palmerini, "Adjuvant and neoadjuvant combination chemotherapy for osteogenic sarcoma," *Current Opinion in Oncology*, vol. 19, no. 4, pp. 341–346, 2007.
- [3] A. J. Chou and R. Gorlick, "Chemotherapy resistance in osteosarcoma: current challenges and future directions," *Expert Review of Anticancer Therapy*, vol. 6, no. 7, pp. 1075–1085, 2006.
- [4] F. G. Barr, N. Galili, J. Holick, J. A. Biegel, G. Rovera, and B. S. Emanuel, "Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma," *Nature Genetics*, vol. 3, no. 2, pp. 113–117, 1993.
- [5] R. J. Davis, C. M. D'Cruz, M. A. Lovell, J. A. Biegel, and F. G. Barr, "Fusion of PAX7 to FKHR by the variant t(1;13)(p36;q14) translocation in alveolar rhabdomyosarcoma," *Cancer Research*, vol. 54, no. 11, pp. 2869–2872, 1994.
- [6] P. H. B. Sorensen, J. C. Lynch, S. J. Qualman et al., "PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the Children's Oncology Group," *Journal of Clinical Oncology*, vol. 20, no. 11, pp. 2672–2679, 2002.
- [7] D. Williamson, E. Missiaglia, A. de Reyniès et al., "Fusion gene-negative alveolar rhabdomyosarcoma is clinically and molecularly indistinguishable from embryonal rhabdomyosarcoma," *Journal of Clinical Oncology*, vol. 28, no. 13, pp. 2151–2158, 2010.
- [8] A. Esquela-Kerscher and F. J. Slack, "Oncomirs—microRNAs with a role in cancer," *Nature Reviews Cancer*, vol. 6, no. 4, pp. 259–269, 2006.
- [9] G. Hutvagner and P. D. Zamore, "A microRNA in a multiple-turnover RNAi enzyme complex," *Science*, vol. 297, no. 5589, pp. 2056–2060, 2002.
- [10] Y. Lee, C. Ahn, J. Han et al., "The nuclear RNase III Drosha initiates microRNA processing," *Nature*, vol. 425, no. 6956, pp. 415–419, 2003.
- [11] J. G. Doench and P. A. Sharp, "Specificity of microRNA target selection in translational repression," *Genes and Development*, vol. 18, no. 5, pp. 504–511, 2004.
- [12] Z. Li, M. Q. Hassan, M. Jafferji et al., "Biological functions of miR-29b contribute to positive regulation of osteoblast differentiation," *Journal of Biological Chemistry*, vol. 284, no. 23, pp. 15676–15684, 2009.
- [13] T. Sugatani and K. A. Hruska, "MicroRNA-223 is a key factor in osteoclast differentiation," *Journal of Cellular Biochemistry*, vol. 101, no. 4, pp. 996–999, 2007.
- [14] K. Chen and N. Rajewsky, "Natural selection on human microRNA binding sites inferred from SNP data," *Nature Genetics*, vol. 38, no. 12, pp. 1452–1456, 2006.
- [15] J. Lu, G. Getz, E. A. Miska et al., "MicroRNA expression profiles classify human cancers," *Nature*, vol. 435, no. 7043, pp. 834–838, 2005.
- [16] S. Volinia, G. A. Calin, C. G. Liu et al., "A microRNA expression signature of human solid tumors defines cancer gene targets," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 7, pp. 2257–2261, 2006.
- [17] G. A. Calin, M. Ferracin, A. Cimmino et al., "A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia," *New England Journal of Medicine*, vol. 353, no. 17, pp. 1793–1801, 2005.
- [18] A. Israel, R. Sharan, E. Ruppman, and E. Galun, "Increased microRNA activity in human cancers," *PLoS One*, vol. 4, no. 6, Article ID e6045, 2009.
- [19] S. Subramanian, W. O. Lui, C. H. Lee et al., "MicroRNA expression signature of human sarcomas," *Oncogene*, vol. 27, no. 14, pp. 2015–2026, 2008.
- [20] J. C. Wittig, J. Bickels, D. Priebat et al., "Osteosarcoma: a multidisciplinary approach to diagnosis and treatment," *American Family Physician*, vol. 65, no. 6, pp. 1123–1136, 2002.
- [21] A. Dutour, D. Leclers, J. Monteil et al., "Non-invasive imaging correlates with histological and molecular characteristics of an osteosarcoma model: application for early detection and follow-up of MDR phenotype," *Anticancer Research*, vol. 27, no. 6B, pp. 4171–4178, 2007.
- [22] A. Dutour, J. Monteil, F. Paraf et al., "Endostatin cDNA/cationic liposome complexes as a promising therapy to prevent lung metastases in osteosarcoma: study in a human-like rat orthotopic tumor," *Molecular Therapy*, vol. 11, no. 2, pp. 311–319, 2005.
- [23] M. Allouche, H. G. Delbruck, and B. Klein, "Malignant bone tumours induced by a local injection of colloidal radioactive ¹⁴⁴Cerium in rats as a model for human osteosarcomas," *International Journal of Cancer*, vol. 26, no. 6, pp. 777–782, 1980.
- [24] K. Arakawa, N. Kono, Y. Yamada, H. Mori, and M. Tomita, "KEGG-based pathway visualization tool for complex omics data," *In Silico Biology*, vol. 5, no. 4, pp. 419–423, 2005.
- [25] A. Gougelet, D. Pissaloux, A. Besse et al., "miRNA profiles in osteosarcoma as a predictive tool for ifosfamide chemotherapeutic treatment response," *International Journal of Cancer*. In press.
- [26] R. E. Brown, "Morphoproteomic portrait of the mTOR pathway in mesenchymal chondrosarcoma," *Annals of Clinical and Laboratory Science*, vol. 34, no. 4, pp. 397–399, 2004.
- [27] Q. Zhou, Z. Deng, Y. Zhu, H. Long, S. Zhang, and J. Zhao, "mTOR/p70S6K signal transduction pathway contributes to osteosarcoma progression and patients' prognosis," *Medical Oncology*, vol. 27, no. 4, pp. 1239–1245, 2010.
- [28] B. H. Jiang and L. Z. Liu, "Role of mTOR in anticancer drug resistance: perspectives for improved drug treatment," *Drug Resistance Updates*, vol. 11, no. 3, pp. 63–76, 2008.
- [29] X. Wan, A. Mendoza, C. Khanna, and L. J. Helman, "Rapamycin inhibits ezrin-mediated metastatic behavior in a murine model of osteosarcoma," *Cancer Research*, vol. 65, no. 6, pp. 2406–2411, 2005.
- [30] I. Beuvink, A. Boulay, S. Fumagalli et al., "The mTOR inhibitor RAD001 sensitizes tumor cells to DNA-damaged induced apoptosis through inhibition of p21 translation," *Cell*, vol. 120, no. 6, pp. 747–759, 2005.
- [31] S. Mabuchi, D. A. Altomare, D. C. Connolly et al., "RAD001 (Everolimus) delays tumor onset and progression in a transgenic mouse model of ovarian cancer," *Cancer Research*, vol. 67, no. 6, pp. 2408–2413, 2007.
- [32] S. Ambbs, R. L. Prueitt, M. Yi et al., "Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer," *Cancer Research*, vol. 68, no. 15, pp. 6162–6170, 2008.
- [33] G. A. Calin, M. Ferracin, A. Cimmino et al., "A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia," *New England Journal of Medicine*, vol. 353, no. 17, pp. 1793–1801, 2005.
- [34] A. J. Schetter, S. Y. Leung, J. J. Sohn et al., "MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma," *Journal of the American Medical Association*, vol. 299, no. 4, pp. 425–436, 2008.

- [35] S. L. Yu, H. Y. Chen, G. C. Chang et al., "MicroRNA signature predicts survival and relapse in lung cancer," *Cancer Cell*, vol. 13, no. 1, pp. 48–57, 2008.
- [36] B. Song, Y. Wang, Y. Xi et al., "Mechanism of chemoresistance mediated by miR-140 in human osteosarcoma and colon cancer cells," *Oncogene*, vol. 28, no. 46, pp. 4065–4074, 2009.
- [37] R. Taulli, F. Bersani, V. Foglizzo et al., "The muscle-specific microRNA miR-206 blocks human rhabdomyosarcoma growth in xenotransplanted mice by promoting myogenic differentiation," *Journal of Clinical Investigation*, vol. 119, no. 8, pp. 2366–2378, 2009.
- [38] R. N. Veedu and J. Wengel, "Locked nucleic acids: promising nucleic acid analogs for therapeutic applications," *Chemistry and Biodiversity*, vol. 7, no. 3, pp. 536–542, 2010.
- [39] J. A. Fletcher, M. C. Gebhardt, and H. P. Kozakewich, "Cytogenetic aberrations in osteosarcomas: nonrandom deletions, rings, and double-minute chromosomes," *Cancer Genetics and Cytogenetics*, vol. 77, no. 1, pp. 81–88, 1994.
- [40] L. Han, P. D. Witmer, E. Casey, D. Valle, and S. Sukumar, "DNA methylation regulates microRNA expression," *Cancer Biology and Therapy*, vol. 6, no. 8, pp. 1284–1288, 2007.
- [41] Y. Saito and P. A. Jones, "Epigenetic activation of tumor suppressor microRNAs in human cancer cells," *Cell Cycle*, vol. 5, no. 19, pp. 2220–2222, 2006.
- [42] I. Bray, K. Bryan, S. Prenter et al., "Widespread dysregulation of MiRNAs by MYCN amplification and chromosomal imbalances in neuroblastoma: association of miRNA expression with survival," *PLoS One*, vol. 4, no. 11, Article ID e7850, 2009.
- [43] S. Mi, Z. Li, P. Chen et al., "Aberrant overexpression and function of the miR-17-92 cluster in MLL-rearranged acute leukemia," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 8, pp. 3710–3715, 2010.
- [44] S. Selvarajah, M. Yoshimoto, O. Ludkovski et al., "Genomic signatures of chromosomal instability and osteosarcoma progression detected by high resolution array CGH and interphase FISH," *Cytogenetic and Genome Research*, vol. 122, no. 1, pp. 5–15, 2008.
- [45] M. Wachtel, M. Dettling, E. Koscielniak et al., "Gene expression signatures identify rhabdomyosarcoma subtypes and detect a novel t(2;2)(q35;p23) translocation fusing PAX3 to NCOA1," *Cancer Research*, vol. 64, no. 16, pp. 5539–5545, 2004.
- [46] E. Davicioni, F. G. Finckenstein, V. Shahbazian, J. D. Buckley, T. J. Triche, and M. J. Anderson, "Identification of a PAX-FKHR gene expression signature that defines molecular classes and determines the prognosis of alveolar rhabdomyosarcomas," *Cancer Research*, vol. 66, no. 14, pp. 6936–6946, 2006.
- [47] A. B. Hui, W. Shi, P. C. Boutros et al., "Robust global microRNA profiling with formalin-fixed paraffin-embedded breast cancer tissues," *Laboratory Investigation*, vol. 89, no. 5, pp. 597–606, 2009.
- [48] K. R. M. Leite, J. M. S. Canavez, S. T. Reis et al., "miRNA analysis of prostate cancer by quantitative real time PCR: comparison between formalin-fixed paraffin embedded and fresh-frozen tissue," *Urologic Oncology: Seminars and Original Investigations*. In press.
- [49] A. Liu, M. T. Tetzlaff, P. VanBelle et al., "MicroRNA expression profiling outperforms mRNA expression profiling in formalin-fixed paraffin-embedded tissues," *International Journal of Clinical and Experimental Pathology*, vol. 2, no. 6, pp. 519–527, 2009.
- [50] X. Chen, Y. Ba, L. Ma et al., "Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases," *Cell Research*, vol. 18, no. 10, pp. 997–1006, 2008.
- [51] P. S. Mitchell, R. K. Parkin, E. M. Kroh et al., "Circulating microRNAs as stable blood-based markers for cancer detection," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 30, pp. 10513–10518, 2008.

3. Caractérisation génomique des rhabdomyosarcomes à transcrits rares

Nous avons montré qu'il était possible d'affiner la classification des RMS à l'aide des profils microARN. A la suite de ce travail, nous avons entrepris de réaliser une caractérisation moléculaire exhaustive de tous les cas ambigus de RMS que nous avons rencontré au cours de notre activité diagnostique, c'est-à-dire des cas situés dans la zone grise entre ARMS et ERMS. Ce travail de caractérisation génomique est toujours en cours, et a pour but d'identifier d'éventuels ARMS à transcrits rares ou de vrais nARMS, afin de mieux comprendre leur biologie. A cette fin, un tri initial de notre cohorte de rhabdomyosarcomes nous a permis d'extraire 47 cas difficiles, sur la base d'un code ADICAP additionnel, propre à notre laboratoire, possédant la signification suivante : « rhabdomyosarcome indifférencié difficile à classer ». Selon nos critères, sont regroupés sous ce code les RMS difficiles de morphologie indifférenciée, les formes mixtes, et les ARMS de morphologie plus classique, mais pour lesquels aucun transcrit courant n'a été détecté par technique FISH et RT-PCRq.

3.1. Recherche de cas rares grâce aux profils d'expression des microARN

A la suite de la publication de notre travail sur les RMS dans la revue *Sarcoma*, nous avons poursuivi nos investigations sur un mode prospectif, en passant en TLDA tous les cas diagnostiques pour lesquels nos médecins anatomopathologistes prescrivaient une demande de recherche des transcrits de fusion PAX3-FOXO1 et PAX7-FOXO1 par RT-PCRq, et pour lesquels il restait du matériel biologique extrait après rendu du résultat complémentaire. Chaque nouveau cas est injecté dans une ACP supervisée, à la recherche d'un comportement inattendu pour l'un des échantillons.

A la date d'aujourd'hui, cette approche nous a permis de détecter la présence de deux cas potentiels d'ARMS sans transcrits ou d'ARMS à transcrits rares, mais dans les deux cas, il n'a pas été possible de poursuivre nos investigations, car le matériel tumoral était épuisé. Cependant, de manière inattendue, les profils d'expression des microARN nous ont permis de détecter que trois tumeurs de type ERMS classique sur la morphologie, ne se comportaient pas du tout comme les autres ERMS de notre cohorte (figure 37).

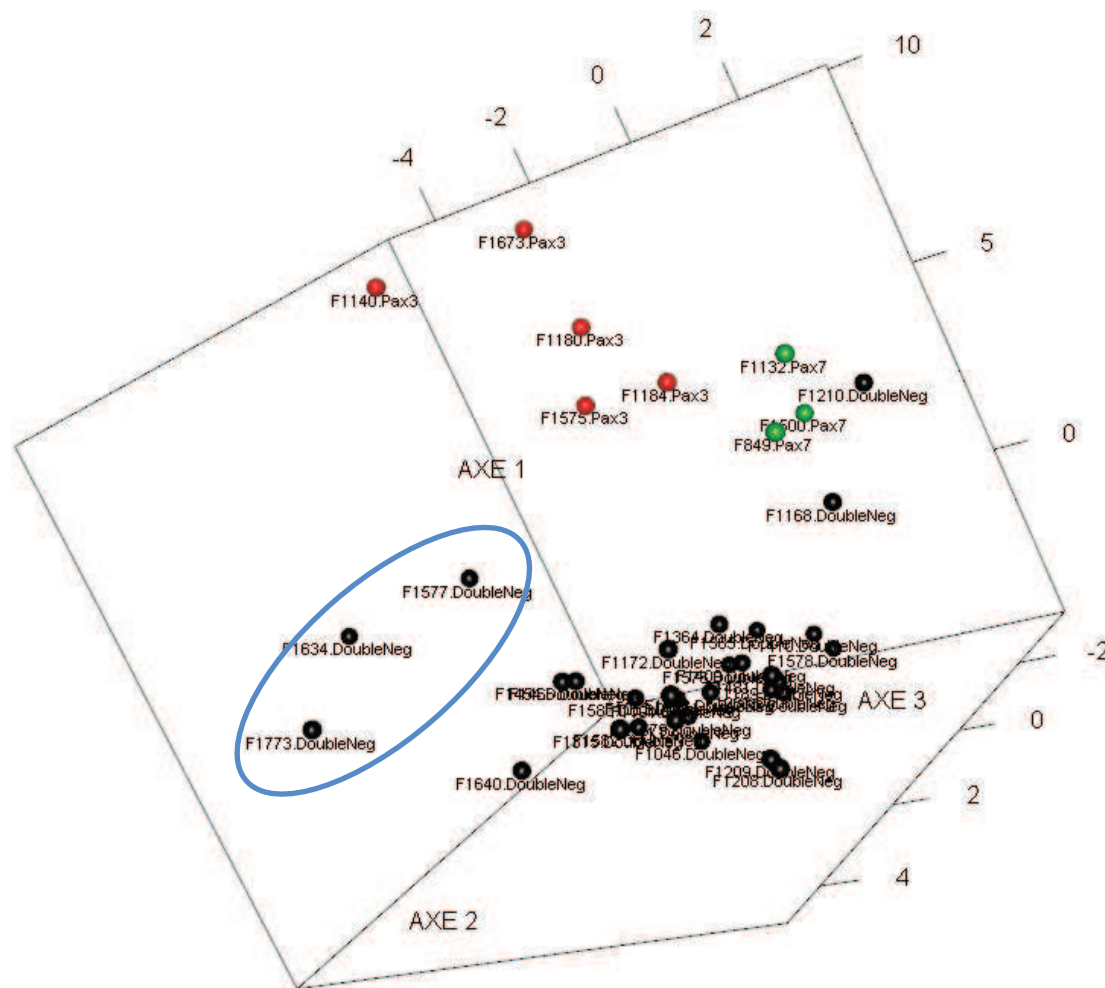


Figure 37 : ACP supervisée de 41 RMS pour lesquels les profils d'expression des microARN ont été établis sur la plateforme TLDA.

Les pARMS de type PAX3-FOXO1 sont représentés en rouge, les pARMS de type PAX7-FOXO1 sont représentés en vert, et tous les autres cas sont représentés en noir (principalement ERMS et RMS difficiles de la zone grise). Trois cas diagnostiqués comme des ERMS sur la morphologie se comportent très différemment des autres ERMS, ils sont entourés en bleu (Analyse réalisée par H. Philippon au centre Léon Bérard).

Nous avons déjà eu l'occasion d'approfondir l'un de ces cas particuliers (échantillon F1773). Par technique FISH, nous n'avons pas trouvé de réarrangement du gène FOXO1. Nous avons alors effectué la recherche de réarrangement du gène PAX3 par technique FISH, qui s'est avérée être positive. Ce cas a ensuite été caractérisé sur le plan génomique par technique CGH (voir partie 3.3. Etude de cas : RMS embryonnaire à transcrit rare).

Nous projetons également d'effectuer en rétrospectif les profils d'expression des microARN des 47 cas de la zone grise entre ARMS et ERMS, si le matériel tumoral restant le permet.

3.2. Recherche des transcrits rares par technique FISH et PCR

Parmi les 47 RMS classés comme difficiles, 38 tumeurs avaient déjà pu bénéficier d'une technique FISH à l'aide de la sonde FOXO1 de type *break-apart*. Parmi ces 38 tumeurs, nous avons obtenus 3 résultats non contributifs, 34 cas pour lesquels le gène FOXO1 n'était pas réarrangé, et 1 cas avec la présence d'un réarrangement du gène FOXO1.

En ce qui concerne l'unique cas possédant un réarrangement du gène FOXO1, une recherche des transcrits de fusion PAX3-FOXO1 et PAX7-FOXO1 par RT-PCRq à l'Institut Bergonié a retrouvé la présence du transcrit de fusion le plus courant, PAX3-FOXO1. Malgré la découverte d'un transcrit courant, ce cas est intéressant car sa morphologie est tout à fait inhabituelle pour un rhabdomyosarcome alvéolaire. En effet, notre médecin anatomopathologiste l'a classé dans une variante embryonnaire à cellules fusiformes, et lui a secondairement attribué notre code ADICAP interne additionnel correspondant aux entités difficiles, en raison de la discordance observée entre la morphologie et la biologie moléculaire. Ce cas nous ayant été adressé pour avis sous forme de coupes colorées HES et de lames blanches, nous n'avons pas poursuivi ces investigations, même s'il aurait été intéressant de connaître le résultat

de sa classification sur la base des profils d'expression des microARN, et de ses anomalies génomiques.

Parmi les 34 cas pour lesquels nous n'avons pas observés de fission du gène FOXO1 par technique FISH, 31 blocs ou lames blanches étaient encore disponibles dans notre laboratoire. Nous avons donc pu rechercher sur ces cas la présence d'éventuels transcrits rares, à l'aide de la sonde PAX3 de type *break-apart*. En effet, FOXO1 n'étant pas réarrangé, cela exclut les deux translocations les plus courantes ; si pour ces mêmes cas le gène PAX3 est réarrangé, cela signifie que nous avons forcément affaire à un transcrit rare de type PAX3-NCOA1, PAX3-NCOA2 ou PAX3-FOXO4. Nous n'avons pas été en mesure d'obtenir une qualité d'hybridation satisfaisante pour 13 cas, et 14 cas ne possédaient pas de réarrangement du gène PAX3. En revanche, il nous a été possible de montrer un réarrangement du gène PAX3 dans 4 cas.

Au final, sur les 47 cas initiaux, grâce à la technique de FISH, nous avons pu mettre en évidence la présence d'une translocation rare dans seulement 4 cas, soit environ 8 % des RMS difficiles. Malheureusement, sur ces 4 cas, du matériel biologique n'est exploitable que pour 1 seul d'entre eux, et nous n'avons pas encore réussi à obtenir des quantités de matériel acceptables après extraction (deux blocs sont actuellement en cours de rappel à l'extérieur).

3.3. Caractérisation génomique par technique CGH sur puce

Sur la base des profils d'expression des microARN, nous avons vu que certains RMS de diagnostic difficile ou considérés comme des nARMS se comportaient différemment des ERMS et des ARMS à transcrits courants. Il a cependant été difficile de rechercher les éventuels transcrits rares de type PAX3-NCOA1, PAX3-NCOA2 et PAX3-FOXO4 sur ces mêmes échantillons. En effet, pour des raisons pratiques et techniques, il a été impossible d'obtenir des résultats complémentaires,

que ce soit par technique PCRq, FISH ou CGH : la paucité des prélèvements et les techniques non contributives ont conduit à ce qu'aucun des échantillons utilisés dans l'étude publiée dans la revue *Sarcoma* ne soit utilisable.

En parallèle à l'étude des profils d'expression des microARN précédemment présentée, ainsi qu'à la caractérisation des transcrits rares par techniques FISH et PCRq, nous avons également essayé de séparer les ARMS des ERMS sur la base de leurs profils génomiques, en utilisant une approche pan-génomique de type *array comparative genomic hybridization* (aCGH). Pour l'instant, seuls deux cas ont été analysés. Le premier cas est l'ERMS à transcrit rare identifié initialement à partir des profils d'expression des microARN, et le deuxième cas est un ARMS évident sur la morphologie, mais placé dans la zone grise puisqu'il ne possédait pas de réarrangement du gène FOXO1 (la technique FISH concernant la détermination du statut du gène PAX3 s'est avérée non contributive).

- Etude de cas : RMS embryonnaire à transcrit rare

Nous avons vu précédemment que sur la base des profils d'expression des microARNs, un cas initialement classé en ERMS par les pathologistes, car de morphologie évidente, semblait se comporter comme un ARMS à transcrit rare. Par technique FISH, il a été retrouvé un transcrit rare impliquant le gène PAX3. Il s'agit donc très certainement d'un cas particulier très rare d'ERMS avec réarrangement du gène PAX3, sans que le gène partenaire soit identifié, tel que l'unique cas rapporté dans la littérature et décrit par l'équipe de Sawada (Hosoi et al., 2009).

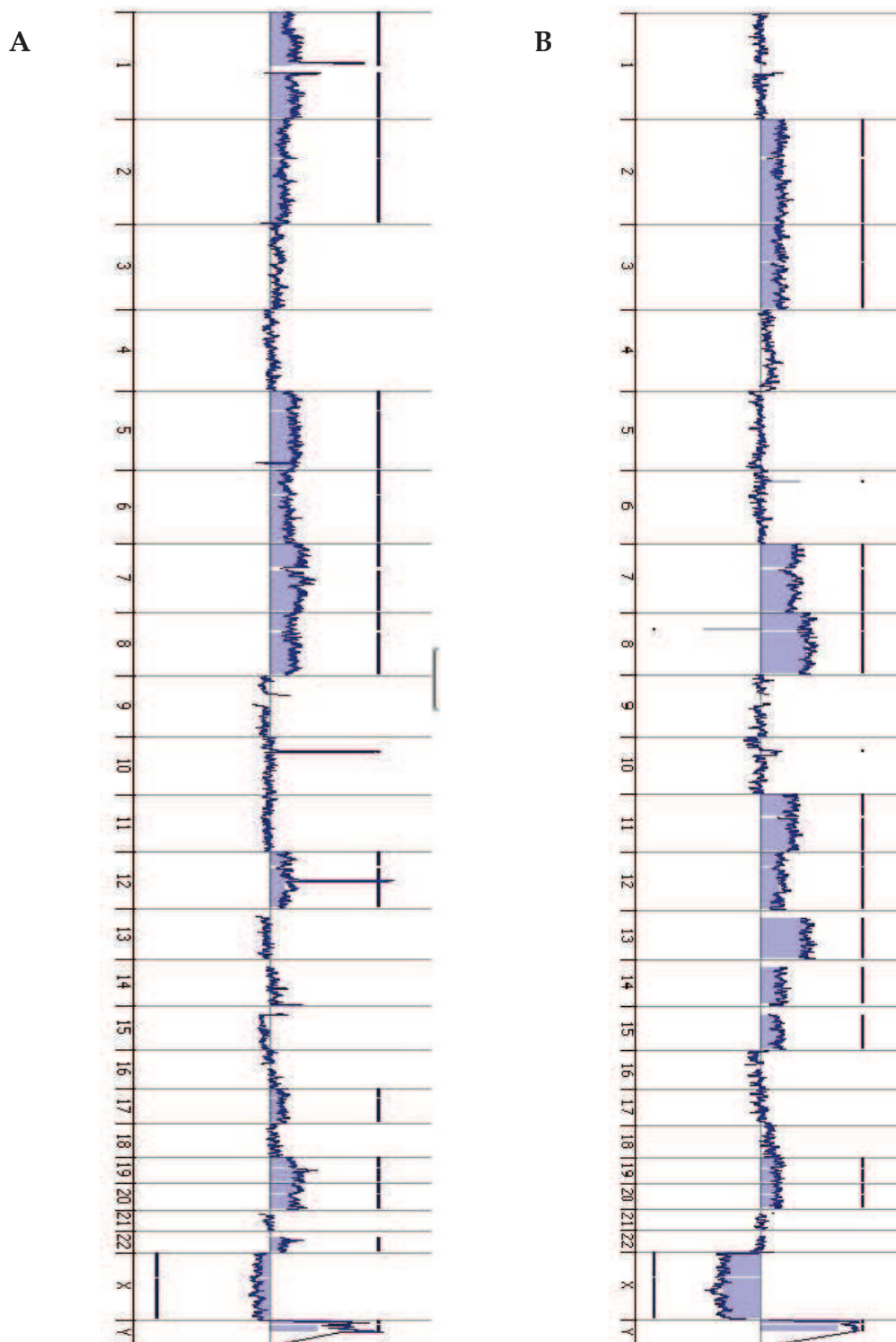


Figure 38 : profils génomiques aCGH obtenus sur plateforme Agilent.

(A) Profil génomique d'un ERMS paratesticulaire plus ou moins différencié avec composante anaplasique focale, présentant une amplification des gènes NOTCH2 et MDM2. (B) Profil génomique d'un ERMS de morphologie classique, mais présentant un transcrit rare impliquant le gène PAX3 (profils génomiques réalisés par S. Paindavoine, au Centre Léon Bérard).

Le profil aCGH obtenu pour cette tumeur, présentant une morphologie classique d'ERMS et un transcrit rare impliquant le gène PAX3, est compatible avec un profil génomique d'ERMS (figure 38). On observe de multiples gains chromosomiques, dont un gain de la totalité du chromosome 8, jamais observé par l'équipe du Dr. Olivier Delattre sur leur cohorte de 50 pARMS à transcrits classiques (Williamson et al., 2010). Aucun point de cassure n'est détectable au niveau des loci des gènes PAX3, PAX7, NCOA1, NCOA2 et FOXO4.

- Etude de cas : RMS alvéolaire à transcrit rare

Parmi les 11 cas de RMS classés en nARMS sur la base d'une morphologie plutôt classique d'ARMS, mais pour lesquels nous n'avions pas retrouvé les transcrits courants PAX3-FOXO1 et PAX7-FOXO1 par technique FISH et RT-PCRq, nous avons eu l'opportunité d'en tester un en technique aCGH, sur matériel fixé en paraffine (figure 39).

Le profil génomique obtenu est compatible avec celui d'un pARMS (Williamson et al., 2010). En effet, on observe de multiples zones d'amplification sur les chromosomes 1 et 2, dont une amplification du gène MDM4 (1q32.1) et du cluster de gènes miR-199A2 / miR-214 (1q24.3). Il existe également une perte du bras long du chromosome 13, impliquant une délétion du gène suppresseur de tumeur RB1 (13q14). En revanche, il n'est pas retrouvé de gain des chromosomes 8 et 11. De manière intéressante, en plus d'une délétion du gène RB1, nous avons trouvé deux altérations génomiques de type amplification ayant pu contribuer à la tumorigenèse de cette tumeur. Il a déjà été montré dans un modèle tumoral de souris que l'amplification du gène MDM4, un homologue du gène MDM2, contribue directement à la tumorigenèse par inhibition de l'activité du gène suppresseur de tumeur P53 (Danovi et al., 2004). Le rôle du cluster de gènes miR-199A2 / miR-214 est plus controversé, mais il est décrit comme amplifié ou surexprimé dans de nombreuses tumeurs (Yin et al., 2010; Tap et al., 2011).

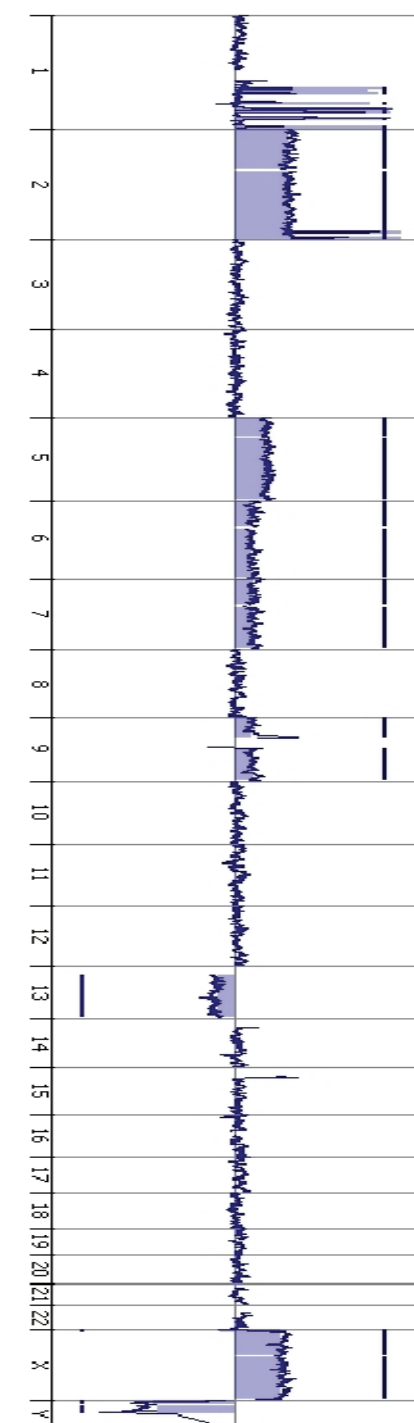


Figure 39 : profil génomique aCGH obtenu sur plateforme Agilent.

Cas d'un RMS pédiatrique classé en ARMS évident sur la morphologie, mais pour lequel il n'a pas été retrouvé les transcrits de fusion courants. Présence de multiples zones d'amplification sur les chromosomes 1 et 2 (profil génomique réalisé par S. Paindavoine, au Centre Léon Bérard).

Malheureusement, la technique FISH à l'aide de la sonde PAX3 break-apart s'est révélée être non contributive. A ce stade, nous ne pouvons donc pas savoir s'il s'agit d'un cas de pARMS possédant un transcrit rare impliquant le gène PAX3, d'un pARMS à transcrit rare non répertorié, ou d'un véritable nARMS.

Comme nous l'avons vu à travers l'étude de ces quelques cas difficiles, la frontière entre ARMS et ERMS reste très difficile à établir. Dans une situation idéale où tous les transcrits rares seraient détectés par des techniques de routine, il ne serait *a priori* pas possible de trancher entre ARMS et ERMS sur ce seul résultat, puisque nous avons montré, par exemple, qu'il existait des ERMS à morphologie et à génomique classiques présentant des transcrits rares. Nous avons prouvé que la morphologie et la biologie moléculaire donnaient parfois des résultats inattendus.

Afin de surmonter cette difficulté, il convient dans un premier temps d'étoffer ces résultats morphologiques et moléculaires sur un nombre de cas plus important, puis de les croiser avec les données cliniques concernant le pronostic et la réponse au traitement de ces formes très originales. En effet, pour les cas où il existe une discordance pour l'instant inexplicable entre la morphologie et la biologie moléculaire, seules les données cliniques de suivi des patients concernés permettront de trancher sur la meilleure façon de les prendre en charge. Ce travail sera particulièrement difficile car ces entités sont rares, mais le bénéfice que nous pourrions apporter aux patients serait indéniable. D'autre part, ces résultats permettront d'affiner la classification nosologique des RMS en permettant une meilleure compréhension de leur biologie et des voies métaboliques dérégulées.

PUBLICATION ANNEXE

Autocrine role for Gas6 with Tyro3 and Axl in leiomyosarcomas

Les léiomyosarcomes (LMS) sont des tumeurs musculaires lisses malignes représentant environ 15 % des sarcomes de l'adulte. Ces tumeurs sont principalement retrouvées au niveau du rétropéritoine et de la cavité abdominale, de l'utérus, des membres et des vaisseaux sanguins. Ce sont des tumeurs agressives de mauvais pronostic, avec un taux de survie à 5 ans de 50 à 55 % et un taux de métastases d'environ 40 % (Enzinger and Weiss, 1995). Les LMS sont généralement traités par exérèse chirurgicale complète et radiothérapie. Les LMS plus avancés ou métastatiques répondent à la chimiothérapie, mais de manière limitée.

Les LMS sont des sarcomes à génomique complexe, caractérisés par de nombreux remaniements chromosomiques (Helman and Meltzer, 2003). Aujourd'hui, aucun marqueur spécifique n'est connu, et il n'existe aucune stratégie thérapeutique efficace. Le but de cette étude était donc d'identifier des cibles potentielles par l'analyse des voies métaboliques altérées dans ces tumeurs, à partir de l'établissement des profils d'expression protéique (puce à anticorps orientée sur la détection des protéines kinases).

Les profils d'expression protéique de 13 LMS ont été comparés au profil d'un échantillon de muscle lisse utérin sain, et quatre protéines ont été retrouvées différemment exprimées. Cette étude a permis d'identifier Tyro3, Axl et Gas6 comme des acteurs d'une boucle autocrine activée dans un sous-groupe de LMS. Ces observations pourraient avoir des conséquences importantes, à la fois pour la compréhension de la biologie de ces tumeurs, mais aussi pour des applications thérapeutiques pour ce groupe de patients où aucune cible thérapeutique n'a été encore identifiée.

Document à suivre :

Article IV en soumission : *Autocrine role for Gas6, Tyro3 and Axl in leiomyosarcomas.*

Title: Autocrine role for Gas6 with Tyro3 and Axl in leiomyosarcomas

Keywords: Leiomyosarcomas, tyrosine kinase, Tyro3, Gas6, Targeted therapy, Sarcoma, FAK.

Authors: Hiba el Sayadi^{1*}, Daniel Pissaloux^{1*}, Laurent Alberti¹, Severine Tabone- Eglinger¹, Dominique Ranchere², Anne Valérie Decouvelaere², Eric Tabone², Isabelle Ray-Coquard², Christophe Caux¹, Jérôme Fayette⁴, Jean-Yves Blay^{1,3,4}.

¹ Pôle des Sciences Cliniques & Centre de Recherche en Cancerologie de Lyon, INSERM UMR5286 Equipe 11, Centre Léon Bérard, Lyon, France.

² Service d'anatomie pathologique, Centre Léon Bérard, Lyon, France.

³ Université Claude Bernard Lyon 1, Lyon, France.

⁴ Département de Médecine Centre Léon Bérard, Lyon, France.

* Both authors contributed equally to this work.

Running title: Autocrine loop with Tyro3 and Gas6 in leiomyosarcomas.

Grant support: Hiba El Sayadi received a grant from CONTICANET Network of Excellence FP6-018806. This work was supported in part by grants from Association pour la Recherche sur le Cancer, Ligue contre le Cancer (Comité de l'Ain et Comité du Rhône), the Breast Cancer Research Foundation, Cancéropôle Lyon Auvergne Rhône Alpes, and support from the French National Cancer Institute, through the Network Netsarc grant, and the LYRIC grant. The research leading to these results has received funding from the European Union Seventh Framework Programme - EuroSarc - (FP7/2007-2013) under Grant Agreement n°278742.

Acknowledgements: We thank Pharmamar for providing the CNIO-AA cell line (Dr M. Aracil). Foretinib was kindly provided by Glaxo Smith Kline (Dr T. Gilmer).

Abstract:***Rationale:***

Leiomyosarcoma (LMS) represent 15% of adult sarcomas. The aim of this work was to identify novel altered pathways in LMS which may be of therapeutic value for patients.

Material and methods

13 fresh frozen samples of soft tissue and visceral LMS were analyzed and compared with normal smooth muscle uterine tissue (NSM) for phosphoproteomic profile. Four proteins were found differentially expressed including Tyro3. The functional role of Tyro3 and its ligand Gas6 was investigated in 2 LMS cell lines, SK-LMS-1 and CNIO-AA.

Results:

Four proteins and phosphoproteins were differentially expressed in LMS samples vs NSM: a loss of FAK Y397 phosphorylation was observed in all LMSs while Tyro3, MSH2, and PKC theta were consistently overexpressed. Gas6, the major ligand of Tyro3 was expressed in 8 of the 13 LMS samples, and Gas6 expression highly correlated to Akt Y473 phosphorylation and to a lesser extent to Erk1/2 phosphorylation. SK-LMS-1 and CNIO-AA LMS expressed Tyro3, Axl and Gas6 at high level in CNIO-AA, while at low levels in SK-LMS-1. Exposure of both cell lines to foretinib, a tyrosine kinase inhibitor of Met, Axl and Tyro3, reduced cell viability and induced caspase 3/7 activation. Transfection of CNIO-AA with siRNA directed against Tyro3 and Axl genes induced a reduction of the expression of the specific proteins, and when combined, significantly reduced CNIO-AA cell viability.

Conclusion:

Leiomyosarcomas overexpress Tyro3. Gas6, a ligand of Tyro3, exerts an autocrine activities though Tyro3 and Axl in a subgroup of LMS.

Introduction:

Leiomyosarcomas (LMS) account for 15% of all adult soft tissue and visceral sarcomas (1-4). Primary sites include limbs, trunk, retroperitoneum, and visceral uterus (3). Despite heterogeneous clinical presentations and biology, the management of LMS is based on a complete surgical excision of the tumour followed or preceded, for most localisations, by radiotherapy on the tumour bed (5). In advanced or metastatic LMS, doxorubicin, trabectedine (8), gemcitabine and anti-angiogenic agents have a significant yet limited antitumour activity (8-10). Leiomyosarcomas exhibit complex and heterogeneous genomic alterations, with probably different nosological entities. It would be useful to identify potential “driver” genes in these tumours, as KIT for GIST (11-17).

The objective of the present work was to investigate the presence of abnormal proteomic patterns in leiomyosarcomas cell lines and fresh frozen tumour samples to screen for possible druggable targets. We report here that Tyro3 and its ligand Gas6 are expressed and activated in LMS cell lines and fresh tissues, and contribute to LMS cell survival and proliferation.

Materials and Methods:

Cell lines and tumour samples:

The SK-LMS-1 uterine leiomyosarcoma cell line was purchased from ATCC (Manassas, VA). The CNIO-AA cell line, derived from a soft tissue leiomyosarcoma, was obtained from Pharmamar (Tres Cantos, Spain). Thirteen fresh frozen samples of LMS and normal smooth muscle (NSM) uterine tissue sample were collected from the Centre Léon Bérard and the ConticaBase tumour bank (conticabase.org). All patients gave a written consent to research use of biological material according to the French laws.

Protein preparation for Kinex™ antibody microarrays:

Tumours and normal tissues were grinded in liquid nitrogen and suspended in lysis buffer. The lysates were then centrifuged at 13000 rpm at 4°C and the amount of proteins in the supernatant fraction was quantified using the Bio-Rad Dc Protein Assay.

Kinex™ antibody microarray analysis:

The experiment was performed by Kinexus Bioinformatics Corporation (Vancouver, BC, Canada). Briefly, tumour or normal tissue proteins were tagged with a fluorescent dye, then incubated with the antibody microarray; target proteins were captured by specific antibodies immobilized on the microarray, and unbound proteins were washed away. Captured proteins were quantified according to the amount of dye signal within each spot. For this study, Kinex™ Service used the KAM-1.1 chip containing over 650 antibodies each spotted in duplicate. Antibodies are directed against 240 protein kinases, 28 phosphatases and 90 other cell signalling proteins that regulate cell proliferation, stress and apoptosis, for a total of 630 proteins tested.

Western blot analysis:

Lysates of tumour cells were rocked for 30 min at 4°C, and centrifuged to remove insoluble material. Supernatant protein contents were determined using Bio-Rad Dc Protein Assay (500-0113-0114-0115). The normalized soluble proteins (supernatant fraction) were directly analyzed by SDS-PAGE Western blotting with antibodies of interest.

Analysis of Tyro3 and Gas6 interaction:

The analysis of Tyro3 and Gas6 in the CNIO-AA cell lines was performed using the Duolink assay (Uppsala, Sweden). Briefly, the two primary antibodies directed against Tyro3 and Gas6 recognize the two antigens of interest in the cells and are recognized by two species-specific secondary antibodies (provided in the Duolink kit) each labelled with a different short DNA strand attached to it. When the secondary antibodies are in close proximity, the DNA strands can interact through the addition of two other circle-forming DNA oligonucleotides, which are then joined by enzymatic ligation. Upon addition of the amplification solution (containing nucleotides and polymerase), several 100-fold replication of the DNA circle occur, and fluorescent probes highlight the product. The resulting high concentration of fluorescence in each single-molecule amplification product is visible as a bright spot in a fluorescence microscope. Nuclei are counterstained with Dapi (blue).

siRNA transfections:

The siRNAs examined in this study were synthesized by Dharmacon, Inc (Lafayette, CO, USA): siGENOME smart pool - Human Axl: M-003104-03-0010, Human TYRO3: M-003183-02-0010, Human GAS6: M-009069-00-0010. Cells were transfected with the DharmaFECT 2 transfection reagent according to the manufacturer's instructions. To determine the efficiency of gene silencing, the whole cell lysates of the co-transfected cells was used for western blotting as described above and probed using specific antibodies for Gas6, Tyro3 and Axl.

Viability assay:

SK-LMS-1 and CNIO-AA leiomyosarcomas cells were plated in duplicate in 96-well plates (625 cells/well and 2000 cells/well respectively). After 24h, the media was replaced by media containing either non-targeting siRNA, siRNA directed against Gas6, Tyro3, or Axl (Dharmacon), alone, in paired combinations or all together for 24, 48, and 72h. Relative numbers of viable cells were measured in comparison to the untreated control and the siRNA control (scramble) using the Cell Titer-Glo® Luminescent Cell Viability Assay (Promega, Charbonnières, France) according to the manufacturer's instructions. Caspase 3/7 assays were performed with the Promega luminescent assay caspase-Glo 3/7 luminescent reagent (Promega, Charbonnières, France).

Sources of antibodies and tyrosine kinase inhibitors:

The following antibodies were used: human Gas6 mouse monoclonal antibody, human Dtk (Tyro3) goat polyclonal antibody, phosphor-Y799-Axl rabbit polyclonal antibody, human Axl goat polyclonal antibody were obtained from R&D systems (France). Phospho-Y397-FAK rabbit polyclonal antibody, FAK mouse monoclonal [12G4] antibody from Abcam (France), phospho-S473-Akt rabbit monoclonal antibody from Epitomics (France), Akt rabbit polyclonal antibody, P44/42 MAP kinase antibody, Phospho-p44/42 Thr 202/Tyr 204 MAP Kinase antibody, PTEN (138G6) rabbit monoclonal antibody were obtained from Cell signaling (France), β -actine: mouse monoclonal antibody from Sigma (France), GAPDH mouse antibody from Biodesign (France). Foretinib, a tyrosine kinase inhibitor of VEGFR2, Met, Axl and Tyro 3 was provided by Glaxo Smith Kline (GSK, North Carolina, USA).

Statistical analysis:

For the statistical analysis of proteomic arrays, Excel® spreadsheets containing raw data, i.e. spot signal and local background intensities were provided by Kinexus, as results for the hybridization experiments and for each microarray. Duplicate values were averaged, and samples were normalized by the mean, first in an intra-slide way (by pairs) to enable direct comparison between samples, then in an inter-slide way to compare data derived from samples that had been assayed on different slides within the same Kinexus batch. The raw intensity data coming from the samples (LMS and NSM as a control) were normalized on an Excel sheet according to Kinexus recommendations. Data of insufficient quality were then removed, upon manufacturer's recommendations by applying the following filters: 1) the normalized intensities should be above 500 in at least one sample. 2) The signal/noise ratio should be above 1.5. 3) In addition, mean values resulting from non consistent duplicates were discarded when the error range was above 15%. 4) Finally, to keep only the most discriminant antibodies, Log₂ (ratio) values between samples were calculated and the antibodies that had a two-fold change ($|\text{Log}_2(\text{ratio})| > 1$) at least one time were selected. An unsupervised clustering was then performed on the remaining 192 proteins selected, after a z-score transformation. Viability rates between samples were compared with Student T test. Protein expression levels between subgroups were compared using a chi square or Fisher's exact test.

Results:

Protein expression of leiomyosarcoma vs normal uterine smooth muscle:

Four (phospho) proteins were most differentially expressed between normal smooth muscle (NSM) and the 13 fresh leiomyosarcoma tumour samples tested. All leiomyosarcoma fresh tissues showed a down regulation of the phosphorylation of FAK on tyrosine 397 (Figure 1). The 3 most overexpressed proteins were Tyro-3, PKC θ , and MSH2 (Figure 1). PKC θ and MSH2 expression will not be further discussed in this paper.

Tyro-3, Axl and Gas6 expression in léiomyosarcome:

Tyro3 is a transmembrane tyrosine kinase receptors, structurally related to Axl and Mer (19). The expression of Gas6, the ligand of Tyro3, and Axl, was investigated in LMS (19). In Western blot, the 140 kDa and 120 kD isoforms of Tyro3 were found expressed at various levels, the latter being detectable only in 8 of the 13 LMS tested (Figure 2). Tyro3 was also detectable at lower level in NSM. No anti phospho-Tyro3 antibody was available for Western blot analysis. Axl was found expressed and phosphorylated in 12 of the 13 samples, at markedly variable levels (Figure 2, Table 1).

Gas6 was found expressed in 8 of the 13 LMS, and in NSM (Figure 2). Interestingly, phosphorylation of Akt on S473 was low or non detectable in 5 of the 5 tumours with low or undetectable Gas6 expression while all 8 tumours with high levels of Gas6 had high levels of P-Akt on S473 ($p=0.0003$, Figure 2, Table 1). High levels of Erk1/2 phosphorylation were observed in 9 of the 13 tumours, including all but 1 of the Gas6-expressing tumours (χ^2 , $p=0.02$) (Table 1). No correlation between Gas6 expression and clinical presentation (grade, gender, location) or outcome was found in this small series.

Protein expression of was investigated in the two LMS cell lines, CNIO-AA (a soft tissue LMS) and SK-LMS-1 (a uterine LMS) using Western blot (Figure 3). CNIO-AA expressed high levels of Gas6, Tyro3, Axl, phosphorylated Akt, phosphorylated Erk1/2, and low levels of phosphorylated FAK Y397. Conversely, SK-LMS-1 expressed low to undetectable levels of phosphorylated Akt but phosphorylated Erk1/2 and Y397-phosphorylated FAK in contrast with what was observed in LMS samples (Figure 3). Gas6 was detectable at low concentrations ($<50\text{pg/mL}/106\text{cells}/24\text{h}$) in CNIO-AA but not in SK-LMS-1 supernatant. Both cell lines had detectable Gas6 in cell lysates, The protein expression profile of CNIO-

AA profile was therefore very similar to that of LMS samples with Gas6 expression, while that of SK-LMS-1 differed of LMS tumour samples, including the uterine LMS sample.

Interaction and colocalisation of Tyro-3 and Gas6 in léiomyosarcome:

To further investigate a putative Gas6/Tyro-3 autocrine pathway in leiomyosarcomas, we investigated the cellular interaction and co-localisation of Tyro3, Axl and Gas6 in SK-LMS-1 and CNIO-AA using the DuoLink assay. As shown in figure 4A, Gas6 and Tyro3 colocalize in CNIO-AA with a dot-like pattern within the tumour cells (Figure 4A). This was also observed with Axl and Gas6 (Figure 4B). Possibly because of the low level of Gas6 expression, the interaction of Gas6 with Tyro3 and Axl was not detectable in SK-LMS-1 (Figure 4C, 4D).

Response of leiomyosarcoma to exogenous Gas6:

Recombinant Gas6 was added to “low” Gas6-producing SK-LMS-1 culture. The addition of exogenous Gas6 on SK-LMS-1 cell line induced a significant increase of Akt phosphorylation in serum free conditions (Figure 5A). In the presence of serum, cell viability was not increased by the addition of Gas6 (Figure 5B). Upon serum deprivation, cell viability increased only marginally after 24 and 48 hours of exposure to Gas6 (Figure 5C). FAK Y397 phosphorylation was detectable in SK-LMS-1, and at low levels in CNIO-AA. Recombinant Gas6 did not alter the phosphorylation of FAK Y397 in serum free conditions in both cell lines (not shown).

Cell viability and caspase activity in cells exposed to Tyro3 and Axl inhibitors:

Anti-Gas6, and/or anti-Tyro3, and/or anti-Axl antibodies did not affect significantly SK-LMS-1 or CNIO-AA cell viability nor Akt phosphorylation (not shown). Cell viability of SK-LMS-1 and CNIO-AA cells was however reduced when cells were then cultured in the presence of foretinib, a tyrosine kinase inhibitor Tyro3, Axl, VEGFR2, and Met. Foretinib induced also an increase in caspase 3 and 7 activity (Figure 6). Phospho Axl expression was reduced upon foretinib exposure. Akt phosphorylation was not reduced in CNIO-AA, which lacks PTEN expression. CNIO-AA was then cultured with siRNA against Tyro3 and Axl. These siRNA significantly reduced the expression of both proteins in cells (Figure 7A). Gas6 expression decreased at 3h in the presence of Gas6 siRNA (not shown), but Akt phosphorylation was not affected in this PTEN negative cell line (Figure 7A). Tyro3 and Axl

siRNA combination reduced cell viability of CNIO-AA after 24, 48 and 72 hours of incubation (Figure 7B).

Discussion:

In the present work, we report that the Tyro3 and Axl receptor tyrosine kinases and their ligand Gas6, act in an autocrine manner, promoting cell survival and activation of Akt in leiomyosarcomas (LMS).

The initial goal was to identify novel targets in LMS by studying differences in proteomic pattern in LMS as compared to the normal smooth muscle (NSM). The choice of uterine NSM tissue as control was dictated by practical considerations (availability of hysterectomy specimens), although the proteomic pattern of uterine smooth muscle cells may be different to that of smooth muscle cells from other organs.

FAK Y397 was consistently hypophosphorylated in all LMS tested. Conversely, three proteins were found significantly up regulated, Tyro3, PKC α , and MSH2. The significance of the overexpression of these 2 proteins is unclear (20), PKC α has been found overexpressed in GIST and PNET (21, 22) However, the present work is focused only on the role of Tyro3. Axl and Tyro3 contribute to neoplastic transformation, tumour progression and resistance to treatment in renal cell carcinoma, pancreatic adenocarcinoma, glioblastoma, breast adenocarcinoma, synovial sarcoma (23-28). Tyro3 has been found activated and to contribute to tumour progression in melanoma and thyroid cancer (29-31).

Tyro3 was found overexpressed as compared to NSM in all the 13 LMS samples and in 2 LMS cell lines. Axl was also expressed and phosphorylated but at variable levels, in LMS samples and also in NSM. The overexpression of Tyro3 is therefore the specific molecular feature of LMS cells. We are exploring with CGH array the specific amplification of these genes on LMS samples.

Gas6 was expressed in 8 of the 13 LMS samples and this correlated with Akt phosphorylation. Similarly, the CNIO-AA cell line, a soft tissue leiomyosarcomas coexpressed Gas6 and phosphorylated Akt at high levels. This is consistent with the involvement of the PI3K/Akt pathway in Tyro3 signalling (32). SK-LMS-1 (a uterine LMS) expressed low levels of both Gas6 and phosphorylated Akt.

The observation of cellular colocalization of Tyro3 and Gas6 in the CNIO-AA cell line confirms the binding of Gas6 to its receptor, showing an autocrine, possible intracrine, pathway in this cell line. Of note while Gas6 was found undetectable in SK-LMS-1

supernatant, and detectable at low level in CNIO-AA supernatant, it was present at high levels in CNIO-AA tumour cell lysates, supporting a possible intracrine loop of Gas6 and its receptors in CNIO-AA, as already reported for KIT in GIST and IL-6 in renal cell carcinomas (33, 34). The lack of detection of co-localisation in the SK-LMS-1 may result from the lower level of expression of the cytokine in this cell line.

Antibodies against Axl, Tyro3 or Gas6 did not affect cellular proliferation in the 2 cell lines possibly because they act in an intracrine loop as reported in other models (33, 34). In the absence of serum, the SK-LMS-1 cell line experienced a dramatic decrease of cell viability, despite the induction of Gas6 proteins. Recombinant Gas6 did not significantly increase cell viability in the absence of serum possibly again because of a preferred intracrine mode of action in this cell lines. Foretinib, a broad spectrum tyrosine kinase inhibitor of Axl and Tyro3, but also VEGFR2 and Met, reduced cell viability and induced caspase 3/7 activity in both cell lines. The pattern of Tyro3, Gas6 and phospho-Akt expression in CNIO-AA is similar to that of LMS samples: combination siRNA against Axl and Tyro3 reduced significantly CNIO-AA viability, as well as the expression of their respective protein, while Gas6 siRNA induced only a transient decrease of Gas6 protein. SiRNA were not tested in the SK-LMS-1 cell line.

Taken together, these results indicate that 1) Tyro3 is overexpressed in all LMS cell lines and tumours, 2) Gas6 is expressed in a subgroup of LMS, 3) that Tyro3 and Gas6 colocalize within an intracellular compartment in CNIO-AA, 4) Gas6 expression correlates with phosphorylation of Akt, 5) Gas6 induces Akt phosphorylation in the SK-LMS-1 cell line, 6) that foretinib induces caspase activation, and impairs cell viability in both cell lines.

Tyro3 may therefore be target for tyrosine kinase inhibitors in leiomyosarcoma patients. A potential biomarker in this perspective may be Gas6 expression by leiomyosarcoma cells.

A downregulation of FAK phosphorylation on the critical Y397 residue was found an important proteomic hallmark of leiomyosarcomas. This is in marked contrast with what is observed in carcinoma models where FAK Y397 phosphorylation correlates to prognosis and aggressiveness in ovarian carcinoma and other carcinoma (35). Loss of FAK Y397 phosphorylation was unrelated to the presence or addition of Gas6. FAK promotes invasion, cell adhesion and matrix adhesion cross talk, as well as matrix adhesion independent cell growth and a protection from anoikis (35-37). FAK has been shown to play an important role in the formation of normal vascular structure by vascular smooth muscle cells (38). The loss of this physiological function of FAK may contribute to neoplastic progression in LMS, pointing out the possible vascular origin of most leiomyosarcomas. Actually, the

identification of Tyro3/Gas6 pathway contributing to LMS cell survival also points to the vascular smooth muscle cell origin of leiomyosarcoma cells in humans. Tyro3, Axl, and Gas6 are important signalling molecules for vascular smooth muscle cells (VSMC) (38). Gas6 is induced following vascular injury, promotes VSMC migration (39), and prevents VSMC apoptosis (40). A subgroup of leiomyosarcomas may well represent a neoplastic transformation of VSMC.

In conclusion, we report on the presence of an autocrine loop of Tyro3 and Gas6 associated with Akt activation in a subset of LMS. The Tyro3 family of receptors may represent an attractive target for tyrosine kinase inhibitors in leiomyosarcomas expressing Gas6.

Legends to the figures:

Figure 1: Four proteins differentially expressed in LMS vs normal smooth muscle

Fig1A: Level of expression of phosphorylated FAK in leiomyosarcoma (LMS) samples vs normal smooth muscle (NSM) according to the Kinex assay, expressed in arbitrary units (normalized intensity); Fig1B: Western blot showing the loss of FAK Y397 in LMS samples vs NSM; Fig1C: Relative expression (fold-increase of normalized intensities) of Tyro3, PKC theta, and MSH2 proteins in LMS samples as compared to NSM using the Kinex assay.

Figure 2: Expression of Tyro3, Axl, Gas6, and Akt proteins in leiomyosarcomas

Western blot evaluating the expression of Gas6, Tyro3, Axl, phosphorylated Axl, Akt, phosphorylated Akt in LMS samples and normal smooth muscle (NSM) as detailed in material and methods. All 5 tumours with low or undetectable levels of expression of Gas6 (LMS samples 1, 6, 7, 8, 13) had consistently low or undetectable levels of Akt protein phosphorylated on S473, while all 8 other LMS tumours with detectable Gas6 have also high levels of Akt protein phosphorylated on S473 (0 of 5 vs 8 of 8 $p=0.0003$).

Figure 3: Expression of Tyro-3, Gas6 and Axl, Erk and Akt in leiomyosarcoma cell lines

Western blot evaluating the expression of Gas6, Tyro3, Axl, Y397-phosphorylated FAK (FAK P-Y397) S473- phosphorylated Akt (Akt P-S473), phosphorylated Erk1 / 2 (P-Erk1/2) in leiomyosarcoma cell lines as detailed in material and methods.

Figure 4: Intracellular colocalisation of Tyro3 and its ligand Gas6

The Duolink assay (see material and methods), was used to demonstrate the interaction of Tyro3 and its ligand Gas6 (Fig4A, Fig4C) as well as Axl and Gas6 (Fig4B, Fig4D) in the CNIO-AA cell line (Fig4A, Fig4B) but not detectable in SK-LMS-1 cell line (Fig4C, Fig4D).

Figure 5: Modulation of cell viability and Akt phosphorylation in SK-LMS-1 by exogenous Gas6

Cell viability of SK-LMS-1 in the presence of exogenous Gas6 in the presence (Fig5A) or absence of serum (Fig 5B) using the Cell Titer-Glo® Luminescent Cell Viability Assay (Promega). LCPS: luminescence counts per second Fig5C: Western blot showing the modulation of Akt phosphorylation in SK-LMS-1 cell line by recombinant Gas6.

Figure 6: Foretinib reduces viability and induces caspase activity in SK-LMS-1 and CNIO-AA cell lines

Cells were cultured 24 hours in presence of the indicated concentrations of foretinib then tested for cell viability using the Cell titer glow assay (Fig6A, Fig6B) and caspase 3/7 activity using the Promega luminescent assay caspase-Glo 3/7 luminescent reagent (Fig6C).

Foretinib blocks the proliferation of both cell lines and induces caspase 3 and 7 activation in a dose dependent manner in the SK-LMS-1 and CNIO-AA cell lines.

Figure 7: siRNA directed against Gas6, Tyro3 and Axl reduce CNIO-AA cell viability

Fig7A: The expression of Tyro3 and Axl, in CNIO-AA exposed to 100nM specific siRNA ; *: for P-Akt, cells were exposed to 100nM of Gas6 siRNA only.

Fig7B: The cell viability of CNIO-AA is decreased when cultured with Tyro3 and Axl siRNA
LCPS: luminescence counts per second.

Conflict of interest statement:

JYB received research support and honoraria not related to this work from Novartis, Roche, GSK, Pharmamar, MSD.

References:

1. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of Soft Tissue and Bone. IARC Press: Lyon, France, 2002.
2. Clark MA, Fisher C, Judson I, Thomas JM. Soft-tissue sarcomas in adults. *N Engl J Med*. 2005;353:701-11.
3. JR. Toro, LB. Travis, HJ Wu, K Zhu, CDM. Fletcher, SS. Devesa. Incidence patterns of soft tissue sarcomas, regardless of primary site, in the Surveillance, Epidemiology and End Results program, 1978–2001: an analysis of 26,758 cases *Int. J. Cancer*: 2006; 119: 2922–2930.
4. Lurkin A, Ducimetière F, Vince DR, Decouvelaere AV, Cellier D, Gilly FN, Salameire D, Biron P, de Laroche G, Blay JY, Ray-Coquard I. Epidemiological evaluation of concordance between initial diagnosis and central pathology review in a comprehensive and prospective series of sarcoma patients in the Rhone-Alpes region. *BMC Cancer*. 2010; 10:150.
5. Casali PG, Jost L, Sleijfer S, Verweij J, Blay JY; ESMO Guidelines Working Group. Soft tissue sarcomas: ESMO clinical recommendations for diagnosis, treatment and follow-up. *Ann Oncol*. 2009; 20 (Suppl 4):132-6.
6. Van Glabekke M, van Oosterom AT, Oosterhuis JW, Mouridsen H, Crowther D, Somers R et al. Prognostic factors for the outcome of chemotherapy in advanced soft tissue sarcoma: an analysis of 2,185 patients treated with anthracycline-containing first-line regimens - a European Organization for Research and Treatment of Cancer Soft Tissue and Bone Sarcoma Group Study. *J Clin Oncol*. 1999; 17:150-7.
7. Blay JY, van Glabbeke M, Verweij J, van Oosterom AT, Le Cesne A, Oosterhuis. JW, Judson I, Nielsen OS. Advanced soft-tissue sarcoma: a disease that is potentially curable for a subset of patients treated with chemotherapy. *Eur J Cancer*. 2003; 39:64-9.
8. Le Cesne A, Blay JY, Judson I, Van Oosterom A, Verweij J, Radford J, Lorigan P, Rodenhuis S, Ray-Coquard I, Bonvalot S, Collin F, Jimeno J, Di Paola E, Van Glabbeke M, Nielsen OS. Phase II Study of ET-743 in Advanced Soft Tissue Sarcomas: A European Organisation for the Research and Treatment of Cancer (EORTC) Soft Tissue and Bone Sarcoma Group Trial. *J Clin Oncol*. 2005; 23:576-584.

9. Maki RG, Wathen JK, Patel SR, Priebe DA, Okuno SH, Samuels B, Fanucchi M, Harmon DC, Schuetze SM, Reinke D, Thall PF, Benjamin RS, Baker LH, Hensley ML. Randomized phase II study of gemcitabine and docetaxel compared with gemcitabine alone in patients with metastatic soft tissue sarcomas: results of sarcoma alliance for research through collaboration study 002. *J Clin Oncol*. 2007; 25:2755-63.
10. Sleijfer S, Ray-Coquard I, Papai Z, Le Cesne A, Scurr M, Schöffski P, Collin F, Pandite L, Marreaud S, De Brauwier A, van Glabbeke M, Verweij J, Blay JY. Pazopanib, a Multikinase Angiogenesis Inhibitor, in Patients With Relapsed or Refractory Advanced Soft Tissue Sarcoma: A Phase II Study From the European Organisation for Research and Treatment of Cancer-Soft Tissue and Bone Sarcoma Group (EORTC Study 62043). *J Clin Oncol*. 2009; 27: 3126-32.
11. Blay JY, Le Cesne A, Alberti L, Ray-Coquard I. Targeted cancer therapies. *Bull Cancer*. 2005 Feb;92(2):E13-8.
12. Helman LJ, Meltzer P. Mechanisms of sarcoma development. *Nat.Rev.Cancer*. 2003;3:685-94.
13. Larramendy ML, Kaur S, Svarvar C, Böhling T, Knuutila S. Gene copy number profiling of soft-tissue leiomyosarcomas by array-comparative genomic hybridization. *Cancer Genet Cytogenet*. 2006;169: 94-101.
14. Meza-Zepeda LA, Kresse SH, Barragan-Polania AH, Bjerkehagen B, Ohnstad HO, Namløs HM, Wang J, Kristiansen BE, Myklebost O. Array comparative genomic hybridization reveals distinct DNA copy number differences between gastro-intestinal stromal tumours and leiomyosarcomas. *Cancer Res*. 2006; 66:8984-93.
15. Larramendy ML, Gentile M, Soloneski S, Knuutila S, Böhling T. Does comparative genomic hybridization reveal distinct differences in DNA copy number sequence patterns between leiomyosarcoma and malignant fibrous histiocytoma? *Cancer Genet Cytogenet*. 2008; 187:1-11.
16. Carneiro A, Francis P, Bendahl PO, Fernebro J, Akerman M, Fletcher C, Rydholm A, Borg A, Nilbert M. Indistinguishable genomic profiles and shared prognostic markers in undifferentiated pleomorphic sarcoma and leiomyosarcoma: different sides of a single coin? *Lab Invest*. 2009;89:668-75.

17. Pérot G, Derré J, Coindre JM, Tirode F, Lucchesi C, Mariani O, Gibault L, Guillou L, Terrier P, Aurias A. Strong smooth muscle differentiation is dependent on myocardin gene amplification in most human retroperitoneal leiomyosarcomas. *Cancer Res.* 2009; 69: 2269-78.
18. F Qian, S Engst, K Yamaguchi, P Yu, K-A Won, L.Mock,T.Lou, J. Tan, C.Li, D.Tam, J.e Loughed, F. M Yakes, F Bentzien, W Xu, T Zaks, R Wooster, J Greshock, AH. Joly. Inhibition of Tumour Cell Growth, Invasion, and Metastasis by EXEL-2880 (XL880, GSK1363089), a Novel Inhibitor of HGF and VEGF Receptor Tyrosine Kinases. *Cancer Res* 2009; 69:: 8009-16
19. Linger RM, Keating AK, Earp HS, Graham DK. TAM receptor tyrosine kinases: biologic functions, signaling, and potential therapeutic targeting in human cancer. *Adv Cancer Res.* 2008; 100:35-83.
20. Nilbert M, Therkildsen C, Nissen A, Akerman M, Bernstein I. Sarcomas associated with hereditary nonpolyposis colorectal cancer: broad anatomical and morphological spectrum. *Fam Cancer.* 2009; 8:209-13.
21. Kang GH, Kim KM, Park CK, Kang DY. PKC-theta expression in Ewing sarcoma/primitive neuroectodermal tumour and malignant peripheral nerve sheath tumour. *Histopathology.* 2009; 55:368-9. .
22. Blay P, Astudillo A, Buesa JM, Campo E, Abad M, García-García J, Miquel R, Marco V, Sierra M, Losa R, Lacave A, Braña A, Balbín M, Freije JM. Protein kinase C theta is highly expressed in gastrointestinal stromal tumours but not in other mesenchymal neoplasias. *Clin Cancer Res.* 2004;10: 4089-95.
23. Gustafsson A, Martuszevska D, Johansson M, Ekman C, Hafizi S, Ljungberg B, Dahlbäck B. Differential expression of Axl and Gas6 in renal cell carcinoma reflecting tumour advancement and survival. *Clin Cancer Res.* 2009; 15: 4742-9.
24. Koorstra JB, Karikari CA, Feldmann G, Bisht S, Rojas PL, Offerhaus GJ, Alvarez H, Maitra A. The Axl receptor tyrosine kinase confers an adverse prognostic influence in pancreatic cancer and represents a new therapeutic target. *Cancer Biol Ther.* 2009; 8:618-26..
25. Hong CC, Lay JD, Huang JS, Cheng AL, Tang JL, Lin MT, Lai GM, Chuang SE. Receptor tyrosine kinase AXL is induced by chemotherapy drugs and overexpression of AXL confers drug resistance in acute myeloid leukemia. *Cancer Lett.* 2008 Sep 18;268(2):314-24.

26. Zhang YX, Knyazev PG, Cheburkin YV, Sharma K, Knyazev YP, Orfi L, Szabadkai I, Daub H, Kéri G, Ullrich A. AXL is a potential target for therapeutic intervention in breast cancer progression. *Cancer Res.* 2008; 68:1905-15.
27. Hutterer M, Knyazev P, Abate A, Reschke M, Maier H, Stefanova N, Knyazeva T, Barbieri V, Reindl M, Muigg A, Kostron H, Stockhammer G, Ullrich A. Axl and growth arrest-specific gene 6 are frequently overexpressed in human gliomas and predict poor prognosis in patients with glioblastoma multiforme. *Clin Cancer Res.* 2008; 14:130-8.
28. Fernebro J, Francis P, Edén P, Borg A, Panagopoulos I, Mertens F, Vallon-Christersson J, Akerman M, Rydholm A, Bauer HC, Mandahl N, Nilbert M. Gene expression profiles relate to SS18/SSX fusion type in synovial sarcoma. *Int J Cancer.* 2006;118:1165-72.
29. Zhu S, Wurdak H, Wang Y, Galkin A, Tao H, Li J, Lyssiotis CA, Yan F, Tu BP, Miraglia L, Walker J, Sun F, Orth A, Schultz PG, Wu X. A genomic screen identifies TYRO3 as a MITF regulator in melanoma. *Proc Natl Acad Sci U S A.* 2009; 106:17025-30. Epub 2009 Sep 23.
30. Loges S, Schmidt T, Tjwa M, van Geyte K, Lievens D, Lutgens E, Vanhoutte D, Borgel D, Plaisance S, Hoylaerts M, Luttun A, Dewerchin M, Jonckx B, Carmeliet P. Malignant cells fuel tumour growth by educating infiltrating leukocytes to produce the mitogen Gas6. *Blood.* 2010; 115:2264-73.
31. Avilla E, Guarino V, Visciano C, Liotti F, Svelto M, Krishnamoorthy G, Franco R, Melillo RM. Activation of TYRO3/AXL Tyrosine Kinase Receptors in Thyroid Cancer. *Cancer Res.* 2011; 71: 1792-804.
32. Lan Z, Wu H, Li W, Wu S, Lu L, Xu M, Dai W. Transforming activity of receptor tyrosine kinase Tyro3 is mediated, at least in part, by the PI3 kinase-signaling pathway. *Blood.* 2000; 95:633-8.
33. Tabone-Eglinger, S, Subra F, El Sayadi H, Alberti L, Tabone E, Michot JP, Théou-Anton N, Lemoine A, Blay JY, Emile JF. KIT Mutations Induce Intracellular Retention and Activation of an Immature Form of the KIT Protein in Gastrointestinal Stromal Tumours. *Clin Cancer Res.* 2008;14:2285-2294
34. Alberti L, Thomachot MC, Bachelot T, Menetrier-Caux C, Puisieux I, Blay JY. IL-6 as an intracrine growth factor for renal carcinoma cell lines. *Int J Cancer.* 2004; 20;111: 653-61.

35. McLean GW, Carragher NO, Avizienyte E, Evans J, Brunton VG, Frame MC. The role of focal-adhesion kinase in cancer - a new therapeutic opportunity. *Nat Rev Cancer*. 2005; 5:505-15.
- 36.. Frisch SM, Vuori K, Ruoslahti E, Chan-Hui PY. Control of adhesion-dependent cell survival by focal adhesion kinase. *J Cell Biol*. 1996; 134:793-9
37. Sieg DJ, Hauck CR, Ilic D, Klingbeil CK, Schaefer E, Damsky CH, Schlaepfer DD. FAK integrates growth-factor and integrin signals to promote cell migration. *Nat Cell Biol*. 2000;2:249-56.
38. Cheng Z, Sundberg-Smith LJ, Mangiante LE, Sayers RL, Hakim ZS, Musunuri S, Maguire CT, Majesky MW, Zhou Z, Mack CP, Taylor JM. Focal adhesion kinase regulates smooth muscle cell recruitment to the developing vasculature. *Arterioscler Thromb Vasc Biol*. 2011;31:2193-202.
39. Fridell, Y. W., Villa, J., Jr., Attar, E. C., and Liu, E. T. (1998). GAS6 induces Axl-mediated chemotaxis of vascular smooth muscle cells. *J. Biol. Chem*. 1998; 273: 7123–7126.
40. Melaragno, M. G., Cavet, M. E., Yan, C., Tai, L. K., Jin, Z. G., Haendeler, J., and Berk, B. C. Gas6 inhibits apoptosis in vascular smooth muscle: Role of Axl kinase and Akt. *J. Mol. Cell Cardiol*. 2004. 37: 881–887.

Table 1: Characteristics of the patients and semi-uantitative expression of molecular markers

Patient Id	Age	Sex	Grade	Site of the primary	Clinical Status	Tyro3	Gas6	Axl	Phospho Axl	Akt	Phospho Akt	Erk1/2	Phospho Erk
LMS 1	80	M	3	Shoulder	NED	++	0	+	++	++	+	++	++
LMS 3	55	F	2	Retroperitoneum	DOD	++	++	+	++	++	++	++	++
LMS 4	64	F	3	Thigh	NED	+	++	++	+	++	++	++	++
LMS 5	75	M	3	Thigh	DOD	++	+	+	+	++	++	++	++
LMS 6	64	F	3	Thigh	DOD	+	0	++	+	++	0	+	0
LMS 2	71	F	2	Thigh	DOD	++	++	++	++	++	++	++	0
LMS 11	69	M	3	Abdominal wall	DOD	++	0	++	+	++	0	++	0
LMS 9	72	F	3	Trunk wall	DOD	++	++	+	+	++	++	++	++
LMS 10	69	M	2	Leg	DOD	+	++	+	++	++	++	++	++
LMS 8	65	F	3	Pelvis	NED	+	0	++	+	++	+	++	++
LMS 7	80	F	2	Arm	AWD	++	0	++	+	++	+	++	0
LMS 12	53	F	3	Uterus	LFU	+	++	+	+	++	++	++	+
LMS 13	67	F	3	Retroperitoneum	NED	++	++	+	++	++	++	++	++

Legend :

AWD : alive with disease, DOD : dead of disease, NED : non evolutive disease, LFU : lost to follow-up,

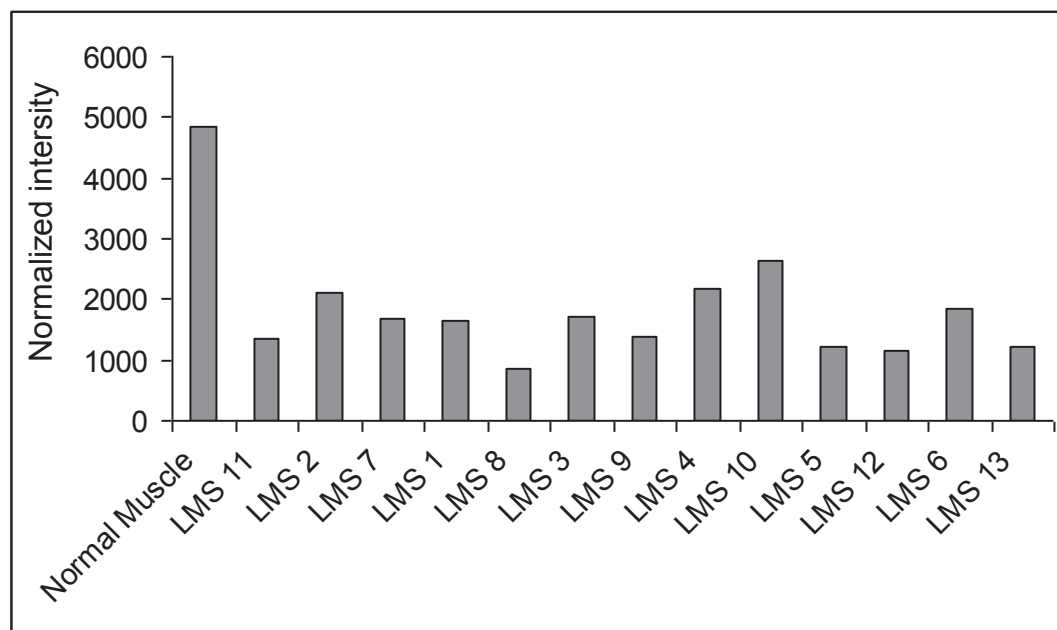
Semi-qantitative evaluation of protein expression: 0: no expression, + low elvel of expression, ++ high level of expression

Figure 1: Loss of FAK P –Y397, overexpression of PKC Θ , MSH2, Tyro3 in LMS

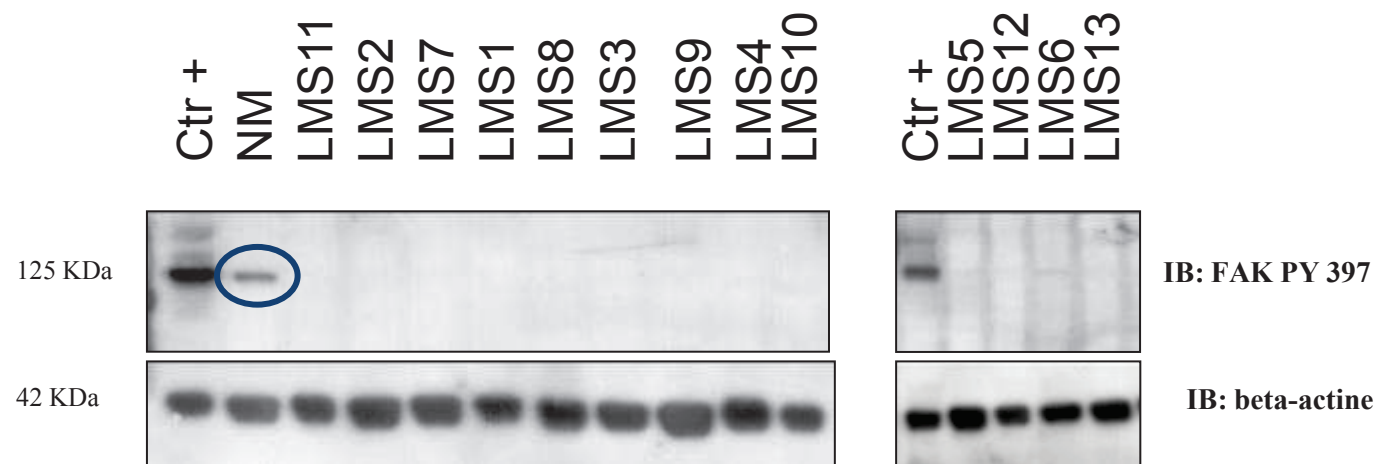
230

Fig. 1A

Kinex Assay

**Fig. 1B**

Western blot



Figures:



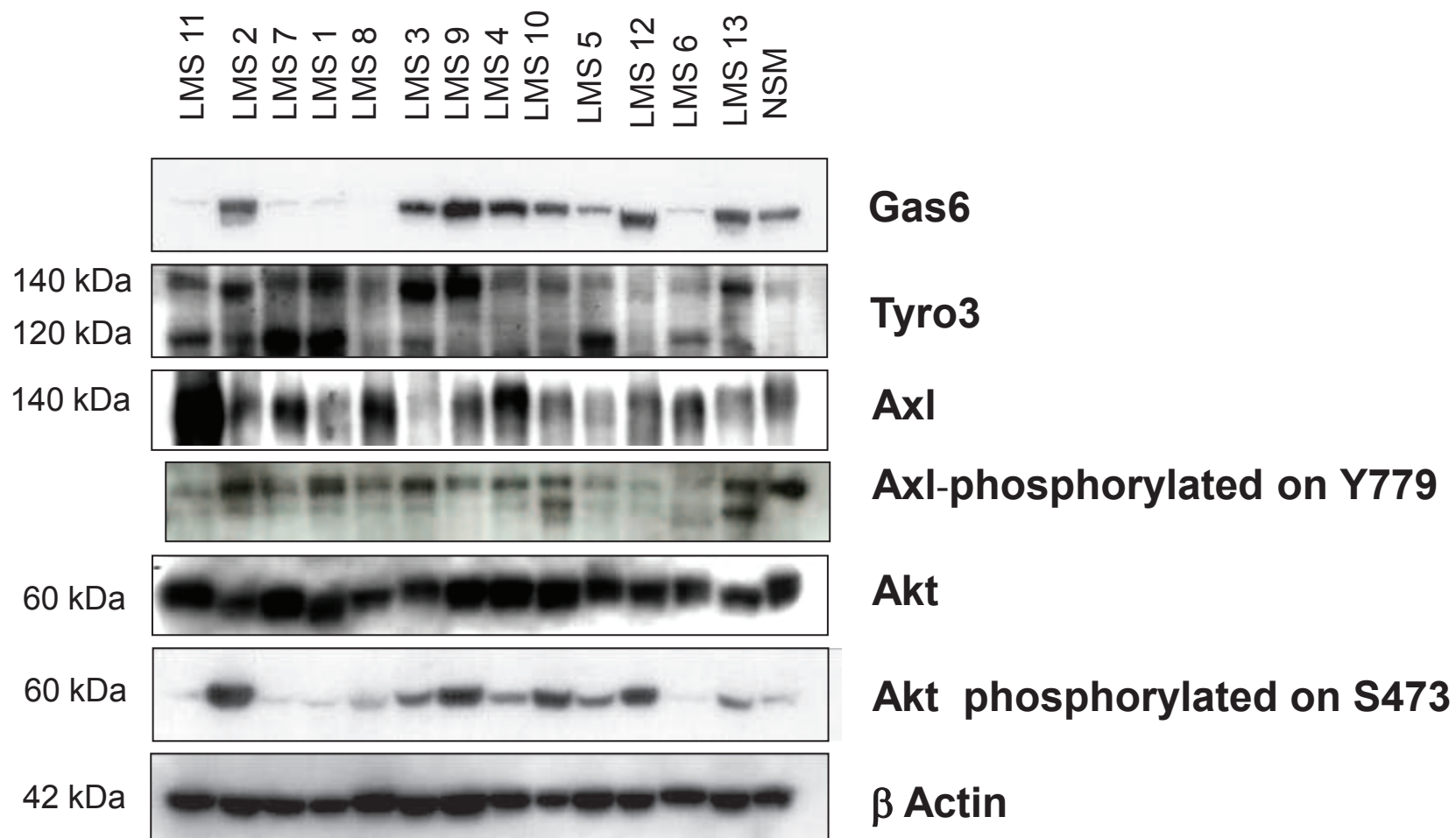
Fig. 2: Expression of Tyro3, Axl, Gas6, and Akt proteins in leiomyosarcomas samples.

Fig. 3: Expression of Gas6, Tyro3 and Axl, FAK, Akt and Erk1/2 proteins in leiomyosarcoma cell lines

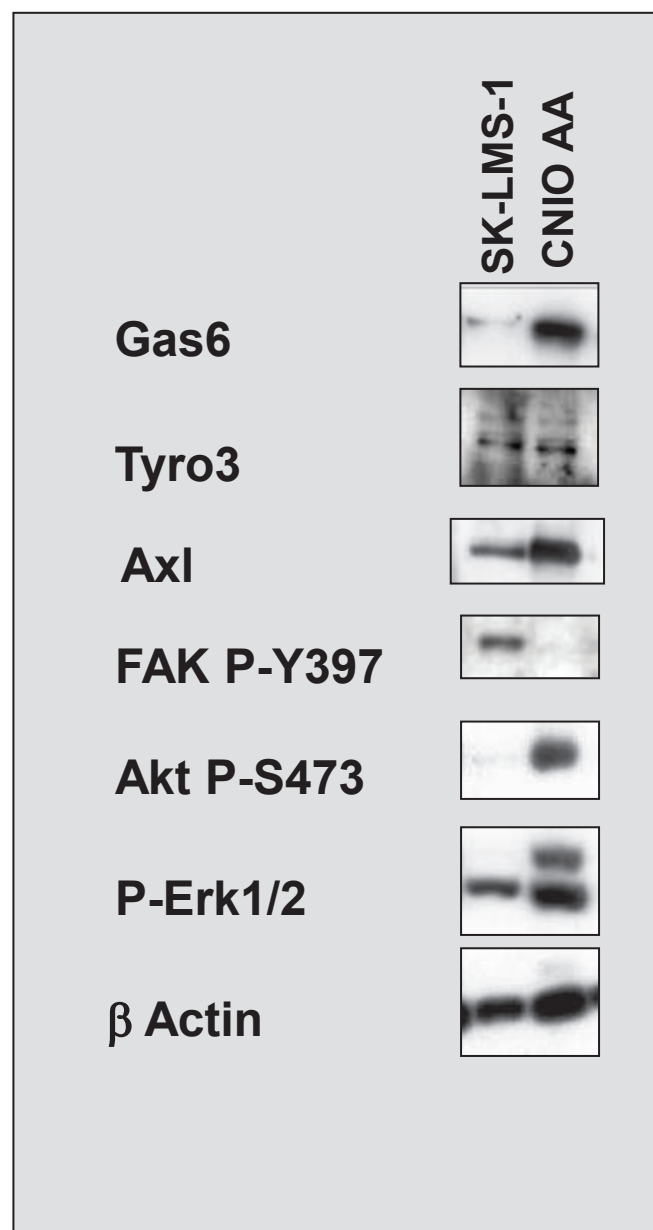


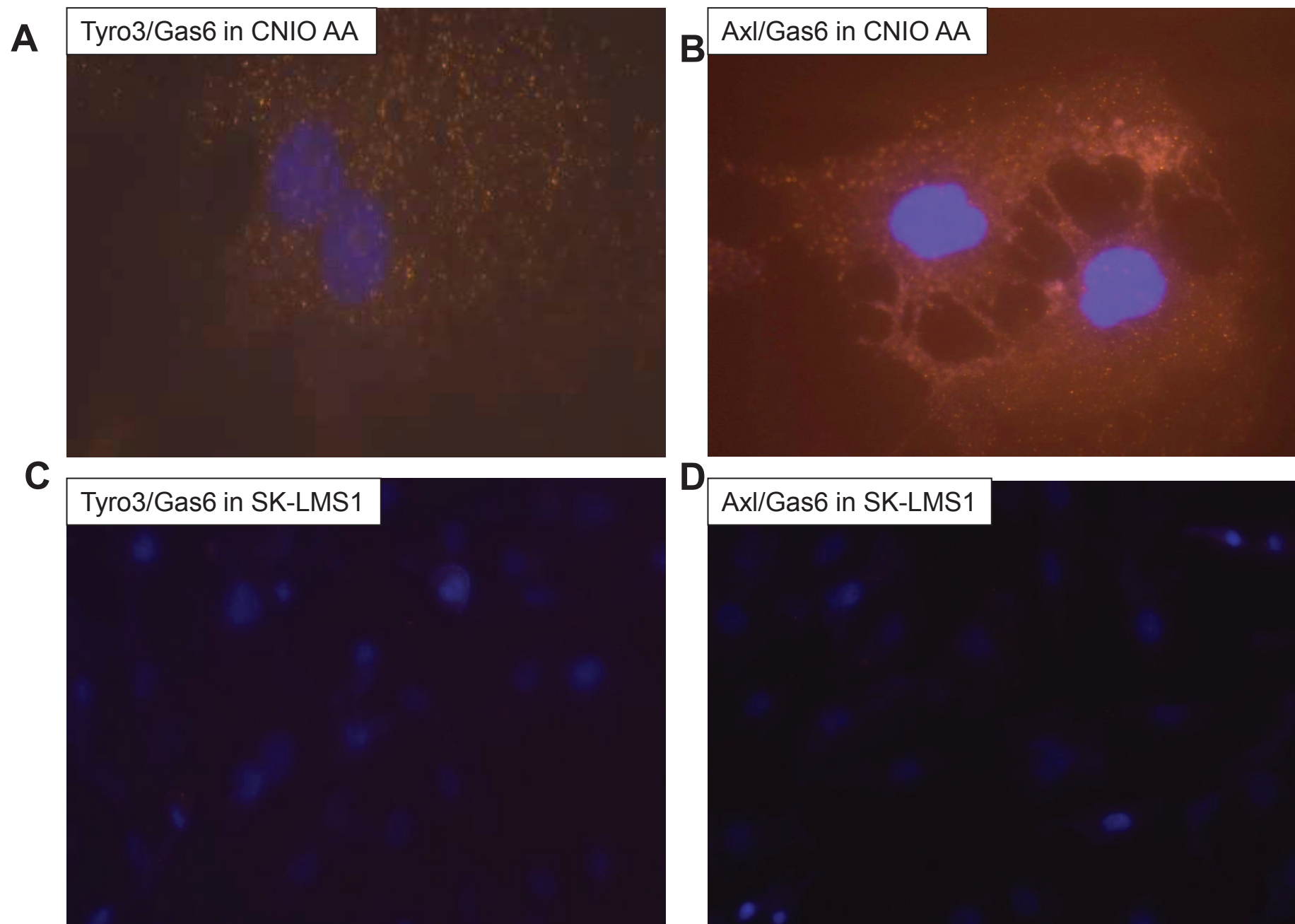
Fig.4: Intracellular colocalisation of Tyro3 and its ligand Gas6

Fig. 5: Modulation of cell viability and Akt phosphorylation in SK-LMS-1 by exogenous Gas6

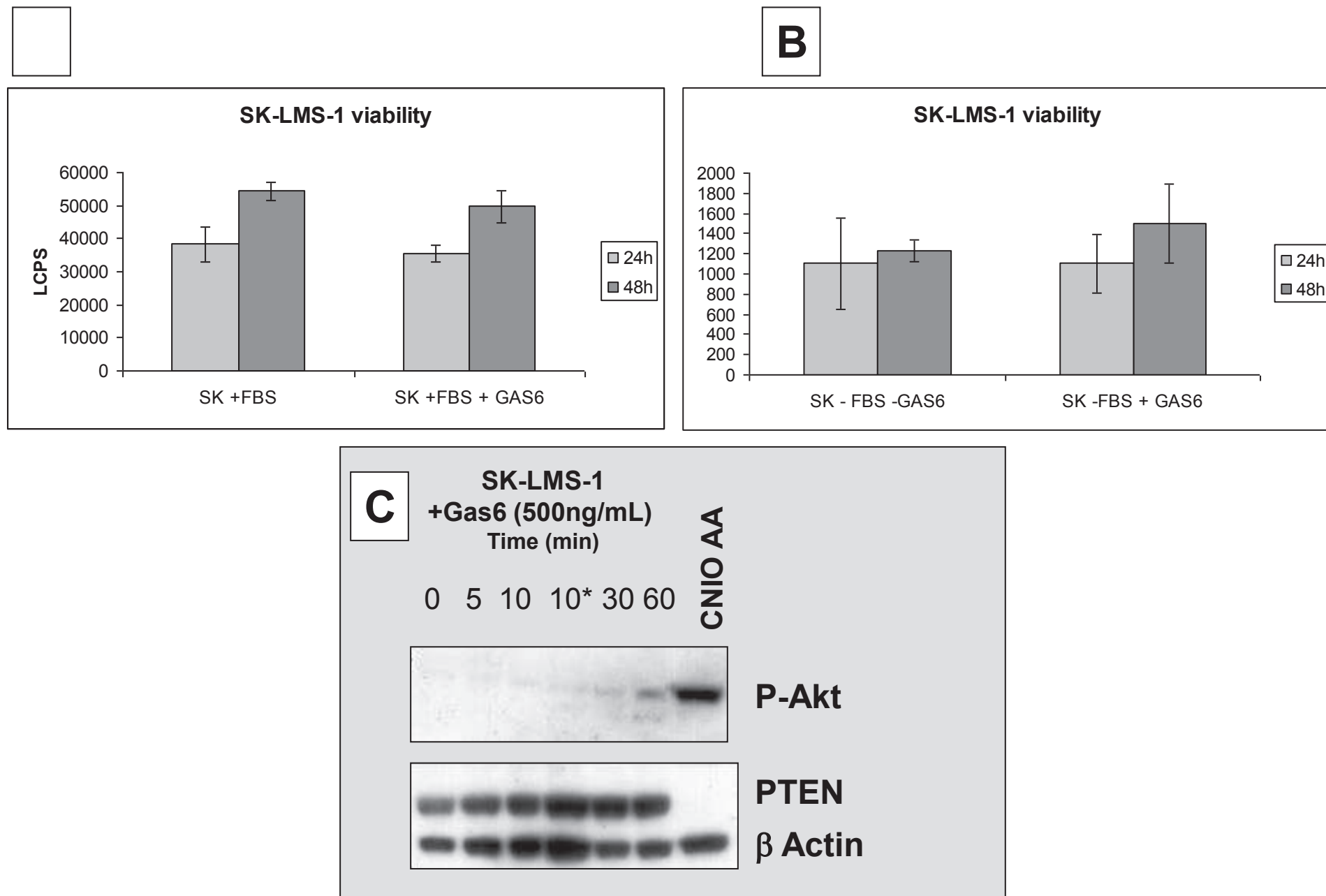


Fig. 6: Modulation of proliferation and apoptosis in SK-LMS-1 and CNIO-AA with and Axl/Tyro3 tyrosine kinase inhibitor

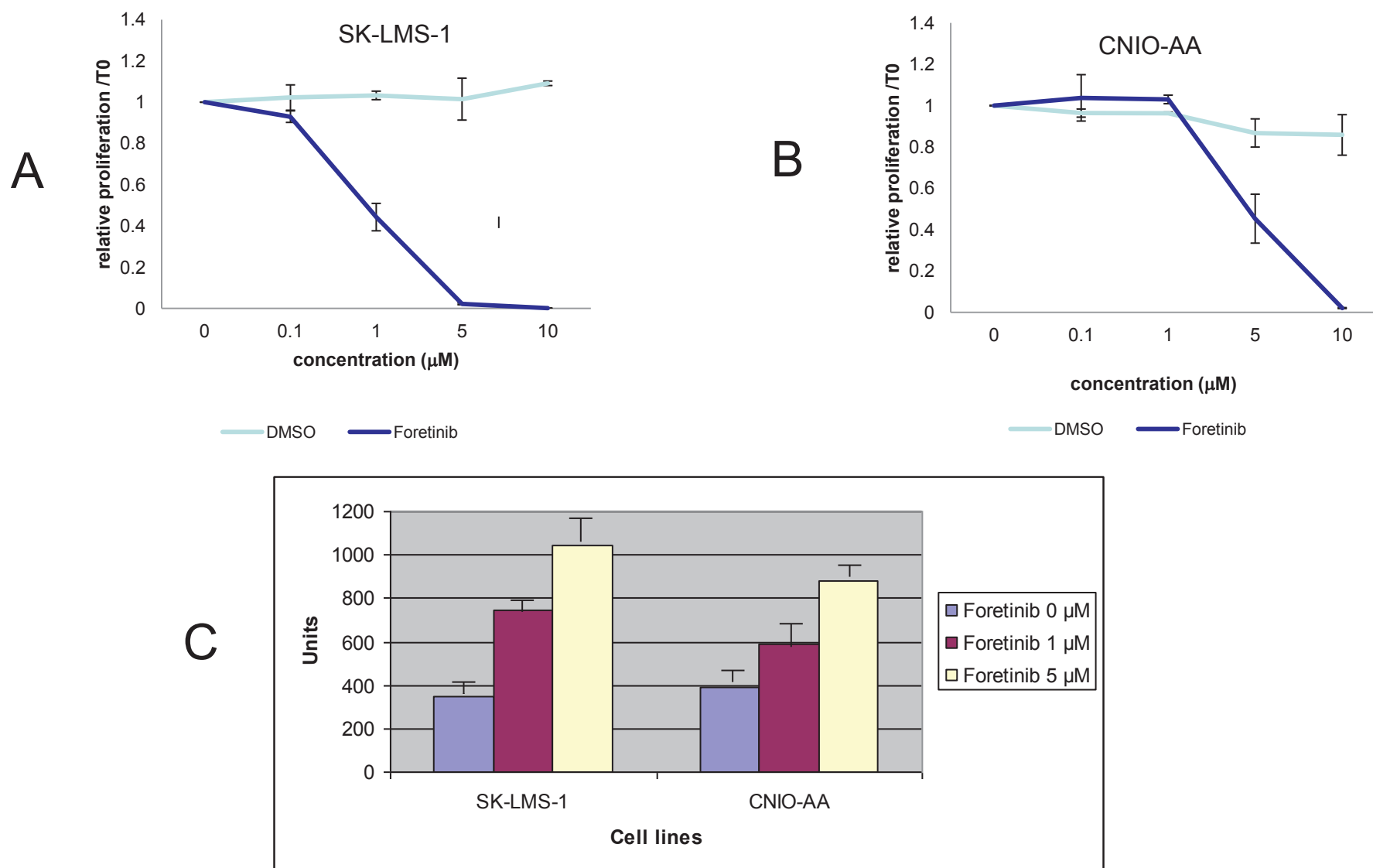
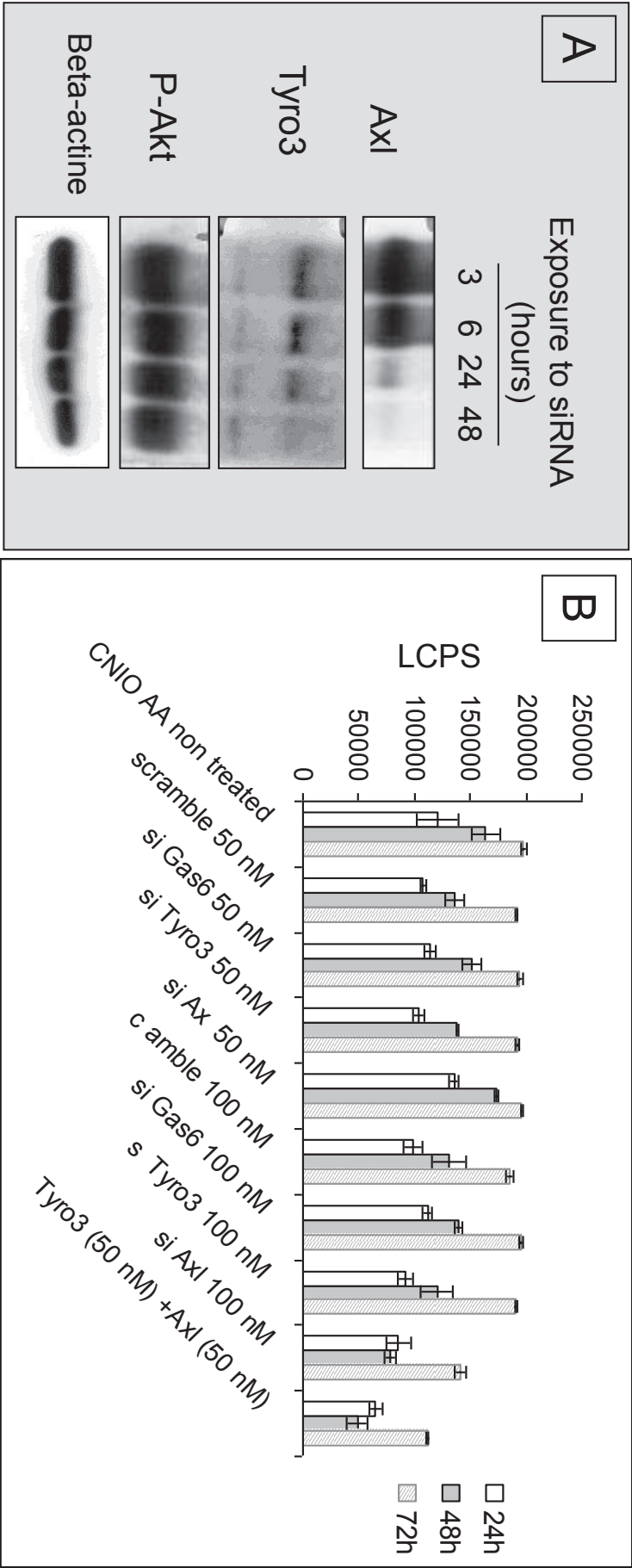


Fig. 7: Gas6, Tyro3 and Axl siRNA reduce CNIO-AA cell viability



DISCUSSION ET PERSPECTIVES

I. IMPORTANCE DE LA DÉTECTION DES BIAIS BIOLOGIQUES ET DES BIAIS D'ANALYSE

Au cours du développement de notre méthodologie d'analyse des données d'expression réalisées à partir de la plateforme TLDA (cf. article I, en soumission au journal *RNA*), nous avons insisté sur les nombreux contrôles qualité qu'il était indispensable de réaliser tout au long du processus d'analyse, de l'extraction des Ct bruts à l'obtention des valeurs de $\log_2(RQ)$. La détection des biais biologiques permet avant tout à l'utilisateur de mieux cerner le comportement de ses données brutes, afin de pouvoir leur appliquer la méthode de normalisation la plus adaptée. La détection des biais de manipulations et d'analyse permet de s'assurer que les modifications appliquées aux données sont cohérentes et statistiquement valides.

1. Biais biologiques

Comme la plupart des techniques de biologie moléculaire, la réalisation des profils d'expression des microARN à partir de la plateforme TLDA est sujette à des biais liés à la nature du prélèvement analysé, et notamment à sa richesse et à son hétérogénéité tumorale.

La richesse en cellules tumorales du prélèvement analysé est primordiale. En effet, les échantillons bruts de tumeurs sont toujours constitués d'un mélange de cellules tumorales et de cellules stromales. Bien que ces cellules stromales jouent un rôle important dans le développement de la tumeur et l'émergence de métastases (Mueller and Fusenig, 2004; Kalluri and Zeisberg, 2006), seule l'analyse des cellules tumorales permet d'appréhender les caractéristiques biologiques intrinsèques d'une tumeur. Les profils d'expression des microARN étant réalisés à partir d'une extraction des ARN de plusieurs milliers de cellules, le profil d'expression observé

est donc une moyenne des profils d'expression de chacune des cellules extraites. De ce fait, il est important que l'échantillon de test possède un pourcentage de cellules tumorales le plus élevé possible (on fixe généralement ce seuil à 50 %). Actuellement, le pourcentage de cellules tumorales est évalué par un médecin anatomopathologiste, généralement à partir d'une coupe colorée HES. Cependant, même pour des pathologistes expérimentés, un système de score est subjectif et imprécis (Furness et al., 2003).

Au démarrage de ce travail de thèse, les profils d'expression des microARN étaient réalisés à partir de solutions d'ARN déjà extraites à partir de copeaux d'un bloc de paraffine. Cette approche concernait en particulier les RMS, puisque la plupart des tumeurs avaient bénéficié au préalable d'une recherche de transcrits de fusion par PCRq. Le pourcentage de cellules tumorales n'a donc pas été évalué lors de cette approche, puisque la méthode de diagnostic employée était une méthode du tout ou rien (présence ou absence de transcrit de fusion). Par la suite, les tumeurs pour lesquelles nous avons réalisé les profils d'expression sur plaques TLDA ont pu bénéficier de grattages du bloc de paraffine dans une zone repérée par un pathologiste, pour laquelle un pourcentage de cellules tumorales avait été évalué sur coupe colorée HES.

En ce qui concerne les prélèvements congelés analysés au cours de notre étude sur la plateforme TLDA, le pourcentage tumoral n'a pas pu être évalué puisque les tumeurs ont été directement cryobroyées. La mesure du pourcentage de cellules tumorales est plus compliquée à réaliser à partir d'un prélèvement congelé. Cela nécessite de débiter le fragment tumoral à l'aide d'un cryostat, afin d'en récupérer des copeaux pour extraction, en veillant à apposer au cours de ce travail une coupe congelée représentative sur lame blanche, pour une évaluation du pourcentage de cellules tumorales après coloration.

Il est très difficile d'évaluer les conséquences précises d'une faible représentativité tumorale d'un échantillon sur la base des profils d'expression de gènes, car les niveaux d'expression résultants dépendent fortement de la nature des

cellules contaminantes (par exemple, cellules de tissus normaux environnants ou cellules inflammatoires). Il nous a tout de même été possible d'observer ce phénomène par une approche de technique CGH, sur une tumeur mélanique analysée dans le cadre de notre activité diagnostique (figure 40). En effet, si la tumeur est homogène, les décalages liés aux anomalies génomiques sont directement proportionnels au pourcentage de cellules tumorales.

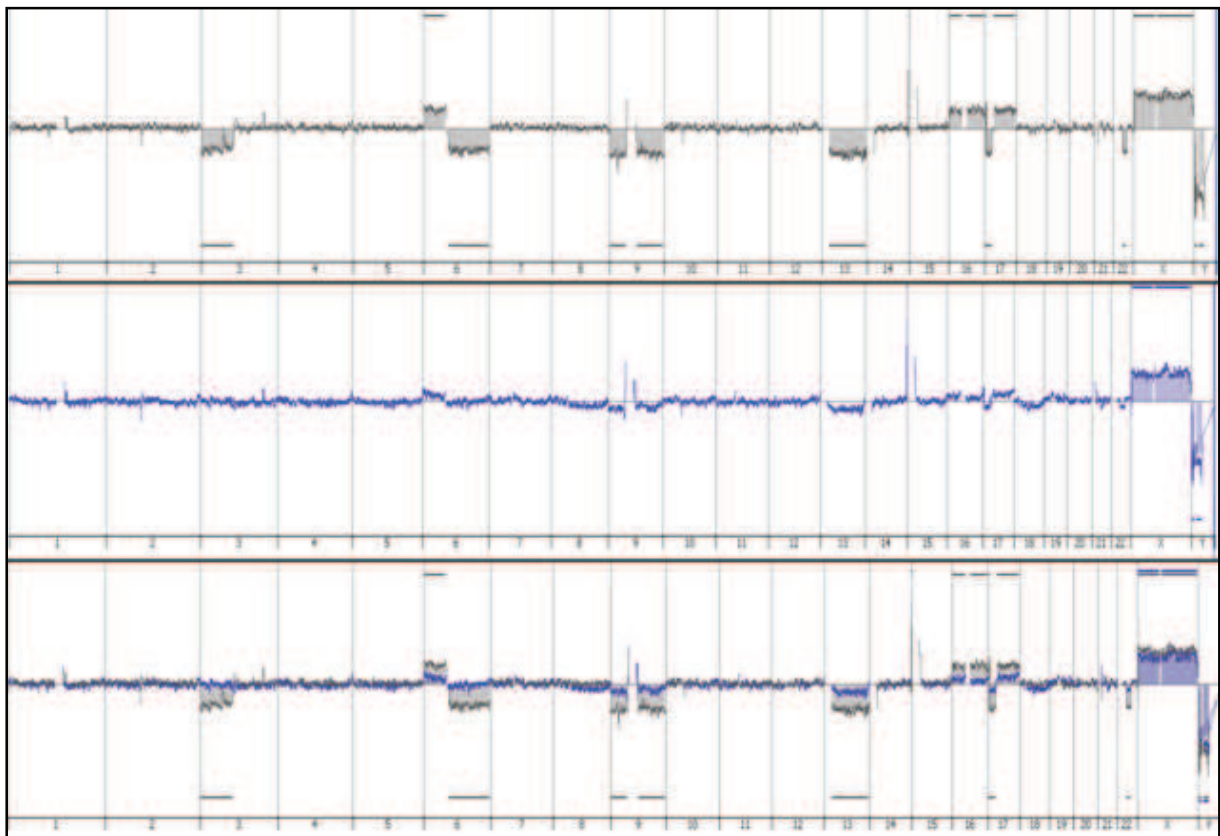


Figure 40 : profils génomiques aCGH obtenus sur plateforme Agilent pour un même échantillon de tumeur mélanique, analysé sur deux zones différentes.

En haut, profil aCGH de la zone riche en cellules tumorales. Au milieu, profil aCGH de la zone pauvre en cellules tumorales. Le graphique du bas est une superposition des deux profils. Lorsque le pourcentage de cellules tumorales est faible, on observe un phénomène de dilution des anomalies génomiques, et seules des « tendances » sont détectables (profils génomiques réalisés par S. Paindavoine, au Centre Léon Bérard).

Ainsi, il nous faut garder à l'esprit que la variabilité des profils d'expression que nous avons pu observer entre des échantillons d'un même type tumoral peut en partie être expliquée par la représentativité en cellules tumorales de chacun de ces prélèvements. Certaines expressions différentielles ont également pu être gommées en raison de ce phénomène de dilution. Dans la mesure du possible, il convient donc de travailler avec des échantillons correctement pré-sélectionnés par un médecin anatomopathologiste. Une augmentation du nombre d'échantillons dans les cohortes d'analyse permet généralement de gommer cette variabilité.

L'hétérogénéité tumorale est une donnée bien connue des médecins anatomopathologistes, puisqu'elle est directement visible au niveau macroscopique et microscopique, et influe fréquemment sur leur interprétation diagnostique. En cytogénétique moléculaire, ce constat est également vérifié *in situ*, par l'intermédiaire de la technique FISH. Un exemple frappant d'hétérogénéité tumorale a été mis à jour au sein de notre équipe, lors d'un test diagnostique. Une tumeur sphérique bien délimitée de type synoviosarcome présentait un réarrangement typique du gène SS18 sur la totalité du prélèvement ; de façon additionnelle, nous avons mis en évidence la présence d'une amplification du gène MDM2 sur une moitié seulement de ce même prélèvement (figure 41).

L'hétérogénéité tumorale est aujourd'hui particulièrement bien étudiée par les biologistes moléculaires (Gerlinger et al., 2012; Ng et al., 2012). Elle pourrait être l'un des mécanismes expliquant l'évolution et l'adaptation des cancers. Selon ce concept, certaines tumeurs seraient une mosaïque complexe formée de plusieurs clones juxtaposés ou intriqués.

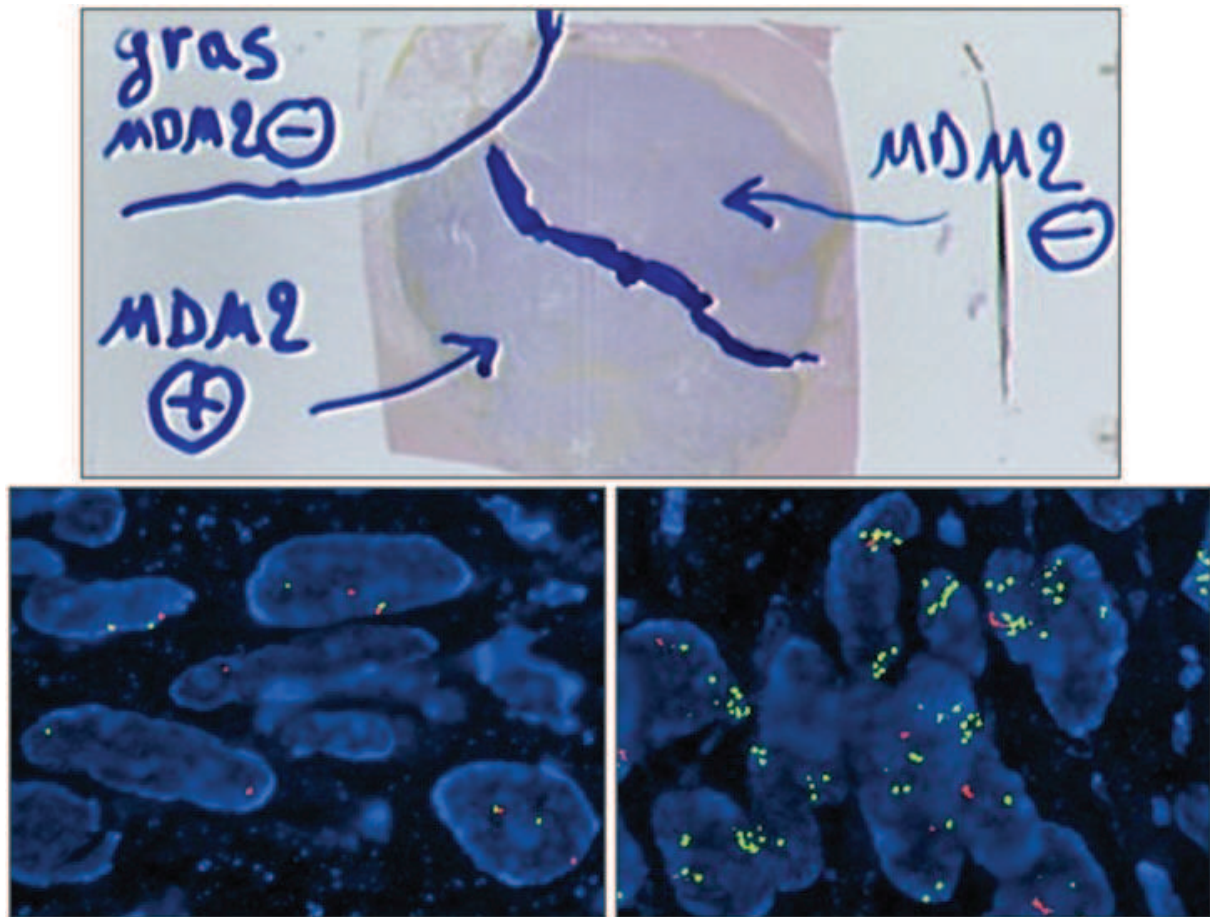


Figure 41 : techniques de FISH réalisées sur coupe paraffine d'un échantillon de synoviosarcome. Une photo de la lame HES représentative est présentée en haut de la figure. On observe la présence d'un réarrangement du gène SS18 sur la totalité du prélèvement (photo en bas à gauche), alors que l'amplification du gène MDM2 (MDM2+) n'est présente que dans une moitié du prélèvement (photos en bas à droite) (techniques FISH réalisées par M. Perrin-Niquet, au Centre Léon Bérard).

Il est très important de bien sélectionner la zone que l'on souhaite analyser, car les biais biologiques liés à la richesse et à l'hétérogénéité tumorale sont difficilement détectables et corrigibles (Roepman et al., 2009). La microdissection laser peut être d'une grande aide pour la réalisation de profils d'expression des microARN à partir de zones d'intérêt de très petites tailles (Hackler, Jr. et al., 2012a; Hackler, Jr. et al., 2012b), puisqu'il est aujourd'hui possible de réaliser des études sur cellule unique (Kalisky and Quake, 2011).

2. Biais de manipulations et d'analyse

Certains biais peuvent être introduits lors de la manipulation des échantillons biologiques. Au moment du prélèvement tumoral chez le patient, certaines étapes sont critiques pour les échantillons destinés aux analyses de biologie moléculaire à partir de l'ARN (Jewell et al., 2002). Une fois l'échantillon obtenu, il est conseillé de le traiter immédiatement, par congélation dans l'azote liquide, par fixation dans le formol ou par immersion dans du RNAlater®. Ceci est primordial pour les échantillons destinés à une extraction ARN, car les RNases, enzymes qui dégradent les ARN, s'activent dès que le prélèvement est en ischémie chaude (temps écoulé entre le clampage vasculaire et l'exérèse), et le phénomène se poursuit lors de l'ischémie froide (délai entre l'exérèse et la stabilisation définitive du tissu par la congélation ou la fixation) (Lazarus, 1974). Dans le cadre de nos études des profils d'expression des microARN sur plaque TLDA, cela pourrait se traduire par de nombreuses valeurs de Ct indéterminées. En outre, il existe une activité métabolique induite par l'ischémie, qui pourrait également venir perturber les profils d'expression des microARN (Ma et al., 2012b). De telles anomalies seraient indétectables au niveau de l'analyse des profils d'expression des microARN. Les temps d'ischémie relatifs à chaque échantillon analysé ne sont généralement pas connus des biologistes moléculaires, et il faut être conscient qu'ils peuvent influencer les profils d'expression des microARN. Cependant, en contrepartie, il a été montré que les microARN seraient moins sujets à la dégradation que les ARNm (Jung et al., 2010).

Dans le cadre de l'analyse de prélèvements osseux, et notamment pour l'établissement des profils d'expression des microARN dans les OS, nous avons été confrontés à des prélèvements ayant subi une étape de décalcification. Cette étape effectuée en routine dans les laboratoires d'anatomopathologie permet de ramollir les tissus osseux, pour permettre les coupes de lames blanches au microtome. Il a été montré que les agents décalcifiants dégradaient fortement les acides nucléiques

(Shibata et al., 2000; Wickham et al., 2000). Notre expérience nous a prouvé que les ARN extraits à partir d'échantillons décalcifiés à l'acide trichloroacétique (TCA) n'étaient pas exploitables pour l'établissement des profils d'expression des microARN (échantillons d'OS traités au CLB), alors que les échantillons décalcifiés à l'EDTA se sont révélés parfaitement exploitables (échantillons d'OS traités à La Timone).

Lors du développement de notre méthodologie d'analyse des données issues de plaques TLDA, nous avons montré que la fixation des échantillons dans le formol introduisait un biais particulier, non homogène sur la gamme dynamique des Ct, et que ce biais devait être corrigé lorsqu'un échantillon FFPE était comparé à un échantillon congelé (cf. article I, en soumission au journal *RNA*). Nous avons montré qu'il est possible de corriger ce biais à l'aide d'une procédure de normalisation locale de type LOESS. Le graphique de corrélation de type MA, développé initialement pour les données d'intensité des expériences d'hybridation sur puce, est également indispensable pour analyser la nature des biais pouvant survenir lors d'expériences issues de plaques TLDA. Cependant, de nombreuses autres conditions non testées pourraient également induire des biais, et il n'est pas certain que ces biais soient visibles sur ce type de graphique, d'où l'importance de multiplier les contrôles à l'aide d'outils variés (boîtes à moustaches, diagramme de Pearson, etc.).

Après une expérience de PCRq sur plaque TLDA, pour obtenir des résultats ayant une signification biologique correcte, les données brutes vont être progressivement modifiées, de manière séquentielle. Chacune de ces étapes peut être l'occasion d'introduire un éventuel biais de manipulation. L'utilisateur doit donc placer de nombreux points de contrôle lui permettant de valider au fur et à mesure ces transformations.

Avant l'extraction des données brutes, il est indispensable de vérifier visuellement les courbes d'amplification pour éliminer les puits aberrants. Sans ce contrôle visuel, nous nous sommes aperçus que le logiciel était capable de générer des Ct cohérents à partir de puits pour lesquels il y avait manifestement eu un problème technique lors de la réaction de PCRq (figure 42). Les puits présentant ce type d'aberration doivent être éliminés de tout le processus d'analyse, et considérés comme des valeurs manquantes. En effet, sans cette précaution, ce type d'erreur stochastique ne peut plus être corrigé par la suite.

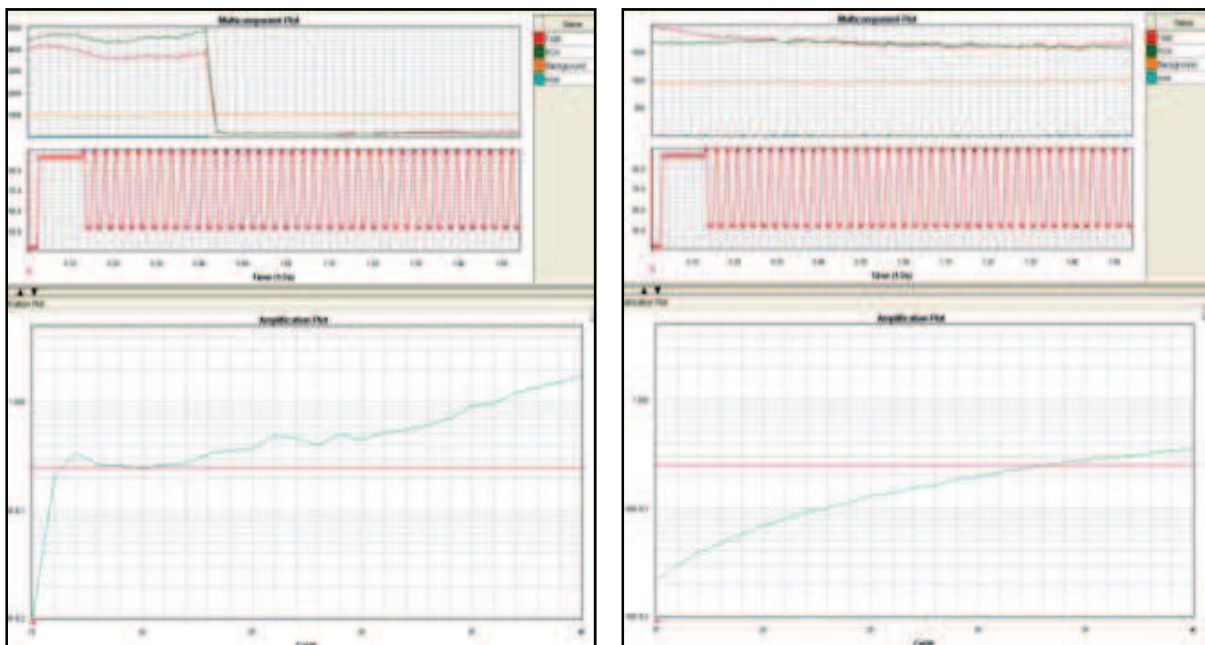


Figure 42 : analyse des intensités brutes de fluorescence avant extraction des données.

Exemples de deux puits anormaux générant une valeur de Ct cohérente (à gauche et à droite). Malgré des intensités brutes de fluorescence ne montrant aucune augmentation (parties supérieures), les algorithmes de calcul intégrant les différents signaux fluorescents génèrent de vraies valeurs de Ct (images extraites à partir du logiciel d'analyse RQ Manager).

Comme nous l'avons décrit dans notre papier méthodologique, la façon dont les données brutes sont extraites peut également être une source de biais générant des données imprécises. Nous avons montré que l'extraction des Cts bruts de plaques TLDA individuelles réalisée à l'aide de seuils automatiquement calculés par le logiciel générait une distribution des Cts en escalier (cf. article I, figure 1C). Cependant, nous n'avons pas détecté ce problème en réalisant des courbes de distribution de manière généralisée, mais plutôt grâce aux graphiques de corrélation de type AM (figure 43). Lors du contrôle qualité initial des données brutes, ce graphique nous a servi également de méthode d'appréciation de la nature des biais. Les nuages de points présentaient des zébrures d'aspect très particulier, traduisant une anomalie de la répartition des données brutes, et une corrélation anormale des deux échantillons testés.

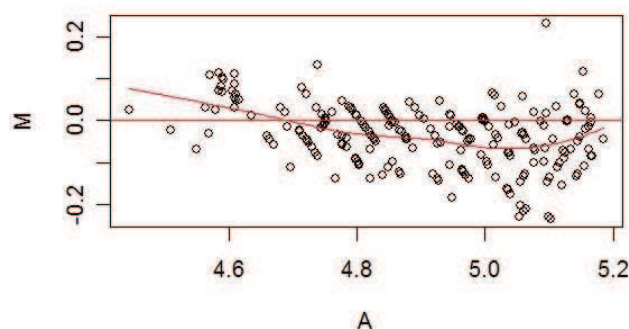


Figure 43 : conséquences du phénomène de distribution des données en escalier.

Exemple de graphique AM réalisé dans le cadre de l'utilisation de la fonction LOESS, montrant une répartition aberrante du nuage de points en arrête de poisson. Les données ont été extraites plaque par plaque, en seuil automatique (image générée par le script R de normalisation LOESS).

En fait, dans ces conditions très particulières (analyse plaque par plaque en seuil automatique), l'algorithme de calcul de ce logiciel aboutit à la création de Ct imprécis, correspondant presque à des arrondis aux valeurs entières supérieures ou inférieures. Ces résultats traduisent vraiment l'importance de réaliser une multitude de contrôles qualité sur ses données, et de ne pas se contenter de suivre à la lettre les

recommandations qui peuvent être données par des solutions commerciales. Il faut toujours garder un esprit critique sur ses données afin d'identifier d'éventuelles anomalies. Malheureusement, de nombreuses études ont été publiées avec des données présentant le biais décrit ci-dessus (Hui et al., 2009; Nissan et al., 2011). Les résultats biologiques générés par ces études sont certainement impactés par ce biais, même s'il est difficile d'en déduire des conséquences précises (figure 44).

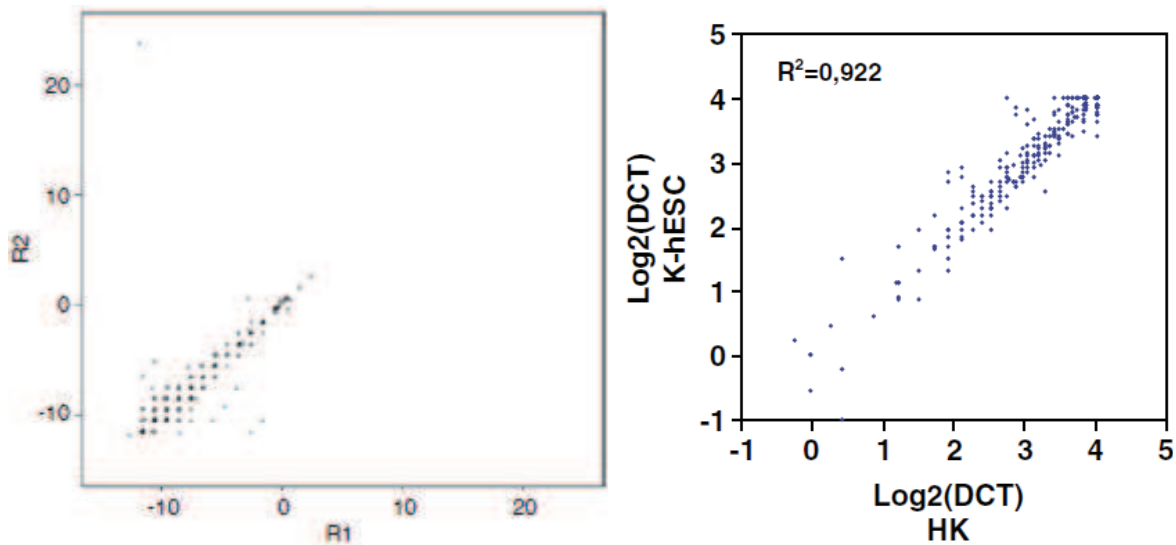


Figure 44 : données de la littérature présentant une distribution en escalier.

Exemples de graphiques de corrélation publiés montrant un biais évident dans la répartition des nuages de points. A gauche, nuage de points publié par l'équipe de Hui, à droite, nuage de points publié par l'équipe de Nissan (figures extraites à partir de (Hui et al., 2009; Nissan et al., 2011)).

Les étapes de normalisation sont également des étapes critiques où il faut corriger efficacement les biais, tout en veillant à ne pas en introduire de nouveaux. Les méthodes de normalisation sont des algorithmes puissants, mais il faut être conscient de leurs limites. Ainsi, nous avons montré que la normalisation par les quantiles était un algorithme particulièrement restructurant, inadaptée à la normalisation des données issues de la plateforme TLDA. En effet, les distributions des données de chaque échantillon sont rendues strictement identiques. Cette

méthode de normalisation ne pouvait donc pas être adaptée à un trop petit nombre de données, car elle est trop transformante. Au contraire, la méthode de normalisation LOESS peut être considérée comme une méthode douce, car elle s'adapte complètement à la forme du nuage de points, et il est possible d'apprécier visuellement le placement de la courbe de régression. Un paramètre de lissage peut être appliqué, et il doit être choisi de telle sorte que la courbe de régression épouse la forme du nuage, sans être affectée par de trop petites variations locales (figure 45). Cette méthode empirique est particulièrement adaptée aux données d'expression issues des plaques TLDA.

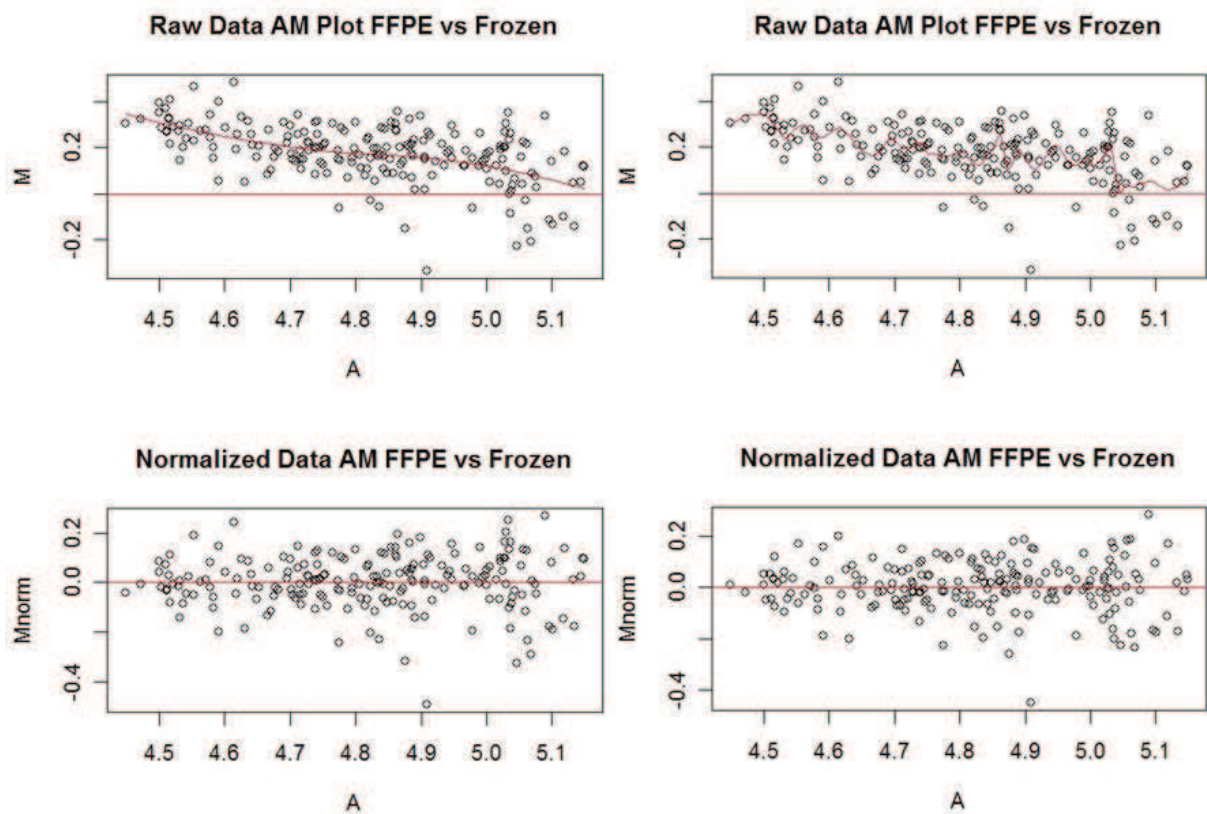


Figure 45 : exemples de courbes LOESS en fonction du paramètre de lissage appliqué.

Colonne de gauche, exemple d'un lissage à 0,9. Colonne de droite, exemple d'un lissage extrême à 0,1. Les deux méthodes semblent corriger efficacement le biais local, mais la méthode de droite introduit des biais supplémentaires en raison d'un paramètre de lissage inadapté, gommant certaines expressions différentielles (graphiques AM réalisés à l'aide du logiciel R).

Au cours de ce travail de thèse, nous avons été confrontés à une difficulté d'interprétation des résultats liée à un autre problème de normalisation, touchant cette fois-ci la plateforme aCGH. Lorsque nous avons réalisé les profils génomiques des RMS, nous avons pour but de les classer dans une des deux principales entités, à savoir dans les ERMS ou dans les ARMS. Les cas de probables ERMS nous ont posé quelques soucis, puisqu'entre nos mains, ces tumeurs ne présentaient que des gains, alors que l'étude publiée par le Dr Olivier Delattre et portant sur une grosse cohorte de 51 ERMS décrivait ces tumeurs comme présentant fréquemment des pertes chromosomiques (figure 46-A) (Williamson et al., 2010).

Il est connu que les puces CGH ne détectent pas les polyploïdies en raison des procédures de normalisation globales qui sont appliquées (Balogh et al., 2011). Ces mêmes procédures de normalisation font que parfois, le niveau de base du profil génomique, correspondant à un nombre de copies normal, n'est pas centré sur la droite des ratios nuls. Cela arrive fréquemment lorsqu'il existe beaucoup de déséquilibres dans un seul sens (du côté des gains ou du côté des pertes). Certains logiciels d'analyse intègrent d'ailleurs des fonctions de « centralisation » permettant de corriger ces biais. Nous avons donc voulu savoir si notre profil d'ERMS était correctement normalisé, puisqu'il semblait possible de lui dessiner une nouvelle ligne de base, qui permettrait alors de le rendre complètement cohérent avec les profils publiés (figure 46-B).

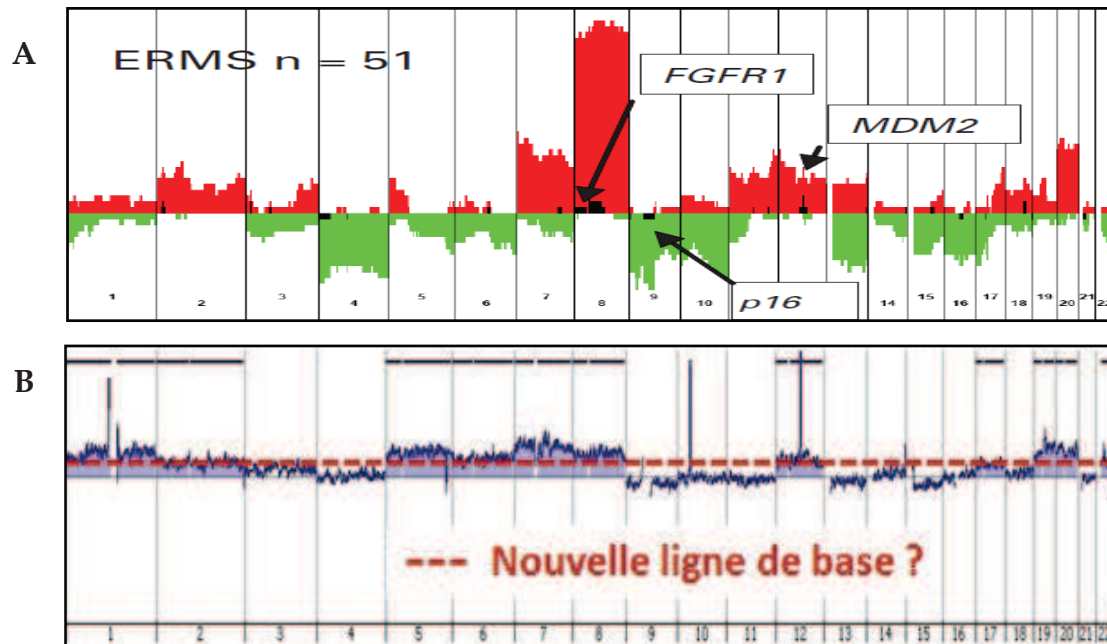


Figure 46 : comparaison de profils aCGH.

A : profils génomiques sommés d'une cohorte de 51 ERMS (Williamson et al., 2010). B : profil génomique de notre cas contrôle d'ERMS. En appliquant une nouvelle ligne de base sur notre cas d'ERMS, les chromosomes étant initialement sur l'ancienne ligne de base se retrouvent en délétion, ce qui correspond alors parfaitement aux anomalies publiées par l'équipe du Dr. Olivier Delattre.

A la suite de ces observations, nous avons réalisé un contrôle du nombre de copies de quelques chromosomes à l'aide des sondes FISH suivantes : DDIT3 de type break-apart en 12q13, FOXO1 de type break-apart en 13q14, et la sonde UroVysion, constituée d'un cocktail de quatre sondes marquant les centromères des chromosomes 3, 7, et 17, et le gène CDKN2A en 9p21. Quelle que soit la sonde utilisée, nous n'avons observé aucune perte de signal. Il semble au contraire exister une polyploïdie hétérogène, certainement liée à des anomalies de ségrégation des chromosomes dans cette tumeur (figure 47).

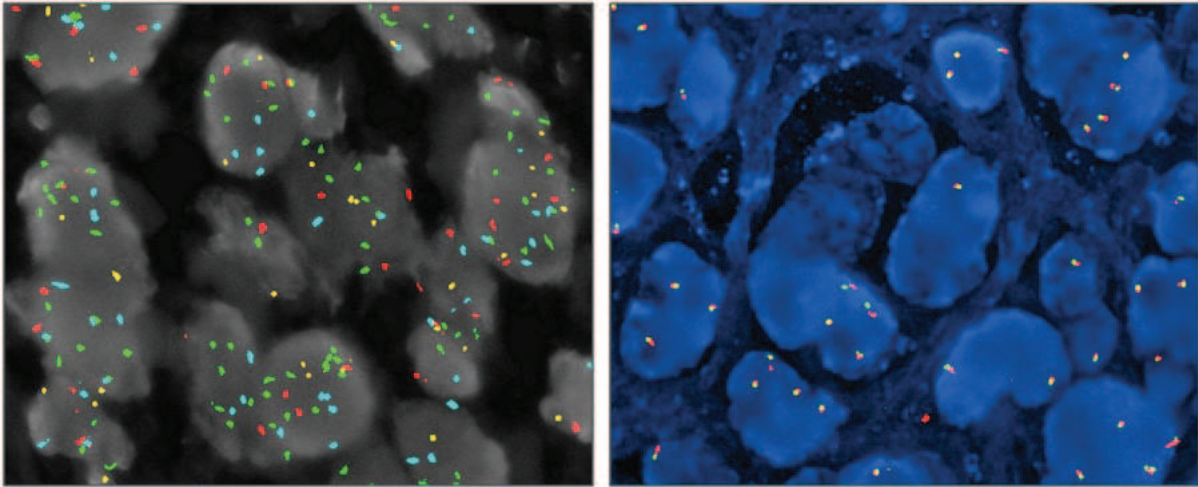


Figure 47 : photos de techniques FISH correspondant au cas contrôle d'ERMS.

A gauche, résultat de l'hybridation à l'aide de la sonde UroVysion : on observe jusqu'à 12 copies du chromosome 7 (sonde verte), et de 2 à 4 copies du chromosome 9 (sonde jaune). A droite, résultat de l'hybridation à l'aide de la sonde FOXO1 de type break-apart, montrant de 2 à 4 signaux de fusion correspondant au nombre de copies du chromosome 13. Au total, les nombres de copies observés sont cohérents avec l'absence de pertes chromosomiques, comme observé sur notre profil CGH correspondant (cf. figure 46-B) (photos réalisées au microscope à fluorescence, au Centre Léon Bérard).

De plus, les études en cytogénétique conventionnelle publiées décrivant des cas d'ERMS confortent toutes nos résultats, puisque les caryotypes décrits sont toujours hyperdiploïdes, et ne comportent que des gains de chromosomes (Gil-Benso et al., 1999; Manor et al., 2012). En technique CGH, les chromosomes à deux copies peuvent donc être vus comme des pertes relatives, comme les résultats d'autres études semblent également le confirmer (Weber-Hall et al., 1996; Bridge et al., 2002).

Enfin, pour revenir aux données issues des plaques TLDA, il existe une dernière étape au cours de laquelle il est possible d'introduire des biais : les analyses statistiques. En effet, ces analyses sont indispensables pour extraire des résultats biologiques à partir de profils d'expression, mais il faut rester prudent quant à leur interprétation, et toujours s'assurer que les conditions pré-requises à l'utilisation d'un test statistique particulier soient remplies (Jafari and Azuaje, 2006). Par exemple, lors d'une analyse supervisée à l'aide d'un T-test, l'algorithme sera toujours capable de trouver des gènes différentiellement exprimés ; toute la difficulté est de bien placer le seuil de *p-value*, pour obtenir des résultats significatifs. En pratique, lorsque l'on travaille sur des tumeurs rares, il est souvent très difficile de réunir un nombre suffisant d'échantillons pour remplir ces conditions. Il faut donc être conscient des limites qui peuvent affecter la robustesse de ses analyses statistiques, et valider si possible ses résultats par d'autres techniques de biologie moléculaire. En ce qui concerne les données générées par les plaques TLDA, il est également possible d'apprécier empiriquement la robustesse des listes de gènes par un retour sur les données normalisées en valeurs de Ct. Sachant qu'une dilution au 1/10^{ème} d'un échantillon correspond à une variation théorique de 3,3 cycles, les valeurs de Ct peuvent nous renseigner sur la significativité des variations observées.

II. LIMITES DES PROFILS D'EXPRESSION DES MICROARN

1. Aléas technologiques et évolutions

La plateforme de TLDA pour microARN présente de nombreux avantages par rapport aux puces microARN (Benes and Castoldi, 2010). Basée sur la PCRq, c'est une technique rapide et sensible. Elle possède également une gamme dynamique considérablement plus grande que celle des puces. Enfin, à la différence des puces, elle ne nécessite que très peu de matériel de départ. Entre nos mains, la plateforme TLDA s'est montrée beaucoup plus robuste que la plateforme de puces (cf. résultats, partie I-2). Cependant, la plateforme TLDA n'est probablement pas exempte de défauts non plus.

Il pourrait exister un premier défaut lié au modèle de conception des plaques TLDA. Les plaques TLDA sont prêtes à l'emploi, puisque chacun des 384 puits contient déjà des amorces et une sonde spécifiques d'un microARN donné. La technologie de micro-canaux utilisée par les TLDA est particulièrement ingénieuse pour faciliter leur remplissage (figure 48). Ainsi, les temps de manipulation sont considérablement réduits, et il n'est pas nécessaire de posséder un robot pour déposer son échantillon d'ADNc dans chacun des 384 puits.

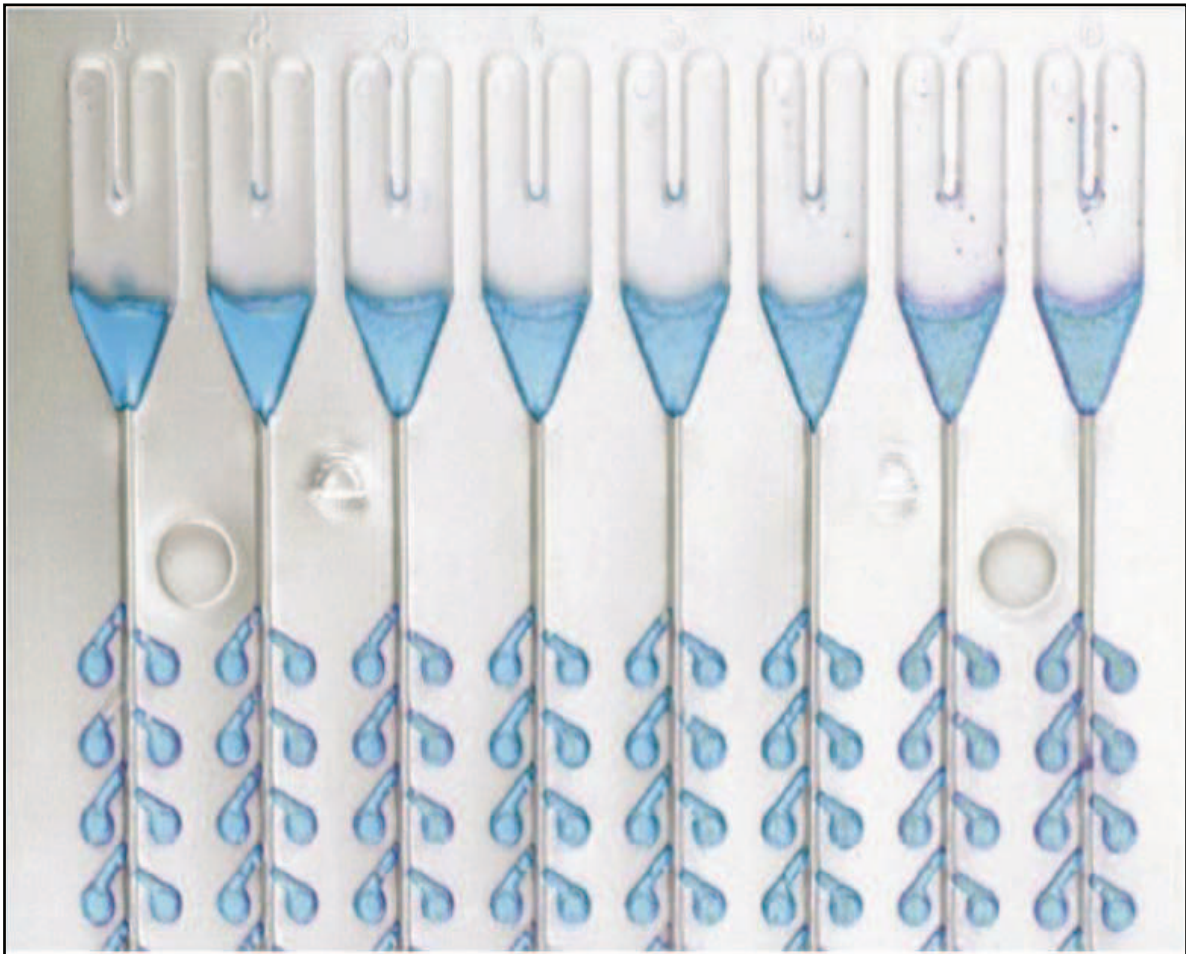


Figure 48 : agrandissement d'une photo montrant la zone de remplissage d'une plaque TLDA.

Exemple d'une plaque déjà remplie et scellée. En dessous de la zone de remplissage, on distingue les huit micro-canaux permettant d'acheminer la solution d'ADNc dans chacun des 384 puits, lors de l'étape de centrifugation.

D'après l'ensemble de nos résultats et de la littérature, cette plateforme donne des résultats robustes. Cependant, un utilisateur est en droit de se demander s'il ne peut pas exister un problème mineur de contamination inter-puits, qui se produirait lors du remplissage par centrifugation de la plaque TLDA. En effet, les puits n'étant évidemment pas scellés à cette étape, une fraction infime de liquide transportant des amorces et des sondes spécifiques pourrait ressortir d'un puits et aller contaminer les puits situés en aval de la même ligne de remplissage (figure 49).

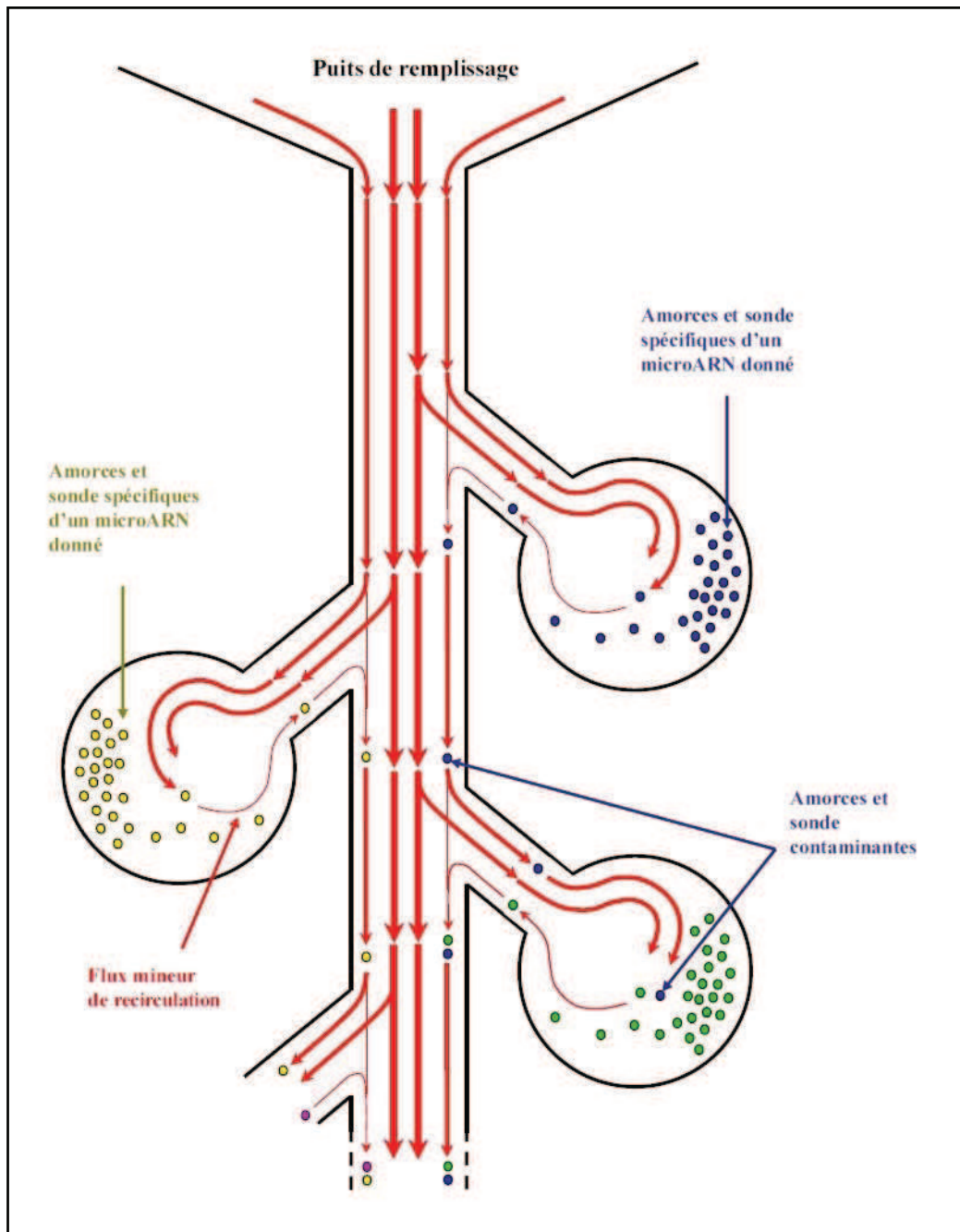


Figure 49 : représentation schématique des flux d'une plaque TLDA lors du remplissage.

Des contaminations inter-puits pourraient éventuellement se produire par transport d'une fraction des amorces et des sondes disposées dans chaque puits vers le flux principal.

D'après cette logique, les premiers puits d'une rangée seraient purs, mais les puits suivants contiendraient une toute petite fraction des amorces et des sondes provenant des puits situés en amont. Nous avons fait part de cette remarque au fabricant des plaques TLDA, qui n'a pas réalisé d'expériences permettant de confirmer ou d'infirmer cette hypothèse de contaminations inter-puits.

Nous avons contacté le Professeur Daniel Hauden de l'Institut FEMTO-ST (UMR 6174), spécialiste en microfluidique et microsystèmes fluidiques, pour lui demander son avis d'expert sur ce sujet. Voici ses commentaires : *« Je suis un peu comme vous assez dubitatif que lorsque le liquide passe dans les cupules rien ne s'échappe vers le flux principal et donc interfère avec les réactions des cupules suivantes. Les dimensions des canaux et des cupules semblent telles que le flux principal est laminaire et déjà la pénétration de liquide dans les cupules doit être "aidée" par un phénomène ou une configuration qui favorise une turbulence au niveau des cupules. Cette turbulence même restreinte suppose un temps de latence du liquide suffisamment long à l'entrée de la cupule pour que le mélange s'obtienne par "diffusion". Qu'en est-il alors de la rétrodiffusion à l'entrée ? »*

Pour valider ou non cette hypothèse de fuites contaminantes, il serait envisageable de réaliser une plaque TLDA à façon, en réalisant des PCRq en duplicats sur deux puits très éloignés sur une même ligne de remplissage. En cas d'obtention de valeurs de Ct trop différentes pour deux puits dupliqués éloignés, par rapport à des puits dupliqués proches, cela prouverait l'existence de contaminations inter-puits. Toutefois, au regard de la cohérence des résultats scientifiques que nous avons obtenus, nous n'avons pas poussé ces investigations plus loin, faute de moyens.

Nous avons également prouvé que ce modèle de conception pouvait entraîner l'émergence de biais stochastiques. En effet, à la fin du remplissage de la plaque TLDA, cette dernière est « scellée » par écrasement des huit canaux de remplissage à l'aide d'un sabot. Un léger dérèglement de ce sabot peut entraîner un mauvais scellement de la plaque TLDA, générant alors des fuites de puits lors de la réaction de RT-PCRq, qui se traduisent ensuite par une absence de Ct ou pire, par des Ct ne suivant aucune logique (figure 50). Bien sûr, un problème de l'ampleur de celui présenté dans la figure 50 est exceptionnel, mais on ne peut exclure la possibilité que de temps en temps, aléatoirement, quelques puits soient mal scellés.

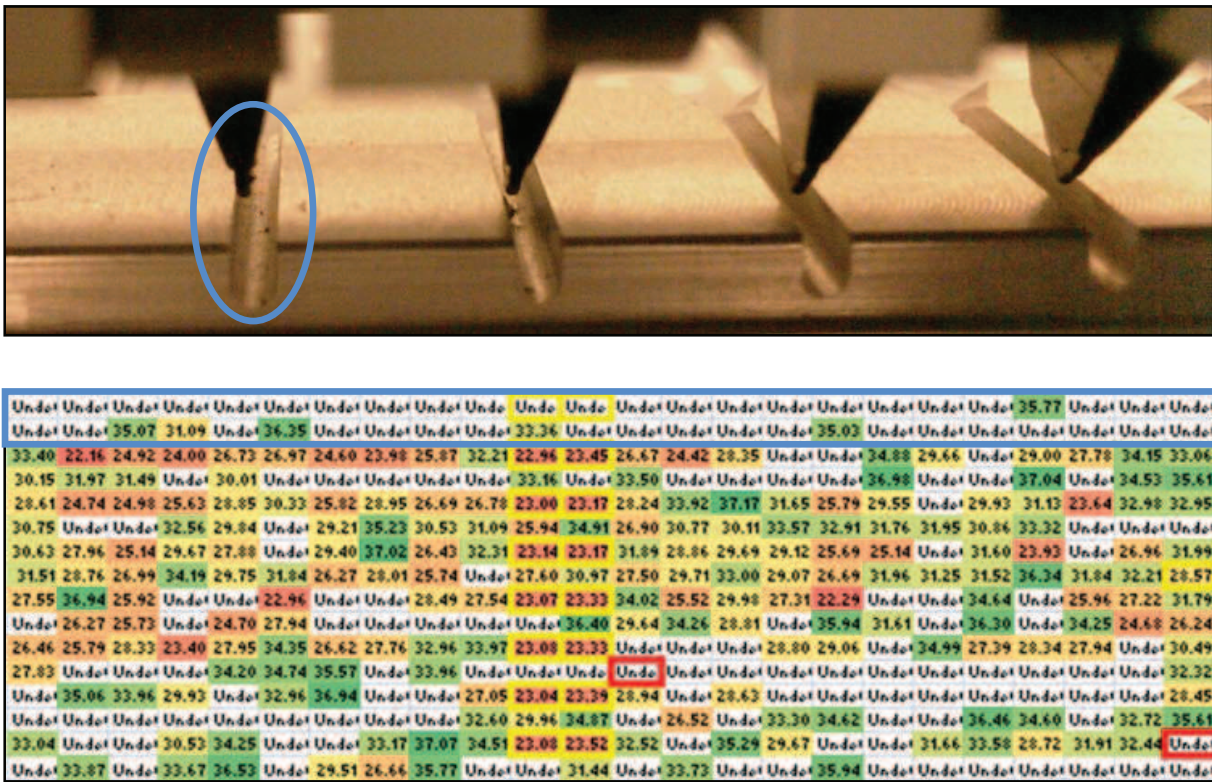


Figure 50 : conséquences du décalage d’une tête du sabot de scellage sur les données brutes.

En haut, photo montrant le décalage de la première tête du sabot par rapport à son rail (tête et rail entourés en bleu). En bas, représentation schématique d’une plaque TLDA, colorisée en fonction des valeurs de Ct. Cette représentation permet de détecter la présence d’anomalies sur la totalité des deux premières lignes, avec de nombreux Ct indéterminés (lignes entourées en bleu). Il s’agit de la conséquence du mauvais scellement du premier des huit canaux de remplissage de la plaque TLDA.

Il pourrait également exister des sources d'imprécision liées aux efficacités des réactions de PCRq, qui émergeraient lors de l'application de la méthode du $\Delta\Delta C_t$. En effet, pour que la méthode de quantification relative soit valide avec l'utilisation de la formule $RQ = 2^{-\Delta\Delta C_t}$, il est nécessaire que les efficacités des PCRq du gène microARN testé et du gène de référence soient égales. La formule précédente est en fait une version simplifiée de la formule exacte $RQ = (1+E)^{-\Delta\Delta C_t}$, où E représente l'efficacité de la PCRq. D'après le fabricant des plaques TLDA, les efficacités des PCRq de tous les microARN et des gènes de contrôle sont égales. Il paraît donc possible d'utiliser la formule simplifiée de la méthode du $\Delta\Delta C_t$, même s'il est difficile de vérifier ces informations.

Lors du développement de notre méthodologie d'analyse des données, nous avons montré que la médiane pouvait être considérée comme un gène de référence donnant de très bons résultats, et nous l'avons ensuite utilisée pour nos études appliquées aux OS et aux RMS. Or, la médiane étant un gène de référence virtuel, elle n'a pas d'efficacité propre quantifiable, mais représente une sorte de moyenne de toutes les efficacités de PCRq de la plaque TLDA. Malgré cela, on peut dire que l'utilisation de la médiane n'a pas engendré d'erreurs majeures de quantification, puisque les résultats obtenus étaient presque toujours meilleurs qu'avec l'utilisation de gènes de référence réels (cf. article I, en soumission au journal *RNA*).

Ainsi, si l'on part du principe que chaque plateforme permettant de générer des profils d'expression possède ses propres biais liés à des caractéristiques purement physiques et techniques, cela pourrait alors expliquer en partie pourquoi il est si difficile de comparer des résultats obtenus sur plusieurs plateformes différentes (Wang et al., 2011a).

De plus, il est bon de rappeler que ces deux plateformes servant à la génération de profils d'expression des microARN sont récentes, et sont toujours en cours de développement. Ainsi, les puces comme les TLDA sont vendues sous forme de versions. Les changements de version sont en partie dus à des évolutions techniques : il peut s'agir d'améliorer certaines sondes présentes sur les puces en jouant sur leur composition en nucléotides LNA, ou dans le cas des TLDA, d'optimiser les séquences des amorces de RT et de PCR pour les rendre plus spécifiques. Enfin, certaines de ces évolutions sont liées à la découverte constante de nouveaux gènes microARN et à la mise à jour consécutive de la miRbase. Ainsi, nous avons commencé notre phase de faisabilité sur la version 1 des TLDA, puis nous sommes rapidement passés à la version 2. Il nous est arrivé de mélanger des plaques TLDA de versions différentes dans certaines analyses de grandes cohortes, mais nous n'avons aucune garantie que les amorces et la sonde d'un puits particulier soient restées les mêmes entre les deux versions, et que cela ne joue pas sur le degré de précision des niveaux d'expression.

Comme nous venons de le voir, lors de l'établissement de profils d'expression, il existe de nombreuses sources d'erreurs potentielles qui pourraient permettre d'expliquer la relative hétérogénéité des résultats de la littérature (Kim et al., 2010b). Malgré ces quelques limites qu'il est bon de rappeler, les plateformes pangénomiques d'établissement de profils d'expression sont des outils formidables pour la découverte de nouveaux biomarqueurs. Elles sont également indispensables pour quantifier des taux d'expression qui présentent souvent des variations relativement faibles. En effet, évaluer le taux d'expression de quelques microARN en simplex est difficile à mettre en œuvre, car les méthodes puissantes de normalisation basées sur l'analyse d'un grand nombre de données ne sont plus applicables. Lorsque le coût de revient de ces techniques chutera, les profils d'expression des microARN pourront sans aucun doute être utilisés en routine pour des applications cliniques.

2. Limites biologiques

2.1. Limites de la classification actuelle des rhabdomyosarcomes

Le diagnostic histologique de RMS est généralement assez facile à réaliser, grâce aux résultats de l'IHC. En effet, lors d'une suspicion de RMS sur des critères morphologiques, deux anticorps ciblant des protéines exprimées lors de la différenciation musculaire sont systématiquement utilisés, et permettent d'affirmer le diagnostic de RMS dans la majorité des cas. En effet, l'anticorps anti-myogénine est spécifique de la lignée musculaire striée, il est toujours positif dans les RMS. L'anticorps anti-desmine est utilisé en confirmation, il est presque toujours positif (Parham et al., 1991; Kumar et al., 2000).

Le principal problème du médecin anatomopathologiste est d'affiner la classification des RMS en les classant dans les deux sous-types principaux, l'ERMS et l'ARMS, en se basant principalement sur les critères définis par la classification de Newton en 1995 (Newton, Jr. et al., 1995), et mise à jour en 2002 par l'OMS (Fletcher et al., 2002). Le pronostic des ARMS étant bien plus défavorable que celui des ERMS, ce classement est fondamental pour l'orientation des patients dans les protocoles thérapeutiques. Cependant, comme nous l'avons montré par notre travail, ce raisonnement dichotomique est certainement trop simpliste.

Il existe une zone grise particulièrement difficile à typer, dont les entités ont été alternativement classées en ARMS ou ERMS au gré de l'évolution des classifications et des pratiques. Cependant, même si la distinction entre ARMS et ERMS reste le principal critère d'inclusion dans un protocole thérapeutique, l'évolution clinique de ces entités est encore mal connue. Aujourd'hui, les médecins anatomopathologistes prennent donc des décisions basées sur leur expertise ainsi que sur l'intégration des différentes techniques complémentaires à leur disposition, sans consensus.

L'étude des résultats des différents protocoles européens de traitement des RMS pédiatriques orchestrés par la SIOP montre bien la difficulté de classement de ces entités de la zone grise, et l'évolution des pratiques (figure 51) (Stevens et al., 2004; Stevens et al., 2005b).

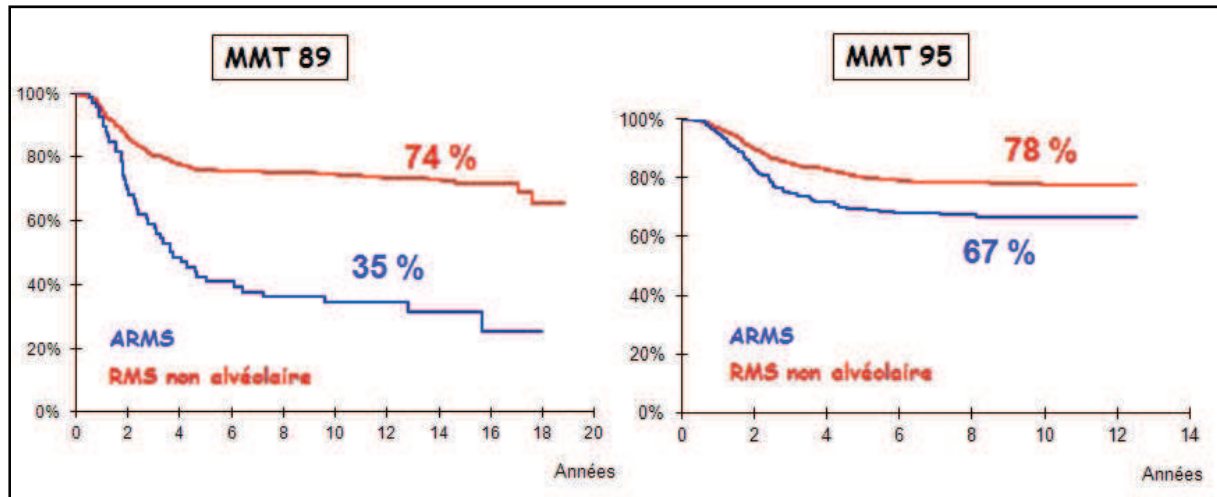


Figure 51 : comparaison dans le temps des courbes de survie globale des RMS.

Courbes de survie globale générées à partir des données cliniques des études *Malignant mesenchymal tumors* (MMT) 89, portant sur 518 patients (à gauche), et MMT 95, portant sur 813 patients (à droite). En rouge, les courbes de survie des RMS non alvéolaires, en bleu, les courbes de survie des RMS alvéolaires (Stevens et al., 2004; Stevens et al., 2005b).

Les protocoles de traitement des formes alvéolaires n'ayant que peu évolué entre 1989 et 1995, le rapprochement des courbes de survie entre ces deux périodes est très probablement dû à l'ajout de formes moins agressives à la cohorte des ARMS. En effet, les critères de classification ont été modifiés entre ces deux études : les formes mixtes, qui étaient autrefois classées dans les ARMS seulement lorsque plus de 50 % du prélèvement présentait une histologie de type alvéolaire (selon la classification SIOP), sont désormais classées en ARMS dès lors qu'une zone, quelle que soit son étendue, présente des caractéristiques de type alvéolaire (selon la classification de Newton) (Parham and Ellison, 2006).

Il est intéressant de noter que sur les 210 RMS de la cohorte présentée dans l'étude publiée par Dr Olivier Delattre (Williamson et al., 2010), presque 20 % des échantillons ont été classés en nARMS sur la base de la morphologie et des résultats de la recherche de transcrits. En comparaison, parmi les 422 RMS de la cohorte du projet RMS2005 que nous utilisons actuellement pour compléter notre étude, 47 cas (< 12 %) ont été placés dans la zone grise en raison d'une morphologie difficile ou de résultats discordants entre morphologie et biologie moléculaire ; parmi ceux-ci, seuls 11 cas (< 3 %) ont été classés formellement en nARMS par notre médecin anatomopathologiste expert, le Dr. Dominique Ranchère-Vince. Ces chiffres sont difficilement comparables, mais ils soulignent les difficultés diagnostiques inhérentes aux entités de la zone grise entre ARMS et ERMS, et le fait que certains critères tenus retenus par les médecins anatomopathologistes ne font pas consensus.

Il est aujourd'hui admis par la plupart des pathologistes référents mondiaux qu'une zone grise existe entre les ERMS et les ARMS, mais il n'existe pas encore de réflexion commune pour améliorer la prise en charge de ces entités rares et difficiles. Il convient de colliger ces cas et d'en faire une caractérisation moléculaire exhaustive, car une meilleure connaissance de leur biologie et de leur évolution clinique ne sera possible que par leur démembrement avec l'aide de techniques fines de biologie moléculaire.

2.2. Complexité génomique des ostéosarcomes

Comme nous l'avons vu dans la partie introductive, les OS sont des tumeurs agressives, dont l'étiologie reste encore inconnue. Les OS sont avant tout caractérisés par une complexité génomique importante et hétérogène, puisqu'il est très difficile de définir des anomalies récurrentes (Sadikovic et al., 2008). Des mécanismes d'instabilité génomiques pourraient être à l'origine de cette complexité.

Partant de ces observations dans les OS, il est intéressant de rentrer plus en détail dans la composition des signatures d'expression des microARN que nous avons identifiées comme étant discriminantes vis-à-vis de la réponse à la l'IFO, selon les modèles d'étude (cf. article II) (Gougelet et al., 2011b) (Tableau 3).

Modèle d'étude	Rat	Homme	Lignées d'OS
Nombre de tumeurs	9	24	2
Bagage génétique	Similaire	Différent	Différent
Signature MicroARN	let-7c ↑	-	-
	miR-15b ↑	-	miR-15b ↑
	miR-19b ↑	-	miR-19b ↑
	miR-24 ↑	-	miR-24 ↑
	miR-29c ↑	-	miR-29c ↑
	miR-30c ↑	-	miR-30c ↑
	miR-30d ↑	-	-
	miR-92a ↑	miR-92a ↑	-
	miR-93 ↑	-	miR-93 ↑
	miR-99b ↑	miR-99b ↑	miR-99b ↑
	-	miR-132 ↓	miR-132 ↓
	miR-142-3p ↑	-	-
	miR-152 ↑	-	-
	miR-193a ↑	miR-193a ↑	miR-193a ↑
	-	miR-422a ↑	-

Tableau 3 : synthèse des différentes signatures microARN retrouvées comme discriminantes pour la réponse à la chimiothérapie, en fonction du modèle d'étude.

↑ = gène surexprimé dans les tumeurs sensibles, ↓ = gène sous-exprimé dans les tumeurs sensibles.

Nous avons tout d'abord travaillé sur le modèle de rat, chez lequel nous avons trouvé une liste de 10 microARN différentiellement exprimés, à partir d'une cohorte de 9 tumeurs possédant toutes le même bagage génétique, puisqu'issues d'une même tumeur originelle radio-induite (Allouche et al., 1980). En analysant une cohorte de 24 OS humains, nous avons identifié une nouvelle liste de 5 microARN différentiellement exprimés. Trois d'entre eux ont finalement été retrouvés comme différentiellement exprimés chez le rat, en appliquant un seuil de *p-value* moins stringeant. Enfin, en travaillant sur les deux lignées cellulaires humaines d'OS, nous

avons retrouvé 9 microARN différentiellement exprimés appartenant à la signature trouvée chez le rat ou chez l'homme.

En travaillant initialement sur le rat, cela nous a permis de réduire considérablement la variabilité inter-tumorale, puisque toutes les tumeurs testées possédaient le même bagage génétique initial. La signature microARN trouvée chez le rat a été modifiée lors de notre passage chez l'homme, dans les tumeurs sporadiques comme dans les lignées cellulaires d'OS. Deux hypothèses permettent d'expliquer ce changement de signature d'expression. Il est possible que ce changement soit dû à la grande hétérogénéité génomique des OS, comme décrite dans la littérature. Ainsi, en passant d'un modèle d'étude simple à un modèle d'étude hétérogène, il est probable que nous soyons passés de la signature spécifique d'une tumeur « unique », à une signature plus universelle correspondant aux tumeurs de type OS. Cependant, il est également possible que les signatures d'expression soient différentes en raison du choix d'un modèle de rat très particulier, puisque les tumeurs sont générées à partir d'une tumeur originelle radio-induite. En effet, il a déjà été montré que les tumeurs radio-induites pouvaient présenter des anomalies génomiques spécifiques aux radiations, différentes des tumeurs sporadiques (Guo et al., 2011; Dom et al., 2012).

Il est intéressant de remarquer que dans le modèle de rat, des tumeurs ayant initialement le même bagage génétique se divisent finalement en deux groupes, en fonction de leur sensibilité à la chimiothérapie. Cette observation est toutefois cohérente avec l'hypothèse selon laquelle la tumorigenèse des OS serait due à une déficience dans les mécanismes de stabilité chromosomique ou de réparation de l'ADN, menant ensuite à d'autres anomalies désordonnées (Khanna, 2008). Grâce à l'analyse des profils génomiques des tumeurs de rat par technique CGH, nous avons également prouvé que la dérégulation des microARN en fonction de la sensibilité à la chimiothérapie n'était pas liée à des anomalies génomiques de type variations du nombre de copies, qui seraient divergentes entre les deux groupes de tumeurs. Pour expliquer ce phénomène, il doit donc exister d'autres mécanismes biologiques à un

niveau différent, tels que des mutations ponctuelles, des modifications épigénétiques, ou encore un rôle de l'environnement tumoral. Il serait intéressant de regarder le statut de méthylation de l'ADN autour de la région comprenant le gène miR-193a. En effet, ce microARN était surexprimé dans les tumeurs chimiosensibles dans nos trois modèles d'étude, alors qu'il a été récemment décrit comme étant méthylé et donc inactivé dans les carcinomes hépatocellulaires résistants à la chimiothérapie (Ma et al., 2012a). L'hypothèse d'un rôle de l'environnement tumoral ne peut pas non plus être écartée, car certains microARN différentiellement exprimés ciblent des voies métaboliques impliquées dans la réponse immunitaire, comme l'avait montré l'équipe de S. Salas dans le même contexte, à partir de puces d'expression des ARNm (Salas et al., 2009).

Malgré des signatures d'expression divergentes entre les modèles, nous avons également montré que toutes ces signatures ciblaient potentiellement les mêmes voies métaboliques, en particulier la voie de signalisation du calcium, des MAP kinases, du TGF β et de Wnt. Ces résultats sont cohérents avec le reste de la littérature, puisqu'il a déjà été montré que ces voies métaboliques étaient d'une part impliquées dans l'ostéolyse et l'ostéogénèse (Kingsley et al., 2007), et d'autre part dérégulées dans les OS pédiatriques résistants à la chimiothérapie (Mintz et al., 2005).

Nous avons également montré que les signatures d'expression des microARN permettant de différencier les tumeurs résistantes des tumeurs sensibles étaient dépendantes du type de médicament utilisé, qu'il s'agisse de chimiothérapies ou de thérapies ciblées (cf. article III) (Gougelet et al., 2011a). Tout cela traduit bien la complexité biologique des OS. Pour ces tumeurs, dans un futur proche, la solution passera très certainement par une thérapie à la carte, non pas en fonction du sous-type d'OS, mais en fonction de chaque couple tumeur-patient. Dans ce but, il convient dès aujourd'hui d'intégrer le maximum de techniques de génomique fonctionnelle et d'adopter une approche de biologie systémique (Tian et al., 2012).

III. VOIES PEU EXPLORÉES ET PERSPECTIVES

Lors de ce travail de thèse, nous avons validé la plateforme TLDA comme moyen d'établissement de profils d'expression fiables, via une méthodologie d'analyse rigoureuse développée en interne. Les résultats obtenus sur nos deux projets plus appliqués d'analyse des profils d'expression des microARN chez les OS et les RMS sont prometteurs et nous encouragent à poursuivre nos investigations.

1. Profils d'expression des microARN et voies métaboliques

Ces dernières années, l'analyse des voies métaboliques est devenue indispensable pour l'interprétation biologique des techniques moléculaires à haut débit (Khatri et al., 2012). En effet, typiquement, ces techniques permettent d'établir une ou plusieurs listes de gènes différentiellement exprimés, et il est indispensable de pouvoir les replacer dans un contexte métabolique pour connaître leurs implications dans la biologie des échantillons analysés.

Comme les gènes classiques, les microARN sont dérégulés dans de nombreux types tumoraux (Tie and Fan, 2011), et peuvent agir en tant qu'oncogènes ou gènes suppresseurs de tumeurs (Zhang et al., 2007). Cependant, leur intégration sur des voies métaboliques est moins directe, puisqu'il est tout d'abord indispensable de faire appel à des logiciels de prédiction de cibles. De plus, cette intégration est très complexe à réaliser, puisque chaque microARN peut posséder une centaine de cibles, et qu'une cible donnée peut être régulée par plusieurs microARN.

La plupart des outils d'analyses des voies métaboliques utilisés pour exploiter des données d'expression sont basés sur les répertoires en ligne de la Gene Ontology (GO) ou de la Kyoto Encyclopedia of Genes and Genomes (KEGG) (Ashburner et al., 2000; Kanehisa et al., 2012). Afin de pouvoir utiliser ces outils, nous avons tout d'abord développé une base de données comportant les résultats des quatre principaux logiciels de prédiction de cibles (MiRanda, TargetScan, PicTar et la miRBase), que nous mettons à jour régulièrement (figure 52).

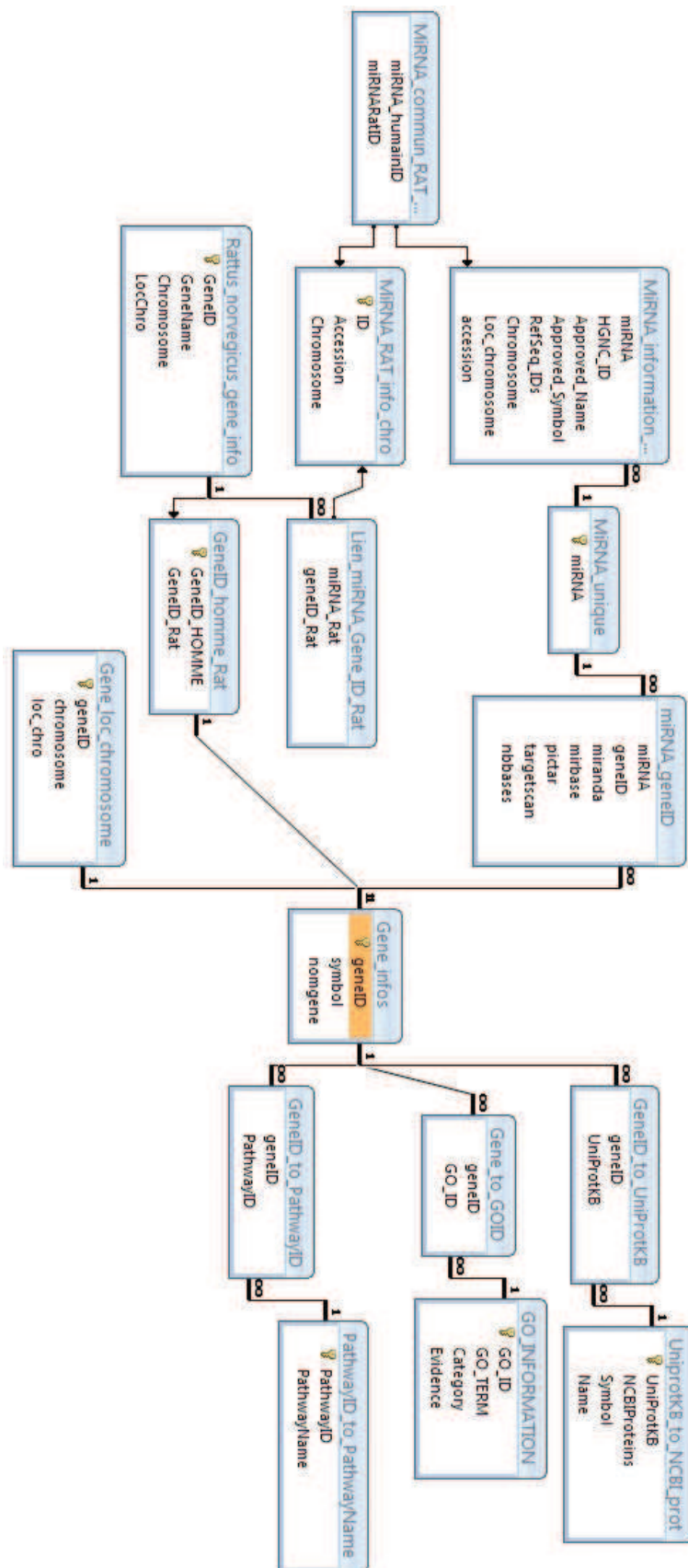


Figure 52 : schéma de la base de données de prédiction de cibles et d'annotations ontologiques.

(Base de données développée par A. Besse, au Centre Léon Bérard).

Cet outil nous a déjà permis d'extraire des résultats biologiques intéressants (figure 53), mais nous pensons qu'il est encore possible de l'améliorer. Nous avons quelquefois été confrontés au placement de plusieurs microARN différemment exprimés, qui ciblaient un même ARNm avec des effets antagonistes (un des microARN était surexprimé, l'autre sous-exprimé). A l'inverse, il arrive aussi que plusieurs microARN ciblant le même ARNm puissent avoir des effets cumulatifs. Il existe donc actuellement une difficulté d'intégration des niveaux d'expression conjointement aux résultats de prédiction de cibles, lorsque plusieurs microARN partagent une cible commune.

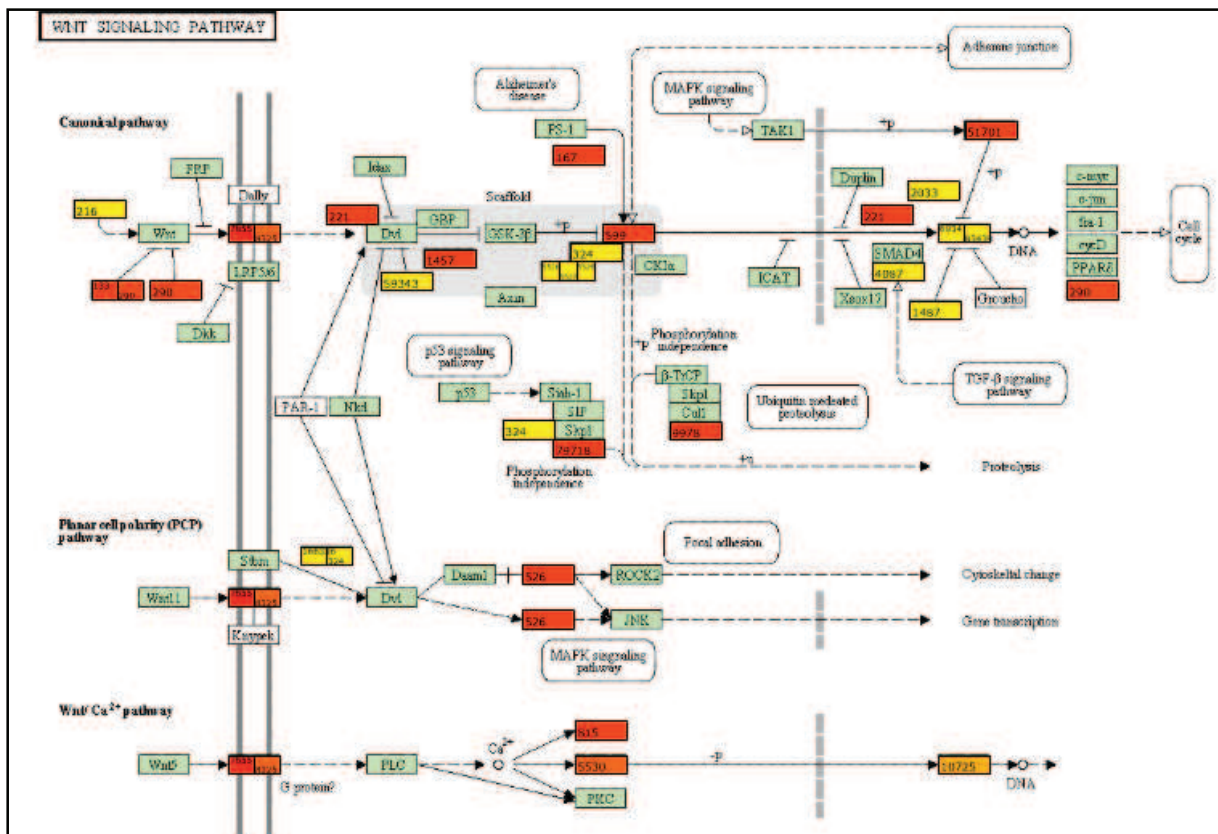


Figure 53: exemple d'intégration des microARN différemment exprimés entre OS chimiosensibles et OS chimiorésistants, sur la voie métabolique de Wnt.

Seules les cibles prédites par au moins deux des logiciels de prédiction sont conservées. Les carrés jaunes représentent les cibles des microARN faiblement exprimés, et les carrés rouges représentent les cibles fortement réprimées par les microARN surexprimés (données générées par A. Besse, au Centre Léon Bérard).

Afin de prendre en compte ces considérations, lorsque plusieurs microARN partagent une cible commune, nous souhaiterions ajouter dans notre base de données les paramètres d'affinités donnés par les logiciels de prédiction de cibles. Ainsi, lors d'une compétition entre plusieurs microARN pour une même cible, nous pourrions intégrer et pondérer l'effet de chaque microARN. Le principe de cette approche est facile à comprendre grâce à l'exemple suivant :

- Soit un ARNm donné, ciblé par deux microARN différentiellement exprimés et antagonistes miR-a et miR-b.
- On estime que miR-a est surexprimé d'un niveau A et possède une affinité pour sa cible égale à 70 %, et miR-b est sous exprimé d'un niveau B et son affinité est de 20 % pour la même cible.
- On peut alors déduire l'effet général de ces deux microARN sur l'ARNm cible par le calcul suivant :

$$\text{Effet d'inhibition intégré} = A \times 0,7 - B \times 0,2$$

A notre connaissance, cette approche n'a jamais été évaluée, même si des études récentes ont souligné les difficultés d'interprétation du réseau complexe que forment les microARN (Yang et al., 2012). Elle nous paraît être fondamentale pour la compréhension générale des dérégulations croisées des voies métaboliques.

2. Développement d'outils cliniques de routine

Grâce aux profils d'expression des microARN réalisés sur plaques TLDA, nous avons extrait des signatures d'expression corrélées à la classification diagnostique des RMS, et à la prédiction de la réponse des OS à la chimiothérapie. L'étape suivante consiste tout naturellement à essayer de détecter ces signatures d'expression par des techniques de routine, moins onéreuses. Les plaques TLDA étant juste une façon de réaliser de la PCRq à moyen débit, les résultats obtenus sont directement

transférables dans des techniques de PCRq simplex. Dans ce cas, comme nous ne pouvons plus normaliser nos données à l'aide de la médiane, nous réfléchissons actuellement au meilleur compromis concernant le choix d'un ou plusieurs gènes de ménage pour la normalisation des ces expériences de PCRq simplex. Ce choix est très difficile, car il n'existe pas de gènes de ménage idéal. Nous avons en effet montré qu'ils sont très dépendants de la composition de la cohorte analysée (cf article I, en soumission au journal *RNA*).

Une autre application pourrait être la quantification des niveaux d'expression de façon non invasive pour le patient, à partir d'échantillons de sérum (Zampetaki et al., 2012; Allegra et al., 2012). Récemment, des avancées technologiques ont été faites pour permettre la détection de très petites quantités de microARN (Chapin and Doyle, 2011; Wang et al., 2011b).

Nous avons également en projet de tester les niveaux d'expression des microARN d'intérêt *in situ*, par technique d'hybridation fluorescente, car de nombreuses études ont récemment prouvé qu'il était possible de les détecter de manière précise à partir de matériel FFPE (Soe et al., 2011; Shi et al., 2012).

Enfin, si les approches précédentes ne donnent pas satisfaction, il sera encore possible d'essayer de quantifier par IHC les protéines cibles des microARN différentiellement exprimés. Par exemple, l'équipe de J. Zavadil a réalisé une étude très complète portant sur la signature microARN des léiomyomes utérins (Zavadil et al., 2010). A partir des profils d'expression des microARN, ils ont montré que les profils d'expression protéiques associés étaient inversement corrélés, par technique d'IHC sur TMA.

De manière générale, une application clinique à plus large échelle nous permettrait de valider définitivement nos résultats.

3. Poursuite générale des projets en cours

Même si nous avons prouvé que les profils d'expression des microARN permettaient de répondre à de nombreuses questions cliniques dans les OS et les RMS, leur établissement est un travail de longue haleine, qu'il nous faut encore poursuivre.

En effet, la complexité génomique et l'hétérogénéité des OS nous poussent à augmenter la taille de nos cohortes d'analyse, et à réaliser des études fonctionnelles supplémentaires dans des modèles d'étude simplifiés (par exemple, le modèle de rat et les lignées cellulaires humaines). Sur ces modèles, nous essaierons en particulier de restaurer la sensibilité à la chimiothérapie par des approches de siARN, avec pour but ultime le développement d'une nouvelle stratégie thérapeutique ciblée.

En ce qui concerne l'étude des RMS, les principales limitations que nous avons rencontrées viennent du fait que les tumeurs difficiles que nous souhaiterions démembrer sont très rares, précieuses pour le diagnostic et de petite taille. S'agissant de tumeurs pédiatriques, souvent difficiles d'accès (sphère oto-rhino-laryngée), les microbiopsies sont fréquemment privilégiées. Afin d'affiner nos résultats, nous sommes contraints de poursuivre cette étude de manière prospective sur du plus long terme. Ce travail de thèse nous a permis de développer dans notre laboratoire de nombreux outils, de façon séquentielle : TLDA, puces protéiques, FISH PAX3 *break-apart* et CGH. Nous sommes également en train de valider une nouvelle approche de RT-PCRq pour la recherche des transcrits de fusion rares.

Nous avons appris à gérer au mieux les prélèvements précieux de RMS pour réaliser toutes ces techniques en parallèle, à partir du matériel restant après diagnostic, et sans l'épuiser. En effet, nous pouvons réaliser la plupart de ces techniques à partir d'un prélèvement microbiopsique FFPE coupé sur quelques lames blanches, à l'exception des profils protéiques. En ce qui concerne les cas de

RMS difficiles de la zone grise, nous approfondirons plus particulièrement les formes mixtes (les formes présentant un contingent de type ERMS et un contingent à petites cellules rondes beaucoup plus indifférenciées), dans le but de mieux comprendre le phénomène d'hétérogénéité tumorale. Par macrodissection sur lame blanche, il sera possible de déterminer si les profils d'expression des microARN (TLDA) et les profils génomiques (aCGH) sont différents. Nous avons également mis au point une technique de FISH permettant de réutiliser des coupes colorées HES ou les lames d'IHC, ce qui nous permettra de faire des corrélations fines entre les deux zones d'intérêt (figure 54). Cette technique pourra être utilisée pour confirmer des résultats génomiques de CGH, ou pour des expériences de FISH sur microARN.

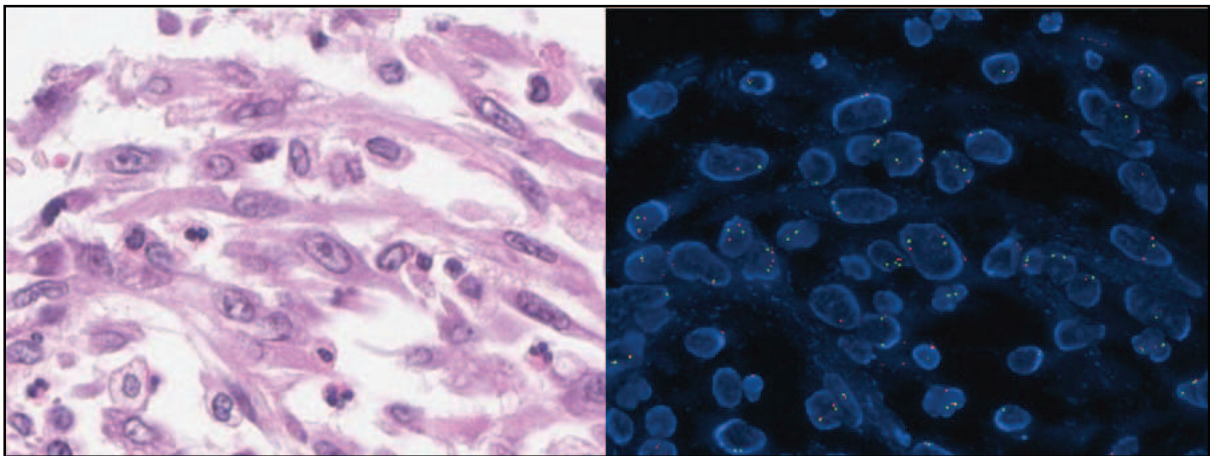


Figure 54 : corrélation entre coupe colorée HES et technique FISH, réalisés sur une même lame.

Exemple de corrélation entre une coupe colorée HES (à gauche) et cette même coupe réutilisée pour une technique FISH (à droite), à l'aide de la sonde correspondant au gène MDM2. La sonde correspondant au gène MDM2 est verte, le contrôle centromérique du chromosome 12 est rouge (photo de lame HES extraite à partir d'une sauvegarde sous forme de lame virtuelle, recadrée sur le même champ que la photo de FISH réalisée au microscope à fluorescence, au grossissement x63).

Enfin, nous projetons d'appliquer nos méthodes d'analyse des profils d'expression des microARN à d'autres pathologies. Dans les léiomyosarcomes, nous essaierons de compléter les profils protéiques que nous avons obtenus (cf. article IV, en soumission à *EJC*) par l'étude des profils d'expression des microARN (TLDA), et

l'étude des profils génomiques (aCGH en cours d'analyse). Dans les GIST, nous disposons déjà d'une grande cohorte analysée en plaques TLDA (cf. article I, en soumission à *RNA*). La plupart de ces tumeurs étant incluses dans l'essai clinique multi-centrique BFR14, nous aurons bientôt accès aux données cliniques et nous pourrons alors essayer de trouver une signature d'expression relative aux différentes mutations du gène KIT. De plus, dans les tumeurs desmoïdes, nous sommes actuellement en train de montrer que les profils d'expression des microARN ont une valeur pronostique.

Enfin, disposant d'une plateforme CGH compatible avec les puces d'expression ARNm, nous n'excluons pas la possibilité d'analyser le transcriptome de tumeurs pour certaines de ces pathologies, afin d'avoir la vue la plus complète possible de la biologie des systèmes de ces tumeurs et des phénomènes de régulations en jeu. L'intégration de toutes les anomalies moléculaires retrouvées sur des voies métaboliques sera alors un nouveau défi à relever.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Abruzzo,L.V., Lee,K.Y., Fuller,A., Silverman,A., Keating,M.J., Medeiros,L.J., and Coombes,K.R. (2005). Validation of oligonucleotide microarray data using microfluidic low-density arrays: a new statistical method to normalize real-time RT-PCR data. *Biotechniques* 38, 785-792.
- Ahrendt,S.A., Halachmi,S., Chow,J.T., Wu,L., Halachmi,N., Yang,S.C., Wehage,S., Jen,J., and Sidransky,D. (1999). Rapid p53 sequence analysis in primary lung cancer using an oligonucleotide probe array. *Proc. Natl. Acad. Sci. U. S. A* 96, 7382-7387.
- Ali,H.M., Urbinati,G., Raouane,M., and Massaad-Massade,L. (2012). Significance and applications of nanoparticles in siRNA delivery for cancer therapy. *Expert. Rev. Clin. Pharmacol.* 5, 403-412.
- Allegra,A., Alonci,A., Campo,S., Penna,G., Petrungaro,A., Gerace,D., and Musolino,C. (2012). Circulating microRNAs: New biomarkers in diagnosis, prognosis and treatment of cancer (Review). *Int. J. Oncol.*
- Allouche,M., Delbruck,H.G., Klein,B., Masse,R., Matar,A., Morin,M., Lafuma,J., and Jasmin,C. (1980). Malignant bone tumours induced by a local injection of colloidal radioactive ¹⁴⁴Cerium in rats as a model for human osteosarcomas. *Int. J. Cancer* 26, 777-782.
- Altman,D.G. (1991). *Practical Statistics for Medical Research*. (London.
- Amendola,B.E., Amendola,M.A., McClatchey,K.D., and Miller,C.H., Jr. (1989). Radiation-associated sarcoma: a review of 23 patients with postradiation sarcoma over a 50-year period. *Am. J. Clin. Oncol.* 12, 411-415.
- Andersen,C.L., Jensen,J.L., and Orntoft,T.F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res.* 64, 5245-5250.
- Anderson,J., Gordon,A., McManus,A., Shipley,J., and Pritchard-Jones,K. (1999a). Disruption of imprinted genes at chromosome region 11p15.5 in paediatric rhabdomyosarcoma. *Neoplasia.* 1, 340-348.
- Anderson,J., Gordon,A., Pritchard-Jones,K., and Shipley,J. (1999b). Genes, chromosomes, and rhabdomyosarcoma. *Genes Chromosomes. Cancer* 26, 275-285.
- Anderson,J., Ramsay,A., Gould,S., and Pritchard-Jones,K. (2001). PAX3-FKHR induces morphological change and enhances cellular proliferation and invasion in rhabdomyosarcoma. *Am. J. Pathol.* 159, 1089-1096.
- Ando,K., Mori,K., Verrecchia,F., Marc,B., Redini,F., and Heymann,D. (2012). Molecular alterations associated with osteosarcoma development. *Sarcoma.* 2012, 523432.
- Aoki,Y., Niihori,T., Kawame,H., Kurosawa,K., Ohashi,H., Tanaka,Y., Filocamo,M., Kato,K., Suzuki,Y., Kure,S., and Matsubara,Y. (2005). Germline mutations in HRAS proto-oncogene cause Costello syndrome. *Nat. Genet.* 37, 1038-1040.
- Arndt,C.A. and Crist,W.M. (1999). Common musculoskeletal tumors of childhood and adolescence. *N. Engl. J. Med.* 341, 342-352.

- Arroyo,J.D., Chevillet,J.R., Kroh,E.M., Ruf,I.K., Pritchard,C.C., Gibson,D.F., Mitchell,P.S., Bennett,C.F., Pogosova-Agadjanyan,E.L., Stirewalt,D.L., Tait,J.F., and Tewari,M. (2011). Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc. Natl. Acad. Sci. U. S. A* 108, 5003-5008.
- Ashburner,M., Ball,C.A., Blake,J.A., Botstein,D., Butler,H., Cherry,J.M., Davis,A.P., Dolinski,K., Dwight,S.S., Eppig,J.T., Harris,M.A., Hill,D.P., Issel-Tarver,L., Kasarskis,A., Lewis,S., Matese,J.C., Richardson,J.E., Ringwald,M., Rubin,G.M., and Sherlock,G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25-29.
- Ayroles,J.F. and Gibson,G. (2006). Analysis of variance of microarray data. *Methods Enzymol.* 411, 214-233.
- Bacci,G., Mercuri,M., Longhi,A., Ferrari,S., Bertoni,F., Versari,M., and Picci,P. (2005). Grade of chemotherapy-induced necrosis as a predictor of local and systemic control in 881 patients with non-metastatic osteosarcoma of the extremities treated with neoadjuvant chemotherapy in a single institution. *Eur. J. Cancer* 41, 2079-2085.
- Balogh,Z., Szemlaky,Z., Szendroi,M., Antal,I., Papai,Z., Fonyad,L., Papp,G., Changchien,Y.C., and Sapi,Z. (2011). Correlation between DNA ploidy, metaphase high-resolution comparative genomic hybridization results and clinical outcome of synovial sarcoma. *Diagn. Pathol.* 6, 107.
- Barad,O., Meiri,E., Avniel,A., Aharonov,R., Barzilai,A., Bentwich,I., Einav,U., Gilad,S., Hurban,P., Karov,Y., Lobenhofer,E.K., Sharon,E., Shibolet,Y.M., Shtutman,M., Bentwich,Z., and Einat,P. (2004). MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. *Genome Res.* 14, 2486-2494.
- Barr,F.G. (2001). Gene fusions involving PAX and FOX family members in alveolar rhabdomyosarcoma. *Oncogene* 20, 5736-5746.
- Barr,F.G., Fitzgerald,J.C., Ginsberg,J.P., Vanella,M.L., Davis,R.J., and Bennicelli,J.L. (1999). Predominant expression of alternative PAX3 and PAX7 forms in myogenic and neural tumor cell lines. *Cancer Res.* 59, 5443-5448.
- Barr,F.G., Galili,N., Holick,J., Biegel,J.A., Rovera,G., and Emanuel,B.S. (1993). Rearrangement of the PAX3 paired box gene in the paediatric solid tumour alveolar rhabdomyosarcoma. *Nat. Genet.* 3, 113-117.
- Barr,F.G., Qualman,S.J., Macris,M.H., Melnyk,N., Lawlor,E.R., Strzelecki,D.M., Triche,T.J., Bridge,J.A., and Sorensen,P.H. (2002). Genetic heterogeneity in the alveolar rhabdomyosarcoma subset without typical gene fusions. *Cancer Res.* 62, 4704-4710.
- Baskerville,S. and Bartel,D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. *RNA.* 11, 241-247.
- Baum,E.S., Gaynon,P., Greenberg,L., Krivit,W., and Hammond,D. (1981). Phase II trial cisplatin in refractory childhood cancer: Children's Cancer Study Group Report. *Cancer Treat. Rep.* 65, 815-822.
- Beißbarth, T. Microarray: Quality Control, Normalization and Design. 2005. 20-10-2012.
Ref Type: Online Source

- Bellis,M. and Casellas,P. (1997). La puce ADN : un multi-réacteur de paillasse : Les nouveaux outils de l'analyse génétique et cytogénétique. *MS. Médecine sciences* 13, 1317-1324.
- Benes,V. and Castoldi,M. (2010). Expression profiling of microRNA using real-time quantitative PCR, how to use it and what is available. *Methods* 50, 244-249.
- Benjamini,Y. (1988). Opening the box of a boxplot. *The American Statistician* 42, 257-262.
- Biegel,J.A., Meek,R.S., Parmiter,A.H., Conard,K., and Emanuel,B.S. (1991). Chromosomal translocation t(1;13)(p36;q14) in a case of rhabdomyosarcoma. *Genes Chromosomes. Cancer* 3, 483-484.
- Bielack,S.S., Kempf-Bielack,B., Delling,G., Exner,G.U., Flege,S., Helmke,K., Kotz,R., Salzer-Kuntschik,M., Werner,M., Winkelmann,W., Zoubek,A., Jurgens,H., and Winkler,K. (2002). Prognostic factors in high-grade osteosarcoma of the extremities or trunk: an analysis of 1,702 patients treated on neoadjuvant cooperative osteosarcoma study group protocols. *J. Clin. Oncol.* 20, 776-790.
- Bolstad,B.M., Collin,F., Simpson,K.M., Irizarry,R.A., and Speed,T.P. (2004). Experimental design and low-level analysis of microarray data. *Int. Rev. Neurobiol.* 60, 25-58.
- Bolstad,B.M., Irizarry,R.A., Astrand,M., and Speed,T.P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics.* 19, 185-193.
- Bos,J.L. (1989). ras oncogenes in human cancer: a review. *Cancer Res.* 49, 4682-4689.
- Boyerinas,B., Park,S.M., Hau,A., Murmann,A.E., and Peter,M.E. (2010). The role of let-7 in cell differentiation and cancer. *Endocr. Relat Cancer* 17, F19-F36.
- Brennecke,J., Stark,A., Russell,R.B., and Cohen,S.M. (2005). Principles of microRNA-target recognition. *PLoS. Biol.* 3, e85.
- Bridge,J.A., Liu,J., Qualman,S.J., Suijkerbuijk,R., Wenger,G., Zhang,J., Wan,X., Baker,K.S., Sorensen,P., and Barr,F.G. (2002). Genomic gains and losses are similar in genetic and histologic subsets of rhabdomyosarcoma, whereas amplification predominates in embryonal with anaplasia and alveolar subtypes. *Genes Chromosomes. Cancer* 33, 310-321.
- Bruland,O.S., Hoifodt,H., Saeter,G., Smeland,S., and Fodstad,O. (2005). Hematogenous micrometastases in osteosarcoma patients. *Clin. Cancer Res.* 11, 4666-4673.
- Bullough,P.G. (2004). Orthopaedic pathology. (Edinburgh ; London.
- Caillaud,J.M., Gerard-Marchant,R., Marsden,H.B., van Unnik,A.J., Rodary,C., Rey,A., and Flamant,F. (1989). Histopathological classification of childhood rhabdomyosarcoma: a report from the International Society of Pediatric Oncology pathology panel. *Med. Pediatr. Oncol.* 17, 391-400.
- Cain,J.W. and Bender,C.M. (1995). Ifosfamide-induced neurotoxicity: associated symptoms and nursing implications. *Oncol. Nurs. Forum* 22, 659-666.
- Calin,G.A., Dumitru,C.D., Shimizu,M., Bichi,R., Zupo,S., Noch,E., Aldler,H., Rattan,S., Keating,M., Rai,K., Rassenti,L., Kipps,T., Negrini,M., Bullrich,F., and Croce,C.M. (2002). Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc. Natl. Acad. Sci. U. S. A* 99, 15524-15529.

- Calin,G.A., Sevignani,C., Dumitru,C.D., Hyslop,T., Noch,E., Yendamuri,S., Shimizu,M., Rattan,S., Bullrich,F., Negrini,M., and Croce,C.M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc. Natl. Acad. Sci. U. S. A* *101*, 2999-3004.
- Calvert,G.T., Randall,R.L., Jones,K.B., Cannon-Albright,L., Lessnick,S., and Schiffman,J.D. (2012). At-risk populations for osteosarcoma: the syndromes and beyond. *Sarcoma*. *2012*, 152382.
- Carter,N.P. (2007). Methods and strategies for analyzing copy number variation using DNA microarrays. *Nat. Genet.* *39*, S16-S21.
- Castanotto,D. and Rossi,J.J. (2009). The promises and pitfalls of RNA-interference-based therapeutics. *Nature* *457*, 426-433.
- Castello,G., Scala,S., Palmieri,G., Curley,S.A., and Izzo,F. (2010). HCV-related hepatocellular carcinoma: From chronic inflammation to cancer. *Clin. Immunol.* *134*, 237-250.
- Castoldi,M., Schmidt,S., Benes,V., Noerholm,M., Kulozik,A.E., Hentze,M.W., and Muckenthaler,M.U. (2006). A sensitive array for microRNA expression profiling (miChip) based on locked nucleic acids (LNA). *RNA*. *12*, 913-920.
- Chan,J.A., Krichevsky,A.M., and Kosik,K.S. (2005). MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. *Cancer Res.* *65*, 6029-6033.
- Chapin,S.C. and Doyle,P.S. (2011). Ultrasensitive multiplexed microRNA quantification on encoded gel microparticles using rolling circle amplification. *Anal. Chem.* *83*, 7179-7185.
- Chen,C., Ridzon,D.A., Broomer,A.J., Zhou,Z., Lee,D.H., Nguyen,J.T., Barbisin,M., Xu,N.L., Mahuvakar,V.R., Andersen,M.R., Lao,K.Q., Livak,K.J., and Guegler,K.J. (2005). Real-time quantification of microRNAs by stem-loop RT-PCR. *Nucleic Acids Res.* *33*, e179.
- Chendrimada,T.P., Finn,K.J., Ji,X., Baillat,D., Gregory,R.I., Liebhaber,S.A., Pasquinelli,A.E., and Shiekhattar,R. (2007). MicroRNA silencing through RISC recruitment of eIF6. *Nature* *447*, 823-828.
- Cheon,G.J., Kim,M.S., Lee,J.A., Lee,S.Y., Cho,W.H., Song,W.S., Koh,J.S., Yoo,J.Y., Oh,D.H., Shin,D.S., and Jeon,D.G. (2009). Prediction model of chemotherapy response in osteosarcoma by 18F-FDG PET and MRI. *J. Nucl. Med.* *50*, 1435-1440.
- Chibon,F., Lagarde,P., Salas,S., Perot,G., Brouste,V., Tirode,F., Lucchesi,C., de,R.A., Kauffmann,A., Bui,B., Terrier,P., Bonvalot,S., Le,C.A., Vince-Ranchere,D., Blay,J.Y., Collin,F., Guillou,L., Leroux,A., Coindre,J.M., and Aurias,A. (2010). 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.
- Chin,L.J., Ratner,E., Leng,S., Zhai,R., Nallur,S., Babar,I., Muller,R.U., Straka,E., Su,L., Burki,E.A., Crowell,R.E., Patel,R., Kulkarni,T., Homer,R., Zelterman,D., Kidd,K.K., Zhu,Y., Christiani,D.C., Belinsky,S.A., Slack,F.J., and Weidhaas,J.B. (2008). A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. *Cancer Res.* *68*, 8535-8540.
- Cho,W.C. (2011). Circulating MicroRNAs as Minimally Invasive Biomarkers for Cancer Theragnosis and Prognosis. *Front Genet.* *2*, 7.

- Cho,W.C. (2012). Great potential of miRNAs as predictive and prognostic markers for cancer. *Expert. Rev. Mol. Diagn.* *12*, 315-318.
- Chou,A.J. and Gorlick,R. (2006). Chemotherapy resistance in osteosarcoma: current challenges and future directions. *Expert. Rev. Anticancer Ther.* *6*, 1075-1085.
- Churchill,G.A. (2004). Using ANOVA to analyze microarray data. *Biotechniques* *37*, 173-5, 177.
- Cimmino,A., Calin,G.A., Fabbri,M., Iorio,M.V., Ferracin,M., Shimizu,M., Wojcik,S.E., Aqeilan,R.I., Zupo,S., Dono,M., Rassenti,L., Alder,H., Volinia,S., Liu,C.G., Kipps,T.J., Negrini,M., and Croce,C.M. (2005). miR-15 and miR-16 induce apoptosis by targeting BCL2. *Proc. Natl. Acad. Sci. U. S. A* *102*, 13944-13949.
- Clark,J.C., Dass,C.R., and Choong,P.F. (2008). A review of clinical and molecular prognostic factors in osteosarcoma. *J. Cancer Res. Clin. Oncol.* *134*, 281-297.
- Cleton-Jansen,A.M., Horst,B., and Pancras,C.W.H. (2005). Central high-grade osteosarcoma of bone: Diagnostic and genetic considerations. *Current diagnostic pathology* *11*, 390-399.
- Cleveland WS. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* *74*, 829-836.
- Cleveland WS. and Devlin SJ. (1988). Locally-Weighted Regression: An Approach to Regression Analysis by Local Fitting. *Journal of the American Statistical Association* *83*, 596-610.
- Corcoran,D.L., Pandit,K.V., Gordon,B., Bhattacharjee,A., Kaminski,N., and Benos,P.V. (2009). Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS. One.* *4*, e5279.
- Cortes,E.P., Holland,J.F., Wang,J.J., Sinks,L.F., Blom,J., Senn,H., Bank,A., and Glidewell,O. (1974). Amputation and adriamycin in primary osteosarcoma. *N. Engl. J. Med.* *291*, 998-1000.
- Costa,F.F., Bischof,J.M., Vanin,E.F., Lulla,R.R., Wang,M., Sredni,S.T., Rajaram,V., Bonaldo,M.F., Wang,D., Goldman,S., Tomita,T., and Soares,M.B. (2011). Identification of microRNAs as potential prognostic markers in ependymoma. *PLoS. One.* *6*, e25114.
- Crist,W., Gehan,E.A., Ragab,A.H., Dickman,P.S., Donaldson,S.S., Fryer,C., Hammond,D., Hays,D.M., Herrmann,J., Heyn,R., and . (1995). The Third Intergroup Rhabdomyosarcoma Study. *J. Clin. Oncol.* *13*, 610-630.
- Danovi,D., Meulmeester,E., Pasini,D., Migliorini,D., Capra,M., Frenk,R., de,G.P., Francoz,S., Gasparini,P., Gobbi,A., Helin,K., Pelicci,P.G., Jochemsen,A.G., and Marine,J.C. (2004). Amplification of Mdmx (or Mdm4) directly contributes to tumor formation by inhibiting p53 tumor suppressor activity. *Mol. Cell Biol.* *24*, 5835-5843.
- Darido,C., Georgy,S.R., Wilanowski,T., Dworkin,S., Auden,A., Zhao,Q., Rank,G., Srivastava,S., Finlay,M.J., Papenfuss,A.T., Pandolfi,P.P., Pearson,R.B., and Jane,S.M. (2011). Targeting of the tumor suppressor GRHL3 by a miR-21-dependent proto-oncogenic network results in PTEN loss and tumorigenesis. *Cancer Cell* *20*, 635-648.

- Davis,A.M., Bell,R.S., and Goodwin,P.J. (1994a). Prognostic factors in osteosarcoma: a critical review. *J. Clin. Oncol.* 12, 423-431.
- Davis,M.E., Zuckerman,J.E., Choi,C.H., Seligson,D., Tolcher,A., Alabi,C.A., Yen,Y., Heidel,J.D., and Ribas,A. (2010). Evidence of RNAi in humans from systemically administered siRNA via targeted nanoparticles. *Nature* 464, 1067-1070.
- Davis,R.J., D'Cruz,C.M., Lovell,M.A., Biegel,J.A., and Barr,F.G. (1994b). Fusion of PAX7 to FKHR by the variant t(1;13)(p36;q14) translocation in alveolar rhabdomyosarcoma. *Cancer Res.* 54, 2869-2872.
- de Franca,N.R., Mesquita,J.D., Lima,A.B., Pucci,F.V., Andrade,L.E., and Silva,N.P. (2010). RNA interference: a new alternative for rheumatic diseases therapy. *Rev. Bras. Reumatol.* 50, 695-702.
- De,P.M. and Naldini,L. (2010). Antagonizing metastasis. *Nat. Biotechnol.* 28, 331-332.
- Dedes,K.J., Natrajan,R., Lambros,M.B., Geyer,F.C., Lopez-Garcia,M.A., Savage,K., Jones,R.L., and Reis-Filho,J.S. (2011). Down-regulation of the miRNA master regulators Drosha and Dicer is associated with specific subgroups of breast cancer. *Eur. J. Cancer* 47, 138-150.
- Deo,A., Carlsson,J., and Lindlof,A. (2011). How to choose a normalization strategy for miRNA quantitative real-time (qPCR) arrays. *J. Bioinform. Comput. Biol.* 9, 795-812.
- Dias,P., Kumar,P., Marsden,H.B., Gattamaneni,H.R., Heighway,J., and Kumar,S. (1990). N-myc gene is amplified in alveolar rhabdomyosarcomas (RMS) but not in embryonal RMS. *Int. J. Cancer* 45, 593-596.
- Diller,L., Sexsmith,E., Gottlieb,A., Li,F.P., and Malkin,D. (1995). Germline p53 mutations are frequently detected in young children with rhabdomyosarcoma. *J. Clin. Invest* 95, 1606-1611.
- Doench,J.G. and Sharp,P.A. (2004). Specificity of microRNA target selection in translational repression. *Genes Dev.* 18, 504-511.
- Dom,G., Tarabichi,M., Unger,K., Thomas,G., Oczko-Wojciechowska,M., Bogdanova,T., Jarzab,B., Dumont,J.E., Detours,V., and Maenhaut,C. (2012). A gene expression signature distinguishes normal tissues of sporadic and radiation-induced papillary thyroid carcinomas. *Br. J. Cancer* 107, 994-1000.
- Driman,D., Thorner,P.S., Greenberg,M.L., Chilton-MacNeill,S., and Squire,J. (1994). MYCN gene amplification in rhabdomyosarcoma. *Cancer* 73, 2231-2237.
- Dudoit S., Yang YH., Callow MJ., and Speed TP. (2002). Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica* 12, 111-139.
- Duhamel,L.A., Ye,H., Halai,D., Idowu,B.D., Presneau,N., Tirabosco,R., and Flanagan,A.M. (2012). Frequency of Mouse Double Minute 2 (MDM2) and Mouse Double Minute 4 (MDM4) amplification in parosteal and conventional osteosarcoma subtypes. *Histopathology* 60, 357-359.
- Dutour,A., Leclers,D., Monteil,J., Paraf,F., Charissoux,J.L., Rousseau,R., and Rigaud,M. (2007). Non-invasive imaging correlates with histological and molecular characteristics of an osteosarcoma model: application for early detection and follow-up of MDR phenotype. *Anticancer Res.* 27, 4171-4178.

- Dutour,A., Monteil,J., Paraf,F., Charissoux,J.L., Kaletta,C., Sauer,B., Naujoks,K., and Rigaud,M. (2005). Endostatin cDNA/cationic liposome complexes as a promising therapy to prevent lung metastases in osteosarcoma: study in a human-like rat orthotopic tumor. *Mol. Ther.* **11**, 311-319.
- Eilber,F.R. and Rosen,G. (1989). Adjuvant chemotherapy for osteosarcoma. *Semin. Oncol.* **16**, 312-322.
- Eisen,M.B., Spellman,P.T., Brown,P.O., and Botstein,D. (1998). Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A* **95**, 14863-14868.
- Enzinger,F.M. and Weiss,S.W. (1995). *Soft Tissue Tumors*. (St Louis.
- Epstein,J.A., Lam,P., Jepeal,L., Maas,R.L., and Shapiro,D.N. (1995). Pax3 inhibits myogenic differentiation of cultured myoblast cells. *J. Biol. Chem.* **270**, 11719-11722.
- Eulalio,A., Behm-Ansmant,I., and Izaurralde,E. (2007a). P bodies: at the crossroads of post-transcriptional pathways. *Nat. Rev. Mol. Cell Biol.* **8**, 9-22.
- Eulalio,A., Behm-Ansmant,I., Schweizer,D., and Izaurralde,E. (2007b). P-body formation is a consequence, not the cause, of RNA-mediated gene silencing. *Mol. Cell Biol.* **27**, 3970-3981.
- Eulalio,A., Huntzinger,E., Nishihara,T., Rehwinkel,J., Fauser,M., and Izaurralde,E. (2009). Deadenylation is a widespread effect of miRNA regulation. *RNA.* **15**, 21-32.
- Ferguson,W.S. and Goorin,A.M. (2001). Current treatment of osteosarcoma. *Cancer Invest* **19**, 292-315.
- Ferrari,S. and Palmerini,E. (2007). Adjuvant and neoadjuvant combination chemotherapy for osteogenic sarcoma. *Curr. Opin. Oncol.* **19**, 341-346.
- Ferrari,S., Palmerini,E., Staals,E., Abate,M.E., Longhi,A., Cesari,M., Balladelli,A., Pratelli,L., and Bacci,G. (2009). Sex- and age-related chemotherapy toxicity in patients with non-metastatic osteosarcoma. *J. Chemother.* **21**, 205-210.
- Feugeas,O., Guriec,N., Babin-Boilletot,A., Marcellin,L., Simon,P., Babin,S., Thyss,A., Hofman,P., Terrier,P., Kalifa,C., Brunat-Mentigny,M., Patricot,L.M., and Oberling,F. (1996). Loss of heterozygosity of the RB gene is a poor prognostic factor in patients with osteosarcoma. *J. Clin. Oncol.* **14**, 467-472.
- Filipowicz,W., Bhattacharyya,S.N., and Sonenberg,N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat. Rev. Genet.* **9**, 102-114.
- Fire,A., Xu,S., Montgomery,M.K., Kostas,S.A., Driver,S.E., and Mello,C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806-811.
- Fletcher,C.D.M., Unni,K.K., and Mertens,F. (2002). *World Health Organization Classification of Tumours. Pathology and Genetics of Tumours of Soft Tissue and Bone*. Lyon, IARC Press).
- Fodor,S.P., Rava,R.P., Huang,X.C., Pease,A.C., Holmes,C.P., and Adams,C.L. (1993). Multiplexed biochemical assays with biological chips. *Nature* **364**, 555-556.

- Folkman,J. (1990). What is the evidence that tumors are angiogenesis dependent? *J. Natl. Cancer Inst.* 82, 4-6.
- Folkman,J. (2006). Antiangiogenesis in cancer therapy--endostatin and its mechanisms of action. *Exp. Cell Res.* 312, 594-607.
- Fontana,L., Fiori,M.E., Albini,S., Cifaldi,L., Giovinazzi,S., Forloni,M., Boldrini,R., Donfrancesco,A., Federici,V., Giacomini,P., Peschle,C., and Fruci,D. (2008). Antagomir-17-5p abolishes the growth of therapy-resistant neuroblastoma through p21 and BIM. *PLoS. One.* 3, e2236.
- Frades,I. and Matthiesen,R. (2010). Overview on techniques in cluster analysis. *Methods Mol. Biol.* 593, 81-107.
- Friend,S.H., Bernards,R., Rogelj,S., Weinberg,R.A., Rapaport,J.M., Albert,D.M., and Dryja,T.P. (1986). A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. *Nature* 323, 643-646.
- Furlanut,M. and Franceschi,L. (2003). Pharmacology of ifosfamide. *Oncology* 65 Suppl 2, 2-6.
- Furness,P.N., Taub,N., Assmann,K.J., Banfi,G., Cosyns,J.P., Dorman,A.M., Hill,C.M., Kapper,S.K., Waldherr,R., Laurinavicius,A., Marcussen,N., Martins,A.P., Nogueira,M., Regele,H., Seron,D., Carrera,M., Sund,S., Taskinen,E.I., Paavonen,T., Tihomirova,T., and Rosenthal,R. (2003). International variation in histologic grading is large, and persistent feedback does not improve reproducibility. *Am. J. Surg. Pathol.* 27, 805-810.
- Gabriely,G., Wurdinger,T., Kesari,S., Esau,C.C., Burchard,J., Linsley,P.S., and Krichevsky,A.M. (2008). MicroRNA 21 promotes glioma invasion by targeting matrix metalloproteinase regulators. *Mol. Cell Biol.* 28, 5369-5380.
- Galili,N., Davis,R.J., Fredericks,W.J., Mukhopadhyay,S., Rauscher,F.J., III, Emanuel,B.S., Rovera,G., and Barr,F.G. (1993). Fusion of a fork head domain gene to PAX3 in the solid tumour alveolar rhabdomyosarcoma. *Nat. Genet.* 5, 230-235.
- Gasparini,G. and Harris,A.L. (1995). Clinical importance of the determination of tumor angiogenesis in breast carcinoma: much more than a new prognostic tool. *J. Clin. Oncol.* 13, 765-782.
- Gasparini,M., Rouesse,J., van,O.A., Wagener,T., Somers,R., Russel,J.A., Voute,P.A., Bramwell,V., Thomas,D., Sylvester,R., and . (1985). Phase II study of cisplatin in advanced osteogenic sarcoma. European Organization for Research on Treatment of Cancer Soft Tissue and Bone Sarcoma Group. *Cancer Treat. Rep.* 69, 211-213.
- Gentleman,R. and Huber,W. (2007). Making the most of high-throughput protein-interaction data. *Genome Biol.* 8, 112.
- Gerlinger,M., Rowan,A.J., Horswell,S., Larkin,J., Endesfelder,D., Gronroos,E., Martinez,P., Matthews,N., Stewart,A., Tarpey,P., Varela,I., Phillimore,B., Begum,S., McDonald,N.Q., Butler,A., Jones,D., Raine,K., Latimer,C., Santos,C.R., Nohadani,M., Eklund,A.C., Spencer-Dene,B., Clark,G., Pickering,L., Stamp,G., Gore,M., Szallasi,Z., Downward,J., Futreal,P.A., and Swanton,C. (2012). Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* 366, 883-892.

- Gibbs,C.P., Weber,K., and Scarborough,M.T. (2002). Malignant bone tumors. Instr. Course Lect. *51*, 413-428.
- Gil-Benso,R., Carda-Batalla,C., Navarro-Fos,S., Pellin-Perez,A., and Llombart-Bosch,A. (1999). Cytogenetic study of a spindle-cell rhabdomyosarcoma of the parotid gland. *Cancer Genet. Cytogenet.* *109*, 150-153.
- Goldberg,N.S. and Collins,F.S. (1991). The hunt for the neurofibromatosis gene. *Arch. Dermatol.* *127*, 1705-1707.
- Gougelet,A., Perez,J., Pissaloux,D., Besse,A., Duc,A., Decouvelaere,A.V., Ranchere-Vince,D., Blay,J.Y., and Alberti,L. (2011a). miRNA Profiling: How to Bypass the Current Difficulties in the Diagnosis and Treatment of Sarcomas. *Sarcoma.* *2011*, 460650.
- Gougelet,A., Pissaloux,D., Besse,A., Perez,J., Duc,A., Dutour,A., Blay,J.Y., and Alberti,L. (2011b). Micro-RNA profiles in osteosarcoma as a predictive tool for ifosfamide response. *Int. J. Cancer* *129*, 680-690.
- Greene,F.L. (2002). The American Joint Committee on Cancer: updating the strategies in cancer staging. *Bull. Am. Coll. Surg.* *87*, 13-15.
- Gregory,R.I., Yan,K.P., Amuthan,G., Chendrimada,T., Doratotaj,B., Cooch,N., and Shiekhattar,R. (2004). The Microprocessor complex mediates the genesis of microRNAs. *Nature* *432*, 235-240.
- Griffiths-Jones,S. (2004). The microRNA Registry. *Nucleic Acids Res.* *32*, D109-D111.
- Grimson,A., Farh,K.K., Johnston,W.K., Garrett-Engle,P., Lim,L.P., and Bartel,D.P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. *Mol. Cell* *27*, 91-105.
- Gripp,K.W., Scott,C.I., Jr., Nicholson,L., McDonald-McGinn,D.M., Ozeran,J.D., Jones,M.C., Lin,A.E., and Zackai,E.H. (2002). Five additional Costello syndrome patients with rhabdomyosarcoma: proposal for a tumor screening protocol. *Am. J. Med. Genet.* *108*, 80-87.
- Grufferman,S., Ruymann,F., Ognjanovic,S., Erhardt,E.B., and Maurer,H.M. (2009). Prenatal X-ray exposure and rhabdomyosarcoma in children: a report from the children's oncology group. *Cancer Epidemiol. Biomarkers Prev.* *18*, 1271-1276.
- Grufferman,S., Schwartz,A.G., Ruymann,F.B., and Maurer,H.M. (1993). Parents' use of cocaine and marijuana and increased risk of rhabdomyosarcoma in their children. *Cancer Causes Control* *4*, 217-224.
- Guillou,L. and Aurias,A. (2010). Soft tissue sarcomas with complex genomic profiles. *Virchows Arch.* *456*, 201-217.
- Gunderson,K.L., Steemers,F.J., Lee,G., Mendoza,L.G., and Chee,M.S. (2005). A genome-wide scalable SNP genotyping assay using microarray technology. *Nat. Genet.* *37*, 549-554.
- Guo,T., Zhang,L., Chang,N.E., Singer,S., Maki,R.G., and Antonescu,C.R. (2011). Consistent MYC and FLT4 gene amplification in radiation-induced angiosarcoma but not in other radiation-associated atypical vascular lesions. *Genes Chromosomes. Cancer* *50*, 25-33.

- Hachitanda,Y., Toyoshima,S., Akazawa,K., and Tsuneyoshi,M. (1998). N-myc gene amplification in rhabdomyosarcoma detected by fluorescence in situ hybridization: its correlation with histologic features. *Mod. Pathol.* **11**, 1222-1227.
- Hacia,J.G. (1999). Resequencing and mutational analysis using oligonucleotide microarrays. *Nat. Genet.* **21**, 42-47.
- Hackler,L., Jr., Masuda,T., Oliver,V.F., Merbs,S.L., and Zack,D.J. (2012a). Use of laser capture microdissection for analysis of retinal mRNA/miRNA expression and DNA methylation. *Methods Mol. Biol.* **884**, 289-304.
- Hackler,L., Jr., Masuda,T., Oliver,V.F., Merbs,S.L., and Zack,D.J. (2012b). Use of laser capture microdissection for analysis of retinal mRNA/miRNA expression and DNA methylation. *Methods Mol. Biol.* **884**, 289-304.
- Halimi,J.M., Azizi,M., Bobrie,G., Bouche,O., Deray,G., des,G.G., Lecomte,T., Levy,B., Mourad,J.J., Nochy,D., Oudard,S., Rieu,P., and Sahali,D. (2008). [Vascular and renal effects of anti-angiogenic therapy]. *Nephrol. Ther.* **4**, 602-615.
- Han,J., Lee,Y., Yeom,K.H., Kim,Y.K., Jin,H., and Kim,V.N. (2004). The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* **18**, 3016-3027.
- Hansen,M.F. and Cavenee,W.K. (1987). Retinoblastoma and osteosarcoma: the prototypic cancer family. *Acta Paediatr. Jpn.* **29**, 526-533.
- Haque,F., Shu,D., Shu,Y., Shlyakhtenko,L.S., Rychahou,P.G., Evers,B.M., and Guo,P. (2012). Ultrastable synergistic tetravalent RNA nanoparticles for targeting to cancers. *Nano. Today* **7**, 245-257.
- Hartley,A.L., Birch,J.M., McKinney,P.A., Teare,M.D., Blair,V., Carrette,J., Mann,J.R., Draper,G.J., Stiller,C.A., Johnston,H.E., and . (1988). The Inter-Regional Epidemiological Study of Childhood Cancer (IRESCC): case control study of children with bone and soft tissue sarcomas. *Br. J. Cancer* **58**, 838-842.
- Hayashi,Y., Inaba,T., Hanada,R., and Yamamoto,K. (1988). Translocation 2;8 in a congenital rhabdomyosarcoma. *Cancer Genet. Cytogenet.* **30**, 343-345.
- He,L. and Hannon,G.J. (2004). MicroRNAs: small RNAs with a big role in gene regulation. *Nat. Rev. Genet.* **5**, 522-531.
- Heegaard,N.H., Schetter,A.J., Welsh,J.A., Yoneda,M., Bowman,E.D., and Harris,C.C. (2012). Circulating micro-RNA expression profiles in early stage nonsmall cell lung cancer. *Int. J. Cancer* **130**, 1378-1386.
- Heinsohn,S., Evermann,U., Zur,S.U., Bielack,S., and Kabisch,H. (2007). Determination of the prognostic value of loss of heterozygosity at the retinoblastoma gene in osteosarcoma. *Int. J. Oncol.* **30**, 1205-1214.
- Helman,L.J. and Meltzer,P. (2003). Mechanisms of sarcoma development. *Nat. Rev. Cancer* **3**, 685-694.
- Ho,R.H., Johnson,J., Dev,V.G., and Whitlock,J.A. (2004). A novel t(2;20)(q35;p12) in embryonal rhabdomyosarcoma. *Cancer Genet. Cytogenet.* **151**, 73-77.

- Hollenbach,A.D., Sublett,J.E., McPherson,C.J., and Grosveld,G. (1999). The Pax3-FKHR oncoprotein is unresponsive to the Pax3-associated repressor hDaxx. *EMBO J.* 18, 3702-3711.
- Hoogerwerf,W.A., Hawkins,A.L., Perlman,E.J., and Griffin,C.A. (1994). Chromosome analysis of nine osteosarcomas. *Genes Chromosomes. Cancer* 9, 88-92.
- Horn,R.C. and Enterline,H. (1958). Rhabdomyosarcoma: a clinicopathological study and classification of 39 cases. *Cancer* 11, 181-199.
- Hosoi,H., Kakazu,N., Konishi,E., Tsuchihashi,Y., Hada,S., Amaya,E., Nakabayahi,Y., Misawa-Furihata,A., Tabata-Maruyama,H., Iehara,T., Sugimoto,T., Yamane,H., Yamasaki,M., Shiwaku,K., Yanagisawa,A., Ono,M., Tokiwa,K., Iwai,N., Hashiba,M., Abe,T., and Sawada,T. (2009). A novel PAX3 rearrangement in embryonal rhabdomyosarcoma. *Cancer Genet. Cytogenet.* 189, 98-104.
- Howlader, N., Noone, A. M., Krapcho, M., Neyman, N., Aminou, R., Waldron, W., Altekruse, S. F., Kosary, C. L., Ruhl, J., Tatalovich, Z., Cho, H., Mariotto, A., Eisner, M. P., Lewis, D. R., Chen, H. S., Feuer, E. J., and Cronin, K. A. SEER Cancer Statistics Review, 1975-2009 (Vintage 2009 Populations). National Cancer Institute.Bethesda, MD. 20-8-2012. 20-9-2012.
Ref Type: Online Source
- Hughes,T.R., Mao,M., Jones,A.R., Burchard,J., Marton,M.J., Shannon,K.W., Lefkowitz,S.M., Ziman,M., Schelter,J.M., Meyer,M.R., Kobayashi,S., Davis,C., Dai,H., He,Y.D., Stephanians,S.B., Cavet,G., Walker,W.L., West,A., Coffey,E., Shoemaker,D.D., Stoughton,R., Blanchard,A.P., Friend,S.H., and Linsley,P.S. (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat. Biotechnol.* 19, 342-347.
- Huh,W.W. and Skapek,S.X. (2010). Childhood rhabdomyosarcoma: new insight on biology and treatment. *Curr. Oncol. Rep.* 12, 402-410.
- Hui,A.B., Shi,W., Boutros,P.C., Miller,N., Pintilie,M., Fyles,T., McCready,D., Wong,D., Gerster,K., Waldron,L., Jurisica,I., Penn,L.Z., and Liu,F.F. (2009). Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. *Lab Invest* 89, 597-606.
- Hutvagner,G., McLachlan,J., Pasquinelli,A.E., Balint,E., Tuschl,T., and Zamore,P.D. (2001). A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293, 834-838.
- Hutvagner,G. and Zamore,P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. *Science* 297, 2056-2060.
- Iolascon,A., Faienza,M.F., Coppola,B., Rosolen,A., Basso,G., della,R.F., and Schettini,F. (1996). Analysis of cyclin-dependent kinase inhibitor genes (CDKN2A, CDKN2B, and CDKN2C) in childhood rhabdomyosarcoma. *Genes Chromosomes. Cancer* 15, 217-222.
- Iorio,M.V., Ferracin,M., Liu,C.G., Veronese,A., Spizzo,R., Sabbioni,S., Magri,E., Pedriali,M., Fabbri,M., Campiglio,M., Menard,S., Palazzo,J.P., Rosenberg,A., Musiani,P., Volinia,S., Nenci,I., Calin,G.A., Querzoli,P., Negrini,M., and Croce,C.M. (2005). MicroRNA gene expression deregulation in human breast cancer. *Cancer Res.* 65, 7065-7070.
- Jafari,P. and Azuaje,F. (2006). An assessment of recently published gene expression data analyses: reporting experimental design and statistical factors. *BMC. Med. Inform. Decis. Mak.* 6, 27.

- Jaffe,N., Carrasco,H., Raymond,K., Ayala,A., and Eftekhari,F. (2002). Can cure in patients with osteosarcoma be achieved exclusively with chemotherapy and abrogation of surgery? *Cancer* 95, 2202-2210.
- Jaffe,N., Frei,E., III, Traggis,D., and Bishop,Y. (1974). Adjuvant methotrexate and citrovorum-factor treatment of osteogenic sarcoma. *N. Engl. J. Med.* 291, 994-997.
- Jensen,T.B., Langkjaer,N., and Wengel,J. (2008). Unlocked nucleic acid (UNA) and UNA derivatives: thermal denaturation studies. *Nucleic Acids Symp. Ser. (Oxf)* 133-134.
- Jewell,S.D., Srinivasan,M., McCart,L.M., Williams,N., Grizzle,W.H., LiVolsi,V., MacLennan,G., and Sedmak,D.D. (2002). Analysis of the molecular quality of human tissues: an experience from the Cooperative Human Tissue Network. *Am. J. Clin. Pathol.* 118, 733-741.
- Jiang,J., Lee,E.J., Gusev,Y., and Schmittgen,T.D. (2005). Real-time expression profiling of microRNA precursors in human cancer cell lines. *Nucleic Acids Res.* 33, 5394-5403.
- Jo,V.Y., Marino-Enriquez,A., and Fletcher,C.D. (2011). Epithelioid rhabdomyosarcoma: clinicopathologic analysis of 16 cases of a morphologically distinct variant of rhabdomyosarcoma. *Am. J. Surg. Pathol.* 35, 1523-1530.
- John,B., Enright,A.J., Aravin,A., Tuschl,T., Sander,C., and Marks,D.S. (2004). Human MicroRNA targets. *PLoS. Biol.* 2, e363.
- Johnson,S.M., Grosshans,H., Shingara,J., Byrom,M., Jarvis,R., Cheng,A., Labourier,E., Reinert,K.L., Brown,D., and Slack,F.J. (2005). RAS is regulated by the let-7 microRNA family. *Cell* 120, 635-647.
- Jung,M., Schaefer,A., Steiner,I., Kempkensteffen,C., Stephan,C., Erbersdobler,A., and Jung,K. (2010). Robust microRNA stability in degraded RNA preparations from human tissue and cell samples. *Clin. Chem.* 56, 998-1006.
- Kalisky,T. and Quake,S.R. (2011). Single-cell genomics. *Nat. Methods* 8, 311-314.
- Kalluri,R. and Zeisberg,M. (2006). Fibroblasts in cancer. *Nat. Rev. Cancer* 6, 392-401.
- Kanehisa,M., Goto,S., Sato,Y., Furumichi,M., and Tanabe,M. (2012). KEGG for integration and interpretation of large-scale molecular data sets. *Nucleic Acids Res.* 40, D109-D114.
- Kansara,M. and Thomas,D.M. (2007). Molecular pathogenesis of osteosarcoma. *DNA Cell Biol.* 26, 1-18.
- Kaste,S.C., Pratt,C.B., Cain,A.M., Jones-Wallace,D.J., and Rao,B.N. (1999). Metastases detected at the time of diagnosis of primary pediatric extremity osteosarcoma at diagnosis: imaging features. *Cancer* 86, 1602-1608.
- Kaya,M., Wada,T., Nagoya,S., Kawaguchi,S., Isu,K., and Yamashita,T. (2004). Concomitant tumour resistance in patients with osteosarcoma. A clue to a new therapeutic strategy. *J. Bone Joint Surg. Br.* 86, 143-147.

- Keleti,J., Quezado,M.M., Abaza,M.M., Raffeld,M., and Tsokos,M. (1996). The MDM2 oncoprotein is overexpressed in rhabdomyosarcoma cell lines and stabilizes wild-type p53 protein. *Am. J. Pathol.* **149**, 143-151.
- Keller,C., Arenkiel,B.R., Coffin,C.M., El-Bardeesy,N., DePinho,R.A., and Capecchi,M.R. (2004). Alveolar rhabdomyosarcomas in conditional Pax3:Fkhr mice: cooperativity of Ink4a/ARF and Trp53 loss of function. *Genes Dev.* **18**, 2614-2626.
- Kelly,K.M., Womer,R.B., Sorensen,P.H., Xiong,Q.B., and Barr,F.G. (1997). Common and variant gene fusions predict distinct clinical phenotypes in rhabdomyosarcoma. *J. Clin. Oncol.* **15**, 1831-1836.
- Khanna,C. (2008). Novel targets with potential therapeutic applications in osteosarcoma. *Curr. Oncol. Rep.* **10**, 350-358.
- Khatri,P., Sirota,M., and Butte,A.J. (2012). Ten years of pathway analysis: current approaches and outstanding challenges. *PLoS. Comput. Biol.* **8**, e1002375.
- Kim,H.J., Chalmers,P.N., and Morris,C.D. (2010a). Pediatric osteogenic sarcoma. *Curr. Opin. Pediatr.* **22**, 61-66.
- Kim,K., Zakharkin,S.O., and Allison,D.B. (2010b). Expectations, validity, and reality in gene expression profiling. *J. Clin. Epidemiol.* **63**, 950-959.
- Kingsley,L.A., Fournier,P.G., Chirgwin,J.M., and Guise,T.A. (2007). Molecular biology of bone metastasis. *Mol. Cancer Ther.* **6**, 2609-2617.
- Kiriakidou,M., Tan,G.S., Lamprinaki,S., De Planell-Saguer,M., Nelson,P.T., and Mourelatos,Z. (2007). An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**, 1141-1151.
- Klastersky,J. (2003). Side effects of ifosfamide. *Oncology* **65 Suppl 2**, 7-10.
- Kong,Y.W., Ferland-McCollough,D., Jackson,T.J., and Bushell,M. (2012). microRNAs in cancer management. *Lancet Oncol.* **13**, e249-e258.
- Kosaka,N., Iguchi,H., and Ochiya,T. (2010). Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis. *Cancer Sci.* **101**, 2087-2092.
- Koscielniak,E., Rodary,C., Flamant,F., Carli,M., Treuner,J., Pinkerton,C.R., and Grotto,P. (1992). Metastatic rhabdomyosarcoma and histologically similar tumors in childhood: a retrospective European multi-center analysis. *Med. Pediatr. Oncol.* **20**, 209-214.
- Kozomara,A. and Griffiths-Jones,S. (2011). miRBase: integrating microRNA annotation and deep-sequencing data. *Nucleic Acids Res.* **39**, D152-D157.
- Krek,A., Grun,D., Poy,M.N., Wolf,R., Rosenberg,L., Epstein,E.J., MacMenamin,P., da,P., I, Gunsalus,K.C., Stoffel,M., and Rajewsky,N. (2005). Combinatorial microRNA target predictions. *Nat. Genet.* **37**, 495-500.
- Kruse,J.P. and Gu,W. (2009). Modes of p53 regulation. *Cell* **137**, 609-622.

- Kuimelis,R.G., Livak,K.J., Mullah,B., and Andrus,A. (1997). Structural analogues of TaqMan probes for real-time quantitative PCR. *Nucleic Acids Symp. Ser.* 255-256.
- Kumar,S., Perlman,E., Harris,C.A., Raffeld,M., and Tsokos,M. (2000). Myogenin is a specific marker for rhabdomyosarcoma: an immunohistochemical study in paraffin-embedded tissues. *Mod. Pathol.* 13, 988-993.
- Kusafuka,T., Fukuzawa,M., Oue,T., Komoto,Y., Yoneda,A., and Okada,A. (1997). Mutation analysis of p53 gene in childhood malignant solid tumors. *J. Pediatr. Surg.* 32, 1175-1180.
- Ladanyi,M., Cha,C., Lewis,R., Jhanwar,S.C., Huvos,A.G., and Healey,J.H. (1993). MDM2 gene amplification in metastatic osteosarcoma. *Cancer Res.* 53, 16-18.
- Lagoda,P. and Regad,F. (2000). Les puces à ADN : outils pour une nouvelle vision de la diversité et des ressources génétiques. *Cahiers Agriculture* 9, 329-340.
- Lagos-Quintana,M., Rauhut,R., Lendeckel,W., and Tuschl,T. (2001). Identification of novel genes coding for small expressed RNAs. *Science* 294, 853-858.
- Lagutina,I., Conway,S.J., Sublett,J., and Grosveld,G.C. (2002). Pax3-FKHR knock-in mice show developmental aberrations but do not develop tumors. *Mol. Cell Biol.* 22, 7204-7216.
- Lau,N.C., Lim,L.P., Weinstein,E.G., and Bartel,D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* 294, 858-862.
- Lawrence,W., Jr., Hays,D.M., and Moon,T.E. (1977). Lymphatic metastasis with childhood rhabdomyosarcoma. *Cancer* 39, 556-559.
- Lazarus,H.M. (1974). Alterations of high molecular weight nuclear RNA following kidney storage. *Exp. Mol. Pathol.* 21, 155-165.
- Le,L.Q. and Parada,L.F. (2007). Tumor microenvironment and neurofibromatosis type I: connecting the GAPs. *Oncogene* 26, 4609-4616.
- Leaphart,C. and Rodeberg,D. (2007). Pediatric surgical oncology: management of rhabdomyosarcoma. *Surg. Oncol.* 16, 173-185.
- Lee,R.C. and Ambros,V. (2001). An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* 294, 862-864.
- Lee,R.C., Feinbaum,R.L., and Ambros,V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* 75, 843-854.
- Lee,Y., Ahn,C., Han,J., Choi,H., Kim,J., Yim,J., Lee,J., Provost,P., Radmark,O., Kim,S., and Kim,V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425, 415-419.
- Lee,Y., Jeon,K., Lee,J.T., Kim,S., and Kim,V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* 21, 4663-4670.
- Lee,Y., Kim,M., Han,J., Yeom,K.H., Lee,S., Baek,S.H., and Kim,V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 23, 4051-4060.

- Lennon,G.G. and Lehrach,H. (1991). Hybridization analyses of arrayed cDNA libraries. *Trends Genet.* **7**, 314-317.
- Leskela,S., Leandro-Garcia,L.J., Mendiola,M., Barriuso,J., Inglada-Perez,L., Munoz,I., Martinez-Delgado,B., Redondo,A., de,S.J., Robledo,M., Hardisson,D., and Rodriguez-Antona,C. (2011). The miR-200 family controls beta-tubulin III expression and is associated with paclitaxel-based treatment response and progression-free survival in ovarian cancer patients. *Endocr. Relat Cancer* **18**, 85-95.
- Lewis,B.P., Burge,C.B., and Bartel,D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* **120**, 15-20.
- Lewis,B.P., Shih,I.H., Jones-Rhoades,M.W., Bartel,D.P., and Burge,C.B. (2003). Prediction of mammalian microRNA targets. *Cell* **115**, 787-798.
- Li,F.P. and Fraumeni,J.F., Jr. (1969). Soft-tissue sarcomas, breast cancer, and other neoplasms. A familial syndrome? *Ann. Intern. Med.* **71**, 747-752.
- Li,L. (2010). Dimension reduction for high-dimensional data. *Methods Mol. Biol.* **620**, 417-434.
- Li,W. and Ruan,K. (2009). MicroRNA detection by microarray. *Anal. Bioanal. Chem.* **394**, 1117-1124.
- Liang,R.Q., Li,W., Li,Y., Tan,C.Y., Li,J.X., Jin,Y.X., and Ruan,K.C. (2005). An oligonucleotide microarray for microRNA expression analysis based on labeling RNA with quantum dot and nanogold probe. *Nucleic Acids Res.* **33**, e17.
- Lin,R.J., Lin,Y.C., Chen,J., Kuo,H.H., Chen,Y.Y., Diccianni,M.B., London,W.B., Chang,C.H., and Yu,A.L. (2010). microRNA signature and expression of Dicer and Drosha can predict prognosis and delineate risk groups in neuroblastoma. *Cancer Res.* **70**, 7841-7850.
- Linardic,C.M. and Counter,C.M. (2008). Genetic modeling of Ras-induced human rhabdomyosarcoma. *Methods Enzymol.* **438**, 419-427.
- Liu,J., Guzman,M.A., Pezanowski,D., Patel,D., Hauptman,J., Keisling,M., Hou,S.J., Papenhausen,P.R., Pascasio,J.M., Punnett,H.H., Halligan,G.E., and de Chadarevian,J.P. (2011). FOXO1-FGFR1 fusion and amplification in a solid variant of alveolar rhabdomyosarcoma. *Mod. Pathol.* **24**, 1327-1335.
- Liu,N., Chen,N.Y., Cui,R.X., Li,W.F., Li,Y., Wei,R.R., Zhang,M.Y., Sun,Y., Huang,B.J., Chen,M., He,Q.M., Jiang,N., Chen,L., Cho,W.C., Yun,J.P., Zeng,J., Liu,L.Z., Li,L., Guo,Y., Wang,H.Y., and Ma,J. (2012). Prognostic value of a microRNA signature in nasopharyngeal carcinoma: a microRNA expression analysis. *Lancet Oncol.* **13**, 633-641.
- Livak,K.J. and Schmittgen,T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402-408.
- Lu,J., Getz,G., Miska,E.A., Alvarez-Saavedra,E., Lamb,J., Peck,D., Sweet-Cordero,A., Ebert,B.L., Mak,R.H., Ferrando,A.A., Downing,J.R., Jacks,T., Horvitz,H.R., and Golub,T.R. (2005). MicroRNA expression profiles classify human cancers. *Nature* **435**, 834-838.
- Lund,E., Guttinger,S., Calado,A., Dahlberg,J.E., and Kutay,U. (2004). Nuclear export of microRNA precursors. *Science* **303**, 95-98.

- Ma,K., He,Y., Zhang,H., Fei,Q., Niu,D., Wang,D., Ding,X., Xu,H., Chen,X., and Zhu,J. (2012a). DNA methylation-regulated miR-193a-3p dictates resistance of hepatocellular carcinoma to 5-fluorouracil via repression of SRSF2 expression. *J. Biol. Chem.* **287**, 5639-5649.
- Ma,Y., Dai,H., and Kong,X. (2012b). Impact of warm ischemia on gene expression analysis in surgically removed biosamples. *Anal. Biochem.* **423**, 229-235.
- Mahajan,A., Woo,S.Y., Kornguth,D.G., Hughes,D., Huh,W., Chang,E.L., Herzog,C.E., Pelloski,C.E., and Anderson,P. (2008). Multimodality treatment of osteosarcoma: radiation in a high-risk cohort. *Pediatr. Blood Cancer* **50**, 976-982.
- Manor,E., Joshua,B.Z., Nash,M., Brennan,P.A., and Bodner,L. (2012). Cytogenetics of primary embryonal rhabdomyosarcoma of the cheek. *Br. J. Oral Maxillofac. Surg.*
- Mansky,P., Arai,A., Stratton,P., Bernstein,D., Long,L., Reynolds,J., Chen,D., Steinberg,S.M., Lavende,N., Hoffman,K., Nathan,P.C., Parks,R., Augustine,E., Chaudhry,U., Derdak,J., Wiener,L., Gerber,L., and Mackall,C. (2007). Treatment late effects in long-term survivors of pediatric sarcoma. *Pediatr. Blood Cancer* **48**, 192-199.
- Margue,C.M., Bernasconi,M., Barr,F.G., and Schafer,B.W. (2000). Transcriptional modulation of the anti-apoptotic protein BCL-XL by the paired box transcription factors PAX3 and PAX3/FKHR. *Oncogene* **19**, 2921-2929.
- Marti,C., Kroner,T., Remagen,W., Berchtold,W., Cserhati,M., and Varini,M. (1985). High-dose ifosfamide in advanced osteosarcoma. *Cancer Treat. Rep.* **69**, 115-117.
- Matsui,I., Tanimura,M., Kobayashi,N., Sawada,T., Nagahara,N., and Akatsuka,J. (1993). Neurofibromatosis type 1 and childhood cancer. *Cancer* **72**, 2746-2754.
- Maurer,H.M., Beltangady,M., Gehan,E.A., Crist,W., Hammond,D., Hays,D.M., Heyn,R., Lawrence,W., Newton,W., Ortega,J., and . (1988). The Intergroup Rhabdomyosarcoma Study-I. A final report. *Cancer* **61**, 209-220.
- Maurer,H.M., Gehan,E.A., Beltangady,M., Crist,W., Dickman,P.S., Donaldson,S.S., Fryer,C., Hammond,D., Hays,D.M., Herrmann,J., and . (1993). The Intergroup Rhabdomyosarcoma Study-II. *Cancer* **71**, 1904-1922.
- Medina,P.P., Nolde,M., and Slack,F.J. (2010). OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. *Nature* **467**, 86-90.
- Mehra,S., de la Roza,G., Tull,J., Shrimpton,A., Valente,A., and Zhang,S. (2008). Detection of FOXO1 (FKHR) gene break-apart by fluorescence in situ hybridization in formalin-fixed, paraffin-embedded alveolar rhabdomyosarcomas and its clinicopathologic correlation. *Diagn. Mol. Pathol.* **17**, 14-20.
- Meister,G., Landthaler,M., Patkaniowska,A., Dorsett,Y., Teng,G., and Tuschl,T. (2004). Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs. *Mol. Cell* **15**, 185-197.
- Merlino,G. and Helman,L.J. (1999). Rhabdomyosarcoma--working out the pathways. *Oncogene* **18**, 5340-5348.
- Merritt,W.M., Lin,Y.G., Han,L.Y., Kamat,A.A., Spannuth,W.A., Schmandt,R., Urbauer,D., Pennacchio,L.A., Cheng,J.F., Nick,A.M., Deavers,M.T., Mourad-Zeidan,A., Wang,H., Mueller,P., Lenburg,M.E., Gray,J.W., Mok,S., Birrer,M.J., Lopez-Berestein,G., Coleman,R.L., Bar-Eli,M., and

- Sood,A.K. (2008). Dicer, Drosha, and outcomes in patients with ovarian cancer. *N. Engl. J. Med.* 359, 2641-2650.
- Meyer,S.U., Kaiser,S., Wagner,C., Thirion,C., and Pfaffl,M.W. (2012). Profound effect of profiling platform and normalization strategy on detection of differentially expressed microRNAs--a comparative study. *PLoS. One.* 7, e38946.
- Meyer,S.U., Pfaffl,M.W., and Ulbrich,S.E. (2010). Normalization strategies for microRNA profiling experiments: a 'normal' way to a hidden layer of complexity? *Biotechnol. Lett.* 32, 1777-1788.
- Meza,J.L., Anderson,J., Pappo,A.S., and Meyer,W.H. (2006). Analysis of prognostic factors in patients with nonmetastatic rhabdomyosarcoma treated on intergroup rhabdomyosarcoma studies III and IV: the Children's Oncology Group. *J. Clin. Oncol.* 24, 3844-3851.
- Michael,A., Bajracharya,S.D., Yuen,P.S., Zhou,H., Star,R.A., Illei,G.G., and Alevizos,I. (2010). Exosomes from human saliva as a source of microRNA biomarkers. *Oral Dis.* 16, 34-38.
- Miele,E., Spinelli,G.P., Miele,E., Di,F.E., Ferretti,E., Tomao,S., and Gulino,A. (2012). Nanoparticle-based delivery of small interfering RNA: challenges for cancer therapy. *Int. J. Nanomedicine.* 7, 3637-3657.
- Miller,C.W., Aslo,A., Campbell,M.J., Kawamata,N., Lampkin,B.C., and Koeffler,H.P. (1996a). Alterations of the p15, p16, and p18 genes in osteosarcoma. *Cancer Genet. Cytogenet.* 86, 136-142.
- Miller,C.W., Aslo,A., Won,A., Tan,M., Lampkin,B., and Koeffler,H.P. (1996b). Alterations of the p53, Rb and MDM2 genes in osteosarcoma. *J. Cancer Res. Clin. Oncol.* 122, 559-565.
- Miller,P.J. and Hollenbach,A.D. (2007). The oncogenic fusion protein Pax3-FKHR has a greater post-translational stability relative to Pax3 during early myogenesis. *Biochim. Biophys. Acta* 1770, 1450-1458.
- Mintz,M.B., Sowers,R., Brown,K.M., Hilmer,S.C., Mazza,B., Huvos,A.G., Meyers,P.A., Lafleur,B., McDonough,W.S., Henry,M.M., Ramsey,K.E., Antonescu,C.R., Chen,W., Healey,J.H., Daluski,A., Berens,M.E., Macdonald,T.J., Gorlick,R., and Stephan,D.A. (2005). An expression signature classifies chemotherapy-resistant pediatric osteosarcoma. *Cancer Res.* 65, 1748-1754.
- Miser,J.S., Kinsella,T.J., Triche,T.J., Tsokos,M., Jarosinski,P., Forquer,R., Wesley,R., and Magrath,I. (1987). Ifosfamide with mesna uroprotection and etoposide: an effective regimen in the treatment of recurrent sarcomas and other tumors of children and young adults. *J. Clin. Oncol.* 5, 1191-1198.
- Missiaglia,E., Shepherd,C.J., Patel,S., Thway,K., Pierron,G., Pritchard-Jones,K., Renard,M., Sciot,R., Rao,P., Oberlin,O., Delattre,O., and Shipley,J. (2010). MicroRNA-206 expression levels correlate with clinical behaviour of rhabdomyosarcomas. *Br. J. Cancer* 102, 1769-1777.
- Missiaglia,E., Williamson,D., Chisholm,J., Wirapati,P., Pierron,G., Petel,F., Concordet,J.P., Thway,K., Oberlin,O., Pritchard-Jones,K., Delattre,O., Delorenzi,M., and Shipley,J. (2012). PAX3/FOXO1 fusion gene status is the key prognostic molecular marker in rhabdomyosarcoma and significantly improves current risk stratification. *J. Clin. Oncol.* 30, 1670-1677.
- Mitchell,P.S., Parkin,R.K., Kroh,E.M., Fritz,B.R., Wyman,S.K., Pogosova-Agadjanyan,E.L., Peterson,A., Noteboom,J., O'Briant,K.C., Allen,A., Lin,D.W., Urban,N., Drescher,C.W., Knudsen,B.S., Stirewalt,D.L., Gentleman,R., Vessella,R.L., Nelson,P.S., Martin,D.B., and Tewari,M. (2008). Circulating microRNAs as stable blood-based markers for cancer detection. *Proc. Natl. Acad. Sci. U. S. A* 105, 10513-10518.

- Miyachi,M., Tsuchiya,K., Yoshida,H., Yagyu,S., Kikuchi,K., Misawa,A., Iehara,T., and Hosoi,H. (2010). Circulating muscle-specific microRNA, miR-206, as a potential diagnostic marker for rhabdomyosarcoma. *Biochem. Biophys. Res. Commun.* **400**, 89-93.
- Molenaar,W.M., Oosterhuis,J.W., Oosterhuis,A.M., and Ramaekers,F.C. (1985). Mesenchymal and muscle-specific intermediate filaments (vimentin and desmin) in relation to differentiation in childhood rhabdomyosarcomas. *Hum. Pathol.* **16**, 838-843.
- Monteys,A.M., Spengler,R.M., Wan,J., Tecedor,L., Lennox,K.A., Xing,Y., and Davidson,B.L. (2010). Structure and activity of putative intronic miRNA promoters. *RNA.* **16**, 495-505.
- Morlando,M., Ballarino,M., Gromak,N., Pagano,F., Bozzoni,I., and Proudfoot,N.J. (2008). Primary microRNA transcripts are processed co-transcriptionally. *Nat. Struct. Mol. Biol.* **15**, 902-909.
- Mourelatos,Z., Dostie,J., Paushkin,S., Sharma,A., Charroux,B., Abel,L., Rappsilber,J., Mann,M., and Dreyfuss,G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* **16**, 720-728.
- Mueller,M.M. and Fusenig,N.E. (2004). Friends or foes - bipolar effects of the tumour stroma in cancer. *Nat. Rev. Cancer* **4**, 839-849.
- Mullis,K., Faloona,F., Scharf,S., Saiki,R., Horn,G., and Erlich,H. (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb. Symp. Quant. Biol.* **51 Pt 1**, 263-273.
- Napoli,C., Lemieux,C., and Jorgensen,R. (1990). Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* **2**, 279-289.
- Nascimento,A.F. and Fletcher,C.D. (2005). Spindle cell rhabdomyosarcoma in adults. *Am. J. Surg. Pathol.* **29**, 1106-1113.
- Nettleton, D. Normalization Methods for Two-Color Microarray Data. 2009. 20-10-2012.
Ref Type: Online Source
- Newton,W.A., Jr., Gehan,E.A., Webber,B.L., Marsden,H.B., van Unnik,A.J., Hamoudi,A.B., Tsokos,M.G., Shimada,H., Harms,D., Schmidt,D., and . (1995). Classification of rhabdomyosarcomas and related sarcomas. Pathologic aspects and proposal for a new classification--an Intergroup Rhabdomyosarcoma Study. *Cancer* **76**, 1073-1085.
- Ng,C.K., Pemberton,H.N., and Reis-Filho,J.S. (2012). Breast cancer intratumor genetic heterogeneity: causes and implications. *Expert. Rev. Anticancer Ther.* **12**, 1021-1032.
- Nielsen,G.P., Burns,K.L., Rosenberg,A.E., and Louis,D.N. (1998). CDKN2A gene deletions and loss of p16 expression occur in osteosarcomas that lack RB alterations. *Am. J. Pathol.* **153**, 159-163.
- Nissan,X., Denis,J.A., Saidani,M., Lemaitre,G., Peschanski,M., and Baldeschi,C. (2011). miR-203 modulates epithelial differentiation of human embryonic stem cells towards epidermal stratification. *Dev. Biol.* **356**, 506-515.
- Nolan,T., Hands,R.E., and Bustin,S.A. (2006). Quantification of mRNA using real-time RT-PCR. *Nat. Protoc.* **1**, 1559-1582.

- Nottrott,S., Simard,M.J., and Richter,J.D. (2006). Human let-7a miRNA blocks protein production on actively translating polyribosomes. *Nat. Struct. Mol. Biol.* **13**, 1108-1114.
- O'Donnell,K.A., Wentzel,E.A., Zeller,K.I., Dang,C.V., and Mendell,J.T. (2005). c-Myc-regulated microRNAs modulate E2F1 expression. *Nature* **435**, 839-843.
- Obana,K., Yang,H.W., Piao,H.Y., Taki,T., Hashizume,K., Hanada,R., Yamamoto,K., Tanaka,Y., Toyoda,Y., Takita,J., Tsuchida,Y., and Hayashi,Y. (2003). Aberrations of p16INK4A, p14ARF and p15INK4B genes in pediatric solid tumors. *Int. J. Oncol.* **23**, 1151-1157.
- Obernosterer,G., Leuschner,P.J., Alenius,M., and Martinez,J. (2006). Post-transcriptional regulation of microRNA expression. *RNA*. **12**, 1161-1167.
- Ohuchida,K., Mizumoto,K., Kayashima,T., Fujita,H., Moriyama,T., Ohtsuka,T., Ueda,J., Nagai,E., Hashizume,M., and Tanaka,M. (2011). MicroRNA expression as a predictive marker for gemcitabine response after surgical resection of pancreatic cancer. *Ann. Surg. Oncol.* **18**, 2381-2387.
- Okamura,K., Phillips,M.D., Tyler,D.M., Duan,H., Chou,Y.T., and Lai,E.C. (2008). The regulatory activity of microRNA* species has substantial influence on microRNA and 3' UTR evolution. *Nat. Struct. Mol. Biol.* **15**, 354-363.
- Olive,V., Bennett,M.J., Walker,J.C., Ma,C., Jiang,I., Cordon-Cardo,C., Li,Q.J., Lowe,S.W., Hannon,G.J., and He,L. (2009). miR-19 is a key oncogenic component of mir-17-92. *Genes Dev.* **23**, 2839-2849.
- Olsen,P.H. and Ambros,V. (1999). The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev. Biol.* **216**, 671-680.
- Orr,M.S. and Scherf,U. (2002). Large-scale gene expression analysis in molecular target discovery. *Leukemia* **16**, 473-477.
- Ottaviani,G. and Jaffe,N. (2009). The etiology of osteosarcoma. *Cancer Treat. Res.* **152**, 15-32.
- Ozaki,T., Schaefer,K.L., Wai,D., Buerger,H., Flege,S., Lindner,N., Kevric,M., Diallo,R., Bankfalvi,A., Brinkschmidt,C., Juergens,H., Winkelmann,W., Dockhorn-Dworniczak,B., Bielack,S.S., and Poremba,C. (2002). Genetic imbalances revealed by comparative genomic hybridization in osteosarcomas. *Int. J. Cancer* **102**, 355-365.
- Ozen,M., Creighton,C.J., Ozdemir,M., and Ittmann,M. (2008). Widespread deregulation of microRNA expression in human prostate cancer. *Oncogene* **27**, 1788-1793.
- Pandita,A., Zielenska,M., Thorner,P., Bayani,J., Godbout,R., Greenberg,M., and Squire,J.A. (1999). Application of comparative genomic hybridization, spectral karyotyping, and microarray analysis in the identification of subtype-specific patterns of genomic changes in rhabdomyosarcoma. *Neoplasia*. **1**, 262-275.
- Parham,D.M. (2001). Pathologic classification of rhabdomyosarcomas and correlations with molecular studies. *Mod. Pathol.* **14**, 506-514.
- Parham,D.M. and Ellison,D.A. (2006). Rhabdomyosarcomas in adults and children: an update. *Arch. Pathol. Lab Med.* **130**, 1454-1465.

- Parham,D.M., Webber,B., Holt,H., Williams,W.K., and Maurer,H. (1991). Immunohistochemical study of childhood rhabdomyosarcomas and related neoplasms. Results of an Intergroup Rhabdomyosarcoma study project. *Cancer* 67, 3072-3080.
- Park,N.J., Zhou,H., Elashoff,D., Henson,B.S., Kastratovic,D.A., Abemayor,E., and Wong,D.T. (2009). Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. *Clin. Cancer Res.* 15, 5473-5477.
- Park,T., Yi,S.G., Kang,S.H., Lee,S., Lee,Y.S., and Simon,R. (2003). Evaluation of normalization methods for microarray data. *BMC. Bioinformatics.* 4, 33.
- Pasquinelli,A.E., Reinhart,B.J., Slack,F., Martindale,M.Q., Kuroda,M.I., Maller,B., Hayward,D.C., Ball,E.E., Degnan,B., Muller,P., Spring,J., Srinivasan,A., Fishman,M., Finnerty,J., Corbo,J., Levine,M., Leahy,P., Davidson,E., and Ruvkun,G. (2000). Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature* 408, 86-89.
- Pastore,G., Peris-Bonet,R., Carli,M., Martinez-Garcia,C., Sanchez de,T.J., and Steliarova-Foucher,E. (2006). Childhood soft tissue sarcomas incidence and survival in European children (1978-1997): report from the Automated Childhood Cancer Information System project. *Eur. J. Cancer* 42, 2136-2149.
- Patino-Garcia,A., Pineiro,E.S., Diez,M.Z., Iturriagaitia,L.G., Klusmann,F.A., and Ariznabarreta,L.S. (2003). Genetic and epigenetic alterations of the cell cycle regulators and tumor suppressor genes in pediatric osteosarcomas. *J. Pediatr. Hematol. Oncol.* 25, 362-367.
- Paulson,V., Chandler,G., Rakheja,D., Galindo,R.L., Wilson,K., Amatruda,J.F., and Cameron,S. (2011). High-resolution array CGH identifies common mechanisms that drive embryonal rhabdomyosarcoma pathogenesis. *Genes Chromosomes. Cancer* 50, 397-408.
- Pease,A.C., Solas,D., Sullivan,E.J., Cronin,M.T., Holmes,C.P., and Fodor,S.P. (1994). Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci. U. S. A* 91, 5022-5026.
- Pecot,C.V., Calin,G.A., Coleman,R.L., Lopez-Berestein,G., and Sood,A.K. (2011). RNA interference in the clinic: challenges and future directions. *Nat. Rev. Cancer* 11, 59-67.
- Pedersen-Bjergaard,J., Andersen,M.K., Andersen,M.T., and Christiansen,D.H. (2008). Genetics of therapy-related myelodysplasia and acute myeloid leukemia. *Leukemia* 22, 240-248.
- Pelgrims,J., De,V.F., Van den Brande,J., Schrijvers,D., Prove,A., and Vermorken,J.B. (2000). Methylene blue in the treatment and prevention of ifosfamide-induced encephalopathy: report of 12 cases and a review of the literature. *Br. J. Cancer* 82, 291-294.
- Peltier,L. (1993). Tumors of bone and soft tissues. In *Orthopedics: A History and Iconography*, (Novato, CA: Norman Publishing), pp. 264-292.
- Petersen,C.P., Bordeleau,M.E., Pelletier,J., and Sharp,P.A. (2006). Short RNAs repress translation after initiation in mammalian cells. *Mol. Cell* 21, 533-542.
- Philip,T., Blay,J.Y., Brunat-Mentigny,M., Carrie,C., Chauvot,P., Farsi,F., Fervers,B., Gentet,J.C., Giammarile,F., Kolher,R., Mathoulin,S., Patricot,L.M., and Thiesse,P. (1999). [Standards, options and recommendations (SOR) for diagnosis, treatment and follow-up of osteosarcoma. Groupe de travail SOR]. *Bull. Cancer* 86, 159-176.

- Picci,P. (2007). Osteosarcoma (osteogenic sarcoma). Orphanet. J. Rare. Dis. 2, 6.
- Picci,P., Sangiorgi,L., Rougraff,B.T., Neff,J.R., Casadei,R., and Campanacci,M. (1994). Relationship of chemotherapy-induced necrosis and surgical margins to local recurrence in osteosarcoma. J. Clin. Oncol. 12, 2699-2705.
- Pratt,C.B., Howarth,C., Ransom,J.L., Bowles,D., Green,A.A., Kumar,A.P., Rivera,G., and Evans,W.E. (1980). High-dose methotrexate used alone and in combination for measurable primary or metastatic osteosarcoma. Cancer Treat. Rep. 64, 11-20.
- Pratt,C.B. and Shanks,E.C. (1974). Doxorubicin in treatment of malignant solid tumors in children. Am. J. Dis. Child 127, 534-536.
- Provenzano,M. and Mocellin,S. (2007). Complementary techniques: validation of gene expression data by quantitative real time PCR. Adv. Exp. Med. Biol. 593, 66-73.
- Provisor,A.J., Ettinger,L.J., Nachman,J.B., Krailo,M.D., Makley,J.T., Yunis,E.J., Huvos,A.G., Betcher,D.L., Baum,E.S., Kisker,C.T., and Miser,J.S. (1997). Treatment of nonmetastatic osteosarcoma of the extremity with preoperative and postoperative chemotherapy: a report from the Children's Cancer Group. J. Clin. Oncol. 15, 76-84.
- Punzalan,M. and Hyden,G. (2009). The role of physical therapy and occupational therapy in the rehabilitation of pediatric and adolescent patients with osteosarcoma. Cancer Treat. Res. 152, 367-384.
- Qian,P., Zuo,Z., Wu,Z., Meng,X., Li,G., Wu,Z., Zhang,W., Tan,S., Pandey,V., Yao,Y., Wang,P., Zhao,L., Wang,J., Wu,Q., Song,E., Lobie,P.E., Yin,Z., and Zhu,T. (2011). Pivotal role of reduced let-7g expression in breast cancer invasion and metastasis. Cancer Res. 71, 6463-6474.
- Ramsay,G. (1998). DNA chips: state-of-the art. Nat. Biotechnol. 16, 40-44.
- Ranade,A.R., Cherba,D., Sridhar,S., Richardson,P., Webb,C., Paripati,A., Bowles,B., and Weiss,G.J. (2010). MicroRNA 92a-2*: a biomarker predictive for chemoresistance and prognostic for survival in patients with small cell lung cancer. J. Thorac. Oncol. 5, 1273-1278.
- Rao,Y., Lee,Y., Jarjoura,D., Ruppert,A.S., Liu,C.G., Hsu,J.C., and Hagan,J.P. (2008). A comparison of normalization techniques for microRNA microarray data. Stat. Appl. Genet. Mol. Biol. 7, Article22.
- Reichek,J.L., Duan,F., Smith,L.M., Gustafson,D.M., O'Connor,R.S., Zhang,C., Dunlevy,M.J., Gastier-Foster,J.M., and Barr,F.G. (2011). Genomic and clinical analysis of amplification of the 13q31 chromosomal region in alveolar rhabdomyosarcoma: a report from the Children's Oncology Group. Clin. Cancer Res. 17, 1463-1473.
- Reik,W., Brown,K.W., Slatter,R.E., Sartori,P., Elliott,M., and Maher,E.R. (1994). Allelic methylation of H19 and IGF2 in the Beckwith-Wiedemann syndrome. Hum. Mol. Genet. 3, 1297-1301.
- Reiner,A., Yekutieli,D., and Benjamini,Y. (2003). Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics. 19, 368-375.
- Reinhart,B.J., Slack,F.J., Basson,M., Pasquinelli,A.E., Bettinger,J.C., Rougvie,A.E., Horvitz,H.R., and Ruvkun,G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403, 901-906.

- Relaix,F., Montarras,D., Zaffran,S., Gayraud-Morel,B., Rocancourt,D., Tajbakhsh,S., Mansouri,A., Cumano,A., and Buckingham,M. (2006). Pax3 and Pax7 have distinct and overlapping functions in adult muscle progenitor cells. *J. Cell Biol.* *172*, 91-102.
- Ren,Y.X., Finckenstein,F.G., Abdueva,D.A., Shahbazian,V., Chung,B., Weinberg,K.I., Triche,T.J., Shimada,H., and Anderson,M.J. (2008). Mouse mesenchymal stem cells expressing PAX-FKHR form alveolar rhabdomyosarcomas by cooperating with secondary mutations. *Cancer Res.* *68*, 6587-6597.
- Resnick,K.E., Alder,H., Hagan,J.P., Richardson,D.L., Croce,C.M., and Cohn,D.E. (2009). The detection of differentially expressed microRNAs from the serum of ovarian cancer patients using a novel real-time PCR platform. *Gynecol. Oncol.* *112*, 55-59.
- Robinson,J.C., Richardson,M.S., Neville,B.W., Day,T.A., and Chi,A.C. (2012). Sclerosing Rhabdomyosarcoma: Report of a Case Arising in the Head and Neck of an Adult and Review of the Literature. *Head Neck Pathol.*
- Rodriguez,A., Griffiths-Jones,S., Ashurst,J.L., and Bradley,A. (2004). Identification of mammalian microRNA host genes and transcription units. *Genome Res.* *14*, 1902-1910.
- Roeb,W., Boyer,A., Cavenee,W.K., and Arden,K.C. (2007). PAX3-FOXO1 controls expression of the p57Kip2 cell-cycle regulator through degradation of EGR1. *Proc. Natl. Acad. Sci. U. S. A* *104*, 18085-18090.
- Roeb,W., Boyer,A., Cavenee,W.K., and Arden,K.C. (2008). Guilt by association: PAX3-FOXO1 regulates gene expression through selective destabilization of the EGR1 transcription factor. *Cell Cycle* *7*, 837-841.
- Roepman,P., Schuurman,A., Delahaye,L.J., Witteveen,A.T., Floore,A.N., and Glas,A.M. (2009). A gene expression profile for detection of sufficient tumour cells in breast tumour tissue: microarray diagnosis eligibility. *BMC. Med. Genomics* *2*, 52.
- Rosenberg,Z.S., Lev,S., Schmahmann,S., Steiner,G.C., Beltran,J., and Present,D. (1995). Osteosarcoma: subtle, rare, and misleading plain film features. *AJR Am. J. Roentgenol.* *165*, 1209-1214.
- Rosenfeld,N., Aharonov,R., Meiri,E., Rosenwald,S., Spector,Y., Zepeniuk,M., Benjamin,H., Shabes,N., Tabak,S., Levy,A., Lebanony,D., Goren,Y., Silberschein,E., Targan,N., Ben-Ari,A., Gilad,S., Sion-Vardy,N., Tobar,A., Feinmesser,M., Kharenko,O., Nativ,O., Nass,D., Perelman,M., Yosepovich,A., Shalmon,B., Polak-Charcon,S., Fridman,E., Avniel,A., Bentwich,I., Bentwich,Z., Cohen,D., Chajut,A., and Barshack,I. (2008). MicroRNAs accurately identify cancer tissue origin. *Nat. Biotechnol.* *26*, 462-469.
- Ruby,J.G., Jan,C.H., and Bartel,D.P. (2007). Intronic microRNA precursors that bypass Drosha processing. *Nature* *448*, 83-86.
- Ruymann,F.B., Newton,W.A., Jr., Ragab,A.H., Donaldson,M.H., and Foulkes,M. (1984). Bone marrow metastases at diagnosis in children and adolescents with rhabdomyosarcoma. A report from the intergroup rhabdomyosarcoma study. *Cancer* *53*, 368-373.
- Sadikovic,B., Yoshimoto,M., Al-Romaih,K., Maire,G., Zielenska,M., and Squire,J.A. (2008). In vitro analysis of integrated global high-resolution DNA methylation profiling with genomic imbalance and gene expression in osteosarcoma. *PLoS. One.* *3*, e2834.

- Salas,S., Jezequel,P., Campion,L., Deville,J.L., Chibon,F., Bartoli,C., Gentet,J.C., Charbonnel,C., Gouraud,W., Voutsinos-Porche,B., Brouchet,A., Duffaud,F., Figarella-Branger,D., and Bouvier,C. (2009). Molecular characterization of the response to chemotherapy in conventional osteosarcomas: predictive value of HSD17B10 and IFITM2. *Int. J. Cancer* *125*, 851-860.
- Saleh,R.A., Graham-Pole,J., Cassano,W., Abbot,F., Vander,G.R., Dickson,N., Metha,P., Heare,M., Kedar,A., Heare,T., and . (1990). Response of osteogenic sarcoma to the combination of etoposide and cyclophosphamide as neoadjuvant chemotherapy. *Cancer* *65*, 861-865.
- Salzer-Kuntschik,M., Brand,G., and Delling,G. (1983). [Determination of the degree of morphological regression following chemotherapy in malignant bone tumors]. *Pathologie* *4*, 135-141.
- Sarver,A.L., Phalak,R., Thayanithy,V., and Subramanian,S. (2010). S-MED: sarcoma microRNA expression database. *Lab Invest* *90*, 753-761.
- Schajowicz,F., Sissons,H.A., and Sobin,L.H. (1995). The World Health Organization's histologic classification of bone tumors. A commentary on the second edition. *Cancer* *75*, 1208-1214.
- Schamberger,A., Sarkadi,B., and Orban,T.I. (2012). Human mirtrons can express functional microRNAs simultaneously from both arms in a flanking exon-independent manner. *RNA. Biol.* *9*.
- Scheidler,S., Fredericks,W.J., Rauscher,F.J., III, Barr,F.G., and Vogt,P.K. (1996). The hybrid PAX3-FKHR fusion protein of alveolar rhabdomyosarcoma transforms fibroblasts in culture. *Proc. Natl. Acad. Sci. U. S. A* *93*, 9805-9809.
- Schetter,A.J., Leung,S.Y., Sohn,J.J., Zanetti,K.A., Bowman,E.D., Yanaihara,N., Yuen,S.T., Chan,T.L., Kwong,D.L., Au,G.K., Liu,C.G., Calin,G.A., Croce,C.M., and Harris,C.C. (2008). MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* *299*, 425-436.
- Schimke,R.N., Lowman,J.T., and Cowan,A.B. (1974). Retinoblastoma and osteogenic sarcoma in siblings. *Cancer* *34*, 2077-2079.
- Schmittgen,T.D., Jiang,J., Liu,Q., and Yang,L. (2004). A high-throughput method to monitor the expression of microRNA precursors. *Nucleic Acids Res.* *32*, e43.
- Schuchhardt,J., Beule,D., Malik,A., Wolski,E., Eickhoff,H., Lehrach,H., and Herzelt,H. (2000). Normalization strategies for cDNA microarrays. *Nucleic Acids Res.* *28*, E47.
- Schwarz,D.S., Hutvagner,G., Du,T., Xu,Z., Aronin,N., and Zamore,P.D. (2003). Asymmetry in the assembly of the RNAi enzyme complex. *Cell* *115*, 199-208.
- Scrabble,H., Witte,D., Shimada,H., Seemayer,T., Sheng,W.W., Soukup,S., Koufos,A., Houghton,P., Lampkin,B., and Cavenee,W. (1989). Molecular differential pathology of rhabdomyosarcoma. *Genes Chromosomes. Cancer* *1*, 23-35.
- Sedlacek,H.H. (2001). Pharmacological aspects of targeting cancer gene therapy to endothelial cells. *Crit Rev. Oncol. Hematol.* *37*, 169-215.
- Selbach,M., Schwanhaussner,B., Thierfelder,N., Fang,Z., Khanin,R., and Rajewsky,N. (2008). Widespread changes in protein synthesis induced by microRNAs. *Nature* *455*, 58-63.

- Serra,M., Pasello,M., Manara,M.C., Scotlandi,K., Ferrari,S., Bertoni,F., Mercuri,M., Alvegard,T.A., Picci,P., Bacci,G., and Smeland,S. (2006). May P-glycoprotein status be used to stratify high-grade osteosarcoma patients? Results from the Italian/Scandinavian Sarcoma Group 1 treatment protocol. *Int. J. Oncol.* 29, 1459-1468.
- Shapiro,D.N., Sublett,J.E., Li,B., Downing,J.R., and Naeve,C.W. (1993). Fusion of PAX3 to a member of the forkhead family of transcription factors in human alveolar rhabdomyosarcoma. *Cancer Res.* 53, 5108-5112.
- Shen,Y.M., Hung,G.Y., Yen,H.J., Hsieh,M.Y., and Hsieh,T.K. (2009). Early development of acute myeloid leukemia following treatment of osteosarcoma: a case report and review of the literature. *Pediatr. Neonatol.* 50, 239-244.
- Sherr,C.J. (1996). Cancer cell cycles. *Science* 274, 1672-1677.
- Sherr,C.J. (2004). Principles of tumor suppression. *Cell* 116, 235-246.
- Sherr,C.J. and Weber,J.D. (2000). The ARF/p53 pathway. *Curr. Opin. Genet. Dev.* 10, 94-99.
- Shi,R. and Chiang,V.L. (2005). Facile means for quantifying microRNA expression by real-time PCR. *Biotechniques* 39, 519-525.
- Shi,Z., Johnson,J.J., and Stack,M.S. (2012). Fluorescence In Situ Hybridization for MicroRNA Detection in Archived Oral Cancer Tissues. *J. Oncol.* 2012, 903581.
- Shibata,Y., Fujita,S., Takahashi,H., Yamaguchi,A., and Koji,T. (2000). Assessment of decalcifying protocols for detection of specific RNA by non-radioactive in situ hybridization in calcified tissues. *Histochem. Cell Biol.* 113, 153-159.
- Shimada,H., Newton,W.A., Jr., Soule,E.H., Beltangady,M.S., and Maurer,H.M. (1987). Pathology of fatal rhabdomyosarcoma. Report from Intergroup Rhabdomyosarcoma Study (IRS-I and IRS-II). *Cancer* 59, 459-465.
- Shin,C., Nam,J.W., Farh,K.K., Chiang,H.R., Shkumatava,A., and Bartel,D.P. (2010). Expanding the microRNA targeting code: functional sites with centered pairing. *Mol. Cell* 38, 789-802.
- Shukla,N., Ameer,N., Yilmaz,I., Nafa,K., Lau,C.Y., Marchetti,A., Borsu,L., Barr,F.G., and Ladanyi,M. (2012). Oncogene mutation profiling of pediatric solid tumors reveals significant subsets of embryonal rhabdomyosarcoma and neuroblastoma with mutated genes in growth signaling pathways. *Clin. Cancer Res.* 18, 748-757.
- Si,M.L., Zhu,S., Wu,H., Lu,Z., Wu,F., and Mo,Y.Y. (2007). miR-21-mediated tumor growth. *Oncogene* 26, 2799-2803.
- Smith,A.C., Squire,J.A., Thorner,P., Zielenska,M., Shuman,C., Grant,R., Chitayat,D., Nishikawa,J.L., and Weksberg,R. (2001). Association of alveolar rhabdomyosarcoma with the Beckwith-Wiedemann syndrome. *Pediatr. Dev. Pathol.* 4, 550-558.
- Soe,M.J., Moller,T., Dufva,M., and Holmstrom,K. (2011). A sensitive alternative for microRNA in situ hybridizations using probes of 2'-O-methyl RNA + LNA. *J. Histochem. Cytochem.* 59, 661-672.
- Sommelet,D., Clavel,J., and Lacour,B. (2009). Epidémiologie des cancers de l'enfant.

- Sorensen,P.H., Lynch,J.C., Qualman,S.J., Tirabosco,R., Lim,J.F., Maurer,H.M., Bridge,J.A., Crist,W.M., Triche,T.J., and Barr,F.G. (2002). PAX3-FKHR and PAX7-FKHR gene fusions are prognostic indicators in alveolar rhabdomyosarcoma: a report from the children's oncology group. *J. Clin. Oncol.* 20, 2672-2679.
- Soussi,T. (2011). TP53 mutations in human cancer: database reassessment and prospects for the next decade. *Adv. Cancer Res.* 110, 107-139.
- Standart,N. and Jackson,R.J. (2007). MicroRNAs repress translation of m7Gppp-capped target mRNAs in vitro by inhibiting initiation and promoting deadenylation. *Genes Dev.* 21, 1975-1982.
- Stegmaier,S., Poremba,C., Schaefer,K.L., Leuschner,I., Kazanowska,B., Bekassy,A.N., Bielack,S.S., Klingebiel,T., and Koscielniak,E. (2011). Prognostic value of PAX-FKHR fusion status in alveolar rhabdomyosarcoma: a report from the cooperative soft tissue sarcoma study group (CWS). *Pediatr. Blood Cancer* 57, 406-414.
- Stempak,D., Gammon,J., Halton,J., Moghrabi,A., Koren,G., and Baruchel,S. (2006). A pilot pharmacokinetic and antiangiogenic biomarker study of celecoxib and low-dose metronomic vinblastine or cyclophosphamide in pediatric recurrent solid tumors. *J. Pediatr. Hematol. Oncol.* 28, 720-728.
- Stevens, M., Rey, A., Bouvet, C., Ellershaw, J., Sanchez de Toledo, J., and Oberlin, O. SIOP MMT 95: Intensified (6 drug) versus standard (IVA) chemotherapy for high risk non metastatic rhabdomyosarcoma (RMS). *Journal of Clinical Oncology. Oral Abstract Session, Pediatric Cancer II* . 2004 ASCO Annual Meeting Proceedings 22[14S], 8515. 15-7-2004.
Ref Type: Conference Proceeding
- Stevens,M.C. (2005). Treatment for childhood rhabdomyosarcoma: the cost of cure. *Lancet Oncol.* 6, 77-84.
- Stevens,M.C., Rey,A., Bouvet,N., Ellershaw,C., Flamant,F., Habrand,J.L., Marsden,H.B., Martelli,H., Sanchez de,T.J., Spicer,R.D., Spooner,D., Terrier-Lacombe,M.J., van,U.A., and Oberlin,O. (2005a). Treatment of nonmetastatic rhabdomyosarcoma in childhood and adolescence: third study of the International Society of Paediatric Oncology--SIOP Malignant Mesenchymal Tumor 89. *J. Clin. Oncol.* 23, 2618-2628.
- Stevens,M.C., Rey,A., Bouvet,N., Ellershaw,C., Flamant,F., Habrand,J.L., Marsden,H.B., Martelli,H., Sanchez de,T.J., Spicer,R.D., Spooner,D., Terrier-Lacombe,M.J., van,U.A., and Oberlin,O. (2005b). Treatment of nonmetastatic rhabdomyosarcoma in childhood and adolescence: third study of the International Society of Paediatric Oncology--SIOP Malignant Mesenchymal Tumor 89. *J. Clin. Oncol.* 23, 2618-2628.
- Stout,A. (1946). Rhabdomyosarcoma of the skeletal muscles. *Ann. Surg.* 123, 447-472.
- Streiner,D.L. and Norman,G.R. (2011). Correction for multiple testing: is there a resolution? *Chest* 140, 16-18.
- Subramanian,S., Lui,W.O., Lee,C.H., Espinosa,I., Nielsen,T.O., Heinrich,M.C., Corless,C.L., Fire,A.Z., and van de Rijn,M. (2008). MicroRNA expression signature of human sarcomas. *Oncogene* 27, 2015-2026.
- Sumegi,J., Streblow,R., Frayer,R.W., Dal,C.P., Rosenberg,A., Meloni-Ehrig,A., and Bridge,J.A. (2010). Recurrent t(2;2) and t(2;8) translocations in rhabdomyosarcoma without the canonical PAX-FOXO1

fuse PAX3 to members of the nuclear receptor transcriptional coactivator family. *Genes Chromosomes. Cancer* 49, 224-236.

Takahashi,Y., Oda,Y., Kawaguchi,K., Tamiya,S., Yamamoto,H., Suita,S., and Tsuneyoshi,M. (2004). Altered expression and molecular abnormalities of cell-cycle-regulatory proteins in rhabdomyosarcoma. *Mod. Pathol.* 17, 660-669.

Tan,X., Qin,W., Zhang,L., Hang,J., Li,B., Zhang,C., Wan,J., Zhou,F., Shao,K., Sun,Y., Wu,J., Zhang,X., Qiu,B., Li,N., Shi,S., Feng,X., Zhao,S., Wang,Z., Zhao,X., Chen,Z., Mitchelson,K., Cheng,J., Guo,Y., and He,J. (2011). A 5-microRNA signature for lung squamous cell carcinoma diagnosis and hsa-miR-31 for prognosis. *Clin. Cancer Res.* 17, 6802-6811.

Tang,F., Hajkova,P., Barton,S.C., O'Carroll,D., Lee,C., Lao,K., and Surani,M.A. (2006). 220-plex microRNA expression profile of a single cell. *Nat. Protoc.* 1, 1154-1159.

Tap,W.D., Eilber,F.C., Ginther,C., Dry,S.M., Reese,N., Barzan-Smith,K., Chen,H.W., Wu,H., Eilber,F.R., Slamon,D.J., and Anderson,L. (2011). Evaluation of well-differentiated/de-differentiated liposarcomas by high-resolution oligonucleotide array-based comparative genomic hybridization. *Genes Chromosomes. Cancer* 50, 95-112.

Taylor,A.C., Shu,L., Danks,M.K., Poquette,C.A., Shetty,S., Thayer,M.J., Houghton,P.J., and Harris,L.C. (2000). P53 mutation and MDM2 amplification frequency in pediatric rhabdomyosarcoma tumors and cell lines. *Med. Pediatr. Oncol.* 35, 96-103.

Thomson,J.M., Parker,J., Perou,C.M., and Hammond,S.M. (2004). A custom microarray platform for analysis of microRNA gene expression. *Nat. Methods* 1, 47-53.

Tian,Q., Price,N.D., and Hood,L. (2012). Systems cancer medicine: towards realization of predictive, preventive, personalized and participatory (P4) medicine. *J. Intern. Med.* 271, 111-121.

Tie,J. and Fan,D. (2011). Big roles of microRNAs in tumorigenesis and tumor development. *Histol. Histopathol.* 26, 1353-1361.

Tiwari,M. (2012). Microarrays and cancer diagnosis. *J. Cancer Res. Ther.* 8, 3-10.

Torres,A., Torres,K., Paszkowski,T., Jodlowska-Jedrych,B., Radomanski,T., Ksiazek,A., and Maciejewski,R. (2011). Major regulators of microRNAs biogenesis Dicer and Drosha are down-regulated in endometrial cancer. *Tumour. Biol.* 32, 769-776.

Trang,P., Wiggins,J.F., Daige,C.L., Cho,C., Omotola,M., Brown,D., Weidhaas,J.B., Bader,A.G., and Slack,F.J. (2011). Systemic delivery of tumor suppressor microRNA mimics using a neutral lipid emulsion inhibits lung tumors in mice. *Mol. Ther.* 19, 1116-1122.

Tsokos,M., Webber,B.L., Parham,D.M., Wesley,R.A., Miser,A., Miser,J.S., Etcubanas,E., Kinsella,T., Grayson,J., Glatstein,E., and . (1992). Rhabdomyosarcoma. A new classification scheme related to prognosis. *Arch. Pathol. Lab Med.* 116, 847-855.

Tsuchiya,T., Sekine,K., Hinohara,S., Namiki,T., Nobori,T., and Kaneko,Y. (2000). Analysis of the p16INK4, p14ARF, p15, TP53, and MDM2 genes and their prognostic implications in osteosarcoma and Ewing sarcoma. *Cancer Genet. Cytogenet.* 120, 91-98.

- Tsunemi,T., Nagoya,S., Kaya,M., Kawaguchi,S., Wada,T., Yamashita,T., and Ishii,S. (2003). Postoperative progression of pulmonary metastasis in osteosarcoma. *Clin. Orthop. Relat Res.* 159-166.
- Turc-Carel,C., Lizard-Nacol,S., Justrabo,E., Favrot,M., Philip,T., and Tabone,E. (1986). Consistent chromosomal translocation in alveolar rhabdomyosarcoma. *Cancer Genet. Cytogenet.* 19, 361-362.
- Udvardi,M.K., Czechowski,T., and Scheible,W.R. (2008). Eleven golden rules of quantitative RT-PCR. *Plant Cell* 20, 1736-1737.
- Vahey,M., Nau,M.E., Barrick,S., Cooley,J.D., Sawyer,R., Sleeker,A.A., Vickerman,P., Bloor,S., Larder,B., Michael,N.L., and Wegner,S.A. (1999). Performance of the Affymetrix GeneChip HIV PRT 440 platform for antiretroviral drug resistance genotyping of human immunodeficiency virus type 1 clades and viral isolates with length polymorphisms. *J. Clin. Microbiol.* 37, 2533-2537.
- Valadi,H., Ekstrom,K., Bossios,A., Sjostrand,M., Lee,J.J., and Lotvall,J.O. (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* 9, 654-659.
- van de Vijver,M.J., He,Y.D., van't Veer,L.J., Dai,H., Hart,A.A., Voskuil,D.W., Schreiber,G.J., Peterse,J.L., Roberts,C., Marton,M.J., Parrish,M., Atsma,D., Witteveen,A., Glas,A., Delahaye,L., van,d., V, Bartelink,H., Rodenhuis,S., Rutgers,E.T., Friend,S.H., and Bernards,R. (2002). A gene-expression signature as a predictor of survival in breast cancer. *N. Engl. J. Med.* 347, 1999-2009.
- Vandesompele,J., De,P.K., Pattyn,F., Poppe,B., Van,R.N., De,P.A., and Speleman,F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3, RESEARCH0034.
- Varkonyi-Gasic,E., Wu,R., Wood,M., Walton,E.F., and Hellens,R.P. (2007). Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods* 3, 12.
- Vester,B. and Wengel,J. (2004). LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. *Biochemistry* 43, 13233-13241.
- Vickers,M.M., Bar,J., Gorn-Hondermann,I., Yarom,N., Daneshmand,M., Hanson,J.E., Addison,C.L., Asmis,T.R., Jonker,D.J., Maroun,J., Lorimer,I.A., Goss,G.D., and Dimitroulakos,J. (2012). Stage-dependent differential expression of microRNAs in colorectal cancer: potential role as markers of metastatic disease. *Clin. Exp. Metastasis* 29, 123-132.
- Volinia,S., Calin,G.A., Liu,C.G., Ambs,S., Cimmino,A., Petrocca,F., Visone,R., Iorio,M., Roldo,C., Ferracin,M., Prueitt,R.L., Yanaihara,N., Lanza,G., Scarpa,A., Vecchione,A., Negrini,M., Harris,C.C., and Croce,C.M. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc. Natl. Acad. Sci. U. S. A* 103, 2257-2261.
- Voute,P.A., van den Berg,H., Behrendt,H., Michiels,E., and de,K.J. (1996). Ifosfamide in the treatment of pediatric malignancies. *Semin. Oncol.* 23, 8-11.
- Wachtel,M., Dettling,M., Koscielniak,E., Stegmaier,S., Treuner,J., Simon-Klingenstein,K., Buhlmann,P., Niggli,F.K., and Schafer,B.W. (2004). Gene expression signatures identify rhabdomyosarcoma subtypes and detect a novel t(2;2)(q35;p23) translocation fusing PAX3 to NCOA1. *Cancer Res.* 64, 5539-5545.

- Wade,M., Wang,Y.V., and Wahl,G.M. (2010). The p53 orchestra: Mdm2 and Mdmx set the tone. *Trends Cell Biol.* 20, 299-309.
- Wang,B., Howel,P., Bruheim,S., Ju,J., Owen,L.B., Fodstad,O., and Xi,Y. (2011a). Systematic evaluation of three microRNA profiling platforms: microarray, beads array, and quantitative real-time PCR array. *PLoS. One.* 6, e17167.
- Wang,W., Slevin,M., Kumar,S., and Kumar,P. (2005). The cooperative transforming effects of PAX3-FKHR and IGF-II on mouse myoblasts. *Int. J. Oncol.* 27, 1087-1096.
- Wang,Y., Zheng,D., Tan,Q., Wang,M.X., and Gu,L.Q. (2011b). Nanopore-based detection of circulating microRNAs in lung cancer patients. *Nat. Nanotechnol.* 6, 668-674.
- Weber,C.O. (1854). Anatomische Untersuchung einer hypertrophischen Zunge nebst Bemerkungen uber die Neubildung quergestreifter Muskelfasem. *Virchows Arch A Pathol Anat* 7, 115-125.
- Weber,J.A., Baxter,D.H., Zhang,S., Huang,D.Y., Huang,K.H., Lee,M.J., Galas,D.J., and Wang,K. (2010). The microRNA spectrum in 12 body fluids. *Clin. Chem.* 56, 1733-1741.
- Weber-Hall,S., Anderson,J., McManus,A., Abe,S., Nojima,T., Pinkerton,R., Pritchard-Jones,K., and Shipley,J. (1996). Gains, losses, and amplification of genomic material in rhabdomyosarcoma analyzed by comparative genomic hybridization. *Cancer Res.* 56, 3220-3224.
- Weiss,L. (2000). Early concepts of cancer. In *Cancer and Metastasis Reviews*, Kluwer Academic Publishers), pp. 205-217.
- Werner,H. and Le Roith D. (1997). The insulin-like growth factor-I receptor signaling pathways are important for tumorigenesis and inhibition of apoptosis. *Crit Rev. Oncog.* 8, 71-92.
- Westermarck,U.K., Wilhelm,M., Frenzel,A., and Henriksson,M.A. (2011). The MYCN oncogene and differentiation in neuroblastoma. *Semin. Cancer Biol.* 21, 256-266.
- Wexler,L.H., Meyer,W.H., and Helman,L.J. (2006). *Rhabdomyosarcoma and the undifferentiated sarcomas. Principles and Practice of Pediatric Oncology*, 5th Edition).
- Whelan,J. (2005). First clinical data on RNAi. *Drug Discov. Today* 10, 1014-1015.
- Whelan,J., Patterson,D., Perisoglou,M., Bielack,S., Marina,N., Smeland,S., and Bernstein,M. (2010). The role of interferons in the treatment of osteosarcoma. *Pediatr. Blood Cancer* 54, 350-354.
- Wickham,C.L., Sarsfield,P., Joyner,M.V., Jones,D.B., Ellard,S., and Wilkins,B. (2000). Formic acid decalcification of bone marrow trephines degrades DNA: alternative use of EDTA allows the amplification and sequencing of relatively long PCR products. *Mol. Pathol.* 53, 336.
- Wightman,B., Ha,I., and Ruvkun,G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855-862.
- Williamson,D., Missiaglia,E., de,R.A., Pierron,G., Thuille,B., Palenzuela,G., Thway,K., Orbach,D., Lae,M., Freneaux,P., Pritchard-Jones,K., Oberlin,O., Shipley,J., and Delattre,O. (2010). Fusion gene-negative alveolar rhabdomyosarcoma is clinically and molecularly indistinguishable from embryonal rhabdomyosarcoma. *J. Clin. Oncol.* 28, 2151-2158.

- Williamson,D., Selfe,J., Gordon,T., Lu,Y.J., Pritchard-Jones,K., Murai,K., Jones,P., Workman,P., and Shipley,J. (2007). Role for amplification and expression of glypican-5 in rhabdomyosarcoma. *Cancer Res.* 67, 57-65.
- Winkler,K., Beron,G., Delling,G., Heise,U., Kabisch,H., Purfurst,C., Berger,J., Ritter,J., Jurgens,H., Gerein,V., and . (1988). Neoadjuvant chemotherapy of osteosarcoma: results of a randomized cooperative trial (COSS-82) with salvage chemotherapy based on histological tumor response. *J. Clin. Oncol.* 6, 329-337.
- Winqvist,R., Mannermaa,A., Alavaikko,M., Blanco,G., Taskinen,P.J., Kiviniemi,H., Newsham,I., and Cavenee,W. (1993). Refinement of regional loss of heterozygosity for chromosome 11p15.5 in human breast tumors. *Cancer Res.* 53, 4486-4488.
- Witkos,T.M., Koscianska,E., and Krzyzosiak,W.J. (2011). Practical Aspects of microRNA Target Prediction. *Curr. Mol. Med.* 11, 93-109.
- Witlox,M.A., Lamfers,M.L., Wuisman,P.I., Curiel,D.T., and Siegal,G.P. (2007). Evolving gene therapy approaches for osteosarcoma using viral vectors: review. *Bone* 40, 797-812.
- Wong,F.L., Boice,J.D., Jr., Abramson,D.H., Tarone,R.E., Kleinerman,R.A., Stovall,M., Goldman,M.B., Seddon,J.M., Tarbell,N., Fraumeni,J.F., Jr., and Li,F.P. (1997). Cancer incidence after retinoblastoma. Radiation dose and sarcoma risk. *JAMA* 278, 1262-1267.
- Wu,L., Fan,J., and Belasco,J.G. (2006). MicroRNAs direct rapid deadenylation of mRNA. *Proc. Natl. Acad. Sci. U. S. A* 103, 4034-4039.
- Yang,H., Zhang,H., Zhu,L., Wang,J., Zhang,C., and Li,D. (2012). Pathway analysis of cancer-associated microRNA targets. *Int. J. Oncol.*
- Yang,J.S., Phillips,M.D., Betel,D., Mu,P., Ventura,A., Siepel,A.C., Chen,K.C., and Lai,E.C. (2011). Widespread regulatory activity of vertebrate microRNA* species. *RNA.* 17, 312-326.
- Yang,Y.H., Dudoit,S., Luu,P., Lin,D.M., Peng,V., Ngai,J., and Speed,T.P. (2002). Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 30, e15.
- Yin,G., Chen,R., Alvero,A.B., Fu,H.H., Holmberg,J., Glackin,C., Rutherford,T., and Mor,G. (2010). TWISTing stemness, inflammation and proliferation of epithelial ovarian cancer cells through MIR199A2/214. *Oncogene* 29, 3545-3553.
- Yoda,M., Kawamata,T., Paroo,Z., Ye,X., Iwasaki,S., Liu,Q., and Tomari,Y. (2010). ATP-dependent human RISC assembly pathways. *Nat. Struct. Mol. Biol.* 17, 17-23.
- Yoo,J., Lee,H.K., Kang,C.S., Park,W.S., Lee,J.Y., and Shim,S.I. (1997). p53 gene mutations and p53 protein expression in human soft tissue sarcomas. *Arch. Pathol. Lab Med.* 121, 395-399.
- Yoshida,A., Ushiku,T., Motoi,T., Shibata,T., Beppu,Y., Fukayama,M., and Tsuda,H. (2010). Immunohistochemical analysis of MDM2 and CDK4 distinguishes low-grade osteosarcoma from benign mimics. *Mod. Pathol.* 23, 1279-1288.
- Zampetaki,A., Willeit,P., Drozdov,I., Kiechl,S., and Mayr,M. (2012). Profiling of circulating microRNAs: from single biomarkers to re-wired networks. *Cardiovasc. Res.* 93, 555-562.

- Zavadil,J., Ye,H., Liu,Z., Wu,J., Lee,P., Hernando,E., Soteropoulos,P., Toruner,G.A., and Wei,J.J. (2010). Profiling and functional analyses of microRNAs and their target gene products in human uterine leiomyomas. *PLoS. One.* 5, e12362.
- Zhan,S., Shapiro,D.N., and Helman,L.J. (1994). Activation of an imprinted allele of the insulin-like growth factor II gene implicated in rhabdomyosarcoma. *J. Clin. Invest* 94, 445-448.
- Zhang,B., Pan,X., Cobb,G.P., and Anderson,T.A. (2007). microRNAs as oncogenes and tumor suppressors. *Dev. Biol.* 302, 1-12.
- Zhang,H., Kolb,F.A., Jaskiewicz,L., Westhof,E., and Filipowicz,W. (2004). Single processing center models for human Dicer and bacterial RNase III. *Cell* 118, 57-68.
- Zheng,D., Haddadin,S., Wang,Y., Gu,L.Q., Perry,M.C., Freter,C.E., and Wang,M.X. (2011). Plasma microRNAs as novel biomarkers for early detection of lung cancer. *Int. J. Clin. Exp. Pathol.* 4, 575-586.
- Zhou,H., Randall,R.L., Brothman,A.R., Maxwell,T., Coffin,C.M., and Goldsby,R.E. (2003). Her-2/neu expression in osteosarcoma increases risk of lung metastasis and can be associated with gene amplification. *J. Pediatr. Hematol. Oncol.* 25, 27-32.
- Zipper,H., Brunner,H., Bernhagen,J., and Vitzthum,F. (2004). Investigations on DNA intercalation and surface binding by SYBR Green I, its structure determination and methodological implications. *Nucleic Acids Res.* 32, e103.
- Zuo,Z., Calin,G.A., de Paula,H.M., Medeiros,L.J., Fernandez,M.H., Shimizu,M., Garcia-Manero,G., and Bueso-Ramos,C.E. (2011). Circulating microRNAs let-7a and miR-16 predict progression-free survival and overall survival in patients with myelodysplastic syndrome. *Blood* 118, 413-415.